Translating and Evolving: Towards a model of language change in DisCoCat

Tai-Danae Bradley, Martha Lewis, Jade Master, and Brad Theilman

SYCO3, March 2019

Oxford, UK

Background

Adjoint School, ACT 2018

Overview

The Workshop on Applied Category Theory 2018 takes place in May 2018. A principal goal of this workshop is to bring early career researchers into the applied category theory community. Towards this goal, we are organising the Adjoint School.

The Adjoint School will run from January to April 2018. By the end of the school, each participant will:

www.appliedcategorytheory.org

How can we incorporate language change and language learning into DisCo?

Translation - construed in a broad sense

- Translation between two different languages - French to English and back
- Translation between different levels of complexity of the same language
- Translation between two users of one language - updating each other's language models
- The aim is to provide a categorical description of translation that encompasses these three different notions.

Preliminary concepts

- We introduce the notion of a *language model* that unifies the product space representation of [Coecke et al., 2010] and the functorial representation of [Kartsaklis et al., 2013]
- This allows us to formalize the notion of *lexicon* which had previously been only loosely defined in the DisCoCat framework
- We then describe how to build a *dictionary* between two lexicons

Categorical Language Models and Translations

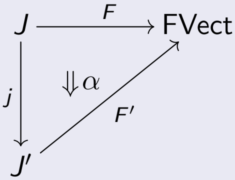
Definition

Let J be a category which is freely monoidal on some set of grammatical types. A **distributional categorical language model** *or* **language model** *for short is a strong monoidal functor*

$$F\colon (J, \cdot) \to (\mathsf{FVect}, \otimes)$$

Definition

A **translation** $T = (j, \alpha)$ *from a language model* $F\colon J \to \mathrm{FVect}$ *to a language model* $F'\colon J' \to \mathrm{FVect}$ *is a monoidal functor* $j\colon J \to J'$ *and a monoidal natural transformation* $\alpha\colon F \Rightarrow F' \circ j$. *Pictorially,* (j, α) *is the following 2-cell*



Categorical Language Models and Translations

Definition

Let DisCoCat *be the category with distributional categorical language models as objects, translations as morphisms. Composition of morphisms runs as follows: Given translations* $(T = j, \alpha)$ *and* $T' = (j', \alpha')$*, the composite translation is computed pointwise. That is,* $T' \circ T$ *is the translation* $(j' \circ j, \alpha' \circ \alpha)$ *where* $\alpha' \circ \alpha$ *is the vertical composite of the natural transformations* α *and* α'*.*

Product Space Representation

Definition

Let $F\colon J \to \mathrm{FVect}$ *be a language model and let* $K\colon \mathrm{FVect} \to \mathrm{Cat}$ *be a faithful functor. The* **product space representation** *of* F *with respect to* K*, denoted* $\mathtt{PS}_K(F)$*, is the Grothendieck construction of* $K \circ F$*. Explicitly,* $\mathtt{PS}_K(F)$ *is a category where*

- *an object is a pair* $(g, \vec{u})$ *where* g *is an object of* J *and* $\vec{u}$ *is an object of* $K \circ F(g)$
- *a morphism from* $(g, \vec{u})$ *to* $(h, \vec{v})$ *is a tuple* (r, f) *where* $r\colon g \to h$ *is a morphism in* J *and* $f\colon K \circ F(r)(\vec{u}) \to \vec{v}$ *is a morphism in* $K \circ F(h)$
- *the composite of* $(r', f')\colon (g, \vec{u}) \to (h, \vec{v})$ *and* $(r, f)\colon (h, \vec{v}) \to (i, \vec{x})$ *is defined by*

$$(r, f) \circ (r', f') = (r \circ r', f \circ (K \circ F)(r)(f'))$$

Product Space Representation

What should we use for the functor K?

Definition

Let V be a finite dimensional real vector space. Then, the **free chaotic category** *on V, denoted $C(V)$, is a category where*

- *objects are elements of V and,*
- *for all $\vec{u}, \vec{v}$ in V we include a unique arrow $d(\vec{u}, \vec{v}) \colon \vec{u} \to \vec{v}$ labeled by the Euclidean distance $d(\vec{u}, \vec{v})$ between $\vec{u}$ and $\vec{v}$.*

This construction extends to a functor $C \colon \mathrm{FVect} \to \mathrm{Cat}$. For $M \colon V \to W$, define $C(M) \colon C(V) \to C(W)$ to be the functor which agrees with M on objects and sends arrows $d(\vec{u}, \vec{v})$ to $d(M\vec{u}, M\vec{v})$.

The morphisms in $C(V)$ for a vector space V allow us to keep track of the relationships between different words in V.

Product Space Representation

What should we use for the functor K?
When $K = C$ as in Definition 5 the product space representation is as follows:

- objects are pairs $(g, \overrightarrow{u})$ where g is a grammatical type and $\overrightarrow{u}$ is a vector in $F(g)$.
- a morphism $(r, d) \colon (g, \overrightarrow{u}) \to (h, \overrightarrow{v})$ is:
  - a type reduction $r \colon g \to h$
  - a positive real number $d \colon C \circ F(r)(\overrightarrow{u}) \to \overrightarrow{v}$

Product Space Representation

Proposition ($\mathtt{PS}_K(F)$ is monoidal)

For $K = C$ and $K = D$, $\mathtt{PS}_K(F)$ is a monoidal category with monoidal product given on objects by

$$(g, \overrightarrow{u}) \otimes (h, \overrightarrow{v}) = (g \cdot h, \Phi_{g,h}(\overrightarrow{u} \otimes \overrightarrow{v}))$$

and on morphisms by

$$(r, f) \otimes (r', f') = (r \cdot r', \Phi_{g,h}(f \otimes f'))$$

where $\Phi_{g,h} \colon K \circ F(g) \overrightarrow{\otimes} K \circ F(h) \to K \circ F(g \cdot h)$ is the natural isomorphism included in the data of the monoidal functor $K \circ F$.

Product Space Representation

Proposition (Translations are monoidal)

Let $K\colon \mathsf{FVect} \to \mathsf{Cat}$ *be a fully faithful functor. Then there is a functor* $\mathtt{PS}_K\colon \mathsf{DisCoCat} \to \mathsf{MonCat}$, *where* MonCat *is the category where objects are monoidal categories and morphisms are strong monoidal functors, that sends*

- *language models* $F\colon J \to \mathsf{Cat}$ *to the monoidal category* $\mathtt{PS}_K(F)$
- *translations* $T = (j, \alpha)$ *to the strong monoidal functor* $\mathtt{PS}_K(T)\colon \mathtt{PS}_K(F) \to \mathtt{PS}_K(F')$ *where the functor* $\mathtt{PS}_K(T)$ *acts as follows:*
  - *On objects,* $\mathtt{PS}_K(T)$ *sends* $(g, \overrightarrow{u})$ *to* $(j(g), \alpha_g \overrightarrow{u})$.
  - *Suppose* $(r, f)\colon (g, \overrightarrow{u}) \to (h, \overrightarrow{v})$ *is a morphism in* $\mathtt{PS}_K(F)$ *so that* $r\colon g \to h$ *is a reduction in* J *and* $f\colon F(r)(\overrightarrow{u}) \to \overrightarrow{v}$ *is a morphism in* $F(h)$*. Then* $\mathtt{PS}_K(T)$ *sends* (r, f) *to the pair* $(j(r), \alpha_h \circ f)$.

Defining the lexicon

Definition

Let F be a categorical language model and let W be a finite set of words, viewed as discrete category. Then a **lexicon for** *F is a functor $\ell\colon W \to \mathtt{PS}(F)$. This corresponds to a function from W into the objects of $\mathtt{PS}(F)$.*

NB: We have now fixed $K = C$ and dropped the subscript in $\mathtt{PS}(F)$

- We extend this to phrases, i.e. finite sequences of words $v_1 \ldots v_n \in W^*$ where W^* is the free monoid on W.
- This defines a unique object in $\mathtt{PS}(F)$:

$$(g, \overrightarrow{v}) := \otimes_{i=1}^{n} \ell(v_i) = (g_1, \overrightarrow{v}_1) \otimes \ldots \otimes (g_n, \overrightarrow{v}_n)$$
$$= (g_1 \cdots g_n, \overrightarrow{v}_1 \otimes \ldots \otimes \overrightarrow{v}_n)$$

- The extension of ℓ to W^* will be denoted by

$$\ell^*\colon W^* \to \mathtt{PS}(F).$$

Defining a dictionary

Definition

Let $\ell\colon W \to \mathtt{PS}(F)$ *and* $m\colon V \to \mathtt{PS}(G)$ *be lexicons and let* T *be a translation from* F *to* G. *Then, the* F-G **dictionary** *with respect to* T *is the comma category*

$$(\mathtt{PS}(T) \circ \ell^*) \downarrow m^*$$

denoted by Dict_T. *Since* W *and* V *are discrete categories,* $(\mathtt{PS}(T) \circ \ell^*) \downarrow m^*$ *is a set of triples* $(p, (r, d), q)$ *where* $p \in W^*$, $q \in V^*$ *and* $(r, d)\colon (\mathtt{PS}(T) \circ \ell)(p) \to m(q)$ *is a morphism in* $\mathtt{PS}(G)$.

Defining a dictionary

Explicitly, let

$$\ell(p) = (g, \overrightarrow{p}) \quad \text{and} \quad m(q) = (g', \overrightarrow{q})$$

then (r, d) is

- a type reduction $r\colon j(g) \to g'$ in the grammar category J
- a morphism d in $C \circ G(g')$. Recall from definition 5 that this corresponds to a real number $d(\overrightarrow{p}', \overrightarrow{q})$ denoting the distance between $\overrightarrow{p}'$ and $\overrightarrow{q}$ in $G(g')$.

Example

Example (Translation at the phrase level)

- $J_{En} = \mathscr{C}(\{n_E, s_E\})$, $J_S = \mathscr{C}\{n_S, s_S\}$
- Consider distributional categorical language models

$$F_{En} \colon J_{En} \to \text{Cat and } F_{Sp} \colon J_{Sp} \to \text{Cat}$$

- We take the fragment consisting just of nouns and intransitive verbs. Let $F_{En}(n) = N_{En}$, $F_{En}(s) = S_{En}$, $F_{Sp}(n) = N_{Sp}$ and $F_{Sp}(s) = S_{Sp}$.
- To specify $\mathtt{PS}(T)$ we set:
  - $j(n_E) = n_S$ and $j(s_E) = j(s_S)$
  - $\alpha_{n_E}, \alpha_{s_E}$ to be linear transformations that 'behave in the right way'

Example

What does ‘behave in the right way’ mean?

- Backpedal: Suppose we are interested only in nouns. Then we can learn a linear transformation from lists of noun vectors and their translations.
- However, α is a monoidal natural transformation, so $\alpha_{gh} = \alpha_g \otimes \alpha_h$ for every product type *gh*.
- This holds if $\alpha_{n_E}, \alpha_{s_E}$ are unitary
- In general, $\alpha_{n_E}, \alpha_{s_E}$ won't be unitary

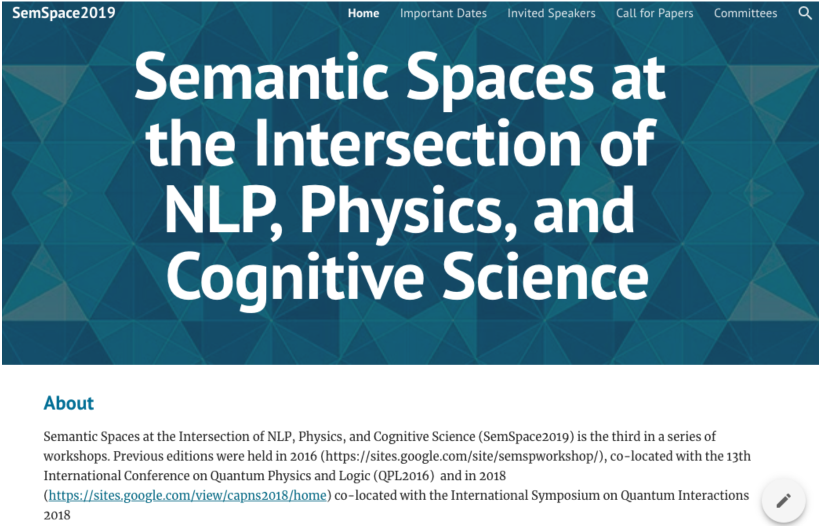
Future Work

- Our current model doesn't deal with changing the order of words, like adj-noun/noun-adj. This is a matter of priority
- Use the metric structure of vector spaces to form dictionaries
- Work with the diagrammatic calculus
- Investigate meaning change and negotiated meaning
- Investigate the implementation of compositional translation matrices.

If you liked this...

... try this!

https://sites.google.com/view/semspace2019/home



SemSpace2019 Home Important Dates Invited Speakers Call for Papers Committees

Semantic Spaces at the Intersection of NLP, Physics, and Cognitive Science

About

Semantic Spaces at the Intersection of NLP, Physics, and Cognitive Science (SemSpace2019) is the third in a series of workshops. Previous editions were held in 2016 (https://sites.google.com/site/semspworkshop/), co-located with the 13th International Conference on Quantum Physics and Logic (QPL2016) and in 2018 (https://sites.google.com/view/capns2018/home) co-located with the International Symposium on Quantum Interactions 2018

References I

Coecke, B., Sadrzadeh, M., and Clark, S. (2010).
Mathematical foundations for a compositional distributional model of meaning.
arXiv preprint arXiv:1003.4394.

Kartsaklis, D., Sadrzadeh, M., Pulman, S., and Coecke, B. (2013).
Reasoning about meaning in natural language with compact closed categories and frobenius algebras.
Logic and Algebraic Structures in Quantum Computing, page 199.