Enacted Cognition as Adjunction: A Topos-Theoretic Approach to Enactivism

Conor L. Mahany

September 21, 2024

Abstract

This paper introduces a novel framework for understanding cognition by leveraging category theory and topos theory. Departing from traditional cognitivist models that emphasize internal representations of an external world, we conceptualize cognition as a dynamic interplay between action and perception, modeled as an adjunction between categories. By treating the self as the initial object in the category of contexts, we draw parallels with the Buddhist concept of *śūnyatā* (emptiness), offering an interdisciplinary perspective that bridges Western mathematical frameworks and Eastern philosophical thought. We illustrate our framework with concrete examples and discuss its implications for cognitive science, philosophy, and future research.

If I were to say that the so-called philosophy of this fellow Hegel is a colossal piece of mystification which will yet provide posterity with an inexhaustible theme for laughter at our times, that it is a pseudo-philosophy paralyzing all mental powers, stifling all real thinking, and, by the most outrageous misuse of language, putting in its place the hollowest, most senseless, thoughtless, and, as is confirmed by its success, most stupefying verbiage, I should be quite right.

Further, if I were to say that this summus philosophus ... scribbled nonsense quite unlike any mortal before him, so that whoever could read his most eulogized work, the so-called *Phenomenology of Spirit*, without feeling as if he were in a madhouse, would qualify as an inmate for Bedlam, I should be no less right.

—Arthur Schopenhauer, *On the Basis of Morality* (1839)

The slogan is "Adjoint functors arise everywhere".

—Saunders Mac Lane, *Categories for the Working Mathematician* (1971)

Form is emptiness, emptiness is form.

—Heart Sūtra of the Mahayana Buddhist tradition

Contents

1 Introduction

Understanding cognition remains one of the most challenging and profound quests in both science and philosophy. Traditional Western approaches often model cognition as the manipulation of internal representations of an external world, emphasizing a computational view of the mind. In contrast, enactivism, an emerging paradigm in cognitive science, posits that cognition arises through a dynamic interaction between an organism and its environment, without reliance on internal representations [1].

In this paper, we aim to provide a rigorous mathematical framework for enactivism by employing tools from category theory and topos theory. These mathematical structures allow us to model the co-evolution of action and perception as an adjunction between categories, capturing the essence of enaction in a formal setting. Furthermore, by conceptualizing the self as the initial object in the category of contexts, we find a natural correspondence with the Buddhist concept of śūnyatā or emptiness [2].

Our goals in this paper are threefold:

1. To introduce and justify the use of category theory and topos theory in modeling cognitive processes.

2. To provide concrete examples that illustrate how action and perception can be modeled as functors between categories.

3. To discuss the philosophical implications of this framework, especially the connection between the empty self and śūnyatā, and its impact on cognitive science.

1.1 Structure of the Paper

The paper is organized as follows:

- In Section 2, we provide the philosophical motivation for our approach, discussing the limitations of representational models and introducing enactivism.

- Section 3 introduces the necessary mathematical background in category theory and topos theory, with intuitive explanations and examples.

4

- In Section 4, we develop our main framework, modeling enaction as an adjunction between the categories of contexts and behaviors.

- Section 5 revisits the concept of the self as the initial object and explores its connection to *śūnyatā*.

- Section 6 discusses the implications of our framework for cognitive science and philosophy.

- In Section 7, we relate our work to existing literature, highlighting similarities and differences.

- Section 8 outlines future research directions, including empirical validation and further theoretical development.

- We conclude in Section 9 by summarizing our findings and their significance.

2   Philosophical Motivation

Wittgenstein had notable objections, especially in his *Philosophical Investigations*, to the law of excluded middle (LEM). One promising feature of a topos is that the logic of its internal language is intuitionistic — it does not in general include LEM. Mac Lane and Moerdijk provide a topos-theoretic expression—an *internalization*—of the ZFC axioms and show their weak equivalence to ZFC in set theory. Notably, the topos they use to demonstrate this is *well-pointed* and therefore Boolean (i.e., the logic of its internal language is Boolean).

2.1   Limitations of Representational Models

Traditional cognitive science often relies on the assumption that cognition involves internal representations of an external world. This view, rooted in computationalism, treats the mind as a symbol-manipulating system [4]. However, this model faces several challenges:

- Frame Problem: Determining which aspects of a complex environment are relevant to a given context is computationally intractable [6].

- Embodiment: It neglects the role of the body and the environment in shaping cognitive processes [3].

- Dynamic Interaction: It fails to account for the continuous and reciprocal interaction between an organism and its environment [7].

2.2   Enactivism

Enactivism offers an alternative by proposing that cognition arises through an organism's active engagement with its environment. Key principles include:

- Autonomy: Cognitive systems are self-organizing and self-maintaining.

- Embodiment: The body shapes the mind, and cognition cannot be separated from the physical form.

- Embeddedness: Cognition is situated within and cannot be isolated from the environment.

5

- Dynamic Co-emergence: Organism and environment co-determine each other in a continuous feedback loop [1].

2.3 The Need for a Formal Framework

While enactivism provides a compelling philosophical perspective, it lacks a rigorous mathematical formalism that can model the dynamic interplay between action and perception. Category theory and topos theory offer powerful tools for abstracting and formalizing such complex interactions.

2.4 Preview

Topos theory offers a generalization of set theory and its foundations. Here, the fundamental relation is not the structural, hierarchical relation $\in$ of set membership but the arrow $X \to Y$ between objects. Topos theory also provides an abstract description and generalization of point-set topology, encapsulating the notion of local determination through (pre)sheaves and context-dependent truth. The quintessential image is that of a function $f\colon X \to \mathbb{R}$ which is continuous when restricted to some open subset $U \subset X$.

In what follows, I will illustrate that

$$
\begin{array}{ccc}
\{x \mid \varphi(x)\} & \longrightarrow & 1 \\
\Big\downarrow & & \Big\downarrow {\scriptstyle \text{true}} \\
U \xrightarrow{\ \alpha\ } X & \xrightarrow{\ \varphi(x)\ } & \Omega
\end{array}
$$

is a representation of local truth. Let $\mathcal{C}$ denote the category of contexts and let $\mathcal{B}$ denote the category of behaviors. Then, enaction is modeled as an adjunction $\mathcal{A} \dashv \mathcal{P}$, where action is a functor $\mathcal{A}\colon \mathcal{C} \to \mathcal{B}$ and perception is a functor $\mathcal{P}\colon \mathcal{B} \to \mathcal{C}$. That is, enaction consists of a bijection

$$
\phi_{CB}\colon \operatorname{Hom}(\mathcal{A}(C), B) \to \operatorname{Hom}(C, \mathcal{P}(B))
$$

that is natural in both C and B, for any context C and behavior B. In other words, each morphism $\mathcal{A}(C) \to B$ of behaviors is *enacted* by a unique morphism $C \to \mathcal{P}(B)$ of contexts.

3 Mathematics

In this section, we introduce the mathematical concepts necessary for our framework, aiming to make them accessible through intuitive explanations and examples.

3.1 Category Theory

Definition 3.1. *A **category** $\mathcal{C}$ consists of:*

(a) *A collection of **objects** $X, Y, Z, \ldots$.*

(b) *For each pair of objects X and Y, a collection of **morphisms** (or **arrows**) $f\colon X \to Y$.*

(c) *For each object X, an **identity morphism** $1_X\colon X \to X$.*

(d) A **composition law** that assigns to each pair of morphisms $f\colon X \to Y$ and $g\colon Y \to Z$ a morphism $g \circ f\colon X \to Z$.

These must satisfy the following axioms:

1. **Associativity**: For morphisms $f\colon W \to X$, $g\colon X \to Y$, and $h\colon Y \to Z$, we have $h \circ (g \circ f) = (h \circ g) \circ f$.

2. **Identity**: For every morphism $f\colon X \to Y$, $1_Y \circ f = f$ and $f \circ 1_X = f$.

A prototypical example of a category is **Set**, the category whose objects are sets and whose morphisms are functions between sets.

3.1.1 Specific Categories

Definition 3.2. *Let* **Con** *denote the* **category of contexts**, *where:*

1. **Objects**: *Each object is a moment of conscious experience.*

2. **Morphisms**: *Arrows represent causal structures or transitions between contexts.*

Definition 3.3. *Let* **Beh** *denote the* **category of behaviors**, *where:*

1. **Objects**: *Each object is a behavior in lived experience.*

2. **Morphisms**: *Arrows represent causal structures or transitions between behaviors.*

These categories serve as the foundational structures for modeling the dynamics of cognition as an interplay between contexts and behaviors.

3.1.2 Functors

Definition 3.4. *A* **functor** $F\colon \mathcal{C} \to \mathcal{D}$ *between categories* $\mathcal{C}$ *and* $\mathcal{D}$ *assigns:*

- *To each object X in $\mathcal{C}$, an object $F(X)$ in $\mathcal{D}$.*

- *To each morphism $f\colon X \to Y$ in $\mathcal{C}$, a morphism $F(f)\colon F(X) \to F(Y)$ in $\mathcal{D}$.*

Such that:

1. $F(id_X) = id_{F(X)}$.

2. $F(g \circ f) = F(g) \circ F(f)$.

3.2 Limits and Colimits

Understanding limits and colimits is essential for grasping the structural aspects of categories that model cognitive processes.

Definition 3.5. *A* **limit** *of a diagram $D\colon \mathcal{J} \to \mathcal{C}$ in a category $\mathcal{C}$ is a universal cone to D. Concretely, it consists of an object L in $\mathcal{C}$ and a family of morphisms $\{\pi_j\colon L \to D(j)\}_{j \in \mathcal{J}}$ such that for every morphism $f\colon j \to k$ in $\mathcal{J}$, $D(f) \circ \pi_j = \pi_k$. Moreover, for any other cone $(N, \{\phi_j\colon N \to D(j)\})$, there exists a unique morphism $u\colon N \to L$ such that $\pi_j \circ u = \phi_j$ for all $j \in \mathcal{J}$.*

Definition 3.6. *A* **colimit** *is the dual notion of a limit. It is a universal cocone from a diagram $D\colon \mathcal{J} \to \mathcal{C}$.*

3.2.1 Finite Limits

Definition 3.7. *A category $\mathcal{C}$ has **finite limits** if it has all limits of finite diagrams. This includes the existence of products, equalizers, and a terminal object.*

Definition 3.8. *A **pullback** (or fibered product) of two morphisms $f\colon X \to Z$ and $g\colon Y \to Z$ in a category $\mathcal{C}$ is a limit of the diagram consisting of X, Y, and Z with f and g as the morphisms to Z. It is denoted by:*

$$
\begin{array}{ccc}
P & \longrightarrow & Y \\
\downarrow & & \downarrow{\scriptstyle g} \\
X & \xrightarrow{\ f\ } & Z
\end{array}
$$

where P is the pullback object.

In **Set**, the pullback of f and g is the set $\{(x,y) \in X \times Y \mid f(x) = g(y)\}$.

3.3 Monomorphisms and Epimorphisms

Definition 3.9. *A morphism $f\colon X \to Y$ in a category $\mathcal{C}$ is a **monomorphism** (or **mono**) if for all objects Z and all pairs of morphisms $g_1, g_2\colon Z \to X$, $f \circ g_1 = f \circ g_2$ implies $g_1 = g_2$.*

Definition 3.10. *A morphism $f\colon X \to Y$ in a category $\mathcal{C}$ is an **epimorphism** (or **epi**) if for all objects Z and all pairs of morphisms $g_1, g_2\colon Y \to Z$, $g_1 \circ f = g_2 \circ f$ implies $g_1 = g_2$.*

In **Set**, monomorphisms are injective functions, and epimorphisms are surjective functions.

3.4 Equalizers and Coequalizers

Definition 3.11. *Given two parallel morphisms $f, g\colon X \rightrightarrows Y$ in a category $\mathcal{C}$, an **equalizer** of f and g is a morphism $e\colon E \to X$ such that $f \circ e = g \circ e$, and for any morphism $m\colon M \to X$ with $f \circ m = g \circ m$, there exists a unique morphism $u\colon M \to E$ such that $e \circ u = m$.*

Definition 3.12. *A **coequalizer** of two parallel morphisms $f, g\colon X \rightrightarrows Y$ is a morphism $q\colon Y \to Q$ such that $q \circ f = q \circ g$, and for any morphism $m\colon Y \to M$ with $m \circ f = m \circ g$, there exists a unique morphism $u\colon Q \to M$ such that $u \circ q = m$.*

3.5 Subobjects and Subobject Classifier

Definition 3.13. *A **subobject** of an object X in a category $\mathcal{C}$ is an equivalence class of monomorphisms into X, where two monomorphisms $f\colon A \to X$ and $g\colon B \to X$ are considered equivalent if there exists an isomorphism $h\colon A \to B$ such that $f = g \circ h$.*

Definition 3.14. *A **subobject classifier** in a category $\mathcal{C}$ is an object Ω together with a monomorphism $\top\colon 1 \to \Omega$ such that for every monomorphism $m\colon A \to X$, there exists a unique morphism $\chi_m\colon X \to \Omega$ making the following diagram a pullback:*

$$
\begin{array}{ccc}
A & \xrightarrow{\ m\ } & X \\
\downarrow & & \downarrow{\scriptstyle \chi_m} \\
1 & \xrightarrow{\ \top\ } & \Omega
\end{array}
$$

In **Set**, the subobject classifier is the two-element set $1 + 1 = \{\text{true}, \text{false}\}$, and χ_m maps each element of X to true if it is in the subset A and to false otherwise.

3.6 Topos Theory

Definition 3.15. *A **topos** $\mathcal{E}$ is a category that satisfies the following properties:*

1. *$\mathcal{E}$ has all finite limits.*

2. *$\mathcal{E}$ has exponentials; that is, for any objects X and Y in $\mathcal{E}$, there exists an exponential object Y^X.*

3. *$\mathcal{E}$ has a subobject classifier Ω.*

Topos theory generalizes set theory and provides a framework for mathematical logic and geometry. A topos can be seen as a universe of varying contexts, each equipped with its own internal logic.

3.6.1 Internal Logic of a Topos

Each topos $\mathcal{E}$ has an internal language that is intuitionistic, meaning that it does not necessarily satisfy the law of excluded middle (LEM). The internal logic allows for reasoning about objects and morphisms within the topos as if they were sets and functions, but with a more flexible logical foundation.

3.6.2 Sheaves and Presheaves

Definition 3.16. *A **presheaf** on a category $\mathcal{C}$ is a functor $F\colon \mathcal{C}^{op} \to \mathbf{Set}$.*

Definition 3.17. *A **sheaf** on a topological space X is a presheaf that satisfies the sheaf axioms, which ensure that local data can be uniquely glued together to form global data.*

Sheaf theory provides a way to systematically track locally defined data attached to the open sets of a topological space, enabling the formalization of context-dependent truth.

3.7 Adjunctions in Category Theory

Definition 3.18. *An **adjunction** between two categories $\mathcal{C}$ and $\mathcal{D}$ consists of two functors:*

$$\mathcal{A}\colon \mathcal{C} \to \mathcal{D} \quad and \quad \mathcal{P}\colon \mathcal{D} \to \mathcal{C}$$

and a natural bijection

$$\phi_{CB}\colon Hom_{\mathcal{D}}(\mathcal{A}(C), B) \cong Hom_{\mathcal{C}}(C, \mathcal{P}(B))$$

for all objects C in $\mathcal{C}$ and B in $\mathcal{D}$.

In this context, $\mathcal{A}$ is the **left adjoint** and $\mathcal{P}$ is the **right adjoint** and we write $\mathcal{A} \dashv \mathcal{P}$.

4 Enactment as Adjunction

4.1 Defining the Categories

Definition 4.1. *The category* **Con** *(Contexts):*

- **Objects**: *Contexts or situations in which cognition occurs.*

- **Arrows**: *Transitions or relationships between contexts.*

Example 4.1. *A context could be a visual scene, and a morphism could represent the act of shifting attention within that scene.*

Definition 4.2. *The category* **Beh** *(Behaviors):*

- **Objects**: *Behaviors or actions taken by an agent.*

- **Arrows**: *Processes transforming one behavior into another.*

Example 4.2. *A behavior could be reaching for an object, and a morphism could represent the modification of that action due to environmental feedback.*

4.2 Modeling action and perception as functors

We model action and perception as functors between these categories:

- **Action Functor $\mathcal{A}$: Con $\rightarrow$ Beh**: Maps contexts to behaviors.

- **Perception Functor $\mathcal{P}$: Beh $\rightarrow$ Con**: Maps behaviors to resulting contexts.

4.3 Establishing the Adjunction

The adjunction $\mathcal{A} \dashv \mathcal{P}$ captures the reciprocal relationship between action and perception.

Theorem 4.1. *For any context C and behavior B there is a bijection*

$$\phi_{C,B} \colon Hom_{\textbf{Beh}}(\mathcal{A}(C), B) \cong Hom_{\textbf{Con}}(C, \mathcal{P}(B))$$

between such behavioral and contextual transformations that is natural in both C and B.

Proof. This is straight up the definition of $\mathcal{A} \dashv \mathcal{P}$. $\square$

This adjunction formalizes the idea that actions are generated from contexts and, in turn, behaviors influence subsequent contexts through perception. It embodies the enactivist view that cognition is a continuous loop of action and perception.

In this diagram, the dotted arrow represents the process of enaction, combining action and perception into a cohesive cognitive process.

5 The Empty Self and *Śūnyatā*

5.1 Initial Object in Category Theory

Definition 5.1. *An **initial object** in a category $\mathcal{C}$ is an object 0 such that for every object X in $\mathcal{C}$, there exists a unique morphism $0 \to X$.*

Example 5.1. *In* **Set**, *the empty set $\emptyset$ is the initial object.*

5.2 Self as the Initial Object

We propose that the self can be modeled as the initial object in **Con**.

Remark 5.1. *The self, being the origin of all experiences, maps uniquely to every context but lacks inherent structure within this framework, aligning with the notion of an initial object.*

5.3 Connection to *Śūnyatā*

In Buddhist philosophy, *śūnyatā* refers to the emptiness of inherent existence in all phenomena, including the self [2].

"Form is emptiness, emptiness is form." – *Heart Sutra*

Our model reflects this by representing the self as devoid of intrinsic properties, defined only through its relations (morphisms) to contexts.

5.4 Philosophical Implications

This conceptualization challenges the traditional notion of a fixed, representational self and supports a relational understanding of identity.

6 Examples and Applications

6.1 A Simple Cognitive Process

Consider an agent navigating a maze.

- **Contexts**: The agent's current location and perception of the maze.

- **Behaviors**: Movements such as turning left or right.

- **Action Functor** $\mathcal{A}$: Given a context (current location), the functor maps to a behavior (choose direction).

- **Perception Functor** $\mathcal{P}$: Given a behavior (movement), the functor maps to a new context (new location).

The adjunction captures how the agent's decision to move influences its perception and how that perception informs future actions.

6.2 Language Acquisition

In language learning:

- **Contexts**: Linguistic environments, such as exposure to certain vocabulary or grammar structures.

- **Behaviors**: Speech acts or language production attempts.

- **Adjunction**: The interplay between hearing language (perception) and speaking (action), leading to language acquisition.

6.3 Implications for Artificial Intelligence

Modeling cognition in this way can inform the development of AI systems that learn and adapt through interaction, rather than relying solely on pre-programmed representations.

7 Implications for Cognitive Science and Philosophy

7.1 Relational Cognition

Our framework emphasizes that cognition is not a computation over representations but a relational process emerging from interactions.

7.2 Context-Dependence and Situatedness

By modeling contexts explicitly, we highlight the importance of situatedness in cognitive processes, aligning with embodied cognition theories [3].

7.3 Bridging Eastern and Western Thought

The correspondence between the empty self and *śūnyatā* offers a philosophical bridge that enriches both traditions.

8 Related Work

8.1 Enactivism in Cognitive Science

Our work builds upon the foundational ideas of enactivism [1], extending them with mathematical rigor.

8.2 Category Theory in Cognitive Models

Previous attempts to apply category theory to cognition include conceptual spaces [**?**] and neural network architectures [**?**].

8.3 Philosophical Perspectives on the Self

The concept of the self as relational is explored in both Western [9] and Eastern philosophies [2].

9 Future Work

9.1 Formalizing the Categories

Further research will focus on precisely defining the objects and morphisms in **Con** and **Beh**, potentially using empirical data.

9.2 Empirical Validation

We aim to test our model by applying it to specific cognitive tasks and comparing predictions with experimental results.

9.3 Extensions to Other Cognitive Domains

Exploring how this framework applies to social cognition, emotions, and consciousness.

10 Conclusion

We have presented a novel framework that models cognition as an adjunction between contexts and behaviors, capturing the dynamic interplay central to enactivism. By conceptualizing the self as the initial object, we align with the philosophical concept of *śūnyatā*, offering a relational understanding of identity. This interdisciplinary approach provides new insights into cognitive science and opens avenues for future research.

Acknowledgments

References

[1] Varela, F. J., Thompson, E., & Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience.* MIT Press.

[2] Garfield, J. L. (1995). *The Fundamental Wisdom of the Middle Way: Nāgārjuna's Mūlamadhyamakakārikā.* Oxford University Press.

[3] Clark, A. (1997). *Being There: Putting Brain, Body, and World Together Again.* MIT Press.

[4] Fodor, J. A. (1980). *Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology.* Behavioral and Brain Sciences, 3(1), 63-73.

[5] Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind.* Harvard University Press.

[6] Pylyshyn, Z. W. (1987). *The Robot's Dilemma: The Frame Problem in Artificial Intelligence.* Ablex Publishing.

[7] Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind.* Harvard University Press.

[8] Wittgenstein, L. (1953). *Philosophical Investigations.* Blackwell Publishing.

[9] Heidegger, M. (1962). *Being and Time.* (J. Macquarrie, E. Robinson, Trans.). Harper & Row.