Social choice using moral metrics

Kenneth Halpern

Independent, Cambridge, Massachusetts, USA

Abstract

Purpose – This paper aims to develop a geometry of moral systems. Existing social choice mechanisms predominantly employ simple structures, such as rankings. A mathematical metric among moral systems allows us to represent complex sets of views in a multidimensional geometry. Such a metric can serve to diagnose structural issues, test existing mechanisms of social choice or engender new mechanisms. It also may be used to replace active social choice mechanisms with information-based passive ones, shifting the operational burden.

Design/methodology/approach – Under reasonable assumptions, moral systems correspond to computational black boxes, which can be represented by conditional probability distributions of responses to situations. In the presence of a probability distribution over situations and a metric among responses, codifying our intuition, we can derive a sensible metric among moral systems.

Findings – Within the developed framework, the author offers a set of well-behaved candidate metrics that may be employed in real applications. The author also proposes a variety of practical applications to social choice, both diagnostic and generative.

Originality/value – The proffered framework, derived metrics and proposed applications to social choice represent a new paradigm and offer potential improvements and alternatives to existing social choice mechanisms. They also can serve as the staging point for research in a number of directions.

Keywords Social choice, Behavioral economics, Voting theory, Mathematical morality, Mathematical philosophy, Moral geometry

Paper type Research paper

1. Introduction

An ideal social choice mechanism is both fair and perceived as fair. Arrow famously demonstrated that it is impossible to accommodate even three basic tenets of fairness in ranked preference systems (Arrow, 1950), and similar results hold for other systems. Even if anomalies are unavoidable, we can seek to reduce unfairness by minimizing their prevalence and severity. Of equal importance, we can seek mechanisms that reduce the perception of unfairness.

We offer a tool to address both parts of the equation. By inferring the moral systems of individuals and constructing a suitable distance function between them, it is possible to construct a moral geometry with attendant notions of proximity, neighborhoods and clustering.

A metric is a multidimensional structure and far more versatile than the linear orders often employed for social choice. It opens the door to a variety of new approaches, but also can beget new linear orders for use with existing social choice mechanisms. A metric can embody the relationship between entire sets of views, and the very use of a precisely quantified moral geometry may help foster a sense of inclusion and fairness.

We begin by reifying the notion of a "moral system" (MS), equating it with a computational black box that issues responses when presented with situations. Under reasonable inference assumptions, such a black box can be represented by an asymptotic conditional probability distribution (CPD). We subsequently also consider inferred or estimated CPDs as representatives.

After formally defining our assumptions, we turn to the question of metric construction. For a metric among MSs to be useful, it must reflect our intuition in some fashion. Direct assertion of such a metric is untenable, and it must inherit meaning from simpler structures through which we plausibly can codify our intuition.

The natural semantic objects are situations and responses. In a given problem, we understand these and can characterize them in a sensible fashion. We argue that the appropriate *a priori* structures are a probability distribution (PD) over situations and a metric among responses. It is from our intuition for these that the metric among MSs must derive its behavior and meaning.

After discussing the specification of our *a priori* structures, that of the MSs themselves, and a few related issues, we propose several applications of this framework to social choice. We next introduce a number of related concepts, corresponding to notions of hypocrisy, judgment, worldview and moral trajectory, and consider some additional social choice applications involving these.

We also present several concrete, well-behaved derived metrics among CPDs and conclude with a discussion of the use of Euclidean embeddings for selection and specification of the *a priori* metric among responses.

We will refer to both metrics and pseudometrics as "metrics," only drawing the distinction for emphasis or when necessary. Recall that a metric is a nonnegative function $d: X \times X \to X$ such that (1) $d(x, y) = d(y, x)$, (2) $d(x, y) = 0$ iff $x = y$ and (3) $d(x, z) \leq d(x, y) + d(y, z)$. A pseudometric relaxes this to allow $d(x, y) = 0$ when $x \neq y$. For most purposes, the distinction is immaterial. Although metric-derivation procedures almost invariably produce pseudometrics among CPDs, these usually restrict to metrics among MSs.

We will not delve into questions of empirical measurement or experimental construction. Observation most likely would entail case histories, surveys or interviews, carefully curated and with due regard for unreliability.

Note that our use of the term "metric" is topological, and we speak of "geometry" in reference to distances, neighborhoods and clusters. This should not be confused with Riemannian metrics or notions of curvature.

2. Framework

2.1 Central premise

A "moral system" (MS) embodies how an individual, institution or group responds to situations. The central dogma of our approach is that MSs correspond to computational black boxes, which, under reasonable inference assumptions, have CPDs as mathematical proxies. Through these means, the otherwise ill-defined problem of constructing a metric among MSs becomes mathematically well defined.

2.2 Situations and responses

A "situation" is a stimulus provided to a subject, and a "response" is a reaction of a subject to a situation. When working with surveys, questions would be situations, and answers would be responses. When working with judicial sentencing, cases could be situations, and sentences could be responses.

We will denote by S the set of all possible situations, and by R a set containing all possible responses. An MS generates a response in R to any given situation in S. Situations and responses have meaning and are the primary sources of semantics in a problem.

S is not a theoretical universe of situations. It is finite and chosen to capture the behavioral aspects we care about. There is no concept of a basis that spans behaviors (our spaces are not linear), but S sometimes can serve in a similar capacity.

The set of "accessible responses" R_A consists of every response that can arise from the MS under consideration with nonzero probability for some $s \in S$. It is not known to us *a priori*, though we can attempt to infer it.

We require that $R_A \subseteq R$ is finite, though R need not be. It is not unreasonable to assume *a priori* knowledge of R, and that $R_A \subseteq R$, without knowing *which* subset it is. Consequently, we can expand R as needed to admit simple parametrization or other convenient properties.

2.3 Moral systems as black boxes

The only way to probe an MS is through its responses to situations. From our perspective, it is an opaque machine for determining responses to situations.

We do not assume that an MS is deterministic. A person may not always respond the same way to a given situation, perhaps reflecting a true stochastic element, imperfect information or stateful evolution. Because the decision-making process is hidden from us, we cannot attribute apparent randomness to any specific source.

We will refer to a "sample" as a single observed response by a given MS to a specific situation. We assume that MSs act independently of one another, each MS only responds and evolves based on the sequence of situations it encounters, and we are not privy to its initial state. There is no notion of synchronous sampling, and we cannot meaningfully compare the responses of two MSs to a given situation in a single trial. We only can compare statistical behaviors.

We have no notion of time or computational complexity or computability. As machines, MSs are assumed to always halt and to operate in constant time. We only consider discrete sequences of samples.

2.4 Inference assumption

Without inference assumptions, we can say nothing useful, even in the presence of unlimited data. We adopt a form of ergodicity.

Given an MS and some PD $P(S)$ s.t. $P(s) > 0 \; \forall s \in S$, we assume that (1) regardless of the unknown initial state, the histogram of any sequence of samples with situations drawn from $P(S)$ will asymptotically converge to a unique $P(R|S)$ for the MS, and (2) $P(R|S)$ encompasses everything relevant to us about the MS's behavior. We will refer to it as the "true CPD" of the MS.

This says nothing about the rate of convergence, and we implicitly also assume (3) we have (or can produce) adequate sample data for satisfactory inference in the given application. When studying moral trajectories in Section 4.4, we will weaken these assumptions to allow adiabatic variation of the CPD.

Under our inference assumption, $P(R|S)$ is the natural mathematical proxy for an MS. We will denote by $C_{S,R}$ the space of all such CPDs, infinite even for finite S and R. We will denote by X the specific set of MSs under study. This need not be fixed *a priori*, and may expand. For example, new individuals could be surveyed.

Note that a CPD obtained from finite sample data is not the true CPD of an MS, and we will term it an "inferred CPD." For reasons to be discussed, we often confine ourselves to a model subspace $B_{S,R} \subset C_{S,R}$. Rather than the true CPD or an inferred CPD, we estimate an element of $B_{S,R}$. An estimated CPD obtained with unlimited data will be termed the "asymptotic

estimate" for the MS, while any finite data estimate will be termed an "inferred estimate," also in $B_{S,R}$.

2.5 A priori structures

It is on X that we seek to derive a metric. We do so by first building a metric D_M on $C_{S,R}$, then restricting it to the model subspace $B_{S,R} \subset C_{S,R}$ and finally pulling it back along the indexing map $X \rightarrow B_{S,R}$, which associates each MS with its inferred estimate. This is just a fancy way of saying the metric on $C_{S,R}$ induces one on X in the obvious manner.

For convenience, we will use D_M interchangeably for the derived metric on $X, B_{S,R}$, or $C_{S,R}$. For example, $D_M(x, x')$ on X implicitly means $D_M(P_x, P_{x'})$ on $C_{S,R}$ or $B_{S,R}$, where P_x denotes whichever CPD we associate with x. Because X is discrete, we will sometimes write $D_{ij} \equiv D_M(x_i, x_j)$.

Note that we are not simply trying to find *some* metric on $C_{S,R}$. That could be accomplished using the Fisher–Rao metric (Rao, 1945) or a variety of other approaches, but the resulting geometry would be uninformative. The choice of D_M determines the utility of our framework, and it must embody the behavioral aspects we care about.

Directly positing a metric among MSs or CPDs is very difficult. These are complicated objects, and we generally have no intuition for distances between them. We need something simpler and more intuitive. R and S are endowed with semantics, and it is to these we must turn. The sets themselves do not suffice, and we require some sort of structures on them.

Our approach is to require a PD $P(S)$ over S and a metric (or pseudometric) d_R on R. It is from these structures that D_M must derive its behavior. We will now motivate these choices.

Note that we are not simply replacing the problem of defining a metric on $C_{S,R}$ with a comparably difficult one on a different space. R is much smaller than $C_{S,R}$ and is more likely to admit a simple, intuitive metric. We are building D_M from tractable components.

2.6 P(S)

Without a PD over situations, we must confine ourselves to per-situation analysis, and this is inadequate for our purposes. We require some form of aggregation over S, and summation is the natural choice. $P(S)$ provides the necessary measure. It may represent an estimated likelihood of occurrence, an importance weight or a bit of both. For certain purposes, the interpretation of results is easiest when $P(S)$ is a likelihood.

Specifying $P(S)$ usually is straightforward. For example, we could empirically measure observed frequencies of occurrence. We will assume that $P(S)$ is strictly positive on all of S (it is easy to introduce nominal support if not). Note that $P(S)$ may suppress the probability $\sum_{s \in S} P(s)P(r|s)$ of an accessible response r to near zero. Any derived quantity (such as D_M) effectively ignores such responses.

2.7 d_R

We require an *a priori* choice of metric (or pseudometric) d_R on R. There are reasons this is a natural structure to employ.

$P(S)$ only tells us how to weigh each situation when computing distances, but offers no conduit to comparison of MSs. Any meaningful distance between MSs must derive from some comparator on R. The triangle inequality is very difficult to prove from scratch, but sometimes can be inherited. We are more likely to derive a metric on $C_{S,R}$ from a metric on R than from some other structure.

Another reason concerns the type of information present. Unlike MSs, responses often have independent, objective meaning. Distances between them are something people are more likely to agree on.

Section 6 offers a means of parlaying intuition for distances on R into an actual metric. As a rule, d_R is less mutable than $P(S)$. We may change $P(S)$ to reflect new priorities or updated frequency statistics, but d_R rarely would be modified once chosen (except perhaps to test robustness to such changes).

Note that d_R is the core of D_M, and its source of semantics. It must be chosen carefully and reflect our intuition. Although responses can be codified as finite strings, "edit distances" such as the Hamming distance (Hamming, 1950) or Levenshtein distance (Levenshtein, 1966) are devoid of semantics and not suitable for our purpose.

2.8 Specification of moral systems

MSs in X must be specified in some manner, either up front or as they arise. Often, they all naturally sit between S and R, but this need not be the case. Though S, R and X each carry semantics, their generative mechanisms may not be the same. Normalization may be required.

2.8.1 Normalization. If an MS has natural input space I and output space O, we include in its definition two maps: $\alpha : S \rightarrow I$ tells us how to encode S for the MS, and $\beta : O \rightarrow R$ tells us how to decode O. They need not be injective or surjective. If no normalization is needed, $I = S$, $O = R$, $\alpha = id_S$ and $\beta = id_R$.

Requiring α and β for each MS is not pointless or pedantic. It would be impossible to compare MSs without a common meaning for inputs and outputs. α and β plug an MS into the semantics of our framework and attach this common meaning to I and O.

$P(S)$ induces an effective PD $P'(i) \equiv \sum_{x \in \alpha^{-1}(i)} P(x)$ on each I, and we can compare outputs of different MSs via $d_R(\beta_1(o_1), \beta_2(o_2))$. However, d_R is not a metric in this capacity, because o_1 and o_2 are elements of distinct sets (we have pulled back $d_R : R \times R \rightarrow \mathbb{R}$ along $\beta_1 \times \beta_2 : O_1 \times O_2 \rightarrow R \times R$ to $d_* : O_1 \times O_2 \rightarrow \mathbb{R}$, which is a pseudometric only if $O_1 = O_2$ and $\beta_1 = \beta_2$, and a metric if β_1 also is injective). Any CPD $P(O|I)$ for an unnormalized MS induces an effective CPD $\widehat{P}(R|S)$ via $\widehat{P}(r|s) \equiv \sum_{o \in \beta^{-1}(r)} P(o|\alpha(s))$.

When working with provided data, unnormalized samples may be unavoidable. The use of O rather than R is not an issue, and we just apply $\beta(o)$. However, I may pose a problem. If α is not injective we may be unable to determine which $s \in \alpha^{-1}(i)$ to adopt, and if α is not surjective, there may be no corresponding s at all. In the latter case, we could discard the sample, and in the former, we could randomly draw from a uniform distribution over $\alpha^{-1}(i)$. However, this constitutes an additional assumption.

2.8.2 Modes of access. An MS is associated with something real: a person, an institution, a decision system. We require some form of access to it, a way to acquire knowledge of its behavior. We will consider three such modes: (1) full knowledge of the true CPD, (2) a fixed set of sample data and (3) the ability to actively acquire sample data.

Rarely do we have direct access to the true CPD for an MS. It is large and can be difficult to store or work with. Instead, we generally work with samples. We will not distinguish between access modes (2) and (3). Though (3) allows efficient sampling strategies, we remain limited to a relatively small data set.

We compute a distance between two MSs by first inferring the relevant CPDs, then plugging these into D_M. Direct inference of the distance would be preferable from a statistical standpoint, avoiding the undesirable inference of large intermediates (as advised against in Vapnik, 1999). However, devising an algorithm for direct inference of distances is impractical in most cases. The use of estimation is a good compromise, reducing the size of intermediate objects while remaining conceptually simple.

2.9 Estimation

$C_{S,R}$ is very large, and attempting to infer the true CPD is inadvisable. Inference with limited sample data would lead to noisy results and huge hidden correlations. We typically model CPDs using a parametrized subspace $B_{S,R} \subset C_{S,R}$, or perhaps a discrete set of representatives. Standard dimensional reduction techniques (such as regression) can be employed to estimate a point in $B_{S,R}$ from sample data. Estimation of a few model parameters is more tenable than inference of an entire CPD.

Practical considerations must govern the choices of $B_{S,R}$ and estimation procedure. We take both as part of a problem's *a priori* structure. Any sensible procedure will be agnostic to the order in which sample data are processed.

There are reasons other than sound inference to employ estimation. The elements of $B_{S,R}$ could represent idealized or canonical MSs, or we could use $B_{S,R}$ to isolate relevant behavioral factors.

Note that there are two types of approximation at play. The estimation procedure confines consideration to $B_{S,R}$, but we also estimate with limited data. We obtain only an inferred estimate in $B_{S,R}$, approximating the asymptotic estimate.

Statistical learning theory has much to say about the bounds of plausible inference (see Mitchell, 1997; Vapnik, 1999; Mohri, 2018), and we will not digress into such matters here.

2.10 Aggregation

It sometimes is useful to combine individual MSs into larger ones, either because the aggregates are of direct interest or to improve our statistics. There are two ways to accomplish this.

We could treat a set of MSs as a single MS and collate the underlying samples into a single sequence. For example, surveys from everyone in a town could contribute to a single town-wide aggregate. This is the cleanest approach, but rather inflexible. Even a simple weighting of the underlying MSs is difficult to efficiently implement.

Another approach is to aggregate the CPDs representing underlying MSs, and we can do so in many ways. This type of aggregation is more expensive, because inference/estimation must be performed on each underlying MS. However, it has advantages as well. Once those underlying calculations have been performed, there is little cost to adopting or modifying an aggregation scheme. For a model, we may directly aggregate parameters rather than CPDs.

2.11 Units and scaling

It may be tempting to think of distances as taking units, much as Euclidean distances do. However, this need not be the case for a general metric.

For units to make sense, a distance of fixed numeric value must have the same meaning everywhere. A mile in Michigan is the same as a mile in Florida. This amounts to translation invariance, which derives from a linear structure. $C_{S,R}$ is not a vector space, and R need not be. Translation invariance on such spaces must be inherited, and this is accomplished through isometric embedding. If the metric on $C_{S,R}$ or R has a Euclidean embedding, we may define units on that space. These units only make sense in the specific embedding coordinates (or those related by Euclidean isometries), which may not be intuitive or natural for us.

We also may wish to consider the relationship between D_M and d_R. If d_R is a metric, so is $c \cdot d_R$ for any $c > 0$. Ratios of distances will be unchanged (though if d_R is not translation invariant, a given numeric ratio value will not have the same meaning everywhere).

D_M depends on d_R via some derivation procedure, and we can ask whether it scales with d_R. To do so, D_M must be homogeneous to some fixed degree in d_R. This need not be the case, but often is in practice. The D_M candidates presented in Section 5 all are homogeneous in d_R to degree 1. In that two-step procedure, the metric D among PDs over R is homogeneous in d_R,

and the metric D_M is homogeneous in D. Many common operations such as integration preserve homogeneity.

When d_R and D_M both take units and D_M is homogeneous in d_R of degree n, units of $[L]$ for d_R induce units of $[L]^n$ for D_M. If both take units but D_M is *not* homogeneous in d_R, their units are unrelated. We must be cautious interpreting results in that case. The use of unrelated units can be quite counterintuitive, and a change of scale for d_R could affect comparisons, ratios or induced linear orders on D_M.

Our framework is of greatest utility when d_R and D_M both admit units and D_M is homogeneous in d_R, and we will assume this going forward. This constrains the admissible methods of deriving D_M from $P(S)$ and d_R.

In Section 6, we will discuss the use of Euclidean embeddings for visualization and metric construction. The present requirement that d_R and D_M take units is different. All we need are isometric embeddings in some metric vector spaces. We do not require Euclidean embeddings or low-dimensional ones, though these may yield more intuitive coordinates. When an exact isometric embedding of D_M is not possible, an approximate one may suffice. In that case, the approximate D_M is translation invariant and should be used for calculation. We speak here only of embedding for units, not visualization. The latter is just a nicety and does not affect calculation with D_M.

Note that we only require an embedding of D_M on X, not of D_M on all of $C_{S,R}$ or $B_{S,R}$. Nonetheless, an embedding of $B_{S,R}$ is preferable when possible. A single element added to X could drastically alter its embedding, but would not affect that of $B_{S,R}$.

2.12 Proximity, neighborhoods and clusters

D_M endows X with meaningful notions of proximity and neighborhood. An ϵ-ball (or ϵ-neighborhood) of $x \in X$ is $\{y \in X | D_M(x, y) < \epsilon\}$, and these form the basis for a nontrivial topology on X.

The notion of neighborhood brings a wide array of mathematical tools. We have a geometry of MSs, and it sometimes may be visualized using an approximate low-dimensional Euclidean embedding (Section 6).

We also may identify clusters, sets of MSs whose intra-cluster distances are small relative to inter-cluster ones. For example, we could use a cutoff ratio $r \in (0, 1)$ and say that $\{x_1 \ldots x_n\}$ form a cluster iff $D_M(x_i, x_j)/D_M(x_i, y) < r$ for all x_i, x_j in the putative cluster and all y outside it. r is dimensionless, and clusters embody a notion of nearness independent of units. Not all metric spaces exhibit useful (or any) clustering.

The presence of a cluster of MSs does not imply its members form a group in any non-statistical sense. They may be unaware of one another, unaffiliated or comprise many different social groups. In fact, statistical clusters may prove entirely incongruous with preconceived notions of political or ideological alignment.

Because X is finite, distances on it are bounded, and we can derive a number of natural length scales. Any of these may be chosen as the unit, or explicitly serve as the divisor in a dimensionless ratio. Examples are the mean and maximum distances between distinct points: $D_{avg} \equiv \frac{1}{|X|(|X|-1)} \sum_{i \in S} \sum_{j \in S} D_{ij}$ and $D_{max} \equiv \max_{i,j \in S} D_{ij}$. Note that $C_{S,R}$ and $B_{S,R}$ need not be compact, and we generally cannot do something analogous on them.

2.13 Participants and indexing

Most systems we care about have a notion of "participants": individuals, judges, institutions, etc. We will denote the set of these Y, and it may grow if X does.

In the simplest case, each $y \in Y$ is associated with a single MS via a labeling map $Y \to X$. However, it sometimes makes sense to assign multiple labeled MSs to each participant.

We will denote the labeling set J and the labeling map g: $Y \times J \to X$. We always assume g is bijective.

For example, suppose we have two surveys per person, the first asking what they believe and the second asking what they think other people believe. Then, $J = \{self, other\}$, and g would identify the "self" and "other" surveys for every person. This information must be available to us, perhaps as part of the survey label. The map g (and any procedure for adapting it if J or X expand) is part of the specification of a problem.

The notions of participants and indexing will find use when we define hypocrisies and related concepts in Section 4.

3. Social choice

Deferring the question of how to derive D_M, let us consider how our framework could apply to questions of social choice. In this and Section 4, we will not concern ourselves with which CPDs are used to represent MSs. The discussion applies equally well to true CPDs, inferred CPDs, asymptotic estimates or inferred estimates.

There are two primary modes of application to social choice: (1) D_M serves as a diagnostic tool for existing social choice mechanisms, ascertaining whether broad acceptance is attainable, the degree to which compromise is possible, which groups likely would be alienated and the anticipated extent of disaffection; and (2) D_M spawns new social choice mechanisms, for use *in lieu* of or conjunction with existing ones. We offer a few examples below, and there are myriad others.

Our examples are illustrative but simplistic, and any real application must address nontrivial questions of measurement and feasibility. Let us suppose there is a representative set of major political issues, and that S, R, $P(S)$ and d_R have been chosen to sensibly model individuals' views on these issues, perhaps through surveys or interviews.

An MS embodies this behavioral information in some fashion, as does the CPD that represents it. We will assume the true and perceived social choice mechanism are the same and fully visible to all participants. Without yet designating what constitutes a social choice in this context, we will use the terms "approval" to refer to an individual's degree of happiness with a particular outcome and "acceptance" to refer to an individual's perception of the fairness of the mechanism by which it was reached.

3.1 For social choice diagnostics

A geometry of MSs can offer a variety of insights. The clustering or diffuseness of points can signal whether broad approval is possible through any social choice mechanism.

If MSs are arranged in two distant clusters, any outcome either moderately displeases everyone or strongly displeases one cluster, while a diffuse cloud of MSs admits a greater range of compromises. Though the total disapproval may be similar in both cases, the degree of acceptance may differ. Almost any social choice mechanism risks disaffection in the presence of strong clusters, while almost any sensible mechanism is likely to find acceptance in the diffuse case.

This example may seem trite, but without a metric, we could at best speak in terms of a single issue. A well-constructed D_M allows us to incorporate all issues into a unified geometry.

We implicitly treated social choice outcomes as MSs (or perhaps points in $B_{S,R}$). This sometimes makes sense, but often does not. Let us consider an example of each.

Suppose we have a set V of candidates for office. These have MSs, and we will just treat them like a subset of voters $V \subset X$. We have a small set of points in $B_{S,R}$ representing candidates, and a much larger set representing voters. D_M encapsulates all relevant issues,

not just one. If certain issues are expected to be paramount in the election, $P(S)$ could be adjusted so that D_M reflects that emphasis.

A quantity like $D_c \equiv \max_{i,j \in V} D_{ij}$ measures the dispersion of candidates, and we could calculate their diversity relative to voters via D_c/D_{avg}. Candidates tightly clustered relative to voters do not offer much choice, and the election would feel pointless regardless of the voting mechanism. The absence of such clustering does not guarantee acceptance. Candidates still must be suitably distributed relative to the voting population.

Let $f(r)$ denote the fraction of voters within radius r of any candidate, with inverse $r(f)$ denoting the minimum radius r at which a fraction f of voters would be within range of some candidate. These could serve as measures of available choice, or to furnish threshold criteria. For example, we could demand that 80% of voters have a candidate within $0.2D_{avg}$ of them ($r(0.8) < 0.2D_{avg}$ or $f(0.2D_{avg}) > 0.8$), and no more than 5% of voters must pick a candidate $0.5D_{avg}$ away ($r(0.95) < 0.5D_{avg}$ or $f(0.5D_{avg}) > 0.95$). Note that such constraints only address the variety of candidates. Other facets, such as the social choice mechanism itself or aforementioned voter clustering, may play a major role in acceptance.

Let us now consider social choice involving a single issue, perhaps via referendum or legislation. Let O be the set of possible outcomes, and suppose that any element of $B_{S,R}$ favors a single outcome as reflected in some known $f : B_{S,R} \to O$. This could be a complicated function, or as simple as $\operatorname{argmax}_{o \in O} P(o|s)$ if $O \subset R$ and some $s \in S$ directly probes that issue.

Any well-behaved f partitions $B_{S,R}$ into subspaces, each representing adherents to a particular outcome. Let $l_{o,i} \equiv \min_{P' \in f^{-1}(o)} D_M(P', P_i)$ denote the geometric distance from a voter's MS (embodied in CPD P_i) to the surface representing outcome o. A quantity like the mean distance of voters from a given surface ($l_o \equiv (\sum_{i \in X} l_{o,i})/|X|$) could furnish a quality measure for any given outcome o. This in turn could be used to rate the actual performance of various social choice mechanisms.

3.2 As a mechanism for social choice

We also may use D_M to build novel social choice mechanisms. Each diagnostic example above has a corresponding metric-based choice mechanism. In fact, we could use the departure of existing social choice mechanisms from these as a diagnostic tool in itself.

As before, choices involving candidates or schools of thought have outcomes represented by points in $B_{S,R}$. We only can compute quantities such as centroids in the presence of a Euclidean embedding of $B_{S,R}$ (rather than just X), and we will not assume one.

One approach would be to define a utility function $u(i \in V) \equiv \sum_{j \in X} f(D_{ij})$ representing displeasure, where f is some function that maps distance to displeasure (e.g. $f(d) = d^2$), and select the outcome that minimizes it via $\operatorname{argmin}_{i \in V} u(i)$.

We may want to impose constraints and could implement these in several ways. Hard constraints, such as $D_{ij} < t$ ($i \in V, j \in X$) for some distance t and fraction r of voters, can prevent undesirable scenarios, but risk excluding all outcomes. An alternative is to adjust $u(i)$ via a penalty term. These are just two examples, and most of the playbook of general optimization theory can be brought to bear.

Returning to our single-issue example, the corresponding social choice mechanism would pick $\operatorname{argmin}_{o \in O} l_o$, the surface that minimizes the mean distance to voters.

In principle, it may be possible to entirely replace voting with a metric-based social choice mechanism. The MSs encompass full sets of views on major issues. Once we know them (with possible adiabatic adjustment as needed), each election differs only in the choice of candidates and the relative prominence of issues. We once again could account for the latter via changes to $P(S)$, if accomplished in a manner that defies objection.

Given a snapshot of the voters' and candidates' MSs, the selection process then becomes automatic. We have ignored obvious practical concerns (such as how to obtain the MSs and

potential gaming of the system), and this approach would prove utterly impractical in real elections. However, it may have other uses. Comparison of derived outcomes with actual election results serves as an additional diagnostic tool and can help identify whether an existing social choice mechanism is fair or representative.

4. Related concepts

Many applications have some additional structure that allows us to define quantities reflecting notions of hypocrisy, judgment of others, worldview and moral trajectory. These can serve as aids to social choice or furnish additional mechanisms. Throughout this section, we will assume the concept of participants, as discussed in Section 2.13.

4.1 Hypocrisies

The purpose of our framework is not to judge MSs as better or worse than one another, but to measure distances between them. In this sense, it is agnostic to the MSs involved. Even within the confines of this moral relativism, an individual still may be judged against himself. Given a nontrivial indexing set J and map $g : Y \times J \to X$, we can define a set of $\frac{|J|(|J|-1)}{2}$ distances between the MSs of any given participant $y \in Y$. We will term these the "hypocrisies" of y, denoted $h_{ij}(y) \equiv D_M(g(y, i), g(y, j))$ for $i, j \in J$.

For example, suppose each person has three associated MSs ($J = \{\mathbf{p}, \mathbf{a}, \mathbf{b}\}$): (**p**) that they claim, (**a**) that they exhibit and (**b**) that they believe in or aspire to. We will ignore how one practically would ascertain (**p**) or (**b**).

Loosely speaking, h_{pa} corresponds to a notion of true hypocrisy ("Do as I say, not as I do"), h_{pb} could be termed superficial hypocrisy ("Do as I say, but what I say differs for you and me") and h_{ab} relates to courage of one's convictions ("I do as I do, not as I should"). This is a vast oversimplification, but vast oversimplifications often prove useful.

The presence of hypocrisies allows us to define various ratios and linear orderings. We can (1) compare two hypocrisies for a given participant $h_{ij}(y) \overset{?}{\leq} h_{kl}(y)$, (2) compare the same hypocrisy for two different participants $h_{ij}(y) \overset{?}{\leq} h_{ij}(y')$, (3) induce a weak linear order on Y for each hypocrisy using (2) thus ranking participants despite the absence of a linear order on X, (4) compute the dimensionless ratio $h_{ij}(y)/h_{kl}(y)$ (suitably controlled for zeros), (5) compute the dimensionless ratio $h_{ij}(y)/h_{ij}(y')$ (suitably controlled for zeros), (6) construct a pseudometric $D_{J,y}$ on J for each $y \in Y$ by pulling D_M back along $g(y, \cdot) : J \to X$ (unsurprisingly, $D_{J,y}(i, j) = h_{ij}(y)$).

Note that it does not matter where we get Y, J and g. If we have those components, we may define a set of hypocrisies. The essential element of hypocrisies is that they are defined for each participant without regard to any other.

4.2 Judgment

Hypocrisies constitute an inward-facing view of a person. We cannot judge those MSs in isolation, but can judge their constellation for a given participant. Let us now consider an outward-facing view. We will initially assume J is trivial (so $g : Y \to X$).

There is no preferred MS or participant in our framework, but we can ask how the world appears to any given MS or participant. MSs and participants are equivalent here, but will not be when we consider nontrivial J, so we will consider them both.

MS x sees x' at distance $D_M(x, x')$, and we have function $K_x : X \to \mathbb{R}$ given by $K_x(x') \equiv D_M(x, x')$. This is what the world looks like to x, and defines pseudometric $\tilde{D}_{M,x}(x', x'') \equiv |K_x(x') - K_x(x'')|$

on X, the pull-back of the Euclidean metric along K_x. There is nothing special about the Euclidean metric here, but with other metrics on $\mathbb{R}$, the resulting $\tilde{D}_{M,x}$ would not be as intuitive.

To see how $\tilde{D}_{M,x}$ differs from D_M, consider level sets. The level sets of D_M relative to x are (indexed by $l \geq 0$) $\{x' \in X | D_M(x, x') = l\}$. They partition X, and $\tilde{D}_{M,x}$ is a metric among them. $\tilde{D}_{M,x}$ does not care about the directions of x' and x'' relative to x, just their distances from it under D_M.

Analogous definitions hold with respect to Y. We define $\tilde{K}_y(y') \equiv K_{g(y)}(g(y'))$. For given $y \in Y$, we have an induced metric on Y given by $\tilde{D}_{Y,y}(y', y'') \equiv \left| \tilde{K}_y(y') - \tilde{K}_y(y'') \right|$. $\tilde{K}_y$ is how participant y sees the world.

Taking a cue from this, we define a "judgment" to be any non-negative map $\tilde{K} : Y \to \mathbb{R}$. A judgment induces a linear order on Y via $\tilde{K}(y) \overset{?}{\leq} \tilde{K}(y')$ and a distance on Y via $\left| \tilde{K}(y) - \tilde{K}(y') \right|$. A choice of $\tilde{K}$ is a preferred standard of judgment and constitutes extra information. One of our $\tilde{K}_y$'s could serve, but that requires choosing a specific y. Note that we equally well could define a judgment as $K : X \to \mathbb{R}$.

Suppose we have a preferred judgment $\tilde{K}_y$ and denote by F_Y the space of non-negative functions $Y \to \mathbb{R}$. Every $\tilde{K}_y \in F_Y$, as is $\tilde{K}$. For any function $A : F_Y \to \mathbb{R}$, we can define a metric on F_Y as $D_{F_Y}(\tilde{K}, \tilde{K}') \equiv \left| A(\tilde{K}) - A(\tilde{K}') \right|$. We also could compare $\tilde{K}_y$ and $\tilde{K}_{y'}$ for two participants this way. As an example, we could define $A(\tilde{K}) \equiv |Y|^{-1} \sum_{y \in Y} \tilde{K}(y)$. In this case, $D_{F_Y}(\tilde{K}, \tilde{K}_y)$ would represent how closely aligned y's perception of the world is with that of $\tilde{K}$, through the lens of the arithmetic mean. Note that we could do the same with functions $A : F_Y \to \mathbb{R}^n$ for any $n > 0$.

Let us now generalize to nontrivial J. Judgments defined in terms of X are unchanged, and $K : X \to \mathbb{R}$ remains the same under the new definition. In terms of participants, things are a little different. The equivalent map now is $\tilde{K} : Y \times J \to \mathbb{R}$, and this defines a "judgment." $\tilde{K}_y$ is replaced with $\tilde{K}_{y,i}$, and to prefer one requires a choice of both the participant and index. It also is possible to deal with nontrivial J by confining ourselves to a preferred choice of $j \in J$, but this is tantamount to a trivial J with restricted $X_j \equiv \{g(y, j) | y \in Y\}$.

In the presence of nontrivial J, any choice of judgment $\tilde{K}$ yields an alternate set of hypocrisies $\tilde{h}_{ij}(y) \equiv \left| \tilde{K}(y, i) - \tilde{K}(y, j) \right|$, those seen through the lens of $\tilde{K}$ rather than D_M. If $\tilde{K}$ is chosen to be $\tilde{K}_{y,i}$ for some y and i, then $\tilde{h}_{ij}(y) = h_{ij}(y)$ and $\tilde{h}_{jk}(y) = \left| h_{ij}(y) - h_{ik}(y) \right|$.

4.3 Worldview
The $\tilde{K}_y$ are views of the world implied by D_M and may not reflect participants' actual views. We do not know those actual views, and they would have to be supplied. A judgment $\tilde{K}$ for each participant will be termed a "worldview." It is a map $\eta : Y \to F_Y$, assigning a judgment to each participant. Any map $A : F_Y \to \mathbb{R}^n$ induces a map $A \circ \eta : Y \to \mathbb{R}^n$. We will denote by $\hat{\eta}$ the worldview induced by D_M. Clearly, $\hat{\eta}(y) = \tilde{K}_y$. For simplicity, we will once again assume J is trivial.

A choice of η allows us to speak not only of a participant's MS, but how they view other MSs. Though we are unlikely to be supplied with an explicit worldview, a distinct metric D_M' would generate a different worldview $\hat{\eta}'$, and for purposes of comparison, it may be useful to treat $\hat{\eta}'$ as externally imposed relative to D_M.

$\hat{\eta}$ is symmetric and represents the way two participants see one another through the lens of D_M. But worldviews need not be symmetric in general, and participants' views of one another

may not be reciprocal. In that case, it may make sense to define a "difference in mutual esteem" as something like $|\eta(y, y') - \eta(y', y)|$.

4.4 Moral trajectory

It sometimes is useful to relax our inference assumptions and allow slow variation of an MS. In place of ergodicity, we require that (a) our sample data be divisible into cohorts and (b) within each cohort, we have adequate data to infer the relevant CPD to our satisfaction. The cohorts need not be disjoint. We will define a "moral trajectory" to be a sequence of MSs obtained from cohorts of data for a single participant.

Time has not played any role in our framework so far, but this does not matter. We only require a means of segmenting our data into cohorts. External time intervals can serve, but so could other criteria. As long as our relaxed inference assumption applies to the cohorts, we are fine.

As an example, consider judges issuing criminal sentences. Rather than simply comparing MSs of different judges, we may wish to study the evolution of a given judge's MS over time. We could break our case history into year-long intervals and treat each as an independent MS. This opens the door to a variety of time-series tools, and we could study correlations between moral trajectories of different judges, etc.

Suppose we have a timeframe $[0, n\Delta]$ broken into intervals of length Δ and have sufficient sample data in each interval to adequately infer a CPD. Denoting the sequence of inferred CPDs $(v_1 \ldots v_n)$, there is a sequence of distances $(D_M(v_1, v_2), \ldots, D_M(v_{n-1}, v_n))$. From this, we could compute various moments, autocorrelations, etc. Given two such sequences, we also could compute correlations, etc.

4.5 Applications to social choice

Let us briefly consider a few of the many ways these concepts could be applied to social choice theory.

Any single hypocrisy induces a linear order among participants, and this could power any order-based social choice mechanism (e.g. ranking candidates).

We also could use hypocrisy statistics from a limited subset of participants to adjust our global geometry. For example, consider two MSs per person: $x_c(y)$ is that claimed, and $x_a(y)$ is that observed. Let us assume access to $x_c(y)$ for everyone, perhaps through surveys or interviews, but access to $x_a(y)$ only for a small subset of public figures (those with voting records, judicial histories, etc.). We could construct a PD $P(h)$ over hypocrisy from the known subset and apply it to the unknown subset. Note that we cannot infer or sample the unknown $x_a(y)$ itself, only its CPD. Denote by $P_c(y)$ and $P_a(y)$ the CPDs representing $x_c(y)$ and $x_a(y)$. *In lieu* of $P_a(y)$, we have a PD over points in $B_{S,R}$. To sample it, we first draw h according to $P(h)$, then sample uniformly within the level set $\{D_M(P_a(y), P_c(y)) = h\}$.

We could perform Monte Carlo analysis by sampling each $P_a(y)$ in this fashion, either in service of our own social choice mechanism or to test the robustness of another mechanism to such fuzziness. If demographic or other labeling data $l(y)$ are present, we could infer $P(h|l)$ rather than $P(h)$ from the known hypocrisies. Sampling each $P_a(y)$ from $P(h|l(y))$ could mitigate some of the (substantial) selection bias in our example, but at the cost of noisier inference.

Known hypocrisies of politicians also could be used to penalize candidates when a utility function is employed for social choice. More generally, hypocrisies offer a useful measure of fuzziness and may warn us of potentially unreliable results. Quantities such as judgment, worldview and mutual esteem can be employed to measure polarization and the likelihood of disaffection or to test the robustness of results to geometric fuzziness as in the hypocrisy scenario above.

They also can be used to emulate individual decision-making. Any given D_M generates a worldview, a judgment for each participant. Though this incorporates information about the individual's MS, it may not reflect their true judgment. This is not a matter of inference or hypocrisy. A worldview contains far more information than a metric, and D_M necessarily distills out certain aspects. Survey-based knowledge of MSs is plausible, but knowledge of true judgments is not. D_M may be all we have to work with, and a noisy D_M at that. However, this is not a dealbreaker. A metric-based emulation mechanism need not precisely reflect each individual's voting choice. It only must do so statistically and arrive at the correct overall election result. For example, we could emulate individual voting by ranking candidates according to each individual's $\tilde{K}_y$. In principle, we could replace voting altogether with a mechanism for MS acquisition and maintenance. This likely would be an awful idea in practice, but suggests that automated mechanisms along these lines may be worth exploring.

Moral trajectories could be used to detect convergent or divergent judicial behaviors, the impact of structural changes to mechanisms informing or effecting social choice, or sudden changes in behavior. Large changes in the geometry of voters or candidates or (most alarmingly) the two relative to one another could signal tectonic social shifts, which merit careful examination and possibly even reconsideration of the mechanism of social choice.

5. Metrics among conditional probability distribution

We now offer several methods of deriving a metric (or pseudometric) on $C_{S,R}$ from $P(S)$ and d_R. In this section, we will be careful to distinguish metrics from pseudometrics. Our approach is to break the problem into two parts: (1) from d_R derive a metric or pseudometric D on P_R, the space of PDs over R, and (2) from $P(S)$ and D, derive a pseudometric D_M on $C_{S,R}$.

5.1 Pseudometric vs metric

Note that almost any plausible methods, including our own, result in pseudometrics rather than metrics on $C_{S,R}$. We almost always must sum, integrate or average over $P(S)$ in some fashion, and this introduces degeneracy with near certainty. It turns out this is not a problem for two reasons.

A pseudometric suffices for most applications of our framework. Coincident points do not pose a problem for the techniques mentioned, and rarely do at all. It also turns out that we almost always end up with a metric on X, the set we actually care about. This is because X is small. D_M on X is just the restriction of D_M on $C_{S,R}$ to the particular set of CPDs representing X. Unless D_M is enormously degenerate on $C_{S,R}$, or some aspect of a given problem conspires to retain degeneracy, the probability that degeneracies will survive restriction to X is tiny. The same holds if d_R is a pseudometric, and this is one reason why a pseudometric is sufficient in that role.

5.2 Some metrics on $C_{S,R}$

The central obstruction to deriving metrics or pseudometrics ab initio is the triangle inequality. This is part of what motivated d_R as an *a priori* structure. Though we will not include proofs here, we observe that they depend heavily on two principles: (1) the pull-back of a metric is a pseudometric or metric and (2) the weighted average of a family of metrics is a pseudometric or metric.

In addition to being metrics or pseudometrics, our candidates pass certain sanity tests. For strongly peaked distributions, we require that D resembles d_R, and D_M resembles D. In what follows, w can be any strictly positive weight on R.

The following two candidates for D are pseudometrics. Here, P, Q are PDs over R (i.e. elements of P_R), and the $\pm$ are in tandem.

$$D(P, Q) \equiv \left| \int_R dx \int_R dy \, w(x) \, w(y) \, d_R(x, y) \left((P(x) \pm P(y))^2 - (Q(x) \pm Q(y))^2 \right) \right|$$

Given $P(S)$ and a choice of D, there are two straightforward choices for D_M. Here, f, g are the function forms of CPDs $P(R|S)$. For example, $f: S \to P_R$ yields a PD over R for each $s \in S$.

$$D_M^{(1)}(f, g) \equiv \sum_{s \in S} P(s) D(f(s), g(s))$$

$$D_M^{(2)}(f, g) \equiv D\left(\sum_{s \in S} P(s) f(s), \sum_{s \in S} P(s) g(s) \right)$$

When D is a metric, $D_M^{(1)}$ is a metric, and $D_M^{(2)}$ is a pseudometric. Note that the dependence on d_R is implicit in D.

6. Euclidean embeddings of d_R and D_M

An isometric embedding of metric space (Z, d) in metric space (Z', d') is an injection $i: Z \to Z'$ that is metric-preserving $\left(d(z_1, z_2) = d'(i(z_1), i(z_2)) \right)$, and a Euclidean embedding is an isometric embedding in $\mathbb{R}^n$ (endowed with the Euclidean metric). We will refer to a Euclidean space or embedding as "low-dimensional" if $n = 1, 2$, or 3. Our ability to visualize is limited to low-dimensional Euclidean spaces, and it is easiest to work with these.

Our framework features two metrics: d_R and D_M. A Euclidean embedding can assist in the *a priori* choice and specification of d_R and in the visualization of a derived D_M. The assumption that both metrics take units implies they have isometric embeddings in metric vector spaces, but these need not be low-dimensional or Euclidean.

6.1 Euclidean embeddings

Young and Householder identified the criterion for a Euclidean embedding to exist (Young and Householder, 1938). Let $Z = \{z_1 \ldots z_n\}$ and d_{ij} be the distance matrix for $(z_1 \ldots z_{n-1})$ relative to z_n. A Euclidean embedding of d exists iff the $(N - 1) \times (N - 1)$ matrix $B_{ij} \equiv \frac{1}{2} \left(d_{in}^2 + d_{jn}^2 - d_{ij}^2 \right)$ has only nonnegative eigenvalues, in which case rank B is the minimal embedding dimension. More efficient methods exist for actual calculation (see Crippen, 1978).

Exact Euclidean embeddings are rare, and low-dimensional ones are rarer. In most cases, an approximate embedding must suffice. Metric multidimensional scaling (MDS) is a method that replaces Young and Householder's B matrix with a lower-rank surrogate in a manner closely resembling principal component analysis (PCA). Details can be found in Eckart and Young (1936), and an alternate approach is offered in Matousek (2002, 2013). We can measure the quality of an approximate embedding in a variety of ways, such as the fraction of absolute eigenmass captured.

6.2 Visualization of D_M

D_M is derived rather than chosen, and we cannot expect it to have an exact low-dimensional Euclidean embedding. An approximate low-dimensional Euclidean embedding is possible, but may be of low quality. If the top three eigenvalues do not comprise most of the eigenmass, then too much information may have been lost. Since our purpose is visualization, this determination is subjective. We can produce a picture, but it may not be representative or useful.

6.3 Specification of d_R

The selection of d_R often is the most difficult aspect of our setup. It is less mutable and more critical than the choice of $P(S)$, and there is no obvious way to go about it, except in the simplest cases.

d_R is not just a pretty face. It is the core structure from which D_M derives, and the utility of the framework relies on d_R embodying a sensible intuition for distances on R.

Rarely does a natural d_R present itself, and R may be a complicated space. We often must leverage piecemeal intuition for distances into a precise metric, and a Euclidean embedding can help.

Were there a single correct d_R, this would not be the case. A general d_R is unlikely to have a reasonable-dimensional exact embedding or a sufficiently high-quality approximate one, and the crucial role of d_R will not brook lower quality.

However, our imperfect intuition bestows a degree of flexibility. We are doing something akin to heuristic embedding, much as an artist may render vague visual concepts into a cogent scene. Rather than a single correct d_R, there usually is a set of plausible candidates that fit our intuition. Practical or other considerations may further restrict this set, but it generally remains well populated. We will denote it S_d. Absent other criteria, any element of S_d may be selected as d_R. Robustness to that choice is a good test of the framework.

The larger S_d, the more likely there exists a reasonable-dimensional exact (or high-quality approximate) Euclidean embedding of at least one metric $d \in S_d$. This may not be low-dimensional, but sometimes can be constructed from low-dimensional component embeddings in a fashion we now describe.

To avoid excess verbiage, we will define a "Euclidean proxy," to be either an exact Euclidean embedding or a sufficiently high-quality approximate Euclidean embedding (one that does not lose relevant information). We do not assume Euclidean proxies are low-dimensional, but do require them to be of manageable dimension (i.e. not intractably large).

6.3.1 Subdivisible Euclidean embedding of d_R*: example.* It sometimes is possible to construct a higher-dimensional Euclidean proxy from easily visualized pieces. Certain systems, including many that arise in practice, have an R that naturally decomposes into semantically distinct components. We can try to construct a low-dimensional Euclidean proxy for each component, and then glue these together.

Consider a judicial sentencing framework where MSs arise from judges, S is a set of crimes and R is a set of punishments. Judges are presented with crimes, and they issue fines and/or jail terms. A point in R has natural coordinates (f, p), where f is in dollars and p is in years. Note that dollars and years are not units. R is not a vector space, and we have not posited translation invariance. f and p happen to be numeric labels, but have no more structure than lexical labels would.

We may not have direct intuition for the distance between ($5000, 4y$) and ($30000, 1y$), but we do have a sense of distances between two fines or two jail terms. Among other things, fines and jail terms each have a meaningful linear ordering.

Let us assume that fines are translation invariant in coordinates of dollars, corresponding to an exact embedding in $\mathbb{R}$ via $h(f) = f$ (i.e. assigning the numeric label its numeric value), with corresponding metric $d_1(f_1, f_2) = |f_1 - f_2|$. It now makes sense to refer to dollars as the "unit" for fines.

For jail terms, let us suppose this is not the case. Perhaps a one-year difference in jail term does not have the same marginal impact on a one-year sentence as on a ten-year sentence. Instead, we will assume a doubling of sentence has uniform significance (an unlikely perspective but suitable for illustration). This corresponds to an embedding in $\mathbb{R}$ via $h_2(p) = (c_1 + c_2 \ln x)$ for constants c_1, c_2 (in practice, we would probably employ something like

ln(p + 1) to avoid singularities near the origin). Choosing $c_1 = 0$ and $c_2 = \frac{1}{\ln 2}$, we get $d_2(p_1, p_2) = |\log_2 \frac{p_1}{p_2}|$.

Translation invariance only holds in the embedding coordinates, and $|p_1 - p_2|$ has no universal meaning. Only $|\log_2 \frac{p_1}{p_2}|$ has the same meaning everywhere, allowing the use of units. We will define unit $1T$ (for "term-doubling") to be $\Delta \log_2 p = 1$. Taking $f = 0$ and $\log_2 p = 0$ as the coordinate origins, (\$5000, 4y) becomes (\$5000, $2T$).

To obtain a Euclidean proxy for d_R, we must relate the scales of the two coordinates. If we deem \$20000 equivalent to one term-doubling, we can write our point as (\$5000, \$40000) in unified units of dollars.

We now have an embedding in $\mathbb{R}^2$. In terms of our original coordinates, it is $d_R([f, p], [f', p']) = \sqrt{(f - f')^2 + 20000^2 (\log_2(p/p'))^2}$.

6.3.2 Subdivisible Euclidean embedding of d_R*: general case.* Suppose in a problem, (1) every response can be decomposed into n distinct conceptual components: $R \subseteq \prod_1^n R_i$, (2) we have clear intuition for distances within each R_i and (3) we have some sense of how much each R_i should contribute to d_R. Note that R need only be separable semantically, not statistically or structurally.

For each R_i, we attempt to build a low-dimensional Euclidean proxy $h_i : R_i \to \mathbb{R}^{n_i}$ (with $n_i \leq 3$). If the R_i are small, simple spaces, such proxies are quite plausible. The corresponding metrics then are $d_i(x, x') \equiv |h_i(x) - h_i(x')|$.

To combine the d_i into d_R, we require a set of distance conversion factors. Let $c_{ij} > 0$ denote the distance in R_j corresponding to unit distance in R_i. These must satisfy $c_{ik} = c_{ij}c_{jk}$, $c_{ii} = 1$ and $c_{ij} = 1/c_{ji}$, and they comprise $n - 1$ independent values. Though their effect is simply to scale the embeddings $h_i \to c_{i1}h_i$ for $i = 2\ldots n$ (with d_i adjusted accordingly), they are not superfluous. The Euclidean proxies for the R_i are built in isolation, and their scales are arbitrary. We must adjust them to reflect our intuition for relative contributions, and the c_{ij} provide the necessary lever.

The resulting metric is $d_R([r_1 \ldots r_n], [r'_1 \ldots r'_n]) \equiv \sqrt{\sum_{i=1}^n c_{i1}^2 (h_i(r_i) - h_i(r'_i))^2}$ where $r_i \in R_i$ and $r'_i \in R'_i$ are expressed in unembedded coordinates.

This approach still is very restrictive, and we only can represent a small fraction of metrics this way. Aside from the need for a semantic decomposition of R, and low-dimensional Euclidean proxies for all the R_i, the conversion factors also impose a big constraint. They require that the relative meaning of distances in R_i and R_j be the same everywhere. Otherwise, we could not glue them with a simple, global scale factor. If such a constraint is unacceptable, this method cannot be used.

Fortunately, conceptual decomposition is organic to many problems. In applications where we have the flexibility to choose R, this may motivate our choice. Also, we always can try to expand an existing R into a suitable space.

Although we could try something similar with non-Euclidean embeddings, we made implicit use of a special property of Euclidean spaces: Euclidean metrics can be combined using a Euclidean metric. All p-norm metrics have this property, but most other families of metrics do not.

7. Conclusion

The framework described has broad applicability. Although our discussion centered on MSs and social choice applications, any decision system that can be framed in suitable terms may be analyzed using our methods. Examples could include customer satisfaction, political intelligence, judicial analysis and business planning.

There are many possible directions of future research. Our exploration of derived metrics (distilled to the selection presented in Section 5) was by no means exhaustive. Each metric captures certain facets of behavior, and additional candidates would mean greater flexibility.

Questions of stability relative to changes in underlying assumptions and components are important and deserve attention in any real application. Lack of robustness of D_M to small

changes in $S, R, P(S)$ and d_R can impair its utility. It also should be stable in the face of minor changes to $B_{S,R}$, the estimation method or aggregation procedure. Conceptually small changes in the framing of a problem should not drastically alter results.

We have said little about practical issues of data acquisition, cleaning or curation. These are of critical importance in any application, as is the relevance of those data. In addition to standard empirical issues, there may be specific ones surrounding our particular combinations of inference, estimation and Euclidean embedding. Our earlier comments notwithstanding, direct inference of distances also may be worth exploring.

The potential applications to social choice are myriad. We mentioned a few, briefly and imprecisely. Each of these could prove beneficial or interesting. The idea of using static or adiabatic knowledge of MSs to automate decisions may have applications in diverse fields, replacing frequent, burdensome social choices with upfront data acquisition and some periodic maintenance. A great deal more can be said about optimization of utility functions, constraints and cluster analysis as well. All these topics may provide fruitful avenues of inquiry.

References

Arrow, K.J. (1950), "A difficulty in the concept of social welfare", *Journal of Political Economy*, Vol. 58 No. 4, pp. 328-346.

Crippen, G. (1978), "Rapid calculation of coordinates from distance matrices", *Journal of Computational Physics*, Vol. 26, pp. 449-452.

Eckart, C. and Young, G. (1936), "The approximation of one matrix by another of lower rank", *Psychometrika*, Vol. 1 No. 3, pp. 211-218.

Hamming, R.W. (1950), "Error detecting and error correcting codes", *The Bell System Technical Journal*, Vol. 29 No. 2, pp. 147-160.

Levenshtein, V.I. (1966), "Binary codes capable of correcting deletions, insertions and reversals", *Soviet Physics–Doklady*, Vol. 10 No. 8, pp. 707-710.

Matousek, J. (2002), *Lectures on Discrete Geometry*, Springer-Verlag, New York, NY, available at: https://link.springer.com/book/10.1007/978-1-4613-0039-7#about.

Matousek, J. (2013), "Lecture notes on metric embeddings", available at: https://kam.mff.cuni.cz/~matousek/ba-a4.pdf.

Mitchell, T.M. (1997), *Machine Learning*, McGraw-Hill, New York, NY, available at: https://www.worldcat.org/title/machine-learning/oclc/36417892.

Mohri, M. (2018), *Foundations of Machine Learning*, 2nd ed., MIT Press, Cambridge, MA, available at: https://dl.acm.org/doi/10.5555/2371238.

Rao, C.R. (1945), "Information and accuracy attainable in the estimation of statistical parameters", *Bulletin of the Calcutta Mathematical Society*, Vol. 37 No. 3, pp. 81-91.

Vapnik, V.N. (1999), *The Nature of Statistical Learning Theory*, 2nd ed., Springer-Verlag, New York, NY, available at: https://link.springer.com/book/10.1007/978-1-4757-3264-1#about.

Young, G. and Householder, A. (1938), "Discussion of a set of points in terms of their mutual distances", *Psychometrika*, Vol. 3, pp. 19-22.

Corresponding author

Kenneth Halpern can be contacted at: khalpern@alum.mit.edu