Stochastic and Contingent-Payment Auctions

Christopher Meek
Microsoft Research
Redmond, WA 98052
meek@microsoft.com

David Maxwell Chickering
Microsoft Research
Redmond, WA 98052
dmax@microsoft.com

David Bruce Wilson
Microsoft Research
Redmond, WA 98052
dbwilson@microsoft.com

ABSTRACT

We describe a family of stochastic auctions that are incentive-compatible (i.e., "truthful"), and which generalize Vickrey auctions. We describe situations in which such stochastic auctions are useful. We then discuss contingent-payment auctions and their incentive-compatible counterparts. Finally, we discuss how to utilize both reinforcement-learning techniques and stochastic auctions to learn probabilities for contingencies.

Keywords

Stochastic mechanisms, auctions, incentive-compatible auctions, contingent payments.

1. INTRODUCTION

Contingent-payment auctions are widely used in online advertising. Such auctions are used for sponsored search advertising and for contextual advertising. In both of these examples, advertisers bid to show an advertisement to a user. For the sponsored search auctions, the ad is presented on a search results page that matches some criterion (e.g., the search query contained the phrase "plasma TV"). Similarly, for contextual advertising, the ad is presented to the user on a "content" page that matches some criterion (e.g., the web-page contains "investment advice"). In both of these examples, the contingency is typically a click on the advertisement, and there is uncertainty regarding this contingency. Two natural ranking or allocation rules for advertisements in this scenario are "by bid", in which one allocates the impressions to the advertisers with the highest bids, and "by revenue", where one allocates the impressions to the advertisers with the highest expected revenue (their payment or bid multiplied by an estimate of the probability of the contingency).

One issue with deterministic allocation rules (e.g., "by bid") in repeated auctions is that the auction can be become prone to vindictive bidding by antisocial participants. In Vickrey auctions, the price that the winner is charged is determined by the bids of the other participants. An antisocial participant may figure out a competitor's bid, and then deliberately bid a large value to increase the price for the winner, while not bidding high enough to be the winner himself (Brandt & Weiß 2001). While vindictive bidding may initially increase revenue for the auctioneer, the presence of vindictive bidding can encourage the participants to bid less than their true value, which reduces revenue.

If one uses a stochastic auction for allocation then the auction is typically no longer efficient; that is, the items are not allocated to the bidders with the highest value. Nonetheless, there are a number of reasons for considering stochastic auctions, especially in the repeated auction setting. These include: (1) increasing the diversity of bidders who are allocated items, (2) allowing bidders with low bids to win occasionally, (3) limiting vindictive bidding (discussed further in the next section), and (4) adapting the auction (e.g., learning the click-through rates, discussed in section 3). Several of these considerations could potentially lead to greater participation in the auction. In addition, using a stochastic incentive-compatible auction can improve revenue as compared to a deterministic incentive-compatible auction.

We describe allocation rules that include stochastic mechanisms inspired by reinforcement learning algorithms for exploration and reduction of uncertainties associated with contingent events, and give pricing rules for these stochastic contingent-payment auctions that make them incentive-compatible ("truthful"). These stochastic auctions include as a special case the usual Vickrey auctions with by-bid and by-revenue allocations.

2. STOCHASTIC AUCTIONS

Allocation rules used in auctions are typically deterministic and only resort to stochastic allocation in the event of identical bids. In this section, we describe a family of stochastic allocation rules with incentive-compatible pricing rules that generalize Vickrey auctions.

Consider a stochastic auction for k items with N bidders. We denote the set of bidders by $\mathcal{N} = \{1, \ldots, N\}$ and associate with each bidder i a bid $b_i \in [0, \infty)$ and a private value $v_i \in [0, \infty)$. We assume that there is either no reserve price or that the bids and values are given in excess of the reserve price (i.e., player i's bid is $b_i + r$ and value is $v_i + r$ where r is the reserve price). We denote the collection of all bids by $\mathbf{b}$, the collection of bids excluding bidder i by $\mathbf{b}_{-i}$, and the collection of all values by $\mathbf{v}$.

An *allocation* of the k items is a vector $\mathbf{a} \in \{0, 1, \ldots\}^N$ that indicates how many items are allocated to each bidder, where $\sum_i a_i = \min(k, N)$. Note that the sum is equal to the minimum of k and N because all bidders have bids at or above the reserve price. We concentrate on *exclusive allocations* in which no bidder can obtain more than one item, that is, $\mathbf{a} \in \{0,1\}^N$. The set of exclusive allocations is $\mathcal{A} = \{\mathbf{a} \in \{0,1\}^N : \sum_i a_i = \min(k, N)\}$. A *trivial* auction is one in which there are not more bidders than items (i.e., $N \leq k$).

A *stochastic auction* (or stochastic mechanism) $(\mathcal{B}, \pi, \mu)$ has the following components: (1) a set of possible messages, one from each player $\mathcal{B} = \mathcal{B}_1 \times \cdots \times \mathcal{B}_N$, (2) an allocation rule $\pi : \mathcal{B} \to p_{\mathcal{A}}$ from messages to a probability distributions over allocations, and (3) a pricing rule $\mu : \mathcal{B} \times \mathcal{A} \to \mathbb{R}^N$ that maps each possible bid vector and allocation to a price for each bidder. A *direct auction* is an auction in which the message from each player is precisely their bid and thus $\mathcal{B} = \mathbb{R}^N$.

2.1 Incentive-Compatible Stochastic Auctions

A direct stochastic auction is *incentive compatible* (or "truthful") if each bidder can maximize their *expected payoff* by bidding their value (see, e.g., Milgrom 2004), or more formally, if for each bidder i and for all possible bids of the other players $\mathbf{b}_{-i}$, the expected value of $v_i - \mu_i((\mathbf{b}_{-i}, b_i), a)$ is maximized by choosing $b_i = v_i$. We say that the auction is *strictly* incentive compatible if for each bidder i and for all possible bids of the other players $\mathbf{b}_{-i}$, the bid $b_i = v_i$ is the *only* bid for player i that maximizes $v_i - \mu_i((\mathbf{b}_{-i}, b_i), a)$. The usual Vickrey auction with the highest bidder winning the auction is incentive compatible, but it is not strictly incentive compatible, because an antisocial participant i, if he knows the other bids $\mathbf{b}_{-i}$, might bid $b_i > v_i$ with no loss of revenue. Strictly incentive-compatible auctions do not rule out vindictive bidding altogether, but any bidder who chooses $b_i \neq v_i$ will have strictly less expected revenue.

Define $p^i_{\mathbf{b}_{-i}}(b_i)$ to be the probability that bidder i wins exactly one item when bidding b_i and the other players bid $\mathbf{b}_{-i}$. In other words,

$$p^i_{\mathbf{b}_{-i}}(b_i) = \sum_{\mathbf{a}: a_i = 1} \pi_{\mathbf{b}_{-i}, b_i}(\mathbf{a}).$$

We say that an allocation rule π is non-decreasing (or strictly increasing) if, for all bidders i and all bids from other bidders $\mathbf{b}_{-i}$, $p^i_{\mathbf{b}_{-i}}(b_i)$ is non-decreasing (or strictly increasing). When the allocation rule π is non-decreasing (or strictly increasing), we present a pricing rule that is incentive compatible (or strictly incentive compatible).

For convenience, let us assume for the moment that $p^i_{\mathbf{b}_{-i}}(b_i)$ is differentiable. We define the *condex* pricing rule for a non-decreasing allocation rule π as follows: bidder i's price given bids $\mathbf{b}$ is

$$\mu_i(\mathbf{b}) = \frac{\int_0^{b_i} s \frac{dp^i_{\mathbf{b}_{-i}}(s)}{ds}\, ds}{p^i_{\mathbf{b}_{-i}}(b_i)}. \tag{1}$$

Note that in the event that there are not more bidders than items (i.e., $N \leq k$), if every participant always gets an item, then $\mu_i(\mathbf{b}) = 0$ regardless of the bids.

The name "condex" is short for conditional expectation and is due to the following interpretation: the price $\mu_i(\mathbf{b})$ is the expected value of the minimum winning bid for bidder i given that the bidder i wins the auction with a bid of b_i (and given the remaining bids $\mathbf{b}_{-i}$). From this we see that

$$\mu_i(\mathbf{b}) \leq b_i.$$

THEOREM 1. *If a direct stochastic auction has a non-decreasing (or strictly increasing) allocation rule and the corresponding condex pricing rule, then the stochastic auction is incentive compatible (or strictly incentive compatible).*

Proof: Player i wins the auction with probability $p^i_{\mathbf{b}_{-i}}(b_i)$ for a profit of $(v_i - \mu_i(\mathbf{b}_{-i}, b_i))$, and therefore has expected revenue

$$\begin{aligned}
&(v_i - \mu_i(\mathbf{b}_{-i}, b_i))\; p^i_{\mathbf{b}_{-i}}(b_i) \\
&= \left(v_i - \frac{\int_0^{b_i} s \frac{dp^i_{\mathbf{b}_{-i}}(s)}{ds}\, ds}{p^i_{\mathbf{b}_{-i}}(b_i)} \right) p^i_{\mathbf{b}_{-i}}(b_i) \\
&= v_i p^i_{\mathbf{b}_{-i}}(b_i) - \int_0^{b_i} s \frac{dp^i_{\mathbf{b}_{-i}}(s)}{ds}\, ds.
\end{aligned}$$

If player i bids $b_i \neq v_i$, then the expected change in revenue (*vis-à-vis* bidding v_i) is

$$\int_{v_i}^{b_i} (v_i - s) \frac{dp^i_{\mathbf{b}_{-i}}(s)}{ds}\, ds,$$

which is nonpositive, and is furthermore strictly negative if the allocation rule π is strictly increasing. □

If one uses a deterministic allocation rule for allocating the items then $p^i_{\mathbf{b}_{-i}}(x)$ is non-decreasing but is not continuous and in particular does not have a derivative. One can relax this differentiability assumptions through the use of measure theory where one replaces the Lebesgue integrals above with their measure theoretic counterparts:

$$\mu_i(\mathbf{b}) = \frac{\int_0^{b_i} s\, dp^i_{\mathbf{b}_{-i}}(s)}{p^i_{\mathbf{b}_{-i}}(b_i)}.$$

When this is done, the auction with the deterministic allocation rule in which the k items are allocated to the k highest bidders and its corresponding condex pricing rule is equivalent to the standard Vickrey multiple-item auction (see, e.g., Krishna 2002, page 171 for a discussion of multiple-item Vickrey auctions).

Example 1: Consider a stochastic auction for $k = 1$ item to N bidders. We allocate the item to bidder i with probability $p^i_{\mathbf{b}_{-i}}(b_i) = b_i^\beta / \sum_j b_j^\beta$, where b_j^β is bidder j's bid to the β^{th} power. As $\beta \to \infty$ the stochastic assignment rule approaches the deterministic assignment to the bidder with the highest bid. We give the condex pricing rule for a few choices of β. For convenience, let $x = b_i$ and $y^\beta = \sum_{j \neq i} b_j^\beta$.

When $\beta = 1$, we may substitute $p^i_{\mathbf{b}_{-i}}(x) = x/(x+y)$ in (1) and apply some basic algebra and calculus to obtain the pricing rule

$$\mu_i(\mathbf{b}_{-i}, b_i = x) = y(-1 + (1 + y/x)\log(1 + x/y)).$$

We note that the pricing rules for $\beta = 1/2$ is

$$\mu_i(\mathbf{b}_{-i}, b_i = x) = \sqrt{xy} + 2y - 2y(1 + \sqrt{y/x})\log(1 + \sqrt{x/y}),$$

and the pricing rule for $\beta = 2$ is

$$\mu_i(\mathbf{b}_{-i}, b_i = x) = -y^2/x + y(1 + y^2/x^2)\arctan(x/y).$$

More generally, the pricing rule for $\beta > 0$ can be expressed in terms of the hypergeometric function ${}_2F_1$:

$$\mu_i(\mathbf{b}_{-i}, b_i = x) = -x\frac{y^\beta}{x^\beta} + x\Big(1 + \frac{y^\beta}{x^\beta}\Big) F\left(\begin{matrix} 1,\ 1/\beta \\ 1 + 1/\beta \end{matrix} \;\middle|\; -\frac{x^\beta}{y^\beta}\right).$$

2.2 Stochastic Prices

For repeated auctions it may sometimes be desirable to use stochastic prices. If the actual price never exceeds the bid, and the expected price is given by the condex pricing rule (1), then the auction remains incentive compatible. One reason to use a stochastic price is that it may be easier to select a random price with the desired expected value than to actually compute the expected value.

In the course of making an allocation, the auctioneer flips some coins that are summarized in the random variable U and then computes the allocation as a deterministic function ϕ of U and the bids $\mathbf{b}$: $\mathbf{a} = \phi(U, \mathbf{b})$. A *realizably monotone* allocation rule is one with the following property: for every U, for every bidder i, and for every set of bids $\mathbf{b}_{-i}$, the allocation a_i to bidder i is monotone increasing in b_i. In other words, an allocation rule is realizably monotone if bidding more can only help the bidder win. If the allocation rule is realizably monotone, then, for a given U and $\mathbf{b}_{-i}$ there is some critical bid b_i^* above which the bidder wins. We can define the price $\mu_i(U, \mathbf{b}_{-i}, b_i)$ for player i to be this critical bid b_i^*:

$$\mu_i(U, \mathbf{b}_{-i}, b_i) = \inf\{s \in \mathbb{R} \;:\; \phi_i(U, \mathbf{b}_{-i}, s) = 1\}. \qquad (2)$$

In expectation, this price is a condex pricing rule for a realizably monotone allocation rule— that is, one can show that $\mathbb{E}[\mu_i(U, \mathbf{b}_{-i}, b_i) | \phi_i(U, \mathbf{b}_{-i}, b_i) = 1]$ satisfies (1)— so the resulting stochastic auction with stochastic prices is incentive-compatible.

Remark: One feature of using stochastic prices is that, conditional on the auctioneer's random coins U and the other bids $\mathbf{b}_{-i}$, bidder i's price is independent of b_i. In contrast, if deterministic prices are used for the stochastic auction, the price for bidder i depends on b_i.

Example 2: Consider a stochastic auction for k items to N bidders. The auctioneer picks N independent positive random numbers $u_1, \ldots, u_N$ distributed in some fashion. The distribution of u_i may depend on i. Let i_j be the identity of the bidder with the j^{th} largest value of $b_i u_i$. The winners of the auction are $i_1, \ldots, i_k$. Then the stochastic pricing rule assigns a price of

$$\mu_i(U, \mathbf{b}_{-i}, b_i) = \frac{b_{i_{k+1}} u_{i_{k+1}}}{u_i} \qquad (3)$$

to a winning bidder i and zero to every non-winning bidder. Note that $\mu(\mathbf{b}_{-i}, b_i) \leq b_i$ if bidder i wins the auction. Also note that this price is easier to compute than, say, a hypergeometric function.

Remark: We recover the Vickrey auction when the u_i's deterministically take the same value.

Example 1 revisited: Recall that in example 1, bidder i wins with probability $b_i^\beta / \sum_j b_j^\beta$. This allocation scheme can be realized by using N independent standard exponential random variables $X_1, \ldots X_N$, and picking the i which minimizes X_i / b_i^β, or equivalently, the i which maximizes $b_i X_i^{-1/\beta}$. We see then that this allocation scheme is a special case of Example 2 where $k = 1$ and $u_i = X_i^{-1/\beta}$, and we may use the stochastic price given in (3).

2.3 Vindictive Bidding and Expected Revenue

In a one-time incentive-compatible auction, bids among competitors need not accurately reflect the bidders' private values for the items. For instance, consider a second-price sealed-bid auction for one item in which there are two bidders $\{1, 2\}$ with private values $v_1 = 11$ and $v_2 = 1$. Let us further assume that bidder 2 knows that bidder 1 values the item highly and will bid more than 9 in the auction. If bidder 2 is a vindictive bidder then bidder 2 can safely bid 9 without risk. If bidder 1 bids his value then he will pay 9 for the item. If we use the simple lottery scheme describe in Example 1 above ($\beta = 1$), there is a risk to the vindictive bidder that he will win the auction and have to pay more than he values the item.

Table 1 provides three bidding scenarios. In all scenarios, bidder 1 bids truthfully and has the same value, and bidder 2 has the same value (but different than bidder 1). The three scenarios differ in bidder 2's bidding behavior: he either bids much above his own value (vindictive), slightly above his own value (mean), or his value (truthful). For each bidding scenario, the prices are determined using the stochastic auction from the example above and the expected revenue for the seller is provided.

In the vindictive scenario, bidder 2 actually has an expected loss of 1.2 with the stochastic auction. In the mean scenario, bidder 2 reduces his bid to 2.1 so that his expected loss is zero.

While this stochastic auction doesn't completely remove the possibility of vindictive bidding, it does reduce the incentive for this behavior. Using alternative stochastic allocation rules (such as $\beta < 1$) can further reduce this incentive.

It is interesting to compare the revenue of stochastic auctions with alternative deterministic auctions. For this example it is natural to compare with the second-price Vickrey auction. As noted earlier, auctions from the family of stochastic auctions described above approach a Vickrey auction as $\beta \to \infty$. For smaller values of β the behavior is interesting. Consider the expected revenue from the truthful scenario. In this case, the stochastic auction has an expected revenue that is 60% higher than a Vickrey auction. However, if the bids of the two bidders are more tightly coupled, for instance in the vindictive scenario, the expected revenue drops dramatically. This is not surprising as the allocation rule is far from efficient for $\beta = 1$.

3. CONTINGENT-PAYMENT AUCTIONS

A contingent-payment auction is the same as a regular auction except that a winning bidder only pays for the item when a bidder-specific contingency occurs; otherwise, the item is free. Contingent-payment auctions are used to sell impressions of advertisements on the web, typically using a pay-per-click model in which the advertiser-specific contingency is that the displayed advertisement is clicked. For web advertising, contingent-payment auctions also work for other types of arrangements such as pay-per-acquisition and

Scenario	Values	Bids	Prices	$\mathbb{E}$[revenue]
vindictive	$v_1 = 11; v_2 = 1$	$b_1 = 11; b_2 = 9.0$	$\mu_1 = 4.0; \mu_2 = 3.6$	3.9
mean	$v_1 = 11; v_2 = 1$	$b_1 = 11; b_2 = 2.1$	$\mu_1 = 2.5; \mu_2 = 1.0$	2.2
truthful	$v_1 = 11; v_2 = 1$	$b_1 = 11; b_2 = 1.0$	$\mu_1 = 1.7; \mu_2 = 0.5$	1.6

Table 1: Bidding scenarios for a two bidder one item stochastic auction using the allocation scheme from Example 1 with $\beta = 1$ and deterministic prices.

pay-per-call. The special case of pay-per-impression is the contingency that happens with probability one.

In this section, we extend the results about incentive-compatible payments from the previous section to contingent-payment auctions, and we discuss methods for learning contingency probabilities when those probabilities are used by the allocation rule. For the remainder of the section, we consider auctions in which there is a *single* contingency (e.g., pay-per-click); more complicated scenarios allow bidders to submit different bids based on the outcome of multiple contingencies. For example, an advertiser might submit a bid for which they agree to pay \$1 each time an advertisement is clicked, and they agree to pay an additional \$4 if a sale is made as a result of the click.

3.1 Incentive-Compatible Auctions

We let v_i° denote bidder i's value for the item if the contingency c_i does not occur, and $v_i^\checkmark$ be the additional value if the contingency does occur. Let v_i denote bidder i's *a priori* value for the item (without knowing whether or not the contingency will occur). We have $v_i = v_i^\circ + q_i(c_i)v_i^\checkmark$, where $q_i(c_i)$ is bidder i's estimate of the probability $q(c_i)$ that the contingency c_i occurs. Since the bidder only pays if the contingency c_i occurs, the average value received per payment is $v_i^\circ/q_i(c_i) + v_i^\checkmark$. We say that the contingent-payment auction is incentive compatible if each bidder i can maximize their expected payoff by bidding $v_i/q_i(c_i) = v_i^\circ/q_i(c_i) + v_i^\checkmark$.

In some scenarios, the bidder may only value the item if the contingency occurs (e.g., an advertiser may only value impressions that get clicked). In these scenarios, we have

$$v_i = q_i(c_i)v_i^\checkmark,$$

and thus incentive compatibility requires the bidder to have maximum expected return whenever $b_i = v_i^\checkmark$, which eliminates the need for the bidder to estimate a contingency probability.

As we show in the following corollary, the results from the previous section apply to contingent-payment auctions using a slightly modified version of the argument in the proof of Theorem 1.

Corollary 1. *If a direct stochastic contingent-payment auction has a non-decreasing (or strictly increasing) allocation rule and the corresponding condex rule, then the stochastic auction is incentive compatible (or strictly incentive compatible).*

Proof: As in the proof of Theorem 1, we express the expected return to the bidder for submitting a bid of b_i. The only change is that the value is received and the payment is made only when the contingency occurs:

$$\mathbb{E}[\text{profit for bidder } i] = (v_i - q_i(c_i)\mu_i(\mathbf{b}_{-i}, b_i))p^i_{\mathbf{b}_{-i}}(b_i).$$

Using the same approach as in the proof of Theorem 1, it is easy to show that this expectation is maximized when $b_i = v_i/q_i(c_i)$, and if we have a strictly increasing allocation rule, that this maximum is unique. □

It is natural for auctioneer to take into account his or her estimates $q_\star(c_i)$ of the contingency probabilities $q(c_i)$ when making the allocation. We may for instance modify the scheme from Example 1 so that bidder i wins with probability

$$p^i_{\mathbf{b}_{-i}}(b_i) = \frac{q_\star(c_i)b_i}{\sum_j q_\star(c_j)b_j}.$$

For convenience we let $x = b_i$ and $y = \sum_{j \neq i} q_\star(c_j)b_j/q_\star(c_i)$. Then $p^i_{\mathbf{b}_{-i}}(b_i) = x/(x+y)$, and the same derivation from before yields the pricing rule

$$\mu_i(\mathbf{b}_{-i}, b_i) = y(-1 + (1 + y/x)\log(1 + x/y)).$$

Another natural choice is to use the allocation and pricing rule from Example 2, and take for instance $u_i = q_\star(c_i) \times \text{Uniform}(0,1)^\beta$. In the limit when $\beta \to 0$, this yields the "by revenue" allocation rule with Vickrey pricing.

3.2 Learning Contingency Probabilities

In this section, we consider situations where bidders participate in a single auction multiple times, and therefore the auction system can use repeated observations about contingencies to estimate the contingency probability $q(c_i)$ for each bidder i. When the allocation rule in a contingent-payment auction depends on the probability of each bidder contingency, it is important for the auction system to have an accurate estimate of these probabilities. For example, in a (non-stochastic) second-price auction for selling web impressions to advertisers, the allocation rule might be to award impressions to those advertisers with highest expected value per impression: $b_i q_\star(\text{AdGetsClick})$. If, for a particular advertiser, the auction system underestimates the true click rate of the advertisement by too much, there is the risk that the advertisement will never be shown and thus the system will never be able to update the estimate with new data.

In general, if a non-stochastic contingent-payment auction system can only update the estimate for $q(c_i)$ if bidder i wins the auction, then the system is vulnerable to the *starvation problem*: if the system has inaccurate estimates of contingency probabilities, or if contingency probabilities can change over time, then an underestimate can result with a bidder never winning and thus the underestimate cannot be corrected. Stochastic contingent-payment auctions are less vulnerable to the starvation problem whenever the allocation rule assigns *some* probability to each bidder winning; this way, an underestimate of a contingency probability will eventually be corrected.

Intuitively, the auction system should trade off (1) the immediate reward of optimally allocating items according to the current probability estimates, and (2) the future reward of gaining better estimates by performing a non-optimal allocation. The starvation problem is well studied within the reinforcement-learning community; we discuss two well-known algorithms and describe how they can be applied to an auction system.

Let $\theta_i = q(c_i)$ denote the true probability that the contingency will occur if bidder i wins the auction. Since we are learning θ_i, we assume that θ_i has a posterior distribution governed by a beta distribution:

$$\mathrm{pdf}(\theta_i) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} {\theta_i}^{a-1}(1-\theta_i)^{b-1}.$$

A standard point estimate for θ_i is the expected value $a/(a+b)$ of this distribution. There are a number of advantages to having a distribution over θ_i. First, we have an explicit representation for our uncertainty of θ_i that can be used by a reinforcement-learning algorithm. Second, assuming $\mathrm{pdf}(\theta_i)$ is a beta distribution allows us to easily update the distribution whenever bidder i wins: if the contingency occurs, we increment a by one; otherwise, we increment b by one. Finally, we can initialize $\mathrm{pdf}(\theta_i)$ for new bidders using prior knowledge about both the expected value (e.g., we know the click rates of similar advertisements) and the variance (e.g., there is a wide variety of click rates of similar advertisements).

Using the *interval-estimation* algorithm (Kaelbling, 1993), we calculate a 95% confidence interval (or more generally, any confidence interval) for θ_i. Then, instead of using the expected value of θ_i to allocate and price items, we use the high bound on the confidence interval. The result is that we "inflate" our estimates of contingency probabilities about which we are uncertain. As a bidder wins more and more auctions, the variance of θ_i decreases, and thus the high bound will tend toward the expectation of θ_i. This algorithm can still lead to the starvation problem if θ_i can change over time. We note that Kitts and Leblanc (2004) use the interval-estimation algorithm for optimizing bidder behavior in continuous pay-per-click auctions.

In an alternative reinforcement-learning algorithm, originally due to Thompson (1933), we choose the winners for the auction based on the probability that they would win if we knew the true contingency probabilities. For example, in a single-item auction in which the winner is the one for which $b_i q(c_i)$ is largest, we use $\mathrm{pdf}(\theta_i)$ to determine the probability that $b_i \cdot \theta_i$ is larger than for any other bidder. This algorithm changes the deterministic allocation rule into a stochastic one, although as the variances of $\mathrm{pdf}(\theta_i)$ decrease it converges to the original (deterministic) rule. To mitigate the problem of contingency probabilities changing over time, we can artificially keep the variance of $\mathrm{pdf}(\theta_i)$ from getting too low by, for example, renormalizing a and b such that their sum never exceeds some constant.

4. REFERENCES

[1] F. Brandt and G. Weiß. Antisocial agents and Vickrey auctions. In J.-J. Meyer and M. Tambe, editors, *Pre-proceedings of the Eighth International Workshop on Agent Theories, Architectures, and Languages (ATAL-2001)*, pages 120–132, 2001.

[2] L. P. Kaelbling. *Learning in Embedded Systems*. The MIT Press, Cambridge, MA, 1993.

[3] B. Kitts and B. Leblanc. Optimal bidding on keyword auctions. *Electronic Markets*, 14(3):186–201, 2004.

[4] V. Krishna. *Auction Theory*. Academic Press, New York, New York, 2002.

[5] P. Milgrom. *Putting Auction Theory to Work*. Cambridge University Press, Cambridge, UK, 2004.

[6] W. R. Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometricka*, 25:285–294, 1933.