Changing minds

Max Collard

Medical Scientist Training Program, Neuroscience Graduate Program
University of California, San Francisco, USA

Abstract

We build worlds inside—by looking at and acting on our environment, and watching as our surroundings look at and act on us. The content of this knowledge within an embodied entity and its dynamics are an important object of study, independent of any underlying material substrate.

Viewing knowledge content as relational at its core, the natural data structure for capturing the world model of an entity C at any moment is a *category* $\mathcal{C}$, whose objects $c : \mathcal{C}$ are the *concepts* that the entity has knowledge of, and whose *morphisms* $f : c \to d$ are all of the ways in which C knows c and d to be related. A functor $F : \mathcal{D} \to \mathcal{C}$ between two such knowledge categories then encapsulates a way in which concepts and relationships in $\mathcal{D}$ trace their origin back to $\mathcal{C}$, in a way that preserves knowledge content, the arrangement of relationships.

Is it possible that investigating well-controlled schemes by which world models—*i.e.*, categories—can be built can give insight into how knowledge-bases grow?

How do I contain such rich knowledge inside?

Neuroscience has historically been focused on a very specific set of experimental questions, aimed at understanding how neurons respond when an organism is presented with exogenous stimuli or engages in some observable activity. However, mental life mostly occurs out of view; and furthermore, because the dynamical structure of a brain is learned nonlinearly over a span of many years, there is no *a priori* reason to believe that the organization of observed activity will provide comprehensible, disentangleable information about the structure of relational knowledge within a brain.

How, for example, would I even know where to *begin* to ask about whether the planarian *C. elegans* can process sequences? Historically, this has been done by copying heuristics of what "looks like" sequences in canonical model systems; and historically, this approach leads to many false negatives, because experimenters seldom know what es ecologically relevant to the organisms they study—after all, the umwelt of a *C. elegans* stretches the naïve imagination.

In the context of category theory—and particularly, of **Poly**—a clearer picture of how to understand a more general framework of cognition and computation emerges, based on understanding the building blocks—like $+$, $\times$, $\otimes$, $\triangleleft$, $[-, -]$, and so on—that can compose to build a large conceptual repertoire.

In the case of *C. elegans* above, a suitably-defined monad arising from an adjunction to **Mon** (perhaps weakened or approximate in a way that quantifies the degree to which biological systems cannot recapitulate the mathematical perfection of "listness"; more on this later), and its corresponding Eilenberg-Moore category, seem to be a more complete picture of not just whether, but *all of* the ways in which *C. elegans* manifests this computational essence.

How is knowledge created?

One process of crucial interest is that of *semiosis*, the process by which new knowledge is created from existing concepts. Starting with a category $\mathcal{C}_0$, we may envision a process through which we sequentially build new categories by extension:

$$\mathcal{C}_0 \xleftarrow{F} \mathcal{C}_1$$

where F is taken from a suitable class of functors that trace the lineage of concepts in $\mathcal{C}_1$ back to $\mathcal{C}_0$. The dynamics of these processes may be interesting.

How does a baby come to understand a pen?

Imagine an entity's dynamics as

$$\phi : \sum_{c:\mathbf{Cat}} y^{F[c]} \to B\, y^{A}$$

where F is a suitable class of category extensions, as above; that is, the entity can be directed to choose an extension of its current knowledge.

Given another set of generalized dynamics,

$$\psi : d \to B'\, y^{A'}$$

is there a choice of ϕ and a way of wiring together the two systems that enables the entity's category $c^{(k)}$ after running for some number of steps k to posess an image of the dynamics d?

How is a concept communicated?

Given two entities' present knowledge-schemas (categories, hence comonoids) c and d, a bicomodule between them (seen as a prafunctor)

$$p : d\mathbf{Set} \to c\mathbf{Set}$$

provides a way of migrating a *generalized* d-concept (a copresheaf $x(-)$, giving the ways that x relates to all concepts in the schema d) to a *generalized* c-concept (the copresheaf $p(x)(-)$, giving the ways that the transmitted data relates to all concepts iin the schema c).

Are there more restrictive or inclusive notions of transmission that are useful? How would c know what $p(x)$ is? Does c having an internal concept of d help?

And a few questions more:

How do I change when you tell me a concept?

How does the dynamics of a shared world state mediate communication?

How do we find the right language together by changing p over time?

How are concepts contained, or communicated, weakly? "Strongly enough"?

What are the semiospheres built from hereditary processes of knowledge extension and communication?

What is it for knowledge to be implemented on a substrate?