Universal and Symmetric Scoring Rules for Binary Relations

William S. Zwicker

Abstract

Are Plurality voting, the Kemeny rule, Approval voting, and the Borda Mean Dichotomy rule all versions of the same voting rule? Yes, in a sense. In an extension of work by Barthélemy and Monjardet [3], we consider functions F that assign real number scoring weights $F(R_1, R_2)$ to pairs of binary relations on a finite set A of alternatives, serving as symmetric measures of similarity between R_1 and R_2. Any such F induces a *symmetric binary relational scoring rule* $\mathcal{F}$ – an abstract aggregation rule with arbitrary binary relations as ballots R_1 and as aggregated outcomes R_2. The level of generality is surprisingly effective. By restricting the classes of relations allowed as ballots and elections outcomes, $\mathcal{F}$ yields more familiar and concrete scoring rules. The symmetric assignment F^H, for example, arises from an inner product in a simple and natural way, and restrictions of the induced scoring rule $\mathcal{F}^H$ yield *all* the aforementioned familiar voting rules. Moreover, the inner product formulation yields a Euclidean form of *distance rationalization* for $\mathcal{F}^H$, resulting in a universal distance rationalization for all concrete scoring rules obtained as restrictions.

1 Introduction

In a *scoring rule*, as defined traditionally, each voter specifies as her ballot a ranking (a linear order, or sometimes a weak order) of all alternatives in some finite set A. Each such ballot contributes points to individual alternatives, according to some fixed table of contributions wherein the contribution of a ballot to an alternative x depends on how highly x is ranked by that ballot. The outcome of the election is the alternative(s) x awarded the greatest point total (the sum of individual contributions made to x by the voters). Alternatively, the outcome might be the *social ranking* of alternatives in which x is ranked over y when x has a greater point total than does y. Such a social ranking is a binary relation on A.

In a *relational scoring rule*, "ballots" come from some class $\mathcal{C}_1$ of binary relations on A: dichotomous weak orders, or partial orders, or equivalence relations, or ... Each ballot R_i awards $F(R_i, R)$ points directly to each binary relation R from some possibly different class $\mathcal{C}_2$. The outcome is the relation $R \in \mathcal{C}_2$ with most points. The idea has been studied in the case of linear order relations in [20] and [5], where *ranking scoring rules* are shown to be strictly more general than traditional ones; [6] takes a related approach towards judgment aggregation.

Our most important example F^H of a universal scoring assignment represents the same aggregation rule as the *median procedure* of Barthélemy and Monjardet [3]. While this manuscript has significant overlap with [3], it also contains substantial novel content. We discuss these similarities and differences in Appendix II.

If we imagine an *axis of abstraction* or *generality* for scoring rules, then traditional scoring rules (with ballots limited to the class of linear orders) sit at the least abstract end – at the left, in Fig. 1 – with ranking scoring rules slightly to their right. The *generalized scoring rules*, considered implicitly by Meyerson in [15] and defined explicitly in [20], allow ballots from a completely abstract set $\mathcal{I}$ and election outcomes from a second abstract set

$\mathcal{O}$. They sit at the opposite end of our axis.[1] The *pairwise scoring rules* of Xia and Conitzer [18] allow partial orders as ballots, with election outcomes that are individual alternatives, or linear orders, or sets of alternatives having a specified size k.

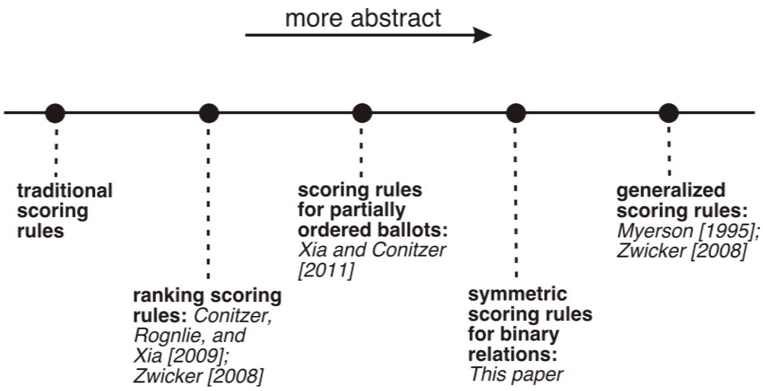

Figure 1: An axis of abstraction for scoring rules.

The *symmetric binary relational scoring rules* (aka *symmetric BRSR*s) introduced here are situated, on our axis of abstraction, between generalized scoring rules and the scoring rules for partially ordered ballots in [18], to which they are closely related. The meaning of "pairwise" rule is essentially the same in both contexts. But there are significant differences. A principal focus of [18] is the *Maximum Likelihood Estimator* interpretation that applies to certain aggregation rules, and we do not consider *MLE* interpretation here. The partial orders that serve as ballots for the pairwise rules in [18] allow two distinct alternatives to be incomparable, but do not allow them to be equivalent (as we do here).

A more important distinction is that complete symmetry between inputs (ballots) and outputs (potential election outcome) becomes apparent only at the BRSR level of generality. The value $F(R_1, R_2)$ can be viewed as a symmetric measure of similarity between the two relations, so that the points awarded to the potential election outcome R_2 by the ballot R_1 can equally well be considered an award by the ballot R_2 to the potential election outcome R_1. That symmetry is usually broken when we induce a concrete scoring rule by restricting ballots to members of one subclass $\mathcal{C}_1$ of binary relations on A and potential outcomes to members of a possibly different subclass $\mathcal{C}_2$. But the underlying symmetry of the unrestricted scoring assignment F tells us that this restriction is, in a sense, the *same* concrete rule as that obtained by switching the roles of $\mathcal{C}_1$ and $\mathcal{C}_2$. We will see that Approval voting and the Borda Mean Dichotomy rule are, in this *switching* sense, the same (see also [7], [8]).

We've all seen examples of *false*, or *empty*, generalization, wherein the enlarged class of objects adds none that are inherently interesting, the idea itself yields no new understanding, and the proofs and constructions similarly fail to be novel. Its antipode, a *consequential* generalization, yields insights that would not be apparent at either higher or lower levels of abstraction. Our view is that symmetric binary relational scoring rules are consequential in this sense – a view we hope will be shared by the reader, when they have finished reading.[2]

The remaining sections are organized as follow: after the technical preliminaries of §2, we introduce the universal and symmetric binary relational scoring rule $\mathcal{F}^H$ in §3. Properties of $\mathcal{F}^H$ and its restrictions are the principal focus of this paper. Although we initially define $\mathcal{F}^H$ via an inner product, we show in §4 that it may also be described as a sum via a table

[1] In [19], Xia studies "generalized scoring rules" of a very different kind. While Fig 1. depicts *an* axis of generalization for scoring rules, it is not the only such axis possible.

[2] Had someone suggested to me, six months ago, that scoring rules for arbitrary binary relations was a worthwhile topic for exploration, my private reaction would probably have been "False generalization!"

of *pairwise* contributions. Section 5 discusses $\mathcal{F}^S$, which is related to but distinct from $\mathcal{F}^H$. A Euclidean version of *distance rationalization* for $\mathcal{F}^H$ (and for all its restrictions) is demonstrated in §6, with some brief concluding remarks (that mention a BRSR for the Borda count) in §7. Appendix I contains two proofs, and Appendix II discusses the relationship of this paper to [3].

2 Technical preliminaries

We'll use A to denote a finite set of *alternatives* − alternatives might be candidates in a multicandidate election (for example) or objects that must be grouped into several clusters based on similarity (a very different example). A *binary relation* R on A is a subset of the cartesian product $A \times A$ – equivalently, R is an element of the power set $2^{A\times A}$. A *class of binary relations* for A is a set $\mathcal{C} \subseteq 2^{A\times A}$ of binary relations on A. Such classes are typically specified in terms of the properties required for membership. The following definitions for these properties are not completely standard, in that they mention only those ordered pairs (x, y) that are *distinct*, satisfying $x \neq y$.[3]

Definition 1 *A binary relation R on a set A is*

- symmetric *if $(x, y) \in R \Rightarrow (y, x) \in R$ for each $x,\ y \in A$ with $x \neq y$,*
- antisymmetric *if $(x, y) \in R \Rightarrow (y, x) \notin R$ for each $x,\ y \in A$ with $x \neq y$,*
- complete *if $(x, y) \in R$ or $(y, x) \in R$ for each $x,\ y \in A$ with $x \neq y$,*
- transitive *if $(x, y),\ (y, z) \in R \Rightarrow (x, z) \in R$ for each $x,\ y,\ z \in A$ with $x \neq y,\ x \neq z$, and $y \neq z$*

Definition 2 *Some classes of binary relations:*

- *$\mathcal{W}$, the* weak orders*: binary relations on A that are complete and transitive*
- *$\mathcal{L}$, the* linear orders*: binary relations on A that are antisymmetric, complete and transitive*
- *$\mathcal{P}$, the* partial orders*: binary relations on A that are antisymmetric and transitive*
- *$\mathcal{E}$, the* equivalence relations*: binary relations on A that are symmetric and transitive.*

Definition 3 *Given a binary relation R on A, the three* primitive derived relations *are:*

- $\overline{R} = \{(x, y) \in A \times A \mid (x, y) \notin R\}$,
- $R^{\text{reverse}} = \{(x, y) \in A \times A \mid (y, x) \in R\}$
- $\overline{R^{\text{reverse}}} = \{(x, y) \in A \times A \mid (y, x) \notin R\}$, *also known as the* dual *relation* R^d

and the four classifying derived relations *are:*

1. $R \cap \overline{R^{\text{reverse}}} = \{(x, y) \in A \times A \mid (x, y) \in R$ *and* $(y, x) \notin R\}$

[3] We have chosen, as a simplifying assumption, to ignore differences between two relations based on which pairs of form (x, x) are members. However, most natural classes of relations contain exclusively *reflexive* relations – satisfying $(x, x) \in R$ for all $x \in A$ – or exclusively *antireflexive* relations – satisfying $(x, x) \notin R$ for all $x \in A$, so those pairs (R_1, R_2) to which we award points will typically agree on all (x, x) pairs, or all disagree on all (x, x) pairs. Thus, if we did award points based on (x, x) differences, those points might shift all totals up or down by a common amount, but would usually contribute nothing to the *differences* among the totals.

2. $\overline{R} \cap R^{\text{reverse}} = \{(x, y) \in A \times A \mid (x, y) \notin R$ *and* $(y, x) \in R\}$

3. $R \cap R^{\text{reverse}} = \{(x, y) \in A \times A \mid (x, y) \in R$ *and* $(y, x) \in R\}$*, also known as the* indifference *relation* $\sim_R$ *(or just* $\sim$*)*

4. $\overline{R} \cap \overline{R^{\text{reverse}}} = \{(x, y) \in A \times A \mid (x, y) \notin R$ *and* $(y, x) \notin R\}$*, also known as the* incomparability *relation* $\downarrow_R$ *(or just* $\downarrow$*)*

Definition 4 *Let* R *be a binary relation on* A *and* $(x, y) \in A \times A$ *be a distinct pair of alternatives. Then* (x, y) *is a member of exactly one of the four classifying relations, numbered 1-4 in definition 3). We'll use* pair-case$_R(x, y)$ *to denote the classifying relation to which* (x, y) *belongs.*

The following proposition contains observations that, while self-evident, are useful for what follows.

Proposition 1 *Let* R *be any transitive binary relation on* A*. Then:*

1. *R is a weak order* $\Leftrightarrow$ *no distinct pairs* $(x, y) \in R$ *fall into pair case 4*
2. *R is a partial order* $\Leftrightarrow$ *no distinct pairs* $(x, y) \in R$ *fall into pair case 3*
3. *R is a linear order* $\Leftrightarrow$ *no distinct pairs* $(x, y) \in R$ *fall into pair cases 3 or 4*
4. *R is an equivalence relation* $\Leftrightarrow$ *no distinct pairs* $(x, y) \in R$ *fall into pair cases 1 or 2*
5. $\sim_R$ *is an equivalence relation, with equivalence classes we call* I-classes

Definition 5 *Two additional classes of binary relations:*

- $\mathcal{D}$*, the* dichotomous orders*: weak orders on* A *having exactly two I-classes, known as* T *– or "top" – and* B *– or "bottom." When* $\{T, B\}$ *partitions* A *into non-empty pieces, we will write* $T > B$ *to denote the dichotomous order relation* R *wherein* $(x, y) \in R \Leftrightarrow x \in T$ *or* $y \in B$*.*
- $\mathcal{U}$*, the* plurality orders*: dichotomous orders* $\{t\} > N \setminus \{t\}$ *on* A *whose top contains a unique element* $t \in A$*.*

An *Approval voting* ballot (in which $X \subset A$ is the set of approved alternatives) can be identified with a dichotomous order (in which $T = X$ and $B = A \setminus X$). A *Plurality voting* ballot (cast in favor of the alternative t) can be identified with a plurality order (in which $T = \{t\}$). We are now ready to introduce the principal objects of study:

Definition 6 *Let* A *be a finite set of alternatives. A* binary relation scoring assignment*, or* BRSA*, is a function* F *that assigns a real number scoring weight* $F(R_1, R_2)$ *to each pair* (R_1, R_2) *of binary relations in its domain* $\text{Dom}(F) \subseteq 2^{A \times A} \times 2^{A \times A}$*. If* $\text{Dom}(F) = 2^{A \times A} \times 2^{A \times A}$ *(*F *assigns scoring weights to* every *pair of binary relations), then* F *is* universal*. If* $\text{Dom}(F) = \mathcal{C}_1 \times \mathcal{C}_2$*, where* $\mathcal{C}_1$ *and* $\mathcal{C}_2$ *are classes of binary relations, then* F *is Cartesian. If* $\text{Dom}(F) = \mathcal{C} \times \mathcal{C}$ *and* $F(R_1, R_2) = F(R_2, R_1)$ *for each* $(R_1, R_2) \in \mathcal{C} \times \mathcal{C}$*, then* F *is* symmetric.

Any Cartesian BRSA F *with domain* $\mathcal{C}_1 \times \mathcal{C}_2$ *induces a corresponding* binary relation scoring rule*, or* BRSR $\mathcal{F}$*, as follows. Given a finite set* N *of* voters *and any profile* $P = \{B_i\}_{i \in N}$ *of ballots* $B_i \in \mathcal{C}_1$*, each voter* $i \in N$ *awards* $f(B_i, O)$ *points to each* $O \in \mathcal{C}_2$*. The election outcome* $\mathcal{F}(P)$ *is the set* $\mathcal{O} \subset \mathcal{C}_2$ *of binary relations* $O \in \mathcal{C}_2$ *that amass the greatest* *point total,*[4] *as summed over all voters.* $\mathcal{F}$ *is* universal *(respectively,* Cartesian, symmetric*) if the inducing BRSA* F *is universal (respectively, Cartesian, symmetric).*

When $\mathcal{C}_1 \times \mathcal{C}_2 \subset \mathrm{Dom}(F)$ *we'll use* $F_{\mathcal{C}_1,\mathcal{C}_2}$ *to denote the BRSA obtained by restricting* F*'s domain to the subdomain* $\mathcal{C}_1 \times \mathcal{C}_2$*, and* $\mathcal{F}_{\mathcal{C}_1,\mathcal{C}_2}$ *to denote the corresponding restriction of the induced scoring rule.*[5]

We have in mind that a single universal and symmetric BRSR $\mathcal{F}$ can, via restriction to various Cartesian subdomains $\mathcal{C}_1 \times \mathcal{C}_2$, yield an assortment of more familiar and concrete aggregation rules. When $\mathcal{C}_1 \neq \mathcal{C}_2$, symmetry of the mother rule $\mathcal{F}$ is lost in the restriction $\mathcal{F}_{\mathcal{C}_1,\mathcal{C}_2}$, if only because of the asymmetry of $\mathcal{C}_1 \times \mathcal{C}_2$. This factor may explain why scoring rules have not previously been recognized as arising from symmetric measures of similarity between binary relations.

3 The universal and symmetric rule $\mathcal{F}^H$

Suppose our finite set A of alternatives has cardinality m. Choose some reference enumeration $\{(x_i, y_i)\}_{i=1}^{(m)(m-1)}$ of all pairs $(x, y) \in A \times A$ that are *distinct* (satisfy $x \neq y$).

Definition 7 *Let* $\mathcal{J}^H : 2^{A\times A} \to \Re^{(m)(m-1)}$ *be defined as follows: for each binary relation* $R \in 2^{A\times A}$ *on* A *and each* $i = 1, 2, \ldots, (m)(m-1)$*:*

- $\mathcal{J}^H(R)_i = 1$*, if* $(x_i, y_i) \in R$
- $\mathcal{J}^H(R)_i = -1$*, if* $(x_i, y_i) \notin R$.

Thus $\mathcal{J}^H$ embeds the set $2^{A\times A}$ of all binary relations on A into Euclidean space $\Re^{(m)(m-1)}$ of dimension $(m)(m-1)$. Geometrically, $\mathcal{J}^H$ identifies binary[6] relations with vertices of a certain hypercube H in $\Re^{(m)(m-1)}$, with $\mathcal{J}^H(R)$ serving as a symmetric version of R's *characteristic vector* (as qualified by footnote 6). H is centered at the origin, aligned with the coordinate axes, with side-length 2.

Note that the norm $\|\mathcal{J}^H(R)\|$ of each embedded relation has the same value (here, $\sqrt{(m)(m-1)}$), so these vertices also sit on the sphere of radius $\sqrt{(m)(m-1)}$ centered at the origin. $\mathcal{J}^K$ thus satisfies the following definition:

Definition 8 *A function* $\mathcal{J} : 2^{A\times A} \to \Re^k$ *is* spherical *if the Euclidean norm* $\|\mathcal{J}(R)\|_2$ *has the same value for every binary relation* $R \in 2^{A\times A}$.

Our initial formulation of F^H is via the inner product (aka dot product) of these symmetric characteristic functions. More precisely:

Definition 9 *The universal and symmetric binary relational scoring assignment* F^H*, with corresponding scoring rule* $\mathcal{F}^H$*, is given by*

$$F^H(R_1, R_2) = \mathcal{J}^H(R_1) \cdot \mathcal{J}^H(R_2) = \sum_{i=1}^{(m)(m-1)} (\mathcal{J}^H(R_1)_i)\,(\mathcal{J}^H(R_2)_i). \tag{1}$$

[4] As one would expect, $\mathcal{O}$ can contain a single relation O, or more than one (tied) relations, when several $O \in \mathcal{C}_2$ share a greatest total.

[5] It does not matter whether one first induces a scoring rule from a scoring assignment and then restricts the domain, or restricts first and induces second. The result is the same, and we'll be sloppy about the distinction.

[6] This identification is bijective if we restrict $\mathcal{J}^H$'s domain to reflexive relations only (or to antireflexive relations only), but maps any two relations to the same vertex if they differ only on pairs (x, x) on the diagonal of $A \times A$.

and equivalently by a symmetric count of pairs on which R_1 and R_2 agree and disagree:

$$F^H(R_1, R_2) = |\{(x,y) \in A \times A \mid x \neq y \text{ and } (x,y) \in R_1 \Leftrightarrow (x,y) \in R_2\}| \\ - |\{(x,y) \in A \times A \mid x \neq y \text{ and } (x,y) \in R_1 \Leftrightarrow (x,y) \notin R_2\}|. \tag{2}$$

We leave the proof of equivalence to the reader. The standard geometric interpretation of the Euclidean inner product tells us that

$$F^H(R_1, R_2) = \underline{\|\mathcal{J}^H(R_1)\|}\ \underline{\|\mathcal{J}^H(R_2)\|}\ cos(\theta) \tag{3}$$

and we know that the underlined terms are constant. Thus $F^H(R_1, R_2)$ is a scaled version of the cosine of the angle between the spatial locations (as determined by $\mathcal{J}^H$) of the two relations; this interpretation supports the view that $F^H(R_1, R_2)$ is a symmetric measure of similarity between binary relations. The equivalent formulation (2) supports this view as well, and also yields a particularly simple characterization of the induced scoring rule as a form of pairwise majority rule:

Proposition 2 *Let $P = \{B_k\}_{k \in N}$ be a profile of binary relations, and R be any binary relation on A. For $(x,y) \in A \times A$ with $x \neq y$ let* $\text{supp}(x,y)$ *denote* $|\{k \in N \mid (x,y) \in B_k\}|$. *Then $R \in \mathcal{F}^H(P)$ if and only if R meets majoritarian condition (4)*

$$\text{supp}(x,y) > \frac{|N|}{2} \implies (x,y) \in R \implies \text{supp}(x,y) \geq \frac{|N|}{2} \tag{4}$$

for each $(x,y) \in A \times A$ with $x \neq y$.

Proposition 2 shows that at the level of generality of a universal BRSR, it is possible to reconcile the conflict between being a scoring rule and satisfying (an abstract version of) Condorcet's principle; one could say that this conflict arises from the restriction process. This proposition may also seem to suggest that rule $\mathcal{F}^H$ is both trivial and uninteresting, so we turn next to arguing otherwise, by noting the remarkable variety of familiar aggregation rules that can be obtained as restrictions of $\mathcal{F}^H$. We'll have more to say, in the next section, about the intuition behind definition 9.

In perusing the following Theorem 1 the reader should note both the earlier comments on plurality and approval ballots (immediately preceding definition 6), and the remarks that follow the theorem's statement. The proof (parts of which are in the Appendix) is elucidated by material in the next section.

Theorem 1 (Restrictions of $\mathcal{F}^H$) *The following aggregation rules are restrictions $\mathcal{F}^H_{\mathcal{C}_1,\mathcal{C}_2}$ of $\mathcal{F}^H$ to the indicated classes $\mathcal{C}_1$ (for ballots) and $\mathcal{C}_2$ (for election outcomes) of binary relations.*

- $\mathcal{F}^H_{\mathcal{L},\mathcal{L}}$ = *the Kemeny rule*
- $\mathcal{F}^H_{\mathcal{D},\mathcal{D}}$ = *the Mean rule of Duddy and Piggins [7] (defined below)*
- $\mathcal{F}^H_{\mathcal{D},\mathcal{U}}$ = *Approval voting$_1$ (the election outcome contains the individual alternative(s) with greatest Approval score)*
- $\mathcal{F}^H_{\mathcal{D},\mathcal{L}}$ = *Approval voting$_2$ (the election outcome is the weak order induced by Approval score)*
- $\mathcal{F}^H_{\mathcal{L},\mathcal{D}}$ = *the Borda Mean Dichotomy rule (defined below)*

- $\mathcal{F}^H_{\mathcal{U},\mathcal{U}}$ = *Plurality voting*$_1$ *(the election outcome contains the individual alternative(s) with greatest Plurality score)*

- $\mathcal{F}^H_{\mathcal{U},\mathcal{L}}$ = *Plurality voting*$_2$ *(the election outcome is the weak order induced by Plurality score)*

How do these restricted rules behave when there are ties? In Plurality voting$_1$ and Approval voting$_1$, a tie means that several alternatives t yield plurality orders $\{t\} > N \setminus \{t\}$ sharing a highest score. With Plurality voting$_2$ and Approval voting$_2$ ties imply that the order $\geq$ induced by score is weak and not linear. The election outcome represents such a weak order $\geq$ in the form of the set containing *all* linear orders that extend $\geq$ (by breaking all ties); these extensions all achieve the same greatest score. (In the Kemeny rule, as we know, the set of linear orders with highest score need not correspond, in this way, to any weak order.)

These restrictions need not satisfy Proposition 2, because if one applies majority rule on a pair-by-pair basis to aggregate ballots chosen from some specified class $\mathcal{C}_1$ of binary relations, there is no reason to expect the result to belong to a second specified class $\mathcal{C}_2$. Whether or not the restriction process retains some remnant of Condorcet's principle depends on the particular restriction.

Definition 10 *The* Mean rule *of Duddy and Piggins [7] aggregates dichotomous ballots $\{T_i > B_i\}_{i \in N}$ into a dichotomous social order $T > B$ by considering the average*

$$q = \frac{\sum\limits_{a \in A} |\{i \in N : a \in T_i\}|}{m} \tag{5}$$

number of approvals taken over all m alternatives. An alternative $a \in A$ is placed in T if a is approved more than q times, and is placed in B if approved fewer than q times. If there are alternatives approved exactly q times, each may be placed in T or in B, resulting in a tie among all dichotomous orders obtained by making such choices freely and independently for all such alternatives.

Definition 11 *When $\{R_i\}_{i \in N}$ is a profile of linear orders and $a \in A$ is an alternative, we'll use a^β to denote the Borda score of a, in the standard sense. The* Borda Mean Dichotomy rule *[8] aggregates linear ballots into a dichotomous social order $T > B$ by considering the average*

$$q^\beta = \frac{\sum\limits_{a \in A} a^\beta}{m} \tag{6}$$

Borda score taken over all m alternatives. An alternative $a \in A$ is placed in T if $a^\beta > q^\beta$ and is placed in B if $a^\beta < q^\beta$. If there are alternatives with $a^\beta = q^\beta$, each may be placed in T or in B, resulting in a tie among all dichotomous orders obtained by making such choices freely and independently for all such alternatives.

Note that Borda Mean Dichotomy is not a new rule – it is used as a single step in the iterative elimination process of *Nanson's rule.*

One method for calculating a Borda score in the context of weakly ordered ballots is to extend each such ballot to a linear order, and then award to each alternative x the average number of points awarded (according to the standard Borda score, applied to the linear extension) to all alternatives in x's I-class. If one applies that method to the Borda mean Dichotomy rule with dichotomous ballots, the result is the Mean rule of Definition 10.

The seven restrictions mentioned in Theorem 1 are not the only possibilities. For example, the restriction $\mathcal{F}^H_{\mathcal{E},\mathcal{E}}$ seems to be an interesting and plausible rule for aggregating equivalence relations, and could be adjusted via further restrictions that impose an endogenously specified number of equivalence classes (among ballots, outcomes, or both). For that matter, the restrictions of $\mathcal{F}^H_{\mathcal{W},\mathcal{W}}$ and $\mathcal{F}^H_{\mathcal{P},\mathcal{P}}$ also seem worth exploring, but we leave that for the most part to the future, except for a brief comment in section 5.[7]

4 Pairwise Scoring Rules

One way to generate a scoring weight $F(R_1, R_2)$ from a pair of binary relations is to sum independent contributions from each pair $(x, y) \in A \times A$. We will think of (x, y)'s contribution $K_{R_1,R_2}(x, y)$ as measuring the extent to which R_1 and R_2 "agree" on the two alternatives x and y, and will assume that each contribution depends only on the values of *pair-case*$_{R_1}(x, y)$ and *pair-case*$_{R_2}(x, y)$ (see definition 4) – equivalently, $K_{R_1,R_2}(x, y)$ depends only on the vector of answers to the following four membership questions:

$$(x, y) \in R_1? \quad (y, x) \in R_1? \quad (x, y) \in R_2? \quad (y, x) \in R_2? \tag{7}$$

More precisely:

Definition 12 *A* pairwise contribution function K *is a partial function*[8] *that assigns a real number output* $K_{R_1,R_2}(x, y)$ *to a pair* (R_1, R_2) *of binary relations on* A *together with a distinct pair* $(x, y) \in A \times A$*, and that satisfies the following condition:*

$$\begin{array}{c} \text{pair-case}_{R_1}(x, y) = \text{pair-case}_{R_1}(x', y') \text{ and } \text{pair-case}_{R_2}(x, y) = \text{pair-case}_{R_2}(x', y') \\ \Longrightarrow \\ K_{R_1,R_2}(x, y) = K_{R_1,R_2}(x', y') \end{array} \tag{8}$$

for all distinct pairs (x, y), $(x', y') \in A \times A$. K *satisfies* relational symmetry *if* $K_{R_1,R_2}(x, y) = K_{R_2,R_1}(x, y)$ *holds for all* R_1, R_2, *and* $x \neq y$, *and satisfies* pair-reversal symmetry *if* $K_{R_1,R_2}(x, y) = K_{R_1,R_2}(y, x)$ *holds for all* R_1, R_2, *and* $x \neq y$.[9]

Definition 13 *A BRSA* F *is* pairwise *if is a restriction of some BRSA*

$$F_K(R_1, R_2) \quad = \sum_{(x,y)\,\in\, A\times A,\; x\neq y} K_{R_1,R_2}(x, y) \tag{9}$$

induced as in equation (9) by a pairwise contribution function K*. The domain of* F_K *is* $\{(R_1, R_2) \in 2^{A\times A} \times 2^{A\times A} \mid K_{R_1,R_2}(x, y)$ *is defined for all* $(x, y) \in A \times A$ *with* $x \neq y\}$.

As each contribution $K_{R_1,R_2}(x, y)$ is eventually added to the contribution $K_{R_1,R_2}(y, x)$ of the reverse pair, it seems that one might as well replace a K that fails pair-reversal symmetry with a symmetrized version $K^{Sym}_{R_1,R_2}(x, y) = \frac{1}{2}[K_{R_1,R_2}(x, y) + K_{R_1,R_2}(y, x)]$, but some K will arise in an initially asymmetric form.

An pairwise contribution function K may be presented in the form of a *pairwise contribution table* similar to Table 1A (below). Any such table is read as follows: given any

[7] Observe that when R_1 is a linear order and R_2 is an equivalence relation, $\mathcal{J}^H(R_1) \cdot \mathcal{J}^H(R_2) = 0$ (this can be seen more directly in Table 1B, section 4). Such relations are orthogonal, according to $\mathcal{J}^H$, implying that $\mathcal{F}^H$ has no useful restrictions that aggregate equivalence relations to obtain linear orders (or vice-versa). Is this consequence peculiar to $\mathcal{J}^H$, or reflective of some more general sense in which these two classes of relations are somehow orthogonal? We suspect the latter, but as yet have no formulation expressing that sense precisely.

[8] We allow $K_{R_1,R_2}(x, y)$ to be undefined for some values of R_1, $R_2 \in 2^{A\times A}$ and $x, y \in A$ with $x \neq y$.

[9] Our convention for either of these symmetry equations will be that they hold when both sides are undefined, but fail when exactly one side is defined.

distinct pair (x, y) of alternatives, we choose the row i labeled with $\textit{pair-case}_{R_1}(x, y)$ and the column j labeled with $\textit{pair-case}_{R_2}(x, y)$; $K_{R_1,R_2}(x, y)$ given by the number in the (i, j) cell of the table. The particular entries of Table $1A$ represent a pairwise contribution function K^H that induces the same BRSA F^H defined in the previous section. We can see this by matching the table entries with terms of the sum representing F^H (in definition 9) as an inner product. For example, suppose (a, b) is a pair of distinct alternatives, with $(a, b) = (x_i, y_i)$ and $(b, a) = (x_j, y_j)$, in the enumeration fixed at the start of section 3. Further assume $(a, b) \in R_1 \cap R_1^{reverse}$ and $(a, b) \in \overline{R_2} \cap R_2^{reverse}$. Then the i^{th} term in the inner product sum (of equation (1)) is $(+1)(-1) = -1$, matching the $(3, 2)$ entry $K^H_{R_1,R_2}(a, b)$ in Table $1A$. The j^{th} term makes a contribution of $(+1)(+1) = 1$, and this is the $(3, 1)$ entry (as dictated by the pair-cases of (b, a) for the two relations).

(x, y)	$R_2 \cap \overline{R_2^{reverse}}$	$\overline{R_2} \cap R_2^{reverse}$	$R_2 \cap R_2^{reverse}$	$\overline{R_2} \cap \overline{R_2^{reverse}}$
$R_1 \cap \overline{R_1^{reverse}}$	1	-1	1	-1
$\overline{R_1} \cap R_1^{reverse}$	-1	1	-1	1
$R_1 \cap R_1^{reverse}$	1	-1	1	-1
$\overline{R_1} \cap \overline{R_1^{reverse}}$	-1	1	-1	1

Table 1A: The table for the pairwise contribution function K^H, inducing the universal and symmetric BRSA F^H.

In particular, $K^H_{R_1,R_2}(a, b) \neq K^H_{R_1,R_2}(b, a)$, and so the K^H given by Table $1A$ fails to satisfy pair-reversal symmetry. More generally, pair-reversal symmetry of a pairwise contribution function K is equivalent to the following property for K's table: the combined effect of switching the first row of entries with the second row and also switching the first column of entries with the second column (without switching labels for these rows and columns) is to leave the entries unchanged.[10]

Thus, when we construct Table $1A$ so as to match the inner product formula on a term-by-term basis, the consequence is a table that violates the condition for pair-reversal symmetry. However, if we then apply to K^H the averaging process described earlier (immediately after definition 13), the result is the reversal-symmetric function $K^{H\star}$ of Table $1B$. For example, the $(3, 2)$ entry of -1 and $(3, 1)$ entry of $+1$ (from Table $1A$) are averaged, yielding 0 as the Table $1B$ entries in the $(3, 2)$ and $(3, 1)$ positions. Thus while $K^{H\star}$ does satisfy pair-reversal symmetry, and induces the same BRSA F^H as Table $1A$, it no longer matches the inner product formula on a term-by-term basis.

(x, y)	$R_2 \cap \overline{R_2^{reverse}}$	$\overline{R_2} \cap R_2^{reverse}$	$R_2 \cap R_2^{reverse}$	$\overline{R_2} \cap \overline{R_2^{reverse}}$
$R_1 \cap \overline{R_1^{reverse}}$	1	-1	0	0
$\overline{R_1} \cap R_1^{reverse}$	-1	1	0	0
$R_1 \cap R_1^{reverse}$	0	0	1	-1
$\overline{R_1} \cap \overline{R_1^{reverse}}$	0	0	-1	1

Table 1B: The table for the pairwise contribution function $K^{H\star}$, which satisfies pair-reversal symmetry and induces the same BRSA F^H.

[10]Symmetry of the table about the diagonal (i.e., equality of the (i, j) entry with the (j, i) entry) corresponds to relational symmetry of the corresponding pairwise contribution function K. Table $1A$ *is* symmetric in this sense, as are the other tables we provide.

What role do these two tables play? Table 1A suffices to establish that the universal and symmetric BRSA F^H is a *pairwise* rule (Definition 13, which is consistent with the definition of "pairwise" in [18]). Of course, it is also the case that F^H arises as an inner product, so it is worth remarking that these properties are independent; a symmetric BRSA can be pairwise without arising as an inner product, and can arise as an inner product without being pairwise (and such examples exist even among universal BRSAs), Nonetheless, the natural examples that we know of (at the time of writing) are pairwise and also arise as inner products (hence, they are symmetric), although not all are universal.

Table 1B may be thought of as a simplified version of 1A. We turn to Table 1B to gain insight when thinking about the nature of various restrictions of the scoring rule $\mathcal{F}^H$ (in the proof of theorem 1, for example) or when comparing corresponding restrictions for two different BRSAs (as in the next section). As one example, consider the scoring rule $\mathcal{F}^H_{2^{A\times A},\mathcal{E}}$ which aggregates arbitrary relations[11] into an equivalence relation and is represented by the eight entries of columns 3 and 4 of Table 1B. What do the four zeros (among these eight entries) tell us one thing about the behavior of $\mathcal{F}^H_{2^{A\times A},\mathcal{E}}$ ...and what do the four nonzero entries tell us?

5 $\mathcal{F}^S$ – an alternative to $\mathcal{F}^H$

We next consider a different unrestricted pairwise extension $\mathcal{F}^S$, which shares some, but not all, of $\mathcal{F}^H$'s properties.

Definition 14 *Let $\mathcal{J}^S : 2^{A\times A} \to \Re^{(m)(m-1)}$ be defined as follows: for each binary relation $R \in 2^{A\times A}$ on A and each $i = 1, 2, \ldots, (m)(m-1)$:*

- $\mathcal{J}^S(R)_i = 1$, *if* $(x_i, y_i) \in R \cap \overline{R^{\text{reverse}}}$
- $\mathcal{J}^S(R)_i = -1$, *if* $(x_i, y_i) \in \overline{R} \cap R^{\text{reverse}}$
- $\mathcal{J}^S(R)_i = 0$, *if* $(x_i, y_i) \in R \cap R^{\text{reverse}}$ *or* $(x_i, y_i) \in \overline{R} \cap \overline{R^{\text{reverse}}}$

Thus $\mathcal{J}^S$ is an alternative embedding of binary relations on A into $\Re^{(m)(m-1)}$. Unlike the case for $\mathcal{J}^H$, the norm $\|\mathcal{J}^S(R)\|$ of an embedded relation varies for different R, so $\mathcal{J}^S$ is not spherical. While some relations are still located at vertices of the same hypercube H, others now sit at midpoints of H's lower-dimensional faces (or at the origin). Our first formulation of F^H is again via an inner product:

Definition 15 *The universal and symmetric binary relational scoring assignment F^S, with corresponding scoring rule $\mathcal{F}^S$, is given by*

$$F^S(R_1, R_2) = \mathcal{J}^S(R_1) \cdot \mathcal{J}^S(R_2) = \sum_{i=1}^{(m)(m-1)} (\mathcal{J}^S(R_1)_i)\,(\mathcal{J}^S(R_2)_i). \tag{10}$$

Note that for $(a, b) \in A \times A$ with $a \neq b$, if $(a, b) = (x_i, y_i)$ and $(b, a) = (x_j, y_j)$, then $\mathcal{J}^S_i = \mathcal{J}^S_j$. This was not the case for $\mathcal{I}^H$, and explains why there is no need to apply averaging to Table 2; its initial form – wherein table entries match the corresponding terms of the sum in (10) – is already pair-reversal symmetric:

[11] A ballot that fails itself to be an equivalence relation might be interpreted as a *noisy* observation of some "true" equivalence relation, in the spirit of *Condorcet's Jury Theorem.*

(x, y)	$R_2 \cap \overline{R_2^{reverse}}$	$\overline{R_2} \cap R_2^{reverse}$	$R_2 \cap R_2^{reverse}$	$\overline{R_2} \cap \overline{R_2^{reverse}}$
$R_1 \cap \overline{R_1^{reverse}}$	1	-1	0	0
$\overline{R_1} \cap R_1^{reverse}$	-1	1	0	0
$R_1 \cap R_1^{reverse}$	0	0	0	0
$\overline{R_1} \cap \overline{R_1^{reverse}}$	0	0	0	0

Table 2: The table of pairwise contributions inducing $\mathcal{F}^S$.

In what ways are $\mathcal{F}^S$ and $\mathcal{F}^H$ similar? Different? As Tables 1B and 2 agree in all but 4 of their 16 entries, it is clear that any restriction of $\mathcal{F}^S$ that never makes use of these 4 cells must be equal to the corresponding restriction of $\mathcal{F}^H$. For example, $\mathcal{F}^S_{\mathcal{L},\mathcal{L}} = \mathcal{F}^H_{\mathcal{L},\mathcal{L}}$ (each is the Kemeny rule) and $\mathcal{F}^S_{\mathcal{P},\mathcal{L}} = \mathcal{F}^H_{\mathcal{P},\mathcal{L}}$; more generally, for *any* class $\mathcal{X}$ of binary relations, $\mathcal{F}^S_{\mathcal{X},\mathcal{L}} = \mathcal{F}^H_{\mathcal{X},\mathcal{L}}$ and $\mathcal{F}^S_{\mathcal{L},\mathcal{X}} = \mathcal{F}^H_{\mathcal{L},\mathcal{X}}$.

On the other hand, thanks to the different $(3, 3)$ entries of Tables 1B and 2 (and the different$(4, 4)$ entries), some very simple profiles demonstrate that $\mathcal{F}^S_{\mathcal{W},\mathcal{W}} \neq \mathcal{F}^H_{\mathcal{W},\mathcal{W}}$ (respectively that $\mathcal{F}^S_{\mathcal{P},\mathcal{P}} \neq \mathcal{F}^H_{\mathcal{P},\mathcal{P}}$). Given a profile $\{W_i\}_{i \in N}$ of weak orders and a potential weak order outcome $W \notin \mathcal{L}$, let $W^\dagger$ be a linear extension of W. Then according to F^S, the total score $\sum_{i \in N} F^S(W_i, W)$ awarded W is equal to the score $\sum_{i \in N} F^S(W_i, W^\dagger)$ awarded to $W^\dagger$, so it is impossible for W to be the uniquely highest-scoring member of $\mathcal{W}$. With F^H it is certainly possible for some $W \in \mathcal{W} \setminus \mathcal{L}$ to achieve the uniquely highest score (for example, when the profile consists of unanimous ballots of W). Something similar happens when one substitutes $\mathcal{P}$ for $\mathcal{W}$.

Our tentative interpretation is that $\mathcal{F}^S$ treats weak orders (and partial orders) purely as knife-edge transitions among a set of tied linear orders, while $\mathcal{F}^H$ treats them as credible stand-alone outcomes.[12] Perhaps the only distinction between $\mathcal{F}^S_{2^{A \times A}, 2^{A \times A}}$ and $\mathcal{F}^S_{2^{A \times A}, \mathcal{L}}$ is that the non-linear relations achieving maximal score are explicit for the former and are implied (by the set of linear orders achieving greatest score) for the latter. If so, that might suggest that $\mathcal{F}^S$ has no *raison d'être*, representing a "naive" extension of Kemeny, while $\mathcal{F}^H$ is a more sophisticated extension. But these assertions are speculative at this time.

6 Euclidean Distance Rationalization of Inner Product Scoring Rules

The standard approach to distance rationalization of a voting rule G (see [1], [2], [4], [9], [10], [11], [12], [13], [14], [16], [17]) begins with a metric δ on the space of possible ballots, and extends it via summation to a measure of distance on profiles, with $\overline{\delta}(\{B_i\}_{i \in N}, \{B'_i\}_{i \in N}) = \sum_{i \in N} \delta(B_i, B'_i)$. If $\{B'_i\}_{i \in N}$ represents a *consensus* among the voters that the winner is t, and is closer to $\{B_i\}_{i \in N}$ than is any other such consensus profile, we demand $G(\{B_i\}_{i \in N}) = t$. If one can identify a metric δ as well as a notion $\mathcal{S}$ of consensus such that all such demands are satisfied, we say that G is metric rationalized via δ and $\mathcal{S}$. Reasonable notions of consensus in favor of t include "unanimous rankings (with t top-ranked)," "existence of a Condorcet alternative t," and others.

Euclidean distance rationalization for a voting rule G differs in some respects, but is closely related in spirit (see [20]). We assume that ballots are drawn from set I and election outcomes from another set O, and with functions $\mathcal{J}_I : I \to \Re^k$ and $\mathcal{J}_O : O \to \Re^k$ used to

[12] This suggests the possibility that a *Maximum Likelihood Estimator* interpretation of $\mathcal{F}^S$ views weak or partial orders as noisy observations of an objectively "correct" ground truth consisting of a linear order, while $\mathcal{F}^H$ allows the ground truth itself to be weak or partial.

induce the Euclidean metric (so that $\delta(b,o) = \|\mathcal{J}_I(b) - \mathcal{J}_O(o)\|_2$ for each $(b,o) \in I \times O$, with $\| \ \|_2$ denoting the standard Euclidean norm). The extension $\overline{\delta}$ is now a bit different. If $\{b_i\}_{i\in I}$ is a profile of ballots in I and $o \in O$, we set

$$\overline{\delta}(\{b_i\}_{i\in N}, o) = \big(\textstyle\sum_{i\in N} \delta(b_i,o)^2\big)^{\frac{1}{2}} = \big\|\big(\delta(b_1,o), \delta(b_2,o), \ldots, \delta(b_n,o)\big)\big\|_2 \qquad (11)$$

Notice that in the special case $I = O$, $\mathcal{J}_I = \mathcal{J}_O$ this agrees with the earlier extension of δ to $\overline{\delta}$ with the second profile consisting of unanimous ballots of o . . . except of course that we are using the $\| \ \|_2$-norm rather than the $\| \ \|_1$-norm to extend metrics to cartesian product spaces.[13] Euclidean rationalization of rule G via $\mathcal{J}_I$ and $\mathcal{J}_O$ now demands that $G(\{b_i\}_{i\in N})$ equal the element $o \in O$ minimizing $\overline{\delta}(\{b_i\}_{i\in N}, o)$.[14] Our immediate goal here is:

Proposition 3 *The universal and symmetric scoring rule $\mathcal{F}^H$ is Euclidean rationalizable via the same embedding $\mathcal{J}^H$ used, in definition 9, to represent F^H as an inner product (with $\mathcal{J}^H$ serving both as $\mathcal{J}_I$ and as $\mathcal{J}_O$ in the above definition of Euclidean rationalization).*

Corollary 1 *All restrictions $\mathcal{F}^H_{\mathcal{C}_1,\mathcal{C}_2}$ of $\mathcal{F}^H$ are Euclidean rationalizable via the common embedding $\mathcal{J}^H$ – i.e., via a common induced Euclidean metric.*

There is a limited sense, then, in which the Kemeny rule, plurality voting, the Borda mean dichotomy rule, etc. are all Euclidean rationalized by the same metric.[15] The very short proof of proposition 3 that follows relies on $\mathcal{J}^H$ being spherical.[16]

Proof: For R_i, R any two binary relations,

$$\begin{aligned} \|\mathcal{J}^H(R_i) - \mathcal{J}^H\|_2^2 &= (\mathcal{J}^H(R_i) - \mathcal{J}^H(R)) \cdot (\mathcal{J}^H(R_i) - \mathcal{J}^H) \\ &= \underline{\|\mathcal{J}^H(R_i)\|_2^2 + \|\mathcal{J}^H(R)\|_2^2} - 2\mathcal{J}^H(R_i) \cdot \mathcal{J}^H(R) \end{aligned} \qquad (12)$$

and the underlined term is constant, so that the R minimizing $\sum_{i\in N} \|\mathcal{J}^H(R_i) - \mathcal{J}^H(R)\|_2^2$ is the same as the R maximizing $\sum_{i\in N} \mathcal{J}^H(R_i) \cdot \mathcal{J}^H(R)$. ∎

7 Borda-like Rules and Concluding Remarks

For x an alternative and R a binary relation, let xR denote $\{y \in A \,|\, (x,y) \in R\}$, Rx denote $\{y \in A \,|\, (y,x) \in R\}$ and $x^\beta_R = |xR| - |Rx| = |xR \backslash Rx|$. It is easy to see that when R happens to be a linear order, x^β_R is a symmetric version of the Borda scoring weight awarded x by ranking R. Let $\mathcal{J}^B : 2^{A\times A} \to \Re^m$ by $\mathcal{J}^B(R)_i = (x_i)^\beta_R$, where $x_1, x_2, \ldots, x_m$ enumerates A. Thus $\mathcal{J}^B$ is a third embedding of binary relations on A into Euclidean space (of lower dimension than that for the earlier embeddings).

Linear orders are mapped by $\mathcal{J}^B$ to the vertices of the m-permutahedron (see [20]) and these vertices are equidistant from the origin, but other relations are sent closer, and $\mathcal{J}^B$ is not spherical. The universal and symmetric binary relational scoring assignment F^B, with corresponding scoring rule $\mathcal{F}^B$, is given by $F^B(R_1, R_2) = \mathcal{J}^B(R_1) \cdot \mathcal{J}^B(R_2) =$

[13] Except for the historic precedent set by Kemeny in extending Kendall distance between individual rankings, I am unclear why the standard distance-rationalization approach extends via the $\| \ \|_1$-norm only.

[14] As shown in [20] Euclidean rationalizability is equivalent to being a *Mean Proximity Rule* (G always selects the output closest to the mean input) and to being a scoring rule (in the appropriate sense).

[15] "Limited" in that the common embedding $\mathcal{J}^H$ does not map two different classes of binary relations to the same points of $\Re^{(m)(m-1)}$.

[16] Lirong Xia has pointed out (private communication) that the argument goes through providing at least one of $\mathcal{J}_I$, $\mathcal{J}_O$ are spherical.

$\sum_{i=1}^{m}(\mathcal{J}^B(R_1)_i)\ (\mathcal{J}^B(R_2)_i)$. Examples suggest that $\mathcal{F}^B_{\mathcal{L},\mathcal{L}}$ is the Borda count (in the same sense that $\mathcal{F}^H_{\mathcal{D},\mathcal{L}}$ was a form of approval voting), but we do not yet have a proof. Preliminary investigation suggest that certain of the other restrictions of $\mathcal{F}^B$ may agree with the corresponding restrictions of $\mathcal{F}^H$, but that restrictions of $\mathcal{F}^B$ to partial orders may not be interesting.

Is there a broad class of universal and symmetric binary relational scoring rules that are truly interesting, or is $\mathcal{F}^H$ the only really notable example? Does the uniform form of Euclidean distance rationalization for all $\mathcal{F}^H$ restrictions extend to a universal, interesting Euclidean version of maximum likelihood rationalization for all its restrictions? These are matters for further study.

References

[1] Baigent, N., Metric rationalization of social choice functions according to principles of social choice. In *Mathematical Social Sciences* 13, 59-65, 1987.

[2] Baigent, N. and Klamler, C. Transitive closure, proximity and intransitivities. In *Economic Theory* 23 (1), 175-181, 2004.

[3] Barthélemy, J.P. and Monjardet, M. The median procedure in cluster analysis and social choice theory. In *Mathematical Social Sciences* 1 (3), 235-267, 1981.

[4] Campbell, D. and Nitzan, S. Social compromise and social metrics. In *Social Choice and Welfare* 3 (1), 1-16, 1986.

[5] Conitzer, V., Rognlie, M., and Xia, L. Preference functions that score rankings and maximum likelihood estimation. In *Proceedings of the Twenty-First International Joint Conference on Artificial Intelligence (IJCAI-09)*, 109-115, 2009.

[6] Dietrich, F. Scoring rules for judgment aggregation. In *Social Choice and Welfare* 42 (4), 873-911, 2014.

[7] Duddy, C., and Piggins, A. Aggregation of binary evaluations: a Borda-like approach. Preprint dated August 5, 2013. (Newer version with W.S. Zwicker, dated January 16, 2014.)

[8] Duddy, C., Houy, N., Lang, J., Piggins, A., and Zwicker, W.S. Social Dichotomy Functions. Extended abstract for presentation at 2014 meeting of the *Society for Social Choice and Welfare.*

[9] Eckert, D. and Klamler, C. Distance-based aggregation theory. In *Consensual Processes (Studies in Fuzziness and Soft Computing Volume 267)* 3-22, 2011.

[10] Elkind, E., Faliszewski, P., and Slinko, A. On the role of distances in defining voting rules. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: (AAMAS 10)* 1, 375-382, 2010.

[11] Elkind, E., Faliszewski, P., and Slinko, A. Good Rationalizations of Voting Rules. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI-10)*, 774-779, 2010.

[12] Lerer, E., and Nitzan, S. Some General Results on the Metric Rationalization for Social Decision Rules. In *Journal of Economic Theory* 37, 191-201, 1985

[13] Meskanen, T., and Nurmi, H. Distance from consensus: A theme and variations. In *Mathematics and Democracy. Recent Advances in Voting Systems and Collective Choice* (B. Simeone, F. Pukelsheim, eds.), Springer, 117-132, 2007.

[14] Meskanen, T., and Nurmi, H. Closeness counts in social choice. In *Freedom, and Voting* (M. Braham and F.Steffen, eds.), Springer, 289-306, 2008.

[15] Myerson, C. Axiomatic derivation of scoring rules without the ordering assumption. In *Soc. Choice Welfare* 12, 109-115, 1995.

[16] Nitzan, S. Some measures of closeness to unanimity and their implications. In *Theory and Decision* 13, 129-138, 1981.

[17] Pivato, M. Voting rules as statistical estimators. In *Social Choice and Welfare* 40, 581-630, 2013.

[18] Xia, L. and Conitzer, V. A maximum likelihood approach towards aggregating partial orders. In *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI-11) 1*, 446 - 451, 2011.

[19] Xia, L. Generalized scoring rules: a framework that reconciles Borda and Condorcet. To appear in *SIGecom Exchanges.*

[20] Zwicker, W.S. Consistency without neutrality in voting rules: when is a vote an average? In A. Belenky, edior, *Mathematical and Computer Modelling 48, special issue on Mathematical Modeling of Voting Systems and Elections: Theory and Applications*, 1357-1373, 2008.

This material owes much to related joint work with Jérôme Lang, Conal Duddy, Ashley Piggins, and Nicolas Houy, and to recent conversations with Lirong Xia. We are grateful to Xavier Mora for pointing out the connections to the paper [3] by Barthélemy and Monjardet.

8 Appendix I

We provide proofs for a few parts of Theorem 1:

$\mathcal{F}^H_{\mathcal{D},\mathcal{L}}$ = Approval voting$_2$ (the election outcome is the weak order induced by Approval score, expressed as a tie among all linear orders that extend this induced weak order.)

Proof: Recall that A is our set of alternatives, and N is the set of voters. Let $P = \{X_i\}_{i\in N}$ be an approval profile (so that $X_i \subset A$ is the set of alternatives approved by voter i) and $P^\dagger = \{D_i\}_{i\in N}$ be the corresponding profile of dichotomous orders; in terms of the notation introduced in Definition 5, this means that $D_i = X_i > N \setminus X_i$ for each $i \in N$.

For any $x \in A$, let

$$App(x) = |\{i \in N \,|\, x \in X_i\}| \tag{13}$$

denote the conventional approval score of alternative x, and for any linear order relation $R \in \mathcal{L}$, let

$$Score(F^H, R) = \sum_{i\in N} F^H(D_i, R) \tag{14}$$

denote the total F^H-*score* of relation R for the profile $P^\dagger$. We need to establish that a linear order achieves a maximal F^H-score if and only if it is an extension of the weak order induced by approvel score.

Suppose that $>$ is a linear ordering of A and $>$'s F^H-score is at least as great as that for any other linear ordering of A. Rename the alternatives so that $x_1 > x_2 > \cdots > x_j > x_{j+1} > \cdots > x_m$. To show that $>$ is an extension of the weak order induced by approvel score, it suffices to show for each two successive alternatives $x_j > x_{j+1}$ that $App(x_j) \geq App(x_{j+1})$. Let $>_\tau$ be the ordering of A that transposes x_j and x_{j+1} but otherwise agrees with $>$: $x_1 >_\tau x_2 >_\tau \cdots >_\tau x_{j-1} >_\tau x_{j+1} >_\tau x_j >_\tau x_{j+2} \cdots >_\tau x_m$. Let $N_{x>y}$ denote $\{i \in N : x \in X_i \textit{ and } y \notin X_i\}$. Then any single voter who approves x_j and does not approve x_{j+1} contributes 1 to $Score(F^H, >)$ and -1 to $Score(F^H, >_\tau)$, for a total contribution of 2 to the difference $Score(F^H, >) - Score(F^H, >_\tau)$, while any who approves x_j and not x_{j+1} contributes -2 to that score difference. Thus

$$0 \geq Score(F^H, >) - Score(F^H, >_\tau) = 2\big(|N_{x_j > x_{j+1}}| - |N_{x_{j+1} > x_j}|\big) = 2\big(App(x_j) - App(x_{j+1})\big) \quad (15)$$

whence $App(x_j) \geq App(x_{j+1})$, as desired.

Now consider any linear order $>_1$ that extends the weak order induced by approval scores. We'll show it achieves a maximal F^H-score. Choose any linear order $>_2$ that does achieve a maximal F^H-score. Then $>_2$ also extends that weak order (as shown in the previous paragraph). It follows that $>_2$ may be transformed into $>_1$ via a sequence of transpositions of alternatives that are order-adjacent just prior to being transposed. Equation 15 shows that each such transposition has no effect on F^H-score so $>_1$'s F^H-score is also greatest possible. ■

$\mathcal{F}^H_{\mathcal{L},\mathcal{D}}$ = the Borda Mean Dichotomy rule

The proof that follows imports some ideas and results from [8]. Consider the complete directed graph $\mathcal{G}$ on the vertex set A; this means that for every two distinct alternatives $x \neq y$, $\mathcal{G}$ has both a directed edge (x, y) from x to y and a directed edge (y, x) from y to x. An *antisymmetric edge weighting* w assigns a real number edge weight $w(x, y)$ to each such edge and satisfies that $w(y, x) = -w(x, y)$ for each $x, y \in A$ with $x \neq y$.

Proof: (a) Let A, N, $P = \{>_i\}_{i \in N}$ and F be as stated. Recall from [8] that the induced net preference flow $(\mathcal{G}_P, w_P)$ consists of the complete directed graph $\mathcal{G}$ on vertex set A, paired with the antisymmetric edge weighting $w(x, y) = Net_P(x > y) = |\{i \in N : x \geq_i y\}| - |\{i \in N : y \geq_i x\}|$. By the Anti-symmetric Graph Cut Theorem of [8], the dichotomies $\mathcal{X}$ maximizing $ToSep(\mathcal{X}, w_P)$ are the Borda Mean dichotomies. To show that these coincide with the dichotomies accruing the maximal F-score for profile P it suffices to observe that $\mathcal{X}$'s F-score is equal to $ToSep(\mathcal{X}, w_P)$: either sum may be obtained by adding 1 point for each triple (x, y, i) in $X_T \times X_B \times N$ with $x >_i y$, and -1 point for triple (x, y, i) in $X_T \times X_B \times N$ with $y >_i x$. ■

9 Appendix II

Reference [3] of this paper, "The median procedure in cluster analysis and social choice theory" by J. P. Barthélemy and B. Monjardet (1981), discusses the broad applicability of the *median procedure* (which is the same as our $\mathcal{F}^H$) to amalgamation of binary relations from a number of the important classes we discuss here, including linear orders, weak orders (aka *complete preorders*), partial orders, and equivalence relations, as well as others we do not consider, such as tournaments, pointing out that a number of known amalgamation rules can be seen to be special cases of the median rule.

Barthélemy and Monjardet do not consider dichotomous weak orders or plurality orders, and so do not consider approval voting, plurality voting, the mean rule, or the Borda mean rule. The reason may be related to a difference in perspective; while our emphasis here is on scoring rules, with some secondary discussion of distance rationalization, [3] uses the distance approach exclusively. Consequently, it is not concerned with scoring weights that arise as inner products, but does discuss the equivalence of Kendall's *tau* to the squared Euclidean distance (under the "hypercube" embedding of relations into space), as well as the permutahedron embedding we touch on in Section 7, and connections with Maximum Likelihood Estimation. Note, as well, that to represent certain rules (both forms of approval voting that we consider here, one form of plurality voting, and the Borda mean rule) as restrictions of $\mathcal{F}^H$ it is essential that the class $\mathcal{C}_1$ of binary relations from which ballots are drawn be different from the class $\mathcal{C}_2$ of potential election outcomes. The focus in [3] is on examples for which the two classes are the same (although the description on page 238 of the Slater suggests using different classes).

It might be better for the material in this appendix to be integrated with the text, rather than appearing separately as it does here. The current approach is due primarily to shortness of time between our learning of the connections to [3] and the date for submitting revised Comsoc-2014 manuscripts.

William S. Zwicker
Mathematics Department
Union College
Schenectady, NY 12308
USA
Email: zwickerw@union.edu