Some Research Problems in Computational Social Choice

Nicolaus Tideman

It is useful to think of a social choice process as composed of a sequence of sub-processes: nomination, evaluation, message submission, message processing, resolution of ties (if any), and authoritative announcement of the result. Researchers with computational expertise might want to apply their efforts to any of the sub-processes, or to questions of design that involve combinations of the sub-processes. But the questions that stand out as calling for the talents of persons with computational expertise are primarily questions of evaluating the feasibility and attractiveness of message processing rules (vote-counting rules).

1 Questions Associated with the Spatial Model of the Election Universe

Many questions with respect to the evaluation of vote-counting rules require a model of the process that generates election outcomes. Recent evidence suggests that a spatial model is appropriate for this purpose. Consider elections in which voters rank the candidates. (That is, the message that the participants in the social choice process must send is a ranking of the options.) For an election with M candidates, define an "election outcome" as a vector of with $M!$ components, in which each component is the number of voters who placed the candidates in one of the $M!$ possible orders. Three candidates span a space of two dimensions. In this space, assume that voters have ideal points that have a bivariate normal distribution, and that they have circular indifference contours. The space is then divided into six wedges assigned to the six orderings of the candidates. There are five degrees of freedom in the shares of votes going to the different orderings of the candidates, but only four degrees of freedom in the spatial model, so the spatial model is refutable. Evidence indicates that deviations from the spatial model can entirely or nearly entirely be explained by sampling variability. Research questions: Will the results hold for additional data sets? The proportion of voters who know the candidates appears to correlate with how well the spatial model explains the outcome. Can other correlates be identified? What happens when you look at elections with four candidates? With five? With M? Are other versions of the spatial model better? What is the best way to deal with ties that arise in survey data? Implication: Modeling of the consequences of alternative voting rules should be done with the spatial model.

2 Questions Associated with Identifying the Outcome under Rules for Selecting One Candidate from More than Two

A number of voting rules have been proposed for elections with more than two candidates. Some of these rules pose computational problems. Examples: The Condorcet-Kemeny-Young rule potentially requires the evaluation of $M!$ sums. The Ranked Pairs rule (which I devised) poses computational challenges that I could imagine solving only in a very crude and time-consuming way. Are there computationally efficient ways of dealing with the difficult

cases that could occasionally arise under these voting rules? What about the "estimated centrality" rule, which selects the candidate whose estimated spatial location is closest to the center of the distribution of voters' ideal points. Is that rule computationally feasible for more than three candidates? Would someone like to offer a general program that counted votes by a wide variety of rules?

3 Questions Associated with Evaluating the Susceptibility of Voting Rules to Strategizing

The Gibbard-Satterthwaite theorem tells us that all reasonable voting rules are subject to strategy in some instances. There are a number of ways in which the susceptibility of voting rules to strategizing might be measured. What is the best way to measure the susceptibility of voting rules to strategizing? How do different rules compare?

4 Questions Associated with the Single Transferable Vote Form of Proportional Representation

The Single Transferable Vote (STV) is a form of proportional representation in which voters submit rankings of candidates, and votes are counted by a complex algorithm that is intended to identify a winning set of candidates of a specified size that reflects the diversity of preferences in the electorate. There are a number of versions of STV, varying in their sophistication and in their susceptibility to different concerns. There are at least two proposed versions of STV that may be so sophisticated that they might require an unacceptably long time to determine the winners. Thus it is interesting to ask: What are the best computational algorithms for identifying the winning sets of candidates under the highly sophisticated versions of STV? What are the resulting computational times with specified hardware? If the most sophisticated versions of STV pose computational problems that make it impossible to guarantee computability, what are the closest approximations that do permit guarantees of computability?

Nicolaus Tideman
Department of Economics
Virginia Tech
Blacksburg VA 24060, USA
Email: ntideman@vt.edu