Philosophy and Innovation

David Corfield

Philosophy, University of Kent

4 May 2024

Sections

- Category theory and its uses
- Category theory and AI
- Philosophical precursors

Category theory and its uses

Applications (with approximate dates)

- Mathematics (from the 1940s)
- Logic/Foundations (from the 1960s)
- Computer science (from the 1970s)
- Physics (from the 1980s)
- *'Applied Category Theory'* (from the 2010s)

In mathematics

Today category theory is used everywhere in:

- Algebraic topology
- Algebraic geometry
- Langlands program
- ...

In mathematics

Today category theory is used everywhere in:

- Algebraic topology
- Algebraic geometry
- Langlands program
- ...

Its range is extending...

> *Some time ago I expressed skepticism regarding the possibility of recovering hard analysis categorically. I am still not convinced but it seems soft analysis is already there, in this summer's course by Clausen and Scholze. (Michael Harris, personal communication on Condensed Mathematics and Complex Geometry)*

> *We would like to say that our proofs are proofs by "formal nonsense" and in particular analysis-free. (Clausen and Scholze)*

In logic and computer science

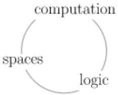
Computational Trinitarianism/Trilogy

1. Idea

A profound cross-disciplinary insight has emerged – starting in the late 1970s, with core refinements in recent years – observing that three superficially different-looking fields of mathematics,

- computation/programming languages
- formal logic/type theory
- ∞-category theory/∞-topos theory (algebraic topology)

are but three different perspectives on a single underlying phenomenon at the foundations of mathematics:



(nLab: computational trilogy)

In physics

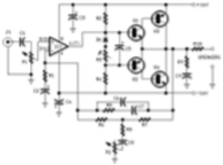
- Topological quantum field theory in 1980s
- String diagrams as morphisms in monoidal categories for quantum mechanics, 2000s
- Higher gauge theory, 2000s
- Modal and linear homotopy type theory in Physics, The Quantum Monadology, now

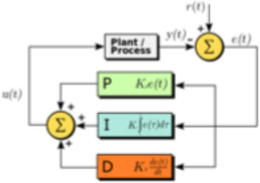
Applied Category Theory today

We see people applying category theory for:

> *Causality, probabilistic reasoning, statistics, learning theory, deep neural networks, dynamical systems, information theory, database theory, natural language processing, cognition, consciousness, active inference, systems biology, genomics, epidemiology, chemical reaction networks, neuroscience, complex networks, game theory, robotics, quantum computing,...*

Systems diagrams





A key idea here is Compositionality.

- Plug together systems in parallel.
- Plug together systems in series.
- Plug one system inside another.

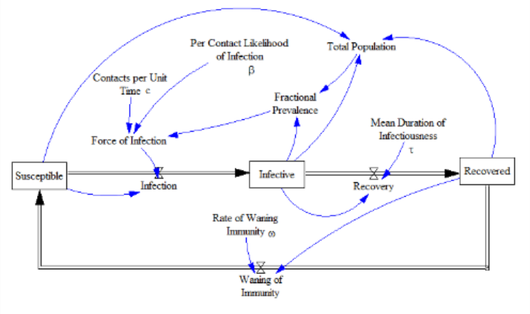
A key idea here is Compositionality.

- Plug together systems in parallel.
- Plug together systems in series.
- Plug one system inside another.

This allows the representation of: open graphs, open Petri nets, open chemical reaction with rates, open electrical circuits, open Markov processes, open dynamical systems,... (c.f. Baez)

A cutting edge development along this line is the applied category-theoretic treatment of Stock and Flow diagrams used to model disease spread and control.

- John C. Baez, Xiaoyan Li, Sophie Libkind, Nathaniel D. Osgood and Eric Redekopp, **A categorical framework for modeling with stock and flow diagrams**, to appear in *Mathematics for Public Health*, Springer.



Category theory and AI

Knowledge representation:

> *The ethos of this research program is that category theory can serve as a general purpose modeling language for science and engineering. Having internalized this perspective, it is but a short step to contemplate a general-purpose knowledge representation system based on category theory.*
>
> *Our philosophy is that category theory is a universal modeling language enabling a more expansive understanding of knowledge representation. (Evan Patterson, Knowledge Representation in Bicategories of Relations, p. 49)*

Category theory and AI

Category theory is well-suited to inferential and learning processes which are

- Compositional
- Structural
- Presentable diagrammatically
- Hybrid

Symbolica – $31 million

Approach

All current state of the art large language models such as ChatGPT, Claude, and Gemini, are based on the same core architecture. As a result, they all suffer from the same limitations.

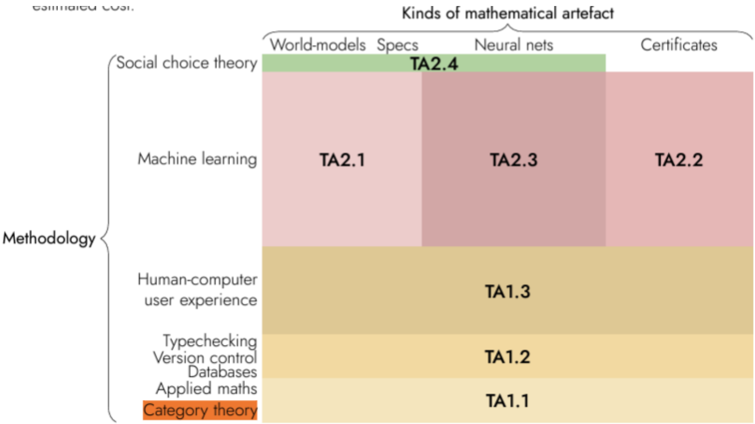
Extant models are expensive to train, complex to deploy, difficult to validate, and infamously prone to hallucination. Symbolica is redesigning how machines learn from the ground up.

We use the powerfully expressive language of category theory to develop models capable of learning algebraic structure. This enables our models to have a robust and structured model of the world; one that is explainable and verifiable.

It's time for machines, like humans, to think symbolically.

ARIA – safeguarded AI - £59 million

TA1.1 **Theory** shall research and construct computationally practical mathematical representations and formal semantics for world-models, specifications, proofs, neural systems, and "version control" (incremental updates or patches) thereof.



Quantinuum – worth lots

www.quantinuum.com/news/quantinuum-is-developing-new-frameworks-for-artificial-intelligence

Quantinuum is developing new frameworks for artificial intelligence

We do this in part by using a type of math called "**category** theory" that has been used in everything from classical computer programming to neuroscience....

View

Philosophical precursors: Peirce

Charles Peirce developed his *existential graphs* over many years in three systems: alpha, beta, gamma. The beta system corresponds to first-order logic.

Philosophical precursors: Peirce

Charles Peirce developed his *existential graphs* over many years in three systems: alpha, beta, gamma. The beta system corresponds to first-order logic.

A recent response:

arXiv > cs > arXiv:2401.07055

Search...

Help |

Computer Science > Logic in Computer Science

[Submitted on 13 Jan 2024]

Diagrammatic Algebra of First Order Logic

Filippo Bonchi, Alessandro Di Giorgio, Nathan Haydon, Pawel Sobocinski

We introduce the calculus of neo-Peircean relations, a string diagrammatic extension of the calculus of binary relations that has the same expressivity as first order logic and comes with a complete axiomatisation. The axioms are obtained by combining two well known categorical structures: cartesian and linear bicategories.

Subjects: **Logic in Computer Science (cs.LO)**; Category Theory (math.CT)

Cite as: arXiv:2401.07055 **[cs.LO]**

(or arXiv:2401.07055v1 **[cs.LO]** for this version)

https://doi.org/10.48550/arXiv.2401.07055

Calculus of relations

With the rise of first order logic, Peirce's calculus was forgotten until Tarski, who in [80] recognised its algebraic flavour. In the introduction to [81], written shortly before his death, Tarski put much emphasis on two key features of CR: (a) its lack of quantifiers and (b) its sole deduction rule of substituting equals by equals. The calculus, however, comes with two great shortcomings: (c) it is strictly less expressive than FOL and (d) it is *not* axiomatisable.

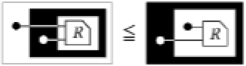
Despite these limitations, CR had —and continues to have— a great impact in computer science, e.g., in the theory of databases [20] and in the semantics of programming languages [2, 38, 45, 47, 74]. Indeed, the lack of quantifiers avoids the usual burden of bindings, scopes and capture-avoid substitutions (see [25, 30, 33, 40, 68, 70] for some theories developed to address specifically the issue of bindings). This feature, together with purely equational proofs, makes CR particularly suitable for proof assistants [43, 71, 72].

In this paper, we introduce the calculus of *neo-Peircean relations*, a string diagrammatic account of FOL that has several key features:

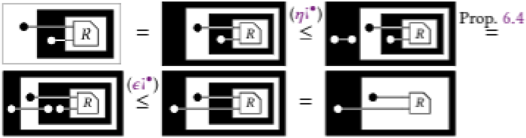
(1) Its diagrammatic syntax is closely related to Peirce's EGs, but it can also be given through a context free grammar equipped with an elementary type system;
(2) It is quantifier-free and, differently than FOL, its compositional semantics can be given by few simple rules: see (8);
(3) Terms and predicates are not treated as separate syntactic and semantic entities;
(4) Its sole deduction rule is substituting equals by equals, like CR, but differently, it features a complete axiomatisation;
(5) The axioms are those of well-known algebraic structures, also occurring in different fields such as linear algebra [11] or quantum foundations [21];
(6) It allows for compositional encodings of FOL, CR and PFL;
(7) String diagrams disambiguate interesting corner cases where traditional FOL encounters difficulties. One perk is that we allow empty models —forbidden in classical treatments— leading to (slightly) more general Gödel completeness;
(8) The corner case of empty models coincides with *propositional* models and in that case our axiomatisation simplifies to the deep inference Calculus of Structures [15, 34].

By returning to the algebraic roots of logic we preserve CR's benefits (a) and (b) while overcoming its limitations (c) and (d).

Let R be a symbol with arity 2 and coarity 0. The two diagrams on the right correspond to FOL



formulas $\exists x. \forall y. R(x,y)$ and $\forall y. \exists x. R(x,y)$: see § 9 for a dictionary of translating between FOL and diagrams. It is well-known that $\exists x. \forall y. R(x,y) \models \forall y. \exists x. R(x,y)$, i.e. in any model, if the first formula evaluates to true then so does the second. Within our calculus, this statement is expressed as the above inequality. This can be proved by mean of the axiomatisation we introduce in this work:



(1)

Peirce's existential graphs were devised as central to his system of philosophy

> *The System of Existential Graphs may be characterized with great truth as presenting before our eyes a moving picture of thought. Provided this characterization be taken not as a flatly literal statement, but as a simile, it will, I venture to predict, surprise you to find what a strain of detailed comparison it will bear without snapping. A picture is visual representation of the relations between the parts of its object; a vivid and highly informative representation, rewarding somewhat close examination. (MS, 1908)*

Should we look to understand it in its context?

Philosophical precursors: Geometric deep learning and Cassirer

Cassirer's 1944 paper The concept of group and the theory of perception (translated from an earlier article in French) brings together Klein's Erlangen Program with the findings of the Gestalt school of psychology.

Cassirer

If perception is to be compared to an apparatus at all, the latter must be such as to be capable of "grasping intrinsic necessities." Such intrinsic necessities are encountered everywhere. It is only with reference to such "intrinsic necessity" that the "transformation" to which we subject a given form is well defined, inasmuch as the transformation is not arbitrary and executed at random but proceeds in accordance with some rule that can be formulated in general terms. (p. 26) ("grasping intrinsic necessities" is due to Max Wertheimer)

By their reference to such "good points", the particular impressions receive a new kind of determination. They lose, so to speak, their "atomicity", their uniqueness as mere particular items; they unite into groups and totals. (p. 28)

The "images" that we receive from objects, the "impressions" which sensationalism tried to reduce perception to, exhibit no such unity. Each and every one of these images possesses a peculiarity of its own; they are and remain discrete as far as their contents are concerned. But the analysis of perception discloses a formal factor which supersedes this particularity and disparity. Perception unifies and, as it were, concentrates the manifolds of particular images with which we are supplied at every moment...Each invariant of perception is ... a scheme toward which the particular sense-experiences are orientated and with reference to which they are interpreted. (p. 32)

Erlangen Program

Felix Klein (1872) brings order to the panoply of geometries, organising classification around group actions.

(Higher Klein geometry finds its use in physics.)

Geometric deep learning

In this text, we make a modest attempt to apply the Erlangen Programme mindset to the domain of deep learning, with the ultimate goal of obtaining a systematisation of this field and 'connecting the dots'. We call this geometrisation attempt 'Geometric Deep Learning', and true to the spirit of Felix Klein, propose to derive different inductive biases and network architectures implementing them from first principles of symmetry and invariance.

(Geometric Deep Learning: Grids, Groups, Graphs, Geodesics, and Gauges, Michael M. Bronstein, Joan Bruna, Taco Cohen, Petar Veličković)

This is generalized in an article, Categorical Deep Learning: An Algebraic Theory of Architectures timed to launch with Symbolica.

> *We present our position on the elusive quest for a general-purpose framework for specifying and studying deep learning architectures. Our opinion is that the key attempts made so far lack a coherent bridge between specifying constraints which models must satisfy and specifying their implementations. Focusing on building a such a bridge, we propose to apply category theory – precisely, the universal algebra of monads valued in a 2-category of parametric maps – as a single theory elegantly subsuming both of these flavours of neural network design.*

Extended Erlangen

Category theory can be “regarded as a continuation of the Klein Erlangen Programme, in the sense that a geometrical space with its group of transformations is generalized to a category with its algebra of mappings” (Eilenberg and Mac Lane 1945)

Here the writings of the precursor Cassirer on invariance have not been tapped.

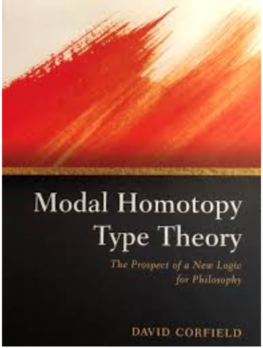
Philosophical precursors: Quantum AI

Quantum cognition meets category theory from Quantinuum:

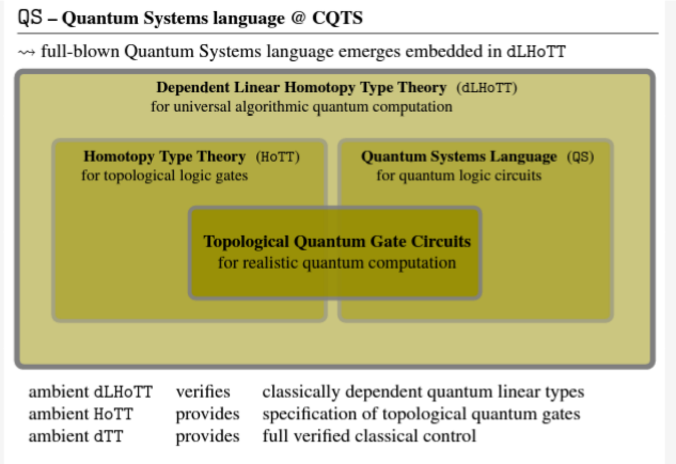
- Sean Tull, Razin A. Shaikh, Sara Sabrina Zemljič and Stephen Clark, From Conceptual Spaces to Quantum Concepts: Formalising and Learning Structured Conceptual Models

In this article we present a new modelling framework for concepts based on the mathematical formalism used in quantum theory, and demonstrate how the conceptual representations can be learned automatically from data, using both classical and quantum-inspired models. A contribution of the work is a thorough category-theoretic formalisation of our framework, following Bolt et al. (2019) and Tull (2021). Formalisation of conceptual models is not new (Ganter & Obiedkov, 2016), but we claim that the use of category theory (Fong, 2019), and in particular the use of string diagrams to describe quantum processes (Coecke & Kissinger, 2017), helps elucidate some of the most important features of our approach to concept modelling. This aspect of our work also fits with the recent push to introduce category theory into machine learning and AI more broadly. The motivation is to make deep learning less ad-hoc and less driven by heuristics, by viewing deep learning models through the compositional lens of category theory (Shiebler et al., 2021).

Naturally, I'd be keen for a variety of modal homotopy type to feature:



Exciting work from the physics/computing end: Linear HoTT and QS



QS claimed there to improve on existing quantum programming languages: Quipper, QWIRE, etc.

Philosophers of physics should take note:

Besides these technical properties, the logical language QS is curiously satisfying on quantum-philosophical grounds: For example, the internal language-construct in QS for quantum measurement via the modal logic of necessity is *verbatim* the same as for classical measurement, only now applied to (dependent) linear types where it happens to *imply* the collapse of the wavefunction in the categorical semantics. But in the internal logic this effect is just standard conditioning of expectations. In this sense the notorious "measurement problem" of quantum physics disappears when we speak proper QS. (This is analogous to what happens internal to proper quantum probability theory, see there.) Moreover, the *deferred measurement principle* verified in QS implies that even this collapse of the wavefunction in subsystems may be arbitrarily postponed, by observers who have access to the system at large (the "bath").

Philosophers of physics should take note:

Besides these technical properties, the logical language QS is curiously satisfying on quantum-philosophical grounds: For example, the internal language-construct in QS for quantum measurement via the modal logic of necessity is *verbatim* the same as for classical measurement, only now applied to (dependent) linear types where it happens to *imply* the collapse of the wavefunction in the categorical semantics. But in the internal logic this effect is just standard conditioning of expectations. In this sense the notorious "measurement problem" of quantum physics disappears when we speak proper QS. (This is analogous to what happens internal to proper quantum probability theory, see there.) Moreover, the *deferred measurement principle* verified in QS implies that even this collapse of the wavefunction in subsystems may be arbitrarily postponed, by observers who have access to the system at large (the "bath").

Very much computational trinitarian territory.

If Linear HoTT can be used in Quantum AI, we have a great many philosophical precursors to revisit, including Hegel, Kant, Husserl,...

To conclude

Three examples out of many possibilities.

- Unless these are merely isolated philosophical ideas which fortuitously happen to play a heuristic role in current data science and AI, there should be interest in understanding their settings.
- Philosophers can then contribute to AI by explaining the contexts of such ideas.
- But category theory relates and unifies different formalisms, so that perhaps we could use category-theoretic AI to explore connections within philosophy itself.