Artificial Intelligence and Category Theory

Nick Rossiter
Computer Science
University of Northumbria
nick.rossiter1@btinternet.com
https://nickrossiter.org.uk/process/

Presented at

- ANPA 46
  - Alternative Natural Philosophy Association
- 12th August 2025
- Department of Materials
- University of Oxford

Should be a Connection

- AI (Artificial Intelligence)
  - Mainly based on neural nets
  - Connections
  - Multi-level directed graph
  - Based on models
- CT (Category Theory)
  - Based on the arrow
  - Multi-level
- ANNs (Artificial Neural Net models) are complex, a challenge to any theory

Why do we need a theory?

- Aids teaching and understanding
- Adds confidence in results through proofs
- Can emphasise commonality with other approaches
- Can aid extensions
- Can emphasise assumptions
- Can ease implementation
- But theory has to be appropriate for the target

Artificial Neural Nets (ANN)

- Fundamental object is the artificial neuron
  - A receptor or a transmitter of signals, often an algorithm
- A signal between neurons (synapses) can be weighted
- Neurons are assembled in layers
    - The opening layer receives inputs
    - The closing layer provides an output
    - Middle layers are hidden
- A neuron can have many inputs
  - But usually just one output (which can branch into multiple neurons)

Theories for ANNs

- Inspired by biology
  - Nodes with connections
  - Connections have weights
  - Weights are adjusted in training
- In training
  - Universal function approximator
    - Approximate any continuous function to an arbitrary degree of accuracy
    - Learn and represent a vast range of functions
- Deep learning – multiple hidden layers
- Output of each neuron computed by some non-linear function of the totality of its inputs
- Output is a real number

Suggested Formalisms

- Machine Learning
  - Adaptive learning
- Statistical
  - Probability, statistics, Bayesian with weights
  - Linear and non-linear algebra
  - Dynamical systems theory

Category Theory for ANNs

- In representing information and other systems, used
  - Cartesian products for relating data
    - Relates whole entities to one another
  - Monads for transactions
    - Relates functors as process via adjunctions (two-way mappings)
- For ANNs, need:
  - More flexible ways of linking and processing data
    - Require tensor products
    - With monoids for basic unit

Monad vs Monoid

- A monad is a monoid in the category of endofunctors:
  - An endofunctor is a functor from a category to itself
  - Often the endofunctor is actually a pair of adjoint functors
    - G –! F (F: C → D; G: D → C; G o F: C → D → C)
    - As in much of our work on information systems and music
- So a monad is a monoid in a particular setting
- In category theory a monad is an important level of abstraction

Monoid Definition (set)

- In algebra a monoid is a set
  - With an associative binary operation
  - And an identity (neutral) element (leaves an object unchanged)
- Example 1
  - $M \times M \rightarrow M$
  - $1 \in M$
- Example 2
  - $\mathrm{M} + \mathrm{M} \rightarrow \mathrm{M}$
  - $0 \in M$

Monoid Definition (category)

- The monoid is an object
  - a hom-set (identity arrows as in say a set)
  - composition morphism (x, +, ...)
  - a unit (neutral) element
- Example
  - $(C, +, I)$
- Enriched monoid category (tensor product operation)
  - $(C, \otimes, I)$

Tensor Products

- The operation ⊗
  - e.g. $A \otimes B$
  - If A and B are vectors, performs pairwise multiplication of each component of each vector
  - The Cartesian product would pair each whole vector with another whole vector
- Tensor products are usually formalised in
  - A monoidal category
    - which can also handle Cartesian products
- Tensor products give n-tuples: can have many arguments

Monoidal Category 1 (Basics)

- A monoidal category is a category $\mathcal{C}$ equipped with
- 1) a functor
  - $\otimes : \mathcal{C} \times \mathcal{C} \longrightarrow \mathcal{C}$
  - out of the product category of $\mathcal{C}$ with itself, called the tensor product,
- 2) an object
  - $1 \in \mathcal{C}$
- called the unit object or tensor unit

Monoidal Category 2 (natural)

- 3) a natural isomorphism
  - $a:((-)\otimes(-))\otimes(-) \overset{\simeq}{\longrightarrow} (-)\otimes((-)\otimes(-))$
  - with components of the form
    - $a_{x,y,z}:(x\otimes y)\otimes z \to x\otimes(y\otimes z)$
  - called the associator,
- 4) a natural isomorphism
  - $\lambda:(1\otimes(-)) \overset{\simeq}{\longrightarrow} (-)$
  - with components of the form
    - $\lambda_x:1\otimes x \to x$
  - called the left unitor

Monoidal Category 3 (natural)

- 5) a natural isomorphism
  - $\rho:(-)\otimes 1 \xrightarrow{\simeq} (-)$
  - with components of the form
    - $\rho_x : x \otimes 1 \to x$
- called the right unitor

Monoidal Category 4 (identities, Strict)

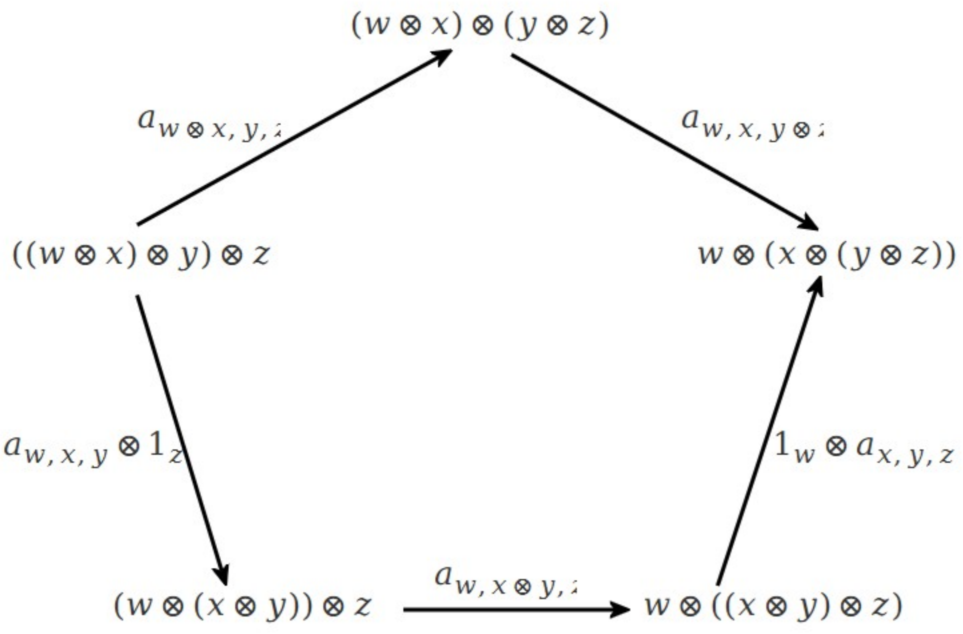
Such that the following two kinds of diagrams commute for all objects:

1. triangle identity

$$
\begin{array}{ccc}
(x \otimes 1) \otimes y & \xrightarrow{a_{x,1,y}} & x \otimes (1 \otimes y) \\
\rho_x \otimes 1_y \searrow & & \swarrow 1_x \otimes \lambda_y \\
 & x \otimes y &
\end{array}
$$

Monoidal Category 5 - Pentagon associativity

2. pentagon identity



Further facilities beyond Tensor Products

- Non-linear is not achievable within basic categories
- Monoidal categories are fundamental building blocks for higher categories, which are used to model more complex mathematical structures.
  - Monoidal categories involve an extra layer of arrows for internal structure so arrow structure is more layered
- We need to move to category of category
  - Higher-order category theory
- In particular 2-category

n-Cells

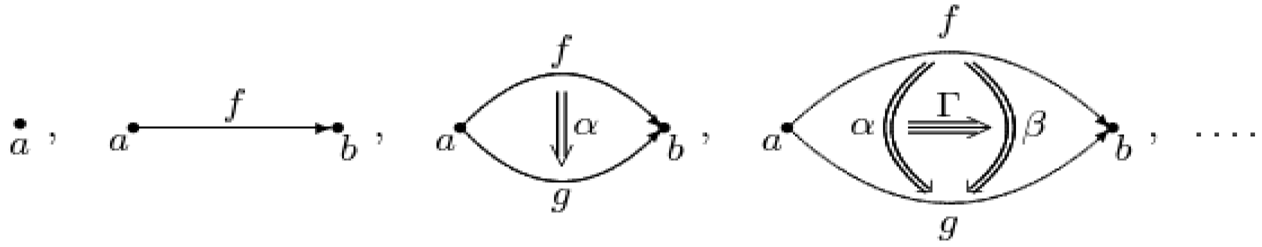
- Category of category or double category
- We have met the concept before as adjunction
- It is also now known as a lens category
- It is a type of n-category, containing the basic cells:
  - 0-cell is objects (identity arrows)
  - 1-cell is arrows between objects (category)
  - 2-cell is arrows between arrows (natural transformation, adjunction)
  - 3-cell is (arrows between arrows) between (arrows between arrows)
    - Or adjunction arrows mapped onto adjunction arrows (modification)
    - We (Sisiaridis) used 3-cells for security in information systems
  - 4-cell and beyond have been used in knot theory

n-Category

- No upper limit to n, which is the highest ranking cell in the category
- A category with just 0-cells (objects) is a set, a discrete category
- A category with just 1-cells (arrows between objects) is an ordinary category
- A category holding up to 2-cells (arrows between arrows) is a 2-category
- A category holding up to 3-cells -- (arrows between arrows) between (arrows between arrows) -- is a 3-category
- And so on

Lens Diagram

Cell types



0-cell 1-cell 2-cell 3-cell

Isomorphism and Associativity

- In 2-category (and above)
  - unique up to natural Isomorphism (composition) and associativity are no longer in general enforced
- Except in strong (strict) 2-category
  - Composition and associativity are strictly applied throughout
  - These do not occur often in nature
- In weak 2-category (multicategory)
  - Composition and associativity are restricted up to a particular level of isomorphism (1-isomorphisms for instance)

Multicategory

- In a multicategory the objects of an arrow
  - are not single objects but a finite sequence of them
- So an object can be a vector space
- Respecting this structure requires a tensor product for multiplication

Operads

- An operad is an operation taking a multicategory as input and producing a single output
  - So it can convert a vector space to a single output
  - Or a tensor product of a vector space to a single output

Lax Monoidal Categories

- A lax monoidal category or skew monoidal category is a monoidal category in which the associator- and unitor-transformations are not required to be invertible, i.e. are not required to be natural isomorphisms (as they are for strict monoidal categories).
- This allows more flexibility in the mappings beyond isomorphisms

Lax Monoidal Category 2

- There are still some constraints
  - Typically there is a bias in one direction (left or right)
- A left-biased lax monoidal category has a binary tensor product
  - A $\otimes$ B, a unit I and constraints:

$$a \otimes (b \otimes c) \to (a \otimes b) \otimes c \qquad I \otimes a \to a \qquad a \otimes I \to a$$

  - But no inverse for the associator

Linear-Non-linear adjunction

- The superposition principle, also known as superposition property, states that, for all linear systems, the net response caused by two or more stimuli is the sum of the responses that would have been caused by each stimulus individually.
- Does category theory have a non-linear construction?
- For example, a composition of adjoints:
  - One tensor
  - One Cartesian
- Is termed linear-non-linear (LNL) adjunction, as is
  - Any well-formed category including topos
  - LNL is property of intuitionistic logic
    - No closed world assumption
- Hyland 2020 (Cambridge), Shulman 2023 (San Diego)

2-Category: Linear-non-linear Adjunction

A linear-non-linear adjunction is for example an adjunction between a Cartesian (M) category and a non-Cartesian monoidal category (L)

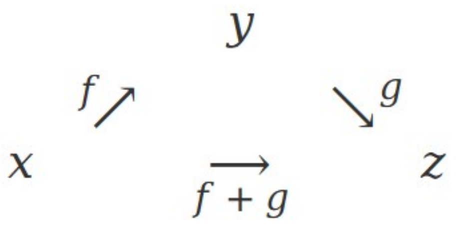
$$! \circlearrowright (\mathbf{L}, \otimes, 1) \underset{M}{\overset{L}{\underset{\longrightarrow}{\overset{\longleftarrow}{\perp}}}} (\mathbf{M}, \times, \top)$$

Weights

- Where are the weights?
  - Lawvere was the first to investigate this:
    - can select real number to represent the distance between two points, with curvature permitted.
    - Topology context so non-Euclidean
- Lawvere metric space
  - a metric space is defined to be a set X equipped with a distance function d: $X \times X \rightarrow [0, \infty]$
  - Triangle inequality
- Now viewed more as an example of an enriched category

Triangle Inequality

- The triangle inequality, in its most basic form, states that for any triangle, the sum of the lengths of any two sides must be greater than or equal to the length of the third side.
  - The distance d(x,z) between any two points x,z is no larger than the sum of the distances d(x,y) and d(y,z) via any third point y



Weighted Categories

- A category where each morphism has a number (for example a ‘cost’ or ‘length’), and composition comes with a triangle inequality
- A weighted category is a category, where to each morphism f: A → B
  - Is assigned a non-negative real number (up to ∞)
    - the weight w(f)
  - Such that
    - $w(ID_X) = 0$
    - w(g o f) ≤ (w)g + w(f) (triangle inequality)

Enriched Weighted Categories

- Define a weighted set to be a set X equipped with a function:
  - w: $X \rightarrow [0, \infty]$
- Take a (short) function f: $(X, w_X) \rightarrow (Y, w_Y)$ with rule:
  - $w_Y(f_X) \leq (w_X)X$
- A closed monoidal category is constructed by taking the tensor product of weighted sets in the cartesian product of the sets (x,y):
  - For all (x,y) in (X,Y): $w_{(x \otimes y)}((x,y)) = w_X(x) + w_Y(y)$
- A symmetric closed monoidal category has LNL

Relation between a topos and a closed monoidal category

- In earlier work, we used as our data structure, the topos or at least a cartesian closed category
- A closed monoidal category is equivalent to a cartesian closed category
  - When the object is a discrete set with no internal structure
  - The product is cartesian rather than tensor
- They are not equivalent if say the object is a vector space
  - The monoidal category involves a tensor product

Limits in Categories

- Cartesian closed and topos are based on limits and colimits for universal constructions
- Do we have limits for monoidal categories?
- Yes, limits can be defined, which respect the monoidal structure (tensor product and identity).
- The limit functor should be at least a lax monoidal functor, meaning it preserves the monoidal structure up to specified isomorphisms.

Applications of Monoidal Categories 1

- Neural nets
  - Monoidal Coalgebraic Metrics, Filippo Bonchi, University of Pisa,
    - Filippo and team intend to establish a robust mathematical framework that extends beyond the metrics expressible in quantitative algebraic theories and coalgebras over metric spaces. **By shifting from Cartesian to a monoidal setting**
    - johncarlosbaez.wordpress.com/2025/02/08/category-theorists-in-ai/
  - Filippo Bonchi
    - The theory of traces for systems with nondeterminism, probability, and termination
    - Logical Methods in Computer Science 18 (2) 1-21 (2022)
    - Most of our results are based on the exciting interplay between monads and their presentations via algebraic theories.
    - lmcs.episciences.org/9713/pd

Applications of Monoidal Categories 2

- Back propagation (training, neural nets)
- B. Fong, D. Spivak, and R. Tuyeras, “Backprop as functor: A compositional perspective on supervised learning,” pp. 1–13, 2019.
    - defines a monoidal functor from a category of parametrised functions to this category of update rules
        - arXiv:1711.10455 [2019]

Applications of Monoidal Categories 3

- Matthew Varughese, Categorical Quantum Computing with Finite Fields, Supervisor: Dr Jamie Vicary, Oxford University MSc (2009)
  - 1) The link between classical structures in Rel the category of finite relations and single valued classical structures in FdVectF2 ...,
  - 2) Monoidal structures in FdVectFp (dimension n) that are isomorphic to the finite field Fpn .
  - The main result here is that such monoidal structures always exist and each admits a special Frobenius algebra. This work also provides a background to the use of category theory in quantum computing. We cover the use of symmetric monoidal categories …
  - www.cs.ox.ac.uk/people/bob.coecke/Varughese.pdf

Applications of Monoidal Categories 4

- Topos UK, n-categories in AI
- The Topos Institute is a math research institute in Berkeley with a focus on category theory. Three young category theorists there—Sophie Libkind, David Jaz Myers and Owen Lynch—wrote a proposal called “Double categorical systems theory for safeguarded AI”, which got funded by ARIA. So now they are moving to Oxford, where they will be working with Tim Hosgood, José Siqueira, Xiaoyan Li and maybe others at a second branch of the Topos Institute, called Topos UK.
- Double category is n-category with n=2

Applications of Monoidal Categories 5

- Topos Institute Oxford, maths, relations
- José Siqueira: Double functorial representation of indexed monoidal structures (2025)
  - Adding a monoidal structure to the fibers (and the double pseudofunctor) allows us to capture both the Beck-Chevalley and Frobenius conditions. We will concern ourselves with the monoidal analogues of regular hyperdoctrines and similar structures (where the objects of predicates are not necessarily posets, but rather live in some 2-category), and show how they are equivalent to (lax symmetric monoidal) double pseudofunctors between spans and quintet double categories.
  - www.youtube.com/watch?v=glQUcRImOkw

Applications of Monoidal Categories 6

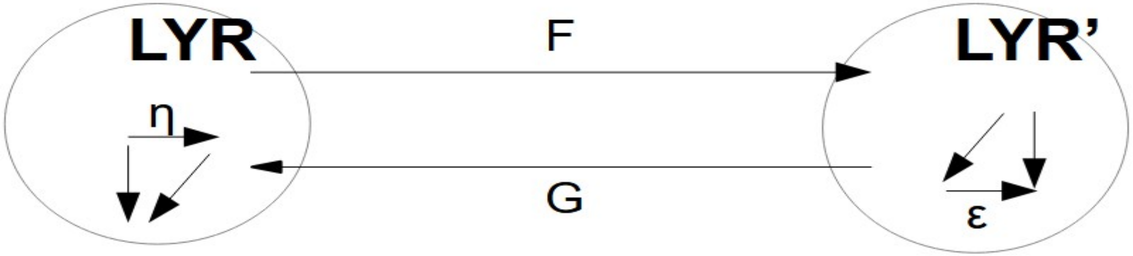
- Maths Stack Exchange, knots
  - A categorical approach to knot theory?
  - Khovanov homology is a "categorification" of the Jones polynomial. Polynomials are replaced with modules (chain complexes in particular), and certain kinds of maps between knots (knot concordance) are sent to homomorphisms of chain complexes. There is a paper by Lauda and Pfeiffer that seems to give a construction of Khovanov homology for tangles, but I'm not sure if this gives a functorial construction in the above sense (I haven't really read it).
  - Another book(s) that might be interesting to you are ones by Kauffman. He likes to work with knots as formal combinatorial objects, and that lends itself well to categorical treatment.
  - math.stackexchange.com/questions/4206898/a-categorical-approach-to-knot-theory
  - Khovanov, Mikhail (2000), "A categorification of the Jones polynomial", Duke Mathematical Journal, 101 (3): 359–426, arXiv:math.QA/9908171

Relation to our Earlier Work

- We used a monad to represent transactional process on a cartesian closed structure with certain ‘nice’ properties such as limits and a subobject classifier
  - A topos
- The monad represented activity transactionally through an adjointness between current state and next state
- The topos represented data structure

One 'Cycle' for adjointness F -| G

- One ‘cycle’ for GF
  - Assessing unit η in LYR and counit ε in LYR' to ensure overall consistency

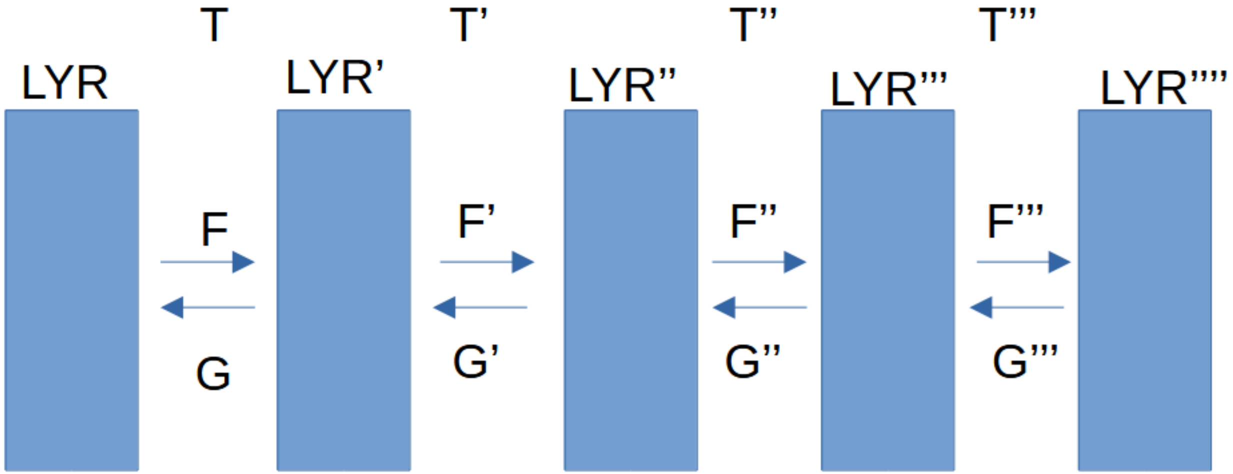


$\eta: 1_{LYR} \rightarrow GF(LYR)$ $\varepsilon: FG(LYR') \rightarrow 1_{LYR'}$

Neural Net as a Transactional Monad

- Each layer is a closed monoidal category
  - With tensor product, weights.
  - Within a 2-category, with lax symmetry,
    - relaxing isomorphism and associativity requirements
    - giving non-linearity
- Between each layer is a
  - Free functor, goal-directed
  - Underlying functor, enforcing back propagation (training)
- Such a construction is a 2-monad
  - Provided some transactional conditions are satisfied
  - Which can be composed with some housekeeping conditions (distributive e.g. Kleisii-lift in Cartesian)

Composition of 2-Monads



T''' o T'' o T' o T   (T= G o F)

Each layer LYR is a 2-category: arrows between arrows

Each 2-monad is a 3-category: (arrows between arrows) between (arrows between arrows)

2-Monads

- Exist within a 3-category
- May be strict, fully weak, lax, pseudo, colax
- Pseudo preserves the structure up to a specified isomorphism
- So more scope for flexibility in strictness of associativity, isomorphism rules

To conclude

- Category theory has developed from its simple cartesian basis with natural transformations and adjointness
- Monoidal categories provide tensor products
- n-categories provide controlled relaxation of rules
- 2-monads provide composition of adjoints between 2-categories, within a 3-category
- Such 3-categories begin to match the complexity of neural net systems

3-category

With 3-cells