From *SIAM News*, Volume 42, Number 9, November 2009

Math and Congressional Redistricting

By James Case

In 2004, Supreme Court Justice David Souter expressed the opinion that "the increasing efficiency of partisan redistricting has damaged the democratic process to a degree that our predecessors only began to imagine." His colleague Anthony Kennedy concurred: "Because there are as yet no agreed upon substantive principles of fairness in districting, we have no basis on which to define clear, manageable, and politically neutral standards." Furthermore, he added, "If workable standards do emerge . . . courts should be prepared to order relief." As no such standards emerged in 2005, the court was obliged in 2006 to uphold the 2003 redistricting of Texas that enabled Republicans to gain an additional 6 of that state's 32 congressional seats, for a total of 21.

Since then, numerous redistricting methods have been proposed. Recently, for example, at the City College of New York, Zeph Landau (who is now at the University of California, Berkeley), along with then CCNY undergraduates Ilona Yershov and Oneil Reid, devised an algorithm that exploits the mathematics of fair division [3].

Fair Division Revisited

The "fair division problem" seems to have been treated mathematically for the first time in a 1946 paper by B. Knaster and H. Steinhaus [2]. A particularly simple version concerns the location of the only two hot dog stands licensed to operate on a particular stretch of beach. If one stand is already in place, the second should be located as close as possible to the first, on either the right or the left side, assuming that customers automatically patronize the nearer stand. Hence, the owner of the first stand should choose a location that makes the choices available to the competitor equally acceptable. This is little more than the age-old "you cut, I choose" protocol for splitting a piece of pie or cake two ways. It assumes that no customers from the far ends of the beach are lost because of travel distance.

Fair division problems are more complicated when the property is to be divided among more than two individuals, or when different individuals have different preferences, so that one may prefer share A to share B while another prefers B to A. The mathematics of fair division has grown to considerable proportions in recent years as researchers have worked to incorporate these and other complicating effects. An important milestone was a 1993 proof by the late David Gale that fair divisions ordinarily exist.

Gale defined an allocation of shares to each of n individuals as "fair" if it meets two criteria:

- It is "envy-free," in the sense that no individual prefers another's share to his or her own.
- It is "undominated," in the sense that no alternative allocation assigns each individual a share that is at least equally desirable and, in at least one case, strictly preferable.

Gale then asked whether, for any given set of individual preference relations, there must always exist a fair allocation of cake or pie. The answer is affirmative if $n = 2$, and more nuanced otherwise.

There is no shortage of algorithms for congressional redistricting. Winning papers in the 2007 MCM, for example, described a method based on Voronoi diagrams, a cluster-theoretic approach, and others. All the winning teams tested their methods by partitioning New York State into the required number of plausible districts.

For Gale, a "pie" is the perimeter of the disc $\{z: |z| \leq 1/2\pi\}$, and a "cake" is the unit interval $I = [0,1]$. Hence, pies and cakes are simply the wrapped and unwrapped versions of I. The distinction is important because—under mild technical assumptions—every cake can be divided fairly among n individuals (call them $1, 2, \ldots, n$), whereas some pies cannot.

To express individual preferences, Gale assumed the existence of n finitely additive measures $v_1(\cdot) \geq 0, \ldots, v_n(\cdot) \geq 0$ such that individual i prefers subinterval J of I to J' if and only if $v_i(J) > v_i(J')$, with equality indicating indifference. Gale discovered that for $n > 2$, it suffices to know whether the measures $v_1(\cdot), \ldots, v_n(\cdot)$ are "mutually absolutely continuous," in the sense that if $v_i(J) = 0$ for some $J \subset I$ and $i \in \{1, \ldots, n\}$, then $v_j(J) = 0$ for each $j \neq i$. The answers to Gale's question for $n > 2$ are:

- affirmative for cake if the individuals' preference measures are mutually absolutely continuous,
- negative for cake in the absence of mutual absolute continuity, and
- negative for pie in either case.

Because Gale's proofs are non-constructive, the search is still on for fair division-finding algorithms. Not everyone finds the currently popular "moving knife" procedures satisfactory. An up-to-date exposition of these results can be found in [1].

Recent Algorithms

Because the congressional redistricting process is dominated by just two political parties, Zeph Landau and his colleagues could ignore the difficulties that accompany $n > 2$. But because the 50 states are irretrievably two-dimensional, the algorithms for finding fair divisions of the (wrapped or unwrapped) unit interval were of no obvious value to them. To design a procedure of their own, they began by constructing an example illustrating the means by (and extent to) which redistricting can alter outcomes. In most states, the majority party at census time is entitled to redraw the district boundaries. As shown in Figure 1, the Democrats gain 3 of the 5 available seats by transferring just 6 of the 25 voting precincts between districts.

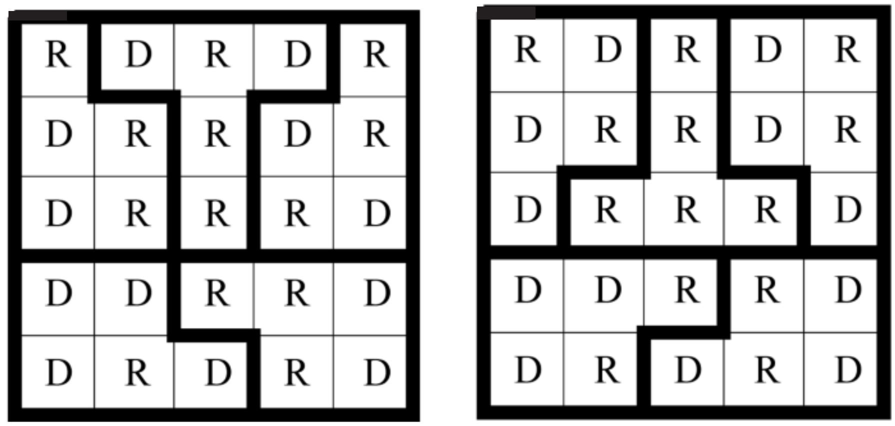

Figure 1. *At stake: a total of 5 districts (surrounded by heavy lines). Before redistricting (left), the Republicans win 4 of the 5; afterward, the Democrats win 4 of the 5. From [3].*

Landau's group sought a districting algorithm that would assign the population of a state to n districts, each containing d people, in a manner equitable to political parties A and B. They began by defining a k-split of a state as a division into two contiguous pieces X and Y in such a way that the population of X is kd; the pair (X,Y) is then called a k-split of the state. The method they propose requires that a disinterested party ∇ first construct, for each k between 1 and $n-1$, a k-split (X_k,Y_k) such that

$$X_1 \subset X_2 \subset \ldots \subset X_{n-1}. \qquad (1)$$

For each k, ∇ then asks both political parties to choose be-tween option (a), which allows party A to divide X_k into k districts and B to divide Y_k into $n-k$ districts, and option (b), which allows A to divide Y_k into $n-k$ districts and B to divide X_k into k districts. Should there exist a k such that the parties A and B both prefer option (a) to option (b) for (X_k,Y_k), or vice-versa, then a division applying the preferred option to that k-split should be created. Should no such k exist, a k^* must be found between 1 and $n-2$ such that party A prefers option (b) to (a) for $k = k^*$ and option (a) to (b) for $k = k^* + 1$. Such a k^* must exist, because A prefers (b) to (a) for $k = 1$ and (a) to (b) for $k = n - 1$. Finally, a fair coin is tossed twice, to decide first whether to set k equal to k^* or to $k^* + 1$, and then whether to apply option (a) or option (b) to (X_k,Y_k). In either case, the result will be an ostensibly fair division of the state into the required number of districts, each containing the required number of citizens.

The obvious objection is that the "disinterested party" ∇ bears a great deal of responsibility for the outcome, in that he or she can choose the finite sequence (1) in many different ways. Landau and colleagues work through an example in which X_k consists of the left-most k columns of the 5×5 grid from Figure 1. But they could as easily have chosen the top-most k rows, or any number of other options. They might, for instance, have selected any Hamiltonian path through the graph whose vertices are the centers of the 25 cells and whose edges connect each vertex to its four (or fewer, for boundary cells) nearest neighbors, allowing X_k to consist of the first $5k$ cells along the path, thereby meeting the contiguity requirements. With many such paths, and thus many ways in which the finite sequence (1) can be chosen, choosing that sequence sounds suspiciously like setting an agenda, which exposes it to the meta-theorem whereby "he who sets the agenda controls the outcome." Donald Saari, in 1995, actually proved a theorem to that effect in a voting context.

There is no shortage of algorithms for congressional redistricting. One of the problems in the 2007 Mathematical Contest in Modeling, for instance, asked the competing teams to devise a method for dividing a state into contiguous districts containing the same number of people, and to apply their methods to New York State. Several of the winning papers are reproduced in the fall 2007 issue of the *UMAP Journal* [4]. One such paper describes a method based on Voronoi diagrams, and another uses a cluster-theoretic approach. A third implements two methods, one of which minimizes the sum of the squares of the distances between the centroids of the various districts and their component census tracts, weighted by population size. All the winning teams tested their methods by partitioning New York State—and in some cases other states as well—into the required number of plausible districts. It would have been instructive if Landau's group had pursued their investigation to that length.

Richard Burkhart—recently retired from Boeing—reports taking a pattern-recognition approach to the problem of identifying voting blocks from election-day returns.* Mathematically speaking, a voting block is a cluster of significant size and density well separated from other such clusters.

A cluster need not reflect political party affiliation, but might correspond to race, religion, gender, education, economic status, or position on such perennial hot-button issues as abortion, health care, gun control, and capital punishment. Using actual ballot data from Northern Ireland, Burkhart employs a fixed-point algorithm to obtain convergence of the cluster sets, represented internally by their centroids. He reports that his algorithm often, but not always, manages to converge to the same final clustering from different initial clusterings, as individual clusters merge. On some occasions the results are remarkably clear-cut; on others the outcome seems to be unduly sensitive to algorithmic details.

In any case, it is far from clear that Justice Kennedy was asking for an algorithmic approach to redistricting when he deplored the lack of "clear, manageable, and politically neutral standards" against which to judge "fairness in redistricting." A non-algorithmic approach to the redistricting problem was proposed in the November 2007 issue of *SIAM News*.

*Personal communication.

References

[1] J. Barbanel, S.J. Brams, and W. Stromquist, *Cutting a pie is not a piece of cake*, Amer. Math. Monthly, 116:6 (2009).

[2] B. Knaster and H. Steinhaus, *Untitled*, Ann. de la Soc. Polonaise de Math., 19 (1946), 228–31.

[3] Z. Landau, O. Reid, and I. Yershov, *A fair division solution to the problem of redistricting*, Social Choice and Welfare, 32:3 (2009), 479–492.

[4] UMAP Journal, Fall 2007.

James Case writes from Baltimore, Maryland.