Coherent rejection functions for arbitrary things

Kevin Blackwell[1]

[1]Department of Philosophy, University of Bristol, United Kingdom

Abstract

This paper investigates how to characterize (axiomatize) coherent rejection functions for arbitrary objects. In this very general setting, we assume that there is some sensible notion of preference over these objects, the existence of an "objective" background order treated as a constraint on preferences, and that the preferences constitute a strict partial order; we assume nothing else about the structure of the objects. The first insight is that we can represent binary comparisons of objects by ordered pairs – which are just another kind of thing; it is simple to represent coherent preference orders directly in terms of these higher-order objects. Once we have coherence axioms for sets of desirable things of this kind, we can immediately see what the corresponding coherence axioms are for sets of desirable sets (SDS) of these things. But these are not in one-to-one correspondence with rejection functions for the original things; they express more. The second insight is that rejection functions do correspond (almost) exactly to objects we will call "set preferences". So I give coherence axioms for set preferences, which equivalently fully characterize coherence for rejection functions. I present two main results: (1) coherence axioms for set preferences, and (2) the connection between coherent rejection functions for the original things and coherence for SDS of the ordered-pair things.

Keywords. coherence, rejection functions, desirability, sets of desirable things, sets of desirable sets of things

1. Background and setting up the question

Suppose we have an arbitrary set of things X; these could be, for instance: pizzas, penguins, protractors, probability mass functions, . . . – even things whose names don't begin with 'p'. We are assuming nothing about the kinds of objects that might be in X and nothing about the mathematical structure of X; in particular, X can be of any cardinality.

In this very general setting, there are clearly sensible ways of formulating notions of both coherent preference and coherent desirability. Jasper De Bock explores desirability in [1]; Gert de Cooman, Arthur Van Camp, and Jasper De Bock build on this work in [4].[1] However, the standard way of *relating* preference and coherent desirability needs additional structure: the set of things must be a vector space. For essentially the same reasons: there are clearly sensible ways of defining both coherent choice/rejection functions (which we can think of as a kind of *uncertainty model on binary preference*) and clearly sensible ways of defining coherent sets of desirable sets of things (which are, in parallel, uncertainty models over *sets of desirable things*), but no obvious way of relating these things.

De Bock poses this question, of how to characterize coherent choice functions in terms of desirability for arbitrary things, as a problem for future research in [1, p. 150]. In this paper, I provide an answer to the equivalent question of how to characterize *rejection* functions in terms of desirability for arbitrary things. In this section, I review some relevant background: the definitions of coherent preference and coherent rejection functions for arbitrary things; the definitions of a coherent set of desirable things and a coherent set of desirable sets of things; and the standard connections between desirability and preference, between rejection functions and sets of desirable sets of things, in the special case where our set of things X is a vector space.

I'll conclude this introduction by saying *just a bit* about the motivation for this kind of general study of rejection functions. I don't have any particular applications in mind, but this formalism should be useful in any context where we have enough structure / information to sensibly talk about preferences that an agent has, but not enough structure to represent the objects the agent has preferences over in terms of a vector space. This formalism can be usefully applied to practical decisions,

[1]There is a related vein of research which readers might also be interested in which generalizes desirability in a different way: assume the standard structure of gambles, but generalize the concept of coherence. E.g., Miranda and Zaffalon study coherent desirability of gambles in terms of general closure operators in [6]. We can view this vein as intermediate in generality between the standard treatments of sets of desirable gambles and the general study of desirability introduced by De Bock [1].

ethical judgments, aesthetic preferences, and epistemic rationality. Of course, there is a tradeoff here: with more assumptions on structure of the objects, we can express more interesting constraints and do things like drawing interesting connections between practical and epistemic rationality. On the other hand, the general notion of coherence for rejection functions studied here is applicable in an *enormous* range of contexts.

1.1. Coherent preferences for arbitrary things. Suppose we have a "background" strict partial order $\rhd$ (i.e., $\rhd$ is transitive and irreflexive) on X which we are interpreting as an "objective" constraint on our agent's preferences. In certain contexts, we might take this background order seriously as a normative commitment (e.g., representing a genuine rational or moral constraint); however, for the purposes of this investigation, all that really matters is that the agent in question *regards* this as a constraint, in the sense that they will treat only preference orders which accept all judgments in $\rhd$ as well-posed. Suppose that our agent's preferences are to be represented by a strict partial order, $>$.[2] In this context, there is a very natural notion of coherent preference:

Definition 1.1 (Coherent Preference Order on X)**.** $> \subset X \times X$ is *coherent* iff:

P_1. $x \rhd y \to x > y$ (Background Order Inheritance)

P_2. $x \not> x$ (Irreflexivity)

P_3. $x > y$ and $y > z \to x > z$ (Transitivity)

Under this definition, of course, $\rhd$ is itself the smallest coherent preference order on X.

1.2. Coherent rejection functions, characterized in terms of preference orders. Rejection functions are a very general imprecise model of preference originally introduced by Kadane, Schervish, and Seidenfeld [5]. Suppose an agent is presented with some set of options O; even with great uncertainty about which options they most prefer, they may be able to decisively *reject* some options as determinately worse than some other option on offer. Call the set of all such rejected options, $R(O)$, the *rejection set* for O.

Definition 1.2 (Rejection Function)**.** A rejection function, $R : \mathcal{P}(X) \to \mathcal{P}(X) : O \subseteq X \mapsto R(O) \subseteq O$,[3] maps each subset of X, O, to some subset $R(O)$.

The intended interpretation is that $R(O)$ represents the agent's rejection set for the option set $O \subseteq X$. It is common in the literature to also refer to choice functions, which are *dual* to rejection functions in the sense that: for any rejection function R, we can define a choice function

$$C_R : \mathcal{P}(X) \to \mathcal{P}(X) : O \mapsto O \setminus R(O) \tag{1}$$

which represents exactly the same imprecise preferences as R.

Intuitively, we want rejection functions to represent *uncertainty* over binary preferences. This motivates notions of *consistency* and *coherence* in terms of coherent preference orders. For a rejection function to be *consistent*, it should be "sharpenable into" some coherent preference order; that is, there should be some coherent preference order which is *at least as decisive* in the sense that it makes *every rejection judgment* that the rejection function does (and possibly more). We can formalize this in terms of a dominance concept:

Definition 1.3 (Binary rejection function corresponding to a preference order)**.** For any preference order $> \subseteq X \times X$, define $R_>$ s.t., $\forall O \subseteq X$, $R_>(O) := \{a \in O : (\exists b \in O)(b > a)\}$.

[2] In the rest of the paper, I'm going to indulge in the standard practice of taking these orders to be extensionally defined. That is, we'll take a preference order *to be* a set of ordered pairs specifying preferences the agent has, $> := \{(x, y) : x > y\}$. There are good reasons to be dubious about *identifying* an agent's preferences with this kind of extensionally-defined relation.

A familiar general worry about extensional definitions of relations is the problem of coextensionality. If a relation *is defined* in terms of its extension, then any two relations which have the same extension are the same relation. But it seems, intuitively, that there can be relations which happen to be coextensive at the actual world, but which *could* well be separable. In particular, we often think of preferences as being sensitive to certain features or properties of objects; agents can make the same preference judgments over some set of objects, but because of *different* preferences over these features. E.g., suppose that in some collection of ice cream flavors, only the pistachio ice cream is made with honey. Two agents might both prefer this flavor to all of the others – but for one, that's because of the pistachio; while for the other, that's because of the honey. In such a case, there seems to be something deeply misleading about saying that the agents have the same preferences, even though they exhibit the same (extensionally-defined) preference order on these objects. There are also some familiar ways to try to deal with this problem: intensional definitions, or taking preferences to be extensionally defined over *all possible* objects of some kind.

Despite these worries, working with extensional definitions will greatly simplify the notation throughout the paper, and repeated references to the extension of a preference order (as possibly distinct from the preference order itself) are cumbersome; thank you to an anonymous reviewer for suggesting I streamline the paper this way. Another way of putting the point might be: throughout the paper, I'll be assuming that a preference order is just an extensionally-defined object; but this probably means that there are important aspects of what *preferences* are that preference orders fail to capture.

[3] Throughout this paper, for any set S, $\mathcal{P}(S)$ denotes the power set of S. There is a subtlety about the definition of rejection functions that I am not going to worry about in this paper. It's often required that $R(O)$ is a *proper* subset of O, which reflects the idea that our agent should be able to permissibly choose *something* from any option set. There are substantive philosophical debates to be had about this assumption, especially in the context of infinite option sets. In particular, coherence in the sense of Definition 1.6 is entirely consistent with every option being eliminated by some other option in an infinite chain of preference. One trick is to insist that option sets must be finite; then coherence in this sense *does* guarantee that there's always some non-rejected option. Alternatively, we could require that the choice set is always non-empty even for infinite option sets. For present purposes, I leave the question of what we should think about this open. If we take the former tack, it would affect the results of this paper very little. If we take the latter tack, then a dominating preference order representation is *necessary* but not *sufficient* for a rejection function to be coherent.

Definition 1.4 (Dominance for rejection functions)**.** Rejection function R' *dominates* rejection function R iff $(\forall O \subseteq X)(R'(O) \supseteq R(O))$.

Definition 1.5 (Consistent rejection function)**.** For any rejection function $R : \mathcal{P}(X) \to \mathcal{P}(X) : O \subseteq X \mapsto R(O) \subseteq O$, R is *consistent* iff there is at least one *coherent* preference order $> \subseteq X \times X$ s.t. $R_>$ dominates R.

The intuitive idea behind *coherence* is that R is not only *consistent*, but "follows through on its commitments" – e.g., from Transitivity and Background Order Inheritance (BOI). There might be multiple coherent preference orders which dominate a consistent rejection function; when they *"disagree"* about whether to reject a from O or not, there's nothing wrong with our agent failing to reject a from O. But if they *all agree* that a should be rejected, then rejecting a is really a consequence from coherent preference – which can be inferred even with R's uncertainty, given only the judgments R does make. Formally:

Definition 1.6 (Coherent rejection function)**.** Let R be a consistent rejection function, and let Λ be the (nonempty) set of all coherent preference orders which dominate it. Then R is *coherent* iff $(\forall O \subseteq X)(\forall a \in O)(a \in R(O) \leftrightarrow (\forall > \in \Lambda)(\exists b \in O)(b > a))$.

Lemma 1.1. *A rejection function R is coherent iff there is* any *nonempty set of coherent preference orders, Λ, s.t. $R = \bigcap_{>\in\Lambda} R_>$.*

Proof. If R is coherent, then by definition, we have that there is a nonempty Λ consisting of coherent preference orders which dominate it, and that $(\forall O \subseteq X)(\forall a \in O)$, $a \in R(O) \leftrightarrow (\forall > \in \Lambda)(\exists b \in O)(b > a)$. Viz., $(\forall O \subseteq X)(R(O) = \bigcap_{>\in\Lambda} R_>(O))$.

For the other direction: suppose there is any (nonempty) Λ consisting of coherent preference orders s.t. $R = \bigcap_{>\in\Lambda} R_>$. It's immediately clear that every $> \in \Lambda$ is a coherent preference order which dominates R (and Λ isn't empty), so R is consistent.

It is not necessarily true that Λ contains *every* coherent preference order which dominates R; call Λ' the set of *all* coherent preference orders which dominate R. By definition, $\Lambda' \supseteq \Lambda$, and so $\bigcap_{>\in\Lambda'} R_> \subseteq \bigcap_{>\in\Lambda} R_> = R$. But also by definition, every $> \in \Lambda'$ dominates R, so $\bigcap_{>\in\Lambda'} R_> \supseteq R$. So, $\bigcap_{>\in\Lambda'}$ must also *equal* R. Equivalently, $(\forall O \subseteq X)(\forall a \in O)(a \in R(O) \leftrightarrow (\forall > \in \Lambda')(\exists b \in O)(b > a))$. And so, R is coherent. □

1.3. Coherent desirability for arbitrary things. Suppose we have some binary feature called *desirability* which a thing may or may not have. This framework generalizes the fundamental idea behind sets of desirable gambles (SDGs); in that case, desirability is interpreted as preference to the status quo (or the 0 gamble). We may have some set of objects $X^+ \subset X$ which are "objectively" desirable – as I remarked for the "background" preference order, we might have different interpretations for what this means depending on the kinds of things and purposes of the representation. E.g., in the SDGs special case, X^+ will contain all *weakly positive gambles* (nonnegative in all values, positive in at least one) and this is intended to represent a *rationality constraint*. Similarly, we may have some set of things $X^- \subset X$ which are "objectively" *not desirable*; in the SDGs special case, this will include all of the *weakly negative gambles*. Finally, we have a *closure operator*, which can be interpreted as representing a kind of inference mechanism: given that an agent finds some set of things $A \subseteq X$ desirable, they *should* (in some sense) also find $\mathrm{cl}(A)$ desirable.

Definition 1.7 (Closure operator)**.** $\mathrm{cl} : \mathcal{P}(X) \to \mathcal{P}(X)$ is a *closure operator* iff it satisfies all of the following properties for all $A, B \subseteq X$:

1. cl is extensive: $A \subseteq \mathrm{cl}(A)$.
2. cl is monotone: if $A \subseteq B$, then $\mathrm{cl}(A) \subseteq \mathrm{cl}(B)$.
3. cl is idempotent: $\mathrm{cl}(\mathrm{cl}(A)) = \mathrm{cl}(A)$.

Definition 1.8 (Coherence for sets of desirable things)**.** [cf. 1, Definition 2] A set of desirable things (SDT) $D \subset X$ is *coherent* iff:

D_1. $X^+ \subseteq D$

D_2. $X^- \cap D = \emptyset$

D_3. $D = \mathrm{cl}(D)$

1.4. Coherence for sets of desirable sets of things. A set of desirable things represents a *conjunctive* judgment: *each* $x \in D$ is desirable. One way of glossing what a set of desirable sets of things (SDST) is that it represents a conjunction of *disjunctive* judgments. Let $K \subset \mathcal{P}(X)$; so, each $A \in K$ will be a set of things (or the empty set). We interpret each $A \in K$ as representing a set of things which our agent is certain contains *at least one* desirable thing. (In this interpretation, of course, it never makes sense to include the empty set: it contains no things, hence no desirable things.)

This also yields a natural interpretation of a set of desirable sets of things as an *uncertainty* or *imprecision* model *over desirability*. Just as we can motivate coherence for rejection/choice functions in terms of coherent preferences, we can similarly motivate the coherence axioms for sets of desirable sets of things by appeal to coherent desirability.

Definition 1.9 (Coherence for sets of desirable sets of things)**.** [cf. 4, p. 4] For any $W \subseteq \mathcal{P}(X)$, a *selection map* $\sigma : W \to X : S \mapsto \sigma(S) \in S$ is a function that selects exactly one element of each $S \in W$. Call Φ_W the set of all selection maps on W. Define $\sigma(W) := \{\sigma(S) : S \in W\}$, the set of all things selected by σ across W.

A set of desirable sets of things $K \subset \mathcal{P}(X)$ is *coherent* iff:

K_1. $\emptyset \notin K$.

K$_2$. $S_1 \in K$ and $S_2 \supseteq S_1 \rightarrow S_2 \in K$.

K$_3$. $S \in K \rightarrow S \setminus X^- \in K$.

K$_4$. For all $x^+ \in X^+$, $\{x^+\} \in K$.

K$_5$. Take any nonempty $W \subseteq K$. For each $\sigma \in \Phi_W$, choose $x_\sigma \in \mathrm{cl}(\sigma(W))$. Then, $X_\Phi := \{x_\sigma : \sigma \in \Phi_W\} \in K$.

We've already discussed the motivation for K$_1$: the empty set cannot contain a desirable thing. K$_2$: if S_1 contains at least one desirable thing, then certainly any superset of S_2 does, too. K$_3$: since we are assuming that the elements of X^- are objectively undesirable, coherent desirability requires that a set can only be desirable in virtue of some elements not in X^-; so S must be desirable only if it would still be desirable with all elements of X^- removed. K$_4$: since every element of X^+ is objectively desirable, a set containing any element from X^+ is sufficient to make that set desirable. This actually motivates a stronger condition, viz.: if $A \cap X^+ \neq \emptyset$, then $A \in K$. But this stronger condition is derivable from K$_4$ and K$_2$, so we only need K$_4$ as an axiom.

K$_5$ is the trickiest of the axioms. Here's a quick sketch of the intuitive motivation: we want our coherence axioms to guarantee both that (1) any coherent set of desirable set of things, K, is dominated by at least one coherent set of desirable things; and (2) if every coherent set of desirable set of things dominating some coherent K "agrees" that a set is desirable, K must contain that set. Each selection map $\sigma \in \Phi_W$ represents a particular *conjunctive* way of making all of the disjunctive judgments encoded by W true. Viz., for each $A \in W$, we can interpret $\sigma(A)$ as a particular way that A could be desirable – A is desirable (as a set) because $\sigma(A)$ is desirable (as a thing). If we interpret $\sigma(W)$ as a set of desirable things, then by construction, $W \subseteq K_{\sigma(W)}$. In particular, when $W = K$, $\sigma(K)$ is, by construction, a set of desirable things which dominates K (viz., $K_{\sigma(K)} \supseteq K$). So, every selection map on K corresponds to a SDT which dominates K.[4]

If these dominating sets of desirable things were coherent, then they would be closed under cl. An X_Φ is constructed to represent a judgment that *every* dominating SDT is committed to, assuming closure; it's built to contain at least one thing that each of these (closed) selection maps finds desirable. Then, K$_3$ and K$_1$ together with K$_5$ guarantee that at least one of these dominating SDTs is coherent; if not we could construct an $X_\Phi \subseteq X^-$, which would then force $\emptyset \in K$, contradicting K$_1$.

K$_5$ (with K$_3$ and K$_2$) also guarantees (2). The basic idea is that we can show that any A that all of the coherent dominating selection functions represent as desirable must be in K by using K$_5$ to show that $B \cup R$ is in K – where $B \subseteq A$ and $R \subseteq X^-$; then, by K$_3$, $B \in K$; and, by K$_2$, $A \in K$.

1.5. Normal connection between desirability and preference. Suppose that X is a real vector space with a finite basis (this can also be extended to countable settings, but I'll focus on the simpler case here). Let 0 denote the zero vector (the additive identity element). We ordinarily take the "background" order in this context to be *weak dominance*: $a \rhd b$ iff $a \geq b$ and $a \neq b$ (where $a \geq b$ represents a *component-wise* comparison). The exposition in this subsection follows ideas developed, in much more detail, in [2] and [3].

Desirability is normally defined as *preference to the zero vector*: $a \in D_\succ$ iff $a \succ 0$. And it's typically assumed that $a \succ b$ iff $a - b \succ 0$. Together, these assumptions have the immediate consequence that $a \succ b$ iff $a - b \in D_\succ$; in the same way, for any set of desirable things D, we can define the corresponding preference order $\succ_D$.

Interpreting sets of desirable sets of things as an imprecision model over desirability and rejection functions as an imprecision model over preferences, the natural connection is to relate a SDST K to corresponding rejection function R_K as: for any $O \subseteq X$ and any $x \in O$,

$$x \in R_K(O) \leftrightarrow O - x \in K, \tag{2}$$

where $O - x$ denotes the set $\{a - x : a \in O\}$. In this interpretation, this means that x is rejected from O iff the agent is certain that some element of O is preferred to x, without necessarily knowing which one; which is equivalent to saying that the agent is certain that there is some $a \in O$ such that $a - x$ is desirable, without necessarily knowing which one. Again, just as for the connection between preference and desirability, we can define K_R in terms of any rejection function R in precisely the same way.

In that context, it's possible to prove equivalence between coherence axioms on preference and coherence axioms on desirability; similarly, between coherence axioms on rejection functions and coherence axioms on sets of desirable sets of things.

However, this way of proceeding is only possible because of special features of being a vector space. In the present paper, we're considering sets of objects on which operations like subtraction will not even be definable (what would "subtracting" a penguin from a protractor mean?) – *let alone* having any reason to think this operation (if definable) corresponds in any useful way to facts about preferences. We are not assuming that X has any kind of privileged element (like a zero element), so there is also no general way of translating facts about desirability as facts about preference comparisons to some privileged element.

How can we characterize rejection functions in terms of something like sets of desirable sets of things in this

[4] But this isn't a one-to-one correspondence, because according to a single SDT, multiple elements of a single $A \in K$ might be desirable; also, SDTs could be more informative (make more judgments) than required by the selection maps. For both of these reasons, there can be multiple selection maps corresponding to a single SDT.

very general setting?

2. First insight: binary comparisons of things are things, too

Let $Q = X \times X$ be the set of all ordered pairs of X-things. We can interpret any "Q-thing" (x, y) as representing the preference $x > y$. Then take *desirability* of Q things to represent that the preference in question holds.

Definition 2.1 (Sets of desirable Q-things representation of preferences on X)**.** Let $D \subseteq Q$ be a set of desirable "Q-things". We can interpret any such D as describing a preference order on X, $>_D := D$. And we can interpret any preference order on X as a set of desirable Q-things, $D_> := >$.

Then the coherence axioms for sets of desirable Q-things which correspond to our coherence axioms for preference orders on X (P_1-P_3) are obvious:

Definition 2.2 (Coherent sets of Q-things)**.** Let $Q^+ := D_\triangleright = \triangleright$ and $Q^- = \{(x, x) : x \in X\}$ be the set of all reflexive preference judgments.

$\mathrm{cl}(D)$ is the transitive closure of D, which we can define as follows: let $\mathbb{D}_T(D) = \{D' \supseteq D : >_{D'}$ is transitive$\}$ be the set of all transitive sets of Q-things which dominate D. Then $\mathrm{cl}(D) = \bigcap_{D' \in \mathbb{D}_T(D)} D'$, the unique smallest transitive set of Q-things which includes D.

$D \subset Q$ is a *coherent set of desirable Q-things* iff:

D_1. $Q^+ \subseteq D$ (Background Order Inheritance)

D_2. $Q^- \cap D = \emptyset$ (Irreflexivity)

D_3. $D = \mathrm{cl}(D)$ (Transitivity)

Proposition 2.1 (Equivalence of $>$ and D coherence)**.** *Any preference order $>$ on X satisfies axioms P_1-P_3 iff $D_>$ satifies axioms D_1-D_3. And any set of desirable Q-things satisfies D_1-D_3 iff $>_D$ satisfies P_1-P_3.*

The proof is trivial, merely unpacking definitions.

In a way, this reframing of coherent preference orders on X as coherent desirable sets of Q-things is trivial; $D_>$ is just the extension of $>$, which we are taking to be what $>$ *is*. But there's a conceptual shift here from thinking about properties (like desirability) of the X-things to thinking about properties of the ordered pairs. And although the recasting of coherent preference orders is trivial, this shift to thinking about "desirability" of Q-things also opens up the possibility of thinking about sets of desirable sets of Q-things; as we will see, the expressive power of SDSTs will be extremely useful in figuring out how to characterize coherent rejection functions.

2.1. Sets of desirable sets of Q-things and rejection functions. Here's a natural thought:[5] in the "normal" context (see subsection 1.5), once we've established the equivalence between coherence for desirability and coherence for preference orders, there is similarly an equivalence between coherence for sets of desirable sets of things and equivalence for rejection functions. We've identified an equivalence between coherent sets of desirable Q-things and coherent preference orders on X; perhaps we should also hope that there will be an equivalence between coherent sets of desirable sets of Q-things and coherent rejection functions for X-things?

Given the work of Gert de Cooman, Arthur Van Camp, and Jasper De Bock, we can quite easily obtain the coherence axioms for sets of desirable sets of Q-things which correspond (in the sense explained in subsection 1.4) to our coherence axioms for sets of desirable Q-things (D_1-D_3); this is a special case of Definition 1.9.

Definition 2.3 (Coherence for sets of desirable sets of Q-things)**.** A set of desirable sets of things $K \subset \mathcal{P}(Q)$ is *coherent* iff:

K_1. No empty set: $\emptyset \notin K$.

K_2. Supersets: $S_1 \in K$ and $S_2 \supseteq S_1 \to S_2 \in K$.

K_3. Irreflexivity: $S \in K \to S \setminus Q^- \in K$.

K_4. BOI: For all $q^+ \in Q^+, \{q^+\} \in K$.

K_5. Transitivity: Take any nonempty $W \subseteq K$. For each $\sigma \in \Phi_W$, choose $q_\sigma \in \mathrm{cl}(\sigma(W))$. Then, $Q_\Phi := \{q_\sigma : \sigma \in \Phi_W\} \in K$.

We can also identify a natural analogue of Eq. 2. First, a piece of notation:

Definition 2.4. For any $A \subseteq X$ and $x \in X$, define $A^x = \{(a, x) : a \in A\}$.

From any set of desirable sets of Q-things K, we can construct a rejection function for X-things defined by: for any $O \subseteq X$ and any $x \in O$,

$$x \in R_K(O) \leftrightarrow O^x \in K. \tag{3}$$

Intuitively, just as in the normal case, we interpret a set $A \in K$ as representing the disjunctive judgment that at least one of the Q-things in A is desirable. So, $O^x \in K$ means that the agent is certain there is some $o \in O$ such that (o, x) is desirable, without necessarily knowing which one; given our interpretation of desirability for Q-things, this means that the agent is certain some $o \in O$ is preferable to x, without necessarily knowing which one.

I won't prove it yet (see Theorem 4.1 below), but it turns out that if K is coherent (satisfies K_1-K_5), R_K will be coherent in the sense of Definition 1.6.

However, this correspondence between coherent Ks for Q-things and coherent rejection functions for X-things is *not one-to-one*. One quick way to see this is that a K can contain many sets which are not of the form A^x (for any $A \subseteq X$, $x \in X$). All such sets are irrelevant

[5] Or at least it seemed so to me.

to R_K, and so *incoherent* Ks must be able to generate coherent R_Ks. (E.g., take a coherent K and add a random set of Q-things not of the form A^x. This new K' is obviously incoherent; it, at least, violates K_2. But $R_{K'} = R_K$.) Furthermore, *distinct coherent* K, K' can generate the same $R_K = R_{K'}$.[6]

So, although Equation 3 gives us a recipe for extracting coherent rejection functions for X-things from coherent sets of desirable sets of Q-things, it is not yet clear how to generally *characterize* coherent rejection functions for X-things. One obvious idea is to formulate a kind of mathematical object which is more closely related to rejection functions – this is what we explore in the next section.

3. Second insight: "set preferences"

Seidenfeld, Schervish, and Kadane connect rejection functions to a strict partial order over sets of options:

Definition 3.1 (SSK: strict partial order on option sets from rejection)**.** [8, p. 164] For any $O_1, O_2 \subseteq X$, $O_1 \langle O_2$ iff $O_1 \subseteq R(O_1 \cup O_2)$.[7]

Jason Konek pointed out to me that we can characterize rejection functions in terms of a relation which compares sets of options against *single objects* from X, like so:

Definition 3.2 (Set preference on X)**.** A set preference is a *heterogenous relation* between subsets of X and elements of X; viz., $>^s \subseteq \mathcal{P}(\mathcal{P}(X) \times X)$.

The intended interpretation is that $A >^s x$ whenever at least one element of A is preferred to x, without necessarily being able to determine which element is preferred – this is *very* similar to a rejection function. Given a set preference $>^s$, we can define a corresponding rejection function $R_{>^s}$: $x \in R_{>^s}(A) \leftrightarrow (x \in A$ and $A >^s x)$; and given a set preference $>^s$ we can define a corresponding rejection function $R_{>^s}$: $A >^s_R x \leftrightarrow x \in R(A \cup \{x\})$.

There's a slight subtlety with equating set preferences and rejection functions if we allow (as above) set preferences to make comparisons even when $x \notin A$. When we extract a rejection function from a set preference, we *lose information* about the comparisons like this that the set preference makes. But if the set preference is *coherent*, then there is a unique way of recovering $>^s$ from $R_{>^s}$. Minimally, all we need is that $>^s$ is irreflexive.[8]

Definition 3.3 (Irreflexivity of a set preference)**.** A set preference is *irreflexive* iff for every $A \subseteq X$ and every $x \in X$, $A >^s x \leftrightarrow A \setminus \{x\} >^s x$.

Proposition 3.1 (Equivalence of rejection functions and set preferences under irreflexivity)**.** *For any rejection function R, $R_{>^s_R}$ and R are equivalent.*

If $>^s$ is an irreflexive *set preference, then $>^s_{R_{>^s}}$ and $>^s$ are equivalent.*

Proof. Consider arbitrary rejection function R, $x \in X$, and $A \subseteq X$. $x \in R_{>^s_R}(A) \Leftrightarrow x \in A$ and $A >^s_R x$. Equivalently: $x \in A$ and $x \in R(A \cup \{x\})$, which is true iff $x \in R(A)$.

Consider any irreflexive set preference $>^s$, arbitrary $x \in X$, and arbitrary $A \subseteq X$. $A >^s_{R_{>^s}} x \Leftrightarrow x \in R_{>^s}(A \cup \{x\})$; equivalently: $A \cup \{x\} >^s x$. Because we are assuming $>^s$ is irreflexive, this is true iff $A >^s x$. □

The SSK order on options and set preferences should also clearly be intertranslateable under coherence. Suppose we have an arbitrary relation on option sets $\langle \subseteq \mathcal{P}(X) \times \mathcal{P}(X)$ which may or may not actually be an SSK order (viz., may or may not actually be representable as being constructed from a rejection function as in Definition 3.1). Under what conditions can we translate between $\langle$ and a corresponding set preference, *as if* $\langle$ were an SSK order? The minimal conditions we need are irreflexivity of the set preference and a condition on the option set order which I will call "composition". Composition is a necessary feature for $\langle$ to represent a rejection function in the way an SSK order does, as I sketch in Footnote 9 below.

Definition 3.4 (Composition for option set orders)**.** Let $\langle \subseteq \mathcal{P}(X) \times \mathcal{P}(X)$ be a relation on option sets. $\langle$ satisfies *composition* iff, for any $O_1, O_2 \subseteq X$, $O_1 \langle O_2 \leftrightarrow (\forall x \in O_1)(\{x\} \langle (O_1 \cup O_2))$.[9]

Definition 3.5 (Translation scheme between option set orders and set preferences)**.** For any $\langle$ order on option sets, we we can define the corresponding set preference $>^s_\langle := \{(A, x) \subseteq \mathcal{P}(X) \times X : \{x\} \langle A\}$. Similarly, for any set

[6] E.g.: let $K_V = \{A \subseteq Q : (\exists q^+ \in Q^+)(q^+ \in A)\}$ be the vacuous model – the smallest coherent SDS of Q-things. Suppose $\rhd$ leaves at least 4 distinct X-things *completely unconstrained* (there is no comparison by the background order between any of these 4 things and any other X-thing): call them a, b, c, and d. Let $K' = K_V \cup \{A \subseteq Q : A \supseteq \{(a, b), (c, d)\}\}$. It is easy to see that K' is coherent and distinct from K_V, but nonetheless $R_{K'} = R_{K_V}$.

[7] In this paper, they work with choice/rejection functions in the context of horse lotteries; however, this connection between orderings on sets of options and choice/rejection is general – it's still applicable in this context.

[8] Here's a sketch of the issue: a rejection set is always a subset of the option set. So, the rejection function doesn't directly store any comparisons between a set and any elements not in the set. But, in our intended interpretation of $A >^s x$ as meaning that at least one element of A is preferred to x, we *should* have that $A >^s x$ iff $A \setminus \{x\} >^s x$; it doesn't make sense to prefer x to itself. *If $>^s$ respects this*, we can recover whether $A >^s x$, for some $x \notin A$, from $R_{>^s}$ by checking whether $A \cup \{x\} >^s x$ – viz., by whether $x \in R_{>^s}(A \cup \{x\})$. But without assuming something like this, there's no way to recover that kind of information from $R_{>^s}$.

[9] If $\langle$ were an SSK order, then $O_1 \langle O_2$ means that $O_1 \subseteq R(O_1 \cup O_2)$. I.e., each $x \in O_1$ is rejected from $O_1 \cup O_2$, which $\langle$ must also represent by $\{x\} \langle (O_1 \cup O_2)$ to be generated by a rejection function in this way.

preference $>^s$, we can define $\langle_{>^s} := \{(O_1, O_2) \subseteq \mathcal{P}(X) \times \mathcal{P}(X) : (\forall x \in O_1)((O_1 \cup O_2) >^s \{x\})\}$.

Proposition 3.2 (Equivalence of option set orders and set preferences under minimal coherence requirements)**.** *For any set preference $>^s$ which is irreflexive, $>^s_{\langle_{>^s}}$ and $>^s$ are equivalent.*

If an option set order $\langle$ satisfies composition, then $\langle_{>^s_\langle}$ and $\langle$ are equivalent.[10]

Proof. Consider irreflexive set preference $>^s$, arbitrary $x \in X$, and arbitrary $A \subseteq X$. $A >^s_{\langle_{>^s}} x$ iff $\{x\}\langle_{>^s} A$ iff $A \cup \{x\} >^s x$. Since we are assuming $>^s$ is irreflexive, this is true iff $A >^s x$.

Consider any option set order $\langle$ which satisfies composition and arbitrary $O_1, O_2 \subseteq X$. $O_1 \langle_{>^s_\langle} O_2$ iff $(\forall x \in O_1)((O_1 \cup O_2) >^s_\langle \{x\})$ iff $(\forall x \in O_1)(\{x\}\langle(O_1 \cup O_2))$. Since we are assuming $\langle$ satisfies composition, this is true iff $O_1 \langle O_2$. □

Once again, the intended interpretation of a set preference as an imprecision / uncertainty model over binary preference motivates certain notions of consistency and coherence. From any binary preference order $>$, we can define the corresponding set preference $>^s_>$: for any $A \subseteq X$ and any $x \in X$,

$$A >^s_> x \leftrightarrow (\exists y \in A)(y > x) \tag{4}$$

Definition 3.6 (Consistency for set preferences)**.** A set preference $>^s$ is *consistent* iff it is dominated by at least one *coherent* preference order on X. Viz., there is some coherent binary preference order $>$ s.t. $>^s_>$ dominates $>^s$ in the sense that $>^s_> \supseteq >^s$.

The intuitive idea here is, unsurprisingly, the same as for rejection functions. Since a set preference is supposed to represent uncertainty over binary preference, it can only consistently make judgments which binary preference orders could coherently make; there must be some binary preference order which is *at least as decisive*. And, also much like for rejection functions, to be *coherent*, a set preference must be consistent and must make all judgments that "follow" (via closure, together with required judgments from the background order) from other judgments that it makes. We can formalize this idea in much the same way as for rejection functions:

Definition 3.7 (Coherence for set preferences)**.** Let $>^s$ be a consistent set preference and let Λ be the nonempty set of all coherent preference orders which dominate it. Then $>^s$ is *coherent* iff $\forall A \subseteq X, x \in X, A >^s x \leftrightarrow (\forall > \in \Lambda)(\exists y \in A)(y > x)$.

[10] An immediate corollary is that this translation scheme always works for SSK orders, because SSK orders must satisfy composition.

Lemma 3.1. *A set preference $>^s$ is coherent iff there's any nonempty coherent set of preference orders, Λ, s.t. $>^s = \bigcap_{>\in\Lambda} >^s_>$.*

The proof of Lemma 3.1 is essentially the same as the proof of Lemma 1.1

If we can find coherence axioms for set preferences, we have equivalently axiomatized coherence for rejection functions; here are my coherence axioms for set preferences:

Definition 3.8 (Coherence axioms for set preferences)**.** For any $W \subseteq >^s$, define $W_p := \{A^x : (A, x) \in W\}$.[11] Let Φ_{W_p} be the set of all selection maps, $\sigma : W_p \to Q : S \mapsto \sigma(S) \in S$. Define $\sigma(W_p) := \{\sigma(S) : S \in W_p\}$ – the set of all Q-things selected by σ across W_p.

A set preference $>^s$ is *coherent* iff:

SP_1. Supersets: If $A >^s x$ and $B \supseteq A$, then $B >^s x$.

SP_2. BOI: $(\exists y \in A)(y \rhd x) \to A >^s x$.

SP_3. Consider any nonempty $W \subseteq >^s$ and, for each $\sigma \in \Phi_{W_p}$, choose any $q_\sigma \in \mathrm{cl}(\sigma(W_p))$. Let $Q_\Phi := \{q_\sigma : \sigma \in \Phi_{W_p}\}$. Then:

(a) For any (nonempty) $W \subseteq >^s$, and any Q_Φ constructible as above, $Q_\Phi \setminus Q^- \neq \emptyset$.

(b) For any (nonempty) $W \subseteq >^s$ and any such Q_Φ, if $Q_\Phi \setminus Q^- = A^x$ for some $A \subseteq X, x \in X$, then $A >^s x$.

In comparison to the coherence axioms for sets of desirable sets of things, my SP_3 essentially combines the axioms for Irreflexivity (handled by SP_3(a)), No empty set (a trivial application of SP_3(a)), and Transitivity (handled by SP_3(b)). The motivation for combining the removal axiom (Irreflexivity) and the closure axiom (Transitivity) is the following problem:

A Q_Φ constructible as outlined in SP_3 may or may not represent a judgment that set preferences "know how to understand". To get all of the judgments which *are* binding on the set preference, I'm essentially borrowing the greater expressive power of sets of desirable sets of things (Ks) for Q-things and then filtering (in SP_3(b)) to only Q_Φ sets which have the form of A^x. Judgments of this form are the only ones which set preferences can express. But consider a Q_Φ which has the form $A^x \cup R$, where R contains only reflexive judgments (viz., $R \subseteq Q^- = \{(y, y) : y \in X\}$); Q_Φ will, in general, not have form B^x, because R can contain Q-things other than (x, x). And so we cannot directly "translate" Q_Φ into a set-preference judgment.

But, of course, the fact that this is a Q_Φ which can be generated from the commitments $>^s$ already makes *should* require that $A >^s x$, in line with Definition 3.7. No coherent preference order which dominates $>^s$ can

[11] Recall definition 2.4. Note that $W_p \subseteq \mathcal{P}(Q)$; it's a set of sets of pairs of X-things.

do so by accepting any of the judgments in R – they're all inconsistent. What this strongly suggests is that we should do reflexivity removal *first* – if, after this removal, the set takes a form which can be expressed as a set preference judgment, $>^s$ must make it to be coherent. So, this is what SP_3 does.

Theorem 3.1 **(SP_1-SP_3 and definition 3.7 are equivalent).** *A set preference order $>^s$ is coherent (satisfies axioms SP_1-SP_3) iff there is a (nonempty) set, Λ, of coherent preference orders on X s.t. $A >^s x \leftrightarrow (\forall > \in \Lambda)(\exists a \in A)(a > x)$.*

Proof. Suppose we have such a Λ and construct $>^s_\Lambda$. Let's show that $>^s_\Lambda$ satisfies the axioms:

SP_1:

$$A >^s_\Lambda x \Rightarrow (\forall > \in \Lambda)(\exists a \in A)(a > x). \tag{5}$$

If $B \supseteq A$, then those witnessing elements are all also in B, so

$$(\forall > \in \Lambda)(\exists a \in B)(a > x), \tag{6}$$

hence $B >^s_\Lambda x$, by construction.

SP_2: Every $> \in \Lambda$ satisfies P_1, so

$$(\forall > \in \Lambda)(y \rhd x \to y > x). \tag{7}$$

Thus, if there's some $y \in A$ s.t. $y \rhd x$, then

$$(\forall > \in \Lambda)(y > x), \tag{8}$$

hence $A >^s x$.

SP_3: Consider any $W \subseteq >^s_\Lambda$.

$$(\forall > \in \Lambda)(\exists \sigma \in \Phi_{W_p})(\forall S \in W_p)(\sigma(S) \in >). \tag{9}$$

This is a straightforward consequence of the fact that

$$A >^s_\Lambda x \Leftrightarrow (\forall > \in \Lambda)(\exists (a,x) \in A^x)((a,x) \in >). \tag{10}$$

(But there can be multiple selection maps on W_p corresponding to a single $> \in \Lambda$; sometimes $>$ accepts more than one judgment in an A^x.)

(a) Thus, it's immediately clear that every Q_Φ must satisfy $Q_\Phi \setminus Q^- \neq \emptyset$. (Consider even just a single $> \in \Lambda$ and a single $\sigma_>$ for which $(\forall S \in W_p)(\sigma(S) \in >)$. By definition, any Q_Φ includes an element $q_{\sigma_>} \in \mathrm{cl}(\sigma_>(W_p))$. $\sigma_>(W_p) \subseteq >$, so $\mathrm{cl}(\sigma_>(W_p)) \subseteq \mathrm{cl}(>) = >$, because $>$ is coherent. And, also because $>$ is coherent, we know $Q^- \cap \mathrm{cl}(>) = \emptyset$, hence also $Q^- \cap \mathrm{cl}(\sigma_>(W_p)) = \emptyset$. So, every Q_Φ contains some $q_{\sigma_>} \notin Q^-$.)

(b) Now observe that for any Q_Φ, *every* $> \in \Lambda$ "contributes" at least one element, in the sense that

$$(\forall > \in \Lambda)(\exists q_{\sigma_>} \in Q_\Phi \cap >). \tag{11}$$

Q_Φ might contain some things from Q^-, or it might not. But, again, by coherence of each of the preference orders in Λ,

$$(\forall > \in \Lambda)(\exists q_{\sigma_>} \in (Q_\Phi \setminus Q^-) \cap >). \tag{12}$$

If $Q_\Phi \setminus Q^-$ happens to be representable as an A^x (all of the ordered pairs have a single x as their second element), then we have already shown that $A >^s x$:

$$(\forall > \in \Lambda)(\exists q_{\sigma_>} = (a,x) \in A^x \cap >); \tag{13}$$

equivalently:

$$(\forall > \in \Lambda)(\exists a \in A)(a > x). \tag{14}$$

Now, the other direction: assume we have a coherent $>^s$ and find the Λ. First, we'll show that if $>^s$ satisfies SP_1-SP_3 then there is *at least one* coherent preference order on X which dominates $>^s$:

Let $W = >^s$. Given SP_3(a), there is at least one selection map $\sigma^* \in \Phi_{>^s{}_p}$ s.t. there is no way of choosing $q_{\sigma^*} \in Q^-$. (If not, there would be $Q_\Phi \subseteq Q^-$.) Thus,

$$\mathrm{cl}(\sigma^*(>^s{}_p)) \cap Q^- = \emptyset. \tag{15}$$

Let

$$D := \mathrm{cl}(\sigma^*(>^s{}_p)); \tag{16}$$

we will show that D is a coherent set of desirable Q-things, and hence that $>_D := D$ is a coherent preference order on X. We have just shown that D satisfies D_2.

D_1: $>^s$ satisfies SP_2, so $\{(\{y\},x) : y \rhd x\} \subseteq >^s$ and therefore, $\{\{(y,x)\} : y \rhd x\} \subseteq >^s{}_p$. Thus, for *every* selection map $\sigma \in \Phi_{>^s{}_p}$,

$$Q^+ = \{(y,x) : y \rhd x\} \subseteq \sigma(>^s{}_p). \tag{17}$$

In particular, this is true of σ^*, so

$$Q^+ \subseteq \sigma^*(>^s{}_p). \tag{18}$$

Then, obviously,

$$Q^+ \subseteq \mathrm{cl}(\sigma^*(>^s{}_p)) = D. \tag{19}$$

D_3: $D = \mathrm{cl}(\sigma^*(>^s{}_p))$ and cl is idempotent.

To verify that $>_D$ dominates $>^s$, just recall that we constructed D from a selection map on binary preference judgments which would realize the set preference judgments $>^s$ makes: by construction,

$$(\forall S \in >^s{}_p)(S \cap \sigma^*(>^s{}_p) \neq \emptyset). \tag{20}$$

Thus,

$$(\forall A^x : A >^s x)(\exists (a,x) : a \in A)((a,x) \in D); \tag{21}$$

equivalently,

$$A >^s x \Rightarrow (\exists a \in A)(a >_D x). \tag{22}$$

Let

$$\Phi^{\mathrm{coh}} := \{\sigma \in \Phi_{>^s{}_p} : \mathrm{cl}(\sigma(>^s{}_p)) \cap Q^- = \emptyset\}; \tag{23}$$

$$\Lambda := \{>_{\mathrm{cl}(\sigma(>^s{}_p))} : \sigma \in \Phi^{\mathrm{coh}}\}, \tag{24}$$

which we have just shown to be nonempty. The argument that we just gave for a particular (arbitrary) one of these, generated by σ^*, generalizes: each $> \in \Lambda$ is a coherent preference order on X which dominates $>^s$.

Now obviously, by construction, we already have one direction of the equivalence we want: if $A >^s x$, then $(\forall > \in \Lambda)(\exists a \in A)(a > x)$. So, finally, we show the other direction. Consider any $(A, x) \in \mathcal{P}(X) \times X$ s.t. $(\forall > \in \Lambda)(\exists a \in A)(a > x)$; we show $A >^s x$. First, for each $> \in \Lambda$, choose exactly one of the elements of A witnessing $(\exists a \in A)(a > x)$ – call them $a_>$. Let

$$A_w := \{a_> : > \in \Lambda\}. \tag{25}$$

Clearly, $A_w \subseteq A$, so if we can show $A_w >^s x$, then also $A >^s x$ by SP_1.

Once again, we apply SP_3 with $W = >^s$, but this time we're aiming for (b). Observe that for any $\sigma' \in \Phi_{>^s{}_p} \setminus \Phi^{\mathrm{coh}}$, by definition,

$$\mathrm{cl}(\sigma'(>^s{}_p)) \cap Q^- \neq \emptyset. \tag{26}$$

So, for every $\sigma \in \Phi^{\mathrm{coh}}$, take

$$q_\sigma := (a_{>_{\mathrm{cl}(\sigma(>^s{}_p))}}, x). \tag{27}$$

For every $\sigma' \in \Phi_{>^s{}_p} \setminus \Phi^{\mathrm{coh}}$, choose

$$q_{\sigma'} \in Q^-. \tag{28}$$

By construction,

$$(\forall \sigma \in \Phi_{>^s{}_p})(q_\sigma \in \mathrm{cl}(\sigma(>^s{}_p))). \tag{29}$$

And

$$Q_\Phi = A_w^x \cup R, \tag{30}$$

where R is some subset of Q^-. So,

$$Q^\Phi \setminus Q^- = A_w^x, \tag{31}$$

and thus $A_w >^s x$, by SP_3(b). □

4. SDS of Q-things Representation Revisited

After analyzing coherent set preferences, it is finally clear what the relationship is between coherent rejection functions for X-things and coherent sets of desirable sets of Q-things. Every coherent rejection function is in one-to-one correspondence with a coherent set preference; and a coherent set preference corresponds to a particular subset of the judgments made by *at least one* coherent K.

Proposition 4.1 (SDS of Q-things representation for coherent rejection functions for X-things). *A rejection function for X-things, $R : \mathcal{P}(X) \to \mathcal{P}(X) : O \mapsto R(O) \subseteq O$, is* coherent *iff there is at least one coherent SDS of Q-things, K, s.t. $x \in R(A)$ iff $x \in A$ and $A^x \in K \cap \{B^y : B \in \mathcal{P}(X), y \in X\}$.*

Proof. R is coherent iff the corresponding $>^s_R$ ($x \in R(A)$ iff $x \in A$ and $A >^s_R x$) is coherent. (See Proposition 3.1: rejection functions and set preferences are equivalent under the assumption of irreflexivity, which is weaker than coherence. See Theorem 3.1 for the proof that the coherence axioms for set preferences are indeed equivalent to the notion of coherence in Definition 3.7; and compare with Definition 1.6.)

Suppose we start with a coherent K and construct R_K:

$$x \in R_K(A) \leftrightarrow (x \in A \text{ and } A^x \in K). \tag{32}$$

This is equivalent to the set preference $>^s_K$ defined s.t.:

$$>^s_{K\,p} := K \cap \{B^y : B \in \mathcal{P}(X), y \in X\}. \tag{33}$$

Let's show that if K satisfies K_1-K_5, then $>^s_K$ satisfies axioms SP_1-SP_3.

SP_1:

$$A >^s x \Rightarrow A^x \in >^s_{K\,p} \subset K. \tag{34}$$

$$C \supseteq A \Rightarrow C^x \supseteq A^x, \tag{35}$$

and so

$$C^x \in K, \tag{36}$$

by K_2. But also,

$$C^x \in \{B^y : B \in \mathcal{P}(X), y \in X\}, \tag{37}$$

so

$$C^x \in >^s_{K\,p}. \tag{38}$$

So $C >^s_K x$, by definition.

SP_2 If $(\exists y \in A)(y \rhd x)$, pick some witness and call it y^*. $(y^*, x) \in Q^+$, so by K_4,

$$\{(y^*, x)\} \in K. \tag{39}$$

$A^x \supseteq \{(y^*, x)\}$, so $A^x \in K$ by K_2. But also, $A^x \in \{B^y : B \in \mathcal{P}(X), y \in X\}$, so $A^x \in >^s_{K\,p}$. So $A >^s_K x$, by definition.

SP_3: consider any Q_Φ constructible for any $W \subseteq >^s_K$. By definition, $W_p \subseteq >^s_{K\,p} \subset K$. So Q_Φ is constructible by K_5 and $Q_\Phi \in K$.

(a) Suppose, for reductio, $Q_\Phi \subseteq Q^-$. Then by K_3, $\emptyset \in K$, contradicting K_1; contradiction. So, $Q_\Phi \setminus Q^- \neq \emptyset$, proving SP_3(a).

(b) If $Q_\Phi \setminus Q^-$ is representable as A^x for some $x \in X$, then $Q_\Phi \setminus Q^- \in K \cap \{B^y : B \in \mathcal{P}(X), y \in X\} = >^s_{K\,p}$. So, $A >^s x$, by definition.

The other direction: start with a coherent $>^s$ and use it to construct a coherent K s.t. $K \cap \{B^y : B \in \mathcal{P}(X), y \in X\} = >^s{}_p$. In Theorem 3.1, we have proved that $>^s$ satisfies axioms SP_1-SP_3 iff there is a (nonempty) set of coherent preference orders on X, Λ, s.t. $A >^s x \leftrightarrow (\forall > \in \Lambda)(\exists y \in A)(y > x)$. Let

$$K_\Lambda := \bigcap_{> \in \Lambda} K_>, \tag{40}$$

where

$$K_{>} := \{A \subseteq Q : A \cap > \neq \emptyset\}. \quad (41)$$

Because each $D_{>} = > \in \Lambda$ is a coherent set of desirable Q-things, it follows that each $K_{>}$ is a coherent SDS of Q-things (see [4, Proposition 6]). And the intersection of coherent Ks is always a coherent K. (Satisfying axioms K_1-K_5 is preserved under arbitrary intersections, and so "the intersection of any non-empty family of coherent SDSes is still coherent" [4, p. 6].)

To show that $K \cap \{B^y : B \in \mathcal{P}(X), y \in X\} = >^s{}_p$: by construction:

$$S \in K \Leftrightarrow (\forall > \in \Lambda)(S \cap > \neq \emptyset); \quad (42)$$

$$S \in \{B^y : B \in \mathcal{P}(X), y \in X\} \Leftrightarrow (\exists A \subseteq X, x \in X)(S = A^x). \quad (43)$$

So, we have the following chain of equivalences:

- $S = A^x \in K \cap \{B^y : B \in \mathcal{P}(X), y \in X\}$
- $(\forall > \in \Lambda)(A^x \cap > \neq 0)$
- $(\forall > \in \Lambda)(\exists (y, x) \in A^x)((y, x) \in >)$
- $A >^s x$.

□

5. Conclusion and Further Work

In this paper, I have investigated how to characterize/axiomatize coherent rejection functions in what I take to be one of the most general possible settings: we have no assumptions on the kinds of objects under consideration, and no assumptions on the mathematical structure of the set of objects X from which choice problems might be drawn. The only minimal assumption, which I think is necessary to motivate the present notion of coherent rejection, is that we have some strict partial order (transitive and irreflexive) $\rhd$ that is being interpreted as an "objective" constraint on preferences (also understood as strict partial orders). We understand rejection functions, as is common, as a general uncertainty/imprecision model over binary preference, which directly motivates the initial definitions of consistency (Definition 1.5) and coherence (Definition 1.6).

This paper makes two novel contributions: coherence axioms for set preferences (SP_1-SP_3), which immediately characterize coherence for rejection functions. This result is stated and proved in Theorem 3.1. Coherent set preferences are in one-to-one correspondence with coherent rejection functions, so we have that a rejection function R is coherent iff the corresponding $>^s_R$ satisfies axioms SP_1-SP_3.

The second contribution is to "make good" on the intuition that it should be possible to characterize coherent rejection functions, even in this very general setting, in terms of desirability. Subsection 1.5 briefly explains why this is a challenging problem: the normal connection between desirability and binary preference relies on special features of a vector space which X may well lack. My strategy is to "lift" desirability to ordered pairs which directly represent binary comparisons between X-things; the "desirability" of a Q-thing (x, y) is interpreted as accepting that x is preferred to y. In this framing, it was immediately clear how to extract coherent rejection functions (for X-things) from coherent sets of desirable sets of Q-things; however, coherent rejection functions for X and coherent SDS of Q-things turn out to not be in one-to-one correspondence. After characterizing coherent rejection functions in terms of coherent set preferences, I return to this question in Theorem 4.1: there is a special subset of SDS of Q-things which fully captures the choice/rejection behavior. A rejection function is coherent iff it corresponds to this subset for *at least one* coherent K. As remarked earlier in the paper, distinct Ks (both coherent and incoherent) can generate the same rejection function.

There are several ways in which I think this work could be usefully extended. The future development that I am personally most interested in is exploring coherent rejection in a context allowing for more varied kinds of "objective"/background constraints on preferences than merely those imposed by $\rhd$ and irreflexivity. A very easy extension would be to allow for additional objectively impermissible preferences (in the Q formulation, additional kinds of objects in Q^-); as long as these are disjoint with $\mathrm{cl}(Q^+)$, this would not affect the results of the present paper in any substantive way. More substantively, we could try to extend this to something like an accept- and reject- statement framework (cf. [7]), which is significantly more challenging. I am also interested in thinking more about the best way of representing indifference judgments in something close to this setting. (The point here is: indifference judgments obviously place substantive constraints on rejection functions; but "reading off" indifference judgments from rejection functions seems impossible without a notion of "sweetening the pot".) This also suggests a related issue: what are the minimal assumptions we would need to add to have a sensible notion of pricing of X-things in some currency / *measuring* their value according to some standard? These questions are clearly interconnected.

Additional Author Information

Acknowledgements. Many of the ideas in this paper were developed in consultation with Jason Konek; in particular, it was Jason's idea to try to axiomatize coherence for rejection function by way of set preferences. I would certainly have never found these results without Jason's suggestions.

This work is obviously also deeply indebted to the work on generalized desirability models from Jasper De Bock, Gert de Cooman, and Arthur Van Camp. My initial thought, in fact, was that once an equivalence had been demonstrated between the SDS of comparison-things and rejection functions for the original things, I could simply leverage their results. As it stands, my axiom SP_3 is *heavily* inspired by their closure axiom, K_5. I first learned about this more general approach to desirability from Jasper's talk at the Inaugural EpImp Conference; I've followed Gert's work on this and related approaches with great interest. As always, special thanks to Arthur, who taught me most of what I know about Imprecise Probabilities.

Anonymous reviewers suggested many ways of improving the presentation of this paper; I think that the current version of the paper has benefited substantially from implementing their suggestions.

References

[1] Jasper De Bock. "A theory of desirable things". In: *Proceedings of the Thirteenth International Symposium on Imprecise Probability: Theories and Applications*. Ed. by Enrique Miranda, Ignacio Montes, Erik Quaeghebeur, and Barbara Vantaggi. Vol. 215. Proceedings of Machine Learning Research. PMLR, Nov. 2023, pp. 141–152. URL: https://proceedings.mlr.press/v215/de-bock23a.html.

[2] Jasper De Bock and Gert de Cooman. "Interpreting, Axiomatising and Representing Coherent Choice Functions in Terms of Desirability". In: *Proceedings of the Eleventh International Symposium on Imprecise Probabilities: Theories and Applications*. Ed. by Jasper De Bock, Cassio P. de Campos, Gert de Cooman, Erik Quaeghebeur, and Gregory Wheeler. Vol. 103. Proceedings of Machine Learning Research. Thagaste, Ghent, Belgium: PMLR, Mar. 2019, pp. 125–134.

[3] Gert de Cooman. "Coherent and Archimedean choice in general Banach spaces". In: *International Journal of Approximate Reasoning* 140 (2022), pp. 255–281. DOI: 10.1016/j.ijar.2021.09.005.

[4] Gert de Cooman, Arthur Van Camp, and Jasper De Bock. "The logic behind desirable sets of things, and its filter representation". In: *International Journal of Approximate Reasoning* 172 (2024), p. 109241. DOI: 10.1016/j.ijar.2024.109241.

[5] J. Kadane, Mark Schervish, and Teddy Seidenfeld. "A Rubinesque theory of decision". In: *A festschrift for Herman Rubin* 45 (Jan. 2004). DOI: 10.1214/lnms/1196285378.

[6] Enrique Miranda and Marco Zaffalon. "Nonlinear desirability theory". In: *International Journal of Approximate Reasoning* 154 (2023), pp. 176–199. DOI: 10.1016/j.ijar.2022.12.015. URL: https://www.sciencedirect.com/science/article/pii/S0888613X22002341.

[7] Erik Quaeghebeur, Gert de Cooman, and Filip Hermans. "Accept & reject statement-based uncertainty models". In: *International Journal of Approximate Reasoning* 57 (2015), pp. 69–102.

[8] Teddy Seidenfeld, Mark J. Schervish, and Joseph B. Kadane. "Coherent choice functions under uncertainty". In: *Synthese* 172.1 (2010), pp. 157–176. DOI: 10.1007/s11229-009-9470-7.