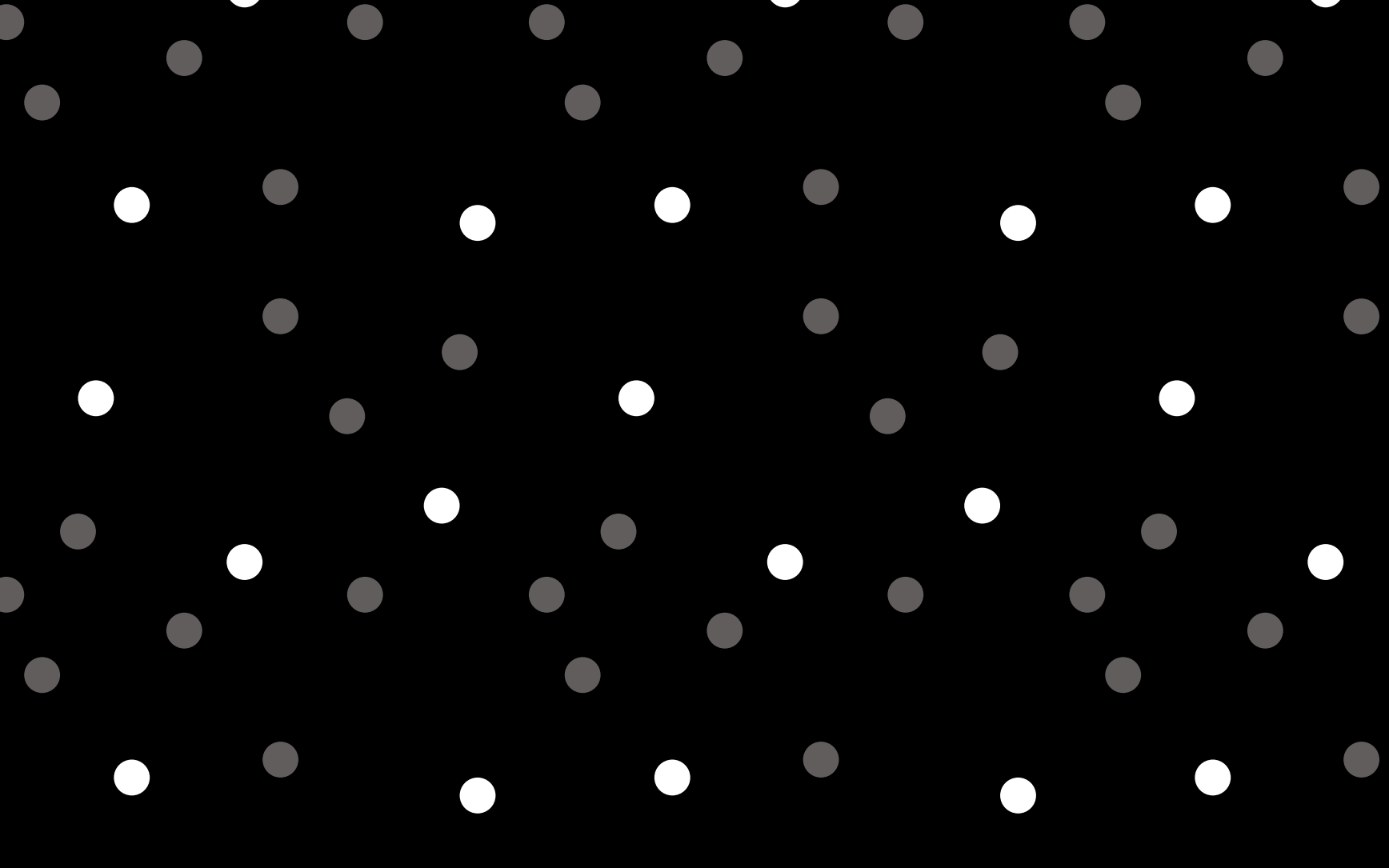


Papers

1. An Enriched Cat. of Language: From Syn. to Sem. by Tai Danae-Bradely, John Terilla, Yiannis Vlassopoulos
2. The Structure of Meaning in Language by TDB, Juan Luis Gastaldi, JT.

Motivation

- Language has structure (lots! not easy to characterise)
- Simplest idea: Corpus of words (a set!)

Common idea: thing w/ structure $\rightsquigarrow$ embed it in a place w/ more structure

e.g. a top. space $X \rightsquigarrow C(X) = \{\text{cont. fun.s } X \to \mathbb{C}\}$

- X is set of words $\rightsquigarrow k^X$ free v. space on X.

Basic idea: Make a matrix. Define $Y = X \times X$.

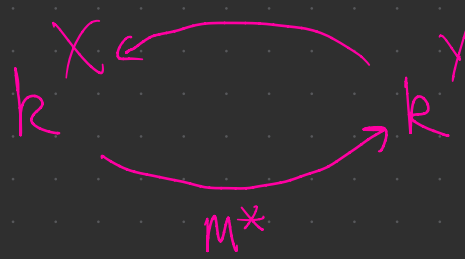
& the matrix is $m : X \times Y \longrightarrow k$.

$\mathbb{R}$

$Y \to k^X$

(firetruck, (red, is))

prob. that red precedes firetruck & is succeeds it in a text.



SVD: MM^*, M^*M are pos. semidef. $\rightsquigarrow$ Spectral Theorem

This gives an orth. basis of k^X

k^Y

$\{u_1, \dots, u_m\}$ eigenv. of MM^*

$\{v_1, \dots, v_n\}$ eigenv. of M^*M

$\lambda_1, \dots, \lambda_r, 0, \dots, 0$ eigenvalues

You have $M = U \Sigma V^*$ where columns of U, V are u_i, v_i.
Σ is diagonal w/ σ_i as entries.

So, for example, first few columns of U give embedding of k^X into low-dim space.

1. Nearby meanings have big inner product.

2. "Linearity of Semantics": Berlin − Germany + Spain $\approx$ Madrid

Queen − Woman + Man $\approx$ King. (Latent Semantic Analysis)

HAL. Formal Concept Analysis.

New Paper

- Move in the direction of category theory.
- Problems w/ previous idea:
  - ↳ Not very sequential.
  - ↳ We're only viewing X as a set.
- Bradley: define a category $\mathcal{L}$ → Objects: Phrases in English.
  - → Poset structure given by inclusion.

red firetruck ⟶ the red firetruck is

- We embed $\mathcal{L} \xhookrightarrow{y} [\mathcal{L}, \mathbf{Set}] = \widehat{\mathcal{L}}$.

$\text{red} \longmapsto y(\text{red})$. $\quad y(\text{red})(c) = \begin{cases} * & \text{red in } c \\ \varnothing & \text{else.} \end{cases}$

$$[y(\text{red}) \times y(\text{blue})](c) = \begin{cases} * & \text{red \& blue in } c \\ \varnothing & \text{otherwise} \end{cases}$$

$$[y(\text{red}) \sqcup y(\text{blue})](c) = \begin{cases} * \sqcup * & \text{if both red \& blue in } c \\ * & \text{if one is in } c \\ \varnothing & \text{otherwise} \end{cases}$$

$$[y(\text{red}), y(\text{blue})](c) \cong \widehat{\mathcal{L}}\,(y(c), [y(\text{red}), y(\text{blue})])$$
$$\cong \widehat{\mathcal{L}}\,(y(c) \times y(\text{red}), y(\text{blue})] \leftarrow \text{define it as this.}$$

Thinking about it, this will only ever be either $*$ or $\emptyset$

$$[y(a), y(b)](c) = \begin{cases} * & \text{whenever a tent contains } c \text{ \& } a, \text{ it contains } b \\ \emptyset & \text{otherwise} \end{cases}$$

↑ She calls this the implication.

This is nice BUT we added the comp. structure of language but not any statistical information ↑

$y(\text{red}) \rightsquigarrow$ the preimage of $*$ is the "principal ideal" associated to red.

- New ingredient: enriched category theory!

Brief Overview of Enr. Categ. Theory

$-\otimes-: \mathcal{V}\times\mathcal{V}\to\mathcal{V}$

You pick your favourite monoidal category $(\mathcal{V}, \otimes, I)$ (I: unit)

We will often want a GOOD one:

- ↳ Symmetric
- ↳ Closed (i.e. $\otimes$ has a right adjoint)

we will want to have categories where instead of $\mathrm{Hom}(X,Y)$ being a set, it's gonna be an object in $\mathcal{V}$.

Classic example: hom-sets are actually abelian groups.

You define a $\mathcal{V}$-category $\mathcal{C}$ as:

1. Objs are as normal.
2. For each X, Y objs, you have $\mathcal{C}(X,Y) \in ob(\mathcal{V})$.
3. "Comp. rule" is a morphism in $\mathcal{V}$ of the form

$$\mathcal{C}(X,Y) \otimes \mathcal{C}(Y,Z) \xrightarrow{\circ_{X,Y,Z}} \mathcal{C}(X,Z)$$

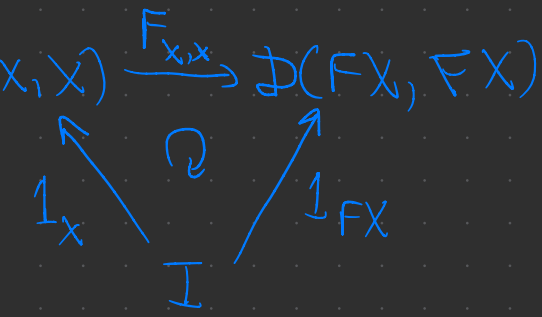
4. Identity morphisms are specified by an arrow $I \to \mathcal{C}(X,X)$
5. You write down various diagrams which explain that composition is assoc. & identities are identities.

$\mathcal{V}$-functors: $F: \mathcal{C} \to \mathcal{D}$

- Map on objects $X \longmapsto F(X) \in ob(\mathcal{D})$
- Give a morphism in $\mathcal{V}$ of the form

$$\mathcal{C}(X,Y) \xrightarrow{F_{X,Y}} \mathcal{D}(FX, FY)$$

- Diagrams.



Here: we will work w/ $\mathcal{V} = [0,1]$ (a categ. via its poset structure)
$\otimes = \cdot$ (multiplication) w/ unit 1.

$$[a,b] = \begin{cases} b/a & (a > b) \\ 1 & \text{else.} \end{cases}$$

New Enriched $\mathcal{L}$

Idea: $\mathcal{L}$ will have same objects (phrases in English)

Now: $\mathcal{L}(x,y) = \pi(y|x)$ ← "the probability that y contains x"

"This works": $\pi(x|x) = 1$ ← (we need to specify an identity for each obj., $1 \in \mathcal{L}(x,x)$)

$$\sum_{y \in Z} \pi(z|y)\,\pi(y|x) = \pi(z|x)$$

The "careful version" that she suggests w/ Vigneaux is we only consider literal extensions "red" ↪ "red firetruck"

$$\pi(y|x) = \begin{cases} 1 & \text{if } y = x \\ 0 & \text{if } y \text{ doesn't contain } x \\ \top & (\text{prob. of adding the next correct token}) \end{cases}$$

"red" ↪ "red men are happy"

"red" $\xrightarrow{p_{men}}$ "red men" $\xrightarrow{p_{are}}$ "red men are" $\xrightarrow{p_{happy}}$ "red men are happy"

& $\pi(\text{red men are happy} \mid \text{red}) = p_{happy}\, p_{are}\, p_{men}$.

If you only look at $T(x) = \{$ terminating strings beg. w/ $x\}$ then they say $\pi(- \mid x)$ assigns a prob. measure on $T(x)$.

We now have $\mathcal{L}$ & we can embed it in $\hat{\mathcal{L}} = [\mathcal{L}, [0,1]]$.

What is in $\hat{\mathcal{L}}$? An obj: $f: \mathcal{L} \to [0,1]$:

- sends phrases to numbers in $[0,1]$
- $\mathcal{L}(\text{red, the red firetruck}) \to [f(\text{red}), f(\text{the red firetruck})]$

$\leq \min\left\{\frac{f(\text{red})}{f(\_\_)}, 1\right\}$

The enriched Yoneda Lemma says

$$\hat{\mathcal{L}}(y(x), f) = f(x)$$

[For $c \in \mathcal{L}$, the rep. functor sends $x \mapsto \pi(x|c)$]

$$[\mathcal{C}, \mathcal{D}](F, G) = \int_{c \in \mathcal{C}} \mathcal{D}(Fc, Gc)$$

In our specific case, for $[0,1]$-copresheaves $f, g : \mathcal{L} \to [0,1]$

we have $\hat{\mathcal{L}}(f, g) = \inf_{x \in \mathcal{L}} [f(x), g(x)]$.

- There's an idea in ling.s "the distr. hypothesis" for words, there's a distr. which governs the phrases it's in. ⇝ the stat. data we're talking about IS the meaning of the word.

this is true in her model b/c it's the Yoneda Lemma.

Speculative Stuff

→ Actual computations (of coprod.s, of magn. etc.)
(let's form concepts!)

→ Enriched internal logic