mathematics

Article

Level-Agnostic Representations of Interacting Agents

Fernando Tohmé * and **Andrés Fioriti**

Departamento de Economía, Universidad Nacional del Sur, Instituto de Matemática de Bahía Blanca-CONICET, Bahía Blanca 8000, Argentina; andres.fioriti@uns.edu.ar
* Correspondence: ftohme@criba.edu.ar

Abstract: The study of the interactions among intentional agents, with rationality being the main source of intentional behavior, requires mathematical tools capable of capturing systemic effects. Here, we choose an alternative toolbox based on Category Theory. We examine potential *level-agnostic* formalisms, presenting three categories: $\mathcal{PR}$, $\mathcal{G}$, and an encompassing one, $\mathcal{I}$. The latter allows for representing dynamic rearrangements of the interactions among different agents. Systems represented in $\mathcal{I}$ capture the dynamic interactions among the interfaces of their sub-agents, changing the connections among them based on their internal states. We illustrate the expressive power of this formalism in four different instances, providing practitioners with a toolbox for representing cases of interest and facilitating their modular analysis.

Keywords: interactions; category theory; game theory; polynomial functors

MSC: 93A16; 18M99; 91A70

1. Introduction

In this paper, we analyze the *interactions* among *intentional* entities. The term "entity" is introduced to refer to individuals as well as other non-human agents, covering all kinds of things capable of exhibiting agency, ranging from social groups to robots. An entity exhibits agency if it can act independently and make its own decisions, rather than being passively determined by external forces.

These intentional entities can be seen as systems composed of other systems. While contemporary disciplines like computer science have embraced this view ([1]), in this contribution, we explore possible formalisms that may support the development of tools for an expanded view of the interactions among agents. We consider here two issues:

- How to deal with the decisions the sub-agents make within a single agent.
- How to scale up the solutions of agents to larger systems, aggregating them.

As an example of the first issue, we can consider a single robot solving two independent problems in parallel. It is natural to conceive the situation as if there were two agents exchanging information and resources to solve the two problems.

In the other direction, the problem of aggregation arises naturally in voting systems. Each voter has a preference, and a government has to be chosen that can be seen as a single agent representing society.

Each of these two issues is an instance of the same problem: one is the *bottom-up* approach and the other is its *top-down* version. Both reveal the need for a *level-agnostic* (or *continuous with respect to subagents*) representation of this "multi-level agency" phenomenon. This paper lays the groundwork for its formalization.

We start by noting that there exists a well-defined notion of an agent defined in terms of a given preference relation over the space of alternatives. Then, the agent is said to be *rational* if it chooses the most preferred alternatives among those that are feasible for it.

In applications, it is customary to reduce the analysis to a subspace of the space of alternatives, simplifying the problem of making a decision. But this comes at the price of

check for updates

Citation: Tohmé, F.; Fioriti, A. Level-Agnostic Representations of Interacting Agents. *Mathematics* **2024**, *12*, 2697. https://doi.org/10.3390/math12172697

Received: 30 July 2024
Revised: 21 August 2024
Accepted: 28 August 2024
Published: 29 August 2024

assuming the independence of the preferences over the subspace from the preferences over the rest of the larger space of alternatives.

In this initial version, we first present a method for ensuring the consistency of the solutions found across different subspaces. Then, we provide another approach to the coordination of independent contexts, involving games with shared players.

The final part of this paper presents a generalization that integrates both models, in which interactions are no longer fixed but can evolve according to the inputs and outputs. In this, as well as in the previous two models, we apply the mathematical framework of Category Theory.

Category Theory provides a high-level abstract representation of formal structures, focusing on their interrelations. It has largely contributed to the advancement of the mathematical sciences by being "*math to scaffold accounts from many disciplines*" [2].

Our contribution can be understood in this sense as a methodology to describe complex systems, using the same formalism for their components as well as for the larger systems they may, in turn, integrate. In this sense, it provides a useful *theoretical* characterization that helps to understand, in a modular form, the interactions among those systems, regardless of their position in the structures in which they participate.

2. Mathematical Preliminaries

As is well known, Category Theory has provided a framework without which most contemporary results in both algebraic geometry and topology would not have been found [3]. As repeatedly shown in actual mathematical practice, the language of Set Theory remains insufficient for capturing the subtleties prevalent in these fields [4]. One reason is that, unlike Set Theory, the categorical approach allows for the maximization of the "external" scope of its formal results *and* controlled "internal" sensitivity to particular differences in content within the representation of mathematical structures. Although Category Theory seems to provide a natural language for representing the decision-making problems outlined above, we should note that some disciplines, like economics, have been reluctant to adopt it. Some notable exceptions are [5–8]. In turn, [9] presented arguments for the adoption of the categorical language in economics.

In this paper, we draw heavily on the literature on Category Theory, although our results are clearly elementary. We now present the basic concepts that will be used in subsequent sections. For further details and clarification, see the excellent general texts on Category Theory by Goldblatt ([10]), Barr and Wells ([11]), Adámek et al. ([12]), Lawvere and Shanuel ([13]), Spivak ([14]), Fong and Spivak ([15]), Southwell ([16]), or Cheng ([17]).

A category **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 denoted by $f : a \to b$. 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 $f : a \to b$, then $f \circ \mathrm{Id}_a = f$ and $\mathrm{Id}_b \circ f = f$ (IDENTITY); and (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 include **SET** (the objects are sets, and the morphisms are functions between sets), **TOP** (the objects are topological spaces, and the morphisms are continuous functions), **POrd** (the objects are preorders, and the morphisms are order-preserving functions), and **Vec** (the objects are vector spaces, and the morphisms linear maps).

The terseness of categories facilitates diagrammatic reasoning. A diagram in which nodes represent objects and arrows represent morphisms allows for the establishment of properties of a category. Diagrams that *commute*, i.e., those in which 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 be 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 $\gamma_{jl} : b_j \to b_l$, then we $\gamma_{jl} \circ f_a^{b_j} = f_a^{b_l}$ (see Figure 1).

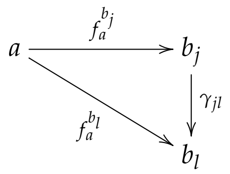


Figure 1. Commutative diagram.

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 2.

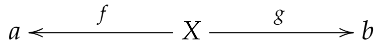


Figure 2. The limit of cones of this shape defines the product $a \times b$.

Then, the limit is the *product* $a \times b$, with arrows p_1 and p_2 representing the projections onto 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 include *direct sums* (in **SET**, disjoint unions) and, somewhat confusingly, *direct limits*, which, in a self-contained description, we use to define *global solutions*.

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** to 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 as $F(B')_{|B}$, 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).

Another categorical notion that is relevant in the next sections is that of a *symmetric monoidal category* (SMC). A category $\mathbf{C}$ is an SMC if the following conditions are fulfilled:

- There exists an object $I \in \mathrm{Ob}(\mathbf{C})$ called the *monoidal unit*.
- There exists a functor $\otimes : \mathbf{C} \times \mathbf{C} \to \mathbf{C}$, called the *monoidal product*, such that the following conditions hold:
  - $I \otimes c \cong c \cong c \otimes I$ for every $c \in \mathrm{Ob}(\mathbf{C})$;
  - $(c \otimes d) \otimes e \cong c \otimes (d \otimes e)$ for every $c, d, e \in \mathrm{Ob}(\mathbf{C})$;
  - $c \otimes d \cong d \otimes c$ for every $c, d \in \mathrm{Ob}(\mathbf{C})$.

Consider two monoidal categories, $\mathbf{C}$ and $\mathbf{D}$, with monoidal products, $\otimes_{\mathbf{C}}$ and $\otimes_{\mathbf{D}}$, and monoidal units, $I_{\mathbf{C}}$ and $I_{\mathbf{D}}$, respectively. A *lax monoidal functor* is a functor $F : \mathbf{C} \to \mathbf{D}$ together with a natural transformation

$$\phi_{XY} : F(X) \otimes_{\mathbf{D}} F(Y) \to F(X \otimes_{\mathbf{C}} Y)$$

and a morphism $\phi : I_{\mathbf{D}} \to F(I_{\mathbf{C}})$.

If $(\mathbf{C}, I, \otimes)$ is a symmetric monoidal category we can define an *operad* $\mathcal{O}_{\mathbf{C}}$ as follows:

- $\mathrm{Ob}(\mathcal{O}_{\mathbf{C}}) = \mathrm{Ob}(\mathbf{C})$.
- A morphism $(X_1, \ldots, X_n) \to Y$ in $\mathcal{O}_{\mathbf{C}}$ is defined as the morphism $X_1 \otimes \cdots \otimes X_n \to Y$ in $\mathbf{C}$.

Equipped with these notions, we can consider a category $\mathcal{WD}$ with the following characteristics:

- Each object is a *box* $X = \left(X^{\text{in}}, X^{\text{out}}\right)$, where $X^{\text{in}}, X^{\text{out}}$ are *typed* finite sets. Each element of $X^{\text{in}} \sqcup X^{\text{out}}$ is called a *port*.
- A morphism between two boxes X and Y is called a *wiring diagram* $\varphi : X \to Y$, such that $\varphi = \left(\varphi^{\text{in}}, \varphi^{\text{out}}\right)$ are defined as follows:

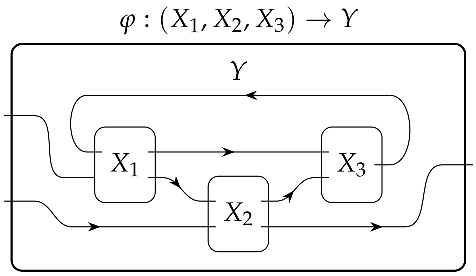
$$\varphi^{\text{in}} : X^{\text{in}} \longrightarrow Y^{\text{in}} \sqcup X^{\text{out}}$$
$$\varphi^{\text{out}} : Y^{\text{out}} \longrightarrow X^{\text{out}}$$

  where $\sqcup$ denotes the disjoint union of sets.
- Given two wiring diagrams, $\varphi : X \to Y$ and $\psi : Y \to Z$, their composition makes the following diagrams commutative:

$$\begin{array}{ccc} X^{\text{in}} & \xrightarrow{(\psi\circ\phi)^{\text{in}}} & Z^{\text{in}} \sqcup X^{\text{out}} \\ \Big\downarrow \phi^{\text{in}} & & \Big\uparrow Z^{\text{in}} \sqcup \phi^{\text{out}} \sqcup X^{\text{out}} \\ Y^{\text{in}} \sqcup X^{\text{out}} & \xrightarrow[\psi^{\text{in}} \sqcup X^{\text{out}}]{} & Z^{\text{in}} \sqcup Y^{\text{out}} \sqcup X^{\text{out}} \end{array} \qquad \begin{array}{ccc} Z^{\text{out}} & \xrightarrow{(\psi\circ\phi)^{\text{out}}} & X^{\text{out}} \\ & \searrow^{\psi^{\text{out}}} & \Big\uparrow \phi^{\text{out}} \\ & & Y^{\text{out}} \end{array}$$

$\mathcal{WD}$ has a symmetric monoidal structure, where $\otimes$ is identified with $\sqcup : \mathcal{WD} \times \mathcal{WD} \to \mathcal{WD}$, and the unit I is $\varnothing$ (the box with an empty set of ports). Then, an operad $\mathcal{O}_{\mathcal{WD}}$ can then be defined to enable the connection of different boxes into a single one.

For example, consider the morphism $\varphi : (X_1, X_2, X_3) \longrightarrow Y$ in $\mathcal{O}_{\mathcal{WD}}$. It can be depicted as follows:



Another categorical formalism to be applied in this paper is that of *polynomial functors*. Since it is quite central to our argument, we leave its presentation for Section 6, where we develop a unified level-agnostic model.

3. Sub-Agents: Local vs. Global

The usual specification of decision making under certainty by an agent starts with a space of possible **options**, $\mathcal{L}$, and a *utility* function, $U : \mathcal{L} \to \mathbb{R}$. Constraints on the set of options limit the available options to $\hat{L} \subseteq \mathcal{L}$. The agent seeks to find $\mathbf{x}^*$, maximizing U over $\hat{L}$.

The space of options, $\mathcal{L}$, is a (real) Hilbert space, i.e., a complete metric space with an inner product. To ensure the existence of a $\mathbf{x}^*$, it is assumed that $\hat{L}$ is a compact subset of $\mathcal{L}$ and that U is a continuous function.

In a category-theoretical treatment of the global optimization of U over $\hat{L}$, $\mathbf{x}^*$ is represented as a *direct limit*. This approach also allows us to analyze the problem of obtaining a global result from local ones.

Consider a family $\{L^k\}_{k=0}^{\kappa}$ of closed linear subspaces of $\mathcal{L}$ and, for any given k, the function

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

such that $\text{Proj}_k(x) = x^k \in L^k$, where x^k is the *projection* of x on L^k (the existence of a projection is ensured by a straightforward application of the Linear Projection Theorem, according to which $|x - x^k| = \min_{y \in L^k} |x - y|$, where $| \cdot |$ is the norm of $\mathcal{L}$ [18]).

Each L^k represents the set of options for a *local* problem. The projection of a global solution $\mathbf{x}^*$ onto L^k returns the point in L^k that is closest to $\mathbf{x}^*$. If the projection does not yield a local solution, another operator, $\Gamma_k : \hat{L} \to \hat{L}^k$, can be defined to provide choices closest to the projection if it does not belong to the subspace:

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

If the global solution is not given, it must be sought by combining local solutions. To formalize this, we introduce a category of local problems ([19]).

Definition 1. *Let $\mathcal{PR}$ be the category of local problems, characterized as follows:*

- *$Obj(\mathcal{PR})$ is the class of objects. Each object, $s^k = \langle \hat{L}^k, u^k, \hat{\mathbf{X}}^k \rangle$, 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 by $\hat{L}^k \subseteq \hat{L}^j$, $u^k = u^j|_{L^k}$, and $dim(L^k) \leq dim(L^j)$. Here, $dim(\cdot)$ denotes the dimension of a subspace of $\mathcal{L}$. 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 $\rho_{jl} : s^j \to s^l$, their composition, $\rho_{jl} \circ \rho_{kl} = \rho_{kl}$, exists 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 $u^j = u^l|_{L^j}$ imply that $u^k = u^l|_{L^k}$.*

We 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 by $A \subseteq B$.

Consider now a functor

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

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 (see Figure 3)

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

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

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

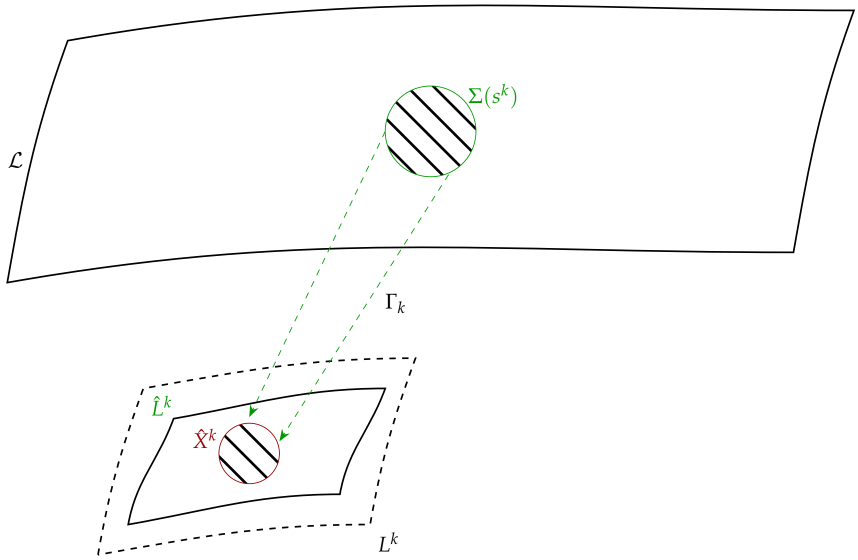


Figure 3. Representation of the relation between Γ_k and $\Sigma(s^k)$.

Given two problems, each identified with a **sub-agent** in charge of solving it, $s^k = \langle \hat{L}^k, u^k, \hat{\mathbf{X}}^k \rangle$ and $s^j = \langle \hat{L}^j, u^j, \hat{\mathbf{X}}^j \rangle$, we write $s^k \lhd s^j$ iff there exists a morphism ρ in $\mathcal{PR}$, $\rho : s^k \to s^j$. 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 it assigns $\Sigma(s^k)$ to $\Sigma(s^j)$. Given a section over s^j, r^j_k yields a section corresponding to its sub-problem s^k.

The proposition below shows that the functor Σ possesses an important property that is crucial for formalizing the possibility of patching up local problems and yielding a "larger" one.

Proposition 1. *Σ is a presheaf.*

Proof. $\Sigma : \mathcal{PR} \to \mathcal{P}(\mathcal{L})$ is a functor. We can analyze its behavior by means of r^j_k:

- For any $s^k \in \mathrm{Obj}(\mathcal{PR})$, 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, $r^j_k \circ r^l_j = r^l_k$.

This means that $\Sigma : \mathcal{PR} \to \mathcal{P}(\mathcal{L})$ is a *contravariant* functor or, in categorical terms, a *presheaf*. □

Consider now a family $\{s^k = \langle \hat{L}^k, u^k, \hat{\mathbf{X}}^k \rangle\}_{k \in K} \subseteq \mathrm{Obj}(\mathcal{PR})$. It is said to be a *cover* of an object $s^j = \langle \hat{L}^j, u^j, \hat{\mathbf{X}}^j \rangle$ of $\mathrm{Obj}(\mathcal{PR})$ 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 is 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 $\Sigma(s^l) = X^l$ (see Figure 4),

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

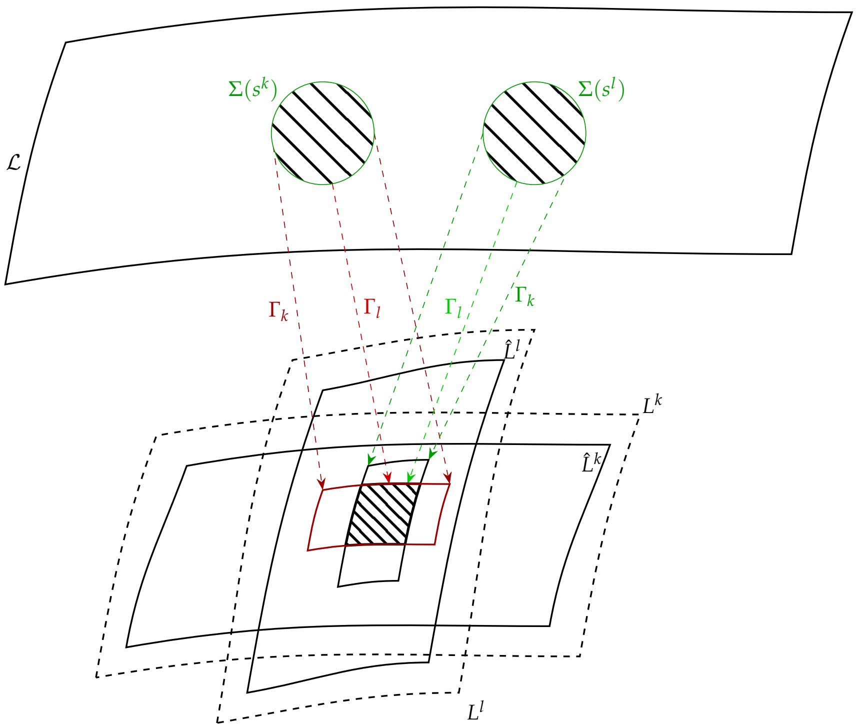


Figure 4. Compatibility of sections.

Given a cover $\{s^k\}_{k \in K}$ of a problem s^j with compatible sections, Σ satisfies the *sheaf property* 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, the sheaf property is satisfied if σ_j in fact "glues" together all the assignments σ_k in $\mathcal{P}(\mathcal{L})$ (see Figure 5).

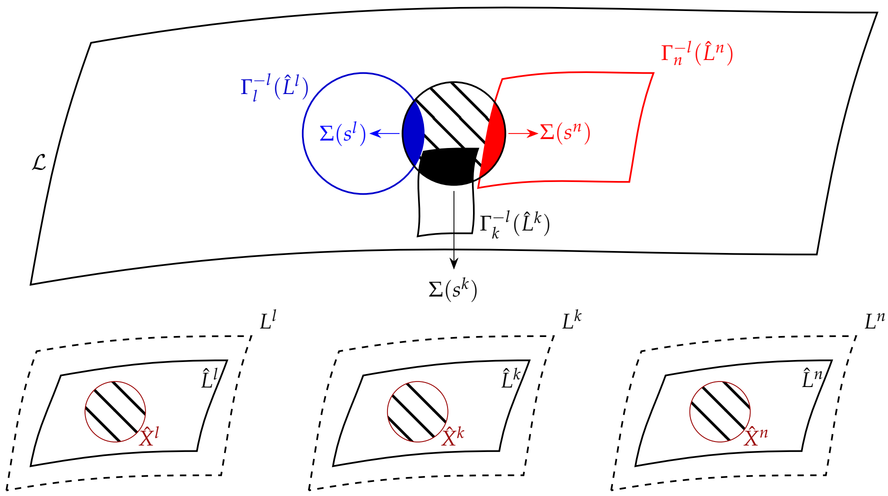


Figure 5. Sheaf property.

Summarizing the discussion up to this point, we can say that given a category of problems $\mathcal{PR}$ over a space $\mathcal{L}$, they can be seen as instances of a global problem if there exists a presheaf $\Sigma : \mathcal{PR} \to \mathcal{P}(\mathcal{L})$ that satisfies the sheaf property. Then, 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 $k \in K$.

That is, the sheaf property ensures that the behavior of the **sub-agents** is consistent with that of the single **agent**.

4. A Categorical Representation of Games

Let us now consider the coordination of games instead of the coordination of different local decision problems, that is, decision problems involving several agents, instead of a single one. Thus, the approach discussed in this section generalizes the sheaf-theoretical framework presented above. Alternative category-theoretical approaches to Game Theory were presented, for instance, in [6,20].

Let us consider a category $\mathcal{G}$ of *games*. Each object G in this category is defined as $G = \langle (I_G, S_G, \mathbf{O}_G, \rho_G), \pi_G \rangle$, where the components are specified as follows:

- The game form $(I_G, S_G, \mathbf{O}_G, \rho_G)$ is characterized by the following components:
  - I_G is the class of players.
  - $S_G = \prod_{i \in I_G} S_i^G$ is the *strategy set* of the game, where $S_i^G \subseteq S_i$ is the set of strategies that player i can deploy in game G for each $i \in I_G$. S_i is the set of all the strategies that player i can play in the games in which he/she participates.
  - $\mathbf{O}_G$ is the class of *outcomes* of the game, and $\rho_G : S_G \to \mathbf{O}_G$ is a one-to-one function that associates each profile of strategies in the game with one of its outcomes.
- $\pi_G = \prod_{i \in I} \pi_i^G$ is a *profile of payoff functions*, where $\pi_i^G : \mathbf{O}_G \to \mathbb{R}^+$ is the payoff function of player i in game G for each $i \in I_G$.

A game is defined in terms of the interactions of *players*. Each player can be seen as described in terms of the strategies he/she can play and the payoffs he/she can receive from the results of his/her actions (jointly with those of the other players).

We can define a category $\mathcal{G}$, where the objects are games. Given two games

$$G = \langle (I_G, S_G, \mathbf{O}_G, \rho_G), \pi_G \rangle \quad \text{and} \quad G' = \langle (I_{G'}, S_{G'}, \mathbf{O}_{G'}, \rho_{G'}), \pi_{G'} \rangle,$$

a morphism of games

$$G \to G'$$

is defined by the following conditions:

- $I_G \subseteq I_{G'}$.
- $S_i^G \subseteq S_i^{G'}$ for each $i \in I_G$.
- $\mathbf{O}_G \subseteq \mathbf{O}_{G'}$.

Thus, if a morphism $G \to G'$ exists, G can be considered a *subgame form* of G'.

To complete the characterization of $\mathcal{G}$, note that it is immediate that we can define *pushouts* and an *initial object* in this category as follows:

- **Pushouts**: Consider three objects G, G', and G'', and morphisms $G \xrightarrow{f} G'$ and $G \xrightarrow{g} G''$. Then, take the coproduct of G' and G'', denoted as $G' + G''$, obtained as the direct sums of the strategy sets and outcomes of both games. By identifying the subgame forms of G' and G'' corresponding to G, we obtain the *pushout* of
$$G' \xleftarrow{f} G \xrightarrow{g} G''$$
- **Initial object**: Consider the *empty game* $G^{\emptyset}$, where $I_{G^{\emptyset}} = \emptyset$ and, consequently, $S_{G^{\emptyset}} = \emptyset$ and $\mathbf{O}_{G^{\emptyset}} = \emptyset$ (thus, $\pi_{G^{\emptyset}}$ must be the empty function). It is immediate to see that $G^{\emptyset} \to G$ for every G in $\mathcal{G}$.

Then, we have that

Proposition 2. *$\mathcal{G}$ is a category with colimits.*

Since $\mathcal{G}$ is a category with colimits, we can define the *cospans* in it. Consider again three objects G, G', and G'', and two morphisms $G \xrightarrow{f} G'' \xleftarrow{g} G'$. This is called a cospan from G to G'. The interpretation of such a cospan is that G and G' are subgame forms of the same game (G'').

We can consider each game G in $\mathcal{G}$ as a *box*, $G = (\text{in}^G, \text{out}^G)$, where in^G and out^G are, respectively, the *input* and *output* ports. in^G has type $\mathbf{O}_G$, i.e., the input is an outcome of G. In turn, the out^G port has type S_G, with each output being a profile in G.

Note that each player i can be represented as a game $(\text{in}^i, \text{out}^i)$, where in^i has type $\cup_{G:i\in I_G}\mathbf{O}_G$ and out^i has type S_i.

Up to this point, our definition of morphisms in $\mathcal{G}$ does not involve the payoffs. They can be incorporated by redefining the games as *modal boxes*, in which an additional component is the set of *internal states* of the game. More precisely, given any G and the class of its internal states, Σ_G, we can identify G as a triple $\langle \text{in}^G, \text{out}^G, \Sigma_G \rangle$, associated with two correspondences:

- **Payoff**: $\phi^1_G : \bar{\text{in}}^G \times \Sigma_G \to \mathbb{R}^{+^{\mathbf{O}_G}}$ such that for the vector $o \in \bar{\text{in}}^G$ (the vector of all possible inputs of G, each entry being an outcome of the game) and state σ, $\phi^1_G(o, \sigma) = (\pi^i_G(o))_{o\in\mathbf{O}_G}$. That is, it yields the vector of payoffs corresponding to all the outcomes of G.
- **Choice**: $\phi^2_G : \Sigma_G \to \bar{\text{out}}^G$ such that for any state σ, $\phi^2_G(\sigma) = s \in \bar{\text{out}}^G$ (the class of all possible strategy profiles in S_G) is a profile of strategies that may be chosen at that state.

Particularly relevant for our analysis is the definition of the internal states of each player i, denoted as Σ_i. Consider a game G such that $i \in I_G$, and a sequence of morphisms in $\mathcal{G}$

$$G^0_i \to G^1_i \to \quad \ldots \quad \to G^{n-1}_i \to G^n_i$$

where G^0_i is a game in which i is the only player and $G = G^n_i$. We denote the state of player i when playing G as a sequence $\sigma^i_G = \langle \sigma^i_0, \ldots, \sigma^i_{n-1} \rangle$, where $\sigma^i_k \in \Sigma_{G^k_i}$, for $k = 0 \ldots, n-1$. Then, a distinguished object $\sigma^i_* \in \Sigma_i$ is defined such that σ^i_G is one of its initial segments. Thus, σ^i_* has a *forest* structure.

Therefore, for each game G, σ^i_* can be instantiated, yielding the corresponding state, and, consequently, the payoffs and the choices of player i in the game. The state σ_G of the entire game is obtained as the profile of the states of its players.

A simple example is $\sigma^i_{G^n}$, yielding, as the payoff for i, the product of the payoffs he/she receives in the subgames of G^n. This case is elaborated on further in Example 1 below.

We can define the category of cospans in $\mathcal{G}$ as $\text{cospan}_{\mathcal{G}}$, which has a symmetric monoidal structure. Its objects are the same as those of $\mathcal{G}$, and a morphism $G \xrightarrow{h} G'$ is a cospan from G to G', indicating that there exists a game of which G and G' are subgame forms. Thus, morphisms in $\text{cospan}_{\mathcal{G}}$ are actually isomorphisms.

Given two morphisms in $\text{cospan}_{\mathcal{G}}$, $G \xrightarrow{f} G'$ and $G' \xrightarrow{g} G''$, there exists a morphism $G \xrightarrow{g\circ f} G''$, which is obtained as a composition of the corresponding cospans.

The monoidal structure of $\text{cospan}_{\mathcal{G}}$ is given by the following:

- The unit is $G^{\emptyset}$, the initial object in $\mathcal{G}$.
- The monoidal product of G and G' is the *coproduct* $G + G'$.

We now present a diagram language for open games. We start by considering the symmetric monoidal category $\mathbf{W}_{\mathcal{G}}$. By definition, we have that

$$\mathbf{W}_{\mathcal{G}} \; = \; \text{cospan}_{\mathcal{G}}.$$

Each object, i.e., a game G, is seen as a $\langle \text{in}^G, \text{out}^G, \Sigma_G \rangle$-labeled *interface*, satisfying ϕ^1_G and ϕ^2_G. On the other hand, morphisms $G \to C \leftarrow G'$ are called $\langle \text{in}, \text{out}, \Sigma \rangle$-labeled *wiring diagrams*. The interpretation is that C is the overarching game that connects the subgames (not just the game forms) G and G'.

We write $\psi : G_1, G_2, \ldots, G_n \rightarrow \bar{G}$ to denote the wiring diagram $\phi : G_1 + G_2 + \ldots + G_n \rightarrow \bar{G}$. This can, in turn, be seen as

$$G_1 + G_2 + \ldots + G_n \xrightarrow{f} C \xleftarrow{\bar{f}} \bar{G}$$

which indicates that, since f and $\bar{f}$ isomorphisms,

Proposition 3. *$\bar{G}$ is the minimal game that includes the direct sum of $G_1, \ldots, G_n$ as a subgame.*

In $\mathbf{W}_{\mathcal{G}}$, the monoidal product of G and G', $G \otimes G'$, is defined as follows (where $\cup$ and $\sqcup$ represent the set union and the disjoint union of *sets*, respectively):

- $I_{G \otimes G'} = I_G \cup I_{G'}$.
- $\mathbf{O}_{G \otimes G'} = \mathbf{O}_G \sqcup \mathbf{O}_{G'}$.
- For each $i \in I_{G \otimes G'}$,

$$S_i^{G \otimes G'} = \begin{cases} S_i^G & \text{if } i \in I_G \backslash I_{G'} \\ S_i^{G'} & \text{if } i \in I_{G'} \backslash I_G \\ S_i^G \times S_i^{G'} & \text{if } i \in I_G \cap I_{G'} \end{cases}$$

- $\pi_i^{G \otimes G'}(s) = \pi_i^G(s^G) + \pi_i^{G'}\left(s^{G'}\right)$, where $s^G, s^{G'}$ are the projections of $s \in \prod_{j \in I_{G \otimes G'}} S_j^{G \otimes G'}$.

5. Hypergraph Categories and Equilibria

We define a *hypergraph category* $\langle \mathcal{G}, \mathbf{Eq} \rangle$ with $\mathbf{Eq} : \mathbf{W}_{\mathcal{G}} \rightarrow \prod_i S_i$, such that for every object G in $\mathbf{W}_{\mathcal{G}}$, $\mathbf{Eq}(G)$ is a class of vectors in $\prod_{i \in I} S_i^G$, the strategy set of game G. We assume that $\mathbf{Eq}(G)$ represents a class of *equilibria* of G for some notion of equilibrium (for instance, dominant strategy equilibrium, admissible strategies, or Nash equilibrium).

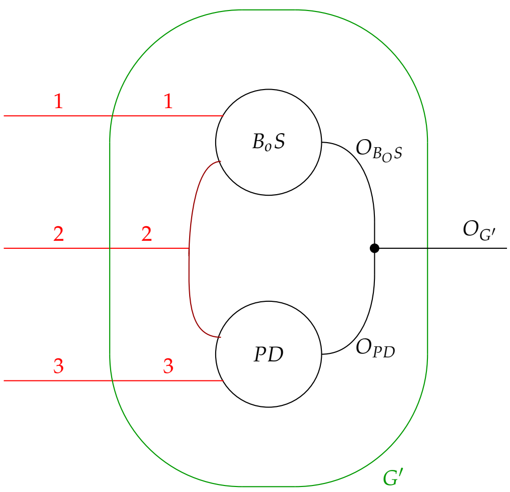
Example 1. *Consider two games, G, between players 1 and 2 (a battle-of-the-sexes game, where $S_1 = S_2 = \{Bx, Bll\}$):*

	Bx	Bll
Bx	2, 1	0, 0
Bll	0, 0	1, 2

and G' between players 2 and 3 (a prisoner's dilemma, where $S_2 = S_3 = \{C, D\}$):

	C	D
C	2, 2	0, 3
D	3, 0	1, 1

The corresponding wiring diagram is as follows:



In red, we have highlighted $\boldsymbol{Eq}(G) = \{(Bx,Bx), (Bll, Bll)\}$ *and* $\boldsymbol{Eq}(G') = \{(D,D)\}$*, where* $\boldsymbol{Eq}$ *corresponds to the Nash equilibrium. Note that here, player 2 participates in two games.*

Let us now represent $G \otimes G'$*. We start by building its corresponding game form. We obtain two tables, where the first one corresponds to player 3 choosing C:*

	$Bx \bowtie C$	$Bx \bowtie D$	$Bll \bowtie C$	$Bll \bowtie D$
Bx	$o_{1,1}$	$o_{1,2}$	$o_{1,3}$	$o_{1,4}$
Bll	$o_{2,1}$	$o_{2,2}$	$o_{2,3}$	$o_{2,4}$

and the second one corresponds to player 3 choosing D:

	$Bx \bowtie C$	$Bx \bowtie D$	$Bll \bowtie C$	$Bll \bowtie D$
Bx	$o'_{1,1}$	$o'_{1,2}$	$o'_{1,3}$	$o'_{1,4}$
Bll	$o'_{2,1}$	$o'_{2,2}$	$o'_{2,3}$	$o'_{2,4}$

For instance, o_{11} *indicates that players 1 and 2 select Box, and players 2 and 3 cooperate. On the other hand,* $o'_{1,1}$ *indicates that, again, players 1 and 2 select Box, but while player 2 keeps cooperating, player 3 defects. The other entries can be interpreted likewise.*

Suppose that the internal states of the players σ^1_*, σ^2_**, and* σ^3_* *are such that, when instantiated on* $G \otimes G'$*, they yield the payoffs and choices described below.*

If player 3 chooses C:

	$Bx \bowtie C$	$Bx \bowtie D$	$Bll \bowtie C$	$Bll \bowtie D$
Bx	$2,1 \times 2,2$	$2,1 \times 3,0$	$0,0 \times 2,2$	$0,0 \times 3,0$
Bll	$0,0 \times 2,2$	$0,0 \times 3,0$	$1,2 \times 2,2$	$1,2 \times 3,0$

If player 3 chooses D:

	$Bx \bowtie C$	$Bx \bowtie D$	$Bll \bowtie C$	$Bll \bowtie D$
Bx	$2,1 \times 0,3$	$2,1 \times 1,1$	$0,0 \times 0,3$	$0,0 \times 1,1$
Bll	$0,0 \times 0,3$	$0,0 \times 1,1$	$1,2 \times 0,3$	$1,2 \times 1,1$

In other words, players 1 and 3 keep the payoffs they receive in the subgames, while player 2 takes the product of the payoffs in G and G'*. In red, we have highlighted the equilibria of* $G \otimes G'$ *under this specification.*

Let us define an operation $\hat{\cup}$ such that, given two equilibria $s \in \mathbf{Eq}(G)$ and $s' \in \mathbf{Eq}(G')$, it yields a new profile $s \bowtie s' \in \mathbf{Eq}(G)\hat{\cup}\mathbf{Eq}(G')$ verifying that, for each player $i \in I_G \cap I_{G'}$, a new strategy is obtained by combining s_i and s'_i, while in all other cases, the individual strategies remain the same as in G and G'. Furthermore, $\pi_i^{G\hat{\cup}G'}(s \bowtie s') = \pi_i^G(s) \times \pi_i^{G'}(s')$ for $i \in I_G \cap I_{G'}$. An alternative yielding Proposition 4 is obtained if, instead, we take $\pi_i^{G\hat{\cup}G'}(s \bowtie s') = \pi_i^G(s) + \pi_i^{G'}(s')$ for $i \in I_G \cap I_{G'}$.

In our example, since $\mathbf{Eq}(G \otimes G') = \{(\mathrm{Bx}, \mathrm{Bx} \bowtie \mathrm{D}, \mathrm{D}), (\mathrm{Bll}, \mathrm{Bll} \bowtie \mathrm{D}, \mathrm{D})\}$, we have that

$$\mathbf{Eq}(G)\hat{\cup}\mathbf{Eq}(G') = \mathbf{Eq}(G \otimes G').$$

This example illustrates the following claim:

Proposition 4. *For any pair of games G and* G'*,* $\boldsymbol{Eq}(G)\hat{\cup}\boldsymbol{Eq}(G') = \boldsymbol{Eq}(G \otimes G')$*.*

Proof. Trivial. If $I_G \cap I_{G'} = \emptyset$, then $G \otimes G' = G \cup G'$ with $G \cap G' = \emptyset$. Thus, each equilibrium of $G \otimes G'$ is simply the disjoint combination of equilibria in G and G'.

If, on the other hand, $I_G \cap I_{G'} \neq \emptyset$, given $i \in I_G \cap I_{G'}$, his/her strategy set in $G \otimes G'$ is $S_i^G \times S_i^{G'}$, where S_i^G and $S_i^{G'}$ are his/her strategy sets in G and G', respectively. Now, suppose that s_i^G and $s_i^{G'}$ are equilibrium strategies of i in the individual games, but that $(s_i^G, s_i^{G'})$ does not belong to an equilibrium in $G \otimes G'$.

Then, there exists an alternative combined strategy $(\hat{s}_i^G, \hat{s}_i^{G'})$ such that, under the new profile, π_i yields a higher payoff. But since this equilibrium can be decomposed into two profiles—one in G and the other in G'—the payoff of i is the product of the payoffs over those two profiles. Therefore, either $\hat{s}_i^G$ yields a higher payoff than s_i^G or $\hat{s}_i^{G'}$ yields a higher payoff than $s_i^{G'}$ (note that all payoffs are positive real numbers).

Thus, either s_i^G or $s_i^{G'}$ is not an equilibrium in the corresponding game. This is a contradiction. □

Proposition 4 indicates that there exists a trivial *natural isomorphism*

$$\mathbf{Eq}(G)\hat{\cup}\mathbf{Eq}(G^{'}) \ \rightarrow \ \mathbf{Eq}(G \otimes G^{'}).$$

Furthermore, taking the unit in $\prod_i S_i$ to be the empty set, we also have that $\emptyset = \mathbf{Eq}(G^{\emptyset})$, where $G^{\emptyset}$ is the initial object in $\mathcal{G}$ and thus in $\mathbf{W}_{\mathcal{G}}$.

Recalling the definition of a *lax monoidal functor* as a functor $F : \mathbf{C} \rightarrow \mathbf{D}$ together with a natural transformation

$$F(X) \otimes_{\mathbf{D}} F(Y) \rightarrow F(X \otimes_{\mathbf{C}} Y)$$

we have, trivially, that

Proposition 5. ***Eq*** *is a lax monoidal functor.*

Thus, the corresponding algebra associates the composition of games with the equilibria of the components.

Proposition 5 depends critically on the possibility of defining $\hat{\cup}$ in terms of a function $\mathbf{f}$, defined as follows. Given a player $i \in I_G \cap I_{G'}$, a combined strategy $s_i \bowtie s_i'$ is such that for $s = (s_i, s_{-i}) \in \text{Eq}(G)$ and $s' = (s_i', s_{-i}') \in \text{Eq}(G')$, satisfying $\pi_i(s \bowtie s') = \mathbf{f}(\pi_i^G(s), \pi_i^{G'}(s'))$ and with $s \bowtie s' \in \text{Eq}(G \otimes G')$. As we saw above, if $\mathbf{f}$ is the arithmetic product or sum, Eq will indeed be a lax monoidal functor.

However, this restricts the compositionality of games to just trivial cases. We are interested in more general and non-obvious cases. To address this, consider an alternative characterization of the hypergraph category $\langle \mathcal{G}, \mathbf{Eq} \rangle$:

$$\mathbf{Eq} : \mathbf{W}_{\mathcal{G}} \rightarrow \prod_i S_i \times \cup_{G \in \mathbf{Obj}(\mathcal{G})} \Sigma_G$$

Furthermore, we need another definition of $\hat{\cup}$:

$$\otimes : (\prod_i S_i \times \cup_{G \in \mathbf{Obj}(\mathcal{G})} \Sigma_G) \ \times \ (\prod_i S_i \times \cup_{G \in \mathbf{Obj}(\mathcal{G})} \Sigma_G) \ \rightarrow \ \prod_i S_i \times \bigcup_{G \in \mathbf{Obj}(\mathcal{G})} \Sigma_G$$

such that given two games G and G' with $s \in \prod_{i \in I_G} S_i$ and σ_G, and $s' \in \prod_{i \in I_{G'}} S_i$ and $\sigma_{G'}$, we have

$$(s, \sigma_G)\hat{\cup}(s', \sigma_{G'}) = (\bar{s}, \sigma_{G+G'}) \in \prod_{i \in I_{G+G'}} S_i \times \Sigma_{G+G'}$$

where $\bar{s} \in S_{G+G'}$ is a Nash equilibrium if and only if s and s' are Nash equilibria of G and G', respectively.

$\hat{\cup}$ is well defined. To see this, recall that, by definition, $G + G'$ is obtained in terms of the game forms of G and G' (the strategy sets and the outcomes), allowing different possible internal states and thus payoffs. The view of games as boxes, as presented in Section 4, indicates that there exist sequences of internal states of games, parallel to sequences of morphisms between games, which allow us to define $\sigma_{G+G'}$, and thus payoffs that make $\bar{s}$ a Nash equilibrium if s and s' are also equilibria.

We can see that $\prod_i S_i \times \bigcup_{G \in \mathbf{Obj}(\mathcal{G})} \Sigma_G$, with $\hat{\cup}$ defined as above, can be viewed as a monoidal category, with morphisms defined in terms of those of $\mathcal{G}$, with $(\emptyset, \emptyset)$ as its initial object. This allows us to define $\mathbf{Eq}$ in such a way that, by definition,

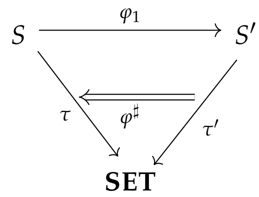
Proposition 6. ***Eq*** *is a lax monoidal functor satisfying* $\boldsymbol{Eq}(G + G') = \boldsymbol{Eq}(G) \hat{\cup} \boldsymbol{Eq}(G')$.

6. A More General Model

$\langle \mathcal{G}, \text{Eq} \rangle$, in either of the two versions of Eq, seems too rigid to capture the dynamics of economic interactions. A more flexible structure is needed.

Let us start with the following category:

- Objects: Pairs (S, τ), where $S \in \text{Ob}(\textbf{Set})$ and $\tau : I \rightarrow \text{Set}$.
- Morphisms: $(S, \tau) \xrightarrow{\varphi} (S', \tau')$ are pairs $\left(\varphi_1, \varphi^\sharp\right)$ such that



That is, $\varphi_1 : S \rightarrow S'$, while $\varphi^\sharp : \tau\prime(s') \mapsto \tau(s)$ for $s' \in S'$ and $s \in S$.

These "two-sided" morphisms generalize the "one-sided" ones we have considered up to this point. The $\varphi^\sharp$ component facilitates the composition of objects that are somehow incompatible. To show precisely what this means, we present a much more evocative and functional presentation of this category, called **Poly** [21]:

- Each object $p \in \text{Ob}(\textbf{Poly})$ is written as

$$p = \sum_{i \in I} y^{p[i]}$$

  where each term $y^{p[i]}$ is a functor with domain $p[i]$ into **Set**. Each i can be considered a *problem*, while $p[i]$ is a set of its *solutions*.
- Given $p = \sum_{i \in I} y^{p[i]}$ and $q = \sum_{j \in J} y^{q[j]}$, a morphism $\phi : p \rightarrow q$ is $\phi = (\phi_{\rightarrow}, \phi^{\leftarrow})$ is defined by the following:
  - $\phi^{\rightarrow} : I \rightarrow J$;
  - $\phi^{\leftarrow} : q[\phi^{\rightarrow}(i)] \mapsto p[i]$.

We can see how this specification captures the previously given definition of **Poly**. Each $y^{p[i]}$ is identified with $\tau : S \rightarrow \textbf{Set}$, where $S \equiv p[i]$. Then, p represents

$$\sqcup_i \{\tau_i : p[i] \rightarrow \textbf{Set}\}$$

Furthermore, $\phi^{\rightarrow}$, which maps problems indexed by I to problems indexed by J, represents φ_1, while $\phi^{\leftarrow}$, which maps the solutions in $q[\phi^{\rightarrow}(i)]$ back to the solutions in $p[i]$, corresponds to $\varphi^\sharp$.

Interestingly, the usefulness of considering this specification of **Poly** is that we can use it to represent a relation between a class of problems, indexed by I, and their solutions $\{p[i]\}_{i \in I}$. Thus, it disregards the codomain of the τ_is, to just focus on the S_is and their indices.

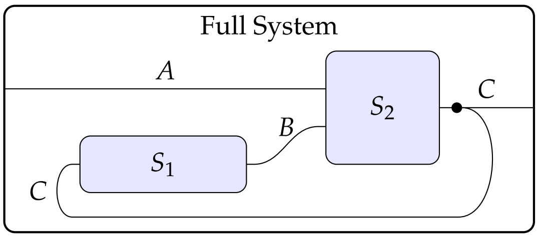
We can consider any $p \in Ob(\textbf{Poly})$ as an *interface* between inputs and outputs, where the inputs are problems and the outputs are their solutions. There are different ways to create new interfaces from other interfaces. We focus on the following construction:

- $[p, q] = \sum_{\phi : p \rightarrow q} y^{\sum_{i \in I} q[\phi^{\rightarrow}(i)]}$, an *internal hom* in **Poly**. It can be seen as a process that takes as inputs (*problems*) the morphisms from p to q and as outputs (*solutions*) all the possible solutions to the images of p in q.
- Given $[p, q]$, a $[p, q] - \textbf{Coalg}$ is a category in which each object is a triple $\langle s, \rho, \mu \rangle$:
  - $s \in S$, where S is a space of *states*, capturing the dynamics of the interface;
  - $\rho : s \mapsto (\phi, i, q[\phi^{\rightarrow}(i)])$, where $\phi : p \rightarrow q$ is a morphism. That is, it assigns to the current state one of the solutions in $[p, q]$;

– μ updates the state in response to that pattern, i.e., $\mu(\phi, i, q[\phi^{\rightarrow}(i)]) = s' \in S$.

Example 2. *Consider a system in which two subsystems, S_1 and S_2, acting in parallel, are described by $p \simeq By^C \otimes Cy^{AB}$, yielding the full system, represented by $q \simeq Cy^A$.*

For any state $s \in S$ of a $[p, q]$-coalgebra (S, ρ, μ), we have that $\rho(s)$ gives a morphism $p \rightarrow q$ in Poly, which can be depicted as follows:



Given $(a, b, c) \in A \times B \times C, \mu(a, b, c)$ is the updated state in S, which, in turn, may yield a new connection between S_1 and S_2.

This example shows that $[\cdot, \cdot]$-coalgebras provide flexible and dynamic connections among subsystems. This inspires the following extension of **Poly**, resulting in the category $\mathbb{O}$**rg**, which can be categorized as follows [22]:

- $Ob(\mathbb{O}\mathbf{rg}) = Ob(\mathbf{Poly})$;
- $Morph(\mathbb{O}\mathbf{rg}) = [p, q] - \mathbf{Coalg}$.

This means that two interfaces (connecting problems with their solutions) p and q are related by dynamic procedures of reconnection between them.

Our generalized model, covering *both* $\mathcal{PR}$ and $\langle \mathcal{G}, \mathrm{Eq} \rangle$, is a category $\mathcal{I}$ based on $\mathbb{O}$**rg** with the following features:

- Each object $a = \left\langle a^{\mathrm{in}}, a^{\mathrm{out}} \right\rangle \in \mathrm{Ob}(\mathcal{I})$ is identified with

$$p_a \simeq a^{\mathrm{out}} y^{a^{\mathrm{in}}} \in \mathrm{Ob}(\mathbb{O}\mathbf{rg})$$

- For objects $a_1, \ldots, a_n, b$, there corresponds a $[p_{a_1} \otimes \ldots \otimes p_{a_n}, p_b] - \mathbf{Coalg}$ of states $S_{a_1,\ldots,a_n,b}$. The operation $p_a \otimes p_b$, where $p_a = \sum_{i \in I} y^{p_a[i]}$ and $p_b = \sum_{j \in J} y^{p_b[j]}$, is such that to each problem $(i, j) \in I \times J$, there corresponds the pair of solutions to i and j, $(p_a[i], p_b[j])$.
- Each object a has an *identity* morphism.
- Pairs of morphisms compose.

The last two requirements indicate, roughly, that morphisms inherit the identity and compositionality properties of $\mathbb{O}$**rg**.

Then, we can prove that $\mathcal{I}$ is a category of level-agnostic dynamic arrangements.

Theorem 1. *There exist two categories $\overline{\mathcal{PR}}$ and $\overline{\mathcal{G}}$, isomorphic to $\mathcal{PR}$ and $\mathcal{G}$, respectively, such that $Ob(\overline{\mathcal{G}}), Ob(\overline{\mathcal{PR}}) \subseteq Ob(\mathcal{I})$, while $Morph(\overline{\mathcal{PR}}), Morph(\overline{\mathcal{G}}) \subseteq Morph(\mathcal{I})$, consist of trivial internal hom coalgebras with single states.*

Proof. Each problem in $\mathcal{PR}$ can be interpreted as an interface between the problem itself and its optimal solutions. The same applies to any interactive decision-making setting in $\mathcal{G}$.

More precisely, a local problem $s^k \in Ob(\mathcal{PR})$ and a game $G \in Ob(\langle \mathcal{G}, Eq \rangle)$ can be represented by polynomial functors p_{s^k} and p_G, respectively. In the former case, p_{s^k} is an interface between the specification of the local problem $(\hat{L}^k, u^k)$ and its solutions $\hat{\mathbf{X}}^k$. In the case of a game, p_G is an interface between the game G and its equilibria $Eq(G)$.

Each state in the morphism between two interfaces p_{s^k} and p_{s^j} represents a particular $r^k_j : \Sigma(s^k) \rightarrow \Sigma(s^j)$ that maps a section of solutions over s^k to a corresponding section over s^j, yielding a *sheaf*.

Analogously, each state in the morphism between two interfaces p_G and $p_{G'}$ represents a particular *wiring*, connecting the games G and G', such that the equilibrium obtains by tensoring those of the two games.

Since in $\mathcal{PR}$ and $\mathcal{G}$, morphisms cannot be rearranged they can be seen as hom coalgebras with a single state. □

Thus, $\mathcal{I}$ incorporates all the representational advantages of $\mathcal{PR}$ and $\mathcal{G}$, adding the possibility of capturing the dynamics of actual systems.

The following two examples exhibit the advantages of formalizing problems in $\mathcal{I}$.

Example 3 ([23]). *Consider a principal-agent problem defined by two functions*

$$\Phi_{\rightarrow} : X \times Y \times \mathbb{R} \rightarrow \mathbb{R} \text{ and } \Pi : X \times Y \times \mathbb{R} \rightarrow \mathbb{R}$$

where X is the compact set of types of the agent; Y is the compact set of possible decisions made by the agent; $\Phi_{\rightarrow}$ is continuous and strictly decreasing in the third argument; $\Phi_{\rightarrow}$ is full range in the third argument, i.e., $\Phi_{\rightarrow}(x, y, \cdot)[\mathbb{R}] = \mathbb{R}$ for every $(x, y) \in X \times Y$; Π is continuous and increasing in the third argument; and Π is full range in the third argument, i.e., $\Pi(x, y, \cdot)[\mathbb{R}] = \mathbb{R}$ for every $(x, y) \in X \times Y$.

Given a type x of the agent, its decision y and v, the money transfer to the principal, $\Phi_{\rightarrow}(x, y, v) = u_A$ is the utility of the agent, while $\Pi(x, y, v) = u_P$ is the utility of the principal.

An inverse generating function is

$$\Phi^{\leftarrow} : Y \times X \times \mathbb{R} \rightarrow \mathbb{R}$$

such that, given $u_A = \Phi_{\rightarrow}(x, y, \Phi^{\leftarrow}(y, x, u_A))$, there exists $v = \Phi^{\leftarrow}(y, x, \Phi_{\rightarrow}(x, y, v))$.

Given $\lambda \in \mathbb{M}$, the class of Borel measures over $X \times Y$ and $\underline{u}$, the reservation utility of the agent, the principal's problem amounts to choosing $\langle \lambda, \bar{u}_A, \bar{v} \rangle$ so as to maximize

$$\int_X \int_Y \Pi(x, y, \Phi^{\leftarrow}(y, x, \bar{u}_A)) d\lambda(x, y)$$

subject to $\bar{v} = \Phi^{\leftarrow}(y, x, \bar{u}_A)$ and $\bar{u}_A \geq \underline{u}$.

This setting can be naturally represented by defining two objects in $\mathcal{I}$: A and P (the agent and the principal, respectively). The corresponding polynomial functors are as follows:

- *p_P takes as input $\underline{u}$ and returns the optimal values λ^*, u^*_A, and $\bar{v}^*$. That is, $p_P = \sum_{\underline{u} \in \mathbb{R}} y^{p_P[\underline{u}]}$, where $p_P[\underline{u}] = \langle \lambda^*, u^*_A, \bar{v}^* \rangle$.*
- *p_A takes as input $\bar{v}$ and returns its decision y and the principal's utility u_P. That is, $p_A = \sum_{\bar{v} \in \mathbb{R}} y^{p_A[\bar{v}]}$, where $p_A[\bar{v}] = \langle y, u_P \rangle$.*

Then, the entire problem can be understood in terms of the identity morphism of $p_A \otimes p_P$, yielding the adjunction between $\Phi^{\rightarrow}$ and $\Phi^{\leftarrow}$.

A promising area of research in which $\mathcal{I}$ could be relevant for the design of mechanisms is described below.

Example 4 ([6,24]). *Mechanisms (and also institutions) can be considered game forms. That is, each mechanism M can be represented as $M = (I_M, S_M, \mathbf{O}_M, \rho_M)$ (see Section 4).*

Each $i \in I_M$ can be given different incentives according to the environment $\mathbf{e} \in E$ in which it interacts with others. Each $\mathbf{e} \in E$ has an associated profile of payoff functions that correspond to the outcomes in M, $\pi^{\mathbf{e}}_M$.

The task of a mechanism designer D is to assign a mechanism $M \in \mathbb{M}$ to a given environment in order to ensure a target $\mathbf{o}^*$. Thus, in $\mathcal{I}$, D has an associated $p_D = \sum_{\mathbf{e} \in E} y^{p_D[\mathbf{e}]}$, where

$$p_D[\mathbf{e}] = \{\langle M, \pi^{\mathbf{e}}_M \rangle : M \in \mathbb{M} \text{ such that } s^*_M \in \mathbf{Eq}(\langle M, \pi^{\mathbf{e}}_M \rangle) \text{ and } \rho(s^*_M) = \mathbf{o}^* \in \mathbf{O}_M\}$$

Each game form $M \in \mathbb{M}$ constitutes a local problem*. The polynomial corresponding to these problems is $p_{\mathbb{M}}$. In turn, given the choice of* nature *(represented by a constant polynomial $p_E = E$), the whole problem can be described by a $[p_D \times p_E, p_{\mathbb{M}}]$-coalgebra, where*

$$[p_D \times p_E, p_{\mathbb{M}}] = \sum_{\phi : p_D \times p_E \rightarrow p_{\mathbb{M}}} y^{\sum_{\mathbf{e} \in E} p_{\mathbb{M}}[\phi^{\rightarrow}(\mathbf{e})]}$$

and $p_{\mathbb{M}}[\phi^{\rightarrow}(\mathbf{e})] = \langle M, \pi^{\mathbf{e}}_M \rangle$.

Another important example in which a level-agnostic description can contribute to rationalizing the behaviors of complex systems is described below. Note that a very similar formalization can be used to represent the behavior of *foundation models*, like those underlying *generative AI* [25].

Example 5 ([26])**.** *Cyber-physical systems combine cyber capabilities with physical capabilities to solve problems that neither part could solve alone. Examples of these range from self-driving cars to robots. They act physically on the world as determined by discrete algorithms that adjust their actuators based on sensor readings of the physical state.*

A CPS is $X = \langle I_X, O_X, S_X, \gamma_X, \rho_X \rangle$, where I_X is a set of inputs, each consisting of a sensor; O_X is a set of outputs, each corresponding to an actuator; S_X is a set of internal states, summarizing the entire information processed by the CPS; and $\gamma_X : I_X \times S_X \rightarrow S_X$ and $\rho_X : S_X \rightarrow O_X$ are functions, where γ_X represents the modification of the internal state upon receiving new inputs and ρ_X sends the adjustments to the actuators, generating new outputs.

One of the main goals in the design of a CPS is to ensure its compositional integration ([26], p. 9). Each cyber-physical system $X = \langle I_X, O_X \rangle$ can be represented by a polynomial

$$p_X = O_X y^{I_X}.$$

Then, a morphism $\phi : p_X \rightarrow p_{X'}$ is defined as follows:

- *To each actuator $o \in O_X$, $\phi^{\rightarrow}$ assigns an actuator $o' \in O_{X'}$.*
- *To the sensors that contribute to activating an actuator o', $\phi^{\leftarrow}$ assigns those that activate o.*

Then, a composition of two CPSs, X and X', yields a new CPS, $\bar{X}$. The behavior of $\bar{X}$ is given by

$$[p_X \otimes p_{X'}, p_{\bar{X}}] - \mathbf{Coalg}.$$

Each element (s, ρ, μ) has the following properties:

- *$s \in S_{\bar{X}}$, where $S_{\bar{X}} = S_X \times S_{X'}$ is a* state*, capturing the dynamics of the composition of X and X';*
- *ρ, given a morphism $\phi : p_X \rightarrow p_{X'}$, assigns to state s and the readings of the sensors a response of the actuators of $\bar{X}$;*
- *μ updates the state.*

A final example shows how this approach can also be applied to symbolic AI.

Example 6 ([27,28])**.** *A Defeasible Logic Program consists of a finite set of facts, strict rules, and defeasible rules $\mathbf{P} = \langle \Pi, \Delta \rangle$, where Π denotes the set of facts and strict rules, while Δ denotes the set of defeasible rules. The set Π is the disjoint union of the sets of facts and of strict rules.*

The behavior of $\mathbf{P}$ is as follows. Given a query from a user, that is, a "question" about the validity of a literal $h \in Lit(\mathbf{P})$ in the language of the program, an argument $\langle A, h \rangle \in 2^{\Delta} \times Lit$ is generated, satisfying the following conditions:

1. *$h \in con(A)$ (where $con(A)$ yields the literals that can be derived using $\Pi \cup A$);*
2. *A is not contradictory (i.e., there is no atom $a \in con(A)$ such that its negation is also $a \in con(A)$);*
3. *If $h \in con(A')$, then $A' \not\subset A$, that is, A' is not a proper subset of A.*

The answer to the query is obtained through a process that can be seen as the play of a zero-sum game. For this, arguments are generated in opposition to the previous argument based on the relations between them. This defines a relation $\leq \subseteq R = \bigcup_{h \in Lit} R_h$ *of defeat among arguments.*

Defeaters can be either proper (using the strict $\prec$ *part of* $\leq$*) or blocking (satisfying the* $\sim$ *part of* $\leq$*). It is important to note that* $\leq$ *is not necessarily a preorder or a partial order, but it is enough to ensure that a game can be played.*

A warrant game for a literal l *is an extensive game with perfect information involving two players. These players are called the proponent and the opponent. We define the game as follows:*

- $P(\emptyset) =$ *proponent.*
- *The actions that the proponent can take at the root of the tree are all the arguments of the form* $\langle A, l \rangle$.
- *The actions after a nonterminal history* y *are the arguments* $\langle A', q \rangle$ *such that* $\langle A, p \rangle \leq \langle A', q \rangle$, *where* $\langle A, p \rangle$ *is the last component in* y. *In this case,* $P(y) =$ *proponent if* y *has an even length, and* $P(y) =$ *opponent if the length is odd.*
- *The payoff for the proponent assumes the value* 1*(win) at a history* $y \in Z$ *if the length of* y *is odd, and* -1 *otherwise. The payoff for the opponent is* -1 *times the payoff of the proponent.*

The composition of two programs, $\mathbf{P}$ *and* $\mathbf{P}'$, *yields all the games for the literals in* $Lit_{\mathbf{P}+\mathbf{P}'} = Lit_{\mathbf{P}} \cup Lit_{\mathbf{P}'}$. *For every* $l \in Lit_{\mathbf{P}+\mathbf{P}'}$, *we have a game* G_l *that can be obtained as the monoidal product of the games in* $\mathbf{P}$ *and* $\mathbf{P}'$.

The composition of DeLP programs can be represented by defining finite DeLP games. Each $G_l = \langle \{l\}, H_{G_l}, S_{G_l}, \pi_{G_l} \rangle$ *is characterized by the following elements:*

- l *is a literal for which the game is played.*
- H_{G_l} *is the class of histories satisfying the conditions of DeLP.*
- $S_{G_l} = S^{G_l}_{Prop} \times S^{G_l}_{Opp}$ *is the strategy set of the game, where* $S^{G_l}_i$ *is the set of strategies that player* i *(*$i \in \{Prop, Opp\}$*) can deploy in game* G_l, *yielding a history in* H_{G_l}.
- $\pi_{G_l} = \pi^{G_l}_{Prop} \times \pi^{G_l}_{Opp}$ *is a profile of payoff functions, where* $\pi^{G_l}_i : \mathbf{O}_{G_l} \to \{-1, 1\}$ *is the payoff function of player* $i\{Prop, Opp\}$ *in* G_l, *representing the idea that one of the two players wins and the other loses.*

Each G_l *can be identified with a polynomial* $p_{G_l} \simeq \pi_{G_l} y^{S_{G_l}}$.

Let us consider the class of warranted literals of a DeLP game p_{G_l}, *denoted as* $W(p_{G_l})$. *Suppose that we have another DeLP game,* $p_{G'_l}$. *We seek to define two operations,* $\oplus$ *and* $\otimes$, *such that* $W(p_{G_l} \oplus p_{G'_l})$ *is identified with* $W(p_{G_l}) \otimes W(p_{G'_l})$.

If both p_{G_l} *and* $p_{G'_l}$ *support* l, $p_{G_l} \oplus p_{G'_l}$ *also supports* l. *It has either a winning strategy for the proponent or the opponent. In the first case, we say then that* $l \in W(p_{G_l}) \otimes W(p_{G'_l})$. *In the latter case,* $\bar{l} \in W(p_{G_l}) \otimes W(p_{G'_l})$. *But only one of these two cases is possible:*

Proposition 7. *If* $l \in W(p_{G_l}) \otimes W(p_{G'_l})$, *then* $\bar{l} \notin W(p_{G_l}) \otimes W(p_{G'_l})$.

7. Conclusions

This paper discussed the question of representing interactions among intentional agents. We utilized the language of Category Theory and, in particular, constructions like *sheaves*, *hypergraph categories*, and *polynomial functors*.

The category defined in terms of the latter, $\mathcal{I}$, has as objects the interfaces between problems and their solutions, while the interaction among them is captured by coalgebras based on the internal homs of the interfaces. These homs represent sets of states that determine the arrangement of connections between the problems and their solutions. Furthermore, the connections are rearranged in response to the outputs obtained previously.

We intend to explore this formalism further and use it to represent specific problems. While the first step involves showing that $\mathcal{I}$ can reformulate known models, the real essence of this development lies in capturing new phenomena and establishing their relationships to the former.

Author Contributions: Conceptualization, F.T.; Formal analysis, F.T.; Investigation, F.T.; Resources, A.F.; Writing—review & editing, F.T. and A.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Agencia Nacional de Promoción de la Investigación, el Desarrollo Tecnológico y la Innovación, grant PICT 2019-01640.

Data Availability Statement: Data are contained within the article.

Acknowledgments: We would like to thank David I. Spivak and all the members of the Topos Institute in Berkeley (CA) for their inspiration and insightful comments that contributed to this paper as well as to a previous version ([29]). We would also like to thank Rocco Gangle and Gianluca Caterina for their deep insights and interesting discussions related to the topics of this paper. The usual disclaimer applies.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Goguen, J.A. Sheaf Semantics for Concurrent Interacting Objects. *Math. Struct. Comput. Sci.* **1992**, *2*, 159–191. [CrossRef]
2. Spivak, D.I. Applied Category Theory: Towards a Hard Science of Interdisciplinarity. Interdisciplinary Summer School 2023 of the Society for Multidisiciplinary and Fundamental Research. Available online: https://www.youtube.com/watch?v=K3NYTZxXbgM (accessed on 27 August 2024).
3. Hatcher, A. *Algebraic Topology*; Cambridge University Press: Cambridge, UK, 2002.
4. Marquis, J.-P. *From a Geometrical Point of View: A Study of the History and Philosophy of Category Theory*; Springer: Berlin/Heidelberg, Germany, 2009.
5. Abramsky, S.; Winschel, V. Coalgebraic Analysis of Subgame-Perfect Equilibria in Infinite Games without Discounting. *Math. Struct. Comput. Sci.* **2017**, *27*, 751–761. [CrossRef]
6. Frey, S.; Hedges, J.; Tan, J.; Zahn, P. Composing Games into Complex Institutions. *PLoS ONE* **2023**, *18*, e0283361. [CrossRef] [PubMed]
7. Ghani, N.; Hedges, J.; Winschel, V.; Zahn, P. Compositional Game Theory. In Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science, Oxford, UK, 9–12 July 2018; pp. 472–481.
8. Rozen, V.; Zhitomirski, G. A Category Approach to Derived Preference Relations in Some Decision-Making Problems. *Math. Soc. Sci.* **2006**, *51*, 257–273. [CrossRef]
9. Crespo, R.; Tohmé, F. The Future of Mathematics in Economics: A Philosophically Grounded Proposal. *Found. Sci.* **2017**, *22*, 677–693. [CrossRef]
10. Goldblatt, R. *Topoi. The Categorical Analysis of Logic*; North-Holland: Amsterdam, The Netherlands, 1984.
11. Barr, M.; Wells, C. Category Theory. Lecture Notes for ESSLLI. 1999. Available online: https://fldit-www.cs.tu-dortmund.de/~peter/barrwells.pdf (accessed on 27 August 2024).
12. Adámek, J.; Herrlich, H.; Strecker, G. Abstract and Concrete Categories. The Joy of Cats. 2004. Available online: http://katmat.math.uni-bremen.de/acc/acc.pdf (accessed on 27 August 2024).
13. Lawvere, F.W.; Schanuel, S. *Conceptual Mathematics: A First Introduction to Categories*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2009.
14. Spivak, D.I. *Category Theory for the Sciences*; MIT Press: Cambridge, MA, USA, 2014.
15. Fong, B.; Spivak, D.I. *An Invitation to Applied Category Theory: Seven Sketches in Compositionality*; Cambridge University Press: Cambridge, MA, USA, 2019.
16. Southwell, R.; Gupta, N. *Categories and Toposes: Visualized and Explained*; KDP Publishing: Seattle, WA, USA, 2021.
17. Cheng, E. *The Joy of Abstraction*; Cambridge University Press: Cambridge, MA, USA, 2022.
18. Luenberger, D. Projection Pricing. *J. Optim. Theory Appl.* **2001**, *109*, 1–25. [CrossRef]
19. Tohmé, F.; Caterina, G.; Gangle, R. Local and Global Optima in Decision-Making: A Sheaf-Theoretical Analysis of the Difference between Classical and Behavioral Approaches. *Int. J. Gen. Syst.* **2017**, *46*, 879–897. [CrossRef]
20. Tohmé, F.; Viglizzo, I. A Categorical Representation of Games. *arXiv* **2023**, arXiv:2309.15981.
21. Niu, N.; Spivak, D.I. Polynomial Functors: A Mathematical Theory of Interaction. *arXiv* **2023**, arXiv:2312.00990.
22. Shapiro, B.; Spivak, D.I. Dynamic Categories, Dynamic Operads: From Deep Learning to Prediction Markets. *arXiv* **2022**, arXiv:2205.03906. [CrossRef]
23. Nöldeke, G.; Samuelson, L. The Implementation Duality. *Econometrica* **2018**, *86*, 1283–1324. [CrossRef]

24. Hurwicz, L. Institutions as Families of Game Forms. *Jpn. Econ. Rev.* **1996**, *47*, 113–132. [CrossRef]
25. Yuan, Y. On the Power of Foundation Models. In Proceedings of the International Conference on Machine Learning, Honolulu, HI, USA, 23–29 July 2023; pp. 40519–40530.
26. Platzer, A. *Logical Foundations of Cyber-Physical Systems*; Springer: Berlin/Heidelberg, Germany, 2018.
27. García, A.; Simari, G.R. Defeasible Logic Programming: An Argumentative Approach. *Theory Pract. Log. Program.* **2004**, *4*, 95–138. [CrossRef]
28. Viglizzo, I.; Tohmé, F.; Simari, G. The Foundations of DeLP: Defeating Relations, Games and Truth Values. *Ann. Math. Artif. Intell.* **2009**, *57*, 181–204. [CrossRef]
29. Tohmé, F. Dynamic Arrangements in Economic Theory: Level-Agnostic Representations. *arXiv* **2023**, arXiv:2309.06383.