Analogy in Decision-Making

Massimiliano Amarante

Université de Montréal et CIREQ

ABSTRACT. In the context of decision making under uncertainty, we 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. We identify the basic structure of a decision problem, and provide a representation of the mappings that preserve this structure. We then consider decision makers who use multiple analogies. Our main results are a representation theorem for "aggregators" of analogies satisfying certain minimal requirements, and the identification of preferences emerging from analogical reasoning. We show that a large variety of multiple-prior preferences can be thought of as emerging from analogical reasoning.

1. Introduction

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. Analogy is one of the cornerstones of human thought ([**10**]). As such it is expected to play a fundamental role in decision-making. The scope of this paper is to formalize the idea of analogical reasoning, and to identify those decision-theoretic models that can be thought of as deriving from analogical reasoning. Our line of reasoning will unfold as follows. Given two problems in decision making, DP_1 and DP_2, a necessary condition for speaking of an analogy between the two is that we 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. We encounter here the familiar mathematical ideas of *structure* and of *structure-preserving mapping* (i.e., homomorphism). Thus, we will say that DP_1 is analogous to DP_2 if there exists a structure-preserving mapping $A : DP_1 \longmapsto DP_2$. Of course, in order to give this definition we will have to identify the basic structure of a decision problem, which we will do in Section 4. Once a satisfactory definition of analogy between two problems has been reached, we go on to study decision makers who solve the problem they face by means of multiple analogies. The main issue here is to find and to characterize "aggregators" of the various analogies. Because of the perversive nature of analogy as a processes of human thinking, these aggregators should only satisfy very minimal requirements.

The main ideas are introduced in the next section, and then gradually elaborated in the subsequent two sections. These developments come together in Section 5, where the concept of Analogy in decision-making is formally stated. This section also uncovers the basic structure of models of decision-making where the solution for the problem at hand is reached by using (possibly) multiple analogies. The main results (Section 7) are a representation theorem for "aggregators" of analogies satisfying certain minimal requirements, and the identification of preferences emerging from analogical reasoning. In particular, it is shown that a large variety of multiple-prior preferences can be thought of as emerging from analogical reasoning.

The issue of the 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.

2. Toward a definition of Analogy

By a problem in decision-making, we 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. We are going to think of a Savage model as of a "small world". Consequently, we are 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 address the following question: What does it mean that a decision-maker solves a (decision) problem by analogy with one or more problems which he solved in the past? We are going to approach the issue step-by-step, and will focus more on ideas rather than on technicalities. Full details will be given in subsequent sections.

2.1. Single analogy. We 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. Let us see what this intuitive description entails in terms of the objects introduced above. 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 can associate to each alternative in DP_1 an alternative in DP_2. That is, there must be a mapping $A : F_1 \longrightarrow F_2$. Then, 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 which 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. We will refer to such mappings as structure-preserving mappings.

Summing up, the basic idea is that we can say that DP_1 is solved by analogy with DP_2 if we can map DP_1 into DP_2 *in a way that preserves the structure of DP_1*, and then we derive the ranking $\succsim_1$ from the ranking $\succsim_2$. Later, we will see that in each decision problem there is a natural relation on the set of alternatives, and we will demand that this relation be preserved whenever we want to talk about analogical reasoning.

2.2. Multiple analogies. 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.

If we are to pursue this idea, then what we need is an "aggregator" of the various analogies. We are going to devote the remainder of this section to understanding what is an aggregator of the analogies. To begin, we notice that, as a consequence of our view above about the analogy between two problems, certain mathematical objects naturally appear in the multi-analogy case. Let us denote by $\mathcal{AP} = \{DP_2, DP_3, ...\}$ the set of problems that are analogous to $DP1$. By definition, for each problem $DP_j \in \mathcal{AP}$, there must exist a structure-preserving mapping $A_j : F_1 \to F_j$. Thus, every alternative $f \in F_1$ is associated to a collection of alternatives, one for each problem that is analogous to today's problem. That is, we have a mapping

$f \longmapsto \{A_j(f)\}_{DPj \in \mathcal{AP}}$. The next step consists of observing that, in the end, all that our individual cares about is to solve DP_1, that is he wants to come up with a ranking of the alternatives in F_1. By definition, each problem $DP_j \in \mathcal{AP}$ has already been solved. Thus, in particular, for each $f \in F_1$ the alternative $A_j(f) \in F_j$ has already been ranked in F_j. Hence, we can assign to $A_j(f)$ an object, $r_f(DP_j)$, which represents the rank that $A_j(f)$ has in problem DP_j, $A_j(f) \longmapsto r_f(DP_j)$. By composing these mappings with the previous one, we get the mapping

$$\kappa : f \longmapsto \{r_f(DP_j)\}_{DP_j \in \mathcal{AP}}$$

For each $f \in F_1$, the collection $\{r_f(DP_j)\}_{DP_j \in \mathcal{AP}}$ can be viewed as a mapping ψ_f on $\mathcal{AP}$ defined by $\psi_f(DP_j) = r_f(DP_j)$. This expresses the following information: the value of mapping ψ_f at DP_j expresses *how alternative f fares if one uses analogy DP_j*. Notice that if DP_j is the only analogy, then

$$f \succsim_1 g \qquad iff \qquad \kappa_j(f) \succsim_j \kappa_j(g)$$

in accordance to what was said in the single-analogy case.

Following this construction, we are thus led to focus on aggregators $V : \psi_f \longmapsto V(\psi_f)$, where $V(\psi_f)$ must represent 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 spits out how f should perform in the problem at hand.

Summing up: 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. A mapping

$$\kappa : f \longmapsto \psi_f$$

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 the analogy with problem 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.

Before we leave this part, one more issue has to be addressed. When we talked about a single analogy, it seemed compelling to impose the condition that the mapping $A_j : F_1 \to F_j$ be structure-preserving. Moving to the multi-analogy case, we have found the mapping κ in point 1 above. Are there any natural requirements to be imposed on the mapping κ? In particular, should we demand that κ be structure-preserving as well? Our answer is yes, and here is why. Ultimately, what an aggregator V does is to produce a ranking of the mappings $\{\kappa(f) = \psi_f\}_{f \in F_1}$, and the ranking of the alternatives in F_1 is reached by using the ranking of the mappings $\{\psi_f\}_{f \in F_1}$:

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

In the same fashion as above, it seems then essential to require that if two alternatives f and g are in a certain relation with each other, then this property would translate into a corresponding relation between the corresponding mappings ψ_f and ψ_g. That is, we believe that the requirement that mappings like κ be structure-preserving should be part of any reasonable definition of analogy.

3. Setting

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

R0: For each j, X_j is a mixture space (see [**3**], [**6**])

R1: For each j, there exists a linear utility $u_j : X_j \longrightarrow \mathbb{R}$ (Axioms on preferences guaranteeing the existence of such an 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, we will

identify the alternatives with the corresponding real-valued functions. We also assume that

> **R2:** 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 $B(\Sigma_j)$.

This is only a simplifying assumption, which serves solely to relate our work to classical models of decision-making, all of which (in some form) make such an assumption (see, for instance, [**3**], [**5**], [**7**]).

4. Structures

In Section 2.1, we put forth the idea that an analogy between two problems $DP_1 = (S_1, X_1, F_1)$ and $DP_2 = (S_2, X_2, F_2)$ should be defined as a mapping $A : F_1 \longrightarrow F_2$ which preserves the problem's structure. For an arbitrary decision problem DP_j, we have just seen that the set of alternatives F_j can be identified to the space $B(\Sigma_j)$ of bounded measurable functions on S_j. Thus, the fundamental structure of the problem DP_j is that of $B(\Sigma_j)$ (clearly, this encodes S_j and, by means of u_j, X_j as well). In turn, this is completely 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. This leads to the following definition.

DEFINITION 1. *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 (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.

4.1. Kernels and their representation. Mappings satisfying the conditions in Definition 1 are called *kernels*. The remainder of this section is devoted to showing that all kernels can be represented in essentially the same way (see [**9**], for more on kernels). This representation will produce valuable insights into the problem of formalizing the idea of analogical reasoning.

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

(4.1) $$\mu^t(A) = (\rho(\chi_A))(t), \qquad for \quad every \quad A \in \Sigma$$

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 1) and is countably additive (by property (3), Definition 1). 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 (4.1) along with properties (1), (2) and (3) of ρ imply that

(4.2) $$\rho(f)(t) = \mu^t(f) \equiv \int_S f d\mu^t$$

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 1; 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 $\rho : B(\Sigma) \longrightarrow B(\Upsilon)$ is 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 $\{\mu^t\}_{t \in T}$. Below, without any loss in generality, we are going to restrict to those kernels whose associated measures are probabilities. This 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.

5. Analogy

Now, we are going to gather what was said in the previous sections. The following definitions should present no surprise.

5.1. Single analogy.

DEFINITION 2. *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 $\theta : F_1 \longrightarrow F_2$ such that*

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

for any two alternatives f, g in DP_1.

By what was established in the previous section, this definition is equivalent to

DEFINITION 3. *Let DP_1 and DP_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 kernel $\kappa : B(\Sigma_1) \longrightarrow B(\Sigma_2)$ such that*

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

for any two alternatives f, g in DP_1.

5.2. Multiple analogies. Of course, the definition that we are going to give will be like: Problem DP_1 is solved by analogy with a collection of problems $\{DP_j\}_{DP_j \in \mathcal{AP}}$ if there exists a structure-preserving mapping $\tilde{\kappa} : f \longmapsto \psi_f$ (the mapping ψ_f was discussed in Section 2.2) and an aggregator of the analogies $\tilde{V}$ (valued in some ordered space) such that

$$f \succsim_1 g \qquad iff \qquad \tilde{V}(\tilde{\kappa}(f)) \geq \tilde{V}(\tilde{\kappa}(g))$$

It will be useful, however, to replace this type of definition with an equivalent one that encodes the representation result which is proven in Lemma 1, below. By virtue of that result, the equivalent definition will be both more manageable and more telling. Throughout, it should be kept in mind that, *by assumption,* we focus on decision makers who reach their ranking for DP_1 by means of an analogical process. Thus, the conditions below are not to be viewed as assumptions but rather as necessary conditions for an analogical procedure.

We have seen that the structure of the decision problem DP_1 is that of $B(\Sigma_1)$. A collection of analogous problems $\{DP_j\}_{DP_j \in \mathcal{AP}}$ produces the mapping $\tilde{\kappa} : f \longmapsto \psi_f$, where the value $\psi_f(DP_j)$ expresses how alternative f fares if the analogy with problem DP_j is used. The requirement that $\tilde{\kappa}$ be structure preserving implies two things: (a) we must have a measurable structure, $\mathcal{Y}$, on $\mathcal{AP}$; and (b) the mappings ψ_f must be $\mathbb{R}$-valued and $\mathcal{Y}$-measurable. Thus, we have a measurable space of analogous problems $(\mathcal{AP}, \mathcal{Y})$ as well as a collection of measurable functions on this space. The lemma below says that this space has an especially useful representation: it can be represented by means of the measurable space $(\mathcal{C}, \mathcal{B})$, where $\mathcal{C} \subset ba(\Sigma_1)$ and $\mathcal{B}$ is the Borel σ-algebra generated by the weak*-topology $(ba(\Sigma_1), B(\Sigma_1))$. In other words, analogies can be represented by means of measures on Σ_1. Let $B(\mathcal{C}, \mathcal{B})$ denote the space of bounded Borel functions on $\mathcal{C}$.

LEMMA 1. *Let $\bar{\kappa}$ be a kernel $B(\Sigma_1) \longrightarrow B(\mathcal{AP}, \mathcal{Y})$. Then, there exists a linear isomorphism λ of $\bar{\kappa}(B(\Sigma_1))$ into a linear subspace of $B(\mathcal{C}, \mathcal{B})$. Moreover, the composition $\kappa = \lambda \circ \bar{\kappa}$ is a kernel $B(\Sigma_1) \longrightarrow B(\mathcal{C}, \mathcal{B})$.*

PROOF. The kernel $\tilde{\kappa}$ defines a mapping $DP_j \longmapsto \mu^{DP_j} \in ba(\Sigma_1)$ (see Section 4.1), with μ^{DP_j} positive and countably additive. Let $\mathcal{C} \subset ba(\Sigma_1)$ denote the collection of all such measures, that is $\mathcal{C} = \{\mu^{DP_j}\}_{DP_j \in \mathcal{AP}}$. For each $f \in B(\Sigma_1)$, define $\hat{\psi}_f : \mathcal{C} \longrightarrow \mathbb{R}$ by $\hat{\psi}_f(\mu^{DP_j}) = \int_S f d\mu^{DP_j}$. Then, $\hat{\psi}_f$ is a weak*-continuous function on $\mathcal{C}$ and, therefore, $\mathcal{B}$-measurable. Next, define $\lambda : \bar{\kappa}(B(\Sigma_1)) \longrightarrow B(\mathcal{C}, \mathcal{B})$ by $\lambda(\psi_f) = \hat{\psi}_f$. Clearly, λ is one-to-one and linear, and it is immediate to see that $\kappa = \lambda \circ \bar{\kappa}$ is a kernel. $\square$

Now, we can give our definition of reasoning by analogy.

DEFINITION 4. *We say that problem DP_1 is solved by analogy with a set of problems $\mathcal{AP}$ if there exists a kernel $\kappa : B(\Sigma_1) \longrightarrow B(\mathcal{C}, \mathcal{B})$ and a functional $V : \kappa(B(\Sigma_1)) \longrightarrow \mathbb{X}$, $\mathbb{X}$ some ordered space, such that*

$$f \succsim_1 g \qquad iff \qquad V(\kappa(f)) \geq V(\kappa(g))$$

Implicit in the definition is the requirement that the kernel κ would factor as $\kappa = \lambda \circ \bar{\kappa}$, with $\bar{\kappa}$ being a kernel $B(\Sigma_1) \longrightarrow B(\mathcal{AP}, \mathcal{Y})$. Equivalently, the space $(\mathcal{C}, \mathcal{B})$ must be a representation (in the sense of Lemma 1) of the space $(\mathcal{AP}, \mathcal{Y})$.

6. Axioms on the aggregator

In this section, we are going to make Definition 4 operational by making three requirements on the aggregator of the analogies $V : \kappa(B(\Sigma_1)) \longrightarrow \mathbb{X}$. The first is certainly restrictive

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

Requirement AA0 is restrictive because it will force the preference on $B(\Sigma_1)$ to be Archimedean. However, most of what we are going to say can be easily extended to non-Archimedean preferences by using well-known methods (see [**2**]). At the same time, restricting to AA0 will help us focus more on substance than on technicalities. The next two assumptions seem quite natural and, we believe, fairly uncontroversial.

AA1 (Positive Affine Equivariance): For any positive affine transformation $a : \mathbb{R} \longrightarrow \mathbb{R}$,

$$a \circ V(\psi_f) = V(a \circ \psi_f)$$

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

AA1 simply states the translation invariance of the functional V since $\psi_f \in \kappa(B(\Sigma_1))$ implies $a \circ \psi_f \in \kappa(B(\Sigma_1))$ for any positive affine transformation $a : \mathbb{R} \longrightarrow \mathbb{R}$. Condition AA1 is motivated by the fact that the utilities u_j of Section 3 are unique only up to a positive affine transformation. Essentially, condition AA1 states that the decision maker's utility can be measured in the same units both in DP_1 and in all analogous problems $DP_j \in \mathcal{AP}$. Next, we have

AA2 (Monotonicity): $\psi_f \geq \psi_g \Longrightarrow V(\psi_f) \geq V(\psi_g)$.

This is clearly self-explanatory: it says that if f fares better than g ($f, g \in B(\Sigma_1)$) with respect to each and any analogy, then f must be preferred to g. There seems to be hardly any doubt that this must be the case.

7. Main result

This section contains our main result: all aggregators satisfying AA0, AA1 and AA2 can be represented in the same way, and every such aggregator is a Choquet integral (Theorem 1). In order to establish this, we first need an easy lemma which states that, without loss in generality, we can assume that the set $\mathcal{C}$ which represents the analogies (Lemma 1) is convex and weak*-compact. For $\mathcal{C} \subset ba(\Sigma_1)$, let $\tilde{\mathcal{C}} = \overline{co}(\mathcal{C})$ (the weak*-closed convex

hull of $\mathcal{C}$). For $f \in B(\Sigma_1)$, denote by ψ_f and $\tilde{\psi}_f$ the mappings defined by

$$\psi_f(\mu) = \int_{S_1} f d\mu \quad , \qquad \mu \in \mathcal{C} \qquad and \qquad \tilde{\psi}_f(\tilde{\mu}) = \int_{S_1} f d\tilde{\mu} \quad , \qquad \tilde{\mu} \in \tilde{\mathcal{C}}$$

Finally, let $\kappa : f \longmapsto \psi_f$ and $\tilde{\kappa} : f \longmapsto \tilde{\psi}_f$.

LEMMA 2. *Let* $V : \kappa(B(\Sigma_1)) \longrightarrow \mathbb{R}$ *be an aggregator satisfying AA1 and AA2. Then, there exists a unique aggregator* $\tilde{V} : \tilde{\kappa}(B(\Sigma_1)) \longrightarrow \mathbb{R}$ *such that* $V(\psi_f) = \tilde{V}(\tilde{\psi}_f)$, *for any* $f \in B(\Sigma_1)$. *Moreover,* $\tilde{V}$ *satisfies both AA1 and AA2.*

PROOF. For $\psi_f \in \kappa(B(\Sigma_1))$, let $z \in \mathbb{R}$ be such that $\inf_{\mathcal{C}}\psi_f + z > 0$. By AA2, $V(\sup_{\mathcal{C}}\psi_f + z) \geq V(\psi_f + z) \geq V(\inf_{\mathcal{C}}\psi_f + z)$. By AA1 and $V(1) = 1$, this implies $\sup_{\mathcal{C}}\psi_f \geq V(\psi_f) \geq \inf_{\mathcal{C}}\psi_f$. Hence, for each ψ_f there exists $\alpha(\psi_f) \in [0,1]$ such that $V(\psi_f) = \alpha(\psi_f)\inf_{\mathcal{C}}\psi_f + (1 - \alpha(\psi_f))\sup_{\mathcal{C}}\psi_f$. Now, observe that

$$\inf_{\mathcal{C}}\psi_f = \inf_{\tilde{\mathcal{C}}}\tilde{\psi}_f \qquad and \qquad \sup_{\mathcal{C}}\psi_f = \sup_{\tilde{\mathcal{C}}}\tilde{\psi}_f$$

and that the mapping $\psi_f \longmapsto \tilde{\psi}_f$ from $\kappa(B(\Sigma_1) \longrightarrow \tilde{\kappa}(B(\Sigma_1))$ is clearly one-to-one and onto. Hence, we can define $\tilde{\alpha} : \tilde{\kappa}(B(\Sigma_1)) \longrightarrow [0,1]$ by $\tilde{\alpha}(\tilde{\psi}_f) = \alpha(\psi_f)$. Then, $\tilde{V} : \tilde{\kappa}(B(\Sigma_1)) \longrightarrow \mathbb{R}$ defined by $\tilde{V}(\tilde{\psi}_f) = \tilde{\alpha}(\tilde{\psi}_f)\inf_{\tilde{\mathcal{C}}}\tilde{\psi}_f + (1 - \tilde{\alpha}(\tilde{\psi}_f))\sup_{\tilde{\mathcal{C}}}\tilde{\psi}_f$ is the unique functional satisfying $V(\psi_f) = \tilde{V}(\tilde{\psi}_f)$, for any $f \in B(\Sigma_1)$. The second part is immediate. $\square$

A short comment is in order. While the lemma says that it is immaterial to assume either the weak*-compactness or the convexity of the set $\mathcal{C}$ representing the analogies, both properties seem quite natural. If $\mu, \nu \in \mathcal{C}$, then – loosely speaking – one can always flip a coin between μ and ν thus effectively creating, at the same time, both another problem and an analogy with that problem. Thus, the idea of mixing appears to be a sound justification for the assumption that $\mathcal{C}$ be convex. Compactness would follow if the set $\mathcal{C}$ representing the analogies is weak*-closed. This also seems to be a natural property, especially when combined with convexity: essentially, it implies that one can always complete the set of analogies by means of an inductive procedure such as that of taking limits.

By virtue of the lemma, we can now assume without any loss in generality that $\mathcal{C}$ is convex and weak*-compact. In such a case, the space $\kappa(B(\Sigma_1)$

coincides with the space $A(\mathcal{C})$, the space of all weak*-continuous affine mappings on $\mathcal{C}$. The following theorem completely characterizes aggregators satisfying AA0, AA1 and AA2.

THEOREM 1. *An aggregator $V : \kappa(B(\Sigma_1)) \longrightarrow \mathbb{R}$ satisfies AA1 and AA2 if and only if there exists a capacity Γ on $\mathcal{B}$, the Borel sets of $\mathcal{C}$, such that for every $\psi_f \in \kappa(B(\Sigma_1))$*

$$V(\psi_f) = \int_{\mathcal{C}} \psi_f d\Gamma$$

where the integral is taken in the sense of Choquet.

PROOF. AA1 states that V is translation invariant. If $\psi, \varphi \in \kappa(B(\Sigma_1)) = A(\mathcal{C})$ and ψ and φ are non-constant, then ψ and φ are comonotonic if and only if they are isotonic ([**1**, Proposition 2]). In such a case, there exist $a > 0$ and $b \in \mathbb{R}$ such that $\varphi = a\psi + b$. Then,

$$V(\psi + \varphi) = V((1 + a)\psi + b) = V(\psi) + V(\varphi)$$

by the translation invariance of V. Thus, V is comonotonic additive on the domain $A(\mathcal{C})$. By AA2, V is monotone as well. In Amarante [**1**, Corollary 1], it was shown that these functionals can be represented by Choquet integrals. The converse is an immediate consequence of the properties of the Choquet integral. $\qquad\square$

The following are all examples of aggregators satisfying AA0, AA1 and AA2:

(1) Every Lebesgue integral;

(2) Every (probabilistic) quantile;

(3) Every generalized quantile (i.e., monotone and ordinally covariant functional $\kappa(B(\Sigma_1)) \longrightarrow \mathbb{R}$, see [**4**]).

The first is obviously a special case of a Choquet integral. The procedure of evaluating alternatives by using a *similarity function* as in Gilboa and Schmeidler [**8**] gives an example of an aggregator of analogies which is a Lebesgue integral (linear weighting). The fact that both (2) and (3) satisfy the conditions of Theorem 1 follows from their representability as Choquet integrals as established in [**4**].

8. Analogical Reasoning preferences

In this section, we are going to use Theorem 1 to identify those preferences that can be thought of as deriving from analogical reasoning. We are going to show that a wide class of preferences can be thought of in this way (Corollaries 1 and 2). Let (S_1, X_1, F_1) be a problem of decision-making under uncertainty, and let $\succsim_1$ be a preference relation on the mappings $S_1 \longrightarrow X_1$.

DEFINITION 5. *We say that preference $\succsim_1$ is an analogical reasoning (AR) preference if $\succsim_1$ satisfies the conditions in Definition 4 with V satisfying AA0, AA1 and AA2.*

The combination of the next two corollaries identifies the class of AR-preferences. We recall that the class of Invariant Biseparable Preferences (IBP) consists of those preferences satisfying the first five axioms in Gilboa-Schmeidler ([**5**], [**7**]).

COROLLARY 1. *Every AR preference is an IBP.*

PROOF. Let $\succsim_1$ be an AR preferences on the mappings $S_1 \longrightarrow X_1$. By Theorem 1, $\succsim_1$ can be represented by a functional $I : B(\Sigma_1) \longrightarrow \mathbb{R}$ defined by

$$I(f) = \int_{\mathcal{C}} \kappa(f) d\Gamma$$

where κ is the mapping $B(\Sigma_1) \longrightarrow A(\mathcal{C})$. It is easily checked that $\succsim_1$ satisfies axioms A1 to A5 in [**7**].   □

The next proposition provides a partial converse to Corollary 1. We recall that an IB preference satisfies the Axiom of Monotone Continuity ([**5**, Sec. B.3]) if and only if all the priors appearing in the representation of the IBP are countably additive and the set $\mathcal{C}$ is weak compact.

COROLLARY 2. *Every IBP which, in addition, satisfies the axiom of Monotone Continuity is an AR preference.*

PROOF. By [**1**, Theorem 2], any IBP preference is represented by a functional of the type $I(f) = \int_{\mathcal{C}} \kappa(f) d\Gamma$, where κ is defined like above and Γ is a capacity on $\mathcal{B}$. So an IB preference is an AR preference if κ is a kernel. This occurs if and only if all the measures in $\mathcal{C}$ are countably additive. In turn, this occurs if and only if the preference satisfies the Axiom of Monotone Continuity ([**5**, Sec. B.3]).   □

There is a gap between the class of IB preferences and that of AR preferences as well as a gap between Monotone Continuous IB preferences (MIBP) and AR preferences; that is, all the inclusions $MIBP \subset ARP \subset IBP$ are strict. Here are two examples:

(a) *An IBP which is not AR*: Take any multiple prior model with $\mathcal{C}$ being a convex combination of two finitely, but not countably, additive measures. Then (by [**1**]), $I(f) = \int_{\mathcal{C}} \kappa(f)d\Gamma$, but κ is not a kernel because it is not a normal mapping;

(b) *An ARP which is not MIBP:* In order to manufacture such an example, it suffices to exhibit a set of priors which is the closure in $ba(\Sigma)$ of a set of countably additive measures, but this closure contains finitely additive measures (equivalently, the closure is weak* but not weak compact).

Overall, these gaps are fairly minor, and occur only when the state space S_1 is infinite. The gap between MIBP and ARP (which, in addition requires the set $\mathcal{C}$ to be infinite-dimensional) can be removed by imposing an additional axiom guaranteeing that the set of measures appearing in the ARP representation be uniformly countably additivity (which is equivalent to imposing the existence of a control measure on the set of priors).

References

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

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

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

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

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

[6] 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.

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

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

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

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

DÉPARTEMENT DE SCIENCES ÉCONOMIQUES, UNIVERSITÉ DE MONTRÉAL ET CIREQ

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