How (not) to metrically rationalize social choice: Distance-based aggregation between impossibility and anything goes results

Daniel Eckert
Karl-Franzens-Universität Graz

July 2012

1 Introduction

Distances play an important role both in belief revision ([1]) resp. belief merging ([6]) and in social choice theory resp. the recent literature on judgment aggregation (for a survey see [8]). Thus, a major formal relationship between these areas consists in the use of distances as objective functions to minimize, which is particularly well established in social choice theory and its recent extensions (see [10], [7]).

In social choice theory, aggregation is the assignment to any list (profile) of individual characteristics (typically, preferences) of a characteristic that is "representative" for the group, typically in view of a collective decision (e.g. the choice of a socially "best" alternative).

Formally, an aggregation rule is a mapping $f : X^n \to Y$ where the domain is a product space of profiles of individual characteristics and the codomain is a space of social characteristics (typically $X = Y$, e.g. in case of an Arrovian social welfare function).

Well known problems of aggregation rules are the classical impossibility results of Arrow and Gibbard-Satterthwaite, roughly stating that every otherwise satisfactory aggregation rule is the dictatorship of a particular individual, resp. that every non-dictatorial social choice rule is manipulable.

These negative results are largely recovered by the recent extensions of Arrovian social choice theory in abstract aggregation theory, judgment aggregation, and computational social choice (for the latter see [11]).

A major, though by far not the only justification for the use of distances is derived from the fact that, typically, for subsets of the profiles in the domain of an aggregation rule the assignment of a collective outcome is uncontroversial (e.g. for unanimous profiles).

This raises the question whether this consensus can be extended to the whole domain by assigning to any profile the outcome of the consensual profile which is closest to it. This intuition was systematically explored in the research program of metric rationalization of social choice ([3]), which essentially consists in the rationalization of a social choice rule by the optimization of a distance-based objective function (typically, the minimization of a distance function). In this way, characterization results for aggregation rules were obtained, e.g. the characterization of the Borda winner by its closeness to being a unanimous winner (for a survey see [9]).

This use of distances suggests the formulation of consistency conditions in terms of distances, but two antagonistic types of problems can emerge with distance-based consistency conditions:

1) "anything goes" results ([5]), stating that (almost) any voting rule can be metrically rationalized by some distance, and

2) impossibility results, e.g. Baigent's ([2]) impossibility of proximity preservation.

2 Formal framework and results

Let Ω denote a set of **possible worlds** (which can be interpreted as the set of all possible complete descriptions of the state of the world by an individual, e.g. linear preference orderings of a set of alternatives). Then, $\Omega/\equiv$ denotes the partition of Ω according to the equivalence relation $\equiv \subset \Omega \times \Omega$ (with the interpretation that the partition of Ω into **equivalence classes** corresponds to the aspects which are relevant for the collective decision, e.g. the top rank of an alternative). Observe that this partition is allowed to be the finest possible (e.g. in the case of a social welfare function).

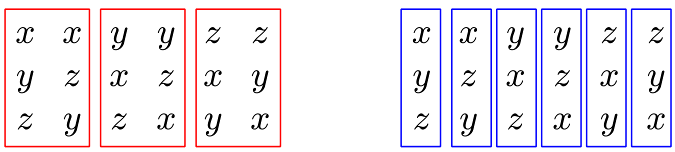

Figure: 1a) "Top rank" partition   1b) Finest partition

For a given partition $\Omega/\equiv$ of Ω into equivalence classes, let $\mathcal{U} = \{\bigcup_{S\in P} |P \in \mathcal{P}(\Omega/\equiv)\backslash\varnothing\}$ denote the set of all **unions of equivalence classes** (the larger any such union of equivance classes, the less valuable is the information for the collective decision).

As usual, N will denote a set of **individuals** such that Ω^N (denoted by Π) is set of **profiles** $\pi = (\pi_1, \pi_2, ..., \pi_{|N|})$ of possible worlds.

Definition 1 *An* ***aggregation rule*** *is a mapping* $F : \Pi \to \mathcal{U}$ *which assigns to each profile* $\pi \in \Pi$ *an equivalence class or a union of equivalence classes* $F(\pi) \in \mathcal{U}$.

Following Elkind, Faliszweski and Slinko [5], a consensus is formalized in the following way:

Definition 2 *A* ***consensus class*** *is a pair* $\mathcal{C} = (C, f)$ *where* $C \subseteq \Pi$ *is the set of* ***consensual profiles*** *and the* ***consensus mapping*** $f : C \to \Omega/\equiv$ *is an onto mapping which assigns to every consensual profile* $\pi \in C$ *an equivalence class in* $\Omega/\equiv$.

Example 3 ***Unanimity*** *is the consensus class* $\mathcal{C} = (C, f)$ *where* $C = \{\pi \in \Pi|(\forall i, j \in N)\pi_i = \pi_j\}$ *is the set of unanimous profiles and* $f : C \to \Omega/\equiv$ *assigns to every unanimous profile of possible worlds the equivalence class of this possible world.*

Consistency with a consensus can now be defined in a natural way.

Definition 4 *An aggregation rule* $F : \Pi \to \mathcal{U}$ *is* ***consistent with a consensus class*** $\mathcal{C}_F = (C, f)$ *if for all profiles* $\pi \in C$, $F(\pi) = f(\pi)$.

Metric rationalization is then a way to extend the consensus mapping to the whole aggregation rule with the help of distances.

Definition 5 *A function* $d : \Pi\times\Pi \to R_+$ *is a* ***distance*** *if for any* $\pi, \pi', \pi'' \in \Pi$
(i) $d(\pi, \pi') = 0$ *if and only if* $\pi = \pi'$ *(identity of indiscernibles)*
(ii) $d(\pi, \pi') = d(\pi', \pi)$ *(symmetry)*
(iii) $d(\pi, \pi') \leq d(\pi, \pi'') + d(\pi'', \pi')$ *(triangle inequality).*

Generalizing Elkind, Faliszweski and Slinko's [5] concept of distance rationalizability, we obtain the following definition:

Definition 6 *An aggregation rule* $F : \Pi \to \mathcal{U}$ *is* ***distance rationalizable*** *via a consensus class* $\mathcal{C}_F = (C, f)$ *and a distance* $d_F : \Pi \times \Pi \to R_+$ *(is* $(\mathcal{C}_F, d_F)$*-rationalizable) if for any profile* $\pi \in \Pi$

$$F(\pi) = \bigcup \left\{ [\omega] \in \Omega/\equiv |[\omega] = f\left(\min_{\pi' \in C} d(\pi, \pi')\right)\right\},$$

i.e. if the outcome is the union of all equivalence classes associated by the consensus mapping f *with the distance minimizing consensual profile(s).*

Unfortunately, distance rationalizability with respect to a consensus does not sufficiently restrict the space of the many aggregation rules that are consistent with it, as the following "anything goes" result shows.

Theorem 7 *For any aggregation rule* $F : \Pi \to \mathcal{U}$ *which is consistent with a consensus class* $\mathcal{C}_F = (C, f)$ *there exists a distance* $d_F : \Pi \times \Pi \to \mathsf{R}_+$ *such that* F *is* $(\mathcal{C}_F, d_F)$*-rationalizable.*

Proof. For the proof of this theorem consider the undirected graph $G = (\Pi, E)$ defined, for any distinct $\pi, \pi' \in \Pi$ by

$$\{\pi, \pi'\} \in E \text{ whenever } F(\pi) \in \Omega/\equiv \text{ and } F(\pi) \subseteq F(\pi')$$

and consider the shortest path distance $d_F : \Pi \times \Pi \to \mathsf{R}_+$ that it induces.
First, assume that $\pi \in C$. Then $\{\pi' \in \Pi | d_F(\pi, \pi') = 0\} = \pi$ (by the identity of indiscernables) and $F(\pi) = f(\pi) \in \Omega/\equiv$ (by consistency of F and f). If $\pi \notin C$, then $d_F(\pi, \pi') \geq 1$ for any profile $\pi' \in \Pi$, and for any profile $\pi' \in C$ we have $d_F(\pi, \pi') = 1$ if $F(\pi') \subseteq F(\pi)$. Finally, for any equivalence class $[\omega] \nsubseteq F(\pi)$ and any profile $\pi' \in C$ such that $f(\pi') = [\omega]$ we have $d_F(\pi, \pi') \geq 2$.
Thus $F(\pi) = \bigcup \left\{ [\omega] \in \Omega/\equiv |[\omega] = f\left(\min_{\pi' \in C} d(\pi, \pi')\right)\right\}$, i.e. the union of all equivalence classes associated by the consensus mapping f with the distance minimizing consensual profile(s), as desired. ■

Obviously, this anything goes result is driven by the non-neutrality of the distance with respect to the consensus class (or, formally: the graph from which the distance is derived is given by the neighborhoods of the consensual profiles).

But is it reasonable to require consistency with respect to an "objective" distance?

A generalization of an older result ([2]) related to metric rationalization which is seen as a discrete analogue to the topological version of Arrow's theorem ([4]) suggests that this requirement might be too strong.

Definition 8 *Let* $d : \Pi \times \Pi \to R_+$ *and* $\delta : \mathcal{U} \times \mathcal{U} \to R_+$ *be a distance on the domain, respectively on the codomain, of the aggregation rule* $F : \Pi \to \mathcal{U}$*.*
F *satisfies* ***proximity preservation*** *if for any profiles* $\pi, \pi', \pi'' \in \Pi$

$$d(\pi, \pi') < d(\pi, \pi'') \Rightarrow \delta(F(\pi), F(\pi')) \leq \delta(F(\pi), F(\pi'')).$$

Unfortunately, the imposition of this property makes any reasonable aggregation rule incompatible with any reasonable neutral distance on profiles as the inconsistency of the following properties suggests.

Definition 9 *An aggregation rule* $F : \Pi \to \mathcal{U}$ *satisfies* ***minimal compensation*** *if it is not the case that for all pairs of i-variants* $\pi = (\pi_1, \pi_2, .., \pi_i, ..., \pi_{|N|})$ *and* $\pi' = (\pi_1, \pi_2, ..., \pi'_i, ..., \pi_{|N|})$ *such that* $F(\pi) \neq F(\pi')$ *there does not exist a profile* $\pi'' = (\pi''_1, \pi''_2, ..., \pi'_i, ..., \pi''_{|N|})$ *such that* $F(\pi'') = F(\pi)$*, i.e. there does not exist a change from* π' *to* π'' *that compensates the change from* π *to* π' *while keeping the pivotal characteristic* π'_i*.*

Observe that minimal compensation is a much weaker condition than anonymity and consider it in combination with the natural property of monotonicity.

Definition 10 *A distance* $d : \Pi \times \Pi \to R_+$ *over profiles is* ***monotonic*** *if for all profiles* $\pi, \pi', \pi'' \in \Pi$
$d(\pi, \pi') < d(\pi, \pi'')$ *whenever* π'' *differs from* π *in more components than* π'*.*

However, the following theorem shows that these two properties are inconsistent.

Theorem 11 *There does not exist an aggregation rule* $F : \Pi \to \mathcal{U}$ *which satisfies minimal compensation and proximity preservation with respect to a monotonic distance on profiles.*

Proof. By **minimal compensation** there exists a pair of i-variants $\pi = (\pi_1, \pi_2, .., \pi_i, ..., \pi_{|N|})$ and $\pi' = (\pi_1, \pi_2, ..., \pi'_i, ..., \pi_{|N|})$ such that $F(\pi) \neq F(\pi')$ and a profile $\pi'' = (\pi''_1, \pi''_2, ..., \pi'_i, ..., \pi''_{|N|})$ such that $F(\pi'') = F(\pi)$. Hence, by monotonicity, $d(\pi, \pi') < d(\pi, \pi'')$, but, by construction, $\delta(F(\pi), F(\pi'')) = 0 < \delta(F(\pi), F(\pi')$, which violates proximity preservation. ■

In conclusion, distance-based approaches are useful for the characterization of existing aggregation rules and eventually for the design of new ones, but metric rationalization in general does not guarantee more "rationality" in aggregation as long as the corresponding consistency conditions can be seen as either too weak or too strong.

References

[1] Alchourron, C.E., Gärdenfors, P., Makinson, D. (1985), On the logic of theory change: patial meet contraction and revision functions, Journal of Symbolic Logic 50: 510-530.

[2] Baigent, N. (1987a), Preference proximity and anonymous social choice, Quarterly Journal of Economics 102: 161-169.

[3] Baigent, N. (1987b), Metric rationalisation of social choice functions according to principles of social choice, Mathematical Social Sciences 13: 59-65.

[4] Chichilnisky, G. (1982), Social aggregation rules and continuity, Quarterly Journal of Economics 96: 337-352.

[5] Elkind, E., Faliszewski, P., Slinko A. (2010), On the Role of Distances in Defining Voting Rules, Proc. 9th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2010).

[6] Gregoire, E., S. Konieczny (2006), Logic-based approaches to information fusion, Information Fusion 7: 4-18.

[7] Lang, J., Pigozzi, G., Slavkovik, M., van der Torre, L. (2011), Judgment aggregation rules based on minimisation, Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge (TARK-2011), Groningen, The Netherlands, July 12-14, ACM, 238-246.

[8] List, C., Puppe, C. (2009), Judgment aggregation, in: Anand, P., Pattanaik, P.K., Puppe, C. eds., Rational and Social Choice: An Overview of New Foundations and Applications. Oxford University Press, Chapter 19.

[9] Meskanen, T., Nurmi, H. (2008), Closeness counts in social choice, in: Braham, M, Steffen, F. eds., Power, Freedom and Voting. Springer.

[10] Miller, M.K., Oshershon, D. (2009), Methods for distance-based judgment aggregation, Social Choice and Welfare 32:575-601.

[11] Rothe, J, Baumeister, D., Lindner, C., Rothe, I. (2011), Einführung in Computational Social Choice, Spektrum: Heidelberg.