# Social Dichotomy Functions
# (extended abstract)

Conal Duddy, Nicolas Houy, Jérôme Lang,
Ashley Piggins, and William S. Zwicker*

February 22, 2014

## 1 What is a Social Dichotomy Function?

A *dichotomy* $\mathcal{A} = (A_T, A_B)$ of a finite set $A$ is a signed partition of $A$ into nonempty subsets $A_T$ ($\mathcal{A}$'s *top* piece) and $A_B$ ($\mathcal{A}$'s *bottom* piece); $\mathcal{A}$ is *signed* in the sense that it is an *ordered* pair of sets, with the first of the two sets identified as the top. Equivalently, any dichotomy $\mathcal{A}$ may be identified with the corresponding dichotomous *weak order* $\geq_{\mathcal{A}}$, wherein $a \geq_{\mathcal{A}} b \Leftrightarrow a \in A_T$ or $b \in A_B$. For any weak order $\geq$ the induced indifference relation $\sim$ (defined by $a \sim b \Leftrightarrow$ both $a \geq b$ and $b \geq a$) is an equivalence relation, with equivalence classes referred to as *indifference classes*, aka *I-classes*, of $\geq$. The weak order $\geq_{\mathcal{A}}$ is *dichotomous* in that it has exactly two nonempty $I$-classes.

Dichotomies of a finite set $A$ of alternatives have been used as the inputs (*ballots*) of a decision rule, in the context of approval voting; a ballot approving a subset $X \subseteq A$ of alternatives can naturally be identified with the dichotomy $(X, A \setminus X)$. Our interest here is with decision rules that yield dichotomies as *outputs* – societal decisions – of an amalgamation process. The underlying goal of such a *social dichotomy function* or *SDF* will be to draw the brightest line possible between a top group and a bottom group.

Shortly, we'll begin to make that goal a bit less fuzzy. But some residual vagueness is, of course, necessary if we wish to leave ourselves open to the possibility of multiple amalgamation rules that operationalize the meaning of "brightest line" differently. Even in its current state of extreme fuzziness, however, this goal can be seen as distinct from that of a *social choice function* – aka *SCF* – as usually conceived. An *SCF* outputs a set of *winners*, with the *losers* being all non-winners, so one might argue that the goal of dichotomizing the set of alternatives is already well studied. However, with most *SCF*s there is an absence of symmetry between the top and bottom groups. For one thing, an *SCF* $\mathcal{F}$ often employs a mechanism that actually yields a weak ranking $\geq$ of $A$, in which

*Corresponding author. My co-authors are not responsible for any errors in this extended abstract.

the winners are those in $\geq$'s top $I$-class, and are thus indistinguishable (in the sense of $\mathcal{F}$'s particular mechanism) from one another. The losers come from all the remaining $I$-classes and as a result often *are* distinguishable via $\mathcal{F}$'s mechanism. In fact, for certain profiles the mechanism itself may suggest that there is little difference between the alternatives in the highest and second-highest $I$-classes, but that a huge gulf lies between the second-highest and all that lie below; in this case, the line immediately below the highest class is hardly the brightest possible. In a scoring rule, for example, the winners are all alternatives tied for highest score, while the losers often achieve very different scores from one another; approval voting behaves similarly (and may be viewed as a scoring rule, under a suitable generalization of the term).

We can make this distinction precise by reading as follows: we might plausibly expect the brightest line between top and bottom groups of alternatives to remain fixed when all voters switch from ranking alternatives in descending order of preference to ranking them in *ascending* order, yet an *SCF* $\mathcal{F}$ typically fails to satisfy the corresponding symmetry axiom:

**Definition 1** *(Ballot Reversal Symmetry) For each profile $P$ (of linear orders, or of weak orders) let $P^\star$ denote the profile obtained from $P$ by reversing the ballot $\geq_i$ of each voter $i$, so that $a \geq_i^\star b \Leftrightarrow b \geq_i a$. A social dichotomy function $\mathcal{F}$ satisfies* ballot reversal symmetry *if $\mathcal{F}(P) = (A_T, A_B) \Rightarrow \mathcal{F}(P^\star) = (A_B, A_T)$ holds for every profile $P$ in $\mathcal{F}$'s domain.*

Our investigations into *SDF*s are not yet sufficiently mature to impose this reversal symmetry axiom as part of the definition of *social dichotomy function*, or to pin down that definition in other ways (such as allowable ballot form). However, the limited number of examples we have in mind do satisfy the axiom, with ballot forms ranging over subclasses of weak orders. So, if we were forced to formulate a definition at this time, we'd say that an SDF is a rule for amalgamating ballots that come from any specified class $C$ of weak orders, and outputting a dichotomy (or several dichotomies, in the event of a tie), in such a way as to satisfy the above reversal symmetry axiom.

# 2 2.5 Examples of Social Dichotomy Functions

The *Mean Rule*, proposed by Duddy and Piggins in (2), amalgamates ballots that are dichotomous weak orders (interpretable as approval ballots), and outputs the following social dichotomy $(A_T, A_B)$: if $q$ denotes the mean approval score of all alternatives, then $A_T$ contains all alternatives achieving an approval score higher than $q$, and $A_B$ contains those with an approval score lower than $q$. (An alternative achieving an approval score of exactly $q$ may be placed in either $A_T$ or $A_B$, resulting in a tie among all dichotomies that can be obtained by making those choices freely and independently for all such alternatives.)

The initial attraction of the Mean Rule is an axiomatization parallel to that given by Young (5) for the Borda rule. The Mean Rule thus suggests a possible opening to the extension of the notion of scoring rule from the context of preference aggregation to the more general domain of judgment aggregation.[1]

More recently, we noticed that the Mean Rule can be characterized alternatively in terms of minimizing a form of total "tension," as summed over all *internal* pairs – pairs of alternatives for which both come from $A_T$ or both come from $A_B$. Equivalently, the dichotomy produced by the mean rule maximizes the total tension "cut" by the separation of $A$ into $A_T$ and $A_B$ (summed over all pairs that cross from $A_T$ to $A_B$).

It turns out that to measure the tension in a pair of alternatives one does not need the ballots in a profile to be dichotomous orders. The *Borda Mean Rule* amalgamates linear order ballots from a profile $P$ into the dichotomy $(A_T, A_B)$ determined by the mean Borda score $q^\beta$ of all alternatives (in the same way that the original Mean Rule utilizes the mean approval score). The Mean Rule and the Borda Mean Rule minimize the same form of tension, and both satisfy the ballot reversal symmetry axiom of Definition 1.

In fact, provided that one uses an appropriate version of Borda score, a Generalized Borda Mean Rule may be used to amalgamate arbitrary weak order ballots into a social dichotomy in the same way, and the Mean Rule then emerges as a special case of the Generalized Borda Mean Rule. We are considering this Generalized version to constitute the "0.5" of the 2.5 social dichotomy functions, discounted because we have not yet explored its behavior.

# 3 Are Approval Voting and the Borda Mean Rule the Same?

We show that the Borda Mean Rule (defined in the previous section) can also be formulated as a scoring rule via a certain scoring function $F$, applied to each ballot $b$ (a linear ordering of $A$) to award $F(b, d)$ points to each dichotomy $d$ of $A$.[2] For each dichotomy $d$, one sums the points awarded to $d$ by all ballots cast, and the dichotomy with the greatest sum coincides with the Borda Mean Dichotomy.

Now suppose we turn this around, by considering ballots that are dichotomies of $A$, and potential election outcomes that are linear orderings of $A$. We may now interpret $F(b, d)$ (using the same $F$ as above) as determining the points awarded by a dichotomous ballot $d$ to each possible linear order $b$. For each ordering $b$, one sums the points awarded to $b$ by all ballots cast, and the ordering with the greatest sum is considered the winner. What rule is this?

We show that the winning ranking coincides with that induced by approval score (when we interpret the ballots as approval ballots and calculate approval score in the conventional way),[3] so we will refer to this voting rule as *Approval Voting*$_2$. The idea of awarding points

[1] This possibility has also been recently explored by Dietrich (1).

[2] Specifically, the value of $F(b, d)$ is the number of ordered pairs $(x, y)$ of alternatives that are ordered strictly by both $b$ and $d$ in the same way ($x >_b y$ and $x >_d y$, or $y >_b x$ and $y >_d x$), minus the number that are ordered strictly by both in the opposite way ($x >_b y$ and $y >_d x$, or $y >_b x$ and $x >_d y$).

[3] Of course, the ordering induced by approval score can be a weak ordering $\geq$, with ties arising from

directly to rankings, rather than to individual alternatives, may seem a bit unusual. But it is important to recall that the Kemeny rule is a scoring rule in precisely this sense. Moreover, as shown independently in (6) and (3), every conventional scoring rule (of the kind that awards points to individual alternatives) is equivalent to a "ranking scoring rule" that awards points directly to rankings.

Thus Approval $\text{Voting}_2$ can be seen as a straightforward adaptation of the "ranking scoring rule" idea to the context of approval ballots. It is surprising, though, that the Borda Mean Rule and Approval $\text{Voting}_2$ employ the same scoring function $F$, albeit in reverse order. Is something more going on here? As we show in (7), there exists an extension $F^H(R_1, R_2)$ of $F(b, d)$ that

- applies to an arbitrary pair $(R_1, R_2)$ of binary relations on $A$
- is symmetric, satisfying $F^H(R_1, R_2) = F^H(R_2, R_1)$
- yields a variety of familiar rules (including plurality and Kemeny, as well as approval and Generalized Borda Mean, but *not* Borda) by restricting $R_1$ to some class $\mathcal{C}_1$ of allowable ballots and $R_2$ to some class $\mathcal{C}_2$ of election outcomes.

It was the relationship between the Borda Mean Rule and approval voting that first hinted at the existence of $F^H$. Xia and Conitzer (4) consider amalgamations of partial order relations that seem closely related to the $F^H$ method.

# 4 What is a Dichotomy Separation Problem?

The techniques used to characterize the Mean Rule, and the Borda Mean Rule, are not limited to the social choice setting, wherein some form of ballot amalgamation is taking place. It seems that there are more abstract contexts in which the question of how best to divide a set $Z$ into a *Top* and *Bottom* makes sense. In each of the settings we consider, there exists some measure $w$ of the separation between two points of a set $Z$. Speaking loosely, the problem is to find the dichotomy $\mathcal{Z} = (Z_T, Z_B)$ that maximizes

$$TS(\mathcal{Z}) = \Sigma_{z_t \in Z_T,\, z_b \in Z_B} w(z_t, z_b),$$

interpreted as a measure of the "total separation" (or "total tension") between $Z_T$ and $Z_B$.

We consider two contexts, of which the first is Euclidean: $Z \subset \Re^k$, with $k = 1$ an important special case. In the second, points arise as the vertices of some complete digraph $\mathcal{G} = (Z, E)$, with $w$ interpreted as an assignment of weights to the directed edges $(x, y) \in E$ of $\mathcal{G}$. These contexts are not entirely distinct – one might, for example, embed $\mathcal{G}$'s vertices $Z$ into $\Re^k$, and use that embedding to induce the edge weights. A further split appears in each of these two contexts according to whether $w$ is a *directed* or *undirected* measure of separation between points. In the Euclidean context, distance between the two points serves as our undirected measure, and vector displacement to the first point from the second

multiple alternatives earning equal approval scores. In this case, all linear orders that extend $\geq$ (by breaking ties within the $I$-classesin all ways possible) will be tied, earning the same greatest sum.

as the directed measure. The corresponding split for graphs arises in the type of symmetry we assume for the function $w$ assigning a weight $w(x, y)$ to each directed edge from a vertex $x$ to a distinct vertex $y \neq x$; we take antisymmetry (i.e., $w(x, y) = -w(y, x)$) as our directed notion and symmetry (i.e., $w(x, y) = w(y, x)$) as the undirected notion. In the antisymmetric case we can discard one of the two edges from each pair of oppositely oriented directed edges, obtaining a weighted *tournament* (which contains the same information as the original weighted digraph); in the symmetric case each such pair of edges could be replaced with a single, undirected edge bearing their common weight. In the general case, of course, no relationship between $w(x, y)$ and $w(y, x)$ is assumed.

For the general case in the graph setting, maximizing $TS(\mathcal{Z})$ is equivalent to solving the *Graph Cut Problem*, which is known to be NP-complete; no simple, analytical solution is likely to exist. The same is true for the symmetric case. However, in the antisymmetric case (which aligns more closely with the goal of the Social Dichotomy Functions considered here) a simple analytical solution does exist, and we show it coincides with the Borda Mean Rule. This relies on the fact that the conventional Borda score can be derived from the weighted tournament induced by a profile, as given by the antisymmetric edge-weighting $w(x, y) =$ [the number of voters who rank x over y] - [the number of voters who rank y over x].

# References


[1] Dietrich, F. Scoring rules for judgment aggregation, forthcoming in Social Choice and Welfare (2013).

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

[3] Conitzer, V., Rognlie, M., and Xia, L. Preference Functions That Score Rankings and Maximum Likelihood Estimation, Proceedings of the Twenty-First International Joint Conference on Artificial Intelligence (IJCAI-09), 109-115 (2009).

[4] 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).

[5] Young,H.P. An axiomatization of Borda's rule, Journal of Economic Theory 9, 43-52 (1974).

[6] Zwicker, W.S.. Consistency Without Neutrality in Voting Rules: When is a Vote an Average?, Mathematical and Computer Modelling 48, special issue on Mathematical Modeling of Voting Systems and Elections: Theory and Applications, A. Belenky, ed., 1357-1373 (2008).

[7] Zwicker, W.S. Universal and Symmetric Scoring Rules for Binary Relations. Preprint (2014).