Risk-Averse Auction Agents

Yaxin Liu
College of Computing
Georgia Tech
Atlanta, GA 30332-0280
yxliu@cc.gatech.edu

Richard Goodwin
IBM T.J. Watson
P.O. Box 218, Route 134
Yorktown Heights, NY 10598
rgoodwin@us.ibm.com

Sven Koenig
College of Computing
Georgia Tech
Atlanta, GA 30332-0280
skoenig@cc.gatech.edu

ABSTRACT

Auctions are an important means for purchasing material in the era of e-commerce. Research on auctions often studies them in isolation. In practice, however, auction agents are part of complete supply-chain management systems and have to make the same decisions as their human counterparts. To address this issue, we generalize results from auction theory in three ways. First, auction theory provides the optimal bidding function for the case where auction agents want to maximize the expected profit. Since companies are often risk-averse, we derive a closed form of the optimal bidding function for auction agents that maximize the expected utility of the profit for concave exponential utility functions. Second, auction theory often assumes that auction agents know the bidder's valuation of an auctioned item. However, the valuation depends on how the item can be used in the production process. We therefore develop theoretical results that enable us to integrate our auction agents into production-planning systems to derive the bidder's valuation automatically. Third, auction theory often assumes that the probability distribution over the competitors' valuations of the auctioned item is known. We use simulations of the combined auction- and production-planning system to obtain crude approximations of these probability distributions automatically. The resulting auction agents are part of a complete supply-chain management system and seamlessly combine ideas from auction theory, utility theory, and dynamic programming.

Categories and Subject Descriptors

I.2.8 [**Computing Methodologies**]: Artificial Intelligence—*Problem Solving, Control Methods, and Search*

General Terms

Algorithms, Economics

Keywords

Auctions, E-Commerce, Supply-Chain Management, Risk Aversion

1. INTRODUCTION

Our goal is to develop agents that are able to make decisions for far more realistic planning objectives than is currently possible. We do this in the context of auction planning. Auctions are an important means for purchasing material in the era of e-commerce. Consequently, agents that automate or support human decision making need to be able to decide whether to participate in auctions and how much to bid. If the auction is won, a company might be able to save money compared to producing the material in house. If the auction is lost, however, the company might incur large penalties for not being able to deliver orders on time. Companies are therefore often risk-averse (conservative) when making these decisions and bid high to increase the chance of winning the auction. In this paper, we investigate how to build auction agents for high-stake single-instance decision situations like these. High-stake decision situations are situations in which large wins or losses are possible. Single-instance decision situations are situations that are faced only once. People typically do not maximize the expected profit in high-stake single-instance decision situations. Consider, for example, that you can participate once in one (and only one) of the following two lotteries without any cost to you:

Choices	Probability	Profit	Expected Profit	Utility	Expected Utility
Choice 1	50 percent	$10,000,000	$5,000,000	0.95	0.475
	50 percent	$0		0.00	
Choice 2	100 percent	$4,500,000	$4,500,000	0.74	0.740

When people have to decide whether they would like to get 4,500,000 dollars for sure (Choice 2) or get 10,000,000 dollars with fifty percent probability and nothing otherwise (Choice 1), many people prefer Choice 2 although its expected profit is clearly lower. They are risk-averse, that is, willing to accept a decrease in the expected profit to reduce the variance and thus the possibility of only a small profit. The recommendations of agents should reflect the risk aversion of people correctly. After all, agents make suggestions for how to act and should make the same suggestions that people would have made themselves, otherwise the agents would not be very helpful.

Utility theory [4, 16] explains why people often do not maximize the expected profit in high-stake single-instance decision situations. Utility theory suggests that they maximize expected utility, where the utility is a strictly monotonically increasing function of the profit. Nonlinear utility functions are necessary to model risk-averse people. Experiments show that utility functions in practice are often approximately logarithmic. We, however, use concave exponential utility function because of their advantageous

properties. The lottery example, for instance, can be explained as follows: Assume that a person has a concave exponential utility function that associates utility (here: pleasure) 0.00 with a wealth of 0 dollars, utility 0.74 with a wealth of 4,500,000 dollars, and utility 0.95 with a wealth of 10,000,000 dollars. Then the (expected) utility of getting 4,500,000 dollars for sure is 0.74, whereas the expected utility of getting 10,000,000 dollars with fifty percent probability is only 0.475. Thus, Choice 2 maximizes the expected utility for this person, which explains why this person chooses it over Choice 1. Other people can have other utility functions and thus arrive at a different conclusion.

In this paper, we use results from auction theory and utility theory to determine the optimal bidding strategy for auction agents that maximize the expected utility for a given concave exponential utility function, both for first- and second-price sealed-bid auctions in the symmetric independent private values model. To the best of our knowledge, only qualitative results were known so far about the resulting optimal bidding strategy. However, qualitative results are insufficient for building risk-averse auction agents. We therefore derive quantitative results about the optimal bidding strategy. We then use dynamic programming to show how one can integrate our auction agents into production-planning systems, which is important because auction theory often studies auctions in isolation but production decisions and results of previous auctions affect the behavior of auction agents in subsequent auctions. Finally, we use simulations of the combined auction- and production-planning system to obtain crude approximations of the probability distribution over the competitors' valuations of the auctioned item, resulting in auction agents that are part of a complete supply-chain management system and seamlessly combine ideas from auction theory, utility theory, and dynamic programming.

2. THE AUCTION MODEL

A suitable auction model for our purposes is the symmetric independent private values (SIPV) model [13]. In this model:

- only one item is for sale, and the seller is willing to sell it to the highest bidder for any positive price;

- the number of bidders $N \geq 2$ is known to all bidders;

- each bidder i knows their own valuation v_i for the auctioned item (that is, the difference in profit between owning and not owning the item);

- no bidder knows the competitors' valuations of the auctioned item, but these valuations are independent random variables drawn from a given continuous probability distribution $F(v)$ with density $f(v)$ over the nonnegative real-valued bids (valuation distribution), and this distribution is known to all bidders;

- and the bidders are indistinguishable.

The SIPV model has been studied in the context of two standard types of auctions, namely the first- and second-price sealed-bid auctions [13]. In these auctions, the bidders submit secret bids. The highest bid wins the item. The winner pays the highest bid ("first price") in the first-price

sealed-bid auction but only the highest losing bid ("second price") in the second-price sealed-bid auction.

Both kinds of auctions are modeled as non-cooperative games with incomplete information, often under the assumption that the behavior of a bidder is determined by a differentiable bidding function. A bidding function $b(\cdot)$ maps the bidder's valuation $v \geq 0$ of the auctioned item to their nonnegative bid $b(v) \geq 0$, resulting in a deterministic bidding strategy. The optimal bidding functions $b^*(\cdot)$ must be in a Bayes-Nash equilibrium, that is, no bidder can do better by changing their bidding function provided that the competitors do not change their bidding functions either. The optimal bidding functions of all bidders are the same due to symmetry in the SIPV model. Obviously, $b^*(0) = 0$. Furthermore, they are strictly monotonically increasing in the valuation for both first- and second-price sealed-bid auctions [13]. It follows immediately that the strategy that maximizes the expected profit for second-price sealed-bid auctions is to bid the bidder's valuation, simply because this strategy dominates all other strategies. Consequently, the optimal bidding function is the identity function. It is more complicated to determine the strategy that maximizes the expected profit for first-price sealed-bid auctions. The optimal bidding function is known to be [13]

$$ b^*(v) = v - \frac{\int_0^v F(t)^{N-1} dt}{F(v)^{N-1}}. \tag{1} $$

Note that the bids have to be smaller than the bidder's valuation for first-price sealed-bid auctions since bidding the valuation results in zero profit.

3. RISK AND UTILITY FUNCTIONS

We suggest that one can use utility theory to model the bidding strategies of risk-averse auction agents in a principled way. Utility theory is a subfield of decision theory that provides a normative framework for rational decision making under uncertainty [4, 16]. Its main claim is that every rational person has a strictly monotonically increasing utility function $u(\cdot)$ that maps profit v into real-valued utility $u(v)$, so that the person always prefers the alternative with the highest expected utility. The form of the utility function determines the risk attitude of the person. A person is risk-averse if they are willing to accept a decrease in the expected profit to reduce the variance and thus the possibility of a large loss. The risk-averse utility functions most often used in utility theory are probably the concave exponential utility functions

$$ u(v) = -\gamma^v, $$

where the parameter $0 < \gamma < 1$ determines the level of risk aversion [17]. If γ approaches one, the person is less and less risk-averse and, in the limit, maximizes the expected profit (under appropriate assumptions) [11]. As γ approaches zero, on the other hand, the person becomes more and more risk-averse. Thus, concave exponential utility functions can model a continuous spectrum of risk aversion. For example, they include the utility function that we used in the context of the lottery example.

4. RISK-AVERSE BIDDING

Research performed in the context of the SIPV model often assumes that auction agents want to maximize the expected profit [13]. We have argued that we can take risk aversion into account by assuming that they maximize the expected utility for a concave exponential utility function. Previous work has shown that the optimal bidding function for risk-averse auction agents remains strictly monotonically increasing in the valuation for both first- and second-price sealed-bid auctions [12]. Again, the strategy that maximizes the expected utility of the profit for second-price sealed-bid auctions is to bid the bidder's valuation, simply because this strategy still dominates all other strategies. Consequently, the optimal bidding function remains the identity function. Thus, the bids of risk-averse auction agents and auction agents that maximize the expected profit are identical for second-price sealed-bid auctions. However, this is not the case for first-price sealed-bid auctions. It is known that the bids of risk-averse auction agents in these auctions are higher than the bids of auction agents that maximize the expected profit [12]. While this result explains why risk-averse auction agents bid high to increase the chance of winning the auction, it is qualitative in nature and does not specify the optimal bidding function. In fact, [12] showed that the optimal bidding function, in general, is the solution of a nonlinear differential equation which can only be solved numerically. However, we show in the following that the optimal bidding function can be obtained in closed form for risk-averse auction agents with concave exponential utility functions in the context of first-price sealed-bid auctions in the SIPV model. We extend the symmetry assumptions of the SIPV model to include the assumption that all bidders have the same concave exponential utility function, and continue to assume that the optimal bidding function $b^*(\cdot)$ is differentiable and thus continuous. Under this assumption, the optimal bidding function is also strictly monotonically increasing [12]. Thus, it has a differentiable inverse that we denote by $\sigma(\cdot)$, and it holds that $\sigma(b) = v$ iff $b = b^*(v)$.

Derivation of the Optimal Bidding Function

We now derive the optimal bidding function for auction agents with concave exponential utility functions in the context of first-price sealed-bid auctions in the SIPV model. Our derivation parallels the one in [13] but maximizes the expected utility of the profit rather than the expected profit. We first determine the probability $\rho(b)$ with which the auction agent wins the auction when it bids b. Without loss of generality, we consider bidder 1. Any competitor $2 \leq i \leq N$ has valuation v_i and thus bids $b^*(v_i)$. Thus,

$$
\begin{aligned}
\rho(b) &= P\left(\max_{2 \leq i \leq N} b^*(v_i) < b\right) \\
&= P(b^*(v_2) < b \wedge \ldots \wedge b^*(v_N) < b) \\
&= \prod_{i=2}^{N} P(b^*(v_i) < b) = \prod_{i=2}^{N} P(v_i < \sigma(b)) \\
&= F(\sigma(b))^{N-1}.
\end{aligned} \tag{2}
$$

Later, we need the derivative of this probability with respect to b. It is

$$
\rho'(b) = (N-1)F(\sigma(b))^{N-2}f(\sigma(b))\sigma'(b). \tag{3}
$$

The expected utility $EU(v,b)$ of the profit obtained from bidding b for an item with valuation v is

$$
EU(v,b) = \rho(b)u(v-b) + (1-\rho(b))u(0). \tag{4}
$$

The optimal bid β maximizes $EU(v,b)$ and thus the derivative of $EU(v,b)$ with respect to b at β is zero. Thus,

$$
\left.\frac{\partial EU(v,b)}{\partial b}\right|_{b=\beta} = 0
$$
$$
\rho'(\beta)u(v-\beta) - \rho(\beta)u'(v-\beta) - \rho'(\beta)u(0) = 0
$$
$$
\rho'(\beta)\left[u(v-\beta) - u(0)\right] = \rho(\beta)u'(v-\beta). \tag{5}
$$

By substituting Equations (2) and (3) as well as $v = \sigma(\beta)$ into Equation (5) and then dividing both sides by $F(\sigma(\beta))^{N-2}$, we obtain

$$
\begin{aligned}
(N-1)&f(\sigma(\beta))\sigma'(\beta)\left[u(\sigma(\beta) - \beta) - u(0)\right] \\
&= F(\sigma(\beta))u'(\sigma(\beta) - \beta).
\end{aligned} \tag{6}
$$

This is a differential equation for $\sigma(\cdot)$, but we are interested in the differential equation for its inverse $b^*(\cdot)$. By substituting $\sigma(\beta) = v$, $\beta = b^*(v)$ and $\sigma'(\beta) = 1/b^{*\prime}(v)$, the relationship between the derivative of a function and the derivative of its inverse, into Equation (6), we obtain

$$
\begin{aligned}
(N-1)&f(v)\left[u(v - b^*(v)) - u(0)\right] \\
&= F(v)u'(v - b^*(v))b^{*\prime}(v).
\end{aligned} \tag{7}
$$

In general, this differential equation cannot be solved in closed form. However, we now show how one can obtain a closed form solution for concave exponential utility functions. In this case, Equation (4) simplifies to

$$
EU(v,b) = -\left[\rho(b)\gamma^{v-b} + (1-\rho(b))\right]. \tag{8}
$$

Similarly, since the derivative of the utility function with respect to v is $u'(v) = -(\ln\gamma)\gamma^v$, Equation (7) simplifies to

$$
(N-1)f(v)\left[-\gamma^{v-b^*(v)} + 1\right] = F(v)\gamma^{v-b^*(v)}(-\ln\gamma)b^{*\prime}(v)
$$

$$
(N-1)f(v)\left[\gamma^v - \gamma^{b^*(v)}\right] = F(v)\gamma^v(\ln\gamma)b^{*\prime}(v). \tag{9}
$$

We can solve Equation (9) with the substitution

$$
B(v) = \gamma^{-b^*(v)}F(v)^{N-1}. \tag{10}
$$

For Equation (10), it holds that

$$
\gamma^{b^*(v)} = \frac{F(v)^{N-1}}{B(v)} \tag{11}
$$

and also that

$$
b^*(v) = (N-1)\log_\gamma F(v) - \log_\gamma B(v). \tag{12}
$$

For Equation (10), it also holds that

$$
\begin{aligned}
B'(v) &= \gamma^{-b^*(v)}F(v)^{N-1}\left(-b^{*\prime}(v)\ln\gamma + (N-1)\frac{f(v)}{F(v)}\right) \\
&= B(v)\left(-b^{*\prime}(v)\ln\gamma + (N-1)\frac{f(v)}{F(v)}\right)
\end{aligned}
$$

and thus

$$
b^{*\prime}(v)\ln\gamma = (N-1)\frac{f(v)}{F(v)} - \frac{B'(v)}{B(v)}. \tag{13}
$$

By substituting Equations (11) and (13) into Equation (9), we obtain

$$(N-1)f(v)\left(\gamma^v - \frac{F(v)^{N-1}}{B(v)}\right)$$
$$= F(v)\gamma^v\left((N-1)\frac{f(v)}{F(v)} - \frac{B'(v)}{B(v)}\right).$$

By expanding both sides of the equation, canceling the first terms and the common factors on both sides, and using $[F(v)^{N-1}]' = (N-1)f(v)F(v)^{N-2}$, we obtain

$$(N-1)f(v)F(v)^{N-2} = \gamma^v B'(v)$$

$$B'(v) = \gamma^{-v}\left[F(v)^{N-1}\right]'$$

$$B(v) = C + \int_0^v \gamma^{-t} dF(t)^{N-1}. \tag{14}$$

Since $F(0) = 0$ and thus $B(0) = \gamma^{-b^*(0)}F(0)^{N-1} = 0$, it follows that $C = 0$. By substituting Equation (14) with $C = 0$ into Equation (12), we obtain a solution of the differential equation for $b^*(v)$ of the form

$$b^*(v) = (N-1)\log_\gamma F(v) - \log_\gamma \int_0^v \gamma^{-t} dF(t)^{N-1} \tag{15}$$

$$= v - \log_\gamma\left[1 + \frac{\ln\gamma}{F(v)^{N-1}}\int_0^v F(t)^{N-1}\gamma^{(v-t)}dt\right]. \tag{16}$$

This is the optimal bidding function for risk-averse auction agents with concave exponential utility functions in the context of first-price sealed-bid auctions in the SIPV model and thus the equivalent of Equation (1) for risk-averse auction agents. Formula (16) shows that the bids are always smaller than the bidder's valuation, as expected. It is also numerically more stable than Formula (15).

Properties of the Optimal Bidding Function

In the following we show that the optimal bidding function from Equation (15) has two desired properties in the limit. First, the bidding function approaches Equation (1) as the auction agent becomes less and less risk-averse and thus becomes more and more interested in maximizing the expected profit. This was to be expected. Second, the optimal bidding function approaches the identity function as the auction agent becomes more and more risk-averse, and the profit is zero in the limit. This is consistent with the qualitative results in [12].

Case 1

As γ approaches one, the auction agent becomes less and less risk-averse. To determine

$$\lim_{\gamma\to 1} b^*(v) = \lim_{\gamma\to 1}\frac{(N-1)\ln F(v) - \ln\int_0^v \gamma^{-t}dF(t)^{N-1}}{\ln\gamma},$$

note that both its numerator and denominator approach zero and we can thus apply L'Hôpital's rule to obtain

$$\lim_{\gamma\to 1} b^*(v) = \lim_{\gamma\to 1}\frac{-\frac{\int_0^v(-t)\gamma^{-t-1}dF(t)^{N-1}}{\int_0^v \gamma^{-t}dF(t)^{N-1}}}{1/\gamma}$$
$$= \frac{\int_0^v t\,dF(t)^{N-1}}{\int_0^v dF(t)^{N-1}} = v - \frac{\int_0^v F(t)^{N-1}dt}{F(v)^{N-1}},$$

which is the same as the optimal bidding function for an auction agent that maximizes the expected profit, as shown in Equation (1).
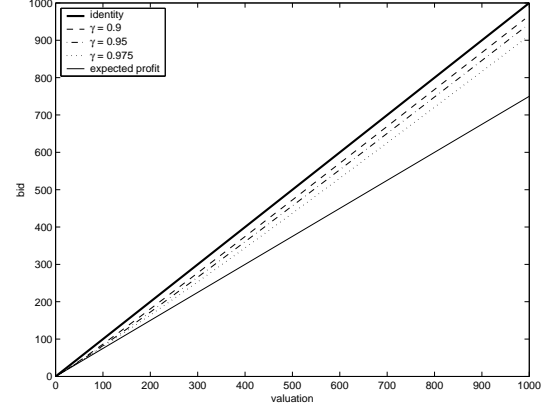

Figure 1: Optimal Bidding Functions

Case 2

As γ approaches zero, the auction agent becomes more and more risk-averse. This time, both the numerator and denominator approach minus infinity and we can thus continue to apply L'Hôpital's rule. We obtain

$$\lim_{\gamma\to 0+} b^*(v) = \lim_{\gamma\to 0+}\frac{-\frac{\int_0^v(-t)\gamma^{-t-1}dF(t)^{N-1}}{\int_0^v \gamma^{-t}dF(t)^{N-1}}}{1/\gamma}$$
$$= \lim_{\gamma\to 0+}\frac{\int_0^v t\gamma^{-t}dF(t)^{N-1}}{\int_0^v \gamma^{-t}dF(t)^{N-1}}$$
$$= \lim_{\gamma\to 0+}\frac{v\gamma^{-v}F(v)^{N-1} - \int_0^v F(t)^{N-1}(\gamma^{-t} - t(\ln\gamma)\gamma^{-t})dt}{\gamma^{-v}F(v)^{N-1} - \int_0^v F(t)^{N-1}(-\ln\gamma)\gamma^{-t}dt}$$
$$= \lim_{\gamma\to 0+}\frac{vF(v)^{N-1} - \int_0^v F(t)^{N-1}(\gamma^{v-t} - t(\ln\gamma)\gamma^{v-t})dt}{F(v)^{N-1} + \int_0^v F(t)^{N-1}(\ln\gamma)\gamma^{v-t}dt}.$$

As γ approaches zero, both integrals in the above expression are well-defined and strictly monotonic with respect to γ. As known from real analysis [1], we can exchange the limit and integral operators. Recall that for any $t < v$, $v - t > 0$, so

$$\lim_{\gamma\to 0+}\gamma^{v-t} = 0 \qquad\text{and}\qquad \lim_{\gamma\to 0+}(\ln\gamma)\gamma^{v-t} = 0.$$

Therefore, it holds that

$$\lim_{\gamma\to 0+} b^*(v) = \frac{vF(v)^{N-1}}{F(v)^{N-1}} = v.$$

Consequently, the optimal bid approaches the bidder's valuation, which decreases the profit to zero but increases the chance of winning the auction.

Example

To illustrate the optimal bidding function for risk-averse auction agents in the context of first-price sealed-bid auctions, we assume that the competitors' valuation distribution is uniform over the interval $[0, 1000]$. In this case, the bidding function that maximizes the expected profit is

$$b^*(v) = (1 - \frac{1}{N})v,$$

which can be derived from Equation (1). The bidding function that maximizes the expected utility of the profit for

risk-averse auction agents is

$$b^*(v) = \log_\gamma \frac{(v \ln \gamma)^{N-1}}{(N-1)! \left[1 - \gamma^{-v} \sum_{j=0}^{N-2} \frac{(v \ln \gamma)^j}{j!} \right]},$$

which can be derived from Equation (15). Figure 1 shows the optimal bidding functions in case there are four bidders. Different bidding functions correspond to different levels of risk aversion. The bidding functions (bids) are always lower than the identity function (bidder's valuations), as expected. If γ is approximately one, then the auction agents maximize the expected profit. As γ decreases, the auction agents become more and more risk averse, which increases the bids. These results are consistent with the qualitative claims from auction theory [12] and the properties of the optimal bidding functions proved in the previous section. Note that, even in this simple case, the bids that maximize the expected utility of the profit are much higher than the bids that maximize the expected profit, to increase the chance of winning the auction.

5. APPLICATION

We study a paper mill that uses both paper machines and cutting machines. Paper machines produce different kinds of paper rolls out of paper pulp. The cost of switching from producing one kind of paper to another kind of paper is substantial, and there are lower and upper limits on the length of paper that the machines can produce. Cutting machines then cut wide paper rolls to produce narrower ones. One of the authors has co-developed a production-planning system that manages the production process of paper mills (that is, determines how to use the machines), reducing the production cost of paper mills by millions of dollars per year [14]. The production process is driven by orders. Whenever a new order is placed, the production-planning system updates the production plan by solving a finite horizon multi-stage optimization problem, where the horizon is either fixed or set to the latest delivery date. If an order cannot be delivered completely by the negotiated date, the paper mill has to pay a penalty to the customer.

We use our auction agent for deciding whether the paper mill should participate in auctions to purchase paper rather than produce it. Consider, for example, a publisher that places an order for the paper needed to print paperback books. Such orders usually contain two different kinds of paper, namely the paper needed for the covers and the paper needed for the pages. It can be advantageous for the paper mill to purchase the paper for the covers, because the amount of paper needed for the covers is much smaller than the amount of paper needed for the pages, and it is therefore expensive to produce the paper for the covers due to both the minimum length constraints and the switching costs. Combined orders of this kind and the resulting participation in auctions are a common scenario in the paper industry and a supply-chain management system must be able to handle them. We therefore demonstrate in the following how our auction agents can be integrated into a simplified version of the existing production-planning system for paper mills.

Assumptions

We assume that bidders are similar, that is, do not only have the same valuation distributions but also the same ob-

jective function, including the same level of risk aversion. This assumption is approximately satisfied for our application: Paper mills are typically of the same size and have similar customers, production capacities, and states of the production process (such as current orders and inventory levels). They also use similar production-planning methods.

We also assume that auctions are not concurrent or combinatorial. This assumption is approximately satisfied for our application: Paper is sold in small bundles of a fixed size. Different kinds of paper are sold at different auctions. Paper mills are typically only members of one auction site and there is typically only one open auction at any point in time for any kind of paper at any one auction site.

Production Planning

So far, we have derived the bidding strategy for risk-averse auction agents in single auctions if they know the bidder's valuations of the auctioned items. We now use dynamic programming to integrate our auction agents into the existing production-planning system. This is important because the bidder's valuation of an item reflects the difference in profit between owning and not owning the item, and thus depends on how the item can be used in the production process. By integrating our auction agents into a production-planning system, we allow them to derive the bidder's valuation of an item automatically.

The production-planning problem can, in principle, be formulated as a discrete-time Markov decision problem (MDP) with a limited time horizon, that can then be solved with standard techniques from dynamic programming [5, 8]. The MDP process starts at time 0. At time t, the process is in some state $s \in S$ and the agent can choose an action $a \in A_t(s)$ for execution. The execution then results in a transition to state $s' \in S_t(s, a)$ with probability $P_t(s'|s, a)$, the agent receives an immediate profit of r according to the profit distribution $r_t(r|s, a, s')$, and the time is increased by one. Execution stops at time $T + 1$, and the agent can no longer obtain any immediate profits, where T is the planning horizon. In the context of production planning for paper mills, for example, the states correspond to the inventory levels and the status of the various machines, the actions correspond to decisions about how to use the machines, and profits correspond to payments as well as production costs.

Suppose the state of the MDP process at time t is s_t and the agent executes action a_t and receives immediate profit r_t. If the agent wants to maximize the expected sum of the immediate profits, that is, the expected profit, the agent has to maximize the following expectation along all possible state sequences $\{s_0, s_1, \ldots, s_{T+1}\}$:

$$E_{r_0, \ldots, r_T} \left[\sum_{t=0}^{T} r_t \right],$$

where the decision variables are the actions $a_t \in A_t(s_t)$ at time $t = 0, \ldots, T$. This can be done with techniques from dynamic programming because the immediate profits are additive, resulting in the one-step decomposition

$$
\begin{aligned}
E_{r_0, \ldots, r_T} \left[\sum_{t=0}^{T} r_t \right] &= E_{r_0, \ldots, r_T} \left[r_0 + \sum_{t=1}^{T} r_t \right] \\
&= E_{r_0} \left[r_0 + E_{r_1, \ldots, r_T} \left[\sum_{t=1}^{T} r_t \right] \right].
\end{aligned}
$$

Based on this decomposition, dynamic programming calculates a value function V_t which maps state s at time t to the largest expected profit that the agent can obtain if execution starts in state s at time t. The value functions are recursively defined by the following equations, known as Bellman Equations [2]:

$$V_t(s) = \max_{a \in A_t(s)} E_{r,s'}[r + V_{t+1}(s')]$$
$$V_{T+1}(s) = 0$$

for all times $t = 0, \ldots, T$ and all states s, where $s' \in S_t(s,a)$. To maximize the expected profit, the agent should then execute the action $a \in A_t(s)$ at time t that maximizes $E_{r,s'}[r + V_{t+1}(s')]$.

Risk-Averse Production Planning

We have assumed that the agent is risk-averse with a concave exponential utility function. Thus, the agent does not maximize the expected profit but rather the expected utility of the profit and thus the expectation

$$E_{r_0,\ldots,r_T}\left[u\left(\sum_{t=0}^{T} r_t\right)\right].$$

In general, this objective cannot be decomposed. However, it is known that it can be decomposed for utility functions with the delta property [9] or, equivalently, constant local risk aversion [15]. The only utility functions that have this property are the linear ones, concave exponential ones and convex exponential ones. The objective can thus be decomposed for concave exponential utility functions, resulting in the one-step decomposition

$$E_{r_0,\ldots,r_T}\left[u\left(\sum_{t=0}^{T} r_t\right)\right] = E_{r_0,\ldots,r_T}\left[-\gamma^{\sum_{t=0}^{T} r_t}\right]$$
$$= E_{r_0,\ldots,r_T}\left[-\prod_{t=0}^{T}\gamma^{r_t}\right] = E_{r_0,\ldots,r_T}\left[\gamma^{r_0}\cdot\left(-\prod_{t=1}^{T}\gamma^{r_t}\right)\right]$$
$$= E_{r_0}\left[\gamma^{r_0}\cdot E_{r_1,\ldots,r_T}\left[-\prod_{t=1}^{T}\gamma^{r_t}\right]\right]$$
$$= E_{r_0}\left[\gamma^{r_0}\cdot E_{r_1,\ldots,r_T}\left[u\left(\sum_{t=1}^{T} r_t\right)\right]\right].$$

Based on this decomposition, one can continue to apply dynamic programming, this time by calculating a value function V_t which maps states s at time t to the largest expected utility of the profit if execution starts in state s at time t. The value functions are recursively defined by the following equations:

$$V_t(s) = \max_{a \in A_t(s)} E_{r,s'}\left[\gamma^r V_{t+1}(s')\right] \quad (17)$$
$$V_{T+1}(s) = u(0) = -1 \quad (18)$$

for all times $t = 0, \ldots, T$ and all states s, where $s' \in S_t(s,a)$. To maximize the expected utility of the profit, the agent should then execute the action $a \in A_t(s)$ at time t that maximizes $E_{r,s'}\left[\gamma^r V_{t+1}(s')\right]$. In previous work, we have applied similar dynamic programming methods to sensor planning [10].

Risk-Averse Bidding and Production Planning

We now integrate risk-averse bidding into the risk-averse production planning system derived so far and show how to determine the bidder's valuation for an auctioned item automatically. We do this by deriving the optimal bid directly and then calculating the bidder's valuation based on the optimal bidding function. It turns out that we can reuse the results that we have derived in the context of isolated auctions, and that the resulting valuations have very intuitive interpretations.

We can easily integrate bidding into the risk-averse dynamic programming formulation of production planning, as given by Equations (17) and (18). Some times t now correspond to auctions. The actions at these times correspond to the possible bids b, and a bid of zero corresponds to the decision not to participate in the auction. The actions have uncertain outcomes. With probability $\rho(b)$ the auction is won. In this case, the immediate profit is the negative of the paid price p and the inventory level increases by the auctioned item (resulting in state s_+). With the complementary probability $1 - \rho(b)$ the auction is lost. In this case, the immediate profit is zero and the inventory level does not change (resulting in an unchanged state s). Consequently, the value functions for the times when the agent participates in auctions are recursively defined by the following equations:

$$V_t(s) = \max_b E_{p,s'}\left[\gamma^{-p} V_{t+1}(s')\right]$$
$$= \max_b E_{s'}\left[E_p[\gamma^{-p}|s'] V_{t+1}(s')\right]$$
$$= \max_b \left[\rho(b) E[\gamma^{-p}|s_+] V_{t+1}(s_+) + (1 - \rho(b)) E[\gamma^0|s] V_{t+1}(s)\right]$$
$$= \max_b \left[\rho(b) E[\gamma^{-p}|s_+] V_{t+1}(s_+) + (1 - \rho(b)) V_{t+1}(s)\right].$$

For simplicity, we define

$$EU_t(s,b) = \rho(b) E[\gamma^{-p}|s_+] V_{t+1}(s_+) + (1 - \rho(b)) V_{t+1}(s)$$

so that $V_t(s) = \max_b EU_t(s,b)$. The price p if the auction is won and thus $EU_t(s,b)$ depend on the kind of auction. We therefore study first- and second-price sealed-bid auctions separately. Note that a similar formulation can also be used to model the option of buying an item for a given price.

First-Price Sealed-Bid Auctions

In this case, $p = b$ if the auction is won and thus $E[\gamma^{-p}|s_+] = \gamma^{-b}$. Therefore,

$$EU_t(s,b) = \rho(b)\gamma^{-b} V_{t+1}(s_+) + (1 - \rho(b)) V_{t+1}(s)$$
$$= \left[\rho(b)\frac{V_{t+1}(s_+)}{V_{t+1}(s)}\gamma^{-b} + (1 - \rho(b))\right] V_{t+1}(s).$$

Comparing the expression with Equation (8) yields

$$EU_t(s,b) = EU\left(\log_\gamma \frac{V_{t+1}(s_+)}{V_{t+1}(s)}, b\right)(-V_{t+1}(s)).$$

$V_{t+1}(s)$ is negative according to Equation (17) and (18). Thus, $-V_{t+1}(s)$ is positive, and the bid that maximizes the first factor also maximizes $EU_t(s,b)$. So far, we have not made any assumption about the competitors. Because we

have no further information, the best we can do is to assume that these valuations are characterized by valuation distributions. We make the simplifying assumption that all valuation distributions and risk attitudes are the same, which allows us to use the optimal bidding function from Equation (16) to determine the optimal bid. It is

$$\beta = b^* \left(\log_\gamma \frac{V_{t+1}(s_+)}{V_{t+1}(s)} \right).$$

The bidder's valuation therefore is

$$
\begin{aligned}
v &= \log_\gamma \frac{V_{t+1}(s_+)}{V_{t+1}(s)} = \log_\gamma[-V_{t+1}(s_+)] - \log_\gamma[-V_{t+1}(s)] \\
&= u^{-1}(V_{t+1}(s_+)) - u^{-1}(V_{t+1}(s)).
\end{aligned}
$$

In utility theory, the expression $u^{-1}(x)$ is called the certainty equivalent of the expected utility x. The bidder's valuation therefore is simply the difference in certainty equivalents of the state that results when the item is won and the state that results when the item is not won.

Second-Price Sealed-Bid Auctions

In this case, assume that the bid of the auction agent is b, and the highest bid made by others is b'. Then $p = b'$ if the auction is won and

$$E[\gamma^{-p}|s+] = E[\gamma^{-b'}|b' < b].$$

So,

$$EU_t(s,b) = \rho(b)E[\gamma^{-b'}|b' < b]V_{t+1}(s_+) + (1 - \rho(b))V_{t+1}(s).$$

Recall that, for a random variable x with distribution $F(x)$, the conditional expectation of a function $g(x)$, given that event A happens is

$$E[g(x)|A] = \frac{\int_A g(x)dF(x)}{P(A)}.$$

Thus, the conditional expectation in $EU_t(s,b)$ can be calculated as

$$
\begin{aligned}
\rho(b)E[\gamma^{-b'}|b' < b] &= \int_0^b \gamma^{-b'} d\rho(b') \\
&= \gamma^{-b}\rho(b) - \int_0^b \rho(b')(-\ln\gamma)\gamma^{-b'} db' \\
&= \gamma^{-b}\rho(b) + \ln\gamma \int_0^b \rho(b')\gamma^{-b'} db'.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
EU_t(s,b) &= \left[\gamma^{-b}\rho(b) + \ln\gamma \int_0^b \rho(b')\gamma^{-b'} db' \right] V_{t+1}(s_+) \\
&\quad + (1 - \rho(b))V_{t+1}(s).
\end{aligned}
$$

The optimal bid β maximizes $EU_t(s,b)$ and thus the derivative of $EU_t(s,b)$ with respect to b is zero:

$$
\begin{aligned}
\left.\frac{\partial EU_t(s,b)}{\partial b}\right|_{b=\beta} &= [(-\ln\gamma)\gamma^{-\beta}\rho(\beta) + \gamma^{-\beta}\rho'(\beta) \\
&\quad + (\ln\gamma)\gamma^{-\beta}\rho(\beta)]V_{t+1}(s_+) - \rho'(\beta)V_{t+1}(s) \\
&= \rho'(\beta)(\gamma^{-\beta}V_{t+1}(s_+) - V_{t+1}(s)) = 0.
\end{aligned}
$$

The solution to the above equation is

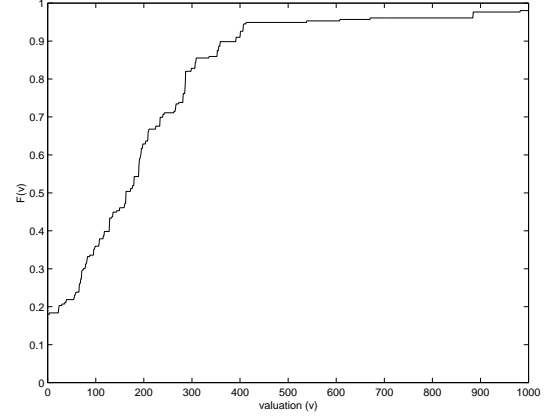$$\beta = \log_\gamma \frac{V_{t+1}(s_+)}{V_{t+1}(s)} = u^{-1}(V_{t+1}(s_+)) - u^{-1}(V_{t+1}(s)).$$

Figure 2: Estimated Valuation Distribution

Since the optimal bidding function for second-price sealed-bid auctions is the identity function [12], the bidder's valuation is again the difference in certainty equivalents of the state that results when the item is won and the state that results when the item is not won.

Estimating the Valuation Distributions

Auction theory often assumes that the competitor's valuation distributions are known. We use simulations to obtain crude approximations of them automatically. This can be done because we assume that all bidders have similar valuation distributions and we can thus approximate the competitors' valuation distributions with the one of the bidder in question, averaged over all states of the production process. We have already shown that the bidder's valuation of an item is the difference in certainty equivalents of the state that results when the item is won and the state that results when the item is not won, in the context of both first- and second-price sealed-bid auctions. We therefore use Monte-Carlo simulations of the combined auction- and production-planning process to sample these certainty equivalents and then approximate the distribution of their differences with the observed frequencies. We only simulate the production process but not the participation in auctions and therefore estimate the valuation distributions under the assumption that there are no subsequent auctions. Our estimates thus provide only crude approximations of the valuation distributions. There are alternatives to our approach. [6] estimates the bid distribution directly and, for risk-averse auction agents with concave exponential utility functions, could then use this information to obtain the valuation distribution. [7] uses fictitious play to obtain the optimal bidding strategy directly without having to estimate the valuation distribution but does so only for auction agents that maximize the expected profit.

Experimental Results

We have implemented our auction agents using heuristic search instead of dynamic programming. Solving combined auction- and production-planning problems optimally is NP-hard in general but heuristic search methods can focus the search and thus are generally somewhat faster than dynamic programming methods. To find optimal solutions, one needs to use heuristics that underestimate the costs, or synony-
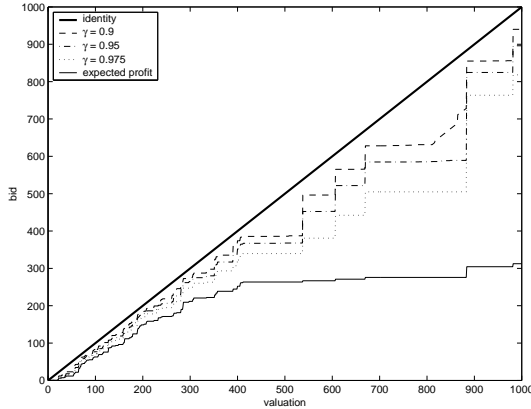
Figure 3: Optimal Bidding Functions

mously, overestimate the rewards. We used heuristics that were based on linear programming relaxations, which overestimate the expected profit and thus also the certainty equivalents for risk-averse auction agents.

We applied the resulting system to a simple problem involving a paper mill that has one paper machine and one cutting machine. Figure 2 shows the valuation distribution that we obtained. It is interesting to note that the auction agents can value an auctioned item as low as 0 dollars if orders can be satisfied easily and higher as 1000 dollars if orders cannot be delivered on time without the auctioned item and the paper mill then incurs large penalties. Figure 3 shows the resulting bidding function in case there are four bidders in a first-price sealed-bid auction whose valuation distributions are given in Figure 2. Different bidding functions correspond to different levels of risk aversion. Similar to our earlier example, the bidding functions (bids) are always lower than the identity function (bidder's valuations). If γ is approximately one, then the auction agents maximize the expected profit. As γ decreases, the auction agents become more and more risk averse, which increases the bids.

6. CONCLUSIONS

In this paper, we developed the theory for auction agents that are part of a complete supply-chain management system and seamlessly combine ideas from auction theory, utility theory, and dynamic programming. First, we used results from utility theory to derive the optimal bidding strategy in closed form for risk-averse auction agents, both for first- and second-price sealed-bid auctions in the symmetric independent private values model. To the best of our knowledge, only qualitative properties were known so far about the optimal bidding strategy, which was insufficient to implement auction agents. Second, we used dynamic programming to integrate the resulting auction agents with a production-planning system. Third, we used simulations of the combined auction- and production-planning system to obtain crude approximations of the competitor's valuations distributions of the auctioned item.

Our work so far has used simple but reasonable models of auctions and production processes. For some applications, our assumptions might be simplifying. For example, we have assumed that there are no constraints on the bids. In practice, however, the bids are often constrained by the avail-

able budget [3]. The theory developed in this paper then still provides a heuristic for building supply-chain management systems that combine auction planning and production planning, and might therefore be a good starting point to build auction agents for more complex decision situations. In future work, we intend to relax some of our assumptions and build even more powerful auction agents.

7. REFERENCES

[1] R. Bartle and D. Sherbert. *Introduction to Real Analysis*. John Wiley and Sons, third edition, 1999.

[2] R. Bellman. *Dynamic Programming*. Princeton University Press, 1957.

[3] J. Benoît and V. Krishna. Multiple-object auctions with budget constrained bidders. *Review of Economic Studies*, 68(1):155–179, 2001.

[4] D. Bernoulli. Specimen theoriae novae de mensura sortis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, 5, 1738. Translated by L. Sommer, *Econometrica*, 22: 23–36, 1954.

[5] D. Bertsekas. *Dynamic Programming, Deterministic and Stochastic Models*. Prentice Hall, 1987.

[6] S. Campo, E. Guerre, I. Perrigne, and Q. Vuong. Semiparametric estimation of first-price auctions with risk averse bidders. *submitted to: Econometrica*, 2003.

[7] S. Hon-Snir, D. Monderer, and A. Sela. A learning approach to auctions. *Journal of Economic Theory*, 82:65–88, 1998.

[8] R. Howard. *Dynamic Programming and Markov Processes*. MIT Press, third edition, 1964.

[9] R. Howard and J. Matheson. Risk-sensitive Markov decision processes. *Management Science*, 18(7):356–369, 1972.

[10] S. Koenig and Y. Liu. Representations of decision-theoretic planning tasks. In *Proceedings of the International Conference on Artificial Intelligence Planning Systems*, pages 187–195, 2000.

[11] S. Koenig and R. Simmons. How to make reactive planners risk-sensitive. In *Proceedings of the International Conference on Artificial Intelligence Planning Systems*, pages 293–298, 1994.

[12] E. Maskin and J. Riley. Optimal auctions with risk averse buyers. *Econometrica*, 52(6):1473–1518, 1984.

[13] R. P. McAfee and J. McMillan. Auctions and bidding. *Journal of Economic Literature*, 25:699–738, 1987.

[14] S. Murthy, R. Akkiraju, R. Goodwin, P. Keskinocak, J. Rachlin, F. Wu, S. Kumaran, J. Yeh, R. Fuhrer, A. Aggarwal, M. Sturzenbecker, R. Jayaraman, and B. Daigle. Cooperative multi-objective decision-support for the paper industry. *Interfaces*, 29(5):5–30, 1999.

[15] J. Pratt. Risk aversion in the small and in the large. *Econometrica*, 32(1-2):122–136, 1964.

[16] J. von Neumann and O. Morgenstern. *Theory of games and economic behavior*. Princeton University Press, second edition, 1947.

[17] S. Watson and D. Buede. *Decision Synthesis*. Cambridge University Press, 1987.