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.

Contents

1 Motivation

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:

1.1 Philosophical Background

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 always 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).

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 morphism $X \to Y$ from one object X to another Y. 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

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

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

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

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.

1.2 The Empty Self and *Śūnyatā*

In Buddhist philosophy, *śūnyatā* or emptiness refers to the absence of inherent existence in all phenomena, including the self. This concept resonates with the categorical notion of the self as the initial object in the category of contexts. An initial object is one that has a unique morphism to every other object in the category, implying that the self does not possess intrinsic properties but is defined solely through its relations to various contexts. Notably, the initial object in the category of sets is the empty set $\emptyset$.

This aligns with the enactive view that cognition arises from the dynamic interplay between an agent and its environment, rather than from internal representations. This perspective challenges the traditional cognitivist model that posits a central, representational self mediating between perception and action.

This alignment not only bridges Eastern and Western philosophical traditions but also provides a robust framework for understanding cognition as fundamentally relational and context-dependent (as well as behavior-dependent, in a symmetric sense). Future sections will delve deeper into how this conceptualization impacts our understanding of cognitive processes and the nature of the self.

2 Mathematics

At this point, I introduce the technical framework necessary to formally express the conception of enactivism presented earlier. The language of category theory, and more specifically topos theory, provides the tools to describe the co-evolution of action and perception abstractly and rigorously.

2.1 Category Theory

Definition 2.1. *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.

2.1.1 Specific Categories

Definition 2.2. *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 2.3. *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.

2.2 Limits and Colimits

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

Definition 2.4. *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 2.5. *A **colimit** is the dual notion of a limit. It is a universal cocone from a diagram $D\colon \mathcal{J} \to \mathcal{C}$.*

2.2.1 Finite Limits

Definition 2.6. *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 2.7. *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 $\{(x,y) \in X \times Y \mid f(x) = g(y)\}$.

2.3 Monomorphisms and Epimorphisms

Definition 2.8. *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 2.9. *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.

2.4 Equalizers and Coequalizers

Definition 2.10. *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 2.11. *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 $u \circ q = m$.*

2.5 Subobjects and Subobject Classifier

Definition 2.12. *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 2.13. *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.

2.6 Topos Theory

Definition 2.14. *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.

2.6.1 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.

2.6.2 Sheaves and Presheaves

Definition 2.15. *A **presheaf** on a category $\mathcal{C}$ is a functor $F\colon \mathcal{C}^{op} \to \mathbf{Set}$.*

Definition 2.16. *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 Enactment as Adjunction

Having established the necessary categorical framework, I now formalize the concept of enaction within this context.

3.1 Adjunctions in Category Theory

Definition 3.1. *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**.

3.2 Modeling Enactment as an Adjunction

In the framework of enactivism, enaction can be modeled as an adjunction between the category of contexts **Con** and the category of behaviors **Beh**. Specifically:

- The functor $\mathcal{A}\colon \mathbf{Con} \to \mathbf{Beh}$ represents **action**.

- The functor $\mathcal{P}\colon \mathbf{Beh} \to \mathbf{Con}$ represents **perception**.

The adjunction $\mathcal{A} \dashv \mathcal{P}$ encapsulates the bidirectional relationship between action and perception, embodying the co-evolutionary dynamics central to enactivism.

3.3 Properties of the Adjunction

- **Unit and Counit**: The adjunction is characterized by natural transformations known as the unit $\eta\colon \mathrm{Id}_{\mathbf{Con}} \to \mathcal{P} \circ \mathcal{A}$ and the counit $\epsilon\colon \mathcal{A} \circ \mathcal{P} \to \mathrm{Id}_{\mathbf{Beh}}$.

- **Adjoint Functor Theorem**: Under certain conditions, adjoint functors can be uniquely determined, ensuring the robustness of the enaction model.

3.4 Implications for Cognition

This adjoint relationship implies that actions are not merely responses to perceptions but are intertwined in a reciprocal process. Each action alters the context, which in turn affects subsequent perceptions, leading to a continuous loop of interaction.

4 The Empty Self Revisited

Building upon the categorical framework, I revisit the concept of the self as the initial object in the category of contexts.

4.1 Initial Objects in Category Theory

Definition 4.1. *An **initial object** in a category $\mathcal{C}$ is an object I such that for every object X in $\mathcal{C}$, there exists a unique morphism $I \to X$.*

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

4.2 Self as an Initial Object

Interpreting the self as the initial object implies that the self does not possess inherent properties but is defined entirely through its interactions with various contexts. This mirrors the Buddhist notion of *śūnyatā*, where the self is seen as empty of intrinsic existence.

4.3 Consequences for Cognitive Modeling

This perspective challenges the traditional view of the self as a central, stable entity mediating between perception and action. Instead, it posits the self as fundamentally relational and defined by its engagements with different contexts.

5 Subobject Classifier and Context-Dependent Truth

5.1 Subobject Classifier in Con

Assuming that **Con** is a topos, it possesses a subobject classifier Ω. This allows for the representation of predicates and propositions within the category, facilitating context-dependent truth values.

5.2 Modeling Local Truth

The commutative diagram presented earlier illustrates how local truth is captured within the topos-theoretic framework. Each context U and its relationship to a behavior via α and $\varphi(x)$ can be understood through the lens of the subobject classifier and the internal logic of the topos.

6 Implications for Enactivism and Cognitive Science

6.1 Relational Cognition

By modeling cognition as an adjunction between contexts and behaviors, this framework emphasizes the relational and dynamic nature of cognitive processes. Cognition is not a process of internal computation but a continuous interaction with the environment.

6.2 Context-Dependence

The use of topos theory highlights the importance of context in determining truth and meaning. Cognitive processes are inherently context-dependent, aligning with enactivist views that cognition cannot be fully understood without considering the situatedness of the agent.

6.3 Bridging Eastern and Western Philosophies

The correspondence between the categorical concept of the empty self and the Buddhist notion of *śūnyatā* offers a unique bridge between Eastern and Western philosophical traditions, enriching the theoretical foundations of cognitive science.

7 Related Work

7.1 Enactivism in Cognitive Science

Enactivism has been a growing perspective in cognitive science, emphasizing the role of embodied action in cognition. Key contributors include Varela, Thompson, and Rosch, who argue against the traditional computational view of the mind.

7.2 Category Theory in Cognitive Science

While category theory has found applications in various fields, its use in cognitive science is relatively novel. Previous work has explored categorical models for neural networks, conceptual spaces, and cognitive architectures, laying the groundwork for the present approach.

7.3 Topos Theory and Logic

Topos theory has been instrumental in providing a unifying framework for different areas of mathematics and logic. Its application to cognitive science represents an innovative extension of its utility, offering new perspectives on the foundations of cognition.

8 Future Work

8.1 Formalizing the Categories of Contexts and Behaviors

Further development is needed to precisely define the categories **Con** and **Beh**, including their objects, morphisms, and structural properties. This formalization will facilitate the application of the adjoint model to specific cognitive phenomena.

8.2 Empirical Validation

Connecting the abstract categorical model to empirical data in cognitive science is a crucial next step. This could involve mapping specific cognitive processes to morphisms in the categories or using the framework to generate testable predictions.

8.3 Expanding the Logical Framework

Exploring the internal logic of the topos and its implications for cognitive processes can provide deeper insights. This includes investigating how intuitionistic logic influences the interpretation of cognitive interactions.

9 Conclusion

This paper presents a novel approach to modeling cognition through the lens of category and topos theory, framing enactivism as an adjunction between contexts and behaviors. By positioning the self as an initial object, aligning with the concept of *śūnyatā*, this framework challenges traditional cognitivist models and offers a relational, context-dependent understanding of cognitive processes. Future work will focus on formalizing the categorical structures and bridging the gap between abstract theory and empirical cognitive science.

Acknowledgments

References

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] Mac Lane, S., & Moerdijk, I. (1998). *Sheaves in Geometry and Logic: A First Introduction to Topos Theory*. Springer.

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

[5] Holland, P. (1995). *Hidden Order: How Adaptation Builds Complexity*. Addison-Wesley.