Mathematical Challenges of Combinatorial Auction Design

By James Case

The auction mechanism is older than recorded history—Herodotus mentions it in his *History of the Babylonians*, said to be the first history book ever written—but there seems to have been no attempt to analyze it quantitatively until the 1950s. Among the

most influential of the early publications on the subject was a paper by William Vickrey [2], who was later to share the Nobel Memorial Prize for economics (with James Mirrlees, in 1996) in part for his work on auctions. The auctions he considered were rather special in that each potential buyer was presumed to know exactly what the object at auction was worth to him or her, and that others would value it differently.

Imagine, for instance, three art dealers at an auction. Each has been authorized by a valued client to bid on a specific painting. The first has leave to spend as much as \$7000, the second \$11,000, and the third \$23,000 to acquire the painting in question. If bids are denominated in multiples of \$100, the last mentioned can be

Many modern auctions are "combinatorial" in nature, meaning that a long list of items is offered for sale and bidders are free to submit bids on packages of items from the list.

expected to delight his client by acquiring the painting for at most \$11,100. This led Vickrey to propose that, in sealed-bid auctions, the object be awarded to the highest bidder at the second highest of the received bids, so that no bidder has any incentive to bid less than he or she is willing to pay, and so that the outcome will mimic that of an ordinary (audible) auction. Such auctions are now known as "Vickrey" or "sealed-bid" auctions. They are common in the auction literature, but uncommon in practice.

Winner's Curse

Celebrated though it was, Vickrey's analysis proved quite useless to the oil industry in its hour of need. Production of oil and gas from offshore wells in the Gulf of Mexico began in 1946; firms would lease drilling rights from the federal government, which made new drilling sites available from time to time via sealed-bid auctions. By 1960, although many such wells were already in production, it was well known in the industry that few if any firms were making money offshore. At first, the disappointed firms assumed themselves to be alone in their suffering, and blamed their troubles on inadequate estimation techniques. But they soon concluded that subtleties lurking within the sealed-bid auction process itself were causing them to pay too much for their drilling rights.

Although a barrel of crude (or a volume of gas) is worth more or less the same amount to any firm equipped to operate offshore, the number of barrels lurking beneath a given section of the ocean floor is exceedingly difficult to estimate—so much so that gross errors are all but unavoidable. Thus, in contrast to the situation at a Vickrey auction, each item offered has roughly the same value to each potential bidder, but no one ever knows what that *common value* is.

Three employees of the Atlantic-Richfield Company (ARCO) found a particularly instructive way to illustrate the difficulty of making money on assets purchased at such auctions [1]. They would simply fill an ordinary glass jar with nickels, and auction it off to the members of any sizeable audience they could assemble. Each prospective buyer was allowed to handle the jar, and to compare it with a standard (\$2) roll of nickels. But no one was permitted to open the jar or handle the contents before placing a bid.

Sometimes the trio conducted ordinary (audible) auctions, and sometimes they solicited sealed bids. But they managed, on virtually every occasion, to sell the jar for more than its cash value. To the difference between the sale price and the cash value of the item—which was often substantial—they gave the name "winner's curse."

To discover what made the difference so large, they conducted contests separate from the auctions in which a "valuable prize" was offered to the person who most accurately estimated the value of the jar and its contents. They found the estimates to be more or less log-normally distributed about their geometric mean, which seldom if ever differed significantly from the cash value of the jar and its contents. But the individual estimates did not tend to cluster tightly around that cash value. On the contrary, they were often in error by factors of three or more.

Finally, the trio noticed that inexperienced bidders seldom bid less than 75% of what they imagined the cash value to be. Hence, the eventual purchaser tended to be someone who supposed the object to be worth at least twice its actual cash value, and whose bid accordingly exceeded that value by some 50%. Accurate estimators—like conservative bidders—had little if any chance of winning!

Groups from other firms, focusing on different aspects of the problem, explored a variety of approaches to it. Several sought to exploit the fact that dozens of sites were included in each auction conducted by the Bureau of Land Management, so that losses on one lease might be offset by profits on others. In addition, an after-market sprang up in which firms winning more sites than they really wanted could hand the extras off to less successful bidders. Indeed, because the BLM always published the winning bids, disappointed parties could offer potential resellers a small but certain profit on each such transaction.

Enter the Internet

The Internet, with its plethora of e-auctions, has reawakened interest in the auction process. E-bay and Amazon.com auction off countless items every week. Airports auction departure gates. Airlines offer cut-rate fares on a first-come, first-served basis, making additional low-price seats available on under-booked flights as departure times near. Called "yield management," the practice is nothing more than a progressive posted-price auction for unsold seats. The Federal Communications Commission auctions off segments of the radio frequency spectrum to telecommunications companies. Electric utilities within and without California solicit bids each day for the next day's service on an hour-by-hour, or even minute-by-minute, basis.

Many modern auctions are "combinatorial" in nature, meaning that a long list of items is offered for sale and bidders are free to submit bids on packages of items from the list. A particular itemthen, can (and often does) add much to the price of one package, but little or nothing to the price of another. Combinatorial auctions are typically broken into a succession of "rounds," during each of which participants are free to bid on any unsold package, and at the end of which the auctioneer must either accept or reject each of the bids received. The process continues until a round occurs in which the auctioneer receives no acceptable bids.

At a "hot topics" conference held in early December at the Institute for Mathematics and Its Applications at the University of Minnesota, no fewer than 19 speakers described mathematical problems encountered in the course of their involvement in eauctions. The opening lectures were tutorial in nature. Michael Rothkopf described the history of the subject, and outlined future modeling challenges, while Rakesh Vorha concentrated on recent developments, quite a few of which had to do with auction design. The following account borrows heavily from Vorha's presentations.

The first and most obvious difficulty in designing the format for a given combinatorial auction is that each bidder must submit a bid for *every* package in which he or she is interested. How is so much information even to be transmitted to the auctioneer? All known designs involve narrow restrictions on the number of "admissible" packages for which bidders are permitted to bid. One approach is to specify the language or languages in which bids can be expressed, it being known that no simple language suffices to express more than a few of the possible preference orderings over the subsets of items on a list of appreciable length.

An alternative approach—not yet explored in depth—involves the use of "oracles" that are capable of computing the bid a given bidder would submit for any admissible package. Several of the IMA talks touched on the design of appropriate oracles. In game-theoretic terms, an oracle is but the software realization of a particular bidding strategy, which bidders could be required to submit in place of bids. With oracles in hand, the auctioneer need only invoke a given bidder's oracle to learn what that bidder would quote for a given package.

Combinatorial Auctions and Integer Programming

After deciding on a language in which bids can be submitted, the auctioneer must still decide which bids to accept or reject. This problem, called the combinatorial auction problem (CAP), is often formulated as an integer programming problem: Let N be the set of bidders and L the list of distinct items on offer. For every subset S of L, let $b^{j}(S)$ be the bid that agent $j \in N$ has announced she or he is willing to pay for S. Let $b(S) = \max_{j \in N} b^{j}(S)$. The CAP can then be formulated as:

$$\max \sum_{S \subset L} b(S) x_{S}$$

s.t. $\sum_{S \ni i} x_{S} \leq 1 \forall i \in L$
 $x_{S} \in \{0,1\} \forall S \subset L.$

Here $x_s = 1$ means that the highest bid on the set S is to be accepted, whereas $x_s = 0$ means that all bids on S should be rejected. The constraint on $\sum x_s$ ensures that no item on the list L is assigned to more than one bidder.

The foregoing formulation correctly models the CAP when the bid functions b^{j} are all superadditive, in the sense that $b^{j}(A) + b^{j}(B) \leq b^{j}(A \cup B)$ for all $j \in N$ and all disjoint $A, B \subset L$. This expresses the notion that the items on the list *L* complement one another in the way that left shoes complement right shoes, and vice versa. If shoes were sold separately, one of each would presumably be worth more than either two left ones or two right ones. Other formulations are available for bid functions that are not superadditive. To extend the present formulation to the case in which there are multiple copies of certain items on the list *L* but no bidder wants more than one of each, it suffices to replace the 1 on the right side of the summatory constraint with a larger number.

The CAP, as formulated above, is an instance of the set packing problem (SPP), a well-known integer programming problem: Given a (finite) list *L* of items, and a collection $\mathcal{L} = \{L_j\}_{j \in J}$ of subsets of *L* with non-negative weights $\{w_j\}_{j \in J}$, find the "heaviest" sub-collection $\mathcal{L}^* \subset \mathcal{L}$ of pairwise-disjoint subsets of \mathcal{L} . The SPP can be formulated as an integer program by having $x_j = 1$ if subset L_j is selected, and 0 otherwise, and by letting $m_{ij} = 1$ if $i \in L$ is also in L_j , and 0 otherwise. Given these definitions, the SPP can be formulated as follows:

$$\max \sum_{j \in J} w_j x_j$$

s.t. $\sum_{j \in J} m_{ij} x_j \le 1 \forall i \in L$
 $x_i \in \{0,1\} \forall j \in J.$

To see that the CAP is an instance of the SPP, let L be the list of objects and \perp the set of all subsets of L.

A close relative of the SPP is relevant to procurement auctions in which the auctioneer is a buyer whose wants the bidders compete to satisfy. Many of the auctions used in the transportation industry are of this type. The list *L* then consists of origin/destination pairs, called "lanes." Bidders submit bids on coherent bundles of lanes, which represent offers to service those exact lanes for a

stipulated fee. The auctioneer wishes to select the collection of bids that provides the requisite service at minimum cost. This can be found by solving the following set partitioning problem (SPA):

$$\min \sum_{j \in J} w_j x_j$$

s.t. $\sum_{j \in J} m_{ij} x_j = 1 \forall i \in L$
 $x_i \in \{0,1\} \forall j \in J.$

A close relative of the set partitioning problem is the set covering problem (SCP), the integer programming formulation of which differs from that of the SPA only in that the "=" appearing in the summatory constraint is replaced by a " \geq ". It too is relevant to certain auctions conducted by the transportation industry, as when railway crews must be assigned to "runs," or lanes.

Reduction to Linear Programming

Integer programs are usually NP-hard, and the SPP is no exception. But it sometimes happens that the vertices of the polyhedron $P(M) = \{x : \sum_{j \in J} m_{ij} x_j \le 1 \forall i \in L; x_j \ge 0 \forall j \in J\}$ all have integer coordinates, which indicates that constraint $x_j \in \{0,1\} \forall j \in J$ is redundant and that the corresponding SPP can be solved as a linear program. In such instances, a combinatorial method can usually be found that solves the given problem even more efficiently than linear programming. Nevertheless, the connection with linear programming is a useful one, because it supplies dual variables that can be interpreted as prices for the individual items on the list *L*. A variety of conditions are known to guarantee that P(M) has integer vertices, and a polynomial-time algorithm exists for deciding whether a given matrix *M* with elements m_{ij} from $\{0,1,-1\}$ generates integer vertices. Most texts on integer programming devote at least a chapter to this aspect of the subject.

Generally speaking, the conditions guaranteeing that P(M) has only integer vertices are unduly restrictive, in the sense that the more interesting problems tend to generate constraint matrices M for which P(M) possesses both integer and noninteger vertices. Integer programming has come a long way since the dawn of the computer age, and is now able to handle surprisingly intricate problems. In particular, Brenda Diet-rich—one of the organizers of the IMA workshop—described the results her group at IBM has been obtaining with column-generation methods.

Column generation is a technique for solving linear programs with so many variables—each one corresponding to a column of the constraint matrix M—that M overflows main memory. For a matrix M with n rows, the technique begins with the generation of an $n \times 2n$ submatrix, and the selection from its columns of an optimal basis for the column space of the submatrix. The unused columns of the submatrix can then be discarded and replaced by n newly generated columns of M; a new optimal basis is then found within the modified submatrix. If each step is performed judiciously, each successive optimal basis will strictly excel its predecessor, and the process will terminate in an optimal basis for the entire column space of M, having generated only relatively few of the many columns of M. Dietrich described a few of the successes her group at IBM has had with this technique.

Special-purpose commercial software packages exist for solving auction-related integer programs. Logistics.com's OptiBid[™] has been used in situations in which the number of bidders varied from 12 to 350, the average being about 120. The number of items (typically lanes) has ranged from 500 to 10,000. SAITECH-INC's bidding software SBID, likewise based on integer programming, is said to work on problems of similar size. Combinatorial auctions seem likely to occupy the attention of auction theorists for some time to come.

References

[1] E.C. Capen, R.V. Clapp, and W.M. Campbell, Competitive bidding in high-risk situations, J. Petrol. Tech., June (1971), 641-653.

[2] E. Vickrey, Counterspeculations, auctions, and competitive sealed tenders, J. Finance, 16 (1961), 8–37.

James Case writes from Baltimore, Maryland.