Machine Learning, Game Theory, and Mechanism Design for a Networked World

A. Blum (PI), M. Blum (co-PI), M. Kearns (co-PI), T. Sandholm (co-PI), M.T. Hajiaghayi (senior personnel)

Many of the key algorithmic challenges in the context of the internet require considering the objectives and interests of the different participants involved. These include problems ranging from pricing goods and resources, to improving search, to routing, and more generally to understanding how incentives of participants can be harnessed to improve the behavior of the overall system. As a result, Mechanism Design and Algorithmic Game Theory, which can be viewed as "incentive-aware algorithm design," have become an increasingly important part of algorithmic research in recent years. Along different lines, the area of Machine Learning has made continued progress in developing methods that can generalize from data, adapt to changing environments, and improve performance with experience, as well as progress in understanding fundamental underlying issues. It is clear these capabilities will be critical as well in the context of such large decentralized systems. While these areas seem quite distinct, recent results of the PIs have developed a number of important connections between them: fundamental questions in Mechanism Design and Algorithmic Game Theory turn out to have a strong relation to issues in Machine Learning and vice versa, and techniques from each seem well-poised to help with key problems of the other. The aim of this proposal is to further develop these connections in order to produce powerful mechanisms for adaptive and networked environments, and improve the experience of users of the Web and internet.

In particular, three main thrusts of this proposal include:

1. *Machine Learning and adaptive algorithms for Mechanism Design.* Classic notions of revenue-optimal mechanisms require strong assumptions about players' types being drawn from distributions that are known to the designer. Machine learning methods can provide tools to achieve near-optimal mechanisms without such assumptions, and even potentially produce mechanisms that adapt to changing populations or conditions. Machine learning methods can also potentially be used to extract relevant preference information in combinatorial auction settings where preference functions can be exceedingly complex.

2. *Understanding the behavior of adaptive networked environments.* Game Theory traditionally studies equilibrium states in settings of full information and connectivity. However, in a large distributed network with multiple entities having limited information, it is critical to understand how network structure affects equilibria, and moreover to understand the dynamics. If each agent is adapting its behavior using a learning technique that is good for itself, what conclusions can one draw for the behavior of the overall system?

3. *Games to help machines understand the Web.* While web-search has become highly developed, the objects that make up the Web are only machine-understandable to a very limited extent. In recent work we have developed an approach to addressing this we call *games with a purpose*: online games that people enjoy playing, and as output produce good labels for images, identify features within images, and gather common-sense facts. These outputs can

be then used as inputs to machine learning algorithms. More broadly, can users of the Web be incentivized to provide information that helps improve the experience for all?

This collaborative proposal, involving researchers at Carnegie Mellon Univerity and the University of Pennsylvania, brings together researchers from mechanism design and game theory, machine learning, algorithm design, and protocol design, to address these interdisclipinary topics. The goal of this work more broadly is to investigate and develop what each area can contribute to the other. We now describe the three main thrusts in more detail below.

1 Machine Learning and Adaptive Algorithms for Mechanism Design

Mechanism Design can be thought of as a form of algorithm design, but where the entities supplying the inputs each have a stake in the outcome. As a result, the procedures produced should ideally be *incentive-compatible*, meaning that it is in each agent's best interest to report truthfully or to otherwise act in a well-behaved manner, and agents cannot gain an advantage by misrepresenting their values or "gaming" the system. Many of the problems facing the internet, for example, such as inappropriate hogging of bandwidth, email spam, and web spam, can often best be viewed as problems in the mechanism: the result of imperfections in the incentive structure of the internet.

Unfortunately, having to deal with the incentives of the various participants in a system makes the design of good mechanisms exceedingly difficult, especially in complex or non-static environments. In this proposed work, we aim to use techniques from machine learning and online algorithms to help manage this complexity. In our past work, we have shown how machine learning methods can be used to reduce mechanism design problems to standard algorithmic questions in a wide variety of revenue-maximizing auction settings [BBHM05], and we have shown how adaptive algorithms can be developed to handle incentive issues in auctions for dynamic environments [BH05, HKP04, HKMP05]. In our proposed work we intend to widen these connections, further developing the relationships between machine learning and mechanism design.

1.1 Machine Learning for Mechanism Design

Consider a seller with multiple digital goods or services for sale, such as movies, software, or network services, over which buyers may have complicated preferences. In order to sell these items through an incentive-compatible auction mechanism, this mechanism should have the property that each bidder i is offered a set of prices which do not depend on the value of her bid. The problem of designing a revenue-maximizing auction is known in the economics literature as the *optimal auction design* problem.

The classical model for optimal auction design assumes a Bayesian setting in which players' valuations (types) are drawn from some probability distribution that furthermore is known to the mechanism designer. For example, to sell a single item of fixed marginal cost, one should set the price that maximizes the profit margin per sale times the probability a random person would be willing to buy at that price. However, in complex or non-static environments, these assumptions become unrealistic. In these settings, machine learning can provide a natural approach to the design of near-optimal mechanisms without such strong assumptions.

Specifically, notice that while a truthful auction mechanism should have the property that the prices offered to some bidder i do not depend on the value of her bid, they can depend on the amounts bid by other bidders j. From a Machine Learning perspective, this is very similar to thinking of bidders as "examples" and our objective being to use information from examples $j \neq i$ to produce a good prediction with respect to example i. Thus, without presuming a known distribution over bidders (or even that bidders come from any distribution at all) perhaps if the number of bidders is sufficiently large, enough information can be learned from some of them to perform well on the rest. In recent work [BBHM05] we formalize this idea and show indeed that *sample-complexity* techniques from machine learning can be adapted to this setting to give quantitative bounds for this kind of approach. More generally, we show that sample-complexity analysis can be applied to reduce incentive-compatible mechanism design to more standard algorithm-design questions, in a wide variety of revenue-maximizing auction problems.

The high level idea of the reduction in [BBHM05] is actually fairly simple. For concreteness, let us imagine we are selling a collection of n goods or services of zero marginal cost to us, to m bidders who may have complex preference functions over these items, and our objective is to achieve revenue comparable to the best possible assignment of prices to the various items we are selling. (Thus, we are in the setting of maximizing revenue in an unlimited supply combinatorial auction.) Then given a set of bids S, we perform the following operations. We first randomly partition S into two sets S_1 and S_2. We then consider the purely algorithmic problem of finding the best set of prices p_1 for the set of bids S_1 (which may be difficult but is purely algorithmic), and the best set of prices p_2 for the set of bids S_2. We then use p_1 as offer prices for bidders in S_2, giving each bidder the bundle maximizing revealed valuation minus price, and use p_2 as offer prices for bidders in S_1.[1] We then show that even if bidders' preferences are extremely complicated, this mechanism will achieve revenue close to that of the best fixed assignment of prices to items so long as the number of bidders is sufficiently large compared to the number of items for sale. For example, if all bidders' valuations on the grand bundle of all n items lie in the range $[1, h]$, then $\tilde{O}(hn/\epsilon^2)$ bidders are sufficient so that with high probability, we come within a $1 + \epsilon$ factor of the optimal fixed item pricing. Or, if we cannot solve the algorithmic problem exactly (since many problems of this form are NP-hard), we lose only a $1 + \epsilon$ factor over whatever approximation our method for solving the algorithmic problem gives us. More generally, these methods apply to a wide variety of pricing problems, including those in which bidders have both public and private information, and also give a formal way one can address design issues such as how fine-grained a market segmentation should be.

1.2 Auctions for dynamic environments

The above work considers a static environment in which all bidders are present up front. If bidders instead arrive one at a time, then so long as they cannot game their timing information, we can view this as an *online* learning problem, and apply online learning techniques such as [KV03] for

[1] One might wonder why we do not optimize over the entire set $S - \{i\}$ when choosing offer prices for bidder i. The reason is that such a scheme does *not* in fact produce good revenue guarantees in an adversarial setting, for much the same reason that *leave-one-out cross-validation* can be highly unstable in machine learning.

auction problems that have appropriate structure [BH05, BB06] (in the future work section below we discuss this setting further). However, if bidders can misrepresent their timing information as well as their valuations, then this problem becomes much more difficult.

Specifically, consider the problem of auction design for dynamic environments in which agents arrive and depart dynamically and in which goods are inherently temporal. Examples of such settings include the problem of WiFi allocation, sale of last-minute tickets, scheduling of scientific grid computing and projects such as MoteLab (at Harvard University) and PlanetLab (at Princeton University). In all these settings, a key challenge is to design online auctions that are incentive-compatible with respect to timing information, so users are not motivated to misrepresent their urgency or lack thereof. Here, even social-welfare optimization can be quite difficult. In particular, one of the most important techniques for designing truthful mechanisms (the Vickrey-Clarke-Groves (VCG) scheme) is inapplicable in most online problems because it requires computing an optimal allocation, which is generally impossible in the online setting [FP03].

In our work [HKP04], motivated by auctions of digital goods, say in eBay or Amazon.com, we study a limited-supply online auction problem, in which an auctioneer has k goods to sell and bidders arrive and depart dynamically. We assume that agent valuations are drawn independently from some unknown distribution[2] and construct an adaptive auction that is nevertheless value- and time-strategyproof. For the $k = 1$ problem we have a strategyproof variant on the classic secretary problem. We present a 4-competitive (e-competitive) strategyproof online algorithm with respect to offline Vickrey for revenue (efficiency). We also show (in a model that slightly generalizes the assumption of independent valuations) that no mechanism can be better than 3/2-competitive (2-competitive) for revenue (efficiency). Our general approach considers a *learning* phase followed by an accepting phase, and is careful to handle incentive issues for agents that span the two phases. We extend to the $k > 1$ case, by deriving strategyproof mechanisms which are constant-competitive for revenue and efficiency. We believe this is the first analysis of an adaptive online auction that makes no assumptions about the distribution from which agent valuations are drawn but provides both constant-competitiveness and time-strategyproofness.

In [HKMP05] we study the case of *re-usable* goods, such as processor time or wireless network access, which can be assigned to different agents at different times. Each agent is assumed to arrive and depart dynamically, and in the basic model requires the resource for one unit of time. In this work, we provide characterizations for the class of truthful online allocation rules and also present an online auction for unit-length jobs that achieves total value that is 2-competitive with the maximum offline value. We further prove that no truthful deterministic online mechanism can achieve a better competitive ratio. Finally, we generalize our model to settings with multiple re-usable goods and to agents with different job lengths.

1.3 Preference elicitation in combinatorial auctions

Preference elicitation in combinatorial auctions is another setting in which machine learning and mechanism design intersect. Here, the key problem is that bidders in a combinatorial auction may have extremely complicated valuation functions, and may not be able to easily "dump" them into the mechanism, even ignoring incentive issues. Instead, it may be much more natural for the

[2] For arbitrary valuations, no constant-competitive auction is possible [EYFKT92, LN00].

mechanism to interact with the bidders by asking appropriate questions about their preferences, and this process is called *preference elicitation* [CS01, HS04].

In our recent work [BJSZ04, ZBS03], we develop connections between this problem and problems of learning via queries studied in computational learning theory. In particular, consider a combinatorial auction in which the auctioneer has a set S of n items to be partitioned among a set of k bidders who each have complicated preferences over subsets of these items. Then, as a learning problem, this can be thought of as a setting in which there are multiple target concepts that can each be queried separately, but where the goal is not so much to learn each concept as it is to produce something like an "optimal example". For instance, if there are just two bidders with preference functions f and g, then the goal is to find a partition (S', S'') of the n items to maximize $f(S')+g(S'')$. Thinking in terms of functions over $\{0,1\}^n$, the goal is to find $x \in \{0,1\}^n$ to maximize $f(x) + g(\bar{x})$. In [BJSZ04, ZBS03] we show how techniques from query learning can be applied to solve problems of this type (for example, if preference functions can be assumed to have certain structure) and more generally develop formal connections between these problem areas. These directions have been further developed by Lahaie and Parkes [LP04].

1.4 New directions and proposed work

There are a number of specific open questions related to the above work that we intend to explore. For example, can the results of [BBHM05] be generalized to the limited-supply setting by using *envy-free* pricing functions [GHK$^+$05] on each subset, or in the other direction extended to problems exhibiting economies of scale, or even more broadly to problems where the auctioneer is a service provider who must pay a cost that is a more complex function of the set of agents being served? Can the results on preference elicitation be extended to broader classes of preference functions? We focus here instead, however, on a few broader questions related to these topics.

First, the work described above shows how machine learning methods can be used to design mechanisms for a number of different types of standard auction problems. An intriguing question is whether such methods can be used to iteratively improve mechanisms in broader real-world settings where the capabilities and motivations of the players may be much more complex. For example, there is a substantial amount of money transacted through search-query ad-auctions, and there has been quite a bit of work on how such auctions should be designed [ABF$^+$05, AHKM06, IJMT05]. However, it seems impossible to design a completely truthful mechanism in such settings given the complexity of motivations of the various parties in the system. Instead, machine learning methods would be a natural approach to iteratively adjusting various parameters of the mechanism, in order to optimize performance given the behavior observed.

Along a related line, Conitzer and Sandholm [San03, CS02, CS03] propose the notion of *automated mechanism design* in which the distribution of players' types, as well as the objective function and constraints on the mechanism, are fed into a program that then performs a computation to output the optimal mechanism for those conditions. The idea is that even in settings for which classic impossibility results such as the Gibbard-Satterthwaite theorem apply, one may be able to still perform optimally for the specific problem instance (e.g., the particular distribution of types). Experimentally, this approach has been applied to a number of settings including a simple divorce-settlement model, public-goods problems, and combinatorial auctions, and has empirically been found to produce a substantial increase in revenue in combinatorial auction problems compared to VCG mechanisms [CS03, CS04]. Two natural connections between this work and that described above are:

1. Can machine learning methods be used to replace the assumption of a known type-distribution in this setting? For example, suppose the only access one has to the type space is to propose and "field-test" a mechanism and then observe its behavior. Can one devise a procedure such that after a certain number of iterations, one can make guarantees on the mechanism being close to optimal?

2. Can automated mechanism design be combined with online algorithms to work in the context of dynamic environments, such as those described in Section 1.2 above? One would like also to consider richer temporal preference models in the online setting, rather than just arrival and departure times, and consider *combinatorial* auctions in this setting as well.

Along somewhat different lines, the generic reductions of [BBHM05] for converting mechanism design problems into standard algorithmic ones require a batch setting in which all bidders are present at once. If instead bidders arrive one at a time (and let us now imagine they cannot control or misrepresent their arrival times), then online learning methods for "combining expert advice" [LW94, FS97] and generalizations [KV03] can be applied for simple problems such as the auction of a single digital good [BKRW03, BH05]. These methods can also be applied to cases where the approximation algorithms used to solve the algorithmic problem have a very special structure [BB06]. However, we know of no generic reduction that works for such broad classes of problems as our reduction in the batch setting. In fact, the problem seems similar to, but different from, the issue that arises when attempting to apply the VCG mechanism for *efficiency* maximization in the presence of approximation algorithms. Thus, a key question is: can a generic reduction from unlimited-supply auction problems to their corresponding algorithmic problems be given for such *online* mechanism design problems?

2 Equilibria and adaptive selfish behavior in networked environments

Game Theory traditionally studies equilibrium states in settings of full information and connectivity. However, in a large distributed network with multiple entities having limited information, it is crucial to understand both how network structure affects equilibria, and what can one expect in terms of dynamics when players are using learning to adapt their behavior. In past work we have investigated several problems of this flavor, including analyzing the game-theoretic behavior of no-regret learning algorithms in the context of routing, studying how network structure affects equilibria in market pricing, and examining evolutionary game theory in a network setting.

2.1 Game-theory of adaptive algorithms

Game Theory focuses on properties of equilibrium states. For instance, one highly studied concept in Algorithmic Game Theory is the notion of *price of anarchy*. The price of anarchy for some problem (such as routing in a congested network) is the worst-case gap between a Nash equilibrium configuration (each player is routing in the best possible way given the behavior of the others) and a globally optimal solution (one designed to minimize total traffic delay). However, Nash equilibria are somewhat delicate objects. In a large system with multiple entities having limited information, it is instead much more natural to assume each is selfishly adapting its behavior based on past experience, producing results that may or may not stabilize. One would therefore like to use an understanding of the properties of such adaptive algorithms to draw conclusions for the behavior of the overall system.

In particular, there has been substantial work in learning theory and game theory on adaptive *no-regret* algorithms for a wide class of repeated decision-making problems including online routing [Han57, FS97, TW02, KV03, Zin03, AK04, MB04, DH06]. These are adaptive strategies an individual can use that give strong guarantees on performance even in adversarially-changing environments. Specifically, their average cost per time step is guaranteed to approach that of the best fixed strategy in hindsight (or better) over time. Moreover, the convergence rates of a number of these strategies are quite good, even in situations like routing where the number of alternatives (possible paths to take) may be exponential in the natural description-length of the problem. In particular, the no-regret property (also called Hannan Consistency) can be viewed as a natural *definition* of well-reasoned self-interested behavior over time. Thus, if all players are adapting their behavior in such a way, can we say that the system as a whole will approach Nash equilibrium?

In [BEL06] we show that in the Wardrop setting of multicommodity flow and infinitesimal agents, if players use strategies with the no-regret property, then flows will indeed approach equilibrium. In particular, a $1 - \epsilon$ fraction of the time steps will have the property that at most an ϵ fraction of the players will have more than an ϵ incentive to deviate from their chosen path, where ϵ approaches 0 at a rate that depends polynomially on the size of the graph, the regret-bounds of the algorithms, and the maximum slope of any latency function.[3] We also show that price-of-anarchy results may be applied to these approximate equilibria, and in addition give extensions to the finite-size (non-infinitesimal) load-balancing model of Azar [Aza98].

One way this result can be viewed is as follows. No-regret algorithms are very compelling from the point of view of individuals: if you use a no-regret algorithm to drive to work each day, you will get a good guarantee on your performance no matter what is causing congestion (other drivers, road construction, or unpredictable events). But it would be a shame if, were everyone to use such an algorithm, this produced globally unstable behavior. Our results imply that in the Wardrop routing model, so long as edge latencies have bounded slope, we can view Nash equilibria as not just a stable steady-state or the result of adaptive procedures specifically designed to find them, but in fact as the inevitable result of individually-selfish adaptive behavior by agents that do not necessarily know (or care) what policies other agents are using. Moreover, our results do

[3]A more traditional notion of approximate Nash equilibrium requires that *no* player will have more than ϵ incentive to deviate from her strategy. However, one cannot hope to achieve such a guarantee using just the no-regret property, since no-regret algorithms may occasionally take bad paths, and in fact such experimentation is even *necessary* when players have limited information. For instance, if there is a route you have had bad experience with but have not taken in a long time, you might begin to worry that it could have significantly improved since your last attempt (to much better than your current route), and thus you would need to give it another try to avoid the possibility of substantial regret. For a similar reason, one cannot hope that *all* rounds will be approximate-Nash. Finally, our guarantee may make one worry that some users could always do badly, falling in the ϵ minority on every round, but in fact the no-regret property can be used to further show that no player experiences many days in which her expected cost is much worse than the best path available in that round.

not in fact require that users follows strategies that are no-regret in the worst-case, as long as their behavior satisfies the no-regret property over the sequence of flows actually observed.

2.1.1 Future work

One natural question we intend to explore is to what extent results of this form can be had for broader classes of problems. There are a number of problems for which specific adaptive dynamics are known to converge to natural equilibria, but can we argue results of this form for no-regret algorithms, which players may want to use because they come with strong worst-case guarantees? Alternatively, can we *directly* analyze the gap between the social welfare when players follow no-regret algorithms and the social-welfare of a globally optimal solution? One could argue that players optimizing for themselves in such a self-interested, adaptive way is much more reflective of "true anarchy" than is a Nash equilibrium.

Along somewhat different lines, algorithms with a stronger guarantee of "no internal regret" have the property that in general games, when played against each other, their empirical distributions will tend to correlated equilibrium [FV97, HMC00]. However, this does not necessarily mean that play on any given round necessarily looks like an equilibrium. For example, imagine a routing game with two routes to choose from, Left and Right, such that each route has cost 0 if at most half of the traffic is using it and cost 1 if more than half of the traffic is using it. If players use no-regret algorithms and traffic does not begin at exactly 50/50, then traffic will oscillate: each route will be cheap or expensive on alternate days, each player will experience bad traffic roughly every other day as well, there will be no regret (no fixed route did better than you did), and the empirical distribution of play will be an approximate correlated equilibrium. However, on any *given* day, approximately half the population has envy towards the other half *for that day*.[4] More generally, the problem is that we have wildly fluctuating payoffs for the different actions from day to day. This suggests that from a *mechanism design* perspective, one should consider the question: given a repeated game that a mechanism designer has some power to adjust (e.g., say the mechanism designer can smooth out the edge latency functions, or in an auction can smooth out the rules in some way), can one construct a game such that regret-minimizing learning algorithms will produce desirable behavior?

In a different direction, and relating this topic to that of the previous section, no-regret learning algorithms can also be used by the network itself in settings where a central authority must determine routing tables without knowledge of the exact traffic patterns. For example, in the setting of *oblivious routing*, the network must commit to how flow will be routed between all pairs of nodes in the network before observing the demands. Räcke [Rae02] showed one could obtain oblivious routing schemes with polylogarithmic competitive ratios in terms of the maximum edge congestion for general edge-capacitated undirected graphs, and Azar et al. [ACF$^+$03] showed that one could efficiently compute the minimax optimal strategy for this problem. However, if the central authority has the ability to adjust the routing tables periodically based on past behavior, then no-regret algorithms can potentially be used to adapt these tables in such a way that performance over time is nearly as good as the best fixed routing table in hindsight, which might be much better than minimax optimal if traffic is not adversarial. In fact, in past work

[4] This example explains why in our results described above we needed edge latency functions to have bounded slope.

[BBCM03] we showed how a no-regret algorithm of Zinkevich [Zin03] could be used to this effect. However, the overall algorithm here is highly computationally intensive (requiring solution of a semidefinite program at each iteration) and perhaps a simpler, more efficient approach can be found.

One drawback in the above work is that the competitive ratio guarantees require undirected networks and uncapacitated vertices. Motivated by applications in optical networks, we [HKLR05] show that for any vertex-capacitated (undirected) graph or more generally any vertex- or edge-capacitated directed graph (such as the Internet in which the capacities of edges in different directions are different or wireless networks in which different nodes have different transmission powers), if the demands for different node-pairs are not completely adversarial but instead drawn i.i.d from some distribution, then there is an oblivious routing algorithm that is within $O(\log^2 n)$ of the optimum congestion, with high probability. Our approach is based on sampling from the distribution, solving the problem optimally, and then taking the average of the optimums. It is an interesting question whether these bounds can be improved and to what extent the assumptions can be generalized.

2.2 Effects of network structure on equilibria

For several years, co-PI Kearns and his colleagues have been systematically investigating how network structure affects equilibria in a variety of game-theoretic, strategic and economic settings. Given that virtually all of the Internet-based settings described in this proposal feature interactions taking place over some network topology (whether that of the Internet itself, or some overlay such as the peers in a file-sharing system), it is of fundamental interest to develop a theory that relates the topology and other features of the network to the collective outcome, as it is described by various equilibria notions and concepts such as the Price of Anarchy. Significant progress in this direction has been made in the last five years, but there are notable areas of importance that remain essentially untouched.

In 2001, graphical games were introduced [KLS01] to model large-population games in which each player's payoff is determined only by its own actions and that of its neighbors in an arbitrary undirected graph or network, and it was shown in a series of papers [KLS01, KO02, KKLO03] that such a potentially succinct representation may have significant algorithmic benefits as well. It was also shown [KKLO03] that there is a very direct relationship between the structure of the underlying graph and the amount and nature of "shared randomization" required in any correlated equilibrium of the game (as represented by a Markov network or undirected probabilistic graphical model). In addition to being one of the first strong ties between network structure and equilibrium properties, this result is especially intriguing in the current context of learning in networked games, since correlated equilibrium appears to be the most natural general convergence notion for the powerful no-regret learning algorithms discussed above. An example of a natural and fundamental open problem in this area is whether a large population of players in a graphical game playing in a no-regret fashion — but from only *local* information (the actions of their neighbors) — will in fact converge to a correlated equilibrium, and one that can be implemented with only local randomization to achieve the required correlations. More generally, the correlated equilibrium result for graphical games raises a number of other interesting connections to learning theory, including its use of maximum entropy methods.

In more recent work, strong relationships between graph topology and equilibrium properties have been established for price (Arrow-Debreu) equilibria in exchange markets [KKO+04] and for evolutionary stable strategies (ESS) in evolutionary game theory [KS06]. The former include necessary and sufficient conditions for there to be no price or wealth variation (which involve perfect matching and weak expander properties), and the latter include conditions under which random graphs will preserve the ESS of the fully connected (classical) case. Significant areas for further work include generalizations to more complex microeconomic models, such as general equilibrium settings in which production and consumption are permitted rather than only exchange, and examination of richer generative models for the underlying network structure, such as whether classical ESS are preserved under preferential attachment.

3 Utilizing the power of the players: Games to produce a better Web

> *Construction of the Empire State Building: 7 million human-hours. The Panama Canal: 20 million human-hours. Estimated number of human-hours spent playing solitaire around the world in one year: 9 billion.*

The point of truthful mechanism design is to give the players in a system an incentive to act in a way that contributes to an overall desired outcome. In the context of the Web, can we give the users (web-surfers and web-page owners) an incentive to provide help that will improve everyone's experience? Can users be given an incentive to help provide machines an *understanding* of what is out there? In recent work we (and our students) have been exploring an approach to this we call *Games with a Purpose*. These are games (in the traditional colloquial sense of something people enjoy doing) such that by having fun playing, people contribute to our understanding of objects that make up the Web. These include: the ESP Game [vAD04] and Phetch [vAGK+06] which are designed to produce good labelings of images on the Web, Peekaboom [vALB06] which is designed to locate objects inside images, and Verbosity [vAKB06], a game designed to collect common-sense facts. By playing these games, individuals provide machine-usable information that can then be used to improve search and improve our ability to make sense of content on the Web. In addition, their outputs could be naturally used as inputs to a machine learning algorithm, because they address the difficult *labeled data problem*: namely, that machine learning methods typically require large quantities of labeled examples to work with, and yet labeling is often a resource-intensive process. Thus, this direction can also be thought of as combining mechanism design and machine learning, though in a quite different way from the other directions described here. A brief description of each of these games is given below.

ESP: In this game, playable at www.espgame.org, a player is paired with a random partner (many people are online playing at the same time), and the goal of the partners is to agree on a word: i.e., to guess what the other is thinking. However, the only information the partners have in common is an image that is presented to both of them at the beginning of each round. More specifically, players type a sequence of text strings, and once there is some string that both of them have typed, they receive a score based on time and move on to the next image. An example of the progression of a round and well as a screen shot of the game is given in Figure 1 below.

Figure 1: The left side of the figure shows a sample progression of a round of the game, with partners agreeing on an image. Note that neither player can see the other's guesses. The right side shows a snapshot of the ESP game. Taboo words are words the players are not allowed to use for the given image (they correspond to labels that have already been provided for this image in the past). Players try to "agree" on as many images as they can in 2.5 minutes. The thermometer at the bottom measures how many images the partners have agreed on so far.

We have found empirically that strings agreed upon are excellent labels for the images given. The game is also something people want to play: it has already produced over 10 million image labels, with some users (not us) playing over 40 hours per week!

Phetch: This is a game (not yet released to the public but for which we have performed small-scale experiments) that provides a different approach to annotating images on the web [vAGK+06]. Unlike the ESP game, which produces short single-word labels, the objective here is to produce longer explanatory *descriptions* of web images. These are descriptions that, for example, could be read to a blind person so they would have a reasonable idea of what the image was about. In fact, the game proceeds in much this way. Phetch is a game for three to five players. In each round, one player is the "describer" and the others are the "seekers". The describer is shown an image and then must provide a text description to the seekers. The seekers must then find (fetch) the image using a search engine. The first seeker to find the image wins points, and the describer wins points based on how quickly the image is found. This game turns out to be quite difficult for seekers when completely random images and standard search engines are used, so we instead use images collected and annotated by the ESP game described above. An example of an image and the description produced is given in Figure 2.

Peekaboom: Peekaboom (www.peekaboom.org) is a game designed for locating objects inside images [vALB06]. As with ESP, this game is played by two players randomly paired together as partners, but as with Phetch, this game is asymmetric. In Peekaboom, the two partners take turns "peeking" and "booming". The Boom player is shown an image, along with a target word corresponding to some object presumably inside the image. The Peek player is shown just a blank screen. Boom's job is to get Peek to type the target word as quickly as possible , in which event both players score points. To do this, Boom can click on portions of

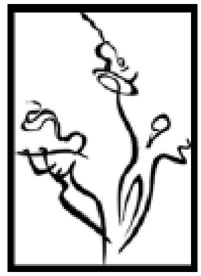

Figure 2: An example of an image described through Phetch and ESP. Under the ESP game, labels given are “man” and “woman”. Description given using Phetch is “an abstract line drawing of a man with a violin and a woman with a flute.”

the image, which then are revealed to Peek. (In addition, there is a provision for providing various types of hints and telling Peek if she is hot or cold.) The parts of the image revealed in this game become good indicators of the location of the target object in the image (see Figure 3).

Figure 3: Given an image and a query word, the result of Peekaboom is a region of the image related to the query word.

Verbosity: This game is designed to collect common-sense facts[vAKB06]. While this has less *immediate* application to web search, this style of game can be useful towards our long-term goal of better understanding more general content on the Web. The high-level idea, as with the games described above, is to convert the process of entering the desired information (in this case, common-sense facts) from a chore into an enjoyable activity. Verbosity is similar to the classic “$25,000 Pyramid” game show or the party game Taboo$^{\mathrm{TM}}$. One player is given a keyword and must help her partner guess the keyword by giving clues. To make these clues more machine-understandable, the game will at various times require these clues to follow certain templates, such as “is a kind of X” or “is used for Y”.

The designs of these games are not completely robust in the sense that if players are too highly motivated to score points, they can often find ways to cheat if they put enough effort into it. For instance, they may agree on some common forum to use simple stop-words like “the” in the ESP game, or subvert the scoring mechanism in various ways in other games. However, we have found that the labels produced as a result of these efforts are often easy to identify and discard. Thus, rather than trying overly hard to penalize all methods of cheating, it is sufficient to simply allow those playing "nontraditionally" to have their fun and ignore their output.

One of the key challenges here, and one reason for this proposed joint investigative effort, is to better formalize the connections between issues raised in these games and issues in traditional mechanism design. We expand on this in more detail below.

3.1 Challenges, open questions and new directions

Challenges and open questions along these lines include:

- Can a game or other mechanism be devised to identify *web-spam*, namely web pages that pretend to be something that they are not, or somehow pretend to be more important than they really are?

- Can a game be developed to help with other forms of information understanding on the web, such as summarizing the content of web pages, or aiding in the organization of information?

- How can one best formalize the connections between this and traditional mechanism design? Traditional mechanism design centers around the problem of private preference information. Here, the preference information is presumed public (people like to play games) and in particular several of the games we have developed are specifically motivated by games people pay money to purchase. The issue instead is that people have other capabilities like the ability to label images, which the mechanism-designer does not have, and the mechanism designer can only induce people to provide this capability by providing some appropriate incentive (such as a game designed to capture the use of this capability). This is related to the issue of *implementation under complete information* studied in economics. In that setting there is state θ (say the vector of players' types) known to all players but not to the mechanism designer who would like this information. The classic solution to that setting is for the mechanism to ask all players to reveal θ and then to give a big penalty for any discrepancies.[5] However, this is more like forcing employees to do a good job by using them to check each others' work and goes against our main constraint that the games must be fun to play. In fact, the most natural theoretical analogy for these games may be that of multi-prover proof systems, except that provers must be given a reason to "want" to provide their output.

  A second issue concerns the notion of incentive-compatibility. In the context of these games, we can split the incentive issues into three parts. First (the most nebulous part) is that the game should be fun so that people want to play. The second, however, is that assuming people are playing and are motivated to score points, the point system should encourage users to exercise the capability we are interested in (the ability to label images, find objects, and so forth) and to provide the results to us. Finally, because we cannot completely

[5] One problem with this approach is that any common output becomes a Nash equilibrium; this issue is rectified by Moore and Repullo [MR88, Moo92] who provide a sequential mechanism under which truth-telling is the only subgame perfect Nash equilibrium.

constrain players to stay within the framework of the game (e.g., not developing some out-of-band communication with their partner) we need to be able to identify the results of cheating and discard them.

This last point brings up a larger issue within mechanism design that we intend to explore further. Namely, can we develop a notion of *robust* mechanism design for information-gathering settings such as these, such that even if we cannot completely stop people from colluding, or even if certain players have motivations that are not captured by our model, we can nonetheless reach our desired goal.

3.2 Behavioral network game theory

The work described above suggests considering the intersection of networked communication and behavioral game theory, which studies the behavior of real organizations. In particular, all of the work in algorithmic game theory to date, including that discussed at length in this proposal, assumes standard rationality on the part of all participants. However, there is a large and rapidly expanding body of work in behavioral game theory documenting the systematic ways in which real humans and organizations deviate (sometimes radically) from classical rationality notions [Cam03]. Given that algorithmic game theory is ultimately meant to be a design and understanding tool for the actual use of human systems (such as the Internet), it is important to begin to develop a behavioral version of the theory.

Co-PI Kearns and colleagues have begun this effort with recent human-subject experiments in what might be called behavioral network science. The basic goal of this work is to understand how an underlying network structure of interaction influences the actual behavior of human subjects in simple game-playing and distributed optimization settings. The specific experiments so far have had roughly 40 participants each controlling the color of a vertex in a collective effort to find a proper coloring of a network from only local information. Findings include the fact that within certain families of graphs, adding edges (which only adds constraints to the problem, but reduces diameter or communication distances) can actually improve collective performance (equivalently, reduce the "behavioral Price of Anarchy"). More generally, it is important and interesting to revisit the results of algorithmic game theory with the lessons of behavioral economics in mind, in both networked and non-networked settings. For instance, behavioral economics has established that in certain games, a modification of the raw monetary payoffs that models altruism (such as "inequality aversion" [Cam03]) better predicts human behavior. How do such modified (ir)rationality notions change the basic theory of algorithmic mechanism design, graphical games and economies, learning in games, and so on?

4 Results from prior NSF support

Results from prior NSF support have been described throughout this proposal. Two most closely related NSF grants involving the PIs (both expiring Summer 2006) are:

- **IIS-0121678 ITR/PE+SY. Collaborative Research: Foundations of Electronic Marketplaces: Game Theory, Algorithms and Systems.** This is a multi-institution collaborative ITR grant involving computer scientists and economists in the study of fundamental issues in electronic marketplaces. Tuomas Sandholm is involved as PI and Avrim Blum as senior personnel. This grant also involved researchers from Northwestern and UCSB. Much of the past work described here building formal connections between machine learning and mechanism design was conducted under this grant.

- **CCR-0122581 ITR/SY+AP+IM. ALADDIN: A Center for ALgorithm ADaptation, Dissemination and INtegration.** This is a multi-institution ITR grant involving multiple algorithm and application-related projects called PROBEs (PROBlem-oriented Explorations) around a central administrative core running workshops bringing different communities together. Avrim Blum and Manuel Blum are both involved as senior personnel. Work described here on *games with a purpose* was conducted under this grant.

5 Outreach and Broader Impact

The proposed work on *Games with a Purpose* has potential for broad impact in at least three distinct ways. First, the outputs of these games, such as labels on images, *descriptions* of images that could be accessed by (or read by computer to) blind web-users, and the outputs of new games to be developed have the potential for improving the experience of all web users, either directly or by having their outputs fed into a learning algorithm that would go the rest of the way (e.g., in labeling images). Second, these games by their nature involve thousands of players, many of whom can then be inclined to learn more about the research underlying these games (links are provided on the game homepage) and perhaps motivated to become more interested in Computer Science. In particular, these games highlight the breadth of Computer Science and while not (yet) directly theoretical can be viewed as a kind of "multi-prover interactive proof system for the real world". Finally, developing these games is an activity that has involved a number of undergraduate REU students (Shiry Ginosar, Mihir Kedia, and Roy Liu, who are co-authors on the resulting research papers) and we believe this is an excellent way to get undergraduates excited about research. Co-PI Kearns has also involved undergraduates in his work on behavioral network game theory, and teaches a prerequisite-free undergraduate course called *Networked Life* at Penn involving many of the topics covered in this grant.

At a broader level, this proposal is intended to weave together several different strands of computer science research: Machine Learning, Mechanism Design and Game Theory, and Games with a Purpose. Though very different, all three can sometimes be used to attack the same problem. A portion of the funding requested is to support three workshops, that would be on topics of this type. For example, consider the problem of *web-spam*, web pages that fool surfers and search-engines into thinking that they are more important or relevant than they truly are. This problem has potential to be addressed from all three angles. From a machine learning point of view, one can attempt to learn characteristics of such pages in terms of content, layout, or link structure; from a mechanism design perspective one can imagine incentive systems that encourage web-page designers to be more "honest" about their content; and from the perspective of Games with a Purpose one can imagine developing some sort of enjoyable activity that as a result aids in identifying such pages or more generally improving web search. These workshops will bring researchers together to think about such problems from all three directions.

References

[ABF$^+$05] K. Asdemir, H.K. Bharghava, J. Feng, G.W. Flake, and D.M. Pennock, editors. *Workshop on Sponsored Search Auctions*. ACM-EC, 2005.

[ACF$^+$03] Yossi Azar, Edith Cohen, Amos Fiat, Haim Kaplan, and Harald Raecke. Optimal oblivious routing in polynomial time. In *STOC '03: Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 383–388, New York, NY, USA, 2003. ACM Press.

[AHKM06] K. Asdemir, J. Hartline, B. Kitts, and C. Meek, editors. *2nd Workshop on Sponsored Search Auctions*. ACM-EC, 2006. http://www.bus.ualberta.ca/kasdemir/ssa2/.

[AK04] B. Awerbuch and R. Kleinberg. Adaptive routing with end-to-end feedback: Distributed learning and geometric approaches. In *Proceedings of the 36th ACM Symposium on Theory of Computing*, 2004.

[Aza98] Y. Azar. *On-line Load Balancing Online Algorithms - The State of the Art*, chapter 8, pages 178–195. Springer, 1998.

[BB06] Maria-Florina Balcan and Avrim Blum. Approximation algorithms for item pricing. In *Proceedings of the 7th ACM conference on Electronic commerce (EC))*, New York, NY, USA, 2006. ACM Press. To appear.

[BBCM03] N. Bansal, A. Blum, S. Chawla, and A. Meyerson. Online oblivious routing. In *ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, pages 44–49, 2003.

[BBHM05] M.F. Balcan, A. Blum, J. Hartline, and Y. Mansour. Mechanism design via machine learning. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 605–614, 2005.

[BEL06] A. Blum, E. Even-Dar, and K. Ligett. Routing without regret: On convergence to nash equilibria of regret-minimizing algorithms in routing games. In *PODC*, 2006.

[BH05] A. Blum and J. Hartline. Near-optimal online auctions. In *Proc. 16th Symposium on Discrete Algorithms*. ACM/SIAM, 2005.

[BJSZ04] Avrim Blum, Jeffrey Jackson, Tuomas Sandholm, and Martin Zinkevich. Preference elicitation and query learning. *Journal of Machine Learning Research*, 5:649–667, 2004. (special issue on Learning Theory).

[BKRW03] A. Blum, V. Kumar, A. Rudra, and F. Wu. Online learning in online auctions. In *Proc. 14th Symposium on Discrete Algorithms*. ACM/SIAM, 2003.

[Cam03] C. Camerer. *Behavioral Game Theory*. Princeton University Press, 2003.

[CS01] Wolfram Conen and Tuomas Sandholm. Preference elicitation in combinatorial auctions: Extended abstract. In *Proceedings of the ACM Conference on Electronic Commerce (ACM-EC)*, pages 256–259, Tampa, FL, 2001. A more detailed description of the algorithmic aspects appeared in the IJCAI-2001 Workshop on Economic Agents, Models, and Mechanisms, pp. 71–80.

[CS02] Vincent Conitzer and Tuomas Sandholm. Complexity of mechanism design. In *Proceedings of the 18th Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 103–110, 2002.

[CS03] Vincent Conitzer and Tuomas Sandholm. Applications of automated mechanism design. In *UAI-03 workshop on Bayesian Modeling Applications*, 2003.

[CS04] Vincent Conitzer and Tuomas Sandholm. Self-interested automated mechanism design and implications for optimal combinatorial auctions. In *Proceedings of the ACM Conference on Electronic Commerce (ACM-EC)*, pages 132–141, 2004.

[DH06] Varsha Dani and Thomas P. Hayes. Robbing the bandit: Less regret in online geometric optimization against an adaptive adversary. In *Proc. 16th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 937–943, 2006.

[EYFKT92] Ran El-Yaniv, Amos Fiat, Richard M. Karp, and G. Turpin. Competitive analysis of financial games. In *Proceedings of the 33rd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 327–333, 1992.

[FP03] Eric J. Friedman and David C. Parkes. Pricing wifi at starbucks: issues in online mechanism design. In *EC '03: Proceedings of the 4th ACM conference on Electronic commerce*, pages 240–241, New York, NY, USA, 2003. ACM Press.

[FS97] Y. Freund and R. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, 1997.

[FV97] Dean P. Foster and Rakesh V. Vohra. Calibrated learning and correlated equilibrium. *Games and Economic Behavior*, 1997.

[GHK+05] Venkatesan Guruswami, Jason D. Hartline, Anna R. Karlin, David Kempe, Claire Kenyon, and Frank McSherry. On profit-maximizing envy-free pricing. In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms (SODA)*, pages 1164–1173, Philadelphia, PA, USA, 2005. Society for Industrial and Applied Mathematics.

[Han57] J.F. Hannan. Approximation to Bayes risk in repeated play. In M. Dresher, A.W. Tucker, and P. Wolfe, editors, *Contributions to the Theory of Games*, volume III, pages 97–139. Princeton University Press, 1957.

[HKLR05] MohammadTaghi Hajiaghayi, Jeong Han Kim, Tom Leighton, and Harald Raecke. Oblivious routing in directed graphs with random demands. In *STOC '05: Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pages 193–201, New York, NY, USA, 2005. ACM Press.

[HKMP05] Mohammad T. Hajiaghayi, Robert D. Kleinberg, Mohammad Mahdian, and David C. Parkes. Online auctions with re-usable goods. In *Proceedings of the 6th ACM conference on Electronic commerce (EC)*, pages 165–174, New York, NY, USA, 2005. ACM Press.

[HKP04] Mohammad T. Hajiaghayi, Robert Kleinberg, and David C. Parkes. Adaptive limited-supply online auctions. In *Proceedings of the 5th ACM conference on Electronic commerce (EC)*, pages 71–80, New York, NY, USA, 2004. ACM Press.

[HMC00] Sergiu Hart and Andreu Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 2000.

[HS04] Benoit Hudson and Tuomas Sandholm. Effectiveness of query types and policies for preference elicitation in combinatorial auctions. In *International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 386–393, 2004.

[IJMT05] N. Immorlica, K. Jain, M. Mahdian, and K. Talwar. Click fraud resistant methods for learning click-through rates. In *Workshop on Internet and Network Economics (WINE)*, 2005.

[KKLO03] S. Kakade, M. Kearns, J. Langford, and L. Ortiz. Correlated equilibria in graphical games. In *ACM Conference on Electronic Commerce*, 2003.

[KKO+04] S. Kakade, M. Kearns, L. Ortiz, R. Pemantle, and S. Suri. Economic properties of social networks. In *NIPS*, 2004.

[KLS01] M. Kearns, M. Littman, and S. Singh. Graphical models for game theory. In *Proceedings of the Conference on Uncertainty in Artificial Intelligence (UAI)*, 2001.

[KO02] M. Kearns and L. Ortiz. Nash propagation for loopy graphical games. In *NIPS*, 2002.

[KS06] M. Kearns and S. Suri. Networks preserving evolutionary stability and the power of randomization. In *ACM Conference on Electronic Commerce*, 2006.

[KV03] Adam Kalai and Santosh Vempala. Efficient algorithms for on-line optimization. In *Proceedings of the The 16th Annual Conference on Learning Theory*, pages 26–40, 2003.

[LN00] Ron Lavi and Noam Nisan. Competitive analysis of incentive compatible on-line auctions. In *Proceedings of the ACM Conference on Electronic Commerce (ACM-EC)*, pages 233–241, Minneapolis, MN, 2000.

[LP04] Sebastién Lahaie and David Parkes. Applying learning algorithms to preference elicitation. In *Proceedings of the ACM Conference on Electronic Commerce (ACM-EC)*, 2004.

[LW94] N. Littlestone and M. K. Warmuth. The weighted majority algorithm. *Information and Computation*, 108(2):212–261, 1994.

[MB04] H.B. McMahan and A. Blum. Online geometric optimization in the bandit setting against an adaptive adversary. In *Proceedings of the 17th Annual Conference on Learning Theory (COLT)*, pages 109–123, 2004.

[Moo92] J. Moore. Implementation in environments with complete information. In J-J Laffont, editor, *Advances in Economic Theory: The Proceedings of the Congress of the Econometric Society.* Cambridge University Press, 1992.

[MR88] J. Moore and R. Repullo. Subgame perfect implementation. *Econometrica*, 56:1191–1220, 1988.

[Rae02] Harald Raecke. Minimizing congestion in general networks. In *FOCS '02: Proceedings of the 43rd Symposium on Foundations of Computer Science*, pages 43–52, Washington, DC, USA, 2002. IEEE Computer Society.

[San03] Tuomas Sandholm. Automated mechanism design: A new application area for search algorithms. In *International Conference on Principles and Practice of Constraint Programming*, pages 19–36, Cork, Ireland, 2003.

[TW02] Eiji Takimoto and Manfred K. Warmuth. Path kernels and multiplicative updates. In *Proceedings of the 15th Annual Conference on Computational Learning Theory*, Lecture Notes in Artificial Intelligence. Springer, 2002.

[vAD04] Luis von Ahn and Laura Dabbish. Labeling images with a computer game. In *ACM CHI*, pages 319–326, 2004.

[vAGK+06] Luis von Ahn, Shiry Ginosar, Mihir Kedia, Ruoran Liu, and Manuel Blum. Improving accessibility of the web with a computer game. In *ACM CHI Notes*, 2006.

[vAKB06] Luis von Ahn, Mihir Kedia, and Manuel Blum. Verbosity: A game for collecting common-sense facts. In *ACM CHI Notes*, 2006.

[vALB06] Luis von Ahn, Ruoran Liu, and Manuel Blum. Peekaboom: A game for locating objects in images. In *ACM CHI*, 2006.

[ZBS03] Martin Zinkevich, Avrim Blum, and Tuomas Sandholm. On polynomial-time preference elicitation with value queries. In *ACM Conference on Electronic Commerce*, 2003.

[Zin03] Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th International Conference on Machine Learning*, pages 928–936, 2003.