

TAI-DANAE BRADLEY SANDBOXAQ / THE MASTER'S UNIVERSITY DEC. 12, 2022





LARGE LANGUAGE MODELS

CHATGPT





Does psi satisfy the area law for entanglement?

Assume it is the correct and complete description of the quantum state I was asking about.

If the quantum state psi that you are asking about is the normalized superposition of \square \square all even parity bit strings of length 4, then it would be written as follows:

Copy code
psi = (1/sqrt(8)) * (|0000> + |0010> + |0100> + |0110> + |1000>
+ |1010> + |1100> + |1110>)

With this information, we can determine that if one subsystem of psi consists of the first two qubits and another subsystem consists of the remaining two qubits, then psi does not satisfy the area law for entanglement.

To see this, we can calculate the entanglement entropy of each subsystem and compare it to the area of the boundary between the subsystems. The entanglement entropy of a subsystem is a measure of the amount of entanglement between that subsystem and the rest of the system, and it is given by the formula S = -tr(rho * log(rho)), where rho is the density matrix of the subsystem and tr is the trace operator.

In this case, the density matrix of the first two qubits would be given by the following:

Copy code rho1 = (1/8) * (|00><00| + |01><01| + |10><10| + |11><11|)

And the density matrix of the remaining two qubits would be given by:



The entanglement entropy of each subsystem can then be calculated by plugging the density matrices into the formula above. This yields S1 = S2 = 1.0 bits.

The area of the boundary between the subsystems can be calculated by simply counting the number of qubits that are shared between the subsystems. In this case, there are two shared qubits, so the area of the boundary is 2 qubits.

Since the entanglement entropy of each subsystem (S1 = S2 = 1.0 bits) is greater than the area of the boundary between them (2 qubits), we can conclude that psi does not satisfy the area law for entanglement.

It's worth noting that this is just one possible way in which the subsystems could be divided, and the area law may hold for different subsystems or different ways of dividing the system. To determine whether psi satisfies the area law in general, we would need to consider all possible subsystems and boundary areas.

Ľ



What mathematical framework describes the passage

probability distributions on text continuations

semantic information

?



- **0. motivation**
 - an analogy
- 1. a category of language
 - some advantages
- 2. an enriched category of language
 - what is enriched category theory?
 - some advantages

La Matematica (2022) 1:551–580 https://doi.org/10.1007/s44007-022-00021-2



ORIGINAL RESEARCH ARTICLE



An Enriched Category Theory of Language: From Syntax to Semantics

Tai-Danae Bradley¹ · John Terilla² · Yiannis Vlassopoulos³

Received: 24 June 2021 / Revised: 7 February 2022 / Accepted: 7 February 2022 / Published online: 11 March 2022 © The Author(s), under exclusive licence to Springer Science+Business Media LLC, part of Springer Nature 2022

Abstract

State of the art language models return a natural language text continuation from any piece of input text. This ability to generate coherent text extensions implies significant sophistication, including a knowledge of grammar and semantics. In this paper, we propose a mathematical framework for passing from probability distributions on extensions of given texts, such as the ones learned by today's large language models, to an enriched category containing semantic information. Roughly speaking, we model probability distributions on texts as a category enriched over the unit interval. Objects of this category are expressions in language, and hom objects are conditional probabilities that one expression is an extension of another. This category is syntactical—it describes what goes with what. Then, via the Yoneda embedding, we pass to the enriched category of unit interval-valued copresheaves on this syntactical category. This category of enriched copresheaves is semantic—it is where we find meaning, logical operations such as entailment, and the building blocks for more elaborate semantic concepts.

Keywords Category theory \cdot Yoneda embedding \cdot Compositionality \cdot Natural language \cdot Probability \cdot Logic

 $\textbf{Mathematics Subject Classification} \ 18D20 \cdot 18A25 \cdot 18A30 \cdot 18A35 \cdot 18B25 \cdot 18B35$

\bowtie	Tai-Danae Bradley tai.danae@math3ma.com
	John Terilla jterilla@gc.cuny.edu
	Yiannis Vlassopoulos yiannis@tunnel.tech

- ¹ X, the Moonshot Factory and Sandbox@Alphabet, Mountain View, CA, USA
- ² The City University of New York and Tunnel, New York, NY, USA
- ³ Tunnel, New York, NY, USA

🖄 Springer

LINEAR ALGEBRA VS. CATEGORY THEORY





If "blah" has nice structure, then functions from a set into "blah" have nice structure, too.

 $f: X \to \mathbb{R}$



If "blah" has nice structure, then functions from a set into "blah" have nice structure, too.

Vector spaces are an example. A similar story holds in category theory....





CATEGORY



FUNCTION

$X \rightarrow \mathbb{R}$



$C \rightarrow Set$



INTERSECTIONS

PRODUCTS

DIRECT SUMS

MEETS

GREATEST COMMON DIVISORS

KERNELS





UNIONS

COPRODUCTS

DIRECT SUMS

JOINS

LEAST COMMON MULTIPLES

COKERNELS



VECTOR SPACE

 \mathbb{R}^X

(CO)PRESHEAF CATEGORY

Set

ONE-HOT ENCODINGS

 $x_i \in X$

 $\langle x_i, - \rangle \colon X \to \mathbb{R}$

 $\langle x_i, x_j \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$

 $x_2 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$

REPRESENTABLE FUNCTORS

$c \in ob(C)$

$C(c, -): C \rightarrow Set$

 $C(c,d) = \{\text{morphisms } c \to d\}$

:= hom(c, d)



AS A CATEGORY

Consider all strings from some finite set of atomic symbols. (Think: free monoid.)

Substring containment defines a preorder on this set.

 $x \leq y$

This gives rise to a category L with a morphism $x \rightarrow y$ whenever x is a substring of y.

Nice, but limited.



YONEDA PERSPECTIVE

Consider functors $L \rightarrow Set.$

Representable functors are akin to a first approximation to "meaning."

$$L($$
blue, $-): L \rightarrow Set$

$$L(\mathbf{blue}, x) = \begin{cases} * & \text{if blue} \le x \\ \emptyset & \text{otherwise} \end{cases}$$

Think: Yoneda lemma / John Firth.



VONEDA PERSPECTIVE

 \emptyset

*

*

Ø

*

Ø

•

 $\mathsf{L}(\mathsf{blue},-) \stackrel{\text{sort of}}{=}$

deep red Bing cherries small blue marble beautiful blue ocean did you put the kettle on red and blue fireworks Sencha green tea



Representable functors also behave like building blocks. We can use them to construct *new* copresheaves, since Set^L has nice structure. It is has all limits, colimits, and is Cartesian closed.

This means we have notions of conjunction, disjunction, and implication.



Representable functors also behave like building blocks. We can use them to construct *new* copresheaves, since Set^L has nice structure. It is has all limits, colimits, and is Cartesian closed.

This means we have notions of conjunction, disjunction, and implication.

EXAMPLE

Coproducts are computed pointwise.

The functor $L(red, -) \sqcup L(blue, -)$ assigns to an expression *x* the union of the sets L(red, x) and L(blue, x).





op

blue















SYNTAX 2.0







ENRICHED CATEGORY THEORY

In category theory, arrows $x \rightarrow y$ in a category C form a set, denoted C(x, y).



In enriched category theory, C(x, y) may not be (just) a set.



If this is a vector space, then C is said to be enriched over Vect.

If this is a topological space, then C is said to be enriched over Top.

If this is a truth value (0 or 1), then C is said to be enriched over Truth and is a preorder. If this is a group, then C is said to be enriched over Group.

C(x, y)



If this a conditional probability, then C is enriched over [0,1].



SETS *X, Y, ...*

FUNCTIONS $X \rightarrow Y$

UNIT INTERVAL

NUMBERS $0 \le a \le 1$

PREORDER $a \leq b$



SETS *X*, *Y*, ...

FUNCTIONS $X \to Y$

CARTESIAN PRODUCT $X \times Y$

UNIT

*

NUMBERS $0 \le a \le 1$

PREORDER

 $a \leq b$

MULTIPLICATION $a \cdot b$

UNIT 1



SETS *X, Y, ...*

FUNCTIONS $X \rightarrow Y$

CARTESIAN PRODUCT $X \times Y$

UNIT

*

ALL LIMITS

ALL COLIMITS

UNIT INTERVAL

NUMBERS $0 \le a \le 1$

PREORDER

 $a \leq b$

MULTIPLICATION $a \cdot b$

UNIT 1

ALL LIMITS

 $a \wedge b = \min\{a, b\}$

ALL COLIMITS

 $a \lor b = \max\{a, b\}$



SETS *X, Y, ...*

FUNCTIONS $X \rightarrow Y$

CARTESIAN PRODUCT $X \times Y$

UNIT

*

ALL LIMITS

ALL COLIMITS

 $\begin{array}{c} \textbf{CLOSURE} \\ X^Y \end{array}$

UNIT INTERVAL

NUMBERS $0 \le a \le 1$

PREORDER

 $a \leq b$

MULTIPLICATION $a \cdot b$

UNIT 1

ALL LIMITS

 $a \wedge b = \min\{a, b\}$

ALL COLIMITS

 $a \lor b = \max\{a, b\}$

CLOSURE $[a, b] = \min\{b/a, 1\}$

OBJECTS $X, Y, \ldots \in ob(C)$

A [0,1]-CATEGORY

OBJECTS $x, y, \ldots \in ob(\mathscr{C})$

OBJECTS $X, Y, \ldots \in ob(C)$

A "HOM SET" C(*X*, *Y*)

A [0,1]-CATEGORY

OBJECTS $x, y, \ldots \in ob(\mathscr{C})$

A "HOM OBJECT" $\mathscr{C}(x, y) \in [0, 1]$

OBJECTS $X, Y, \ldots \in ob(C)$

A "HOM SET" C(*X*, *Y*)

A FUNCTION $C(Y, Z) \times C(X, Y) \rightarrow C(X, Z)$

A [0,1]-CATEGORY

OBJECTS $x, y, \ldots \in ob(\mathscr{C})$

A "HOM OBJECT" $\mathscr{C}(x, y) \in [0, 1]$

A MORPHISM $\mathscr{C}(y, z) \cdot \mathscr{C}(x, y) \leq \mathscr{C}(x, y)$

OBJECTS $X, Y, \ldots \in ob(C)$

A "HOM SET" C(*X*, *Y*)

A FUNCTION $C(Y, Z) \times C(X, Y) \rightarrow C(X, Z)$

A FUNCTION

 $* \rightarrow \mathbf{C}(X, X)$

SATISFYING AXIOMS...

A [0,1]-CATEGORY

OBJECTS $x, y, \ldots \in ob(\mathscr{C})$

A "HOM OBJECT" $\mathscr{C}(x, y) \in [0, 1]$

A MORPHISM $\mathscr{C}(y, z) \cdot \mathscr{C}(x, y) \leq \mathscr{C}(x, y)$

A MORPHISM $1 \leq \mathscr{C}(x, x)$

SATISFYING AXIOMS...

FUNCTOR



$F: ob(C) \rightarrow ob(D)$

 $F_{X,Y}$: $C(X, Y) \to D(F(X), F(Y))$

[0,1]-FUNCTOR



$f: \operatorname{ob}(\mathscr{C}) \to \operatorname{ob}(\mathscr{D})$

 $\mathscr{C}(x,y) \leq \mathscr{D}(f(x),f(y))$

COPRESHEAVES



CATEGORY

Set

COPRESHEAF CATEGORY

[0,1]-COPRESHEAVES



 $[0,1]^{6}$

[0,1]-CATEGORY OF [0,1]-COPRESHEAVES



AS AN ENRICHED CATEGORY

"SYNTAX 2.0"

- Let ${\mathscr L}$ denote the $[0,\!1]\mbox{-}{\rm category}$ with
- objects: strings of symbols (as before)
- hom-objects:

$$\mathscr{L}(x, y) = \begin{cases} p(y \mid x) & \text{if } x \leq y \\ 0 & \text{else} \end{cases}$$

"the probability that y extends x" This defines a [0,1]-category: $1 = p(x \mid x)$ & $p(y \mid x) \cdot p(z \mid y) = p(z \mid x)$

Nice, but limited.



SEMANTICS 2.0

Consider [0,1]-functors $\mathscr{L} \to [0,1]$. **Representable** [0,1]-functors contain same information as before *plus* probabilities. Example: The function $\mathscr{L}(blue, -)$ is supported on all texts that contain "blue."



SEVANTISZO

In enriched category theory, the appropriate notion of limits and colimits are called coproduct of [0,1]-copresheaves:

"weighted" (co)limits. After unwinding the definition, here's an example of a weighted



A GEOMETRIC PERSPECTIVE

Nonnegative extended reals $[0,\infty]$ form a category with $a \to b$ whenever $b \leq a$.



Now one can think about *distances* between copresheaves and much more....



- The map $-\ln : [0,1] \rightarrow [0,\infty]$ is a functor. So, we obtain a category enriched over $[0,\infty]$:









blue



ENRICHED YONEDA [0,1] **ENRICHED** $\mathscr{L}(\mathsf{blue}, -)$

(SUMMARY)

LINEAR ALGEBRA VS. CATEGORY THEORY CONTINUED...



VECTOR SPACE

 \mathbb{R}^X

(CO)PRESHEAF CATEGORY

Set

ONE-HOT ENGODINGS

 $x_i \in X$

 $\langle x_i, - \rangle \colon X \to \mathbb{R}$

 $\langle x_i, x_j \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$

 $x_2 = [0 \ 1 \ 0]$

REPRESENTABLE FUNCTORS

$c \in ob(C)$

$C(c, -): C \rightarrow Set$

 $C(c,d) = \{\text{morphisms } c \to d\}$



EVERY VECTOR IN \mathbb{R}^X IS A LINEAR COMBINATION OF ONE-HOT ENCODINGS.



EVERY <u>COPRESHEAF</u> IN Set^C IS A COLIMIT OF <u>REPRESENTABLE FUNCTORS</u>.

MATRICES

A MATRIX IS A FUNCTION

 $M\colon X\times Y\to\mathbb{R}.$

MATRICES CAN BE MULTIPLIED. $(MN)_{ik} = \sum_{j} M_{ij} \cdot N_{jk}$

EVERY MATRIX HAS SINGULAR VECTORS (SVD). $M^{\dagger}M$

•

•

PROFUNCTORS

A PROFUNCTOR IS A FUNCTOR $F: \mathbb{C} \times \mathbb{D}^{op} \rightarrow \text{Set.}$

PROFUNCTORS CAN BE COMPOSED.

$$F \circ G = \int^d F(-, d) \times G(d, -)$$

EVERY PROFUNCTOR HAS "NUCLEI."

 F^*F_*



ORIGINAL RESEARCH ARTICLE





An Enriched Category Theory of Language: From Syntax to Semantics

Tai-Danae Bradley¹ · John Terilla² · Yiannis Vlassopoulos³

Received: 24 June 2021 / Revised: 7 February 2022 / Accepted: 7 February 2022 / Published online: 11 March 2022 © The Author(s), under exclusive licence to Springer Science+Business Media LLC, part of Springer Nature 2022

Abstract

State of the art language models return a natural language text continuation from any piece of input text. This ability to generate coherent text extensions implies significant sophistication, including a knowledge of grammar and semantics. In this paper, we propose a mathematical framework for passing from probability distributions on extensions of given texts, such as the ones learned by today's large language models, to an enriched category containing semantic information. Roughly speaking, we model probability distributions on texts as a category enriched over the unit interval. Objects of this category are expressions in language, and hom objects are conditional probabilities that one expression is an extension of another. This category is syntactical—it describes what goes with what. Then, via the Yoneda embedding, we pass to the enriched category of unit interval-valued copresheaves on this syntactical category. This category of enriched copresheaves is semantic—it is where we find meaning, logical operations such as entailment, and the building blocks for more elaborate semantic concepts.

Keywords Category theory · Yoneda embedding · Compositionality · Natural language · Probability · Logic

Mathematics Subject Classification 18D20 · 18A25 · 18A30 · 18A35 · 18B25 · 18B35

☑ Tai-Danae Bradley tai.danae@math3ma.com

> John Terilla jterilla@gc.cuny.edu

Yiannis Vlassopoulos yiannis@tunnel.tech

¹ X, the Moonshot Factory and Sandbox@Alphabet, Mountain View, CA, USA

² The City University of New York and Tunnel, New York, NY, USA

³ Tunnel, New York, NY, USA