Analogy in Decision-Making

Massimiliano Amarante

Université de Montréal et CIREQ

Abstract. In the context of decision making under uncertainty, I formalize the concept of analogy: an *Analogy* between two decision problems is a mapping that transforms one problem into the other while preserving the problem's structure. After identifying the basic structure of a decision problem, I introduce the concepts of *Analogical Reasoning Operator* and of *Analogical Reasoning Preference*. The former maps the decision problem at hand into a family of decision problems, which are analogous to the problem under consideration. The latter is the result of aggregating the various analogies. I provide several representations (in decreasing order of generality) of the analogical reasoning operators. After introducing two mild assumptions on the aggregators of analogies, I characterize analogical reasoning (AR) preferences. I give several examples of AR preferences and of the associated aggregators. These include Gilboa-Schmeidler similarities, Choquet integrals and quantiles. Finally, I show that the class of Monotone Continuous Invariant Biseparable (MCIB) preferences (which includes many popular models of decision making under uncertainty) has an important stability property: Any MCIB preference is an AR preference; conversely, every AR preference which results from aggregating MCIB preferences is an MCIB preference.

Key words and phrases. Analogy, kernels, Analogical reasoning operator, Invariant Biseparable preferences, Choquet integral
JEL classification: D81
I am grateful to Mario Ghossoub, Tzachi Gilboa, Marcus Pivato and David Schmeidler for several conversations we shared on this paper's subject.

1. Introduction

Most of the theoretical work on decision making under uncertainty takes a certain type of behavior (preference) as a primitive, and then determines the preference functional which represents that behavior. The chief example is Savage theorem [**15**], which shows that a preference satisfying the rules (axioms) of completeness, transitivity, independence of irrelevant alternatives, etc (see [**15**]) is represented by an Expected Utility functional. The converse problem of determining the type of behavior corresponding to a given functional is also typically considered.

The focus of this paper is different. Here, rather than taking the preference as a primitive, I am interested in studying those processes that lead to the formation of the preference. Loosely speaking, I want to inquire into the reasons that lead a decision maker to choose a over b rather than b over a. Important work in this direction has been done by Gilboa and Schmeidler with their Case Based Decision Theory ([**11**] and [**12**]). In their theory, a decision maker assesses the profitability of an action by recording its performance in each of many decision problems of the same nature (cases). The decision maker keeps choosing the same action as long as the profit realized falls within certain bounds. On occasions, however, he experiments with new actions, which become the new default choice if they turn out to be extremely profitable. After using this procedure over a sufficiently large number of cases, a case-based decision maker will be able to assess (with a certain confidence) the profitability of all actions available, and will rank them accordingly. Thus, the theory of Gilboa and Schmeidler is an explanation of why that preference is formed (and, at the same time, of why we observed certain choices in the various cases). In fact, Case Based Decision Theory is more general than what was just described. Gilboa and Schmeidler allow their decision makers to gather evidence not only from decision problems of the same nature but also from decision problems of a different nature. Given the decision problem at hand, the evidence coming from the various decision problems is weighted by using a (transformation of a) *similarity function*. This expresses, in a synthetic way, the differences between any two decision problems and, therefore, how much the evidence gathered from a certain decision problem is worth for the problem at hand. This procedure is an example of *analogical reasoning*: the decision maker forms assessments about the problem at hand by gathering evidence from other problems that are somehow similar to it. The (transformation of the) similarity function is an example of *aggregator of analogies*: the evidence from each problem is weighted in a way that reflects how similar each problem is to the problem at hand.

Analogy, its mathematical formulation and its role in decision making under uncertainty are the focus of this paper. Analogy is the recognition that A (a phenomenon, a problem, etc.) is like B and that, therefore, consequences (inferences, explanations, solutions, etc.) that can be drawn from A can be drawn from B as well. The literature on the concept of Analogy spans at least from the time of the Sophists to our days and touches on nearly any field of knowledge: to attempt even a summary review would be an unmanageable task. Suffice to say, very synthetically, that Analogy is one of the cornerstones of human thought ([**14**]). As such it is expected to play a fundamental role in decision-making.

1.1. Analogy and aggregators of analogies. Given two problems in decision making, DP_1 and DP_2, a necessary condition for speaking of an analogy between the two is that one ought to be able to transform DP_1 into DP_2. The mapping $A : DP_1 \longmapsto DP_2$ describing this transformation would represent the analogy between the two problems. The idea of analogy, however, demands more: inferences, explanations, solutions for DP_1 must correspond via the mapping A to inferences, explanations, solutions for DP_2. This is a requirement on both the relation between DP_1 and DP_2 and on the mapping A: not only must DP_1 and DP_2 display in some sense the same properties, but also the mapping A must preserve this "alikeness" if it is to represent an analogy between the two. One encounters here the basic mathematical ideas of *structure* and of *structure-preserving mapping* (i.e., homomorphism), which thus dictate what kind of mathematical formulation one should aim for: the definition of analogy should be something like DP_1 *is analogous to* DP_2 *if there exists a structure-preserving mapping* $A : DP_1 \longmapsto DP_2$. Needless to say, a definition of this sort would be useful only when the basic structure of a decision problem is identified. Once this is done, one can go on to study decision makers who solve the problem they face by means of multiple analogies. In this study, the main task is that of characterizing "aggregators" of the various analogies.

The issue of (objective) existence of an analogy between two different problems is not addressed here. A possible view is that the existence of an analogy between two different problems is a subjective statement, that is, it pertains to the decision-maker, and as such is outside the theory. However, future considerations involving dynamics and learning might lead one to alter this point of view.

1.2. Paper outline. The main ideas are introduced in the next section, and then gradually elaborated in the subsequent two sections. In these sections, I show that a certain mathematical environment naturally emerges when studying the role of analogy in decision making under uncertainty. Section 4 is a central section of this paper as it operationalizes the definitions given earlier. In that section, I also formalize the requirements that an aggregator of analogies should satisfy and prove the first representation result: *A collection of analogies can be represented as a collection of multiple prior preferences over the same set of alternatives as the problems at hand* (Theorem 1, Sec. 4). A useful consequence of this representation result is that it allows us to think of aggregators of analogies as of aggregators of preferences, a familiar problem in economic theory. Beginning with Section 5, I restrict myself to a setting where the analogies are representable by Invariant Bi-separable preferences (IB), a wide class of multiple-prior preferences (see [**6**]). This restriction is motivated by three reasons. The first is that I want to give an example of how to apply the concepts of Sec. 4. The second is that this class is still wide enough to include a large variety of preferences that recur in the applications. The third, perhaps the most important, is that IB preferences have an important stability property in relation to the idea of analogical reasoning, which I prove in Section 6. When analogies are representable by means Invariant Bi-separable preferences, the representation theorem of Sec. 4 can be refined: analogies can be represented by capacities and aggregators of analogies can be represented by functionals on a space of functions of those capacities. In Section 5, I give several examples of aggregators of analogies and of the resulting analogical reasoning preferences. The examples of aggregators include Choquet integrals, Gilboa-Schmeidler similarities, quantiles and generalized quantiles. In Section 6, I show that Monotone Continuous Invariant Bi-separable (MCIB) preferences have the following stability property: *every analogical reasoning preference obtained by aggregating MCIB preferences is a MCIB preference; conversely, every MICB preference is an analogical reasoning preference.* Smaller classes of preferences do not have this property: that is, if one starts with a set of analogies represented by preferences that are not MICB, the resulting analogical reasoning preference may not be of the same type as the preferences that generated it. Section 7 concludes the paper. Appendix A contains some background material, Appendix B lists the axioms describing the class of IB preferences and Appendix C contains the proofs omitted from the main text.

1.3. A word about the assumptions. Because of the pervasive nature of analogy as a process of human thinking, a theory of analogy should rest on very minimal assumptions, and this is the case for the present theory. In Sections 3 and 4, nonetheless, I will be making assumptions requiring that certain mappings (utilities, preference functionals, aggregator of analogies) be real-valued. For the most part, the use of the reals $\mathbb{R}$ is only an expositional devise: $\mathbb{R}$ could be replaced by an abstract (non-Archimedean) ordered space or by a product of abstract ordered spaces without too much trouble. The use of $\mathbb{R}$-valued mappings, however, results in a much neater exposition while retaining all the conceptual substance. I will take care to warn the reader when the assumptions imply a substantial restriction.

2. Toward a definition of Analogy

By a problem in decision-making, I mean a Savage-style setting where a decision-maker is called upon to rank a certain set of alternatives F. Each alternative is viewed as a mapping $S \longrightarrow X$, where S is a set of states and X is a set of outcomes. I am going to think of a Savage model as of a "small world". Consequently, I am often going to consider indexed collections, $\{(S_i, X_i, F_i) : i \in I\}$, of decision problems.

The scope of this section is to build the skeleton of my theory of Analogy in decision making. Its four main constituents are: (1) the notion of Single Analogy; (2) the notion of Multiple Analogies and the related one of Analogical Reasoning Operator; (3) the concept of Aggregator of Analogies; and, finally (4) the concept of Analogical Reasoning Preference. Each of these concepts is developed in a specific subsection. The construction presented here is truly bare-bone: flesh and blood (the representation theorems) will be added in Section 4.

2.1. Single analogy. Let us begin with the simplest case. Intuitively, the story behind it goes as follows. Today, an individual faces a problem for which he has to provide a solution. He realizes that the problem (as a whole) "looks like" another problem that he solved yesterday and, therefore, he is going to use yesterday's solution to arrive at a solution for today's problem. As a first step, we must understand what this intuitive description entails in terms of the objects that define a decision problem. Clearly, the crucial issue is to give a precise meaning to the statement *a problem looks like another problem*. Let us label today's problem by DP_1, $DP_1 = (S_1, X_1, F_1)$. The problem consists of ranking the set of alternatives $F_1 = \{f, g, h, ...\}$, where an element of F_1 is a mapping $S_1 \longrightarrow X_1$. A solution to the the problem is a ranking $\succsim_1$ of the alternatives in F_1. Yesterday's problem is labeled $DP_2 = (S_2, X_2, F_2)$, and consisted of ranking a set of alternatives $F_2 = \{\varphi, \gamma, \eta, ...\}$. That was solved by means of a ranking $\succsim_2$. It is pretty clear that a very minimal requirement for us to say that DP_1 looks like DP_2 is that we ought to be able to associate to each alternative in DP_1 an alternative in DP_2. Thus, if DP_2 has to be "analogous" to DP_1, then there must be a mapping $A : F_1 \longrightarrow F_2$. In the terminology of the Introduction, the mapping A is the recognition that DP_2 is like DP_1. The next step consists of realizing that the solution that is drawn from DP_2 can be drawn, via the mapping A, from DP_1 as well. Thus, we can say that DP_1 is *solved by analogy* with the way DP_2 was solved if

$$f \succsim_1 g \qquad iff \qquad A(f) \succsim_2 A(g)$$

That is, DP_1 is solved by analogy with DP_2 if the ranking $\succsim_1$ is derived from the ranking $\succsim_2$, given the mapping A that describes the alikeness between the two problems.

A moment of thought, however, shows that this idea of alikeness is too weak to be fruitful. To see this, suppose, for example, that the same set of consequences occurs both in DP_1 and in DP_2, and that two alternatives $f, h \in F_1$ are such that h produces in each state the same consequences as f but in an order of magnitude twice as big. Then, it seems natural to demand that any reasonable definition of "alikeness" would demand that $A(f)$ and $A(h)$ would be in a similar relation with each other, at least in qualitative terms. For if not, the existence of a mapping $A : F_1 \longrightarrow F_2$ would appear as anything but a mathematical accident. Similarly, if $f, h \in F_1$ are associated to "almost the same consequences", it seems natural to demand that the same would be true for $A(f)$ and $A(h)$. Abstracting from the examples, what seems necessary in order to have a reasonable definition of analogy is that if two alternatives $f, h \in F_1$ are in a certain relation $\mathcal{R}$, $f\mathcal{R}h$, then this relation must be preserved by the mapping A, that is $A(f)\mathcal{R}A(h)$. I will refer to such mappings as structure-preserving mappings. The next definition summarizes the content of this subsection.

DEFINITION 1. *Let* $DP_1 = (S_1, X_1, F_1)$ *and* $DP_2 = (S_2, X_2, F_2)$ *be two decision problems. Denote by* $\succcurlyeq_i$ *the ranking in problem* $i, i = 1, 2$. *We say that* DP_1 *is solved by analogy with* DP_2 *if there exists a structure-preserving mapping* $A : F_1 \longrightarrow F_2$ *such that*

$$f \succsim_1 g \qquad iff \qquad A(f) \succsim_2 A(g)$$

for any two alternatives f, g *in* DP_1.

Clearly, this definition lacks substance since the structure of a decision problem has not been specified. I will study this structure in Section 3, and I will incorporate it in the definition of Analogy in Section 4 (Definition 7).

2.2. Multiple analogies; analogical reasoning operators. Here, the idea is again pretty intuitive but the story is a little more complex. In his life, our individual has already solved many problems, and several of those "look like" the problem DP_1 that he faces today. Thus, multiple analogies are possible. In general, however, it might be that different analogies lead to different solutions for DP_1. So, what our individual wants to do is to collect these multiple analogies together, and use all of them to come up with a solution for DP_1.

Once again, let us see what this intuitive description entails in a formal setting of decision making. Let us denote by $\mathcal{AP} = \{DP_2, DP_3, ...\}$ the set of problems that are analogous to $DP1$. I will index this set by J, with $1 \notin J$. By definition, for each problem $DP_j \in \mathcal{AP}$, there must exist a structure-preserving mapping $A_j : F_1 \longrightarrow F_j$. Each problem $DP_j \in \mathcal{AP}$ has already been solved and, by virtue of the analogy $A_j : F_1 \rightarrow F_j$, its solution is now a candidate solution for the problem $DP1$ to be solved today. A solution for DP_j, a ranking $\succsim_j$ of the alternatives F_j, is always representable by means of a preference functional $I_j : F_j \longrightarrow \mathbb{Y}_j$, where $\mathbb{Y}_j$ is some ordered space; that is

$$h \succsim_j l \qquad iff \qquad I_j(h) \geq I_j(l)$$

for any two alternatives h and l in F_j. By definition, the value $I_j(h)$ in $\mathbb{Y}_j$ represents the rank that alternative h has in problem DP_j.

These observations suggest the following construction. Since for each $DP_j \in \mathcal{AP}$ we have a structure-preserving mapping $A_j : F_1 \longrightarrow F_j$, we can define a mapping

$$f \longmapsto \{A_j(f)\}_{DPj\in\mathcal{AP}}$$

which carries alternative f from today's problem into a collection of alternatives, one for each analogous problem. $A_j(f)$ is a representation of alternative $f \in F_1$ in the analogy DP_j. Since $A_j(f)$ has already been ranked in DP_j, we can assign to $A_j(f)$ its rank $I_j(A_j(f))$. By combining this with the previous mapping, we can then define a mapping

$$f \longmapsto \{I_j(A_j(f))\}_{j\in J}$$

which takes alternative $f \in F_1$ into the collection $\{I_j(A_j(f))\}_{j\in J}$. For each $j \in J$, $I_j(A_j(f))$ represents the rank that the decision maker would assign to f in today's problem if he were to use analogy DP_j.

Example 1. *Suppose that each preference $\succsim_j$ on F_j is representable by means of an $\mathbb{R}$-valued functional, that is there exists an $I_j : F_j \longrightarrow \mathbb{R}$ such that for all $h, l \in F_j$*

$$h \succsim_j l \qquad \Longrightarrow \qquad I_j(h) \geq I_j(l)$$

According to our construction, each $f \in F_1$ is associated to the collection of numbers $\{I_j(A_j(f))\}_{j\in J}$. Intuitively, one can think of this as "If I use analogy j, then f is worth $I_j(A_j(f))$".

It is convenient to view the collection of ranks $\{I_j(A_j(f))\}_{j\in J}$ as a mapping defined on the set $\mathcal{AP}$ of analogous problems. To this end, let us define $\mathbb{Y} = \bigcup_{j\in J} \mathbb{Y}_j$ and, for each $f \in F_1$, let $\psi_f : \mathcal{AP} \longrightarrow \mathbb{Y}$ be the mapping defined by

$$\psi_f(DP_j) = I_j(A_j(f))$$

Summing up: In a decision theoretic setting, the idea of Analogy expressed by Definition 1 coupled with the idea of multiple analogies lead to canonically associating each alternative $f \in F_1$ to a mapping $\psi_f : \mathcal{AP} \longrightarrow \mathbb{Y}$. The latter contains the following information: *the value of the mapping ψ_f at DP_j expresses how alternative f fares if one uses analogy DP_j.*

Definition 2. *Let $f \in F_1$. We say that a mapping*

$$\mathcal{A} : f \longmapsto \psi_f$$

is an analogical reasoning (AR) operator if it is induced by a collection of structure-preserving mappings in the way described above.

EXAMPLE 2. *Assume that everything is as in Example 1. In such a case, the mapping* ψ_f *is* $\mathbb{R}$*-valued,* $\psi_f : \mathcal{AP} \longrightarrow \mathbb{R}$*. The AR operator* $\mathcal{A}$ *carries each* $f \in F_1$ *into the collection of all evaluations that* f *would take according to the various analogies* $DP_j \in \mathcal{AP}$*.*

2.3. Aggregators of Analogies. The AR operator takes each $f \in F_1$ into the mapping ψ_f, which records how f would rank according to the various analogies. Now, this information has to be used to rank f in the problem at hand. Let Ψ be the set of all mappings of the form ψ_f for some $f \in F_1$.

DEFINITION 3. *An aggregator of Analogies is a mapping* $V : \Psi \longrightarrow \mathbb{X}$*, where* $\mathbb{X}$ *is some ordered space.*

The value $V(\psi_f)$ represents the place that f takes in the ranking that our individual provides as a solution for today's problem DP_1. This is pretty intuitive: an aggregator takes into account how alternative f would fare with respect to each analogy, and then determines how f should be evaluated in the problem at hand.

In sum, the problem of our individual is to assign a rank to each alternative f in DP_1. He is going to do so by setting up analogies with problems DP_j that he solved in the past. This procedure is described by two mappings:

1. An AR operator

$$\mathcal{A} : f \longmapsto \psi_f$$

that associates each alternative f in DP_1 to a mapping ψ_f on $\mathcal{AP}$, the set of all problems analogous to DP_1 that have already been solved. The value of the mapping ψ_f at point $DP_j \in \mathcal{AP}$ expresses how alternative f fares if analogy DP_j is used.

2. An aggregator of the analogies

$$V : \psi_f \longmapsto V(\psi_f)$$

The value $V(\psi_f)$ represents the place that f takes in the ranking that our individual provides as a solution for DP_1.

Note that if DP_j is the only analogy, then

$$f \succsim_1 g \qquad iff \qquad A_j(f) \succsim_2 A_j(g)$$

in accordance to what was said in the single-analogy case. Formally, Definition 4. *We say that problem DP_1 is solved by analogy with a collection of problems $\{DP_j\}_{j\in J}$, $1\notin J$, if there exists an analogical reasoning (AR) operator*

$$\mathcal{A}: f \longmapsto \psi_f \qquad , \qquad f\in F_1$$

and an aggregator of analogies V (valued in some ordered space $\mathbb{X}$) such that

$$f \succsim_1 g \qquad iff \qquad V(\psi_f)\geq V(\psi_g)$$

Obviously, the remark at the end of subsection 2.1 applies here as well. Later, Definition 8 in Section 4 will substantiate the definition given here.

2.4. Analogical Reasoning Preferences. The next definition formalizes the idea of a preference that is formed by means of analogical reasoning.

Definition 5. *A preference relation $\succsim_1$ over a set of alternatives F_1 is an analogical reasoning (AR) preference if there exists a collection of decision problems $\{DP_j\}_{j\in J}$, $1\notin J$, such that $\succsim_1$ can be derived from $\{DP_j\}_{j\in J}$ as in Definition 4.*

While trivial, it is useful to notice that if DP_1 is solved by analogy with a collection of problems $\{DP_j\}_{j\in J}$, then the resulting preference $\succsim_1$ is representable by means of the functional $I: F_1 \longrightarrow \mathbb{X}$ defined by

$$I(f)=V(\psi_f)$$

where V is an aggregator of analogies.

3. Structures

By Definition 1, an analogy between two decision problems, $DP_1 = (S_1, X_1, F_1)$ and $DP_2 = (S_2, X_2, F_2)$, is a mapping $A: F_1 \longrightarrow F_2$ which preserves the problem's structure. It is now time to formalize the idea of a *structure of a decision problem.* As a matter of fact, an actual decision problem might display a lot of structure. This might have to do with the fact that the actions available are subject to certain restrictions, that the payoffs achievable must satisfy certain bounds or that some of the states have certain special features. A fruitful application to specific cases of the concept of analogical reasoning must take all of this into account. Here, however, I am interested in the problem in its generality. Thus, I will abstract from the specificity of each particular example, and focus on the structure that is common to all of them. What is more, I am interested in deriving representations of analogical reasoning that would compare, for their level of generality, to the Subjective Expected Utility (SEU) theorem of Anscombe and Aumann [**3**] or to the non-additive theories of Gilboa and Schmeidler [**10**], Schmeidler [**17**] and their extensions (see Gilboa and Marinacci [**9**] for a comprehensive survey).

Each decision problem DP_j (including DP_1) consists of ranking a set of alternatives, which are mappings $S_j \longrightarrow X_j$. I am going to restrict this setting by making the following assumptions.

R0: For each j, X_j is a mixture space (see [**3**], [**7**])

R1: For each j, there exists a linear utility $u_j : X_j \longrightarrow \mathbb{R}$ (Axioms on preferences guaranteeing the existence of such a utility are well-known; see, for instance, [**3**]).

We can use R1 to define a measurable structure on S_j, for each j. In fact, R1 produces an embedding of the set of alternatives F_j into the set of real-valued functions on S_j by means of the mapping $f_j \longmapsto u_j \circ f_j$, $f_j \in F_j$. Then, we can define a σ-algebra Σ_j on S_j as the coarsest σ-algebra which makes all the functions $\{u_j \circ f_j\}_{f_j \in F_j}$ measurable. Thus, for each j, we obtain the measurable space (S_j, Σ_j), and each alternative corresponds to a measurable real-valued function. In the remainder of the paper, I will identify the alternatives with the corresponding real-valued functions (i.e., I will write f_j in the place of $u_j \circ f_j$). I also assume that, for each j, the set of alternatives consists of *all* bounded Σ_j-measurable mappings $S_j \longrightarrow \mathbb{R}$. This set is denoted by $B(\Sigma_j)$.

Two comments are in order. First, the procedure I used to define a measurable structure on S_j is exactly the same procedure that is (often implicitly) used in any paper in decision making which uses an Anscombe-Aumann setting. Second, the assumption that the set of alternatives consists of the whole $B(\Sigma_j)$ is only a simplifying assumption. It serves solely to relate the present work to classical models of decision-making, all of which (in some form) make such an assumption (see, for instance, [**3**], [**6**], [**10**]).

Since for each decision problem DP_j, the set of alternatives F_j can be identified to the space $B(\Sigma_j)$, the fundamental structure of the problem DP_j is precisely that of $B(\Sigma_j)$. Clearly, this encodes S_j and, by means of u_j, X_j as well. In fact, this is precisely the minimal structure that is necessary to derive the SEU theorem of Anscombe and Aumann [**3**] and all other representation theorems mentioned above. In turn, the structure of $B(\Sigma_j)$ is identified by the facts that (i) $B(\Sigma_j)$ is a linear space; (ii) $B(\Sigma_j)$ consists of bounded measurable functions; and (iii) $B(\Sigma_j)$ has a partial order described by its positive cone. The definition of structure-preserving mapping follows automatically from these observations.

Definition 6. *Let (S_i, Σ_i) and (S_j, Σ_j) be two measurable spaces. A mapping $\kappa : B(\Sigma_i) \longrightarrow B(\Sigma_j)$ is structure-preserving if:*

(1) κ is linear;

(2) κ preserves the positive cone, i.e. $\kappa(B_+(\Sigma_i)) \subset B_+(\Sigma_j)$;

(3) κ is normal: $f_n \nearrow f \Longrightarrow \kappa(f_n) \nearrow \kappa(f), \qquad n \in \mathbb{N}$.

Conditions (1) and (2) are self-explanatory. Condition (2), in particular, is necessary in order to be able to talk about monotone preferences. Condition (3) is an important ingredient of the requirement that κ be structure-preserving. For every measurable space (S, Σ), every function in $B(\Sigma)$ is a limit from below ($\nearrow$) of (simple) measurable functions. Thus, a structure-preserving mapping must respect this property.

Mappings satisfying the conditions in Definition 6 are called *kernels*. Appendix A.1 contains a representation result for kernels. While the exposition is technical, it gives valuable insights into the problem of formalizing the idea of analogical reasoning. In particular, it explains why multiple-prior preferences on the same set of alternatives as the problem at hand will appear in the representation theorems of the next sections.

4. Analogy in decision making

Having unveiled the structure of decision problems as well as that of the mappings that preserve that structure (Appendix A.1), we can now turn to the task of substantiating the definitions given in Section 2. In order to ease the comparison, the present section is divided into subsections carrying the same headings as those of Section 2.

4.1. Single analogy. The task of substantiating Definition 1 is straightforward. By incorporating Definition 6, Definition 1 can now be re-stated as follows.

Definition 7. *Let DP_1 and DP_2 be two decision problems. Denote by $\succsim_i$ the ranking in problem $i, i = 1, 2$. We say that DP_1 is solved by analogy with DP_2 if there exists a kernel $A : B(\Sigma_1) \longrightarrow B(\Sigma_2)$ such that*

$$f \succsim_1 g \qquad iff \qquad A(f) \succsim_2 A(g)$$

for any two alternatives f, g in DP_1.

Notice that by means of this definition not only each alternative in DP_1 is associated to an alternative in DP_2, but also each state in DP_1 is associated to a state in DP_2 (by means of indicator functions) and each consequence in DP_1 is associated to a consequence in DP_2 (by means of constant functions).

4.2. Multiple analogies; analogical reasoning operators. The problem of re-stating Definition 4 in a way that would be comparable to Definition 7 is a bit more difficult. Clearly, it requires us to encode in Definition 4 a representation of the AR operators, that is of those mappings that are induced by a collection of structure-preserving mapping as in Section 2. This representation is obtained in Theorem 1 below, which is a consequence of the representation of kernels of Appendix A.1. Its content is quite remarkable: *A preference $\succsim_j$ in an analogous problem DP_j can always be represented as a multiple-prior preference on the same set of alternatives as the problem at hand*. It is worth stressing that this result does not require any assumption on the preferences $\succsim_j$ but it is solely a consequence of the idea that analogies should be represented by mappings that are structure-preserving.

The theorem is stated and proved under the additional assumption that the preferences $\succsim_j$ in all analogous problems admit a $\mathbb{R}$-valued representation. As it is clear from its proof (Appendix C), removing this additional assumption does not entail any conceptual complication but does require a cumbersome notation without adding much to the substance. Thus, from now on I am going to make the assumption that

> **R2:** For each $j \in J$, there exists an $I_j : F_j \longrightarrow \mathbb{R}$ which represents $\succsim_j$.

In the remainder of the paper, I will take care to distinguish the cases where assumption **R2** is made only for expositional convenience from those cases where it implies a substantial restriction.

Since the formulation of the theorem is quite compact, it is probably useful to have some heuristics precede its formal statement. Let us consider an AR preference on F_1 (Definition 5), which is obtained by aggregating a collection of analogies $\{DP_j\}_{j\in J}$. The AR operator associated with this preference takes each alternative $f \in F_1$ into the mapping $\psi_f : \mathcal{AP} \longrightarrow \mathbb{R}$. In turn, this mapping is completely described by its set of values $\{I_j(A_j(f))\}_{j\in J}$, where the functional $I_j : F_j \longrightarrow \mathbb{R}$ represents the preference $\succsim_j$ over F_j. Now, suppose that for each problem DP_j that is analogous to DP_1, there exists another decision problem $\widehat{DP}_j = (\hat{S}_j, \hat{X}_j, \hat{F}_j)$ that is analogous to

DP_j. Intuitively, we ought to be able to replace each DP_j with the corresponding $\widehat{DP}_j$without changing the resulting AR preference. Formally, if $\hat{A}_j : F_j \longrightarrow \hat{F}_j$ is the analogy between DP_j and $\widehat{DP}_j$and $\hat{I}_j$ is the functional that represents the preference in $\widehat{DP}_j$, we ought to be able to replace the function ψ_f with the function $\hat{\psi}_f$, which is defined on the collection $\{\widehat{DP}_j\}_{j\in J}$ and that at point $\widehat{DP}_j$ takes the value $\hat{\psi}_f(\widehat{DP}_j) = \hat{I}_j(\hat{A}_j(A_j(f)))$.

Next, let F be the set of alternatives in a decision problem and let us say that a preference relation $\succsim$ on F is a *multiple-prior preference relation* if, for every $f \in F$, the functional representing it has the form $I(f) = \hat{I}((\int f dP)_{P\in M})$, where M is a set of probabilities on the domain of f. That is, $\succsim$ is a multiple-prior preference if I is a function of a collection of expected utility evaluations. If the preference in $\widehat{DP}_j$ is a multiple-prior preference, then the functional $\hat{I}_j$ is of the form $\hat{I}\left((\int \varphi dP)_{P\in M_j}\right)$, where M_j is a set of probabilities. In such a case, by setting, $\mathcal{M} = \{M_j\}_{j\in J}$, we can express $\hat{\psi}_f$ directly as a function of M_j, that is we can write $\hat{\psi}_f(M_j) = \hat{I}\left(\left(\int \hat{A}_j(A_j(f))dP\right)_{P\in M_j}\right)$. Let Ψ' (resp. $\hat{\Psi}$) be the linear space generated by all the functions of the form ψ_f (resp. $\hat{\psi}_f$).

THEOREM 1. *Let* $\mathcal{A} : f \longmapsto \psi_f$ *be an AR operator. Then, there exists a collection of subsets of (finitely additive) probabilities on* (S_1, Σ_1), $\mathcal{M} = \{M_j\}_{j\in J}$, *such that*

(i) the mapping

$$\hat{T} : \psi_f \longmapsto \hat{\psi}_f$$

can be extended to a linear isomorphism of the spaces Ψ' *and* $\hat{\Psi}$*;*

(ii) the operator

$$\hat{A} = \hat{T} \circ \mathcal{A}$$

is an AR operator.

The first part of the theorem says three things. First, if $\{DP_j = (S_j, X_j, F_j)\}_{j\in J}$ is a collection of problems that are analogous to $DP_1 = (S_1, X_1, F_1)$, then each DP_j is analogous to a problem $\widehat{DP}_j$ of the same form as DP_1, that is for each $j \in J$, $\widehat{DP}_j$ is defined on (S_1, X_1, F_1). Second, for each $j \in J$, the preference in $\widehat{DP}_j$ is a multiple-prior preference. Third, in the notation used above, the collection $\{\widehat{DP}_j\}_{j\in J}$ can be chosen so that $\hat{A}_j(A_j(f)) = f$ for every $f \in F_1$. The second part of the theorem says that our intuition is correct: we can replace the collection $\{DP_j\}_{j\in J}$ with the collection $\{\widehat{DP}_j\}_{j\in J}$ without changing the resulting AR preference. In the remainder of the paper, I will often say that $\widehat{DP}_j$ (resp. the preference in $\hat{DP_j}$) is the *analogous representation* of DP_j (resp. of the preference $\succsim_j$).

Example 3 (maxmin preferences). *When the analogous representation of the preferences $\succsim_j$ are all of the type maxmin, the functional $\hat{I}_j$ is given by*

$$I_j(A_j(f)) = \hat{I}_j\left(\left(\int fdP\right)_{P\in M_j}\right) = \min_{P\in M_j}\int fdP$$

Consequently, the function $\hat{\psi}_f : \mathcal{M} \longrightarrow \mathbb{R}$ is defined by

$$\hat{\psi}_f(M_j) = \min_{P\in M_j}\int fdP$$

Then, the value of alternative $f \in F_1$ in the AR preference is the result of aggregating the values taken by the function $\hat{\psi}_f$. This can be done, for instance, by using a weighting function.

Example 4 (SEU preferences). *When the analogous representation of the preferences $\succsim_j$ are all of the type SEU, both functionals I_j and $\hat{I}_j$ are linear. Thus, the function $\hat{\psi}_f : \mathcal{M} \longrightarrow \mathbb{R}$ can be written*

$$\hat{\psi}_f(M_j) = \int\limits_{ba_1^+(\Sigma_1)}\int\limits_{S_1} fdP\, d\Gamma_j$$

where Γ_j is a probability charge on $ba_1^+(\Sigma_1)$ and M_j =supp(Γ_j) (see App. C). In this case, it is intuitively clear that the problem of aggregating the preferences $\succsim_j$ is the same as that of aggregating the probability charges Γ_j. Formally, this corresponds to representing the set of analogies by means of the set of probability charges $\Gamma = \{\Gamma_j\}_{j\in J}$ on (S_1, Σ_1). In this new representation, the functions $\hat{\psi}_f$ are replaced by the functions $\tilde{\psi}_f : \Gamma \longrightarrow \mathbb{R}$ given by

$$\tilde{\psi}_f(\Gamma_j) = \int\limits_{ba_1^+(\Sigma_1)}\int\limits_{S_1} fdP\, d\Gamma_j$$

Later, we are going to see that a representation of this type is a special case of the more general one of Theorem 3.

4.2.1. *Monotonic, $\mathbb{R}$-valued preferences.* At this point of the exposition, it is useful to notice that the representation provided by Theorem 1 can be quite strengthened if we assume, in addition, that the preferences $\succsim_j$ are all monotonic (i.e., $f \geq g$ implies $f \succsim_j g$). This is an important property that is shared by most preferences that appear both in theoretical works and in applications. Preferences that are monotonic, representable by a $\mathbb{R}$-valued functional and that, in addition, satisfy certain continuity properties are studied in [**8**].

THEOREM 2. *Let $\mathcal{A} : f \longmapsto \psi_f$ be an AR operator, and assume that all the preferences $\succsim_j$ are monotonic. Then, there exists a collection of subsets of probability charges on (S_1, Σ_1), $\bar{\mathcal{M}} = \{\bar{M}_j\}_{j \in J}$, such that*

(i) For each $j \in J$, $\bar{M}_j$ is convex and weak-compact;*

(ii) Let $\{A_j\}_{j \in J}$ denote the family of kernels $A_j : B(\Sigma_1) \longrightarrow B(\Sigma_j)$ which induces $\mathcal{A}$. Then, for each $j \in J$, $A_j(B(\Sigma_1))$ is linearly isomorphic to $AF(\bar{M}_j)$, the space of weak-continuous affine functions on $\bar{M}_j$.*

(iii) For $f \in F_1$, let $\bar{\psi}_f : \bar{\mathcal{M}} \longrightarrow \mathbb{R}$ be defined by $\bar{\psi}_f(\bar{M}_j) = I_j(A_j(f))$, and let $\bar{\Psi}$ be the linear space generated by all such functions. Then, the mapping

$$\bar{T} : \psi_f \longmapsto \bar{\psi}_f$$

can be extended to a linear isomorphism of the spaces Ψ' and $\bar{\Psi}$;

(iv) the operator

$$\bar{A} = \bar{T} \circ \mathcal{A}$$

is an AR operator.

Clearly, the meaning of Theorem 2 is exactly the same as Theorem 1. The value added is twofold. First, the set of analogies is represented as a subset of the collection of convex, weak*-compact subsets of probability charges on Σ_1, which is a mathematical object with very nice properties. Second, the mappings appearing in the representation have some important continuity properties, which will be used in subsequent theorems.

4.3. Aggregators of Analogies. In the Introduction, I observed that the pervasive nature of Analogy as a process of human thinking dictates that a theory of Analogy should rest on very minimal assumptions. Here, I am going to make two assumptions which will restrict the class of admissible aggregators. I believe that these assumptions are somehow unavoidable because they stem directly from the nature of the problem. The first assumption, which is essentially embedded in the definition of AR preference (Definition 5), expresses the idea that an AR preference should depend only on the analogies and on the solutions provided for those. As a consequence of Theorem 1, this is equivalent to the condition that an AR preference should depend only on the preferences to be aggregated and not on the particular representation that one gives of those. This is a familiar requirement in economic theory. It appears, for instance, in the Theory of Social Choice where the preferences to be aggregated are those of different individuals and the aggregator is a Social Choice function. The second assumption is a monotonicity-type condition: if $f \in F_1$ performs better than $g \in F_1$ according to each and every analogy, then f should be preferred to g.

I will formalize these assumptions under the additional requirement that the aggregators are $\mathbb{R}$-valued. Strictly speaking, this is restrictive because it forces the resulting AR preference to be Archimedean. Just like before, however, this restriction could be removed (for instance by using the methods in [**1**]) but at the price of a rather involved notation. Let $\hat{\Psi}$ be as in Theorem 1, and let $V : \hat{\Psi} \longrightarrow \mathbb{R}$. Then, V is an *aggregator of analogies* if it satisfies the following properties:

AA0: V is $\mathbb{R}$-valued.

AA1 (Invariance): Let $\hat{I}_j$ be a preference functional for the preference $\succsim_j$, $j \in J$ and $j \neq 1$. Suppose that $\tau : \mathbb{R} \longrightarrow \mathbb{R}$ has the property that for each $j \in J$, $j \neq 1$, $\tau \circ I_j$ is a preference functional for $\succsim_j$. Then $\forall f \in F_1$

$$V(\tau \circ \hat{\psi}_f) = \tau \circ V(\hat{\psi}_f)$$

AA2 (Monotonicity): $\forall f, g$

$$\hat{\psi}_f \geq \hat{\psi}_g \qquad \Longrightarrow \qquad V(\hat{\psi}_f) \geq V(\hat{\psi}_g)$$

4.4. Analogical Reasoning Preferences. One of the consequences of Theorem 1 is that any AR preference can always be thought of as the outcome of aggregating a collection of multiple-prior preferences over the same set of alternatives as the problem at hand. This result is encoded in the next definition which, thus, parallels Definition 4 just like Definition 7 parallels Definition 1. Let $\hat{\psi}_f$ and $\hat{\Psi}$ be as in Theorem 1.

Definition 8. *A preference relation $\succsim_1$ on F_1 is an AR preference if there exist (i) a collection $\{\succsim_j\}_{j \in J}$, $j \neq 1$, of multiple-prior preferences on F_1; (ii) an AR operator $\hat{A} : f \longmapsto \hat{\psi}_f$ and (iii) a functional $V : \hat{A}(F_1) \longrightarrow \mathbb{X}$, $\mathbb{X}$ some ordered space, such that*

$$f \succsim_1 g \qquad iff \qquad V(\hat{\psi}_f) \geq V(\hat{\psi}_g)$$

Just like we observed in subsection 2.4, AR preferences are, by construction, representable by means of the functional $I : F_1 \longrightarrow \mathbb{X}$ defined by

$$I(f) = V(\psi_f)$$

5. Analogical reasoning with Invariant Bi-separable preferences

So far, I have imposed very few restrictions on the preferences $\succsim_j$ which appear in the definition of AR preference, and have done so mainly for expositional reasons. In contrast, now I am going to make a substantial assumption: I am going to assume that all the preferences $\succsim_j$ admit an analogous representation that is Invariant Bi-separable (IB). I am going to do so for three reasons. First, I want to give an example of how to apply the concepts introduced thus far and of the theorems that can be proved within specific settings. The reader interested in other classes of preferences should have no trouble to adapt the methods presented here. Second, the class of IB preferences is still a very wide class of preferences. In fact, IB preferences, first introduced by Ghirardato et al. in [**6**], have to satisfy only very mild conditions: (i) they are representable by a $\mathbb{R}$-valued functional; (ii) they are constant-independent, and (iii) monotonic. In particular, IB preferences need not be either concave or convex. For the reader's convenience, the five axioms defining IB preferences are reported in Appendix B. Finally, as we shall see in Section 6, IB preferences have an important stability property in relation to analogical reasoning. This alone justifies the special attention given to IB preferences in this paper. In the remainder of the paper, I will often use the shorthand formulation "if an AR preference results from the aggregation of preferences with property Y" in the place of the proper "if an AR preference results from the aggregation of preferences whose analogous representation in Theorem 1 has property Y".

IB preferences are monotonic and representable by a $\mathbb{R}$-valued functional. Thus, Theorem 2 applies. The first result about IB preferences is Theorem 3 below, which refines Theorem 2. It says that *if all the preferences $\succsim_j$ are IB, then analogies can be represented by capacities.*

THEOREM 3. *Let $\mathcal{A} : f \longmapsto \psi_f$ be an AR operator, and assume that all the preferences $\succsim_j$ are IB. Then, there exists a collection of capacities on the Borel sets of $ba_1^+(\Sigma_1)$, $\tilde{\Gamma} = \{\gamma_j\}_{j \in J}$, such that* *(i) For* $f \in F_1$*, let* $\tilde{\psi}_f : \tilde{\Gamma} \longrightarrow \mathbb{R}$ *be defined by*

$$\tilde{\psi}_f(\gamma_j) = \int_{ba_1^+(\Sigma_1)} \int_{S_1} f dP d\gamma_j$$

and let $\tilde{\Psi}$ *be the linear space generated by all such functions. Then, the mapping*

$$\tilde{T} : \psi_f \longmapsto \tilde{\psi}_f$$

can be extended to a linear isomorphism of the spaces Ψ' *and* $\tilde{\Psi}$*;*

(ii) the operator

$$\tilde{A} = \tilde{T} \circ \mathcal{A}$$

is an AR operator.

In [**2**, Theorem 2], I showed that any IB preference functional can be represented in the form

$$I_j(\varsigma) = \int_{ba_1^+(\Sigma_1)} \int_{S_1} \varsigma dP d\gamma_j$$

where γ_j is a capacity on the Borel subset of $ba_1^+(\Sigma_1)$ (weak*-topology) and the outer integral is taken in the sense of Choquet. Then, following a procedure similar to that of Example 4, the representation in Theorem 3 obtains by replacing the functions $\hat{\psi}_f$ in Theorem 2 with the functions $\tilde{\psi}_f$ defined on the set of capacities representing the IB preferences $\succsim_j$. In fact, Example 4 is a special case of this representation, which obtains when all $\succsim_j$ preferences are SEU and, consequently, all the capacities γ_j are probability charges.

When all $\succsim_j$ preferences are IB, the requirements on the aggregators of analogies take the following form:

AA0-IB: V is $\mathbb{R}$-valued.

AA1-IB (Translation Invariance): For any positive affine transformation $a : \mathbb{R} \longrightarrow \mathbb{R}$,

$$a \circ V(\tilde{\psi}_f) = V(a \circ \tilde{\psi}_f)$$

for any $f \in B(\Sigma_1)$.

AA2 (Monotonicity): $\tilde{\psi}_f \geq \tilde{\psi}_g \Longrightarrow V(\tilde{\psi}_f) \geq V(\tilde{\psi}_g)$.

Below, I give some examples of AR preferences which result from aggregating IB preferences. The set Ξ of all capacities on $ba_1^+(\Sigma_1)$ is endowed with the coarsest topology such that all the functions $\Xi \longrightarrow \mathbb{R}$ of the form

$L(\gamma) = \int_{ba_1^+(\Sigma_1)} Z(P)d\gamma(P)$, where $Z(\cdot)$ is a bounded Borel function on $ba_1^+(\Sigma_1)$, are continuous.

EXAMPLE 5 (Choquet Integrals). *Choquet integrals are* $\mathbb{R}$*-valued, translation invariant and monotone. Thus, the functional* $V : \tilde{\Psi} \longrightarrow \mathbb{R}$

$$V(\tilde{\psi}_f) = \int_{\Xi} \tilde{\psi}_f d\Phi$$

where Φ *is a capacity on the Borel subsets of* Ξ*, is an aggregator of analogies. By construction, the resulting AR preference is represented by the functional* $I : F_1 \longrightarrow \mathbb{R}$ *defined by*

$$I(f) = \int_{\Xi} \tilde{\psi}_f d\Phi = \int_{\Xi} \int_{ba_1^+(\Sigma_1)} \int_{S_1} f dP d\gamma d\Phi$$

where the two outer integrals are taken in the sense of Choquet. That is,

(5.1)

$$f \succsim_1 g \qquad iff \qquad \int_{\Xi} \int_{ba_1^+(\Sigma_1)} \int_{S_1} f dP d\gamma d\Phi \geq \int_{\Xi} \int_{ba_1^+(\Sigma_1)} \int_{S_1} g dP d\gamma d\Phi$$

These preferences appear to be an entirely new object as they are defined by means of a Choquet integration over capacities which, in turn, are defined over sets of measures. In the next section, we will see that this is just a different representation of a familiar object.

EXAMPLE 6 (Quantiles; Generalized quantiles). *If* V *is either a probabilistic quantile or a generalized quantile (i.e., monotone and ordinally covariant functional* [**5**]*), then the resulting preference is an AR preference. These are special cases of the previous example as both probabilistic quantiles and generalized quantiles can be represented as Choquet integrals as established in* [**5**].

EXAMPLE 7 (Gilboa-Schmeidler Similarity). *Lebesgue integrals are* $\mathbb{R}$*-valued, translation invariant and monotone. Thus, preferences on* F_1 *of the same form as in (5.1) but with* Φ *being a probability charge rather than a capacity are also AR preferences. The transformation of the similarity function of Gilboa and Schmeidler is an example of an aggregator which is a Lebesgue integral. AR preferences obtained through a Gilboa-Schmeidler aggregator have an important stability property. In fact, this is a special case of Theorem 4, but for Gilboa-Schmeidler aggregators a more direct argument leads to this conclusion.*

PROPOSITION 1. *AR preferences obtained through Gilboa-Schmeidler aggregators are IB preferences.*

EXAMPLE 8 (SEU preferences with fuzzy weighting). *Other notable examples of AR preferences obtain when all the preferences* $\succsim_j$ *are SEU, but the aggregator weights them in a non-linear fashion, for instance like in a Choquet integral. A possible interpretation is that the decision maker is not sure that the problems* $\{DP_j\}$ *are analogous to the problem at hand. Again, a direct argument shows that these AR preferences also have the stability property above.*

PROPOSITION 2. *Assume that all the preferences* $\succsim_j$ *are SEU and that the aggregator is a Choquet integral. Then, the resulting AR preference is IBP.*

EXAMPLE 9 (maxmin preferences). *A special case of the previous example obtains when the Choquet integral is concave (convex). In such a case, by a classic theorem of Schmeidler* [**16**]*, the resulting AR preference is a maxmin (maxmax) preference.*

EXAMPLE 10 (unstable preferences). *Maxmin (or maxmax) preferences are not stable in the sense that an AR preference which results from aggregating a collection* $\{\succsim_j\}_{j\in J}$ *of maxmin preferences need not be a maxmin preference. This follows from the easy, and well-known, observation that an aggregation of concave preferences need not be concave.*

6. The stability property of Monotone Continuous IB preferences

Let $\succsim_1$ be an AR preference on F_1. In Theorem 1, we saw that the preferences $\{\succsim_j\}_{j\in J}$ in all analogous problems can always be represented as multiple-prior preferences on F_1. As a consequence of the representation of kernels (Appendix A.1), the set of priors M_j associated with the preference $\succsim_j$ consists of countably additive probabilities. When the preference $\succsim_j$ is monotone, we know from Theorem 2 that we can replace M_j with its weak*-closed convex hull $\bar{M}_j$. In general, this closure might contain some charges that are not countably additive.

An IB preference satisfies the Axiom of Monotone Continuity (see Appendix B) if and only if all the priors in the representation are countably additive [**6**, Sec. B.3]. If an AR preference results from aggregating a collection $\{\succsim_j\}_{j\in J}$ of IB preferences, then, just like in the general case, the set of charges associated with each $\succsim_j$ might contain some charges that are not countably additive. We will call *Almost Monotone Continuous* all those multiple-prior preferences whose set of charges is the weak*-closure of countably additive probabilities. It is clear that any Monotone Continuous IB preference is also Almost Monotone Continuous. Theorem 4 below shows that class of Monotone Continuous IB preferences (MCIB) is stable under analogical reasoning.

THEOREM 4. *Every AR preference resulting from the aggregation of MCIB preferences is a MCIB preference. Conversely, any MCIB preference can always be represented as an AR preference which aggregates a collection of MCIB preferences.*

A few remarks are in order.

REMARK 1. *The proof of Theorem 4 (Appendix C) shows a little more than what stated: every Almost Monotone Continuous IB preference is an AR preference.*

REMARK 2. *The second part of the proof of Theorem 4 shows that we can always think of a MCIB as the result of aggregating analogies that are represented by SEU preference. However, this is only a device in the proof and does not have to be taken literally. It is always possible to represent a MCIB preference as an aggregation of a family of MICB preferences that are not SEU.*

REMARK 3. *In Example 5, we encountered an AR preference which was represented by means of a triple integral (a Choquet integration over capacities defined over set of measures). By Theorem 4 (and Remark 1), that preference is an IB preference. By* [**2**, Theorem 2]*, it can be represented by the functional*

$$I(f) = \int_{ba_1^+(\Sigma_1)} \int_{S_1} f dP d\gamma^* \qquad \forall f \in B(\Sigma_1)$$

for some capacity γ^* on the Borel subsets of $ba_1^+(\Sigma_1)$.

EXAMPLE 11 (The Barycenter of a Capacity). *From the previous remark it follows that, given a set Θ capacities on $ba_1^+(\Sigma_1)$ and a capacity Φ on Borel sets of Θ (for the topology defined above), there always exists a capacity γ^**

on the Borel subsets of $ba_1^+(\Sigma_1)$ *such that* $\forall f \in B(\Sigma_1)$

$$\int_\Theta \int_{ba_1^+(\Sigma_1)} \int_{S_1} f dP d\gamma d\Phi = \int_{ba_1^+(\Sigma_1)} \int_{S_1} f dP d\gamma^*$$

For Θ *compact and convex,* $\gamma^* \in \Theta$*;* γ^* *can be interpreted as the barycenter of* Φ *on the set* Θ*.*

7. Summary

In this paper, after some preparatory work, we reached a formal definition of single analogy. Building on that, we then defined analogical reasoning preferences as those that aggregate possibly many analogies. An AR preference can always be thought of as the outcome of aggregating a collection of multiple-prior preferences defined on the same set of alternatives as the problem at hand. It is important to stress that this is only an outcome of the representation theorem: *analogous problems can be defined on any space; it is Theorem 1 that allows us to represent them on the same state space as the problem at hand*. We refined this result first in the case of monotone preferences and then in the case of IB preferences. We then introduced the requirements that any aggregator of analogies should satisfy, and gave several examples of these aggregators and of the corresponding AR preferences. Finally, we showed that MCIB preferences have an important stability property in relation to analogical reasoning

APPENDIX A
BACKGROUND MATERIAL

A.1 Kernels and their representation

This section contains the basic facts about the representation of kernels (see [**13**], for more on kernels). Let (S, Σ) and (T, Υ) be measurable spaces, let $ba(\Sigma)$ denote the space of bounded charges on Σ and let $\rho : B(\Sigma) \longrightarrow B(\Upsilon)$ be a kernel. By using ρ, we can define a mapping $T \longrightarrow ba(\Sigma)$ in the following way: to the element $t \in T$ we associate the charge $\mu^t \in ba(\Sigma)$ defined by the equation

$$\mu^t(A) = (\rho(\chi_A))(t), \qquad for \quad every \quad A \in \Sigma \tag{7.1}$$

where χ_A denotes the indicator function of the set $A \in \Sigma$. Notice that since ρ is a kernel, then μ^t is a positive charge (by property (2) in Definition 6) and is countably additive (by property (3), Definition 6). By using the fact that every $f \in B(\Sigma)$ is a limit from below of measurable simple functions, it is easily seen that equation (7.1) along with properties (1), (2) and (3) of ρ imply that

$$\rho(f)(t) = \mu^t(f) \equiv \int_S f d\mu^t \tag{7.2}$$

that is, a kernel ρ sends the function $f \in B(\Sigma)$ into the function $\rho(f) \in B(\Upsilon)$ which is defined by $\rho(f)(t) = \int_S f d\mu^t$. Conversely, let a mapping $t \longmapsto \mu^t$ be given, $t \in T$ and $\mu^t \in ba(\Sigma)$. Let $\varphi : B(\Sigma) \longrightarrow \mathbb{R}^T$ be defined by $f \longmapsto \varphi(f)$, where $\varphi(f)$ is the function $T \longrightarrow \mathbb{R}$ which at point $t \in T$ takes the value $\int_S f d\mu^t$. Then, if $\varphi(f) \in B(\Upsilon)$ for all $f \in B(\Sigma)$ and if all the measures $\{\mu^t\}_{t \in T}$ are positive and countably additive, then φ is a kernel. In fact, φ is clearly a linear mapping. If $\varphi(f) \in B(\Upsilon)$ for all $f \in B(\Sigma)$, then $\varphi : B(\Sigma) \longrightarrow B(\Upsilon)$. If all the measures $\{\mu^t\}_{t \in T}$ are positive, then φ satisfies property (2) in Definition 6; and if all the measures $\{\mu^t\}_{t \in T}$ are countably additive, then φ satisfies property (3). Notice that countable additivity of the measures is necessary because we need the dominated convergence theorem to hold in order to ensure normality.

Summing up, given two measurable spaces (S, Σ) and (T, Υ), a kernel $\rho : B(\Sigma) \longrightarrow B(\Upsilon)$ can always be represented as a mapping that sends the function $f \in B(\Sigma)$ into the function $\rho(f) \in B(\Upsilon)$ which is defined by

$\rho(f)(t) = \int_S f d\mu^t$, where μ^t is a positive, countably additive measure on Σ. Notice, in particular, that each kernel $B(\Sigma) \longrightarrow B(\Upsilon)$ is automatically associated to a set of measures $\{\mu^t\}_{t\in T}$.

A.2 The pushforward of a capacity

The concept of pushforward of a capacity, introduced below, and the property in the next Proposition are used in the proof of Proposition 2. Let (D, Δ) and (T, Θ) be measurable spaces and let $W : D \longrightarrow T$ be measurable.

Definition 9. *Let $\mathcal{K}$ be a capacity on Δ. The pushforward of $\mathcal{K}$ under W is the capacity $W_*\mathcal{K}$ defined by*

$$W_*\mathcal{K}(E) = \mathcal{K}(W^{-1}(E))$$

for all $E \in \Delta$.

It's easy to see that $W_*\mathcal{K}$ is, indeed, a capacity on Θ.

Proposition 3. *For every measurable $f : T \longrightarrow \mathbb{R}$, $f \circ W$ is measurable on (D, Δ) and*

$$\int_D f \circ W d\mathcal{K} = \int_T f dW_*\mathcal{K} \qquad (7.3)$$

where the integrals are taken in the sense of Choquet.

Proof. The measurability of $f \circ W$ is obvious. It suffices to prove the equality (7.3) for simple functions. So let $f : T \longrightarrow \mathbb{R}$ be simple, $f = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i}$, $A_i \in \Delta$, where the A_i's have been indexed so that $a_i > a_{i+1}$. Notice that $f \circ W(d) = a_i$ iff $d \in W^{-1}(A_i)$. We have

$$\begin{aligned}
\int_T f dW_*\mathcal{K} &= \sum_{i=1}^{n} (a_i - a_{i+1}) W_*\mathcal{K}(\bigcup_{j=1}^{i} A_j) \\
&= \sum_{i=1}^{n} (a_i - a_{i+1}) \mathcal{K}(T^{-1}(\bigcup_{j=1}^{i} A_j)) \\
&= \sum_{i=1}^{n} (a_i - a_{i+1}) \mathcal{K}(\bigcup_{j=1}^{i} T^{-1}(A_j)) \\
&= \int_D f \circ W d\mathcal{K}
\end{aligned}$$

□

APPENDIX B
INVARIANT BI-SEPARABLE PREFERENCES

Let F be the collection of all simple mappings $S \longrightarrow X$, X is a mixture space. A preference relation $\succsim$ on F is Invariant Bi-separable [**6**] if it satisfies the following five axioms.

A1 $\succsim$ is complete and transitive.

A2 (C-independence) For all $f, g \in F$, for all constant mappings $h \in F$ and for all $\alpha \in (0, 1)$

$$f \succ g \qquad \Longleftrightarrow \qquad \alpha f + (1-\alpha)h \succ \alpha g + (1-\alpha)h$$

where $\succ$ is the asymmetric part of $\succsim$.

A3 (Archimedean property) For all $f, g, h \in F$, if $f \succ g$ and $g \succ h$ then $\exists \alpha, \beta \in (0, 1)$ such that $\alpha f + (1-\alpha)h \succ g$ and $g \succ \beta f + (1-\beta)h$.

A4 (Monotonicity) For all $f, g \in F$, $f(s) \succsim g(s)$ for any $s \in S \Longrightarrow f \succsim g$.

A5 (Non-degeneracy) $\exists x, y \in X$ such that $x \succ y$.

Let $\succsim$ be a preference relation satisfying axioms 1 to 5. Let $\succsim^*$ denote the unambiguous preference relation associated to $\succsim$ ([**6**], Sec. B.3).

Axiom of Monotone Continuity (see [6]): For all $x, y, z \in X$ such that $y \succ^* z$, and all sequences of events $\{A_n\}_{n \geq 1} \subseteq \Sigma$ with $A_n \downarrow \emptyset$, there exists $\bar{n} \in \mathbb{N}$ such that $y \succ^* x A_{\bar{n}} z$.

The Axiom of Monotone Continuity is equivalent to the property that all the measures in the representation of an IB preference are countably additive ([**6**], Sec. B.3).

APPENDIX C
OMITTED PROOFS

All the kernels appearing in the proofs are such that all associated measures are probabilities. This is without loss of generality. It simply implies that the function $u_1 \circ f : S_1 \longrightarrow \mathbb{R}$ in decision problem $DP_1 = (S_1, X_1, F_1)$ which is identically equal to 1 on S_1 would be evaluated by the number 1 in each and every analogy.

Section 4

Recall that the set Ψ (subsection 2.3) is the set of all mappings of the form ψ_f for some $f \in F_1$. By assumption **R2**, these mappings are $\mathbb{R}$-valued. We are going to be viewing Ψ as a subset of the Banach space (supnorm) $B(\mathcal{AP})$ of bounded, measurable functions on $\mathcal{AP}$, where $\mathcal{AP}$ is endowed with the coarsest σ-algebra which makes all the mappings ψ_f measurable. Let $ba_1^+(\Sigma_1)$ denote the space of probability charges on (S_1, Σ_1) and let $\mathcal{M} = \{M_j\}_{j\in J}$ be a collection of subsets of $ba_1^+(\Sigma_1)$, that is $M_j \subset ba_1^+(\Sigma_1)$. For each $f \in F_1$, define $\hat{\psi}_f : \mathcal{M} \longrightarrow \mathbb{R}$ by $\hat{\psi}_f(M_j) = I_j(A_j(f))$. Each $\hat{\psi}_f$ is an element of of the Banach space (supnorm) $B(\mathcal{M})$ of bounded, measurable functions on $\mathcal{M}$, where $\mathcal{M}$ is endowed with the coarsest σ-algebra which makes all the mappings $\hat{\psi}_f$ measurable. Finally, let Ψ' denote the closed linear subspace generated by Ψ (in $B(\mathcal{AP})$) and let $\hat{\Psi}$ be the linear subspace (in $B(\mathcal{M})$) generated by all the functions of the form $\hat{\psi}_f$.

Proof of Theorem 1. A mapping $\mathcal{A} : f \longmapsto \psi_f$ is an AR operator if it is induced by a family $\{A_j\}_{j\in J}$ of kernels $A_j : B(\Sigma_1) \longrightarrow B(\Sigma_j)$. In Appendix A.1, we saw that each kernel A_j is associated with a family of probability charges $M_j \subset ba_1^+(\Sigma_1)$. As a consequence of the representation of kernels in Appendix A.1, the function $A_j(f) \in B(\Sigma_j)$ is represented by the function $\rho_j(f) : M_j \longrightarrow \mathbb{R}$ defined by $\rho_j(f)(P) = \int f dP$, $P \in M_j$. Let $\hat{A}_j(B(\Sigma_1))$ denote the set of all functions of the form $\rho_j(f)$, $f \in B(\Sigma_1)$, and for each $j \in J$, define $\hat{I}_j : \hat{A}_j(B(\Sigma_1)) \longrightarrow \mathbb{R}$ by $\hat{I}_j(\rho_j(f)) = I_j(A_j(f)$. Now, let $\mathcal{M} = \{M_j\}_{j\in J}$ be the collection of all M_j, and let $\hat{\psi}_f : \mathcal{M} \longrightarrow \mathbb{R}$ be defined by $\hat{\psi}_f(M_j) = I_j(A_j(f)) = \hat{I}_j(\rho_j(f))$. Clearly, the mapping

$$\hat{T} : \psi_f \longmapsto \hat{\psi}_f$$

is one-to-one and can be trivially extended to a linear isomorphism of the spaces Ψ and $\hat{\Psi}$. For the second part, notice that, as a consequence of the representation of Appendix A.1, the mapping $A_j(B(\Sigma_1)) \longrightarrow \hat{A}_j(B(\Sigma_1))$ defined by $A_j(f) \longmapsto \rho_j(f)$ is a kernel. Thus, the operator $\hat{A} = \hat{T} \circ \mathcal{A} : f \longmapsto \hat{\psi}_f$ is induced by a family of kernels and is, therefore, an AR operator. □

Proof of Theorem 2. (i) and (iii): Endow $ba(\Sigma_1)$, the space of charges on Σ_1, with the weak*-topology produced by the duality $(B(\Sigma_1), ba(\Sigma_1))$. Next, let $\hat{\psi}_f(M_j) = I_j(A_j(f)) = \hat{I}_j(\rho_j(f))$ be as in Theorem 1. Since $\succsim_j$ is monotone both I_j and $\hat{I}_j$ are monotone functionals. Thus, for every $f \in F_1$ there exists $\alpha(f) \in [0,1]$ such that

$$\hat{I}_j(\rho_j(f)) = \alpha(f) \inf \rho_j(f) + (1-\alpha(f)) \sup \rho_j(f) = \alpha(f) \inf_{P \in M_j} \int f dP + (1-\alpha(f)) \sup_{P \in M_j} \int f dP \tag{7.4}$$

For each $j \in J$, let $\bar{M}_j = \overline{co}(M_j)$ denote the weak*-closed convex hull of M_j. Clearly,

$$\inf_{P \in M_j} \int f dP = \min_{P \in \bar{M}_j} \int f dP$$
$$\sup_{P \in M_j} \int f dP = \max_{P \in \bar{M}_j} \int f dP$$

Let $\bar{\mathcal{M}} = \{\bar{M}_j\}_{j \in J}$, and define $\bar{\psi}_f : \bar{\mathcal{M}} \longrightarrow \mathbb{R}$ by

$$\bar{\psi}_f(\bar{M}_j) = \alpha(f) \min_{P \in \bar{M}_j} \int f dP + (1-\alpha(f)) \max_{P \in \bar{M}_j} \int f dP$$

where the function $\alpha(\cdot)$ is the same as that appearing in (7.4). Clearly, the mapping

$$\bar{T} : \psi_f \longmapsto \bar{\psi}_f$$

can be extended to a linear isomorphism of the associated linear spaces.

(ii) We had already noticed in the proof of Theorem 1 that the mapping $A_j(f) \longmapsto \rho_j(f)$ is a kernel. Then, it suffices to notice that the same is trivially true for the mapping $A_j(f) \longmapsto \bar{\rho}_j(f)$, where $\bar{\rho}_j(f) : \bar{M}_j \longrightarrow \mathbb{R}$ is given by $\bar{\rho}_j(f)(P) = \int f dP$. Each of such functions is a weak*-continuous affine function and, in turn, any weak*-continuous affine function has this form for some $f \in F_1$.

(iv) Follows exactly as in Theorem 1. □

Section 5

Proof of Theorem 3. Since IB preferences are both monotonic and representable by a $\mathbb{R}$-valued functional, Theorem 2 applies. By [**2**, Theorem 2], an IB preference functional can always be represented in the form

$$I_j(\varsigma) = \int_{ba_1^+(\Sigma_1)} \int_{S_1} \varsigma dP d\gamma_j(P)$$

where γ_j is a capacity on the Borel subset of $ba_1^+(\Sigma_1)$ (weak*-topology) and the outer integral is taken in the sense of Choquet. By using this and by defining $\tilde{\psi}_f$ like in the theorem, (i) and (ii) follow just like in the previous representation theorems. □

PROOF OF PROPOSITION 1. We already know that, since SEU preferences are monotone, we can assume that the support of Φ is a convex (and compact in the topology defined above) subset of Ξ. Moreover, it is easy to see that for any bounded measurable function $Z(p)$ on $ba_1^+(\Sigma_1)$, the function $\xi : \Xi \longrightarrow \mathbb{R}$ defined by $\xi(\gamma) = \int_{ba_1^+(\Sigma_1)} Z(P)d\gamma$ is affine on suppΦ. By [**4**, Proposition 1.5], Φ has a barycenter. Thus, in particular, there exists $\gamma^* \in$suppΦ with the property that $\forall f \in F_1$

$$\int_\Xi \int_{ba_1^+(\Sigma_1)} \int_{S_1} fdPd\gamma d\Phi = \int_{ba_1^+(\Sigma_1)} \int_{S_1} fdPd\gamma^*$$

Hence, an AR preference generated by a Gilboa-Schmeidler type of aggregator can also be represented by the functional

$$f \succsim_1 g \qquad iff \qquad \int_{ba_1^+(\Sigma_1)} \int_{S_1} fdPd\gamma^* \geq \int_{ba_1^+(\Sigma_1)} \int_{S_1} gdPd\gamma^*$$

By [**2**, Theorem 2], this shows that these preferences are IB. □

PROOF OF PROPOSITION 2. When all $\succsim_j$ preferences are SEU, the capacities γ in (5.1) are probability charges. Since, for each γ, suppγ is convex and weak*-compact, each γ has a barycenter in its support. That is, for each γ there exists a P_γ^* such that $\forall f \in F_1$

$$\int_{ba_1^+(\Sigma_1)} \int_{S_1} fdPd\gamma = \int_{S_1} fdP_\gamma^*$$

Let Φ be the capacity in (5.1) and let $H : Supp\Phi \longrightarrow ba_1^+(\Sigma_1)$ denote the mapping which associates each $\gamma \in Supp\Phi$ with its barycenter P_γ^*. Let $H_*\Phi$ be the pushforward of Φ under H. Then (see Appendix A.2), we have that $\forall f \in F_1$

$$\int_\Xi \int_{ba_1^+(\Sigma_1)} \int_{S_1} fdPd\gamma d\Phi = \int_{ba_1^+(\Sigma_1)} \int_{S_1} fdP_\gamma^* dH_*\Phi$$

which shows, by [**2**, Theorem 2], that these preferences are IB. □

Section 6

Proof of Theorem 4. Let $\succsim_1$ be an AR preference on $F_1 = B(\Sigma_1)$. Then, $\succsim_1$ is represented by the functional $\tilde{V} : F_1 \longrightarrow \mathbb{R}$ defined by

$$\tilde{V}(f) = V(\mathcal{A}(f)) = V(\psi_f)$$

where $\mathcal{A} : f \longmapsto \psi_f$ is an AR operator. Since $\succsim_1$ results from aggregating a collection $\{\succsim_j\}_{j\in J}$ of MCIB preferences, the function ψ_f is a function $\psi_f : \Xi \longrightarrow \mathbb{R}$, where Ξ is the set of capacities in Theorem 3. By Theorem 3, ψ_f is defined by

$$\psi_f(\gamma) = \int_{ba_1^+(\Sigma_1)} \int_{S_1} f dP d\gamma$$

By the translation invariance of both the Lebesgue and the Choquet integral, it follows that for all $f \in B(\Sigma_1)$, $a \geq 0$ and $b \in \mathbb{R}$

$$\psi_{af+b}(\gamma) = \int_{ba_1^+(\Sigma_1)} \int_{S_1} (af+b) dP d\gamma = a \int_{ba_1^+(\Sigma_1)} \int_{S_1} f dP d\gamma + b = a\psi_f(\gamma) + b$$

that is, for all $f \in B(\Sigma_1)$, $a \geq 0$ and $b \in \mathbb{R}$

$$\psi_{af+b} = a\psi_f + b$$

From the monotonicity of both the Lebesgue and the Choquet integral, it follows that

$$f \geq g \qquad \Longrightarrow \qquad \psi_f \geq \psi_g$$

By assumption AA1-IB, V is translation invariant. Thus, for all $f \in B(\Sigma_1)$, $a \geq 0$ and $b \in \mathbb{R}$, we have that

$$\begin{aligned} \tilde{V}(af+b) &= V(\psi_{af+b}) = V(a\psi_f + b) \\ &= aV(\psi_f) + b = a\tilde{V}(f) + b \end{aligned}$$

By assumption AA2-IB, V is monotone. Hence,

$$f \geq g \quad \Longrightarrow \quad \psi_f \geq \psi_g \quad \Longrightarrow \quad V(\psi_f) \geq V(\psi_g) \quad \Longrightarrow \quad \tilde{V}(f) \geq \tilde{V}(g)$$

By an elementary argument, these two properties together imply that $\tilde{V}$ is sup-norm continuous. Thus, the AR preference $\succsim_1$ satisfies the Axioms 1 to 5 in Appendix B, and is IB. By the assumption that all preferences $\{\succsim_j\}_{j\in J}$ are MCIB, all measures in $Supp(\gamma)$, $\gamma \in \Xi$, are countably additive. It follows easily that $\succsim_1$ satisfies the Axiom of Monotone Continuity. We conclude that $\succsim_1$ is MCIB.

Conversely, let $\succsim$ be a MCIB preference on $F = B(S, \Sigma)$. Then, by [**2**, Theorem 2], $\succsim$ is represented by the functional $I : B(S, \Sigma) \longrightarrow \mathbb{R}$ defined by

$$I(f) = \int_{ba_1^+(\Sigma_1)} \int_S f dP d\nu$$

for all $f \in B(\Sigma)$, where ν is a capacity on $ba_1^+(\Sigma_1)$. For each $P \in Supp(\nu)$, let $\succsim_P$ be the preference on F which is represented by the functional $E_P(f) = \int f dP$, $f \in F$. Then, $\succsim_P$ is a SEU preference, hence IB. Since $\succsim$ is MCIB, all measures in $Supp(\nu)$ are countably additive. Hence, all the preferences represented by the functionals $E_P(\cdot)$ are monotone continuous. Since Choquet integrals are monotone and translation invariant, it follows that the preference $\succsim$ on F can be viewed as the result of aggregating a family $\{\succsim_P\}_{P \in Supp(\nu)}$ of MCIB preferences. □

References

[1] Amarante M. (2005), Analogical reasoning in decision processes, Columbia University discussion paper 0506-17.

[2] Amarante M. (2009), Foundations of Neo-Bayesian Statistics, *Journal of Economic Theory* **144**, 2146-73.

[3] Anscombe F.J. and R.J. Aumann (1963), A definition of subjective probability, *Annals of Mathematical Statistics* **34**, 199-205.

[4] Bourgin R.D. (1971), Barycenters of measures on certain noncompact convex sets, *Transactions of the American Mathematical Society* **54**, 323-340.

[5] Chambers C. (2007), Ordinal aggregation and quantiles, *Journal of Economic Theory* **137**, 416-31.

[6] Ghirardato P., F. Maccheroni and M. Marinacci (2004), Differentiating ambiguity and ambiguity attitude, *Journal of Economic Theory* **118**, 133-173.

[7] Ghirardato P., F. Maccheroni, M. Marinacci and M. Siniscalchi (2003), Subjective Foundations for Objective Randomization: A New Spin on Roulette Wheels, *Econometrica* **71**, 1897-1908.

[8] Ghirardato P. and M. Siniscalchi (2010), A More Robust Definition of Multiple Priors, Northwestern University mimeo.

[9] Gilboa I. and M. Marinacci (2011), Ambiguity and the Bayesian Paradigm, IGER WP 379.

[10] Gilboa I. and D. Schmeidler (1989), Maxmin expected utility with a non-unique prior, *Journal of Mathematical Economics* **18**, 141-53.

[11] Gilboa I. and D. Schmeidler (1995), Case-Based Decision Theory, *The Quarterly Journal of Economics* **110**, 605-39.

[12] Gilboa I. and D. Schmeidler (2001), A theory of case-based decisions, Cambridge University Press.

[13] Kastler D. (1982), On A. Connes' Noncommutative Integration Theory, *Communications in Mathematical Physics*, **85**.

[14] Dictionary of Philosophy of Mind, www.artsci.wustl.edu/~philos/MindDict/analogy.html

[15] Savage L.J. (1972), The Foundations of Statistics, Dover.

[16] Schmeidler D. (1986), Integral Representation without Additivity, *Proceedings of the American Mathematical Society* **97**, 255-61.

[17] Schmeidler D. (1989), Subjective probability and expected utility without additivity, *Econometrica* **57**, 571-87.

Département de sciences économiques, Université de Montréal et CIREQ

E-mail address: massimiliano.amarante@umontreal.ca