

Enacted cognition as adjunction: a topos-theoretic approach to enactivism

Conor L. Mahany

September 21, 2024

Abstract

In this paper I radically depart from traditionally Western, cognitivist descriptions of cognition as fundamentally involving the representation of some external world by the agent. Instead of focusing solely on how action is determined from perception or vice versa, I use the language of category theory to provide an abstract description of the *co-evolution* of action and perception, e.g., as an *adjunction*. I show that the self is *empty* in the sense that the self is the initial object in the category of contexts. This corresponds directly to *śūnyatā* in Buddhism, and hence this work can be seen as an implementation of cognition as enaction (or *enactivism*).

Contents

1	Motivation	1
1.1	Philosophical background	2
1.2	The empty self and <i>śūnyatā</i>	2
2	Mathematics	3
2.1	Category theory	3
2.2	Topos theory	3

1 Motivation

Here I have the nigh impossible task of motivating a completely foreign and aggressively abstract nonsense¹ for a purpose that can as yet be only very dimly perceived. It is impossible for me to provide a convincing explanation on the “first pass”, as it were. Equivalently, it is impossible for you to obtain a working understanding after your first read (assuming you are brave enough to endure this in its entirety). My hope is that, in your reading this and in my writing this, we both come closer to an understanding.

¹Saunders Mac Lane. “The PNAS way back then”. Proc. Natl. Acad. Sci. USA Vol. 94, pp. 5983–5985, June 1997. “The first of these papers is a more striking case; it introduced the very abstract idea of a “category”—a subject then called “general abstract nonsense!”

To begin, let's consider some reasons why topos theory might be a viable candidate for modelling embodied cognition.

1.1 Philosophical background

Wittgenstein had some notable beef, especially in the *Philosophical Investigations*, with the law of excluded middle (LEM). One very promising feature of a topos is that the logic of its internal language is intuitionistic — that is, it does not always include LEM. Mac Lane and Moerdijk give a topos-theoretic expression — that is, an *internalization* — of the ZFC axioms and show their weak equivalence to ZFC in set theory. It is worth emphasizing that the topos they use to demonstrate this is “well-pointed” and therefore Boolean (e.g., the logic of its internal language is Boolean).

So, topos theory can provide us with a sort of “generalization” of set theory, and of its foundations. In which, as I will emphasize throughout, the fundamental relation is not the structural, hierarchical relation \in of set membership, but, instead, the arrow $X \rightarrow Y$ from one object X to another Y . Topos theory also provides an abstract description and generalization of point-set topology, that is, a notion of how something can be determined locally, through the concept of a (pre)sheaf (e.g., uniquely collatable, e.g. existence of an equalizer), or, equivalently, a notion of *context-dependent truth*.

In what follows I will show that

$$\begin{array}{ccccc}
 & & \{x \mid \varphi(x)\} & \longrightarrow & 1 \\
 & \nearrow m & \downarrow & & \downarrow \text{true} \\
 U & \xrightarrow{\alpha} & X & \xrightarrow{\varphi(x)} & \Omega
 \end{array}$$

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

$$\phi_{CB}: \text{Hom}(\mathcal{A}(C), B) \rightarrow \text{Hom}(C, \mathcal{P}(B))$$

that is natural in both C and B , for any context C and behavior B .

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. Interestingly, 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

I can no longer defer the technical specification of the language that I will be using to formally express my conception of enactivism. It is the language of mathematics in general and topos theory in particular.

2.1 Category theory

Definition 2.1. A *category* \mathcal{C} consists of a collection of objects X, Y, \dots and of a collection of arrows $f: X \rightarrow Y$ that is closed under associative composition and has an identity arrow $1_X: X \rightarrow X$ for every object X of \mathcal{C} .

A prototypical example of a category is **Set**, the category with sets as objects and functions as arrows.

also mention category **Con** of contexts, where each object is a moment of conscious experience, and arrows represent causal structure

and category **Beh** of behaviors, where each object is a behavior in lived experience and arrows also represent causal structure

todo: define finite limits. discuss pullback as fibered product in Set, initial objects, terminal objects todo: define monomorphisms, epimorphisms, equalizers and coequalizers todo: define subobject of X as equivalence class of monomorphisms into X todo: define subobject classifier, show that $1 + 1$ is the subobject classifier for Set

Definition 2.2.

2.2 Topos theory

Definition 2.3. A *topos* \mathcal{E} is a category with all finite limits, equipped with an object Ω , with a function P which assigns to each object Y of \mathcal{E} an object PY of \mathcal{E} , and, for each object X of \mathcal{E} , with two isomorphisms, each natural in Y

$$\text{Sub}_{\mathcal{E}}(X) \cong \text{Hom}_{\mathcal{E}}(X, \Omega), \quad (1)$$

$$\text{Hom}_{\mathcal{E}}(Y \times X, \Omega) \cong \text{Hom}_{\mathcal{E}}(X, PY). \quad (2)$$