GGEN 1355913
17-10-2017
Revision
CE:PG QA:RB

INTERNATIONAL JOURNAL OF GENERAL SYSTEMS, 2017
VOL. 46, NO. 8, 879–897
https://doi.org/10.1080/03081079.2017.1355913

Taylor & Francis
Taylor & Francis Group

Local and global optima in decision-making: a sheaf-theoretical analysis of the difference between classical and behavioral approaches

Fernando Tohmé[a], Gianluca Caterina[b] and Rocco Gangle[b]

[a]INMABB – Conicet and Universidad Nacional del Sur, Bahía Blanca, Argentina; [b]Center for Diagrammatic and Computational Philosophy, Endicott College, Beverly, MA, USA

ABSTRACT

One of the main differences between the traditional and the behavioral approaches to decision-making is that the latter has not yet been captured in a unifying framework. This hampers in a certain way the whole research program and raises the question of whether this competing approach can provide an encompassing alternative to the classical one. We analyze this issue in light of the problem of reconstructing global choices of an agent up from the solutions found for local problems. We show that a representation based on category theory of the conditions for such reconstruction is general and robust enough to represent both the case in which problems are non-contextual and local as well as that, typical in the literature on behavioral decision-making, in which such properties do not hold. In the first case, we show how a sheaf-theoretical representation provides an abstract characterization of the global solution. In the latter case, we show how locality and contextuality generate obstructions toward the reconstruction of global solutions, yielding a possible clue for the intrinsic difference between behavioral and classical decision theory.

ARTICLE HISTORY
Received 19 January 2017
Accepted 15 May 2017

KEYWORDS
Decision-making; behavioral decision theory; optimal choices; sheaves; projections; category theory

1. Introduction

Most of the economic and decision-making literature assumes that the goal of a rational agent is to find the *best* solution to any relevant choice problem. But in recent decades, this view has become increasingly qualified, mostly by pointing out the different possible meanings of "best".

The traditional position is that an agent should be associated with a preferential order, represented by a continuous real-valued function U over some space of alternatives. The goal of the agent would then be to maximize U, that is, to find a solution to a global optimization problem. But even those who conceive economic decision-making in this way, usually represent the agent in a *partial* or *local* setting, where it is assumed that the solution with respect to the local setting does not necessarily result as a constrained version of some global preferential solution. Furthermore, the solution to each partial problem is

CONTACT Fernando Tohmé ftohme@criba.edu.ar

routinely conceived as being independent from the choices the agent could make in other contexts.

Even though in some special cases closed-form global solutions are within reach, the constraints that need to be imposed are usually too restrictive, undermining the robustness of the analysis.[1] Worse yet, it may even be the case in some instances that no global U is theoretically discernible, and only local problems (i.e. those restricted to some subsets of the space of alternatives) are susceptible to being posed and solved.

On the other hand, the vast literature on behavioral Decision Theory implicitly shows that it might be very difficult or even impossible to find a single unifying theoretical framework for real-world decisional contexts, one that would cover all the different heuristics and biases involved in actual cases of decision-making. While this is not a point usually analyzed in the literature,[2] the diverse models and approaches in the literature (see Camerer, Loewenstein, and Rabin 2004; Cartwright 2011) cannot be seen as variants of a single underlying representation of decision-making. This raises the question of the possibility of reconstructing a global U up from local information. We contend that the behavioral approach fails to satisfy the conditions that would allow for such a reconstruction and, furthermore, that the reasons for such failure are the non-locality and contextuality of the models of behavioral decision-making.[3]

The question of finding a unifying framework for behavioral decision-making is far from being merely scholastic. For instance, many economists have agreed upon the need for a single theoretical body for Behavioral Economics, if it is to realize the goal of replacing the classical approach. As pointed out by D. K. Levine (cited in Hartford (2014)):

> There is a tendency to propose some new theory to explain each new fact. The world doesn't need a thousand different theories to explain a thousand different facts. At some point there needs to be a discipline of trying to explain many facts with one theory.

The goal of building parsimonious theories is present in all scientific disciplines as an expression of *Occam's Razor*, but it is certainly not always easily achievable (Burgess 1998). The seemingly more abstract question of whether it is *in principle* possible to find such a unifying theory for some domain may in fact be quite concretely relevant for scientific progress since, if answered definitively in the negative, it would save a lot of efforts that could be better devoted to exploring the necessarily disparate approaches and their exact ranges of validity within that domain. We think that our paper achieves this goal for the case of behavioral decision theory, pointing out the limits for the reconstruction of a global picture of the preferences (and thus the possibility of self-interested decision-making) up from partial pieces of information. An immediate consequence of our main result is that even modest attempts to unify certain behavioral effects will inevitably fail. Our negative result must be seen as a contribution in this direction, being relevant for both the partisans and the critics of the behavioral approach.

1.1. Purpose of this paper

Given the general presentation of the problem of finding a unifying framework for the different models of Behavioral Economics, let us give a general overview of what are the specific problems analyzed in this paper and the results we find.

Let us start by recalling that mainstream Economics has a notion of *agent* that can be described in terms of a given preference order over the space of alternatives. The agent is said to be *rational* if she chooses the most preferred alternatives among those that are feasible for her. While this view of rationality is *normative*, in practice it constitutes the main explanatory principle that mainstream economists use to make sense of real-world phenomena. This specification can be made more complex in many ways, but this basic description is still the cornerstone of this literature (see Mas-Colell, Whinston, and Green 1995, Chapter 1).[4] Yet on the other hand, in actual applications of this model it is customary to reduce the analysis to a subspace of the space of alternatives, simplifying the problem of making a decision. This requires an assumption of the independence of the preferences over the subspace from the preferences over the rest of the larger space of alternatives (see Mas-Colell, Whinston, and Green 1995, Chapter 10). This *ceteris paribus* clause does not prevent the possibility of inconsistencies when the solutions to different problems are combined. This raises the first question we plan to address in this paper:

> *A*: Under what conditions do the solutions to partial decision-making problems allow for the reconstruction of the preferences over the entire space of alternatives?

This question can be posed also, in more pressing terms, for Behavioral Economics, since its starting point is that empirical evidence indicates that economic agents behave in ways that are not compatible with the mainstream model of decision-making. In particular, it shows that factors other than those of the preferences may influence the decisions made by the agents. Among those factors are certain particularly salient ones, related to the ways in which the players *frame* the decision problems and the *heuristics* they use to solve them. But then the solutions that any agent may obtain for the different problems seem again independent one from another. The question of whether they can be seen as part of a single framework can be formulated as follows:

> *B*: Can we ensure the existence of a overarching set of alternatives such that each particular solution is a "projection" of one of those?

This is, of course, an extreme idealization of the question of whether there exists a universal procedure of decision-making such that each particular way of choosing alternatives is just one of its instances.

We see questions *A* and *B* as closely associated. But this association cannot be clearly established in the traditional mathematical framework of Economics, which differentiates between preferences and choices, except under very specific *revealed preferences* conditions (Chambers and Echenique 2016). Thus, we resort to a formalism that unifies the two questions, but still allows them to be distinguished.

Answering these questions is relevant for understanding the way in which Economics addresses empirical problems. In this sense, one of the goals of this paper is to find conditions that must be verified to ensure the existence of preferences over the entire space of alternatives leading to solutions to partial decision-making problems, without necessarily requiring full rationality.

We start by considering as the central notion that of a *decision problem* with a specific domain, say D and a problem-specific function u, focused on the arguments relevant for the question at hand. Finding the optimal ones yields the problem-specific solution.

Given multiple problems sharing a common domain, a global U might nonetheless be hard to state solely in terms of its local instances. But an intuitive idea is that for each particular problem with domain D and specific function u, we can look for a function V

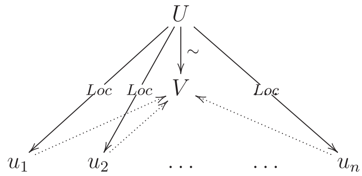

Figure 1. The localization operator *Loc* reduces the global maximization problem to a sequence of local ones.

such that $u = V_{|D}$. Repeated conservative instances of this procedure might promise to yield, in the limit, a V that can be eventually identified with the hypothetical global utility function U [5] (see Figure 1).

Notice that, in order to recover U in this fashion, we must be able to patch together the local restrictions in a consistent way. But this procedure by no means guarantees a global solution. We have to find ways to identify the conditions that allow for the possibility of patching together local pieces of information in a way that remains consistent with their global integration.

This brings us to the choice of a formal approach to answer our questions A and B with appropriate rigor. Our choice is to use the language and techniques of category theory.

1.2. A categorical toolbox

As is well known, this branch of Mathematics has provided a framework without which most of the contemporary results in both Algebraic Geometry and Topology would not have been found (Hatcher 2002). As shown again and again in actual mathematical practice, the language of set theory remains insufficient for capturing perspicuously the nuances prevalent in those fields (Marquis 2009). One reason is that unlike set theory the categorical approach allows for both the maximization of the “external” scope of its formal results *and* the controlled “internal” sensitivity to particular differences in content within the representation of mathematical structures. While category theory might thereby also seem to be a natural choice of formal language for representing the decision-making problems outlined above, we have to note that Economics has been reluctant to adopt it.[6]

In this paper, we draw heavily on the literature on Category Theory, although our results are clearly elementary. We will now present the basic concepts that will be used in subsequent sections. For further details and clarification, see the excellent presentations of Goldblatt (1984), Barr and Wells (1999), Adámek, Herrlich, and Strecker (2004), Lawvere and Schanuel (2009) or Spivak (2014).

A category $\mathbf{C}$ consists of a set of *objects*, $\mathbf{Obj}$ and a class of *morphisms* between pairs of objects. Given two objects $a, b \in \mathbf{Obj}$ a morphism f between them is notated $f : a \to b$. Given another object c and a morphism $g : b \to c$, we have that f and g can be composed, yielding $g \circ f : a \to c$ (COMPOSITION). Additionally, for every $a \in \mathbf{Obj}$, there exists an *identity* morphism, $\mathrm{Id}_a : a \to a$. Morphisms are required to obey two rules: (*i*) if $f : a \to b, f \circ \mathrm{Id}_a = f$ and $\mathrm{Id}_b \circ f = f$ (IDENTITY); (*ii*) given $f : a \to b, g : b \to c$ and $h : c \to d$, $(h \circ g) \circ f = h \circ (g \circ f) : a \to d$ (ASSOCIATIVITY).

Examples of categories are **SET** (the objects are sets, and the morphisms functions between sets), **TOP** (the objects are topological spaces and the morphisms continuous

$$\begin{tikzcd} a \arrow[r, "f_a^{b_j}"] \arrow[dr, "f_a^{b_l}"'] & b_j \arrow[d, "\gamma_{jl}"] \\ & b_l \end{tikzcd}$$

Figure 2. Commutative diagram.

$$a \xleftarrow{f} X \xrightarrow{g} b$$

Figure 3. The limit of cones of this shape defines the product $a \times b$.

functions), **Pord** (the objects are preorders and the morphisms are order-preserving functions), **Vec** (the objects are vector spaces and the morphisms linear maps), etc.

The terseness of categories facilitates diagrammatic reasoning. A diagram in which nodes represent objects and arrows represent morphisms allows to establish properties of a category. Diagrams that *commute*, i.e. such that all different direct paths of morphisms with the same start and end nodes are identified (that is, compose to a common morphism), indicate relations similar to those that can be established by means of equations.

Some of the most interesting constructions that can defined in categories are *limits* and *colimits* (duals of limits). Any limit (or colimit) captures a *universal property* on a family of diagrams with the same basic shape. This basic shape is captured by a *cone*, that is, an object a and a family of arrows $\{f_a^{b_j} : a \to b_j\}_{j \in \mathcal{J}}$, such that for any pair $j, l \in \mathcal{J}$, if there exists a morphism $\gamma_{jl} : b_j \to b_l$, we have that $\gamma_{jl} \circ f_a^{b_j} = f_a^{b_l}$ (see Figure 2).

Then, given a class of cones of a given shape, a limit is an object L in this class such that for every other cone T in the class there exists a single morphism $T \to L$ such that the resulting combined diagram commutes. For instance, consider a family of cones of the shape depicted in Figure 3.
then, the limit is the *product* $a \times b$ and with arrows p_1 and p_2, the projections on the first (a) and second (b) components, respectively. For every other cone, with "apex" X there is a unique morphism $! : X \to a \times b$ such that $f = p_1 \circ !$ and $g = p_2 \circ !$.

Examples of colimits are *direct sums* (in **SET**, disjoint unions) and, somewhat confusingly called, *direct limits*, which in a self-contained description we will use to define *global solutions* in Section 2.

Besides capturing interesting constructions common to many fields of Mathematics, category theory also provides tools for relating different categories to one another. This is achieved by means of mappings called *functors*. Given two categories **C** and **D** a functor F from **C** to **D** maps objects from **C** into objects of **D** as well as arrows from the former to the latter category such that, if

$$f : a \to b$$

in **C**, then:

$$F(f) : F(a) \to F(b)$$

in **D**. Furthermore, $F(g \circ f) = F(g) \circ F(f)$ and $F(\mathrm{Id}_a) = \mathrm{Id}_{F(a)}$ for every object a in **C**.

These functors are called *covariant*. Another class, that of *contravariant* functors, is such that, if

$$f : a \to b$$

in **C**, then:

$$F(f) : F(a) \leftarrow F(b)$$

in **D**. Of particular interest are the contravariant functors $F : \mathbf{C} \to \mathbf{SET}$ (or a category of subsets of a given set), which are called *presheaves*. An intuitive interpretation is that given a morphism $a \to b$ in **C**, the morphism $F(b) \to F(a)$ in **SET** is the *restriction* of the "image" under F of b over the "image" of a. Given an object a in **C**, $F(a)$ is called a *section* of F over a. This can be extended to any family $B = \{b_j\}_{j\in\mathcal{J}}$ of objects in **C**: $F(B)$ is the section over B. In turn, given two families $B \subseteq B'$ and the section over B', namely$F(B')$, we can find its restriction over B, denoted $F(B')_{|B}$, yielding $F(B)$.

Given a presheaf $F : \mathbf{C} \to \mathbf{SET}$, consider a class of objects B in **C** and a cover $\{K_j\}_{j\in\mathcal{J}}$ (i.e. $B \subseteq \bigcup_{j\in\mathcal{J}} K_j$). Let $\{k_j\}_{j\in\mathcal{J}}$ be a sequence such that $k_j \in F(K_j)$ for each $j \in \mathcal{J}$. The presheaf F is said to be a *sheaf* if the following conditions are fulfilled:

- *Locality*: For every pair $i, j \in \mathcal{J}$, $k_{i|_{K_i\cap K_j}} = k_{j|_{K_i\cap K_j}}$ (i.e. the sections a_i, a_j coincide over $V_i \cap V_j$),
- *Gluing*: There exists a unique $\bar{b} \in F(B)$ such that $\bar{b}_{|K_j} = k_j$ for each $j \in \mathcal{J}$ (i.e. there exists a single object in the "image" of B that when restricted to each set in the covering yields the section corresponding to that set).

This brief review of category theory provides the basic concepts necessary for the analysis to be carried out in the rest of the paper.

1.3. Main results anticipated

Once we define *local problems* in decision-making, we consider the category of which they are objects. This category has morphisms reflecting the relations between different problems. We show that our question A (Section 1.1) can be answered positively if we can define a functor from this category into the category of subsets of the space of alternatives, such that this functor is a *sheaf*, i.e. such that the problems are consistent on shared subspaces and there exists a problem such that all other problems are constrained versions of it. Recall that in mainstream Economics, the function to be optimized in each problem is the constrained version of a global U (representing the map of preferences over the entire space of alternatives). *Then, in the classical setting it is indeed the case that A has a positive answer*.

We then generalize this setting, without requiring that the functions to be optimized be special restrictions of some given function. We present two models of behavioral decision-making, *Prospect Theory* (Kahneman and Tversky 1979) and *Case-based Decision Theory* (Gilboa and Schmeidler 2001) formulated in this new category of local problems. The former shows a clear case of dependence upon the context in which the decision is made. The latter, in turn, exhibits non-locality, i.e. the solution depends on how previous problems have been solved. These two features are common to all the models in Behavioral Economics.

Moreover, we show that the only way to give a positive answer to question B (Section 1.1) is if the sheaf property is accompanied by an additional condition, namely that the relation between problems and solutions constitutes a *trivial bundle*. This means basically that each problem is associated in a unique way with the solution of the covering problem (which necessarily exists because of the sheaf condition). More precisely, the solutions of any problem are the projections of the solutions of the covering problem.

Now the answer to question B is obtained by noticing that the properties of behavioral models (contextuality and non-locality) act as obstructions to the definition of a trivial bundle. And thus, *in the behavioral setting the answer to B is negative*. On the other hand, as we show in Lemma 1, in actuality the definition of a trivial bundle is always satisfied by the conditions of the mainstream models, i.e. *if A is true, then B is true, but A is only true in the classical setting but not in the behavioral one*.

The mathematical concepts used in the paper are standard in the literature on categories and (applied) Algebraic Topology (see Ghrist 2014) and on the general question of relating local and global properties (see Macfarlane 2014; Plikynas et al. 2014; Baas 2015, 2016, among others). The proofs are rather straightforward, but the gist is in the definition of the appropriate categorical notions, showing clearly the advantages of adopting this framework in the analysis of economic problems.

The actual presentation of the aforementioned results is as follows. Section 2 introduces a formal description of the problem in the simplest and most straightforward case in which the global function is known and does project well onto the local problems. Section 3 then presents the idea of a sheaf construction of a global outcome induced by a family of local utility functions. Finally, Section 4 shows that only if the decision-making process is local and non-contextual throughout does the sheaf construction yield actual solutions to the maximization of U. Section 5 summarizes the conclusions.

2. Decision-making: local vs. global

The traditional characterization of decision-making under certainty by an individual is as follows. Let $\mathcal{L}_i$ be a space of possible **options** that an agent i may select.[7] Each $x \in \mathcal{L}_i$ is evaluated by means of a *utility* function, $U_i : \mathcal{L}_i \to \Re$. Given a family of constraints limiting the set of options open to the agent to $\hat{L}_i \subseteq \mathcal{L}_i$, the goal of the agent is to find some $\mathbf{x}^*$ that maximizes U_i over $\hat{L}_i$. Since in this work we focus on the possible choices made by a single agent, we drop the subindex i from the notation for $\mathcal{L}$, $\hat{L}$ and U. We will reintroduce the dependence on the agents in a future work in which we will analyze the interaction between different agents.

In order to proceed, we first make some plausible assumptions. The space of options, $\mathcal{L}$, is presumed to be a (real) Hilbert space. That is, it is a complete metric space with an inner product. Furthermore, in order to ensure the existence of a $\mathbf{x}^*$ we will also assume that $\hat{L}$ is a compact subset of $\mathcal{L}$, and that U is a continuous function. Within this very general framework, it is then straightforward to induce a category-theoretical representation of the global optimization of U over $\hat{L}$, that is, of $\mathbf{x}^*$ as a *direct limit*.

To begin, consider first a family $\{L^k\}_{k=0}^{\kappa}$ of closed linear subspaces of $\mathcal{L}$ and, for any given k, let us define the function

$$\mathrm{Proj}_k : \mathcal{L} \to \bigcup_{k=0}^{\kappa} L^k$$

such that $\mathrm{Proj}_k(x) = x^k \in L^k$, where x^k is the *projection* of x on L^k. The existence of such a projection is ensured by a straightforward application of the linear projection theorem.[8]

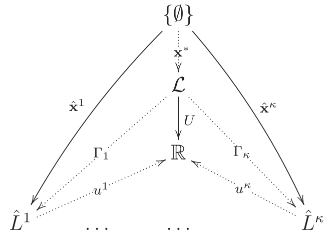

Figure 4. In this diagram $\mathbf{x}^*$ is the global maximum for U, whereas $\hat{\mathbf{x}}^k$ is a local solution (maybe not unique) for the k-restricted problem.

The projector operator Proj_k will play a fundamental role in what follows. The intuition here is that we can think of each L^k as the options set of a local problem. Therefore, the projection of a global solution $\mathbf{x}^*$ onto L^k will return the point in L^k which is the closest (i.e, the best!) to $\mathbf{x}^*$. Analogous approaches have been used successfully in several different contexts.[9]

In case the projection does not return a local solution, however, we can still define an operator, which we call $\Gamma_k(x)$ that formalizes the idea of "best choice" within a local problem (that is, a choice that is the closest to the sought global solution). To analyze this problem, let us define a new correspondence, $\Gamma_k : \hat{L} \to \hat{L}^k$,

$$\Gamma_k(x) = \{x^k \in \hat{\mathbf{X}}^k : x^k \in \mathrm{argmin}_{y \in \hat{\mathbf{X}}^k} |y - \mathrm{Proj}_k(x)|\}.$$

It follows immediately that, if $\mathrm{Proj}_k(x) \in \hat{\mathbf{X}}^k$ then $\Gamma_k(x) = \mathrm{Proj}_k(x)$.

We are interested in the set given by $\Gamma_k(\mathbf{x}^*)$. Figure 4 may help to visualize this setting.

Example 1: Consider a consumer with preferences over three goods. Consumption baskets are denoted (x, y, z), identified as vectors in the non-negative orthant of $\mathbb{R}^3$. The preferences over baskets are represented by the following utility function:

$$U(x, y, z) = x^2 + (y + 1)(z + 1)$$

Suppose that the income of the agent is \$10, while the price of the units of the three goods are, respectively, \$4, \$2 and \$5. This means that the *budget set* is $\mathbf{B} = \{(x, y, z) : 4x + 2y + 5z \leq 10\}$ (a closed and compact set). Since U is continuous and monotonic, its maximal value is achieved at a basket (x, y, z) such that

$$4x + 2y + 5z = 10 \quad (*)$$

Then, a straightforward replacement of x, obtained from $(*)$, into the argument of U, followed by its optimization with respect to y and z yields (notice that U is differentiable) that:

$$x^* = \frac{17}{14} \quad y^* = \frac{57}{28} \quad z^* = \frac{3}{14}$$

Instead, if we restrict the problem to choosing a basket without the z-good, $(*)$ reduces to $4x + 2y = 10$, while the utility becomes $U^a = x^2 + (y + 1)$. A straightforward optimization yields:

$$x^a = 1 \quad y^a = 3$$

If instead, we consider the problem of choosing baskets without the y-good, $(*)$ becomes $4x + 5z = 10$, while the utility becomes $U^b = x^2 + (z+1)$. The solution to this problem is:

$$x^b = \frac{2}{5} \quad z^b = \frac{42}{25}$$

The question is to determine the relations between the global solution (x^*, y^*, z^*) and the local solutions (x^a, y^a) and (x^b, z^b).

A more abstract example, allowing us to analyze those relations in a more rigorous way, is:

Example 2: Consider $\mathcal{L}$ to be $\mathbb{R}^3$ (the three-dimensional real Euclidean space) and the utility function:

$$U(x, y, z) = 3 - 2x^2 - y^2 - 3z^2$$

to be maximized over $\mathcal{L}$. This yields a single global solution $\hat{\mathbf{X}} = \{(0,0,0)\}$.

Consider now two possible "local" problems:

- $L^1 = \{(x,y,z) : z = 0\}$, *with* $u^1(x,y,z) = U_{|L^1} = 3 - 2x^2 - y^2$ *to be maximized over* $\hat{L}^1 = \{(x,y,0) \in L^1 : x^2 + y^2 = 1\}$, *the unit circumference in* L^1. *The class of solutions for this problem is* $\hat{\mathbf{X}}^1 = \{(0,1,0), (0,-1,0)\}$.
- $L^2 = \{(x,y,z) : (x,y,z) \cdot (1,-1,1) = 0\}$ *(i.e. the linear subspace with normal vector* $(1,-1,1)$*), with* $u^2(x,y,z) = 3 - 3x^2 - 4z^2 - 2xz$, *the restriction of* U *on* L^2, *to be maximized over* $\hat{L}^2 = \{(x,y,z) : 2x^2 + 2z^2 + 2xz = 1\}$, *the intersection of the surface of the unit sphere in* $\mathbb{R}^3$ *with* L^2. *Here, the solution set is:* $\hat{\mathbf{X}}^2 = \left\{ \left(-\sqrt{\frac{1}{3}}, -\frac{1}{2\sqrt{3}} - \frac{1}{2}, \frac{1}{2\sqrt{3}} - \frac{1}{2} \right), \left(\sqrt{\frac{1}{3}}, \frac{1}{2\sqrt{3}} + \frac{1}{2}, \frac{1}{2} - \frac{1}{2\sqrt{3}} \right) \right\}$.

It is easy to see that each solution of problem 1 minimizes the distance to the projection of the single global solution $(0,0,0)$ on L^1. More precisely $\Gamma_1(0,0,0) = \hat{\mathbf{X}}^1$. The same is true for problem 2, since all points in L^2 are at a Euclidean distance 1 from the global solution. So, in particular, the elements in $\hat{\mathbf{X}}^2$ minimize the distance to the projection of $(0,0,0)$ on L^2 and thus, $\Gamma_2(0,0,0) = \hat{\mathbf{X}}^2$.

The above example, although very simple, suggests that we can formalize the notion of "local problem" along such lines. We refer accordingly to the triple $s^k = \langle \hat{L}^k, u^k, \hat{\mathbf{X}}^k \rangle$ as the corresponding *local problem* which consists of the maximization of a continuous utility function u^k over a compact set $\hat{L}^k \subseteq L^k$. This in turn yields a non-empty family of solutions $\hat{\mathbf{X}}^k = \{\hat{\mathbf{x}} : u^k(\hat{\mathbf{x}}) \geq u^k(x)$ for every $x \in \hat{L}^k\}$.

Now that we have a precise definition of a local problem, we are ready to take our first leap into a category theory setting. In particular, we begin by showing that the global solution can be regarded as a *direct limit* within an appropriate category.

For each problem s^k, characterized by a compact set $\hat{L}^k$ in a linear space $\mathcal{L}$ and the continuous utility function u^k, we consider the following category $\mathcal{C}^k$:

- the objects are the elements of $\hat{L}^k$.
- a morphism $f_{\alpha\beta}$ from x_α to x_β (where $x_\alpha, x_\beta \in \hat{L}^k$) exists iff $u^k(x_\alpha) \leq u^k(x_\beta)$. Of course, for every $x_\alpha, f_{\alpha\alpha}$ corresponds to the identity $u^k(x_\alpha) = u^k(x_\alpha)$. Furthermore, $f_{\alpha\gamma} = f_{\beta\gamma} \circ f_{\alpha\beta}$ since we have that $x_\alpha \overset{f_{\alpha\beta}}{\rightarrow} x_\beta$ corresponding to $u^k(x_\alpha) \leq u^k(x_\beta)$ and $x_\beta \overset{f_{\beta\gamma}}{\rightarrow} x_\gamma$, corresponding to $u^k(x_\beta) \leq u^k(x_\gamma)$ imply the existence of $x_\alpha \overset{f_{\alpha\gamma}}{\rightarrow} x_\gamma$ since, by transitivity, $u^k(x_\alpha) \leq u^k(x_\gamma)$.

$$\begin{tikzcd}
& \hat{\mathbf{x}}' \arrow[dd, bend left=60, "f"] \\
x_\alpha \arrow[ur, "\psi_\alpha"] \arrow[r, "f_{\alpha\beta}"] \arrow[dr, "\phi_\alpha"'] & x_\beta \arrow[u, "\psi_\beta"] \arrow[d, "\phi_\beta"'] \\
& \hat{\mathbf{x}}
\end{tikzcd}$$

Figure 5. Direct limit construction: f ensures that $\phi_\alpha = f \circ \psi_\alpha$ and $\phi_\beta = f \circ \psi_\beta$.

Consider now an object $\hat{\mathbf{x}}$ and $\Phi = \{\phi_\gamma : x_\gamma \to \hat{\mathbf{x}}\}_{x_\gamma \in \hat{L}^k}$, a family of morphisms such that $x_\gamma \xrightarrow{\phi_\gamma} \hat{\mathbf{x}}$ iff $u^k(x_\gamma) \leq u^k(\hat{\mathbf{x}})$. In the terms of elementary category theory, $\langle \hat{\mathbf{x}}, \Phi \rangle$ is said to constitute a *direct limit* within $\mathcal{C}^k$ if:

(1) It satisfies that $\phi_\alpha = \phi_\beta \circ f_{\alpha\beta}$ i.e. $u^k(x_\alpha) \leq u^k(x_\beta)$ and $u^k(x_\beta) \leq u^k(\hat{\mathbf{x}})$ imply that $u^k(x_\alpha) \leq u^k(\hat{\mathbf{x}})$.
(2) It is universal, in the sense that for any other $\langle \hat{\mathbf{x}}', \Psi \rangle$ such that $\Psi = \{\psi_\gamma : x_\gamma \to \hat{\mathbf{x}}'\}_{x_\gamma \in \hat{L}^k}$ satisfies condition 1, there exists a unique $f : \hat{\mathbf{x}}' \to \hat{\mathbf{x}}$ making commutative the diagram involving $x_\alpha \xrightarrow{f_{\alpha\beta}} x_\beta$, $x_\alpha \xrightarrow{\phi_\alpha} \hat{\mathbf{x}}$ and $x_\beta \xrightarrow{\phi_\beta} \hat{\mathbf{x}}$ as well as $x_\alpha \xrightarrow{\psi_\alpha} \hat{\mathbf{x}}'$ and $x_\beta \xrightarrow{\psi_\beta} \hat{\mathbf{x}}'$ (see Figure 5).

We then have the following preliminary result:

Proposition 1: *Any $\hat{\mathbf{x}} \in \hat{\mathbf{X}}^k$ with the associated morphisms $\{\phi_\gamma : x_\gamma \to \hat{\mathbf{x}}\}_{x_\gamma \in \hat{L}^k}$ constitutes a direct limit in $\mathcal{C}^k$.*

***Proof*:** Trivial. If $\hat{\mathbf{x}} \in \hat{\mathbf{X}}^k$, then $u^k(x_\gamma) \leq u^k(\hat{\mathbf{x}})$, yielding a morphism $\phi_\gamma : x_\gamma \to \hat{\mathbf{x}}$ for every $x_\gamma \in \hat{L}^k$. Then, $\phi_\alpha = \phi_\beta \circ f_{\alpha\beta}$ for every pair $x_\alpha, x_\beta \in \hat{L}^k$. The condition of universality is obtained from the fact that between any two objects $\hat{\mathbf{x}}$ and $\hat{\mathbf{x}}'$, if say $u^k(\hat{\mathbf{x}}) < u^k(\hat{\mathbf{x}}')$ (i.e. a strict inequality), there would exist $\psi_{\hat{\mathbf{x}}} : \hat{\mathbf{x}} \to \hat{\mathbf{x}}'$ but not $\phi_{\hat{\mathbf{x}}'} : \hat{\mathbf{x}}' \to \hat{\mathbf{x}}$. Therefore, to have the full diagram and obtain universality it must be the case that $u^k(\hat{\mathbf{x}}) = u^k(\hat{\mathbf{x}}')$. Then, there exists a unique morphism $f : \hat{\mathbf{x}}' \to \hat{\mathbf{x}}$, namely $f = \psi_{\hat{\mathbf{x}}}$. □

This construction, far from being mere "abstract nonsense", is in fact quite useful for shedding light on the relation between a direct limit $\hat{\mathbf{x}}^k \in \hat{\mathbf{X}}^k$ and $\mathbf{x}^*$, the sought-after global optimum of U on $\hat{L}$. More precisely, the notion of direct limit grasps what we can think of the "best possible scenario", that is, the situation in which all local solutions are actually images under the operator Γ of a consistent global solution for some known global function.

3. A sheaf-theoretical approach

In most cases, however, the scenario presented above is highly unrealistic and overly restrictive. More often the global solution is not given, but must be sought by gluing together local ones "prospectively", in the hope of producing (or better, abducing) a consistent global result. In order to formalize this broadly abductive method for seeking a global solution, we need to take a second, slightly deeper plunge into category theory and start with the definition of a category of local problems (Figures 6 and 7 show the morphisms in the category).

$$\begin{tikzcd} \bullet s_k \arrow[d, "\rho_{kj}"] \\ \bullet s_j \end{tikzcd}$$

Figure 6. Morphism ρ_{kj} from s^k to s^j.

$$\begin{tikzcd} \bullet A \arrow[d, hook] \\ \bullet B \end{tikzcd}$$

Figure 7. Inclusion morphism representing $A \subseteq B$.

Definition 1: Let $\mathbf{P}\mathcal{R}$ be the category of local problems, where

- Obj($\mathbf{P}\mathcal{R}$) is the class of objects. Each one, $s^k = \langle \hat{L}^k, u^k, \hat{\mathbf{X}}^k \rangle$ involves the maximization of the continuous utility function u^k over the compact set $\hat{L}^k \subset L^k$, a closed linear subspace of $\mathcal{L}$, yielding a family of solutions $\hat{\mathbf{X}}^k$.
- a morphism $\rho_{kj} : s^k \to s^j$ is defined as $\hat{L}^k \subseteq \hat{L}^j$, $u^k = u^j|_{L^k}$ and $\dim(L^k) \leq \dim(L^j)$.[10]
It follows from this definition that an identity morphism $\rho_{kk} : s^k \to s^k$ trivially exists for every object s^k. Furthermore, given two morphisms $\rho_{kj} : s^k \to s^j$ and $\rho_{jl} : s^j \to s^l$ there exists their composition $\rho_{jl} \circ \rho_{kl} = \rho_{kl}$, since $\hat{L}^k \subseteq \hat{L}^j \subseteq \hat{L}^l$, $\dim(L^k) \leq \dim(L^j) \leq \dim(L^l)$ and by transitivity of the restrictions $u^k = u^j|_{L^k}$ and $u^j = u^l|_{L^j}$ we have that $u^k = u^l|_{L^k}$.

We can also define $\mathcal{P}(\mathcal{L})$ as the category in which the objects are subsets of $\mathcal{L}$ and a morphism between two objects $f_{AB} : A \to B$ is defined as $A \subseteq B$.

Let us now define now a functor

$$\Sigma : \mathbf{P}\mathcal{R} \longrightarrow \mathcal{P}(\mathcal{L})$$

which assigns to a problem $s^k = \langle \hat{L}^k, u^k, \hat{\mathbf{X}}^k \rangle$ the subset $\Sigma(s^k)$ of $\mathcal{L}$ defined by

$$\Sigma(s^k) = \{y \in \mathcal{L} \mid \Gamma_k(y) \in \hat{\mathbf{X}}^k\}$$

A *section* σ_k over s^k is simply the assignment of the elements of $\Sigma(s^k)$ to s^k:

$$\sigma_k : s^k \mapsto \Sigma(s^k).$$

Given two problems, $s^k = \langle \hat{L}^k, u^k, \hat{\mathbf{X}}^k \rangle$ and $s^j = \langle \hat{L}^j, u^j, \hat{\mathbf{X}}^j \rangle$, let us write $s^k \lhd s^j$ iff there exists a morphism ρ in $\mathbf{P}\mathcal{R}$, $\rho : s^k \to s^j$. That is, s^k is a restriction of s^j.

Let us define $r^j_k : \Sigma(s^j) \to \Sigma(s^k)$ such that to $\Sigma(s^j)$ it assigns $\Sigma(s^k)$. Given a section over s^j, r^j_k yields a section corresponding to its sub-problem s^k.

The following proposition then shows that the functor Σ possesses an important property that will be crucial for formalizing the possibility of patching up local problems and yielding a "larger" one:

Proposition 2: *Σ is a presheaf.*

Proof: $\Sigma : \mathbf{P}\mathcal{R} \rightarrow \mathcal{P}(\mathcal{L})$ is a functor. We can analyze its behavior by means of r_k^j:

- For any $s^k \in \mathrm{Obj}(\mathbf{P}\mathcal{R})$, since $s^k \lhd s^k$, $r_k^k = \mathrm{Id}_{\Sigma(s^k)}$.
- If $s^k \lhd s^j \lhd s^l$ then $s^k \lhd s^l$. Thus, $r_k^j \circ r_j^l = r_k^l$.

This means that $\Sigma : \mathbf{P}\mathcal{R} \rightarrow \mathcal{P}(\mathcal{L})$ is a *contravariant* functor. Or, in categorical terms, a *presheaf.* $\square$

Consider now a family $\{s^k = \langle \hat{L}^k, u^k, \hat{\mathbf{X}}^k \rangle\}_{k\in K} \subseteq \mathrm{Obj}(\mathbf{P}\mathcal{R})$. It is said to be a *cover* of an object $s^j = \langle \hat{L}^j, u^j, \hat{\mathbf{X}}^j \rangle$ of $\mathrm{Obj}(\mathbf{P}\mathcal{R})$ if $s^k \lhd s^j$ for each $k \in K$ and $\hat{L}^j \subseteq \cup_{k\in K} \hat{L}^k$. That is, a problem s^j gets covered by the family $\{s^k\}_{k\in K}$ if the domain of problem s^j is included in the union of the domains of the problems of the family and, furthermore, each s^k is a restriction of s^j.

The family of sections $\{\sigma_k\}_{k\in K}$ is said to be *compatible* if for any pair $k, l \in K$, given $\Sigma(s^k) = X^k$ and $\Sigma(s^l) = X^l$,

$$\Gamma_k(X^k) \cap \Gamma_l(X^k) = \Gamma_k(X^l) \cap \Gamma_l(X^l)$$

Given a cover $\{s^k\}_{k\in K}$ of a problem s^j with compatible sections, Σ is then a K-*sheaf* if there exists a unique $\sigma_j = \Sigma(s^j)$ such that for each $k \in K$,

$$\sigma_k = \sigma_j \cap \Gamma_k^{-1}(\hat{L}^k)$$

That is, intuitively, Σ is a K-sheaf if σ_j in fact “glues” together all the assignments σ_k in $\mathcal{P}(\mathcal{L})$ within the more general framework of their compatibility. Finally, then, if Σ is a K-sheaf for every $\{\sigma_k\}_{k\in K} \subseteq \mathrm{Obj}(\mathbf{P}\mathcal{R})$ it is called a *sheaf.* The following example illustrates how this construction at once generalizes and helps to clarify the previous, more rigid characterization from Section 2.

Example 3: Consider problems 1 and 2 from Example 2, denoted $s^i = \langle \hat{L}^i, u^i, \hat{\mathbf{X}}^i \rangle$ for $i = 1, 2$ as well as a new problem s^0, which is the optimization of U over the surface of the three-dimensional sphere $\hat{L}^0 = \{(x, y, z) : x^2+y^2+z^2 = 1\}$ and thus, $\hat{\mathbf{X}}^0 = \{(0, 1, 0), (0, -1, 0)\}$. Suppose that these are the only objects in $\mathbf{P}\mathcal{R}$. We define $\Sigma : \mathbf{P}\mathcal{R} \rightarrow \mathcal{P}(\mathcal{L})$, summarized by the following table (each row being a section σ_i, $i = 0, 1, 2$):

Problems	a_1	b_1	a_2	b_2
s^1	X	–	X	–
s^2	–	X	–	X
s^0	X	–	X	–

The range of Σ is based only of four elements in $\mathcal{L}$:

$$a_1 = (0, 1, 0) \quad a_2 = (0, -1, 0)$$

and

$$b_1 = \left(-\sqrt{\frac{1}{3}}, -\frac{1}{2\sqrt{3}} - \frac{1}{2}, \frac{1}{2\sqrt{3}} - \frac{1}{2}\right) \quad b_2 = \left(\sqrt{\frac{1}{3}}, \frac{1}{2\sqrt{3}} + \frac{1}{2}, \frac{1}{2} - \frac{1}{2\sqrt{3}}\right)$$

where a_1 and a_2 are the $\mathbb{R}^3$ solutions of problems s^0 and s^1 while b_1 and b_2 are those of s^2.

It is easy to check that $s^i \lhd s^0$ for $i = 1, 2$, since on one hand each problem s^i can be seen as the maximization of U restricted to subsets of the domain of problem s^0. On the other hand, $\Sigma(s^0)$ restricted to each s^i yields $\Sigma(s^i)$. In fact, for s^1 it is clear that this is the case. For s^2, let us note that b_1, b_2 are the solutions of the problem s^0 restricted to $\hat{L}^2$, seen as the inverse projection over the surface $\hat{L}^0$.

Furthermore, $\{\sigma_1, \sigma_2\}$ is a compatible family of sections. Notice that $\hat{L}^1 \cap \hat{L}^2$ does not include the solutions to either problem. But then the projections of $\hat{\mathbf{X}}^1$ and $\hat{\mathbf{X}}^2$ on $\hat{L}^1 \cap \hat{L}^2$ are both $\emptyset$, and thus the sections satisfy, trivially, the compatibility condition.

These arguments indicate that Σ satisfies the sheaf condition.

Summarizing the discussion up to this point, we can say that given a category of problems $\mathbf{P}\mathcal{R}$ over a space $\mathcal{L}$, it is typically desirable to be able to obtain a sheaf $\Sigma : \mathbf{P}\mathcal{R} \to \mathcal{P}(\mathcal{L})$, such that for any problem $\mathbf{s}^j$, covered by any compatible family of sub-problems, $\{s^k\}_{k \in K}$, $\Sigma(\mathbf{s}^j) \cap \Gamma_k^{-1}(\hat{L}^k) = \Sigma(s^k)$ for $k \in K$. The question is whether such a Σ exists and, generally, under what conditions its existence might be either guaranteed or thwarted.

4. Obstructions: non-locality and contextuality

A key idea from behavioral decision-making is that context dependence frequently matters for determining the final choices that agents actually make. Moreover, the real objective of some agent given a particular context may or may not be the maximization of the local utility functions for that agent. A systematic analysis within the less rigid framework of behavioral decision-making therefore requires a slight relaxation of the notion of local problem as presented in Definition 1 above, somewhat related to the approach in Bernheim and Rangel (2007, 2009). More precisely, we have:[11]

Definition 2: Let $\mathbf{G}\mathcal{P}\mathcal{R}$ be the category of *generalized* local problems, where

- Obj($\mathbf{G}\mathcal{P}\mathcal{R}$) is the class of objects. Each one, $s^k = \langle \hat{L}^k, u^k, \tilde{\mathbf{X}}^k \rangle$ is such that $\hat{L}^k$ and u^k are defined as in $\mathbf{P}\mathcal{R}$. But $\tilde{\mathbf{X}}^k \subseteq \hat{L}^k$ is the class of elements in $\hat{L}^k$ that yield the "highest value" for the agent.
- a morphism $\rho_{kj} : s^k \to s^j$ is defined exactly in the same way as morphisms in $\mathbf{P}\mathcal{R}$.

The only difference here is the (admittedly vague) concept of "highest value". If this is understood as achieving the maximum of u^k, we have that $\tilde{\mathbf{X}}^k = \hat{\mathbf{X}}^k$ and $\mathbf{G}\mathcal{P}\mathcal{R}$ becomes the same as $\mathbf{P}\mathcal{R}$. Behavioral decision-making is then absorbed smoothly into the mathematical framework elaborated in Section 2. What is salient then for present purposes are those situations where $\mathbf{G}\mathcal{P}\mathcal{R}$ in fact diverges from $\mathbf{P}\mathcal{R}$. Two cases, studied in the behavioral literature, in which there is a genuine difference between $\tilde{\mathbf{X}}^k$ and $\hat{\mathbf{X}}^k$ are the following:

Example 4: Consider two problems s^1, s^2, with $L^1 = \mathbb{R} = L^2$. Suppose that $u^1 = u = u^2$ (with u a strictly concave function) and that $\omega^1 \in \mathbb{R}$ is understood as the money owned in s^1, while $\omega^2 \in \mathbb{R}$ is owned in s^2. Furthermore, $\hat{L}^1 = \{x \in \mathbb{R} : u(\omega^1 + x) = u(\omega^2)\}$ and $\hat{L}^2 = \{x \in \mathbb{R} : u(\omega^2 - x) = u(\omega^1\}$. Suppose that in both cases we look for the "best" x.

If "best" means the maximization of u^k over $\hat{L}^k$, $\tilde{\mathbf{X}}^1 = \{\omega^2 - \omega^1\} = \tilde{\mathbf{X}}^2$.

On the other hand, according to prospect theory (Kahneman and Tversky 1979), "best" in s^1 means maximizing psychological gain with respect to a reference point (owning ω^1), while in s^2 it means minimizing psychological loss (down from ω^2). Since it assumes an asymmetry in the evaluation of gains and losses, the same amount x has a different value in one case as compared to the other and thus $\tilde{\mathbf{X}}^1 \neq \tilde{\mathbf{X}}^2$.

Example 5: Consider a sequence of problems $s^1, s^2, \ldots, s^K$ with $L^k = L$ and $u^k = u$ for $k = 1, \ldots, K$, and $\cap_{k=1}^K \hat{L}^k \neq \emptyset$. According to sase-based decision theory[12] (Gilboa and Schmeidler 2001) we can build a *memory of cases* $M = \{(s^k, x^k) : x^k \in \tilde{\mathbf{X}}^k, k = 1, \ldots, K - 1\}$. That is, a record of the problems with one element chosen for yielding the highest value in those problems. Furthermore, a *similarity* function sim: $\{s^k\}_{k=1,\ldots K} \times \{s^k\}_{k=1,\ldots K} \to \mathbb{R}$, provides a closeness relation between problems. Then at problem s^K we will have

$$\tilde{\mathbf{X}}^K = \left\{ x : x \in \operatorname{argmax}_{y \in \hat{L}^K} \sum_{(s^k, y)} \operatorname{sim}(s^k, s^K) u^k(y) \right\}$$

i.e. $\tilde{\mathbf{X}}^K$ consists of the elements in $\cup_{k=1}^{K-1} \tilde{\mathbf{X}}^k \cap \hat{L}^K$ that maximize the weighted (by similarity) sum of local utility functions of the previously solved problems.

But then, if problem s^K is solved after an alternative sequence $s^{1\prime}, \ldots, s^{(K-1)\prime}$ (with a different memory M' and with different similarity weights with respect to s^K), we might end up with a set $\tilde{\mathbf{X}}^K$ different from the one found with the sequence $s^1, \ldots, s^{(K-1)}$.

Notice that Example 4 exhibits a dependence on context, while Example 5 involves the non-locality of solutions. These features imply, in our setting, that $\Sigma : \mathbf{G}\mathcal{PR} \to \mathcal{P}(\mathcal{L})$ does not necessarily have to be a sheaf. In effect, given a problem s, the sheaf condition implies that its solution remains independent of other solutions and thus it disregards their contextual relevance. Analogously, if we consider two sequences $s^1, \ldots, s^n$ and $s^{1'}, \ldots, s^{n'}$ in $\mathrm{Obj}(\mathbf{G}\mathcal{PR})$, such that $s^n = s = s^{n'}$, understood as two different paths (of problems previously solved), the sheaf condition implies that the solution to s is independent of the path followed. That is, the solution is purely local.

Indeed, it is possible to demonstrate that the sheaf property, and thus the reconstruction of the global solution up from the local ones, obtains only if two conditions are fulfilled by each problem s^k:[13]

- The elements in $\tilde{\mathbf{X}}^k$ are the maximizers of u^k.
- u^k is the constraint of a single function (U) over L^k.

To establish this claim we start by defining $\Lambda : \mathcal{P}(\mathcal{L}) \to \mathbf{G}\mathcal{PR}$ as follows. For any $X \in \mathcal{P}(\mathcal{L})$:

$$\Lambda(X) = \left\{ s^k = \langle \hat{L}^k, u^k, \tilde{\mathbf{X}}^k \rangle \in \mathrm{Obj}(\mathbf{G}\mathcal{PR}) : X = \Gamma_k^{-1}(\tilde{\mathbf{X}}^k) \right\}$$

That is, given $X \subseteq \mathcal{L}$, Λ yields the problems that have as solutions the projections of X.

It is rather intuitive that Λ has a close connection with Σ:

Proposition 3: *For any $s \in Obj(\mathbf{G}\mathcal{PR})$, $s \in \Lambda(\Sigma(s))$.*

***Proof*:** Given $s = \langle \hat{L}, u, \tilde{\mathbf{X}} \rangle$, any $\tilde{x} \in \tilde{\mathbf{X}}$ is such that the corresponding x in $\Sigma(s)$, restricted to $\hat{L}$ is $\tilde{x}$. Then, by definition s must belong to $\Lambda(\Sigma(s))$. □

On the other hand, if we have a cover $\{s^k = \langle \hat{L}^k, u^k, \tilde{\mathbf{X}}^k \rangle\}_{k \in K}$ of a problem $s = \langle \hat{L}, u, \tilde{\mathbf{X}} \rangle$, and for each $k \in K$ we have $\Gamma_k^{-1} : \hat{L}^k \to \mathcal{L}$, we may then formulate a converse of Proposition 3:

GGEN 1355913
17-10-2017 Revision CE:PG QA:RB

$$\begin{tikzcd} \Sigma(s^k) \arrow[r, "\lambda"] \arrow[dr, "="'] & s^k \times \Sigma(\mathbf{s}) \arrow[d, "p_2 \cap \Gamma_k^{-1}(\hat{L}^k)"] \\ & \Sigma(\mathbf{s}) \cap \Gamma_k^{-1}(\hat{L}^k) \end{tikzcd}$$

Figure 8. Commutative diagram, showing that $\Sigma(s^k) = \Sigma(\mathbf{s}) \cap \Gamma_k^{-1}(\hat{L}^k)$.

Proposition 4: *If $\bigcup_{k \in K} \Gamma_k^{-1}(\tilde{\mathbf{X}}^k) = \tilde{\mathbf{X}}$ then for every $s^k \in \Lambda(\Sigma(s))$, $s^k \lhd s$.*

Proof: Consider a problem s^k in $\Lambda(\Sigma(s))$. Then, there exists $x \in \Sigma(s)$ such that a $\tilde{x}^k \in \tilde{\mathbf{X}}^k$ is $\Gamma_k(x)$. Any such x is thus $x = \Gamma_k^{-1}(\tilde{x}^k)$, and thus $s^k \lhd s$. □

Under the conditions of Proposition 4, Λ can be seen as a *fiber bundle* (Hatcher 2002). This means that, on every problem s, Λ^{-1} is isomorphic to $s \times B_s$, where B_s is a *fiber*. In the particular case that $B_s = B_{s'}$ for any pair of problems s, s', Λ is said to be a *trivial bundle*. More precisely, Λ is trivial if given the global problem $\mathbf{s} = \langle \mathcal{L}, U, \tilde{X} \rangle$ and any problem s^k in $\mathbf{G}\mathcal{PR}$ we have that:

$$\lambda : \Lambda^{-1}(s^k) \to s^k \times \Sigma(\mathbf{s})$$

is an isomorphism.

Lemma 1: *If Λ is such that for every problem s^k, $\Lambda(\Sigma(s^k)) = \{s^k\}$, Λ is trivial iff $\Sigma(s^k) = \Sigma(\mathbf{s}) \cap \Gamma_k^{-1}(\hat{L}^k)$.*

Proof: By the inverse relation between Λ and Σ, we have that $\Lambda^{-1}(s^k) = \Sigma(s^k)$. On the other hand, λ is a functor between categories $\mathcal{P}(\mathcal{L})$ and $\mathbf{G}\mathcal{PR} \times \mathcal{P}(\mathcal{L})$. Below is the diagram corresponding to this relation, where p_2 is the projection over the second component of a product.

Notice that the triviality of Λ implies the commutativity of this diagram, and thus the isomorphism of $\Sigma(s^k)$ and $\Sigma(\mathbf{s}) \cap \Gamma_k^{-1}(\hat{L}^k)$. Since this is an isomorphism in $\mathcal{P}(\mathcal{L})$, $\Sigma(s^k) = \Sigma(\mathbf{s}) \cap \Gamma_k^{-1}(\hat{L}^k)$.

Conversely, if $\Sigma(s^k) = \Sigma(\mathbf{s}) \cap \Gamma_k^{-1}(\hat{L}^k)$, the diagram in Figure 8 is commutative and thus Λ is trivial. □

The triviality of Λ has an important consequence: the absence of *obstructions* to the existence of global sections of Σ. In our setting this may be formulated as follows:[14]

Proposition 5: *If for every $s^k = \langle \hat{L}^k, u^k, \tilde{\mathbf{X}}^k \rangle$ in $\mathbf{G}\mathcal{PR}$, $\Lambda(\Sigma(s^k)) = \{s^k\}$ then Λ is trivial iff there exists $U : \mathcal{L} \to \Re$, such that u^j has the same optimal points as $U_{|L^j}$.*

Proof: $\Leftarrow$) Since $\mathbf{s} = \langle \hat{L}, U, \tilde{\mathbf{X}} \rangle$, it is clear that any $s^k = \langle \hat{L}^k, U_{|L^k}, \tilde{\mathbf{X}}^k \rangle$, and $\Gamma_k(\tilde{\mathbf{X}}) = \tilde{\mathbf{X}}^k$, and thus, $\Sigma(s^k) = \Sigma(\mathbf{s}) \cap \Gamma_k^{-1}(\hat{L}^k)$. Therefore, by Lemma 1, Λ is trivial.

$\Rightarrow$) If Λ is trivial, by Lemma 1 $\Sigma(s^k) = \Sigma(\mathbf{s}) \cap \Gamma_k^{-1}(\hat{L}^k)$ for every s^k. Then, there exists $U : \mathcal{L} \to \Re$ such that for each u^k, $\{x : x = \text{argmax } u^k\} = \{x : x = \text{argmax } U_{|L^k}\}$. Suppose by contradiction that for every possible $U : \mathcal{L} \to \Re$ and for some u^k, $\{x : x = \text{argmax } u^k\} \neq \{x : x = \text{argmax } U_{|L^k}\}$. But then, $\Gamma_k(\tilde{\mathbf{X}}) \neq \tilde{\mathbf{X}}^k$ and thus $\Sigma(s^k) \neq \Sigma(\mathbf{s}) \cap \Gamma_k^{-1}(\hat{L}^k)$. Contradiction. □

This result indicates therefore that locality and non-contextuality as expressed in the categorical framework are conjointly equivalent to the existence of a global utility function U, such that the utility functions of particular problems are restrictions of U over their corresponding option spaces (up to transformations that preserve maxima). Thus, in particular, the economic phenomena modeled routinely by prospect theory and case-based decision theory exclude such global-U models.

5. Conclusions

In this paper, we have shown how framing the problem of maximizing utility functions within a categorical-theoretical language allows for the extraction of enough information from local solutions of a problem to recover a global one if the decision-making process is local and non-contextual. More specifically, we showed that the existence of a sheaf over a category of problems indicates the recoverability of global solutions up from local ones. But a slight generalization of the category of problems (incorporating cases studied by behavioral decision theory) poses obstructions to the existence of global sections yielding global solutions. The requirements for the existence of those global sections amount to seeing local problems as restrictions of the global one. In practice, of course, this is usually not feasible since the local problems present themselves typically without concern for a global solution. This result reveals in a formal register the principle underlying why the different behavioral models may not be captured in a single framework in competition with the traditional one in the discipline. Notice that this is true even if there exist behavioral models not captured in our framework, since at the very least the latter cannot be put in a single framework. That is, behavioral decision theory will always consist of disparate models of effects and biases involved in actual decision-making, without any possible unifying final explanation other than, perhaps, the physical complexities of the human brain. This is by no means a demerit for this field of research, but is rather an indication of a deep and essential difference with respect to the traditional approach, which remains essentially "mindless" (Gul and Pesendorfer 2008).

Notes

1. There is a large literature on how to approach this problem numerically, by the use of more general and flexible forms utilizing approximating functions which are as close as possible to the true solution (see, for instance, Judd (1996)).
2. Notable exceptions are, for instance, Masatlioglu and Ok (2005), Bernheim and Rangel (2007), Salant and Rubinstein (2008), Bernheim and Rangel (2009) and Masatlioglu and Uler (2013), in which different single choice frameworks for behavioral decision-making are developed.
3. This establishes an intriguing analogy with the relation between classical and quantum mechanics (Abramsky and Brandenburger 2011).
4. Behavioral economists weaken this definition by allowing the agent to choose what *it seems* to her to be the most preferred alternative (Kahneman 2013). As such, it is no longer normative nor a general explanatory principle, since it requires specific hypotheses not only about the preferences of the agents but also their possible beliefs and cognitive biases.
5. This yields an *abduction*, as described in Tohmé, Caterina, and Gangle (2015).
6. Some notable exceptions are Ghani and Hedges (2016), Hedges et al. (2016), Abramsky and Winschel (2017) and Rozen and Zhitomirski (2006). In turn Crespo and Tohmé (2016) presents arguments for the adoption of the categorical language in Economics.

7. The meaning of these options depend on the context. If the agent is a consumer in a competitive market with a finite number of goods, she has to choose a vector of those commodities. In a planning problem she has to select a plan specifying the amounts of resources used or consumed at each period of time.
8. That is, $|x - x^k| = \min_{y \in L^k} |x - y|$, where $| \cdot |$ is the norm of $\mathcal{L}$.
9. See, among others, Luenberger (2001), where these methods are employed to model pricing assets whose payoffs are outside the span of marketed assets.
10. $\dim(\cdot)$ yields the dimension of a subspace of $\mathcal{L}$.
11. The different notions defined for $\mathbf{P}\mathcal{R}$ are applied to $\mathbf{G}\mathcal{PR}$ replacing, when necessary, $\tilde{\mathbf{X}}^k$ for $\hat{\mathbf{X}}^k$.
12. According to this theory, individuals make choices based on their previous experiences. More precisely, each agent is endowed with both a preference relation over actions and a memory set which contains his past decisions and their corresponding results valued in "utiles". A utility function over the actions that can be taken to solve a problem is characterized as the sum of utilities derived from past actions weighted up by the similarity between the actual problem and those stored in the memory.
13. We adapt to our setting the sheaf-theoretical analysis of non-locality and contextuality in Quantum Mechanics (Abramsky and Brandenburger 2011).
14. An economic problem, unrelated to Behavioral Economics, in which triviality fails is the following. Consider a policy-maker, who can use three instruments: (a) the exchange rate; (b) the degree of openness in capital movements; and (c) the possibility of implementing an independent monetary policy. If the problem is restricted to choose combinations of (a) and (b), it is optimal to have a fixed exchange rate and free capital flows. In turn, if combinations of (a) and (c) are considered, the optimal solution is to have a fixed exchange rate and an independent monetary policy. Finally, when the problem is restricted to (b) and (c) the solution is to have free flows of capital and an independent monetary policy. On the other hand, it is well known that when the problem involves (a), (b), and (c) together, there is no solution. This is known as the monetary policy *Trilemma* (Obstfeld and Taylor 1997).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Tohmé was supported by the National Research Council of Argentina [grant number PIP 112-200801-00804]; Universidad Nacional del Sur's [grant number PGI 24/E115]. Endicott College provided additional funding and research support during the preparation of this paper.

Notes on contributors

Fernando Tohmé is a principal researcher of CONICET (National Research Council of Argentina) and a full professor at the Department of Economics of the Universidad Nacional del Sur, in Bahía Blanca, Argentina. He holds an undergraduate degree in Mathematics and a PhD in Economics (under the late Rolf Mantel). A former Fulbright Scholar, he held visiting positions at U.C. Berkeley, Washington University in St. Louis and Endicott College. His research has focused on decision problems, game theory, and optimization in socioeconomic settings. He is currently working on a category-theoretical reformulation of game theory. Tohmé has been the thesis advisor of more than fifteen doctoral students in Economics, Engineering, and Mathematics.

He has published research articles in peer-reviewed journals like Theory and Decision, Mathematics of Social Sciences, Journal of Applied Logic, Physica A, Artificial Intelligence, Mathematical and Computational Modeling, and Annals of Operations Research among others.

Gianluca Caterina is an associate professor of Mathematics at Endicott College. His research interests span the fields of Discrete Dynamical Systems, Category Theory, Logic, and Epistemology. His work has appeared on journals which include Physica A, Theoretical Computer Science, Synthese, and the Journal of Applied Logic among others. He coauthored, with Rocco Gangle, the book "Iconicity and Abduction" which was published in 2016 by Springer-Verlag. His research currently focuses on applications of Category Theory in Game Theory.

Rocco Gangle specializes in philosophies of immanence (particularly Spinoza, Deleuze and Laruelle), the semiotics and diagrammatic logic of Peirce, and the mathematics of category theory. He is the translator of Laruelle's Philosophies of Difference (Continuum 2010) and the author of Diagrammatic Immanence: Category Theory and Philosophy (Edinburgh UP 2015) and Francois Laruelle's Philosophies of Difference: A Critical Introduction and Guide (Edinburgh UP 2013). His work has appeared in Philosophy Today, Synthese, the Journal of Applied Logic, Angelaki, SubStance, and Political Theology among other journals. With Gianluca Caterina, he directs the Center for Diagrammatic and Computational Philosophy at Endicott College, where he is also an associate professor of Philosophy. Gangle is also a member of the Global Center of Advanced Studies' Forms and Formalization Research Group.

ORCID

Fernando Tohmé http://orcid.org/0000-0003-2988-4519

References

Abramsky, S., and A. Brandenburger. 2011. "The Sheaf-theoretic Structure of Non-locality and Contextuality." *New Journal of Physics* 13: 113036 (39 pages).

Abramsky, S., and V. Winschel. 2017. "Coalgebraic Analysis of Subgame-perfect Equilibria in Infinite Games without Discounting." *Mathematical Structures in Computer Science* 27: 751–761.

Adámek, J., H. Herrlich, and G. Strecker. 2004. "Abstract and Concrete Categories." In *The Joy of Cats*. http://www.iti.cs.tu-bs.de/

Baas, N. 2015. "Higher Order Architecture of Collections of Objects." *International Journal of General Systems* 44: 55–75.

Baas, N. 2016. "On Higher Structures." *International Journal of General Systems* 45: 747–762.

Barr, M., and Ch. Wells. 1999. *Category Theory*. http://www.ling.ohio-state.edu/.

Bernheim, D., and A. Rangel. 2007. "Toward Choice-theoretical Foundations for Behavioral Welfare Economics." *American Economic Review* 97: 464–470.

Bernheim, D., and A. Rangel. 2009. "Beyond Revealed Preference: Choice Theoretic Foundations for Behavioral Welfare Economics." *Quarterly Journal of Economics* 124: 51–104.

Burgess, J. 1998. "Occam's Razor and Scientific Method." In *The Philosophy of Mathematics Today*, edited by M. Schirn. New York: Clarendon Press.

Camerer, C., G. Loewenstein, and M. Rabin. 2004. *Advances in Behavioral Economics*. New York: Russell Sage Foundation and Princeton University Press.
Cartwright, E. 2011. *Behavioral Economics*. London: Routledge.
Crespo, R., and F. Tohmé. 2016. "The Future of Mathematics in Economics: A Philosophically Grounded Proposal." *Foundations of Science*. doi:10.1007/s10699-016-9492-9.
Chambers, C., and F. Echenique. 2016. *Revealed Preference Theory*. Cambridge: Cambridge University Press.
Ghani, N., and J. Hedges. 2016. "A Compositional Approach to Economic Game Theory." arXiv:1603.04641 [cs.GT].
Ghrist, R. 2014. *Elementary Applied Topology*. Lexington, KY: CreateSpace Independent Publishing Platform.
Gilboa, I., and D. Schmeidler. 2001. *A Theory of Case-based Decisions*. Cambridge: Cambridge University Press.
Goldblatt, R. 1984. *Topoi. The Categorical Analysis of Logic*. Amsterdam: North-Holland.
Gul, F., and W. Pesendorfer. 2008. "The Case for Mindless Economics." In *The Foundations of Positive and Normative Economics*, edited by A. Caplin and A. Shotter. Oxford: Oxford University Press.
Hartford, T. 2014. "Behavioural Economics and Public Policy." *Financial Times*, March 21. http://www.ft.com/intl/cms/s/2/9d7d31a4-aea8-11e3-aaa6-00144feab7de.html#axzz2wcp1Log3.
Hatcher, A. 2002. *Algebraic Topology*. Cambridge: Cambridge University Press.
Hedges, J., E. Shprits, V. Winschel, and P. Zahn. 2016. "Compositionality and String Diagrams for Game Theory." arXiv:1604.06061 [cs.GT].
Judd, K. 1996. Approximation, Perturbation, and Projection Solution Methods in Economics. In *Handbook of Computational Economics*, edited by H. Amman, D. Kendrick and J. Rust. Amsterdam: North Holland.
Kahneman, D. 2013. *Thinking, Fast and Slow*. NY: Farrar, Straus and Giroux.
Kahneman, D., and A. Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47: 263–291.
Lawvere, F. W., and S. Schanuel. 2009. *Conceptual Mathematics: A First Introduction to Categories*. 2nd ed. Cambridge: Cambridge University Press.
Luenberger, D. 2001. "Projection Pricing." *Journal of Optimization Theory and Applications* 109: 1–25.
Macfarlane, A. 2014. "Structures of Systems 1." *Cohomology of Manufacturing and Supply Network-like Systems, International Journal of General Systems* 43: 470–507.
Marquis, J.-P. 2009. *From a Geometrical Point of View: A Study of the History and Philosophy of Category Theory*. Berlin: Springer-Verlag.
Mas-Colell, A., M. Whinston, and J. Green. 1995. *Microeconomic Theory*. New York: Oxford University Press.
Masatlioglu, Y., and E. Ok. 2005. "Rational Choice with Status-quo Bias." *Journal of Economic Theory* 121: 1–29.
Masatlioglu, Y., and N. Uler. 2013. "Understanding the Reference Effect." *Games and Economic Behavior* 82: 403–423.
Obstfeld, M., and A. Taylor. 1997. "The Great Depression as a Watershed: International Capital Mobility over the Long Run." NBER Working Paper 5960.
Plikynas, D., G. Basinskas, P. Kumar, S. Masteika, D. Kezys, and A. Laukaitis. 2014. "Social Systems in Terms of Coherent Individual Neurodynamics: Conceptual Premises." *Experimental and Simulation Scope, International Journal of General Systems* 43: 434–469.
Rozen, V., and G. Zhitomirski. 2006. "A Category Theory Approach to Derived Preference Relations in some Decision Making Problems." *Mathematical Social Sciences* 51: 257–273.
Salant, Y., and A. Rubinstein. 2008. "*(A, f)* Choice with Frames." *Review of Economic Studies* 75: 1287–1296.
Spivak, D. I. 2014. *Category Theory for the Sciences*. Cambridge, MA: MIT Press.
Tohmé, F., G. Caterina, and R. Gangle. 2015. "Abduction: A Categorical Characterization." *Journal of Applied Logic* 13: 78–90.