A Decision Procedure for Bundle Purchasing with Incomplete Information on Future Prices

Buffett, Scott; Spencer, Bruce

Publisher's version / Version de l'éditeur:

International Journal of Electronic Commerce, 8, 4, 2004

National Research Council Canada
Institute for Information Technology

Conseil national de recherches Canada
Institut de technologie de l'information

NRC · CNRC

*A Decision Procedure for Bundle Purchasing with Incomplete Information on Future Prices **

Buffett, S., and Spencer, B.
2004

* published in the International Journal of Electronic Commerce. Volume 8, Number 4. pp. 131-144. 2004. NRC 47143.

Canada

A Decision Procedure for Bundle Purchasing with Incomplete Information on Future Prices

Scott Buffett and Bruce Spencer

ABSTRACT: This paper introduces the on-line bundle-purchasing problem (OBPP) as a new computational challenge induced by e-commerce technology. The task of the OBPP is to decide which of many satisfactory combinations (bundles) of items should be purchased, from whom, and when, to maximize the buyer's overall satisfaction. Satisfaction, formalized as multi-attribute utility, includes attitudes toward quality, reputation, and risk. The Prequote-Quote-Rescind (PQR) protocol communicates probabilistic and temporal information on the future prices and availabilities of items. A comparison set, defined as a set of bundles in which all items are available for a fixed interval and their prices are known, determines future intervals when purchase decisions will be fully informed. A decision procedure is provided that makes effective use of comparison sets and improves a buyer's expected utility compared with a naive decision procedure.

KEY WORDS AND PHRASES: Bundle procurement, comparison shopping, protocols, purchasing-decision procedure, utility theory.

With the high volume of purchasing options available on the Internet today, strategic tools that can cleverly ascertain the true value of several different options are becoming important. More businesses are turning to Web-based pricing tools that sift through large volumes of data on product revenues, inventory levels, and consumer activity to determine how much to charge for items during certain periods [10]. Still others are resorting to dynamic pricing, whereby prices can change in time as well as across consumer markets and across packages of goods or services [9]. Dynamic pricing is utilized not only to maximize profit by responding to changes in supply and demand, but also to discourage the use of price-comparing shopbots by rendering them unreliable. These pricing strategies translate into higher profits for business, mostly at the expense of the consumer. To combat this trend, buyers need decision-analysis technology that can properly assess not only current purchasing options, but also the positive or negative potential of future opportunities. Jacobson and Obermiller support this by demonstrating the importance of considering expected future price in consumer decision-making [8]. When making a purchase decision, consumers choose from alternative courses of action. One alternative is to delay the purchase altogether. The utility of this alternative depends on the expectation of future prices. The problem of assessing such utility becomes more difficult when one considers expected future prices against several alternatives. This paper focuses on the task of computing utility against several options while considering the buyer's preferences for the options as well as the incomplete information on prices.

Consider a buyer in need of one of possibly many acceptable *bundles* of items. In the present context, a bundle is defined as a set of items determined by the buyer. Acquiring all the items in any particular bundle will be considered a success by the buyer (with varying degrees). The buyer must assess his

International Journal of Electronic Commerce / Summer 2004, Vol. 8, No. 4, pp. 131–144.

1086-4415/2004 $9.50 + 0.00.

preferences for the attributes (e.g., total cost, item preference, item suppliers, compatibility of the items) of all the bundles for which items are currently available as well as for bundles containing items that will be available in the future. For these future bundles, the buyer may have only a probability measure on the outcome of the cost. At certain times, the buyer must decide whether to buy a bundle of currently available items or take the risk of letting these pass and waiting for future opportunities. The goal is to buy the single bundle that best meets the buyer's preferences, the so-called *on-line bundle-purchasing problem* (OBPP). A common approach to deciding whether to purchase a given bundle of items is to compare the utility of purchasing the bundle to the expected utility of all other bundle purchases. The bundle in question is purchased if and only if the utility of purchasing it meets or exceeds the expected utility of every other possible bundle purchase. If the purchasing model requires the buyer to choose the aforementioned bundle and commit to it immediately, then the model represents an optimal strategy. However, if the task is to simply decide whether or not to commit to the current bundle, then the expected utility of not committing is almost certainly higher (and never lower) than the expected utility of any future bundle purchase.

This paper proposes an approach that can be used to assess more accurately the expected utility of not purchasing a bundle. The assessment is based on the expected utility of future *time intervals* during which the buyer will have the option of choosing among several possible bundle purchases for which all prices are known. The approach considers the setting when, at any given time, the purchaser has price quotes for some items available for some fixed time as well as knowledge of incoming quotes, availabilities, sales, price fixing, and so forth, for other items available during definite future periods. The purchaser may have only a probability measure on these future prices. Purchasing decisions, therefore, are made on-line with incomplete information.

Literature Review

The approach proposed in this article makes use of *expected utility theory* (EUT) to facilitate the decision process [11, 16]. Basing purchasing decisions on expected utility maximization (as opposed, say, to expected cost minimization) seems appropriate because it is sensitive to the decision-maker's attitude toward risk and preferences for such attributes as item quality, compatibility with other items, and supplier reliability.

The on-line market-clearing problem parallels the OBPP in that transactions are made in a continuous setting where transaction possibilities may pass and information about future possibilities may be incomplete or unknown [1]. The difference is that in on-line market clearing, buy bids and sell bids are matched to maximize some value, such as seller profit or market liquidity. On-line bundle purchasing is concerned with a different problem. Here the task is to match a buyer's needs with items offered by suppliers so that a single complete bundle is purchased. Excess supply and market liquidity are not concerns here. The sole focus is on the satisfaction of the buyer.

Recent work has focused on decision procedures for bundle purchasing where there are multiple auctions in which to bid. Boutilier, Goldszmidt, and Sabata consider the model where a bundle of items must be purchased by participating in a subset of several *sequential* auctions [2, 3].[1] These auctions are first-price sealed-bid, have known start/end times, and do not overlap. At each decision point (auction start time), the optimal bidding strategy is computed and the amount (if any) to bid in the current auction is determined. The approach proposed here differs from this in respect to both auction mechanism and timings. It uses the request-for-quote mechanism and allows quotes to be open in parallel. Byde considers multiple simultaneous auctions, but the purchaser's goal in this case is to buy only one item [4]. The problem of multiple simultaneous auctions has been examined by Byde, Preist, and Jennings [5]. In their model, the purchaser attempts to buy multiple units of a single good. Finally, Preist, Bartolini, and Jennings discuss bundle purchasing in the setting where there are multiple simultaneous auctions [14].[2] Although their problem is more daunting than the one treated in this paper, because they consider English, Dutch, and sealed-bid auctions, their decision-making method only pursues the set of auctions that maximizes expected utility. That is, the expected utility of not participating in the current auctions is judged the expected utility of the optimal future set. Because the algorithm does not truly commit to this set, but instead re-evaluates its options at each decision point, this expected utility is not an accurate account of the true expected utility of the choice. The main idea set forth in this paper is to predict how the algorithm will behave in the future so as to accurately estimate the true expected utility of a choice.

The On-line Bundle Purchasing Problem

Problem Formalization

Let I be a set of items and $B \subseteq 2^I$ be a set of *bundles,* where each $b \in B$ is a combination of items that meets the needs of the buyer. At any given time, let I contain only those items that are known to be available either currently or during some definite future time. I and B may change over time as new availabilities arise and others pass. Each $i \in I$ has a quoted cost $c(i)$, and each $b \in B$ has cost $c(b)$ equal to the sum of its item costs. Note that two instances of the same item are treated as two different items when they are offered by two different suppliers or by the same supplier but as part of two different offers. If an item i is currently available, then assume that the buyer knows $c(i)$. Otherwise, the buyer has a probability measure $p : Z \rightarrow \Re$ on the outcome of the cost of i, where Z is the set of monetary units. This could be any discrete or continuous distribution obtained from market history, the supplier directly, a third party, or even subjectively decided by the buyer. The goal in the OBPP is to make decisions that maximize expected utility, ultimately giving the buyer the greatest chance of purchasing the $b \in B$ that is most preferable in terms of b and $c(b)$. In the present context, the utility of a bundle purchase is denoted by the function $u : B \times Z \rightarrow \Re$, and the expected utility of a bundle purchase by

$E[\tilde{u}(b)]$, where for each $b \in B$, $\tilde{u}(b)$ is a random variable for the unknown outcome of the utility of purchasing b, based on the probability measure for the outcome of $c(b)$.

The PQR Protocol

The Prequote-Quote-Rescind (PQR) protocol is a message-passing protocol for information exchange between a supplier and a purchaser for probabilistic and temporal information. It defines when information will be known by the purchaser about such matters as cost, the distribution of possible outcomes of cost, the time a quote will be offered, and the time a quote will be terminated. This information can then be used when planning purchases. Let $[t_0, t_n] \subseteq \Re$ be the period during which a buyer needs to purchase some bundle b of items I, and let $t_p : I \to \Re$, $t_q : I \to \Re$, and $t_r : I \to \Re$ assign time points to items i $\in I$, where $t_p(i)$ is the *prequote* time, $t_q(i)$ is the *quote* time, $t_r(i)$ is the *rescind* time for i, and $t_p(i) \leq t_q(i) < t_r(i)$. The intervals $[t_p(i), t_q(i)]$ and $[t_q(i), t_r(i)]$ are known as the *prequote interval* and the *quote interval* for i, respectively. The quote time is the time at which the quote will be offered, the rescind time is the time at which the quote expires, and the prequote time is the time at which the buyer learns the quote and rescind times. It is assumed that at the prequote time, the buyer also learns or determines the probability measure on the cost outcome of the item. Note that at time t, $\forall i \in I$, $t_p(i) \leq t < t_r(i)$. In other words, an item is added to I when the prequote is received, and is removed at the rescind time. Table 1 summarizes the periods during which the buyer will have information on the cost, potential cost, and availability of an item.

A Naive Decision Procedure

This section formalizes a naive decision procedure for the On-line Bundle Purchasing Problem. The strategy presented is simple: Each time an item in a bundle b is about to expire, $u(b, c(b))$ is computed as well as the expected utility of all other valid bundle purchases. If $u(b, c(b))$ is higher than all expected utilities, then buy b. Else, let it expire.

Before the formal definition of the decision procedure is given, the following terms are defined:

Definition 1. For a bundle b,

$$pi(b) = [t_q(b), t_r(b)] = \bigcap_{i \in b} [t_q(i), t_r(i)]$$

is known as the *purchase interval* of b. All items in a bundle can be purchased at any time during its purchase interval.

Definition 2. A bundle b *is* valid *if and only if* pi(b) $\neq \phi$.

Let t be the current time, let $B_v \subseteq B$ be the set of valid bundles, and let $\{E[\tilde{u}(b')] \mid b' \in B_v\}$ be the set of expected utilities of valid bundles. If $\exists b \in B_v$ such that $t = t_r(b) - \varepsilon$, then the purchaser decides whether to buy b using the following rule:

Interval	Information
$[t_0, t_p(i)]$	nothing is known about i
$[t_p(i), t_n]$	$t_q(i)$ is known; $t_r(i)$ is known; a probability measure on $c(i)$ (and perhaps $c(i)$ itself) is known
$[t_q(i), t_r(i)]$	i is available for purchase
$[t_q(i), t_n]$	the actual price of i is known
$[t_r(i), t_n]$	i is subject to unavailability or price change

Table 1. Time Periods During Which Buyer Will Have Certain Information About *i*.

If $u(b, c(b)) \geq \max\{E[\tilde{u}(b')] \mid b' \in B_v\}$, then purchase b.
Otherwise, allow b to expire.

This defines the set of decision points to be $\{t_r(b) - \varepsilon \mid b \in B_v\}$. In theory, any bundle is available for purchase at any time during its purchase interval, but it would be unwise to commit to purchasing it much before $t_r(b)$. First, since the cost of the bundle is fixed until $t_r(b)$, and $t_r(b)$ is known by the purchaser, there is no need to commit any earlier. Second, since new information on other bundles may arise, it would be best to wait until the last moment (perhaps leaving a minimal amount of time ε before $t_r(b)$ to perform the transactions or inform the suppliers of the buyer's intentions). Therefore, decisions only need to be made at (or just before) the $t_r(b)$ time points. At such a time, the utility of purchasing b is compared to the utilities and expected utilities of other available and future prospects, and a decision on whether to buy b is made.

This method is a naive decision procedure because it simply pursues the bundle with the greatest expected utility and does not use strategy or consider any other factors. A more intelligent method that takes account of the impact of possible future options is now given.

An Improved Decision Procedure

In the naive decision procedure, each time a valid bundle purchase is about to expire, one must determine whether there is another bundle purchase that is likely to be better. Instead of determining whether *there is a future purchase that is likely to be better*, one should really determine whether *it is likely that a future purchase will be better*. These are two different questions, as can be explained with a simple example:

Example 1. Consider playing a game of chance with a fair six-sided die. You roll it once and get a four. You are then given a decision: either cash in your chips and collect \$4, or give up the \$4 and roll the die twice more, winning the equivalent dollar amount of the higher of your two rolls. Although the expected value of each toss is just 3.5, the *expected higher value* is

$$E(\tilde{x}_h) = \sum_{k=1}^{6} kP(\tilde{x}_h = k) = 4.47 \tag{1}$$

where $\tilde{x}_h$ is the uncertain higher outcome of the two tosses. Therefore, a gambler who chose to continue would expect, on average, to make \$4.47.

The same idea comes up in making decisions about purchases. If a buyer is trying to choose between making a purchase now and waiting until later, and it is known that there is a future time when two or more bundles will be offered, the buyer needs to compare the utility of the current bundle with the *expected highest utility* of the future bundles, because the buyer will have the luxury of comparing them at that time and choosing the one with the highest utility. The notion of a *comparison set*—a set of bundles for which there is a period when the buyer will have complete information—is now introduced.

Comparison Sets

Recall that the purchase interval $pi(b)$ for a bundle b is the period during which the prices of all the items in b are known, and all the items are available for purchase.

Definition 3. Let B_v be a set of valid bundles, and let $CS \subseteq B_v$. CS is a *comparison set* of B_v if and only if it is maximal such that

$$ci(CS) = \bigcap_{b \in CS} pi(b)$$

is nonempty. The interval $ci(CS)$ is called the *comparison interval* of CS. The *comparison set cover* $csc(B_v)$ of B_v is the set of all the comparison sets of B.

Note that $ci(CS)$ is the time during which the prices of all the items in all the bundles in CS are known and all the items are available for purchase. Hence, the buyer has complete information on all the bundles in CS. Note that every bundle in B_v will appear in at least one comparison set—even if by itself—and may appear in more than one. Thus, $csc(B_v)$ is a covering of B_v.

Algorithm 1 (Construction). The comparison set cover $csc(B_v)$ for B_v is constructed by first finding the comparison intervals and then determining the comparison sets from them. Let T be a sorted list of the time points in $\{t_q(b) \mid b \in B_v\} \cup \{t_r(b) \mid b \in B_v\}$ from earliest to latest. Ties between a t_q and a t_r time are broken by placing the t_r time first, and all other ties are broken arbitrarily. For each pair of consecutive elements t_k and t_{k+1} in T, if t_k is a t_q time and t_{k+1} is a t_r time, then $[t_k, t_{k+1}]$ is a comparison interval, and $CS = \{b \in B_v \mid [t_k, t_{k+1}] \subseteq pi(b)\}$ necessarily is a comparison set. The comparison set cover $csc(B_v)$ is then the set of all of these comparison sets.

Example 2. Let $B_v = \{b_1, b_2, b_3, b_4, b_5\}$ where each bundle has a purchase interval as depicted by the horizontal lines in Figure 1 (e.g., the purchase interval for b_1 is [0, 3]). The comparison intervals are indicated by dotted vertical lines. The comparison set cover for B_v is then $csc(B_v) = \{CS_1, CS_2, CS_3\}$, where $CS_1 = \{b_1, b_2\}$, $CS_2 = \{b_2, b_3, b_4\}$, and $CS_3 = \{b_5\}$.

Because all the items in all the bundles in a given comparison set CS are available during a common interval and all the prices are known, a buyer who chooses to buy during this period will choose the bundle in CS with the highest purchase utility. The utility one would expect to achieve during this period is, therefore, equal to the expected highest utility of the bundles in CS, referred to hereafter simply as the expected utility of CS and denoted by $E[\tilde{u}(CS)]$.

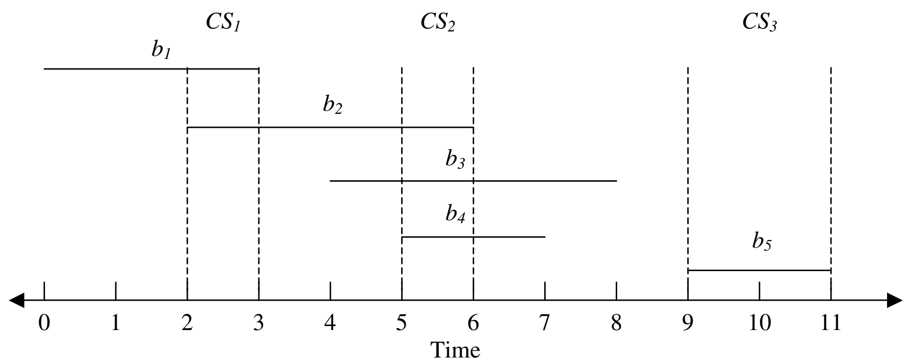

Figure 1. Comparison Set Cover of B_v in Example 2.

The Proposed Decision Procedure

Let b be the bundle currently available at cost $c(b)$ that is about to expire. Let B_v be the set of valid bundles, and let $csc(B_v)$ be the comparison set cover for B_v, each $CS \in csc(B_v)$ with expected utility $E[\tilde{u}(CS)]$.

If $u(b, c(b)) \geq \max\{E[\tilde{u}(CS)] \mid CS \in csc(B_v)\}$, then purchase b.

Otherwise, allow b to expire.

Using this decision procedure provides the buyer with a higher expected utility than using the naive decision procedure (*see Theorem 1*).

Calculating the Expected Utility of a Comparison Set

Unfortunately, computing the expected utilities of comparison sets can be complex. When using continuous random variables to represent item prices, in order to calculate the exact expected utility of a comparison set, one would have to solve the multiple integral

$$E[\tilde{u}(CS)] = \int_0^1 \ldots \int_0^1 \max\{x_1, \ldots, x_n\} \prod_{i=1}^{n} p_i(x_i)\, dx_1 \ldots dx_n \tag{2}$$

where $x_1, \ldots, x_n$ are the utilities of the bundles in CS, and $p_1, \ldots, p_n$ are their respective probability density functions. Since no closed-form expression exists for even the single integral of a normal probability density function [12], if some or all of the p_i are normal (or some other complex form), then it is unlikely that the above can be expressed in closed form. Making this computation even more difficult is the fact that there may be a strong interdependence among many of the p_i, because many bundles will have common items. Although it is possible to compute the value of this integral exactly if it has certain properties (e.g., if the p_i are uniformly distributed or have only a few discrete outcomes), we suggest, in general, the use of a Monte Carlo method to estimate this value.

Monte Carlo methods involve simulation to solve approximately a mathematical problem [6, 13]. Such a method can be used to estimate the expected highest utility of a set of bundles in a comparison set. This is done by first properly modeling the system of items residing in the bundles in question, including the probability distributions for the costs of the items and the interdependencies, if any, among the item costs. The results of several independent simulations of the random elements involved in the system are then obtained. For each simulation, the outcomes of the item prices are used to determine the utility of purchasing each bundle, and the highest is noted. The average of these results is then taken as the unbiased estimate of the expectation. Simulations are run until the standard error $\sigma \sqrt{n}$ is small enough to achieve the desired confidence in the estimate. To help shrink $\sigma \sqrt{n}$ a variance reduction technique, referred to as the *antithetic variate* sampling method, is used [7]. With this method, price outcomes are selected in pairs that mutually compensate for each other's variations. See Hammersley and Handscombe for a more detailed description of the approach [6].

Analysis

The improved decision procedure yields a higher expected utility than the naive decision procedure. The increase in utility one may expect when using the improved decision procedure instead of the naive procedure will be determined for a simple case. An example will then be provided to shed some light on what this increase means.

Improved Decision Procedure Yields Higher Expected Utility

For a comparison set cover $csc(B_v)$ of valid bundles B_v, let $j = \max\{E[\tilde{u}(b)] \mid b \in B_v\}$ and $k = \max\{E[\tilde{u}(CS)] \mid CS \in csc(B_v)\}$.

LEMMA 1. $j \leq k$

Proof. *Let* CS_j *be a comparison set containing* b *such that* $E[\tilde{u}(b)] = j$. *Because the expected utility of choosing from a number of bundles must be at least as high as the expected utility of any one of the bundles,* $E[\tilde{u}(CS_j)] \geq j$. *Thus,* $k = max\{E[\tilde{u}(CS)] \mid CS \in csc(B_v)\} \geq E[\tilde{u}(CS_j)] \geq j$.

LEMMA 2. *The buyer's expected utility,* Eu_{im}*, when using the improved decision procedure is greater than or equal to* k.

Proof. *By induction on the size of* $csc(B_v)$.

Basis step. Let $|csc(B_v)| = 1$. Then, $Eu_{im} = k$.

Induction step. Let $csc(B_v)$ be of arbitrary size, let CS_i be the current comparison set, and let b be the bundle in CS_i available at cost z such that $u(b, z)$ is the maximum in CS_i. Assume that the expected utility when using the informed decision procedure for $csc(B_v) - CS_i$ is at least $\max\{E[\tilde{u}(CS)] \mid CS \in csc(B_v) - CS_i\}$. Consider the two outcomes (note that $u(b, z) > k$ is not possible):

1. $u(b, z) = k$. Buy b and achieve utility k. Thus, $Eu_{im} = k$.
2. $u(b, z) < k$. Then, $k = \max\{E[\tilde{u}(CS)] \mid CS \in csc(B_v) - CS_i\}$. Because b would not be purchased in this case (and thus nothing from CS_i is purchased, it expires) and by induction, the expected utility for $csc(B_v) - CS_i$ is at least $\max\{E[\tilde{u}(CS)] \mid CS \in csc(B_v) - CS_i\}$, then $Eu_{im} \geq k$.

THEOREM 1. *Let* $\mathrm{Eu}_{\mathrm{im}}$ *and* $\mathrm{Eu}_{\mathrm{na}}$ *denote the utilities expected when using the improved and naive decision procedures, respectively. Then,* $\mathrm{Eu}_{\mathrm{im}} \geq \mathrm{Eu}_{\mathrm{na}}$.

Proof. *By induction on the size of* $\mathrm{csc}(\mathrm{B}_v)$.

Basis step. Let $|csc(B_v)| = 1$. Both procedures will choose the bundle purchase with the highest utility, so $Eu_{im} = Eu_{na}$.

Induction step. Let $csc(B_v)$ be of arbitrary size, let CS_i be the current comparison set, and let b be the bundle in CS_i available at cost z such that $u(b, z)$ is the maximum in CS_i. Assume that the expected utility when using the improved decision procedure is greater than or equal to that when using the naive decision procedure for $csc(B_v) - CS_i$. By Lemma 1, there are three cases:

1. $k = u(b, z)$. Both procedures would choose b, thus $Eu_{im}=Eu_{na}$.
2. $j \leq u(b, z) < k$. The naive procedure would choose b, and the improved procedure would not. Thus, $Eu_{na} = u(b, z) < k$, because, by Lemma 2, $Eu_{im} \geq k$, $Eu_{im} > Eu_{na}$.
3. $u(b, z) < j$. Neither procedure would choose b, thus allowing CS_i to pass. By induction, $Eu_{im} \geq Eu_{na}$.

Expected Utility Increase for Improved Decision Procedure

The increase in utility that one would expect to achieve when using the improved procedure as opposed to the naive procedure for a simple case is derived. In the situation considered here, there are two comparison sets, CS_1 and CS_2, and there are no interdependencies between the utilities of the two comparison sets (thus, $CS_1 \cap CS_2 = \phi$), and $ci(CS_1)$ is before $ci(CS_2)$. Let $j = \max\{E[\tilde{u}(b')] \mid b' \in CS_2\}$, and $k = E[\tilde{u}(CS_2)]$. Let an example probability density function for the unknown outcome of $\tilde{u}(CS_1)$, which is the highest utility of all bundles in CS_1, be as depicted in Figure 2. Example points for j and k are also displayed. Using these points, the area under the curve is divided into three regions. Let A_1, A_2, and A_3 represent the areas of these regions. Specifically,

$$A_1 = \int_0^j p(x)\ dx \quad A_2 = \int_j^k p(x)\ dx \quad A_3 = \int_k^1 p(x)\ dx.$$

THEOREM 2. The expected increase in utility achieved by using the improved decision procedure is

$$\int_j^k (k - x)\, p(x)\, dx \qquad (3)$$

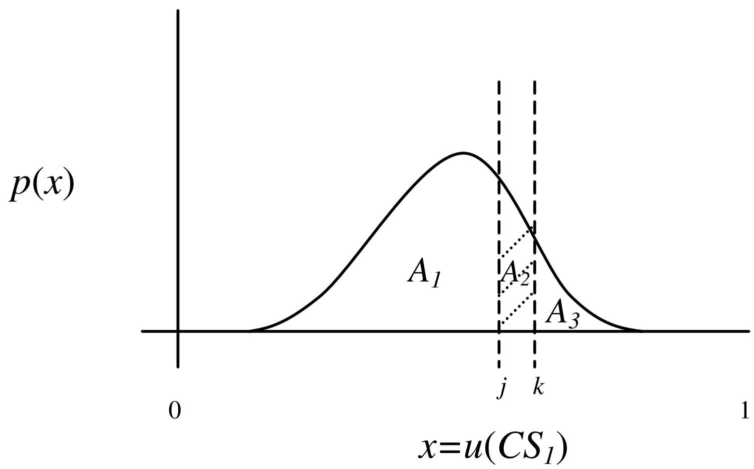

Figure 2 <<Fig title?>>

where there are two independent comparison sets CS_1 and CS_2, $p(x)$ is the probability density function for the outcome of the highest utility over all THE bundles in CS_1, $j = \max\{E[\tilde{u}(b')] \mid b' \in CS_2\}$, and $k = E[\tilde{u}(CS_2)]$.

Proof. Let Eu_{na} be the expected utility of using the naive decision procedure, and let Eu_{im} be the expected utility of using the improved procedure. The naive decision procedure will choose to buy from CS_1 if it contains a bundle with utility higher than j; otherwise it will let CS_1 expire and the expected utility will be k. Then[3]

$$Eu_{na} = A_1 k + A_2 E[\tilde{u}(CS_1) \mid j < \tilde{u}(CS_1) < k].$$
$$+A_3 E[\tilde{u}(CS_1) \mid \tilde{u}(CS_1) > k] \tag{4}$$

The improved decision procedure will choose b if its utility is higher than k; otherwise it will let it expire and the expected utility will be k. Then

$$Eu_{im} = A_1 k + A_2 k - A_3 E[\tilde{u}(CS_1) \mid \tilde{u}(CS_1) > k]. \tag{5}$$

Subtracting gives

$$Eu_{im} - Eu_{na} = A_2 k - A_2 E[\tilde{u}(CS_1) \mid j < \tilde{u}(CS_1) < k]$$

$$= A_2 k - A_2 \int_j^k xp(x)\, dx / A_2$$

$$= k \int_j^k p(x)\, dx - \int_j^k xp(x)\, dx$$

$$= \int_j^k (k - x) p(x)\, dx. \tag{6}$$

b_i	μ_i	σ_i
b_1	0.5	0.06
b_2	0.475	0.13
b_3	0.484	0.1

Table 2. Means and Variances of Bundle Purchase Utilities in the Example.

For the general case involving many comparison sets, if $j = \max\{E[\tilde{u}(b')] \mid b' \in csc(B_v) - CS_1\}$ and $k = \max\{E[\tilde{u}(CS)] \mid CS \in csc(B_v) - CS_1\}$, then this value is a lower bound on the expected increase, with the restriction that comparison set utilities are independent. If there exist some interdependencies, then a lower bound can still be determined if the interdependencies are removed such that j is unaltered. For example, if there exists an item i that resides in bundles in two or more comparison sets, then restrict the quote interval of i to $ci(CS)$, where CS is a comparison set that contains a bundle b such that $i \in b$ and $\tilde{u}(b) = \max\{\tilde{u}(b') \mid i \in b'\}$. Since j will still be the true value, but k could be an underestimate, the value will be a lower bound.

Example

Let $\{CS_1, CS_2\}$ be a comparison set cover with nonintersecting comparison intervals, and let $CS_1 = \{b_1\}$, and $CS_2 = \{b_2, b_3\}$. For simplicity, let the bundle costs be independent and normally distributed, and consider the buyer risk-neutral. Then, the bundle purchase utilities will be normally distributed. Consider the parameters given in Table 2.

At time $t_r(b_1)$, the buyer will know $c(b_1)$ (/) and therefore $u(b_1, c(b_1))$, μ_2, σ_2, μ_3, and σ_3. A decision must be made at that time between purchasing b_1, which will achieve utility $u(b_1, c(b_1))$, and allowing b_1 to expire, which will achieve the higher of the two utility outcomes for b_2 and b_3.

Recall that the naive decision procedure chooses b_1 if and only if $u(b_1, c(b_1)) \geq j$, where $j = \max\{E[\tilde{u}(b)] \mid b \in CS_2\} = 0.484$, and the improved procedure chooses b_1 if and only if $u(b_1, c(b_1)) \geq k$, where $k = E[\tilde{u}(CS_2)]$. Using a Monte Carlo simulation, $k = 0.545$. Then, by Theorem 2, the expected increase in utility is

$$\int_{.484}^{.545} (.545 - x)\, p(x)\, dx \approx .012 \tag{7}$$

where $p(x)$ is the normal probability density function with $\mu = 0.5$ and $\sigma = 0.06$.

This means that one may expect to achieve 0.012 more utility by using the improved decision procedure. It is difficult to draw any conclusions about the significance of this increase without knowing the utility function, because it is the result of some combination of more highly preferred bundles and lower costs. However, to take a simple example, assume that all the bundles are preferred equally. If this is the case, then lower costs will be completely responsible for the increase in utility. In addition, assume that the range of possible outcomes of bundle costs is [\$100, \$200] and the range of utility is normalized to [0, 1], which is commonly done. Therefore, $u(b, \$100) = 1$ and $u(b, \$200) = 0$. If the buyer is risk-neutral, then each 0.01 of bundle utility represents \$1. Thus, an increase in utility of 0.012 represents a savings of \$1.20. This is a significant result in that this is a small-scale example. Consider a large-scale example, such as purchasing materials for a major construction project, where the range of values is around one million dollars rather than one hundred dollars. In this case, the increase in utility represents \$12,000 in savings.

Conclusion

This paper formalizes the on-line bundle purchasing problem (OBPP) as the problem faced by a buyer contemplating purchasing one of several bundles in a setting where at any given time the purchaser has price quotes for some items that are available for some fixed time as well as knowledge of incoming quotes, availabilities, sales, price fixing, and so forth, for other items available during definite future periods. Expected utility is used as the maximization goal. Initially, the Prequote-Quote-Rescind (PQR) protocol is introduced as a set of message-passing rules that provide a framework defining the time intervals during which certain information is known to the purchaser regarding item availabilities and prices. A decision procedure that exploits time intervals during which many options will be available is then proposed for the OBPP, and is proven to yield a higher expected utility than a naive decision procedure that simply pursues the best bundle. This is substantiated with an analysis of the value of considering future choices by deriving a measure of the improvement a purchaser would expect to realize if future choices are considered when making decisions, compared to simply pursuing the best bundle purchase, in a simple case.

If a market is completely known and invariable, then a naive inflexible purchase decision procedure is appropriate. The PQR protocol accommodates dynamic and volatile market conditions, and a robust, flexible purchase-decision procedure is provided that deals with the uncertainty and exploits future options.

In current on-line shopping, the purchaser is offered quotes immediately upon asking for them, but this forces sellers to set high prices to make up for not knowing what the demand for their product will be. The proposed PQR protocol allows sellers to gather quote requests and gives them some idea of demand before they set a price, perhaps allowing them to set a fairer and less risky value. Price setting in response to quote requests is already common in business-to-business (B2B) e-commerce. In future work, the authors will look at the effect of the PQR protocol on vendors as well as the effects of protocols other than the PQR on both vendors and purchasers. This will probably involve the use of game-theoretic techniques, such as those described by Shubik [15].

In another project, the authors plan to explore the computational ramifications of relaxing the restriction that all the items in a bundle must be available at the same time. It may be that the best solution consists of buying items in a bundle at different times. Making a partial bundle purchase is risky, however, because the buyer may be forced to purchase expensive items to complete the bundle if the cost outcomes turn out to be high. Expected utility theory can determine whether the added utility of making a partial purchase outweighs the risk. This is a complex mathematical problem, because in addition to the expected outcomes of the future items, one must also consider what choices are to be made in the future and what new information will be known at those times. All of these factors affect the expected utility of a decision. Straight simulation of the outcomes would not give an accurate result. The authors are currently researching both randomized and nonrandomized algorithms to solve this new problem.

The authors also plan to extend the model to allow the buyer to participate in on-line auctions. Although the addition of various auction mechanisms would greatly magnify the computational burden, it would certainly make the methods described much more useful. New techniques, based on those developed in this paper, would likely be needed to accomplish this goal.

NOTES

1. They refer to items as "resources."
2. They refer to bundle purchasing as "service composition."
3. Here $E[X \mid C]$ is the expected value of X given that condition C holds.

REFERENCES

1. Blum, A.; Sandholm, T.; and M. Zinkevich. Online algorithms for market clearing. In *ACM-SIAM Symposium on Discrete Algorithms (SODA).* San Francisco: ACM Press, 2002, pp. 971–980.

2. Boutilier, C.; Goldszmidt, M.; and Sabata, B. Continuous value function approximation for sequential bidding policies. In *The Fifteenth Annual Conference on Uncertainty in Artificial Intelligence (UAI-99).* Stockholm: Morgan Kaufmann, 1999, pp. 81–90.

3. Boutilier, C.; Goldszmidt, M., and Sabata, B. Sequential auctions for the allocation of resources with complementarities. In *The Sixteenth International Joint Conference on Artificial Intelligence (IJCAI-99).* Stockholm: Morgan Kaufmann, 1999, pp. 527–534.

4. Byde, A. A dynamic programming model for algorithm design in simultaneous auctions. In *WELCOM'01.* Heidelberg: Springer-Verlag, 2001, pp. 152–163.

5. Byde, A.; Preist, C.; and Jennings, N.R. Decision procedures for multiple auctions. In *Proceedings of the First International Joint Conference. on Autonomous Agents and Multi-Agent Systems (AAMAS'02).* Bologna: ACM Press, 2002, pp. 613–620.

6. Hammersley, J.M., and Handscombe, D.C. *Monte Carlo Methods.* New York: Wiley, 1964.

7. Hammersley, J.M., and Morton, K.W. A new Monte Carlo technique: Antithetic variates. In *Proceedings of the Cambridge Philosophical Society, 52* (1956), 449–475.

8. Jacobson, R., and Obermiller, C. The formation of expected future price: A reference price for forward-looking consumers. *Journal of Consumer Research, 16* (March 1990), 420–432.
9. Kannan, P.K., and P.K. Kopalle. Dynamic pricing on the Internet: Importance and implications for consumer behavior. *International Journal of Electronic Commerce, 5,* 3 (spring 2001), 63–83.
10. Keenan, F. The price is really right. *Business Week Online* (March 31, 2003). Available at www.businessweek.com/magazine/content/03_13/b3826046.htm.
11. Keeney, R.L., and Raiffa, H. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs.* New York: John Wiley and Sons, 1976.
12. Mendenhall, W., and Scheaffer, R.L. *Mathematical Statistics with Applications.* Boston: Duxbury Press, 1973.
13. Metropolis, N., and Ulam, S. The Monte Carlo method. *Journal of the American Statistical Association, 44*, 247 (1949), 335–341.
14. Preist, C.; Bartolini, C.; and Byde, A. Agent-based service composition through simultaneous negotiation in forward and reverse auctions. In *Proceedings of the Fourth ACM Conference on Electronic Commerce.* San Diego: ACM Press, 2003, pp. 55–63.
15. Shubik, M. *A Game-Theoretic Approach to Political Economy.* Cambridge, MA: MIT Press, 1984.
16. Von Neumann, J., and Morgenstern, O. *Theory of Games and Economic Behavior*, 2d ed. Princeton: Princeton University Press, 1947.

SCOTT BUFFETT (scott.buffett@nrc.gc.ca) is a research officer in the Internet Logic Group of the National Research Council, Institute for Information Technology—e-Business. He holds a Ph.D. in computer science from the University of New Brunswick. His research interests involve the application of artificial intelligence and economics research to e-commerce and related topics, such as e-procurement, auctions, supply chain management, privacy, and automated negotiation.

BRUCE SPENCER (bruce.spencer@nrc.gc.ca) is the group leader of the Internet Logic Group of the National Research Council, Institute for Information Technology—e-Business, and an adjunct professor of computer science at the University of New Brunswick. He received his Ph.D. from the University of Waterloo in 1990. His research interests center on automated reasoning, applying reasoning systems to the Web, artificial intelligence, logic programming, and software engineering.