Local and Global Optima in Decision-Making: a Sheaf-theoretical Analysis of the Difference between the Classical and the Behavioral Approach to Decision-Making.

Fernando Tohmé.

INMABB - Conicet and Universidad Nacional del Sur
Argentina

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 poses the question of whether this competing approach can provide an encompassing alternative to the classical one. We analyze this issue in the 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, usual in the literature on Behavioral decision making, in which such properties do not hold. In the first case, we show that a sheaf-theoretical representation provides an abstract characterization of the global solution. In the latter case, we show that locality and contextuality generate an obstruction towards the reconstruction of global solutions, yielding a possible clue for the intrinsic difference between Behavioral and classical decision theory.

Palabras Clave: Decision-making; Behavioral Decision Theory; Optimal choices; Sheaves; Projections; Category Theory.

Local vs. Global in Decision-Making: on the Algebraic Topology of Behavioral Economics[1]

[1] Joint with G. Caterina and R. Gangle

Outline of the Paper

- Decision making: local problems and global solutions.
- Sheaves and the reconstruction of global solutions.
- Behavioral economics: non-locality and contextuality.
- Obstructions and a possibility condition.

Local problems

- Consider a family of local *problems* each with its own domain, say D_i and each with a problem-specific function u_i.

- Hypothesis: there exist a global function U over D ($D_i \subseteq D$ for each problem).

- We look for a function V such that $u = V_{|D}$. Furthermore, V should yield the same global solution as U.

- To obtain V, which recovers the hypothetical U, we must be able to patch together the local restrictions in a consistent way.

A localization operator *Loc* reduces the global maximization problem to a sequence of local ones:

$$
\begin{tikzcd}
 & & U \arrow[ddll, "Loc"'] \arrow[ddl, "Loc"] \arrow[d, "\sim"] \arrow[ddrr, "Loc"] & & \\
 & & V & & \\
u_1 \arrow[urr, dotted] & u_2 \arrow[ur, dotted] & \cdots & \cdots & u_n \arrow[ull, dotted]
\end{tikzcd}
$$

Main result

- It is entirely possible that there might not be any global U such that each u is its restriction to the corresponding domain.

- Only if non-contextuality and locality are properties of the decision-making process is it possible to obtain a global result via the patching-up of local solutions.

Decision-making: local vs. global

- Let $\mathcal{L}$ be a space of possible **options** that an agent may select.
- Each $x \in \mathcal{L}$ is evaluated by means of a *utility* function, $U : \mathcal{L} \to \Re$.
- Given a family of constraints limiting the set of options for the agent to $\hat{L} \subseteq \mathcal{L}$, the goal of the agent is to find some $\mathbf{x}^*$ that maximizes U over $\hat{L}$.

Decision-making: local vs. global

- Consider a family $\{L^k\}_{k=0}^{\kappa}$ of closed linear subspaces of $\mathcal{L}$.

- Let us define

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

  such that $\mathrm{Proj}_k(x) = x^k \in L^k$, where x^k is the *projection* of x on L^k.

- The projection of a global solution $\mathbf{x}^*$ onto L^k will return the point in L^k which is the closest to $\mathbf{x}^*$.

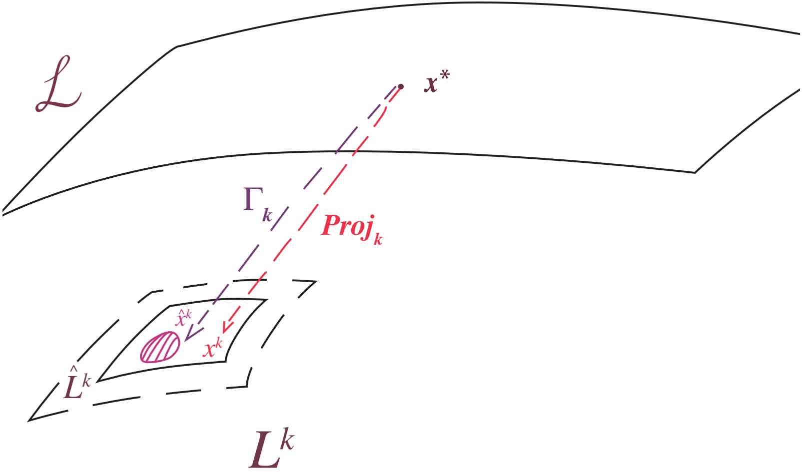
Decision-making: local vs. global

- 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.

- To analyze this problem, let us define a new correspondence, $\Gamma_k : \hat{L} \to \hat{L}^k$:

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

Local vs. global, *Math with bad drawings*-style



Decision-making: local vs. global

We are interested in the set given by $\Gamma_k(\mathbf{x}^*)$. In the following 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.

$$
\begin{tikzcd}
 & & \{\varnothing\} \arrow[d, dotted, "\mathbf{x}^*"] \arrow[dddll, bend right, "\hat{\mathbf{x}}^1"'] \arrow[dddrr, bend left, "\hat{\mathbf{x}}^\kappa"] & & \\
 & & \mathcal{L} \arrow[d, "U"] \arrow[ddll, dotted, "\Gamma_1"'] \arrow[ddrr, dotted, "\Gamma_\kappa"] & & \\
 & & \mathbb{R} & & \\
\hat{L}^1 \arrow[uurr, dotted, "u^1"'] & \dots & \dots & & \hat{L}^\kappa \arrow[uull, dotted, "u^\kappa"]
\end{tikzcd}
$$

Example

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

$$\hat{\mathbf{X}} = \{(0, 0, 0)\}.$$

We will consider two possible "local" problems.

Example

- $L^1 = \{(x,y,z) : z = 0\}$, *where*

$$u^1(x,y,z) \;=\; U_{|L^1} \;=\; 3 - 2x^2 - y^2$$

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

$$\hat{\mathbf{X}}^1 = \{(0,1,0),(0,-1,0)\}.$$

Example

- $$L^2 = \{(x,y,z) : (x,y,z)\cdot(1,-1,1) = 0\}$$

  (i.e. the linear subspace with normal vector $(1,-1,1)$*), where*

  $$u^2(x,y,z) \;=\; 3 - 3x^2 - 4z^2 - 2xz,$$

  is 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:*

  $$\hat{\mathbf{X}}^2 = \{(-\sqrt{\frac{1}{3}}, -\frac{1}{2\sqrt{3}} - \frac{1}{2}, \frac{1}{2\sqrt{3}} - \frac{1}{2}), (\sqrt{\frac{1}{3}}, \frac{1}{2\sqrt{3}} + \frac{1}{2}, \frac{1}{2} - \frac{1}{2\sqrt{3}})\}.$$

Example

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 category of local problems

Definition

A *local problem* is $s^k = \langle \hat{L}^k, u^k, \hat{\mathbf{X}}^k \rangle$. It involves 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

$$\hat{\mathbf{X}}^k = \{\hat{\mathbf{x}} : u^k(\hat{\mathbf{x}}) \geq u^k(x) \text{ for every } x \in \hat{L}^k\}.$$

The category of local problems

Definition

Let $\mathcal{PR}$ be the category of local problems, where

- $\mathrm{Obj}(\mathcal{PR})$ is the class of objects. Each one is a problem s^k.
- a morphism $\rho_{kj} : s^k \to s^j$ exists if two conditions are fulfilled:
  - $\hat{L}^k \subseteq \hat{L}^j$, $u^k = u^j|_{L^k}$ and
  - $\dim(L^k) \leq \dim(L^j)$.
- Given two morphisms $\rho_{kj} : s^k \to s^j$ and $\rho_{jl} : s^j \to s^l$ there exists their composition $\rho_{jl} \circ \rho_{kl} = \rho_{kl}$.

The sheaf of local problems

- 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 $A \subseteq B$.

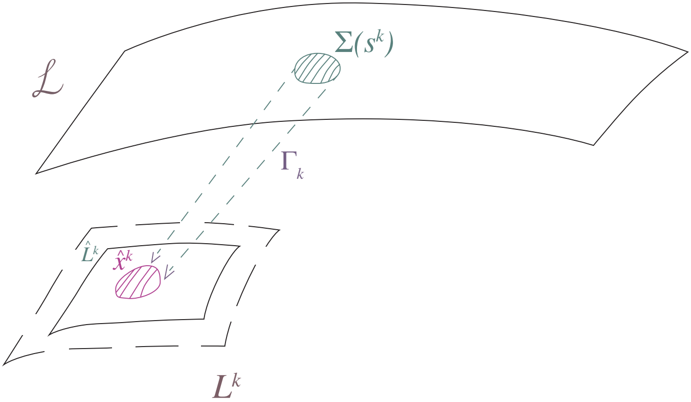
- We can now define a functor

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

which assigns to a problem s^k the subset $\Sigma(s^k) \subseteq \mathcal{L}$:

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

The sheaf of local problems



The sheaf of local problems

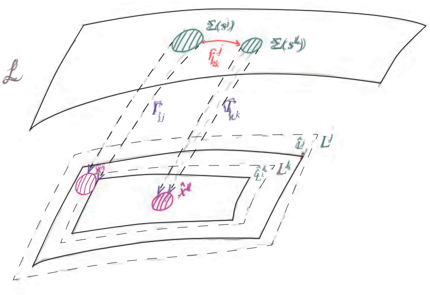
- A *section* σ_k over s^k is simply the assignment of the elements of $\Sigma(s^k)$ to s^k:

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

- Given two problems, s^k and s^j, let us write $s^k \lhd s^j$ iff s^k is a restriction of s^j.

- Finally, given $s^k \lhd s^j$ let us define r^j_k, assigning to section $\Sigma(s^j)$ the section corresponding to its sub-problem s^k.

The sheaf of local problems



The sheef of local problems

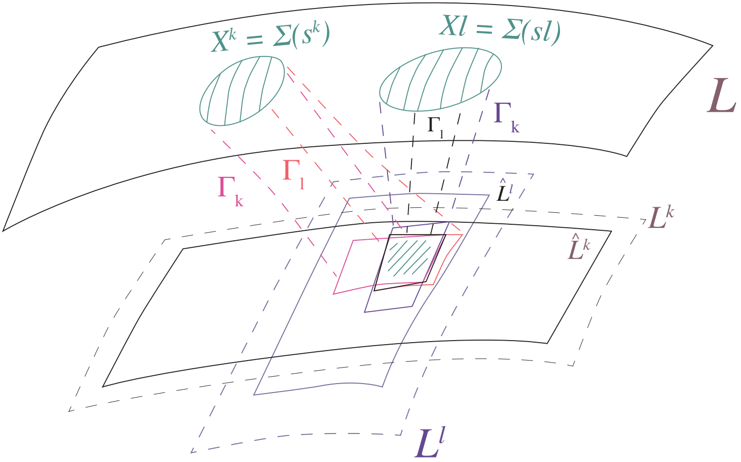
- Σ, so defined is a presheaf (i.e. a *contravariant* functor between $\mathcal{PR}$ and $\mathcal{P}(\mathcal{L})$).

- A family $\{s^k\}_{k\in K} \subseteq \mathrm{Obj}(\mathcal{PR})$ is said to be a *cover* of problem s^j if $s^k \lhd s^j$ for each $k \in K$ and $\hat{L}^j \subseteq \cup_{k\in K} \hat{L}^k$.

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

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

The sheaf of local problems



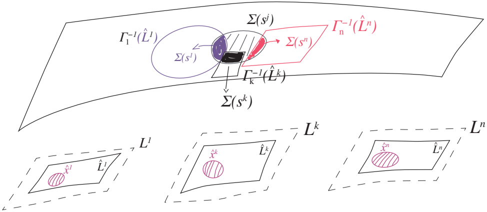
The sheaf of local problems

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

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

- Intuitively, Σ is a sheaf if σ_j in fact "glues" together all the assignments σ_k in $\mathcal{P}(\mathcal{L})$.

The sheaf of local problems



Example

Consider the problems 1 and 2 from the previous example, 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,*

$$\hat{\mathbf{X}}^0 = \{(0, 1, 0), (0, -1, 0)\}.$$

Example

We define $\Sigma : \mathcal{PR} \to \mathcal{P}(\mathcal{L})$, 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, a_2 and b_1, b_2 which are the $\mathbb{R}^3$ solutions of problems s^0 and s^1, respectively.

Example

- *It is easy to check that* $s^i \lhd s^0$ *for* $i = 1, 2$. *On the other hand,* $\Sigma(s^0)$ *restricted to each* s^i *yields* $\Sigma(s^i)$.

- $\{\sigma_1, \sigma_2\}$ *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 either* $\hat{\mathbf{X}}^1$, $\hat{\mathbf{X}}^2$ *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.*

Generalizing the concept of problem

Definition

Let $\mathcal{GPR}$ be the category of *generalized* local problems, where

- $\text{Obj}(\mathcal{GPR})$ 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 $\mathcal{PR}$. 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 $\mathcal{PR}$.

Generalizing the concept of problem

- The only difference between $\mathcal{GPR}$ and $\mathcal{PR}$ is the (admittedly vague) concept of "highest value".
- If this understood as achieving the maximum of u^k we have that $\tilde{\mathbf{X}}^k = \hat{\mathbf{X}}^k$ and $\mathcal{GPR}$ becomes the same as $\mathcal{PR}$.
- Behavioral Economics is then absorbed smoothly into the mathematical framework.
- What is salient then for present purposes are those situations where $\mathcal{GPR}$ in fact diverges from $\mathcal{PR}$.

Example

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-Tversky), "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). Thus*

$$\tilde{\mathbf{X}}^1 \neq \tilde{\mathbf{X}}^2.$$

Example

- *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 Case-based Decision Theory (Gilboa-Schmeidler) 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 and 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.*

Example

- *Then at problem s^K we will have*

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

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 with a set $\tilde{\mathbf{X}}^K$ different from the one found with the sequence $s^1, \ldots, s^{(K-1)}$.*

Contextuality and non-locality

- The prospect theory example exhibits a dependence on context, while the CBDM one involves the non-locality of solutions.

- These features imply, in our setting, that $\Sigma : \mathcal{GPR} \to \mathcal{P}(\mathcal{L})$ does not necessarily have to be a sheaf.

- 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 $s1', \ldots, s^{n'}$ in $\mathrm{Obj}(\mathcal{GPR})$, such that $s^n = s = s^{n'}$, understood as two different paths (of problems previously solved).

- Thus, the sheaf condition implies that the solution to s is independent of the path followed. That is, the solution is purely local.

Contextuality and non-locality

Proposition

If for every s^k in $\mathcal{GPR}$ we have that:

- *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.*

Then $\Sigma : \mathcal{GPR} \to \mathcal{P}(\mathcal{L})$ is a sheaf.

Contextuality and non-locality

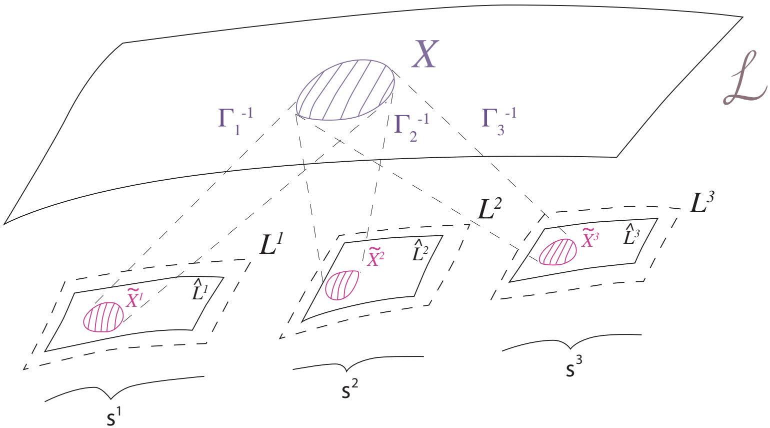
- To establish this claim we start by defining a functor $\Lambda : \mathcal{P}(\mathcal{L}) \rightarrow \mathcal{GPR}$.

- For any $X \in \mathcal{P}(\mathcal{L})$:

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

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

Contextuality and non-locality



$$\Lambda(x) = \{s^1, s^2, s^3\}$$

Proposition

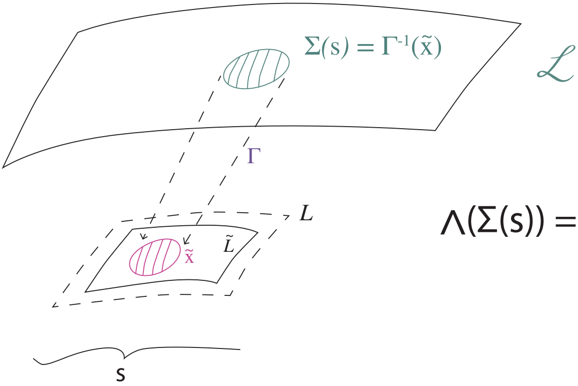
For any $s \in Obj(\mathcal{GPR})$, $s \in \Lambda(\Sigma(s))$.

and

Proposition

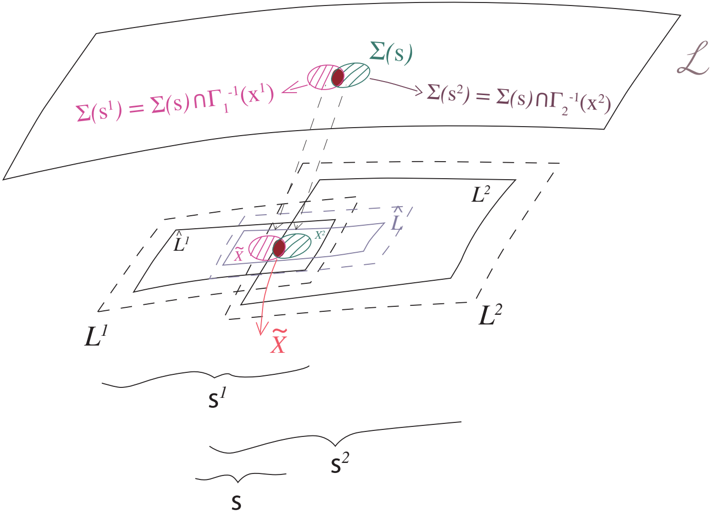
If $\bigcup_{k \in K} \Gamma_k^{-1}(\tilde{\mathbf{X}}^k) = \tilde{\mathbf{X}}$ *then* $\Lambda(\Sigma(s)) \subseteq \{s\}$.

Contextuality and non-locality



$$\Lambda(\Sigma(s)) = \{ \dots , s , \dots \}$$

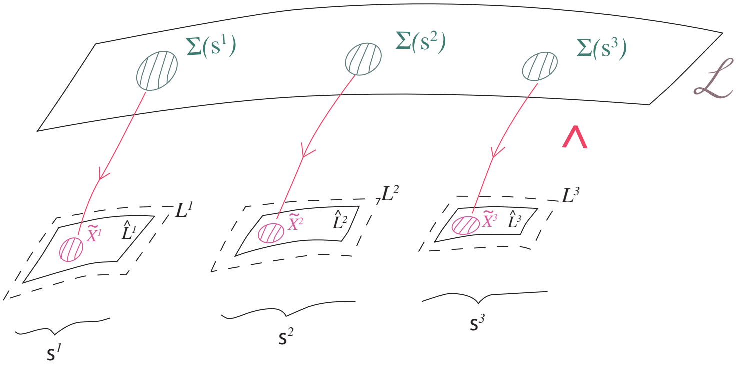
Contextuality and non-locality



Contextuality and non-locality

- Under the conditions of the last Proposition, Λ can be seen as a *fiber bundle*.

Contextuality and non-locality



Contextuality and non-locality

- 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*.

Λ is trivial if given the global problem $\mathbf{s} = \langle \mathcal{L}, U, \tilde{X} \rangle$ and any problem s^k in $\mathcal{GPR}$ we have that:

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

is an isomorphism.

Contextuality and non-locality

Proposition

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

Contextuality and non-locality