Mathematical and Computer Modelling 48 (2008) 1357–1373

Contents lists available at ScienceDirect

Mathematical and Computer Modelling

journal homepage: www.elsevier.com/locate/mcm

Consistency without neutrality in voting rules: When is a vote an average?

William S. Zwicker

Mathematics Department, Union College, Schenectady NY 12308, USA

ARTICLE INFO

Article history:
Received 19 May 2008
Accepted 22 May 2008

Keywords:
Scoring rule
Generalized scoring rule
Mean proximity rule
Mean neat rule
Voting system
Geometry of voting
Convex
Decomposition
Consistent
Connected
Ties
Hyperplane separation
Voronoi

ABSTRACT

Smith [J.H. Smith, Aggregation of preferences with variable electorate, Econometrica 41 (1973) 1027–1041] and Young [H.P. Young, A note on preference aggregation, Econometrica 42 (1974) 1129–1131; H.P. Young, Social choice scoring functions, SIAM J. Appl. Math. 28 (1975) 824–838] characterized scoring rules via four axioms: consistency, continuity, anonymity, and neutrality. In their context a ballot consists of a strict ranking of alternatives, and an election outcome is either a set of (winning) alternatives (Young) or a weak ordering of alternatives (Smith). Many rules fail to fit this context, yet intuitively satisfy one's notion of a ***generalized scoring rule***; this very broad class GSR includes the Kemeny rule, approval voting, and certain grading systems. We show that GSR is identical with the class MPR of ***mean proximity rules*** loosely, rules in MPR are those for which the "average voter" determines the outcome. The techniques in the proof allow us to make some surprisingly direct comparisons between rules (for example, between Kemeny and Borda) that might initially seem to be of completely different sorts. The ***abstract anonymous voting rules*** provide the context for GSR, which is of necessity too general to admit a neutrality axiom. A natural question arises: "What happens to the Smith and Young characterizations in the absence of neutrality?" We discuss one answer in the form of a characterization of the ***rational mean neat voting rules*** (a class closely related to GSR) as those that are consistent and connected. Connectedness is a strong form of continuity that implies a discrete analogue to the Intermediate Value Theorem.

1. Introduction

Our focus is on two new theorems in the mathematics of voting. The *Equivalence Theorem* shows that the wide class of voting rules that can be described in geometric terms via the mean coincides with the class of rules that can be described in terms of scores (in a sense much broader than that traditionally implied by the term "scoring rule"). Our second theorem provides an axiomatic characterization of a closely related class — the rational mean neat rules. In this paper we provide context, definitions, and examples for both theorems, prove the Equivalence Theorem, and indicate some consequences. The long and technical proof of the second theorem appears in [4] (in this issue).

A voting rule in either class is necessarily *consistent* and *continuous*. Loosely, consistency requires of a rule that if two disjoint electorates yield outcomes with some common features, then the rule applied to the combined electorate also yields an outcome with these features. Continuity expresses the idea that sufficiently small changes in the expressed preference of the voters cannot change a loser into a winner. Exact definitions of these properties appear in Sections 3.2 and 3.3. Other technical terms mentioned in the introduction are also defined in subsequent sections.

This line of investigation began with the work of Smith [1] and Young [2,3], who characterized the class of *scoring rules* (in the narrower, traditional sense defined in Example 2.3.2) with four axioms:

E-mail address: zwickerw@union.edu.

0895-7177/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.mcm.2008.05.031

Table 1a
Scoring weights for two "mystery" voting rules (three alternatives)

Ranking	System A weight	System B weight
xyz	4	3
xzy, *yxz*	3	2
zxy, *yzx*	1	1
zyx	0	0

Table 1b
Scoring weights for two "mystery" voting rules (four alternatives)

Ranking	System A weight	System B weight
wxyz	10	6
xwyz, *wyxz*, *wxzy*	9	5
xwzy	8	4
wzxy, *wyzx*, *ywxz*, *xywz*	7	4
wzyx, *yxwz*	6	3
xzwy, *ywzx*	5	3
xyzw, *zwxy*	4	3
xzyw, *yxzw*, *zxwy*, *zwyx*	3	2
yzwx	2	2
yzxw, *zywx*	1	1
zyxw	0	0

- consistency (a.k.a. "reinforcement")
- continuity (a.k.a. "Archimedean property")
- neutrality
- anonymity.

The beauty of these characterizations (along with the proofs) places them among the more mathematically interesting and significant theorems in the theory of voting. These results assume that an individual ballot consists of a strict ranking[1] of some finite set A of alternatives. In Smith's context, an election outcome is a weak ordering of A, which we interpret as the *social ordering*: those at the top of this ordering are the election winner(s) (possibly non-unique, if there is a tie for first place), etc., but each non-winner is also ordered. A voting rule in Smith's context is sometimes called a *social welfare function*. Young dealt with *social choice functions* (also called *social choice correspondences*) in which an outcome consists of a non-empty subset of A, whose member(s) are interpreted as the election winner(s). The different contexts led to slightly different versions of the four axioms.

Myerson [5] was particularly interested in *approval voting*, in which a ballot consists of a set $S \subseteq A$. Each such ballot assigns 1 point to each alternative $a \in S$ and 0 points to each $a \notin S$. The alternative(s) with the greatest point total is the winner. In this sense approval voting may be thought of as a scoring rule, but it does not fit the Smith or Young context because a ballot is not a strict ranking of A. By modifying the definition of neutrality, Myerson was able to extend the Smith and Young characterizations to a broader context that does encompass approval voting. However, Myerson's definition of neutrality implicitly assumes that the points determining the score are awarded to the individual alternatives. That is not necessarily the case for all rules:

Example 1.1. We consider two voting rules, "System A" and "System B". A vote will consist of a strict ranking of the alternatives from some finite set A. Any such vote assigns a scoring weight to each of the strict rankings of A. Table 1a shows these scoring weights for three alternatives when the vote is for *xyz* (short for $x > y > z$, where x is the top-ranked alternative); here "*xyz*" can stand for the three alternatives in A, *taken in any order*. In either system, we sum the weights awarded to a ranking by all voters, and the ranking(s) with the greatest point total is the winning ranking(s).[2] Table 1b provides the weights in the case of four alternatives.

Systems A and B each turn out to be familiar voting rules in disguise. Which ones are they, and what principles lie behind the particular weights? We return to these issues in Section 4.3.

There are a number of other rules that similarly escape Myerson's characterization in [5], and that can be thought of as scoring rules, if one broadens the context suitably:

- The Kemeny rule
- Primitive Pairwise Majority Rule (Condorcet's rule, modified so that an election outcome is a possibly intransitive binary relation)
- Certain grading systems (in which the "election" process averages a student's test scores in a course, and the outcome is a letter grade for that student).

[1] A *strict ranking* of A is a complete (total), transitive, antisymmetric binary relation on A — that is, a linear ordering of A. A *weak ordering* is complete and transitive but not necessarily antisymmetric; there may be "ties" in the ordering.

[2] Conitzer and Rognlie [6] have also considered systems in this context, and refer to them as "ranking scoring systems".

The version of neutrality in Myerson's characterization does not apply to these voting rules, because it requires that the objects to which points are awarded be treated with complete symmetry.[3] In some cases, it seems clear that *no* reasonable version of neutrality can possibly apply.[4] We are confronted by the following situation: if we adopt a context for voting rules that is sufficiently broad to encompass the class GSR of all *Generalized Scoring Rules* (the wide variety of natural amalgamation rules that are based on accumulated scores).[5] We are forced to discard the neutrality axiom. In fact, for the context we have in mind – the *abstract anonymous* voting rules – a satisfying definition of neutrality does not seem to exist. We are led, quite naturally, to the following:

Question. What happens to characterization theorems for scoring rules in the absence of neutrality?

There seem to be two reasonable ways to make this question more precise:

Question A. What class of voting rules is characterized by the remaining three axioms of consistency, continuity, and anonymity?

Question B. How can we characterize the class GSR of generalized scoring rules?

This paper bears on both questions. We do not answer Question B,[6] but in the ***equivalence theorem*** of Section 4.2 we show that GSR coincides with the class MPR of *mean proximity rules*.[7] This result highlights, we feel, the importance and naturality of the class of generalized scoring rules. It is quite broad, including a number of well-known voting rules omitted from the narrower collection of "traditional" scoring rules.

Definition 1.2 (*Preliminary*[8]). A ***mean proximity rule*** is a voting rule that has a ***mean proximity representation***: For some integer n,

- An *input plot function* R locates each possible legal *ballot* v as a point $R(v) \in \mathbf{R}^n$. These points constitute the vertices of the *representing polytope* $P \subseteq \mathbf{R}^n$.
- Each possible *election output* t is plotted as a point in $\mathbf{R}^n$, via an *output plot function* E.
- In any election, we calculate the vector sum of the $R(v)$ for each ballot v cast (counting multiplicity) and divide by the number of ballots cast, to obtain the mean location q of the voters.
- The *election outcome* is the set of those outputs t for which $E(t)$ is closest to q.

When several points $E(t)$ are equally closest[9] to q, we interpret the result as a tie among the several corresponding outputs t. The choice of election inputs (ballots) and outputs, and of n, R, and E, depend on which rule we are representing (see Section 2.3 for a variety of concrete examples). Notice that this definition implies anonymity (symmetric treatment of voters), but not neutrality (symmetric treatment of alternatives).

Some of the specific mean representations of known rules seem interesting in and of themselves, and provide some comparative insight into the nature of these rules (see Section 4.3). But the mean proximity notion has a theoretical appeal as well, in that it admits a natural relaxation to a broader class, the *mean neat rules*. In general, we may think of a mean voting rule as one for which the outcome is determined by the "average voter" — in this sense, a (collective) vote is an average. In mean proximity representations the requirement that outcomes be determined by the closest $E(t)$ to the mean point q can be rephrased as follows: the representing polytope is decomposed into proximity region(s), or *Voronoi cells* (with the cell corresponding to a particular output t consisting of the points of P that are at least as close to $E(t)$ as they are to any other $E(t')$), and the election outcome is determined by which region(s) contain q, the mean point. A simple example is the cutting of the hexagon of Example 2.3.1 (Fig. 2) into six closed "pie slices" via the solid radial lines; each slice is the proximity region for the corresponding vertex of the hexagon.

In any such *Voronoi decomposition*, each pair of regions u and v can be *neatly separated* by some hyperplane h: all points of u lie on h or to one side, all points of v lie on h or to the opposite side, and $u \cap h = v \cap h$ (Fig. 1); see Cervone and Zwicker [8, 4] for further discussion. We obtain the class of mean neat rules when we associate each output t to a region of P, drop the demand that these output regions be determined by relative proximity to a selection of points, but retain the (strictly weaker) requirement that each pair of regions be neatly separable by some hyperplane.

In Section 5, we conjecture that a voting rule is mean neat if and only if it is both consistent and continuous; this would answer Question A. One direction is straightforward: all mean neat rules are consistent and continuous. Our evidence for the converse comes in the form of a closely related characterization theorem:

[3] Myerson [5] requires that for each permutation τ of these objects, there is a corresponding permutation σ of ballots such that for each election, if we permute the ballots by σ the election result is permuted by τ. He refers to the problem of characterizing scoring rules without this neutrality axiom, but not to the issue of assigning points to objects other than alternatives.

[4] See further discussion in Example 2.3.6.

[5] The definition is in Section 3.1

[6] It would be desirable to answer B by providing an axiomatic characterization. With Juan Enrique Martinez Legaz, we have made some progress, in the form of some axioms that hold of all generalized scoring rules. These axioms are, in the absence of neutrality, strictly stronger than consistency.

[7] The *Euclidean Distance-from-Unanimity Rules* provide a third, equivalent characterization; these are rules for which the outcome of profile p is determined by which unanimous profile is closest to p, when closeness is determined via some embedding into Euclidean space. See Meskanen and Nuurmi [7] for a more general treatment of this idea.

[8] See Section 4.1.

[9] "Closest" will always refer to the Euclidean metric, except where we specify otherwise.

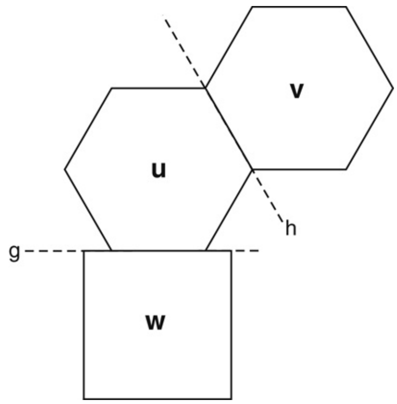

Fig. 1. Regions u and v are neatly separated by hyperplane h. Regions u and w are not neatly separated by g (and cannot be neatly separated).

Theorem 5.4. *An abstract anonymous voting rule is a rational mean neat rule if and only if it is both consistent and connected.*

A mean neat rule is *rational* if the output regions can be neatly separated via hyperplanes that contain a dense set of points with rational coordinates. The proof of 5.4 appears in [4], but the relevant concepts are developed here. *Connectedness*, a new property, can be seen both as a strong form of continuity that implies a certain discrete analogue of the Intermediate Value Theorem; loosely, it asserts that when two elections yield completely different outcomes, there must exist some intermediate profile with a tie in the outcome. It serves as a natural partner to consistency, in that it guarantees the existence of exactly those situations – ties – governed by *down consistency*, one of two complementary weak forms of consistency.

Our results build on two additional bodies of work in the literature. Barthélemy and Monjardet [9] showed that several "median methods", such as the Kemeny voting rule, can be expressed via geometric representations via the mean.[10] In fact a good part of our proof of the Equivalence Theorem in Section 4.2 consists of the observation that their application of a theorem on the mean due to Christiaan Huygens (appearing in their 1981 paper, and stated here as 4.2.2) extends to the general case. Anyone familiar with the work of Saari [10], and of Saari and Merlin [11], will recognize the significant debt we owe to their geometric perspective. One example is our use, in 2.3.4 and 2.3.5, of Saari's *Representation* (hyper) *Cube*.

The rest of the paper is organized as follows. In Section 2 we discuss several examples of voting systems and mean representations. Section 3.1 discusses the *abstract anonymous voting systems* of [5] — the very broad class to which our main result applies. The definition of this class deliberately blurs the distinctions between contexts that are typically treated separately in the literature. While this is essential for our purposes, a consequence is that comparisons between our results and others should be approached cautiously, with attention to the issue of context.

Section 3 continues by introducing the important subclass of generalized scoring rules, and the versions of the consistency, continuity, and connectedness properties that we employ in the *abstract anonymous* context. In Section 4 we introduce the class of mean proximity rules, prove the equivalence theorem, and mention some consequences for representations of certain specific voting rules. Section 5 introduces the mean neat and rational mean neat rules, and states our result and conjecture on axiomatic characterizations of these classes.

2. Examples of voting systems and mean proximity representations

2.1. Social "welfare/preference" functions

Before presenting our examples, we need to settle a somewhat subtle issue of context. A rule such as the Borda count may be thought of either as a ***social choice function*** (in which an election outcome is required to be a non-empty subset S of our finite set A of alternatives with S containing a unique alternative when there are no "ties" for first place), or as a ***social welfare function*** (in which an election outcome is required to be a single weak ordering σ of A, which is strict when there are no "ties" for first place, or for second, etc.). There is a third context as well — that of a ***social preference function***, introduced by Young and Levenglick [12] to describe the situation for the Kemeny rule, in which an election outcome is a set X of strict rankings of A (which contains a single strict ranking when there are no ties in the outcome). Notice that a weak ordering σ can be identified with the set $T(\sigma)$ of all strict extensions of σ : $T(\sigma) = \{\tau | \tau$ is a strict ranking of A satisfying, for all $a, b \in A$, that if $a <_\sigma b$ then $a <_\tau b\}$. In effect, $T(\sigma)$ breaks all ties in all possible ways. Thus any social welfare function can be

[10] Their representation of the Kemeny rule is different from ours. Note that our use of Barthélemy and Monjardet's term "median method" is likely to cause confusion in the current context. The methods they discuss are *medians* with respect to the Kendall metric, yet some prove to be *means* when expressed in terms of the Euclidean metric. We have also considered voting rules that apply a form of spatial median to the Euclidean metric, and this approach yields quite a different kettle of fish.

converted into a social preference function for which every election outcome is a ***tie-generated*** set X of strict rankings: $X = T(\sigma)$ for some strict ranking σ. We will use the term ***social welfare/preference function*** to refer to a social welfare function that has been converted in this way.[11] Of course, not every set X of strict rankings is tie-generated; in fact, it is easy to see that the Kemeny rule is not a social welfare/preference function.

2.2. The permutation proximity lemma

Here we develop a simple result about Euclidean distance that will help us show that our later mean representations do what they purport to do. Suppose $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ are points of $\mathbf{R}^n$ and we would like to find the permutation of y's coordinates closest to x. For any ρ in the symmetric group S_n on n elements let y_ρ denote $(y_{\rho(1)}, y_{\rho(2)}, \ldots, y_{\rho(n)})$, and let Πy be $\{y_\rho | \rho \in S_n\}$. We seek the element(s) of Πy minimizing $\|x - y_\rho\|$.

We will start with the special case in which the coordinates of x are all distinct, as are those of y. We claim that in this case the desired ρ is the one enumerating y's coordinates in the same relative order as x's coordinates, so that

$$x_i < x_j \Leftrightarrow y_{\rho(i)} < y_{\rho(j)} \tag{1}$$

holds for all i and j. For example, if $x = (1, 3, 2)$ and $y = (500, 500.1, -13)$, $y_\rho = (-13, 500.1, 500)$ satisfies condition (1). Any other permutation would fail condition (1) for some $i \neq j$, at which point transposing $y_{\rho(i)}$ and $y_{\rho(j)}$ would reduce the distance to x; for any four numbers with $a < b$ and $c < d$, $(a-c)^2 + (b-d)^2 < (a-d)^2 + (b-c)^2$. To obtain the desired ρ one may enumerate the coordinates of x and of y in strictly decreasing order as $x_{\tau(1)} > x_{\tau(2)} > \cdots > x_{\tau(n)}$ and $y_{\sigma(1)} > y_{\sigma(2)} > \cdots > y_{\sigma(n)}$, and then set $\rho = \sigma \circ \tau^{-1}$.

The situation is a bit more complicated when the coordinates of one or both points fail to be distinct. In this case, of course, condition (1) may be impossible to meet. Instead we ask for a weaker condition.

Definition 2.2.1. If x and y are points of $\mathbf{R}^n$ we will say that a permutation ρ ***enumerates y 's coordinates in an order compatible with x*** if

$$x_i < x_j \Rightarrow y_{\rho(i)} \leq y_{\rho(j)} \tag{2}$$

holds for all i and j.

Condition (2) is equivalent to requiring that for no i and j do we ever have a "crossover" with $x_i < x_j$ and $y_{\rho(i)} > y_{\rho(j)}$; such an offending i and j would yield a transposition providing a different, closer permutation. Any ρ meeting condition (2) can be obtained as $\sigma \circ \tau^{-1}$, where σ and τ^{-1} enumerate x and y, respectively, in non-increasing order: $x_{\tau(1)} \geq x_{\tau(2)} \geq \cdots \geq x_{\tau(n)}$ and $y_{\sigma(1)} \geq y_{\sigma(2)} \geq \cdots \geq y_{\sigma(n)}$. In this case σ and τ need not be unique, so multiple ρ may exist. This can lead to several distinct points y_ρ in Πy that are equally close to x (but it can also happen that several different permutations ρ yield the same point y_ρ). We sum up the situation in the following lemma:

Lemma 2.2.2 (*Permutation Proximity Lemma*). *Let x and y be points of $\mathbf{R}^n$. Then the permutations $y_\rho \in \Pi y$ of y that are closest to x are those for which ρ enumerates y's coordinates in an order compatible with x.*

We leave the detailed proof to the reader.

2.3. Voting systems with mean proximity representations

Example 2.3.1 (*The Borda Count for Three Alternatives p, q, and r*). Each voter submits a strict preference ranking of these three, awarding $+2$ points, 0 points, and -2 points to their favorite, middle, and least favorite alternative, respectively.[12] We will begin by representing the Borda rule as a social welfare/preference function in the sense of Section 2.1; we obtain the weak ordering σ of the alternatives in descending order of point totals and convert it into the corresponding set $T(\sigma)$ of strict rankings.

Here is a mean proximity representation: recalling from Definition 1.2 that the maps R and E locate respectively each election input (ballot) and output (non-tied election result) as points in Euclidean space, fix any arbitrary reference strict ranking σ of A – say $\sigma = pqr$ – and plot σ as the point $R(\sigma) = (2, 0, -2)$ in $\mathbf{R}^3$. Plot any other strict ranking τ as the corresponding vector of points awarded, by τ, to p, q, and r respectively *taken in the reference order*. For example $\tau = qrp$ would yield $R(\tau) = (-2, 2, 0)$. The points $R(\tau)$ for the six possible strict rankings all lie on the plane $x_1 + x_2 + x_3 = 0$, and are the vertices of the regular hexagon appearing in Fig. 2, which also shows the result of the sample election of Table 2.[13]

---

[11] Smith [1] and Young [2] provide a definition of consistency for social welfare functions. Consistency (in their sense) of a social welfare function Y implies consistency (as we define it in Section 3) of the converted rule Y^*, but the converse is false.

[12] See the Observation in Example 2.3.2.

[13] Although Fig. 2 depicts the hexagon flat on the page, it is impossible to draw a regular hexagon in $\mathbf{R}^2$ so that every vertex gets rational coordinates. Given the important role of rational coordinates in the general theory, it is best to think of this hexagon as lying on the plane $x + y + z = 0$ in $\mathbf{R}^3$.

Table 2
Each underlined number counts the votes for the ranking appearing beneath it

Profile #1	0	0	3	0	5	0
	p q r	p r q	r p q	r q p	q r p	q p r

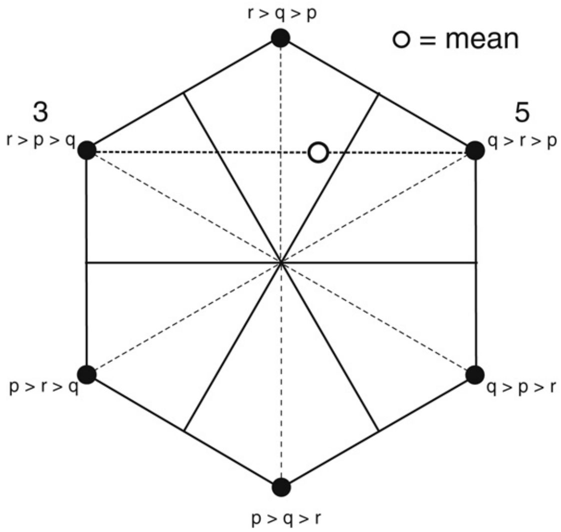

Fig. 2. A mean proximity representation of the Borda count for three alternatives p, q, and r.

The solid radial lines divide the hexagon into six proximity regions. Each region contains the points of the (filled in) hexagon that are at least as close to the vertex of that region as they are to any other vertex. The 8 votes from Table 2 appear as labels on the corresponding vertices. By adding the 8 vectors corresponding to the coordinates of these vertices, and dividing by 8, we get the mean location of the plotted votes. It lies in the proximity region of the ranking rqp, which is thus the "winning" strict ranking. Different profiles can place the mean at any rational point lying in the convex hull of the six vertices. In this representation, each election input and output is a strict ranking, and the functions R and E are the same.

The permutation proximity Lemma 2.2.2 explains why this geometric calculation will agree with that for the standard three-alternative Borda count, for every possible profile. If we identify the mean location of the plotted votes with the point "x" mentioned in the lemma, then x_1 is the average number of points given alternative p by the voters, x_2 is the average for q, etc. Suppose, for example, that q gets the highest Borda score, then p, then r. By Lemma 2.2.2 the closest permutation of the point $y = (2, 0, -2)$ to x will be $(0, 2, -2) = R(qpr)$, because the coordinates have the same relative order as those of x. So $\{qpr\}$ will be the election outcome according to the mean representation. But if r's score is tied with p's as second largest then $(0, 2, -2)$ and $(-2, 2, 0)$ will be equally close, yielding an election outcome of $\{qpr, qrp\} = T(q > [p, r])$ (see Section 2.1).

The mean proximity representation for the Borda count as a social choice function is closely related. We use the same input plot function R, but the election outputs become the individual alternatives. One option for our output plot function E is to locate each alternative at the midpoint of the two hexagon vertices corresponding to the strict rankings that place that alternative on top. Thus

$$E(p) = (2, -1, -1), \qquad E(q) = (-1, 2, -1), \quad \text{and} \quad E(r) = (-1, -1, 2).$$

Alternatively, we can set

$$E^*(p) = (1, 0, 0), \qquad E^*(q) = (0, 1, 0), \quad \text{and} \quad E^*(r) = (0, 0, 1).$$

The proximity regions for E and E^* are the same; they are subregions of the original hexagon, and each is the union of two bordering "pie-slices" of Fig. 2. The argument that this representation agrees with the standard Borda count is again a straightforward application of Lemma 2.2.2.

Example 2.3.2 (*All "Traditional" Scoring Rules*). In a (simple)[14] scoring rule, each voter casts a ballot consisting of a strict ranking of the f alternatives. Each rank is pre-assigned a fixed real number ***scoring weight***, and these weights form the

[14] We are using "simple" to distinguish these systems from the compound scoring systems of [1,2], which allow a hierarchy of tie-breaking scoring vectors. The generalized scoring rules we discuss here are all simple. Saari [10] uses the term "positional voting" for simple scoring systems. However, there are voting methods that are not scoring systems, yet are defined solely in terms of positional information.

scoring vector $w = (w_1, w_2, \ldots, w_f)$, which is required to satisfy $w_1 \geq w_2 \geq \cdots \geq w_f$ and $w_1 > w_f$. Each voter awards to their top-ranked alternative w_1 points, to their second ranked candidate w_2 points, etc. Each alternative receives the total of all points awarded. To obtain a social welfare/preference function we take the weak ordering of alternatives in descending order of point totals, and convert it as in the earlier remark on social welfare/preference systems. For the social choice function version, we choose all alternatives that share the highest total score achieved. The Borda count and the plurality voting rule, with scoring vector (1, 0, 0, …, 0), are important special cases. One useful fact is contained in the following, well-known observation.

Observation. Scoring vectors w and v induce the same social choice (and social welfare) function if and only v is a ***positive affine transform*** of $w : v = \alpha w + (\beta, \beta, \ldots, \beta)$ for some real constants $\alpha > 0$ and β. In particular, the Borda count is induced by *any* w for which $w_i - w_{i+1}$ is a fixed, strictly positive amount for all $i = 1, \ldots, f - 1$.

To obtain a mean proximity representation for a general w, choose an enumeration $\{a_1, a_2, \ldots, a_f\}$ of the set A of alternatives. Let F denote $\{1, 2, \ldots, f\}$. A vote is then a bijective map $\sigma : F \to F$, corresponding to a voter who ranks alternative a_j in the "$\sigma(j)$th position" (where the first position is assigned to that voter's most favored alternative, and the fth position is assigned to her least favored). The input plot function R_w takes any vote σ and positions it at the location $(w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(f)})$. However, regardless of the particular weight vector w at hand, for our output plot function we will always use $E_w = R_{\text{Borda}}$ — the function defined as in the previous sentence, but using any weight vector with equally spaced weights,[15] such as

$$w_{\text{Borda}} = (n, n-1, n-2, \ldots, 1).$$

A strict ranking is then in the winning set if its plotted position according to E_w is among those closest to the mean location q of the plotted votes.

We need a different E, however, if we wish to represent a scoring rule as a social choice function. In this case we set, for each individual alternative a_j, $E_{\text{BordaSC}}(a_j)$ to be the vector mean of all points $R_{\text{Borda}}(\sigma)$ for which σ is a strict ranking in which a_j is top ranked. Alternately, we can use $E^*(a_j) = e_i$, the basis vector with a "1" as its ith component and zeros elsewhere.

These representations are particularly appealing in the case of the Borda count, because of the high level of symmetry. If we apply the above construction to w_{Borda} then we obtain, as the images of all possible ballots under R_{Borda}, the vertices of a certain well-known $(n-1)$-dimensional polytope in $\mathbf{R}^n$ known as the *n-permutahedron*, and we can label each vertex with the corresponding strict ranking of A. The 3-permutahedron is just the regular hexagon (as we have seen in 2.3.1), while the 4-permutahedron is a truncated octahedron (see, for example, http://en.wikipedia.org/wiki/Truncated_octahedron): an Archimedean solid with 8 hexagonal faces and 6 square faces. Each hexagonal face corresponds either to the 6 possible permutations of 4 alternatives that rank some designated alternative a_i first (in which case $E_{\text{BordaSC}}(a_i)$ is located at the center of that face), or to the 6 that rank some designated alternative last.

Thus, the Borda count for four alternatives can be thought of as follows: by strictly ranking the alternatives each voter selects, in effect, a vertex of the truncated octahedron. We find the vector mean q of all the selected vertices (counting multiplicities) and look for the vertices that are closest to q. The labels on these vertices constitute $T(\sigma)$ for some weak ordering σ, and σ coincides with the weak ordering induced by total Borda scores. Alternately, we can seek the individual alternatives whose locations (at the centers of four of the hexagonal faces) are closest to q; these constitute the set of Borda winners.

Example 2.3.3 (*Approval Voting*). In approval voting (see Brams and Fishburn [13]), a vote is a subset S of the set A of alternatives (containing those alternatives of which the voter "approves"). Voters may approve differing numbers of alternatives. The social welfare/preference version ranks alternatives in descending order of the number of approving voters, and then uses the $T(\sigma)$ construction; the social choice version chooses the alternatives sharing the highest number of approving voters.

In our representation, for each $S \subseteq A$ let $R_{\text{Approval}}(S)$ equal the vector average of all points $R_{\text{Borda}}(\sigma)$ taken over all strict rankings σ for which every alternative in S is ranked over every alternative not in S. Once the vector mean q of the plotted votes is found, for the social welfare/preference version of approval voting we again set $E_{\text{Approval}}(\sigma) = R_{\text{Borda}}(\sigma)$, choosing the σ yielding the closest such point to q. For the social choice function representation, we again proceed as in the previous example, by using E_{BordaSC} or E^* as our output plot function E_{Approval}.

Example 2.3.4 (*The Kemeny Rule*). This system, proposed in [14], is closely related to several independent constructions that use a type of discrete median (see [9]). An input vote is again a strict ranking σ of the set of alternatives. The rule employs the (non-Euclidean) ***Kendall metric*** d defined over rankings as follows: $d(\sigma, \tau)$ is equal to the number of pairwise disagreements between the strict rankings — the number of pairs (a_i, a_j) for which $i < j$ and σ and τ differ as to which of these two alternatives a_i and a_j is ranked over the other.

[15] In fact, any weight vector with strictly decreasing weights would do as well; the reason should be clear from Lemma 2.2.2, which applies just as in the previous example.

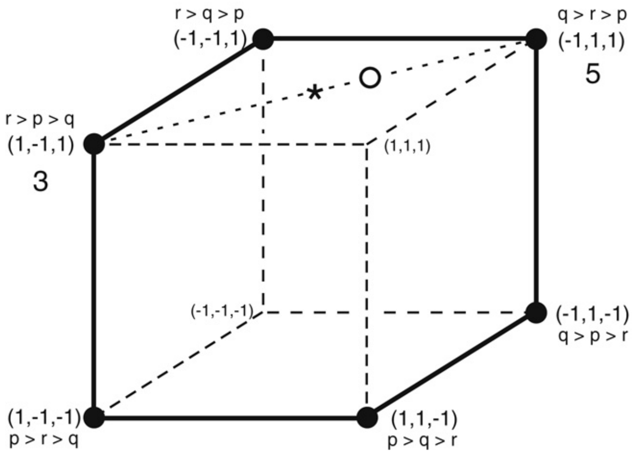

Fig. 3. A mean proximity representation of the Kemeny rule for three alternatives.

This metric extends to a measure D of "total distance" between an individual strict ranking σ and a profile p – that is, an election in which each voter $v \in N$ chooses a strict ranking $p(v)$ of the alternatives – by summing: $D(p,\sigma) = \sum_{v\in N} d(\sigma, p(v))$. The Kemeny election outcome is the set of strict rankings σ that minimizes this total distance $D(p,\sigma)$. (The rule is a "median" in that D is given as the sum of *distances*, rather than of *squared* distances.)

Before we give the details of the geometric representation for an arbitrary number of alternatives, we illustrate the special case for three alternatives, p, q, and r, in Fig. 3. The 6 possible strict rankings are plotted as 6 of the 8 vertices of a $2 \times 2 \times 2$ cube in $\mathbf{R}^3$. These 6 are the *transitive* vertices. The two "unused" *intransitive* vertices of this cube turn out to correspond to the two possible cycles: $p > q > r > p$, and $r > q > p > r$. Each strict ranking σ is plotted at the point $R(\sigma)$ indicated in the figure, the vector mean q of these points is determined, and the outcome is declared to be the set of strict rankings that label the vertices (from among the *transitive* vertices) closest to q.

Fig. 3 also shows the 8-voter election from Table 2, and the mean location of the plotted votes (again shown as a hollow circle). It is easy to see that in this case the (sole) closest transitive vertex is $(-1, +1, +1)$, corresponding to $q > r > p$, which is thus the Kemeny outcome. Recall that with the Borda count, the hexagon in Fig. 2 showed that the outcome of the same election was different: $r > q > p$.

In fact, the 6 transitive vertices of Fig. 3 do form a (non-planar) hexagon, and the order in which the strict rankings are plotted around this hexagon is the same as in Fig. 2, but the Fig. 3 hexagon is bent, so that the vertex angles are 90° rather than 120°; as a result, the vertex labeled $r > q > p$ in Fig. 3 is pulled away from the line joining the two vertices that actually receive any votes. Thus, the $r > q > p$ vertex is now relatively farther from the mean than it was in Fig. 2, and is no longer the closest.

For the general case, assume A is our set containing f alternatives. Choose any set $G \subseteq A \times A$ such that for each pair (x, y) of distinct alternatives of A, exactly one of (x, y) or (y, x) is a member of G. Choose any enumeration $\left((b_1, c_1), (b_2, c_2), \ldots, (b_g, c_g)\right)$ of G, with $g = \binom{f}{2}$. Any strict ranking σ now generates a vector $R_{\text{Kem}}(\sigma)$ of $+1$s and -1s, according to whether σ ranks b_i over c_i:

$$R_{\text{Kem}}(\sigma)_i = \begin{cases} +1, & \text{if } \sigma(b_i) > \sigma(c_i) \\ -1, & \text{if } \sigma(b_i) < \sigma(c_i). \end{cases}$$

Some vectors of $+1$s and -1s are not ***transitive*** — they are not images under R_{Kem} of any transitive σ. Thus our input plot function R_{Kem} places each strict ranking σ at a vertex of a certain hypercube, but only the transitive vertices are used. This same map R_{Kem} is the output plot function, and so the election outcome is a set containing one or more such σ.

Theorem (*Barthélemy and Monjardet [9]*). *The Kemeny rule coincides with the mean proximity rule described above.*

Example 2.3.5 (*The "Primitive" Condorcet Rule*). The Marquis de Condorcet [15,16] suggested that alternatives be ranked by *pairwise majority rule* — that alternative a_i be ranked over alternative a_j when a strict majority of voters place a_i over a_j in the individual strict rankings that they submit as votes. Condorcet knew, of course, that this rule can lead to cycles for some profiles and, more generally, to intransitive binary relations. The question of what to do when this happens – which "Condorcet extension" to use when choosing election outcomes for such profiles – has been considered by many authors.

The method we are calling the "primitive Condorcet rule" ducks this issue entirely. Loosely, it declares the election outcome to be the majority preference relation, whether or not it this relation proves to be transitive. The resulting rule is completely impractical, then, as a real voting system, but its geometric representation sheds some light on the issues raised in this paper, and on the role of the Kemeny rule as the unique consistent and neutral Condorcet extension [12].

In fact, we need to be a bit more careful in defining the primitive Condorcet rule: the outcome of a primitive Condorcet election is the set of all *tournaments* (anti-symmetric and complete binary relations) B such that B contains the strict pairwise majority relation M. Thus, if there are exact ties in majority preference in one or more pairs of alternatives, our rule yields all possible ways to break all such ties.

For our mean proximity representation of this rule in the special case of three alternatives, we again refer the reader to Fig. 3. The input plot function R for the primitive Condorcet representation is that shown in the figure — exactly the same as for the Kemeny rule. The output plot function E for the primitive Condorcet rule is an extension of this R; it employs R for the strict rankings and maps the two possible cycles to the remaining two intransitive vertices of the cube, with the cycle $p > q > r > p$ mapped to $(1, 1, 1)$ and the reverse cycle mapped to $(-1, -1, -1)$.

In some ways the idea is more clear if we imagine starting with this mean proximity representation for the primitive Condorcet rule. To obtain the representation of the Kemeny rule, discard the two cycles as potential outcomes, discarding as well as the corresponding pair of intransitive vertices. Now consider the $1 \times 1 \times 1$ subcube consisting of those points of the bigger cube that are at least as close to the $(1, 1, 1)$ vertex as to any other vertex. Once the intransitive vertices are discarded, the points in this subcube get divided up and parceled out to the proximity regions of three of the six remaining transitive vertices, while the points in the subcube corresponding to the other cycle are similarly reassigned to the proximity regions of the other three.[16]

If we allow an arbitrary number f of alternatives, the mean proximity representation of the primitive Condorcet rule should now be clear. The input plot function R is exactly the same as defined earlier, for the Kemeny rule. The output plot function E extends R by applying the definition given earlier for R to an arbitrary antisymmetric and total binary relation B, in the obvious way:

$$E(B)_i = \begin{cases} +1, & \text{if } (b_i, c_i) \in B \\ -1, & \text{if } (c_i, b_i) \in B. \end{cases}$$

Example 2.3.6 (*A Mean Proximity Grading System*). Suppose that in Mathematics 101 each student takes several exams during the semester, with each graded on a scale from 0 to 100 (and no fractional points awarded). These grades are then amalgamated to determine a letter grade for the course, with the possible letter grades being A (the highest course grade possible), B, C, D, and F (failing).

One way to amalgamate is to assign standard scores to each letter grade (say, $A = 95$, $B = 85$, $C = 75$, $D = 65$, $F = 55$), and to award as course grade the letter whose standard score is closest to a student's average test grade. This is close to what is actually done in many colleges and universities in the United States.[17] The example may not look, at first, like a voting system, but it fits the broad definition we introduce in Section 3; the "voters" are the exams, they "vote" on a particular student's level of knowledge, and the outcome of the "election" is that student's course grade. In fact, related grading systems, along with weighted yes–no voting rules with abstention, provide important examples of voting with multiple levels of approval (as developed in Freixas and Zwicker [18,19] and in [20]) and are mean proximity rules.

For this example, the one-dimensional mean proximity representation is implicit in the original description: the input plot function R is the identity function and the output function E is the assignment of standard scores to letter grades. The set $\{0, 1, \ldots, 100\}$ of inputs is quite different from the set $\{A, B, C, D, F\}$ of outputs, and so neutrality is a meaningless concept; if we consider the transposition τ of the letter grade A with the letter grade D (with all other course grades fixed), there is no corresponding permutation σ of $\{0, 1, \ldots, 100\}$ with the property that whenever σ is applied to a vector of test scores earned by a student, the effect on that student's final grade X is to change it to $\tau(X)$.

Example 2.3.7 (*The Pythagoras System*). There are two alternatives, Pythagorus and Archytas. Each voter casts a ballot for one of the two. Archytas wins if the ratio

$$\frac{\text{number of votes cast for Archytas}}{\text{number of votes cast for Archytas} + \text{number of votes cast for Pythagoras}}$$

exceeds $\sqrt{2}$, and Pythagoras wins otherwise. This example is a "toy" voting method (as is the somewhat similar 2.4.4). It is consistent and continuous but not connected; mean proximity and mean neat, but not rational mean neat.

2.4. Voting systems with no mean neat representation

As we see in Sections 3.2 and 3.3, each of the following rules fails to be consistent or to be continuous. The known direction of Conjecture 5.5 tells us that these rules lack mean neat representations.

---

16 Our debt to [10,11] is clear. There the reader will find related statements, and pictures. However, in their description of the Kemeny rule these authors do not describe these parceled out sections of subcubes in terms of proximity to vertices, but rather in terms of proximity to other subcubes. Consequently (see here [17]) they must use the l_1 metric rather than the Euclidean (l_2) metric. It seems to us that the l_2 version yields additional information as to the geometric relationship between Condorcet and Kemeny, and insight into the connection between the Young and Levenglick characterization and this geometry.

17 Much of the rest of the world employs a single, final examination. In one respect Example 2.3.6 is *not* very realistic — it yields "ties" when test averages fall exactly halfway between two standard scores. See the discussion in Section 3.3.

Example 2.4.1 (*The Copeland Rule*). An alternative a_i is ranked over a_j if the *number* of alternatives a_k for which a strict majority of voters rank a_i over a_k in their individual strict rankings is strictly greater than the number of alternatives a_k for which a strict majority of voters rank a_j over a_k. We obtain a social welfare/preference function as in Remark 2.1, by converting this single weak ordering into an equivalent set of strict rankings. Alternately, as a social choice function the Copeland rule chooses the alternatives who defeat the greatest number of other alternatives in the strict pair-wise majority sense. If there exists an alternative who defeats every other alternative in this sense (a ***Condorcet alternative***, or ***Condorcet winner***), then this alternative will be the sole Copeland social choice; the Copeland rule is a ***Condorcet extension***. The rule is particularly simple in the special case of three alternatives with an odd number of voters (so that the pair-wise majority rule relation M is total). In this case the Copeland *ranking* is either the Condorcet ranking or, in the case of a majority cycle, is a completely tied weak ordering (which gets converted to a six-way tie among all possible strict rankings). The Copeland social *choice* is either the Condorcet alternative or, in the case of a majority cycle, is the set containing all three alternatives. In both social choice and social welfare/preference forms, Copeland fails to be consistent.

Example 2.4.2 (*The Hare System*). This system, proposed by Hare [21] is also known as a special case of "single transferable vote" (or "STV"). It is most common in countries that once belonged to the British Empire and is one of the more popular alternative voting system among amateur societies that favor voting reform. In Hare, an alternative is eliminated from each voter's strict ranking if it is among the alternatives top-ranked least often by the voters; imagine striking such alternatives from each voter's list and "closing up" the list, if need be, to fill the gap. Several alternatives, or even all the alternatives, may be eliminated simultaneously. These elimination rounds are repeated until all alternatives have been eliminated. The Hare social welfare function lists alternatives in reverse order of elimination, with the final batch eliminated ranked at the top. As before, we will convert the resulting weak ordering to a corresponding set of strict rankings to obtain a social welfare/preference function. The Hare social choices are the alternatives eliminated in the last round. As a social choice function, Hare is not consistent; as a social welfare/preference function, it is not connected.[18]

Example 2.4.3 (*The Omninominator Rule*). Each voter ranks all alternatives. An alternative is among the winners if and only if she is top-ranked by at least one voter. This social choice function plays a role in the study of manipulation (see Taylor [22]). It satisfies an important weak form of consistency, but is not fully consistent.

Example 2.4.4 (*The Almost-Dictatorship System*). There are two alternatives, Dictator and Opposition. Each voter casts a ballot for Dictator, or for Opposition. Opposition wins if every voter, without exception, votes for Opposition, and Dictator wins otherwise. This system was constructed to be consistent but not continuous; it has no mean neat representation.

3. Consistency, continuity, and connectedness in abstract anonymous voting systems

3.1. Abstract anonymous voting systems

The following two definitions are implicit in [5]:

Definition 3.1.1. Let $\mathcal{I}$ be a finite set (of ***input*** ballots) and $\mathcal{O}$ be a finite set (of election ***outputs***).[19] Let $\mathbf{Z}^+$ denote the set of non-negative integers. An ***anonymous profile***, henceforth ***profile***, is a function $p : \mathcal{I} \to \mathbf{Z}^+$ satisfying $N(p) > 0$, where $N(p) = \sum_{j \in \mathcal{I}} p(j)$ denotes the number of voters in the profile. We will use $(\mathbf{Z}^+)^{\mathcal{I}}$ to denote the set of all such profiles. An ***outcome*** is a non-empty subset of $\mathcal{O}$, representing a potential election result; $C(\mathcal{O})$ denotes the collection of all such outcomes.

Definition 3.1.2. An ***abstract anonymous voting system*** $\mathcal{V}$ is a triple $(\mathcal{I}, \mathcal{O}, \mathcal{F})$, in which $\mathcal{I}$ and $\mathcal{O}$ are nonempty finite sets, and $\mathcal{F} : (\mathbf{Z}^+)^{\mathcal{I}} \to C(\mathcal{O})$ is a function that specifies an election outcome for each possible profile.

The high level of abstraction encompasses the broad range of examples in the previous section, and is different from what we normally see in the theory of voting. In particular, these definitions impose no structure on the set $\mathcal{I}$ of input ballots or on the set $\mathcal{O}$ of election outputs; a member of either set may be anything whatsoever, including:

- an individual alternative (as in Plurality voting, wherein each voter may vote for a single candidate or alternative), or
- a set of alternatives (as in Approval voting), or
- a strict ranking of alternatives (as in the Borda count, or Kemeny rule), or
- an integer between 1 and 100 (as in some grading systems).

18 An alternative form of Hare in the social preference context has been suggested independently by Merlin [23,24], and by Conitzer and Rognlie [6]. It finesses the pathologies caused by simultaneous elimination, and is a (consistent) compound scoring rule.

19 Examples such as Kemeny, Primitive Condorcet, and grading systems suggest that it is appropriate to distinguish between the terms "alternative" and "election output".

Table 3
Some voting rules placed in the abstract anonymous context

	Voting Rule	I	O	in ***GSR***?	Rational Mean Neat rule?
1	Approval Voting (2.3.3)	$\mathcal{P}(A)$	A or $SR(A)$	Yes	Yes
2	Kemeny Rule (2.3.4)	$SR(A)$	$SR(A)$	Yes[a]	Yes
3	Plurality	A or $SR(A)$	A or $SR(A)$	Yes	Yes
4	Primitive Condorcet (2.3.5)	$SR(A)$	TAB(A)	Yes[a]	Yes
5	Grading System (2.3.6)	$\{1, 2, \ldots, 100\}$	$\{A, B, C, D, F\}$	Yes[a]	Yes
6	Any "Traditional" Scoring Rule	$SR(A)$	A or $SR(A)$	Yes	Yes only if there exists a rational vector of scoring weights
8	Almost Dictator (2.4.4)	{Dictator, Challenger}	{Dictator, Challenger}	No	No
9	Pythagoras (2.3.7)	{Pythagoras, Archytas}	{Pythagoras, Archytas}	Yes[a]	No

Here A is the set of alternatives, $SR(A)$ is the set of strict rankings of A, $\mathcal{P}(A)$ is the power set of A, and TAB(A) is the set of total (complete) and antisymmetric binary relations (tournaments) on A.

[a] For table rows 2, 4, 5, and 9, scoring weights can be derived from the mean proximity representation, as discussed in Section 4.3.

The distinction between outputs and outcomes is important; if $\mathcal{F}(p) = \{g\}$ then g is the sole winning output of the election, and if $|\mathcal{F}(p)| > 1$ then the outcome is a tie among the several outputs in $\mathcal{F}(p)$. Our profiles are anonymous in that they do not keep track of who votes for each input, but only of how many do so, with $p(j)$ representing the number of voters who vote for j. Thus, our approach here is to build anonymity into the context, rather than employ it as one of the characterizing axioms.

Notice that we are in the variable electorate context; we place no upper bound on the number $N(p)$ of voters, and the domain $(\mathbf{Z}^+)^{\mathcal{I}}$ of an abstract anonymous voting system is infinite, or "unbounded". Standard rules that fit this context may be interpretable as abstract anonymous voting rules in more than one way. Table 3 interprets a variety of examples in the way (or ways) that seems most natural. For example, in order for Definition 3.1.2 to work well with the definitions of consistency and connectedness, we are forced to represent a voting rule as a "social welfare/preference" function if it has weak orderings as outcomes (see Section 2.1). This comment applies to rows 1, 3, and 6 of Table 3.

The generalized scoring rules form an important subclass of the abstract anonymous voting systems:

Definition 3.1.3. Given sets $\mathcal{I}$ and $\mathcal{O}$, a ***scoring vector*** for these sets consists of a function $w : \mathcal{I} \times \mathcal{O} \to \mathbf{R}$. Such a scoring vector w constitutes a ***scoring representation*** for an abstract anonymous voting rule $\mathcal{V} = (\mathcal{I}, \mathcal{O}, \mathcal{F})$ if for every profile p and every $u \in \mathcal{O}$, $u \in \mathcal{F}(p)$ if and only if $\text{Score}_{w,p}(u) \geq \text{Score}_{w,p}(v)$ holds for every $v \in \mathcal{O}$, where the ***score*** of an output is given by $\text{Score}_{w,p}(u) = \sum_{i\in\mathcal{I}} p(i)w(i, u)$. Finally, $\mathcal{V}$ is a ***generalized scoring rule*** if it has such a representation. We use ***GSR*** to denote the class of all such rules.

Note that while neutrality is effectively built in to the more standard definition of a scoring rule (in the social choice or social welfare context), that is not the case here.

3.2. Consistency

In addition to its use of in characterizing scoring rules, this property was employed in [12] to show that among social preference functions the Kemeny rule is the unique neutral and consistent Condorcet extension.[20] The implications of consistency appear to be different in these results, because the contexts are not the same. Abstract anonymous voting systems imply no particular context (ballot form, or election output form), and so the question of whether or not a given voting system is consistent according to the definition that follows may depend on how we view that system as an abstract system — that is, on what we declare to be the input set $\mathcal{I}$ and output set $\mathcal{O}$.

Definition 3.2.1. An abstract anonymous voting system $\mathcal{V} = (\mathcal{I}, \mathcal{O}, \mathcal{F})$ is ***consistent*** if for every pair p_1 and p_2 of profiles satisfying $\mathcal{F}(p_1) \cap \mathcal{F}(p_2) \neq \emptyset$, $\mathcal{F}(p_1 + p_2) = \mathcal{F}(p_1) \cap \mathcal{F}(p_2)$.

Here $p_1 + p_2$ denotes the sum of p_1 and p_2 as functions: $(p_1 + p_2)(j) = p_1(j) + p_2(j)$ for each $j \in \mathcal{I}$; kp will denote $p + p + \cdots + p$ (taken k times). Intuitively, $p_1 + p_2$ represents the profile for the disjoint union of the separate electorates giving rise to p_1 and to p_2. The following weak forms, or "components", of consistency, will be useful:

Definition 3.2.2. An abstract anonymous voting system $\mathcal{V} = (\mathcal{I}, \mathcal{O}, \mathcal{F})$ is

- ***homogeneous*** if for every profile p and integer $k \geq 1$, $\mathcal{F}(kp) = \mathcal{F}(p)$,
- ***weakly consistent*** if for every pair p_1 and p_2 of profiles satisfying $\mathcal{F}(p_1) = \mathcal{F}(p_2)$, $\mathcal{F}(p_1 + p_2) = \mathcal{F}(p_1) = \mathcal{F}(p_2)$,

[20] As a corollary of this result, we obtain an alternate short proof that the voting rule "K" given by the geometric representation of Example 2.3.4 is the same as the Kemeny rule. It is clear from the general description that K is anonymous, and the symmetry inherent in the geometry guarantees that K is neutral. Theorem 5.4 tells us that, as K is rational mean neat, it is consistent as a preference function. It is easy to see that K is a Condorcet extension, as follows: divide the original $2 \times 2 \times \cdots \times 2$ cube into $1 \times 1 \times \cdots \times 1$ subcubes, each containing one vertex of the larger cube. If the pairwise majority relation M is a weak ordering σ, then the mean location q will lie in all subcubes whose transitive vertices correspond to strict rankings extending σ (and in no other subcubes), so these vertices will collectively be closest (among transitive vertices) to q. We conclude that K is a neutral, anonymous Condorcet extension. By the theorem in [12], K is the Kemeny rule.

- ***up consistent*** if for every pair p_1 and p_2 of profiles, $\mathcal{F}(p_1 + p_2) \supseteq \mathcal{F}(p_1) \cap \mathcal{F}(p_2)$, and
- ***down consistent*** if for every pair p_1 and p_2 of profiles satisfying $\mathcal{F}(p_1) \cap \mathcal{F}(p_2) \neq \emptyset$, $\mathcal{F}(p_1 + p_2) \subseteq \mathcal{F}(p_1) \cap \mathcal{F}(p_2)$.

Homogeneity is what allows the minimal sort of geometric representation to exist at all, because it tells us that the outcome of an election depends only on the *fraction* of voters choosing each ballot (the relative proportions among the integers $p(i)$ for $i \in \mathcal{I}$) rather than on the absolute vote totals $p(i)$. Smith [1], page 1029, argues that it is an extremely natural property to expect. Surprisingly, however, in Fishburn [25] we learn that the Dodgson and Young procedures can each fail to be homogeneous, depending on some of the details in the precise formulation of these systems. Many other "natural" voting systems (including all examples in Section 2) do satisfy homogeneity. Weak consistency appears in [10].

Proposition 3.2.3. *The Hare system is not weakly consistent*[21] *as a social choice function.*

Proof. Consider the following profile of strict preferences for the four alternatives A, B, C, and D:

Profile p_1	<u>8</u>	<u>3</u>	<u>2</u>	<u>4</u>
	A	B	C	D
	B	C	B	A
	C	D	D	C
	D	A	A	B

Here, the underlined numbers give the number of voters who cast ballots for the corresponding strict rankings.

It is straightforward to check that the Hare procedure, applied to p_1, eliminates C first, then D, then B, then A, and the social choice is $\{A\}$. Let profile p_2 be obtained from profile p_1 by transposing alternatives B and C. Clearly, for p_2 Hare eliminates B first, then D, then C, then A, and the social choice is again $\{A\}$. If we set $p_3 = p_1 + p_2$, weak consistency demands that $\mathcal{F}(p_3) = \{A\}$, yet we find that for p_3, B and C are eliminated simultaneously in the first stage, then A, then D, and $\mathcal{F}(p_3) = \{D\}$. ■

Together, "Up" and "Down" consistency imply full consistency, so we might think of them as the two halves of consistency. A related property called containment consistency appears in Merlin [23,24] (also see Chebotarev and Shamis [26]). Suppose we consider the special case in which $\mathcal{O}$ consists of individual alternatives — candidates, for an elected presidency, for example. Up consistency then tells us that if two separate electorates elect Jill (with one or both possibly producing a tie among several candidates, including Jill) then the combined electorate should also elect Jill. A failure of up consistency would seem fairly outrageous, especially to the voters who supported Jill.

As pointed out in [10], the intuitive content of down consistency is quite different, telling us (mostly) that ties are unstable or "knife-edge".[22] For example, if p_1 and p_2 are profiles satisfying $\mathcal{F}(p_1) = \{a, b\}$ and $\mathcal{F}(p_2) = \{a\}$, then down consistency implies that $\mathcal{F}(p_1 + p_2) = \{a\}$. This holds even when $N(p_1) = 1{,}000{,}000$ and $N(p_2) = 1$, so in this sense even one voter can break a tie in a very large election. Thus, down consistency loses much of its force when a voting system has no ties.

In fact, assume that some unspecified system has no ties whatsoever. While down consistency still has some effect under our assumption – it is equivalent to weak consistency for single-winner situations – it carries no additional force beyond that of up consistency (which also reduces to single-winner weak consistency). Every up consistent system that has no ties is down consistent. In this sense, the down consistency of the Almost Dictatorship and Pythagoras examples is vacuous. In contrast, the Copeland rule illustrates the setting in which the presence of ties makes down consistency "bite". With Copeland, ties are not at all knife-edge. For example, with three alternatives any sufficiently strong cycle (in which each pair-wise majority has a margin of 2 or more voters) results in a Copeland three-way tie. Any addition of a single voter fails to break the cycle, and so fails to break the three-way Copeland tie. The Omninominator rule satisfies $\mathcal{F}(p_1 + p_2) = \mathcal{F}(p_1) \cup \mathcal{F}(p_2)$ and so is a good example of a system that is up consistent and not down consistent. (That Copeland is not fully consistent is already implicit in [3,12].) We have just sketched the proof of the following:

Proposition 3.2.4. *The Copeland rule is not down consistent, either as a social choice function or as a social welfare/preference function. The Omninominator rule is up, but not down, consistent.*

It is not difficult to show that weak consistency and up consistency also fail for Copeland. Favardin et al. [28] compare Copeland to Borda, showing that Copeland is relatively resistant to manipulation. It seems possible that the failure of Copeland to be down consistent may be a partial explanation; this system may owe some of its resistance to manipulation to the large number of ties, which make it relatively resistant to any change at all. The Omninominator rule produces even more ties than Copeland, of course. Consequently, it is almost completely resistant to manipulation (as measured via the methods in Aleskerov and Kurbanov [29], and Smith [30,28]). With three alternatives and an arbitrary number of voters, there is in fact a unique (up to permutations of individuals and voters) manipulable profile.

21 Smith [1] discusses the general inconsistency of scoring run-offs.

22 The strong monotonicity used to characterize majority rule in May [27] can be seen as a special case.

3.3. Continuity and connectedness

Suppose some output $u \in \mathcal{O}$ loses an election. Intuitively, we might expect that any sufficiently small change to the profile would make too small a difference to erase the margin of loss. But even a one-voter change might not be sufficiently small, unless we first enlarge that margin by scaling up the profile.

Definition 3.3.1. Two anonymous profiles p and q satisfy $p \approx q$ if they differ by one vote: for some $i, j \in \mathcal{I}$, $q(i) = p(i)-1$, $q(j) = p(j)+1$, and $q(m) = p(m)$ for each $m \in \mathcal{I}$ with $m \neq i, j$. An abstract anonymous voting system $\mathcal{V} = (\mathcal{I}, \mathcal{O}, \mathcal{F})$ is ***continuous*** if, for each profile p and each $i \in \mathcal{I}$ with $i \notin \mathcal{F}(p)$, there exists an integer k such that each profile q with $q \approx kp$ satisfies $i \notin \mathcal{F}(q)$.

Connectedness guarantees the existence of the ties that allow down consistency to bite.

Definition 3.3.2. An abstract anonymous voting $\mathcal{V}$ is ***connected*** if, given any two anonymous profiles p_1 and p_2 with $\mathcal{F}(p_1) \cap \mathcal{F}(p_2) = \emptyset$, there exist non-negative integers c and d such that $\mathcal{F}(p_1) \cap \mathcal{F}(cp_1 + dp_2) \neq \emptyset$ and $\mathcal{F}(p_1) \neq \mathcal{F}(cp_1 + dp_2)$.

For example, if we had an abstract anonymous voting system in which $\mathcal{O} = \{a, b\}$, with $\mathcal{F}(p_1) = \{a\}$ and $\mathcal{F}(p_2) = \{b\}$, then the profile $cp_1 + dp_2$ provided by the definition would be forced to satisfy $\mathcal{F}(cp_1 + dp_2) = \{a, b\}$. We may think of the profile $cp_1 + dp_2$ as intermediate to p_1 and p_2 — a sort of weighted average of the two. In any mean representation, profile $cp_1 + dp_2$ will be mapped into the line segment joining the images of p_1 and p_2. The intuitive content of connectedness, then, is that as that as voters "gradually change their minds", sliding from profile p_1 to profile p_2, they will pass through one or more "connecting" profiles representing ties. In one sense, this cannot always be literally true, for with an odd number of voters there will never be a tie in majority rule. But there is a precise property – *commensurability* – that expresses this idea, and that holds of voting systems that are both consistent and connected (see Smaoui and Zwicker [31]).

Connectedness does not seem to be very interesting in the absence of consistency. In the presence of consistency, connectedness is strictly stronger than continuity,[23] and there are several equivalent ways to define either property. Many naturally occurring voting systems are connected, including all those with a "Yes" in the rightmost column of Table 3. Exceptions include grading systems that, while roughly similar to Example 2.3.6, are a bit more practical in that they rule out ties. In the normal order of business, an instructor is not allowed to assign a pair of tied grades for a course, so most instructors either adopt a rule that automatically assigns the higher grade for averages falling on a borderline, or break ties using a subjective judgment based on other aspects of a student's performance. In the former case, the system is not connected; in the latter case one might argue that the grading system itself *does* yield ties, and is connected.

Proposition 3.3.3. *The Almost-Dictatorship and Pythagoras systems 2.4.4 and 2.3.7, are both consistent; neither is connected, but Pythagoras is continuous.*

The proof is an easy exercise. While these two examples may seem, at first, to be highly artificial, we show in the next proposition that the Hare social welfare/preference system fails to be connected for essentially the same reason as the Almost Dictatorship system:

- Among the possible profiles, there is a range corresponding to the rational points in the interval $[0, 1]$,
- all profiles corresponding to points in $(0, 1]$ have a common election outcome X, and
- the profile corresponding to 0 has an outcome disjoint from X.

Furthermore, it turns out that the *fundamentally irrational* traditional scoring rules[24] fail to be connected for the same reason as the Pythagoras system [31].

Proposition 3.3.4. *The Hare social welfare/preference function is not connected.*

Proof. Recall profiles p_1, p_2, and $p_3 = p_1 + p_2$ from Proposition 3.2.3. From the order of elimination of the alternatives, we see that, as a social welfare/preference function Hare satisfies $\mathcal{F}(p_1) = \left\{\begin{smallmatrix} A \\ B \\ D \\ C \end{smallmatrix}\right\}$ and $\mathcal{F}(p_3) = \left\{\begin{smallmatrix} D \\ A \\ B \\ C \end{smallmatrix}, \begin{smallmatrix} D \\ A \\ C \\ B \end{smallmatrix}\right\}$. As $\mathcal{F}(p_3) \cap \mathcal{F}(p_1) = \emptyset$, connectedness demands integers $j, k > 0$ with $\mathcal{F}(jp_3 + kp_1) \cap \mathcal{F}(p_3) \neq \emptyset$ (and $\mathcal{F}(jp_3 + kp_1) \neq \mathcal{F}(p_3)$). But in this case, $jp_3 + kp_1 = (j + k)p_1 + jp_2$, and it is easy to see that any profile $sp_1 + tp_2$ with $s > t$ eliminates alternatives in the same order as p_1 (for the same reasons). Thus $\mathcal{F}((j + k)p_1 + kp_2) = \left\{\begin{smallmatrix} A \\ B \\ D \\ C \end{smallmatrix}\right\}$ for every two integers $j, k > 0$, and connectedness fails. ∎

---

[23] This is an immediate consequence of 5.4 and the known direction of 5.5.

[24] A scoring rule is *fundamentally irrational* if no positive affine transform of its scoring weights yields equivalent weights that are (all) rational numbers.

4. Mean proximity representations of voting systems and the equivalence theorem

4.1. Closed decompositions and mean representations

As suggested by the representations in Section 2, after we plot votes as points in space and find their mean location, we use a subdivision of space to determine the outcome:

Definition 4.1.1. An indexed ***decomposition*** of a subset $P \subseteq \mathbf{R}^n$ is a sequence $S = \{r_a\}_{a\in\Delta}$ of subsets of P called ***regions***, with finite index set Δ, satisfying $\bigcup_{a\in\Delta} r_a = P$.

A ***decomposition*** is ***closed*** if its regions are closed sets.

Our interest is in voting systems that can be represented geometrically via the mean:

Definition 4.1.2. A ***mean representation*** of an abstract anonymous voting system $\mathcal{V} = (\mathcal{I}, \mathcal{O}, \mathcal{F})$ consists of a quadruple *(n, R, P, S)*, in which

1. $R : \mathcal{I} \to \mathbf{R}^n$ is an ***input plot function*** locating each possible ballot as a point in Euclidean space,
2. the ***representing polytope*** P is the convex hull of the (finite) set $\{R(j) : j \in \mathcal{I}\}$,
3. $S = \{r_a\}_{a\in\mathcal{O}}$ is a closed decomposition of P indexed by $\mathcal{O}$, and
4. S ***represents*** $\mathcal{V}$: for each $p \in (\mathbf{Z}^+)^{\mathcal{I}}$, $\mathcal{F}(p) = \{a \in \mathcal{O} : \text{Mean}(p) \in r_a\}$, where ***Mean(p)*** denotes the mean location $\frac{\sum_{j\in\mathcal{I}} p(j)R(j)}{N(p)}$ of all plotted votes cast.

The map Mean : $(\mathbf{Z}^+)^{\mathcal{I}} \to P$ can be thought of as an extension, from "one-voter" profiles to arbitrary profiles, of map R. By itself the existence of a mean representation for $\mathcal{V}$ tells us only that $\mathcal{V}$ must be homogeneous. However, further restrictions on the *type* of closed decomposition S employed in the representation are directly tied to the voting-theoretic properties of $\mathcal{V}$.

Definition 4.1.3. A ***mean proximity representation*** of an abstract anonymous voting system $\mathcal{V} = (\mathcal{I}, \mathcal{O}, \mathcal{F})$ is a quintuple *(n, R, P, S, E)* consisting of a mean representation *(n, R, P, S)* of $\mathcal{V}$ together with an output plot function $E : \mathcal{O} \to \mathbf{R}^n$ satisfying that $S = \{r_a\}_{a\in\mathcal{O}}$ is the ***Voronoi decomposition of P induced by E***: for each $a \in \mathcal{O}$, r_a consists of those points x in P satisfying $\|x - E(a)\| \leq \|x - E(b)\|$ for every $b \in \mathcal{O}$ (where $\|x - y\|$ denotes the Euclidean distance between points x and y). Thus $o \in \mathcal{F}(p)$ if and only if o minimizes the distance $\|\text{mean}(p) - E(o)\|$. An abstract anonymous voting system $\mathcal{V}$ is a ***mean proximity voting system*** if it has such a representation.

4.2. The equivalence theorem

Theorem 4.2.1 (*The Equivalence Theorem*). *An abstract anonymous voting system $\mathcal{V} = (\mathcal{I}, \mathcal{O}, \mathcal{F})$ is a generalized scoring rule if and only if it is a mean proximity rule.*

For the proof we follow in the footsteps of Barthélemy and Monjardet [9] by exploiting the following result and its corollary:

Theorem 4.2.2 (*Christiaan Huygens, 1629–1695*). *Let $S = \{s_i | i \in \{1, 2, \ldots, n\}\}$ be a finite sequence*[25] *of points in $\mathbf{R}^k$, and y be any point in $\mathbf{R}^k$. Let* $\text{mean}(S)$ *denote* $\frac{1}{n}\sum_{i=1}^n s_i$. *Then* $\sum_{i=1}^n \left(\|s_i - y\|^2\right) = \sum_{i=1}^n \left(\|s_i - \text{mean}(S)\|^2\right) + n\,\|\text{mean}(S) - y\|^2$.

The following corollary is immediate:

Corollary 4.2.3. *If S and* $\text{mean}(S)$ *are as in 4.2.2, and Y is any finite set of points of $\mathbf{R}^k$, then the elements of Y that minimize the sum of squared distances to S's members (counting multiplicity) coincide with those closest to* $\text{mean}(S)$.

Proof of 4.2.1. [$\Leftarrow$] Assume *(n, R, P, S, E)* is a mean proximity representation of $\mathcal{V}$, and let $w : \mathcal{I} \times \mathcal{O} \to \mathbf{R}$ be given by $w(i, o) = -\|R(i) - E(o)\|^2$, the negative of the squared Euclidean distance between the spatial locations of i and o. We show that w provides the desired scoring representation.

Let p be any profile for $\mathcal{V}$, and o be any element of $\mathcal{O}$.

$$\begin{aligned}
\text{Then } o \in \mathcal{F}(p) &\Leftrightarrow \quad o \text{ minimizes the distance } \|\text{mean}(p) - E(o)\| \\
&\Leftrightarrow \quad o \text{ minimizes the squared distance } \|\text{mean}(p) - E(o)\|^2 \\
&\Leftrightarrow * \; o \text{ minimizes the sum } \sum_{i\in\mathcal{I}} p(i)\,\|R(i) - E(o)\|^2 \\
&\Leftrightarrow \quad o \text{ maximizes the sum } \sum_{i\in\mathcal{I}} p(i)\left(-\|R(i) - E(o)\|^2\right) \\
&\Leftrightarrow \quad o \text{ maximizes Score}_{w,p}(o).
\end{aligned}$$

Note that the third equivalence (asterisked) makes use of Corollary 4.2.3.

25 A sequence allows the same point s to occur multiple times, whereas a set would not "keep track" of the multiplicity of such a point.

[$\Rightarrow$] Assume $w : \mathcal{I} \times \mathcal{O} \to \mathbf{R}$ is a scoring representation for $\mathcal{V}$, and let $\{o_1, o_2, \ldots, o_n\}$ be any fixed enumeration of $\mathcal{O}$. Let $R : \mathcal{I} \to \mathbf{R}^n$ be given by

$$R(i) = (w(i, o_1), w(i, o_2), \ldots, w(i, o_n)).$$

For each k between 1 and n, let e_k denote the point $(0, 0, \ldots, 0, 1, 0, \ldots, 0)$ (where the "1" occurs in the kth coordinate), and let $E : \mathcal{O} \to \mathbf{R}^n$ be given by $E(o_k) = e_k$ for each k. Let P be the convex hull of the image of $\mathcal{I}$ under R, and $S = \{r_a\}_{a \in \mathcal{O}}$ be the Voronoi decomposition of P induced by E. We claim that *(n, R, P, S, E)* is a mean proximity representation of $\mathcal{V}$. To see this, let p be any profile for $\mathcal{V}$, let $Score_{w,p}$ denote $\sum_{i \in \mathcal{I}} p(i)R(i) = (Score_{w,p}(o_1), Score_{w,p}(o_2), \ldots, Score_{w,p}(o_n))$, and let o_k be any member of $\mathcal{O}$. Noting that mean$(p) = (Score_{w,p})/N(p)$, we now obtain the following chain of equivalences:

$$\begin{aligned} o_k \in \mathcal{F}(p) \Leftrightarrow\ & o_k \text{ maximizes } Score_{w,p}(o_k) \\ \Leftrightarrow\ & \text{the } k\text{th component of } Score_{w,p} \text{ is a maximal component of } Score_{w,p} \\ \Leftrightarrow\ & \text{the } k\text{th component of Mean}(p) \text{ is a maximal component of Mean}(p) \\ \Leftrightarrow **\ & k \text{ minimizes the distance from Mean}(p) \text{ to } e_k \\ \Leftrightarrow\ & \text{Mean}(p) \text{ lies in the Voronoi region } r \text{ for } o_k. \end{aligned}$$

and this establishes the desired result. Note that the penultimate (doubly asterisked) equivalence makes use of Lemma 2.2.2. ■

4.3. *A return to* Example 1.1

Our proof of the equivalence theorem is constructive; a mean proximity representation for a rule yields scoring weights in the sense of Definition 3.1.3. In particular, when we do this for the permutahedral representation of the Borda count as a social welfare/preference function, we obtain scoring weights equal to the negative squares of Euclidean distances between the permutahedron vertices. If we then apply a suitable positive affine transform to these weights, we obtain the weights for System A of Example 1.1. In other words, System A *is* the Borda count (albeit in a somewhat novel context) and the System A weights of Tables 1a and 1b are obtained from squared Euclidean distances between vertices of a regular hexagon (1a) or of a truncated octahedron (1b).

Some readers will recognize the System B weights as yielding the Kemeny rule, because they can be obtained as the number of pairs of alternatives that are listed in the same relative order by two strict rankings. Alternatively, the same weights arise via the equivalence theorem from the negative squared distances between the vertices of the hypercube (Example 2.3.4), again after applying a suitable positive affine transform. Perhaps the most surprising thing about Example 1.1 is that we can compare these two well-known rules so directly; the social welfare/preference framework bridges what otherwise might seem to be a significant gap in context. Note also that we obtain different *orderings* of strict rankings — the metrics induce two different ways to list strict rankings in terms of increasing "distance" from a reference strict ranking. These weak orderings of strict rankings are induced by the Euclidean metric applied to two very natural ways to embed strict rankings as points in space. The Kemeny (hypercube) ordering will seem familiar to some readers, but the Borda (permutahedron) ordering offers a compelling alternative notion of what it means for one of two strict rankings to be "more similar" to a third.

From Tables 1a and 1b we observe that these two weak orderings of rankings are identical in the case of three alternatives, while with four alternatives the two weak orderings are compatible (in a sense analogous to Definition 2.2.1), although the Borda order teases apart some ties in the Kemeny version. It turns out that with eight alternatives there are some strict reversals in these ordering of rankings (though we have no evidence that eight is minimal). One important lesson for social choice is that there may not exist any canonical ordering of rankings; for example, if we are comparing voting rules based on how they perform on profiles that have a relatively high level of agreement among the voters then our choice of metric to measure "agreement" may bias the results, possibly in favor of rules that implicitly rely on a similar metric.

5. Mean neat representations

We begin with a key definition from [8].

Definition 5.1. Two sets u and v of points of $\mathbf{R}^n$ are ***weakly separated by a hyperplane*** h of $\mathbf{R}^n$ if every point of u lies either on h or to one side of h, and every point of v lies either on h or to the other side of h. These sets are ***neatly separated*** by h if they are weakly separated and satisfy $u \cap v = h \cap u = h \cap v$. A closed decomposition S of P is ***neat*** if each pair of distinct regions of S is neatly separated by some hyperplane.

In Fig. 1 (in Section 1), regions u and w are weakly separated by hyperplane g (and are properly separated as well — see [8]) but cannot be neatly separated.

Definition 5.2. A ***mean neat representation*** of an abstract anonymous voting system $\mathcal{V} = (\mathcal{I}, \mathcal{O}, \mathcal{F})$ consists of a mean representation *(n, R, P, S)* for which S is a neat decomposition. An abstract anonymous voting system $\mathcal{V}$ is a ***mean neat voting system*** if it has such a representation.

For distinct regions $u \neq v$, neat separability is intermediate in strength to weak and strict separability [8]. Every Voronoi decomposition is neat: we can neatly separate a region u, of points most proximate to point f_u, from a region v, with proximity point $f_v \neq f_u$ via the hyperplane that perpendicularly bisects the line segment $[f_u, f_v]$. It is easy to construct a decomposition S that is neat yet not Voronoi. However, constructing a voting rule that is mean neat without being mean proximity requires something a bit stronger — a decomposition S that is neat yet fails to be Voronoi even if it is embedded in a Euclidean space of higher dimension (that is, S is not ***quasi-Voronoi*** — see [8] for more discussion). The twisted triangle example in [5] provides such an S; one can construct from it a rule that is mean neat but not mean proximity, hence not a generalized scoring rule.

Essentially, consistency of a voting system seems to correspond to neat separability of the regions in the representation, while connectedness corresponds to the requirement that each of the neatly separating hyperplanes h can be taken to be *rational* as an affine subspace, meaning that the set of rational points is dense in h.[26]

Definition 5.3. A ***rational mean neat representation*** of an abstract anonymous voting system $\mathcal{V} = (\mathcal{I}, \mathcal{O}, \mathcal{F})$ consists of a mean neat representation (n, R, P, S) for which any two regions of S may be neatly separated by a rational hyperplane, and $\mathcal{V}$ is a ***rational mean neat voting system*** if it has such a representation.

Theorem 5.4 (*The Rational Mean Neat Characterization Theorem*). *Let $\mathcal{V} = (\mathcal{I}, \mathcal{O}, \mathcal{F})$ be an abstract anonymous voting system. Then $\mathcal{V}$ is consistent and connected if and only if it is a rational mean neat system.*

The proof appears in [4]. What happens when we drop the rationality assumptions of 5.4? It is easy to see that if $\mathcal{V}$ has a mean neat representation then $\mathcal{V}$ must be both consistent and continuous (but not necessarily connected). The proof of 5.4 suggests that the converse may hold as well:

Conjecture 5.5. *Let $\mathcal{V} = (\mathcal{I}, \mathcal{O}, \mathcal{F})$ be an abstract anonymous voting system. Then $\mathcal{V}$ is consistent and continuous if and only if it is a mean neat system.*

Acknowledgements

We gratefully acknowledge support from Grant SAB 2004-0264 of the Spanish Ministry of Education and Science. Some of this work was done while the author was visiting the Departamento Fundamentos del Análisis Económico at the University of Alicante and the Economics Department at the University of Caen. A draft of this paper was circulated privately under the title "When is a vote an average?". We thank Juan Enrique Martinez-Legaz, Vincent Merlin, and Hannu Nurmi for their assistance. Comments by the editor of this special issue and by two referees were very helpful in improving the manuscript.

References

[1] J.H. Smith, Aggregation of preferences with variable electorate, Econometrica 41 (1973) 1027–1041.
[2] H.P. Young, A note on preference aggregation, Econometrica 42 (1974) 1129–1131.
[3] H.P. Young, Social choice scoring functions, SIAM J. Appl. Math. 28 (1975) 824–838.
[4] W.S. Zwicker, A characterization of the rational mean neat voting rules, Math. Comput. Modelling 48 (9–10) (2008) 1374–1384.
[5] R.B. Myerson, Axiomatic derivation of scoring rules without the ordering assumption, Soc. Choice Welf. 12 (1995) 59–74.
[6] V. Conitzer, M. Rognlie, Voting rules that score rankings, 2007 (preprint).
[7] T. Meskanen, H. Nurmi, Distance from consensus: A theme and variations, in: B. Simeone, F. Pukelsheim (Eds.), Mathematics and Democracy: Recent Advances in Voting Systems and Collective Choice, in: Studies in Choice and Welfare, Springer Verlag, Berlin, Heidelberg, 2006, pp. 117–132.
[8] D. Cervone, W.S. Zwicker, Convex decompositions, J. Convex Anal. (2008) (in press).
[9] J.P. Barthélemy, B. Monjardet, The median procedure in cluster analysis and social choice theory, Math. Social Sci. 1 (1981) 235–267.
[10] D.G. Saari, Geometry of Voting, Springer-Verlag, 1994.
[11] D.G. Saari, V.R. Merlin, A geometric examination of Kemeny's rule, Soc. Choice Welf. 17 (2000) 403–438.
[12] H.P. Young, A. Levenglick, A consistent extension of Condorcet's election principle, SIAM J. Appl. Math. 35 (1978) 285–300.
[13] S.J. Brams, P.C. Fishburn, Approval Voting, Birkhäuser, Boston, 1983.
[14] J. Kemeny, Mathematics without numbers, Daedalus 88 (1959) 571–591.
[15] M.J.A.N.C. Condorcet (The Marquis de), Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix, Paris, 1785.
[16] M.J.A.N.C. Condorcet (The Marquis de), in: I. McLean, F. Hewitt (Eds.), Foundations of Social Choice and Political Theory, Elgar Press, Aldershot, UK, 1994.
[17] W.S. Zwicker, Note on a result of Saari and Merlin, 2005 (preprint).
[18] J. Freixas, W.S. Zwicker, Weighted voting, abstention, and multiple levels of approval, Soc. Choice Welf. 21 (2003) 399–431.
[19] J. Freixas, W.S. Zwicker, Anonymous voting games with abstention and multiple levels of approval: Classification and counting, Games Econ. Behav. (2008) (in press).
[20] W.S. Zwicker, Anonymous voting rules with abstention: Weighted voting, in: S. J. Brams, W. V. Gehrlein, F.S. Roberts (Eds.), The Mathematics of Preference, Choice, and Order: Essays in Honor of Peter C. Fishburn, Springer, Heidelberg, 2008 (in press).
[21] T. Hare, The Election of Representatives, Parliamentary and Municipal, Longmans, Green, Reader & Dyer, 1859.
[22] A.D. Taylor, Social Choice and the Mathematics of Manipulation, Cambridge U. Press and Math. Assoc. of America, 2005.
[23] V.R. Merlin, L'agrégation des préférences individuelles: les règles positionnelles itératives et la méthode de Copeland, Ph.D. Diss., U. de Caen, 1996.
[24] V.R. Merlin, Scoring runoff methods are consistent, unpublished mimeo, 1996.

[26] Several equivalent conditions are given in [4]; one of these is that h is given by a single equation $\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = \beta$ in which both β and the components of the normal vector α to h are integers, and another is that h is the affine span of rational points. For background on affine subspaces and related material, see Rockafellar [32] or Webster [33].

[25] P.C. Fishburn, Condorcet social choice functions, SIAM J. Appl. Math. 33 (1977) 469–489.
[26] P.Y. Chebotarev, E. Shamis, Characterization of scoring methods for preference aggregation, Ann. Oper. Res. 80 (1998) 299–332.
[27] K. May, A set of independent, necessary and sufficient conditions for simple majority decision, Econometrica 20 (1952) 680–684.
[28] P. Favardin, D. Lepelley, J. Serais, Borda rule, Copeland method, and stragic manipulation, Rev. Econ. Des. 7 (2002) 213–228.
[29] F. Aleskerov, E. Kurbanov, Degree of manipulability of social choice procedures, in: Alkan, et al. (Eds.), Current Trends in Economics, Springer, 1999.
[30] D. Smith, Manipulability measures of common social choice functions, Soc. Choice Welf. 16 (1999) 639–661.
[31] H. Smaoui, W.S. Zwicker, Consistency, connectedness, and a characterization of rational scoring rules, working paper, 2005.
[32] R.T. Rockafellar, Convex Analysis, Princeton U. Press, Princeton, NJ, 1970.
[33] R. Webster, Convexity, Oxford U. Press, Oxford, UK, 1994.