Coordination in Categorical Compositional Distributional Semantics

Dimitri Kartsaklis

School of Electronic Engineering
and Computer Science
Queen Mary University of London

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

Glasgow, Scotland

11th June 2016

In a nutshell

- Categorical compositional distributional semantics unifies two orthogonal semantic paradigms:
  - The type-logical compositional approach of formal semantics
  - The quantitative perspective of vector space models of meaning
- The goal is to represent sentences as points in some high dimensional metric space.

In this work:

By using Frobenius algebras, we deal with coordination between identical syntactic types, which accounts for the majority of coordination cases in language.

Outline

1. Categorical compositional distributional semantics
2. Frobenius algebras in FdVect and linguistics
3. Coordination in CCDS
4. Non-standard coordination cases
5. Conclusions

Categorical compositional distributional semantics

A pregroup grammar $P(\Sigma, \mathcal{B})$ is a relation that assigns grammatical types from a pregroup algebra freely generated over a set of atomic types $\mathcal{B}$ to words of a vocabulary Σ.

- Pregroup grammars are structurally homomorphic with the category of finite-dimensional Hilbert spaces and linear maps, **FdVect** (both share compact closure)
- In abstract terms, there exists a structure-preserving passage from grammar to meaning:

$$\mathcal{F} : \text{Grammar} \to \text{Meaning}$$

expressed as a strongly monoidal functor:

$$\mathcal{F} : P(\Sigma, \mathcal{B}) \to \textbf{FdVect}$$

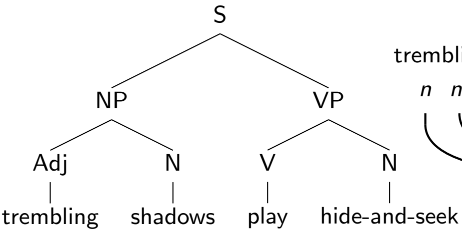
A multi-linear model

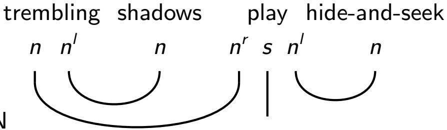
The grammatical type of a word defines the vector space in which the word lives:

- Nouns are vectors in N;
- adjectives are linear maps $N \rightarrow N$, i.e elements in $N \otimes N$;
- intransitive verbs are linear maps $N \rightarrow S$, i.e. elements in $N \otimes S$;
- transitive verbs are bi-linear maps $N \otimes N \rightarrow S$, i.e. elements of $N \otimes S \otimes N$;
- and so on.

- The composition operation is tensor contraction, based on inner product.

Categorical composition: example





Type reduction morphism:

$$(\epsilon_n^r \cdot 1_s) \circ (1_n \cdot \epsilon_n^l \cdot 1_{n^r \cdot s} \cdot \epsilon_n^l) : n \cdot n^l \cdot n \cdot n^r \cdot s \cdot n^l \cdot n \to s$$

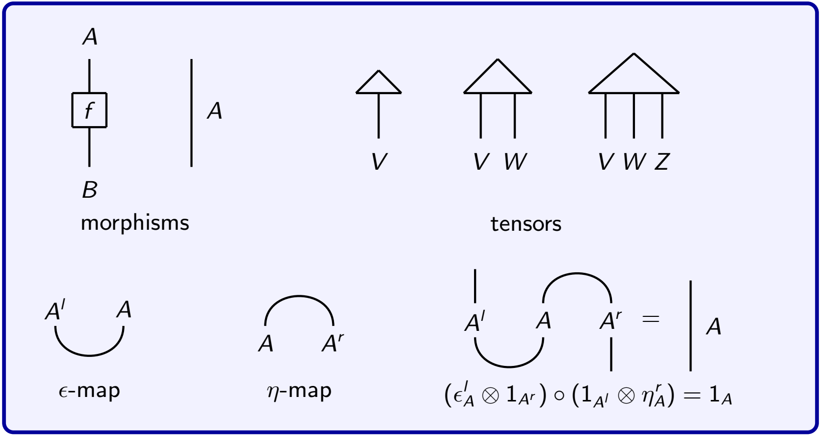
$$\mathcal{F}\left[(\epsilon_n^r \cdot 1_s) \circ (1_n \cdot \epsilon_n^l \cdot 1_{n^r \cdot s} \cdot \epsilon_n^l)\right] \left(\overline{trembling} \otimes \overrightarrow{shadows} \otimes \overline{play} \otimes \overrightarrow{hide\text{-}and\text{-}seek}\right) =$$

$$(\epsilon_N \otimes 1_S) \circ (1_N \otimes \epsilon_N \otimes 1_{N \otimes S} \otimes \epsilon_N) \left(\overline{trembling} \otimes \overrightarrow{shadows} \otimes \overline{play} \otimes \overrightarrow{hide\text{-}and\text{-}seek}\right) =$$

$$\overline{trembling} \times \overrightarrow{shadows} \times \overline{play} \times \overrightarrow{hide\text{-}and\text{-}seek}$$

$\overrightarrow{shadows}, \overrightarrow{hide\text{-}and\text{-}seek} \in N \qquad \overline{trembling} \in N \otimes N \qquad \overline{play} \in N \otimes S \otimes N$

A graphical language for monoidal categories



- Vectors and tensors are states: $\overrightarrow{v} : I \to V$, $\overline{w} : I \to V \otimes V$ and so on.

Outline

1. Categorical compositional distributional semantics
2. Frobenius algebras in FdVect and linguistics
3. Coordination in CCDS
4. Non-standard coordination cases
5. Conclusions

Frobenius algebras in **FdVect**

- Given a symmetric CCC $(\mathcal{C}, \otimes, I)$, an object $X \in \mathcal{C}$ has a Frobenius structure on it if there exist morphisms:

  $\Delta : X \to X \otimes X \; , \; \iota : X \to I$ and $\mu : X \otimes X \to X \; , \; \zeta : I \to X$

  conforming to the Frobenius condition:

$$(\mu \otimes 1_X) \circ (1_X \otimes \Delta) = \Delta \circ \mu = (1_X \otimes \mu) \circ (\Delta \otimes 1_X)$$

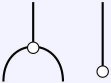
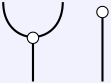
- In **FdVect**, any vector space V with a fixed basis $\{\overrightarrow{v_i}\}_i$ has a commutative special Frobenius algebra over it [Coecke and Pavlovic, 2006]:

$$\Delta : \overrightarrow{v_i} \mapsto \overrightarrow{v_i} \otimes \overrightarrow{v_i} \qquad \mu : \overrightarrow{v_i} \otimes \overrightarrow{v_i} \mapsto \overrightarrow{v_i}$$

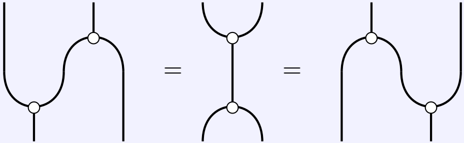
- It can be seen as copying and merging of the basis.

Graphical representation

- Frobenius maps:

$(\Delta, \iota) =$  $(\mu, \zeta) =$ 

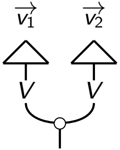
- Frobenius condition:



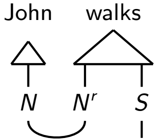
Merging (1/2)

- In **FdVect**, the merging μ-map becomes element-wise vector multiplication:

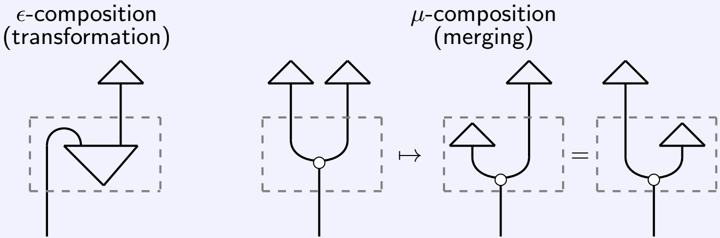
$$\mu(\overrightarrow{v_1} \otimes \overrightarrow{v_2}) = \overrightarrow{v_1} \odot \overrightarrow{v_2} =$$



- An alternative form of composition between operands of the same order; both of them contribute equally to the final result
- Different from standard ϵ-composition, which has a transformational effect. An intransitive verb, for example, is a map $N \to S$ that transforms a noun into a sentence:



Merging (2/2)

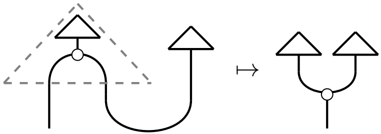


Applications of merging in linguistics:

- Noun modification by relative clauses [Sadrzadeh et al., MoL 2013]
- Modelling intonation at sentence level [Kartsaklis and Sadrzadeh, MoL 2015]
- Modelling non-compositional compounds (e.g. 'pet-fish') [Coecke and Lewis, QI 2015]
- Modelling coordination [Kartsaklis Ph.D. thesis (2015); this work]

Copying

- In **FdVect**, the Δ-map converts vectors to diagonal matrices
- It can be seen as duplication of information; a single wire is split in two; i.e. a maximally entangled state
- A form of type-raising (converts an atomic type to a function) [Kartsaklis et al., COLING 2012]:



- A means of syntactic movement; the same word can efficiently interact with different parts of the sentence [Sadrzadeh et al., MoL 2013]

Outline

1. Categorical compositional distributional semantics
2. Frobenius algebras in FdVect and linguistics
3. Coordination in CCDS
4. Non-standard coordination cases
5. Conclusions

Coordination and Frobenius maps

Coordination

The grammatical connection of two or more words, phrases, or clauses to give them equal emphasis and importance. The connected elements, or conjuncts, behave as one.

Merging and copying are the key processes of coordination:

$$\textit{context}\ c_1\ \textit{conj}\ c_2 \mapsto [\textit{context}\ c_1]\ \textit{conj}\ [\textit{context}\ c_2]$$

(1) Mary studies [philosophy] and [history] $\models$
[Mary studies philosophy] and [Mary studies history]

(2) Men [like sports] and [play football] $\models$
[Men like sports] and [men play football]

(3) John [sleeps] and [snores] $\models$
[John sleeps] and [John snores]

Coordinating identical types

- Translating the usual ternary rule $X\ CONJ\ X \to X$ to pregroups gives:

$$x \cdot (x^r \cdot x \cdot x^l) \cdot x \leq 1 \cdot x \cdot 1 = x$$

- Applying the syntax-to-semantics functor:

$$\mathcal{F}\left[(\epsilon_x^r \cdot 1_x \cdot \epsilon_x^l) \circ (x \cdot x^r \cdot x \cdot x^l \cdot x)\right] = (\epsilon_X^r \otimes 1_X \otimes \epsilon_X^l) \circ (\overrightarrow{x_1} \otimes \overline{conj}_X \otimes \overrightarrow{x_2})$$

We need a way to translate ϵ-composition to μ-composition

From ϵ-composition to μ-composition

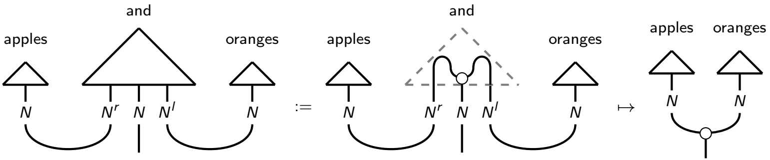
$$
\begin{array}{ccccc}
I & \xrightarrow{\overrightarrow{x_1} \otimes \overrightarrow{x_2}} & X \otimes X & \xrightarrow{1_X \otimes \eta^r_X \otimes \eta^l_X \otimes 1_X} & X \otimes X^r \otimes X \otimes X \otimes X^l \otimes X \\
 & & \Big\downarrow \mu_X & & \Big\downarrow 1_X \otimes 1_{X^r} \otimes \mu_X \otimes 1_{X^l} \otimes 1_X \\
 & & X & \xleftarrow[\epsilon^r_X \otimes 1_X \otimes \epsilon^l_X]{} & X \otimes X^r \otimes X \otimes X^l \otimes X
\end{array}
$$

$$
\begin{aligned}
\mu_X \circ (\overrightarrow{x_1} \otimes \overrightarrow{x_2}) &= (\epsilon^r_X \otimes 1_X \otimes \epsilon^l_X) \circ (1_X \otimes 1_{X^r} \otimes \mu_X \otimes 1_{X^l} \otimes 1_X) \circ (1_X \otimes \eta^r_X \otimes \eta^l_X \otimes 1_X) \circ (\overrightarrow{x_1} \otimes \overrightarrow{x_2}) \\
&= (\epsilon^r_X \otimes 1_X \otimes \epsilon^l_X) \circ \left(1_X \otimes \left[(1_{X^r} \otimes \mu_X \otimes 1_{X^l}) \circ (\eta^r_X \otimes \eta^l_X)\right] \otimes 1_X\right) \circ (\overrightarrow{x_1} \otimes \overrightarrow{x_2}) \\
&= (\epsilon^r_X \otimes 1_X \otimes \epsilon^l_X) \circ \left(\overrightarrow{x_1} \otimes \left[(1_{X^r} \otimes \mu_X \otimes 1_{X^l}) \circ (\eta^r_X \otimes \eta^l_X)\right] \otimes \overrightarrow{x_2}\right)
\end{aligned}
$$

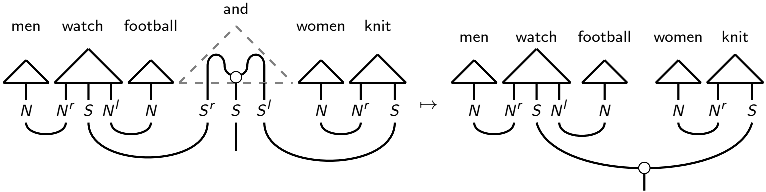
Coordination morphism:

$$
\overline{conj}_X : I \xrightarrow{\eta^r_X \otimes \eta^l_X} X^r \otimes X \otimes X \otimes X^l \xrightarrow{1_{X^r} \otimes \mu_X \otimes 1_{X^l}} X^r \otimes X \otimes X^l
$$

Coordinating atomic types



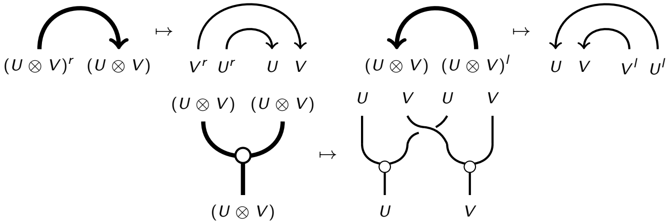
$$(\epsilon^r_N \otimes 1_N \otimes \epsilon^l_N) \circ (\overrightarrow{apples} \otimes \overline{conj}_N \otimes \overrightarrow{oranges}) = \mu(\overrightarrow{apples} \otimes \overrightarrow{oranges}) = \overrightarrow{apples} \odot \overrightarrow{oranges}$$



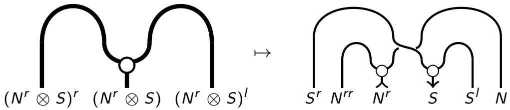
$$(\overrightarrow{men}^{\mathsf{T}} \times \overline{watch} \times \overrightarrow{football}) \odot (\overrightarrow{women}^{\mathsf{T}} \times \overline{knit})$$

Coordinating compound types

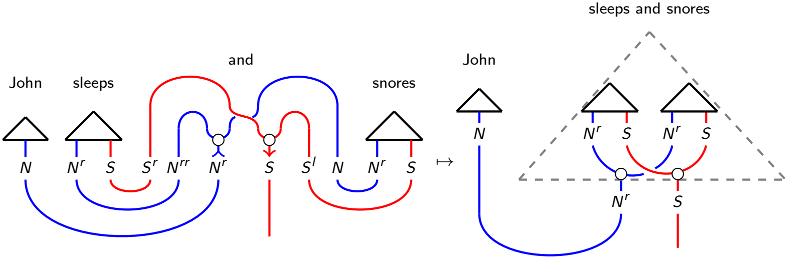
- Lifting the maps to compound objects gives:



- For the case of a verb phrase, we get:



Coordinating verb phrases



1. The subject of the coordinate structure ('John') is copied at the N^r input of the coordinator;
2. the first branch interacts with verb 'sleeps' and the second one with verb 'snores'; and
3. the S wires of the two verbs that carry the individual results are merged together with μ-composition.

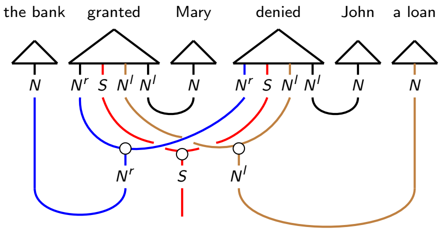
$$\overrightarrow{John}^{\mathsf{T}} \times (\overline{sleep} \odot \overline{snore})$$

($\odot$ here denotes the Hadamard product between matrices)

Moving to higher order tensors

Everything lifts coherently to higher order tensors. Merging and copying in coordination between ditransitive verbs:

(4) The bank [granted Mary] but [denied John] a loan $\models$
[The bank granted Mary a loan] but [the bank denied John a loan]



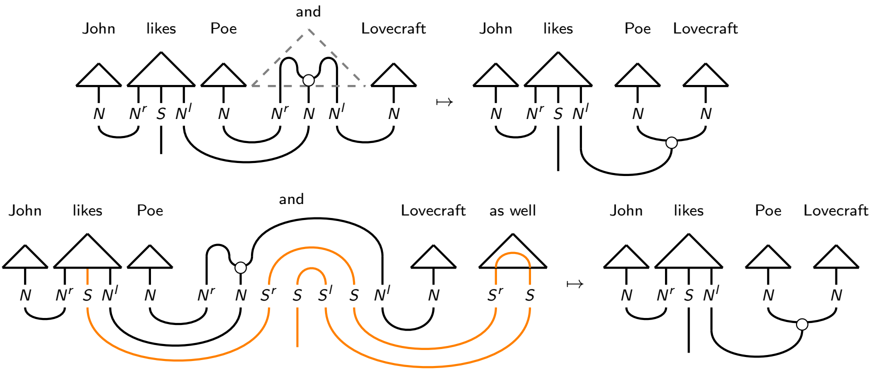
$$\overrightarrow{bank}^{\mathsf{T}} \times \left[(\overline{grant} \times \overrightarrow{Mary}) \odot (\overline{deny} \times \overrightarrow{John}) \right] \times \overrightarrow{loan}$$

Outline

1. Categorical compositional distributional semantics
2. Frobenius algebras in FdVect and linguistics
3. Coordination in CCDS
4. Non-standard coordination cases
5. Conclusions

Stripping

(5) John likes Poe, and$_{\text{[John likes]}}$ Lovecraft as well $\models$ John likes Poe and Lovecraft



(Work in progress with Matt Purver and Mehrnoosh Sadrzadeh on other forms of ellipsis in language.)

Outline

1. Categorical compositional distributional semantics
2. Frobenius algebras in FdVect and linguistics
3. Coordination in CCDS
4. Non-standard coordination cases
5. Conclusions

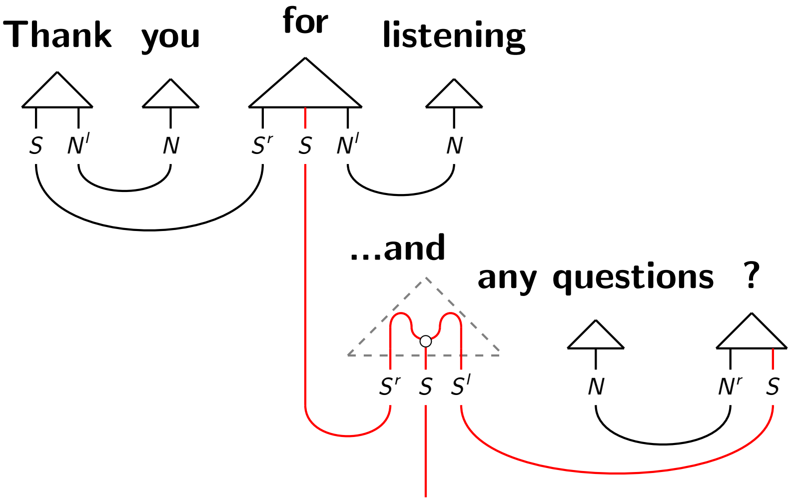
Conclusion and future work

Main points:

- Merging and copying are the key processes of coordination
- Frobenius operators provide a natural account to model those in compositional distributional semantics
- A model with two compositional operators over a distributional setting

Future work:

- Coordination between different types ("John works evenings and on weekends")
- A proper distinction between conjunction and disjunction?
- Investigating other forms of ellipsis, such as verb-phrase ellipsis (joint work with Purver and Sadrzadeh)


Thank you for listening
S N^l N S^r S N^l N
...and
any questions ?
S^r S S^l N N^r S

References I

Coecke, B. and Lewis, M. (2016).
Quantum Interaction: 9th International Conference, QI 2015, Filzbach, Switzerland, July 15-17, 2015, Revised Selected Papers, chapter A Compositional Explanation of the 'Pet Fish' Phenomenon, pages 179–192.
Springer International Publishing, Cham.

Coecke, B. and Pavlovic, D. (2006).
Quantum measurements without sums.
arXiv preprint quant-ph/0608035.

Coecke, B., Sadrzadeh, M., and Clark, S. (2010).
Mathematical Foundations for a Compositional Distributional Model of Meaning. Lambek Festschrift.
Linguistic Analysis, 36:345–384.

Kartsaklis, D. (2015).
Compositional Distributional Semantics with Compact Closed Categories and Frobenius Algebras.
PhD thesis, University of Oxford.

Kartsaklis, D. and Sadrzadeh, M. (2014).
A study of entanglement in a categorical framework of natural language.
In Coecke, B., Hasuo, I., and Panangaden, P., editors, Proceedings 11th workshop on *Quantum Physics and Logic,* Kyoto, Japan, 4-6th June 2014, volume 172 of *Electronic Proceedings in Theoretical Computer Science*, pages 249–261. Open Publishing Association.

Kartsaklis, D., Sadrzadeh, M., and Pulman, S. (2012).
A unified sentence space for categorical distributional-compositional semantics: Theory and experiments.
In *Proceedings of 24th International Conference on Computational Linguistics (COLING 2012): Posters*, pages 549–558, Mumbai, India. The COLING 2012 Organizing Committee.

References II

Kartsaklis, D., Sadrzadeh, M., Pulman, S., and Coecke, B. (2015).
Reasoning about meaning in natural language with compact closed categories and Frobenius algebras.
In Chubb, J., Eskandarian, A., and Harizanov, V., editors, *Logic and Algebraic Structures in Quantum Computing and Information*, Association for Symbolic Logic Lecture Notes in Logic. Cambridge University Press.

Piedeleu, R., Kartsaklis, D., Coecke, B., and Sadrzadeh, M. (2015).
Open system categorical quantum semantics in natural language processing.
In *Proceedings of the 6th Confenrence on Algebra and Coalgebra in Computer Science (CALCO)*, Nijmegen, Netherlands.

Sadrzadeh, M., Clark, S., and Coecke, B. (2013).
The Frobenius anatomy of word meanings I: subject and object relative pronouns.
Journal of Logic and Computation, Advance Access.

Sadrzadeh, M., Clark, S., and Coecke, B. (2014).
The Frobenius anatomy of word meanings II: Possessive relative pronouns.
Journal of Logic and Computation.