

Coordination in Categorical Compositional Distributional Semantics

Dimitri Kartsaklis

Queen Mary University of London
School of Electronic Engineering and Computer Science
Mile End Road, London E1 4NS, United Kingdom
d.kartsaklis@qmul.ac.uk

An open problem with categorical compositional distributional semantics is the representation of words that are considered semantically vacuous from a distributional perspective, such as determiners, prepositions, relative pronouns or coordinators. This paper deals with the topic of coordination between identical syntactic types, which accounts for the majority of coordination cases in language. By exploiting the compact closed structure of the underlying category and Frobenius operators canonically induced over the fixed basis of finite-dimensional vector spaces, we provide a morphism as representation of a coordinator tensor, and we show how it lifts from atomic types to compound types. Linguistic intuitions are provided, and the importance of the Frobenius operators as an addition to the compact closed setting with regard to language is discussed.

1 Introduction

Inspired by quantum mechanics, and specifically by the category-theoretic manifestation of quantum mechanics as set out by Abramsky and Coecke [1], the compositional framework of Coecke, Sadrzadeh and Clark [2] provides an intuitive way to model the interactions between words within a sentence in the context of a distributional model of meaning. In this setting, nouns are represented as simple vectors living in a basic vector space, while words with a relational nature, such as verbs and adjectives, are multi-linear maps living in tensor product spaces and acting on the noun vectors. The model uses the framework of compact closed categories to unify a grammar expressed as a Lambek pregroup [8] with the category of finite-dimensional vector spaces and linear maps (**FdVect**), which is where a distributional model of semantics lives; this is achieved by means of a functorial passage that essentially translates every grammatical derivation to a multi-linear algebraic manipulation on the word vectors and tensors.

A long-standing challenge in compositional distributional models is the representation of functional words, such as prepositions, relative pronouns, or coordinators, since in principle the meaning of these cannot be given by distributional or any other statistical methods. In order to cope with this problem, researchers in the past have exploited the notion of Frobenius algebras which can be canonically induced over the basis of finite-dimensional vector spaces [4]. Frobenius algebras constitute one of the fundamental structures of categorical quantum mechanics [1] (the other one being compact closed categories), where they are used to model classical operations, such as copying or deleting information, that are not allowed in the quantum world. Interestingly, these properties of Frobenius algebras have also been proved useful in language, modelling linguistic aspects for which the representation power of the framework of compact closed categories is insufficient [12, 13, 7, 6, 3].

In this paper we show how Frobenius algebras, in conjunction with the compact closed structure of **FdVect**, can be used to model one of the most ubiquitous phenomena in language, that of coordination.¹

¹A preliminary account of this subject can be found in the doctoral thesis of the author [5].

As we will see in Section 3, from a linguistic perspective Frobenius multiplication can be seen as enforcing the two inputs to contribute equally to the result—an idea that captures the essence of coordination. Furthermore, the Frobenius co-multiplication allows the duplication of information when this makes linguistic sense; as it will become evident in Section 4.2, this action is necessary when coordinating text constituents with compound grammatical types, such as verb phrases. The coordinating morphisms we provide in this paper make use of these two notions; furthermore, they get intuitive linear-algebraic interpretations and come with compact composition formulas that simplify their practical application. Finally, in Section 4.3 we briefly examine the application of the Frobenius machinery on non-standard coordination cases that involve certain forms of ellipsis.

2 Background

This section provides a very short introduction to the compositional framework of Coecke et al. [2] and Frobenius algebras over **FdVect** [4]; the interested reader is encouraged to refer to the original papers for more details. We furthermore assume familiarity with the graphical language of compact closed categories; for an introduction specific to linguistics, see [10], App. A.

2.1 A functorial passage from syntax to semantics

We recall that a pregroup algebra is a partially ordered monoid, each element p of which has a left and a right adjoint, denoted as p^l and p^r respectively, such that:

$$p \cdot p^r \leq 1 \leq p^r \cdot p \quad p^l \cdot p \leq 1 \leq p \cdot p^l \quad (1)$$

Take $P(\mathcal{B})$ to be a pregroup algebra generated over a set of basic (or *atomic*) grammatical types \mathcal{B} , and Σ the vocabulary of the language; then a pregroup grammar is a relation $\Sigma \times P(\mathcal{B})$, denoted as $P(\Sigma, \mathcal{B})$ that assigns a grammar type to every word in the vocabulary [8]. Assuming $\mathcal{B} = \{s, n\}$, where s stands for a well-formed sentence and n for a well-formed noun phrase, we say that a word sequence $w_1 w_2 \dots w_n$ forms a grammatical sentence whenever $t_1 \cdot t_2 \cdot \dots \cdot t_n \leq s$, for $(w_i, t_i) \in P(\Sigma, \mathcal{B})$. For example, given the type assignments (‘Mary’, n), (‘likes’, $n^r \cdot s \cdot n^l$), (‘musicals’, n), the sentence “Mary likes musicals” is grammatical since $n \cdot n^r \cdot s \cdot n^l \cdot n \leq 1 \cdot s \cdot 1 \leq s$, according to Equation 1. Note that the type of the transitive verb ‘likes’ is *compound*, denoting a word that expects noun phrases at both of its sides in order to return a sentence.

Both a pregroup grammar and **FdVect** have compact closed structure. For a pregroup, the structural morphisms of compact closure ε and η become:

$$\varepsilon^r : p \cdot p^r \leq 1, \quad \varepsilon^l : p^l \cdot p \leq 1 \quad \eta^r : 1 \leq p^r \cdot p, \quad \eta^l : 1 \leq p \cdot p^l \quad (2)$$

Unlike a pregroup, **FdVect** is a symmetric compact closed category, meaning that for every pair of objects A, B there is an isomorphism $\sigma : A \otimes B \cong B \otimes A$. In **FdVect**, the ε map becomes the inner product between the involved vectors, and the η map defines identity matrices:

$$\varepsilon^l = \varepsilon^r : V \otimes V \rightarrow \mathbb{R} :: \sum_{ij} c_{ij} (\vec{v}_i \otimes \vec{v}_j) \mapsto \sum_{ij} c_{ij} \langle \vec{v}_i | \vec{v}_j \rangle \quad \eta^l = \eta^r : \mathbb{R} \rightarrow V \otimes V :: 1 \mapsto \sum_i \vec{v}_i \otimes \vec{v}_i \quad (3)$$

A structure-preserving passage from syntax to semantics can be developed between these two categories by using a strongly monoidal functor:

$$\mathcal{F} : P(\mathcal{B}, \Sigma) \rightarrow \mathbf{FdVect} \quad (4)$$

that sends each atomic type x to a vector space X , and compound types to tensor product of spaces, since $\mathcal{F}(p \cdot q) = \mathcal{F}(p) \otimes \mathcal{F}(q)$.

As an example, consider the sentence ‘‘John sleeps’’. The pregroup derivation takes the form $n \cdot n^r \cdot s \leq 1 \cdot s \leq s$, translated to morphism $(\varepsilon^r \cdot 1_s) \circ (n \cdot n^r \cdot s)$ according to Equation 2. Applying our functor to this, and taking $\overrightarrow{John} \in N$ and $\overrightarrow{sleep} \in N^r \otimes S$ to be the semantic representations of the words, will give:

$$\begin{aligned} \mathcal{F}[(\varepsilon^r \cdot 1_s) \circ (n \cdot n^r \cdot s)] &= \mathcal{F}(\varepsilon^r \cdot 1_s) \circ \mathcal{F}(n \cdot n^r \cdot s) = (\varepsilon^r \otimes 1_S) \circ (N \otimes N^r \otimes S) \\ &:= (\varepsilon^r \otimes 1_S) \circ (\overrightarrow{John} \otimes \overrightarrow{sleep}) = \overrightarrow{John}^\top \times \overrightarrow{sleep} \end{aligned}$$

From a linear-algebraic perspective, the above is just the matrix multiplication of the matrix representing the intransitive verb ‘sleep’ with the vector for noun ‘John’. For higher-order tensors (such as a transitive verb living in $N^r \otimes S \otimes N^l$), the composition operation generalizes to tensor contraction.

2.2 Frobenius algebras

An object X in a compact closed category \mathcal{C} has a Frobenius structure on it if there exist morphisms $\Delta: X \rightarrow X \otimes X$, $\iota: X \rightarrow I$ and $\mu: X \otimes X \rightarrow X$, $\zeta: I \rightarrow X$ satisfying (among other associativity and unit conditions) the following:

$$(\mu \otimes 1_X) \circ (1_X \otimes \Delta) = \Delta \circ \mu = (1_X \otimes \mu) \circ (\Delta \otimes 1_X) \quad (5)$$

In the category **FdVect**, any vector space V with a fixed basis $\{\overrightarrow{v}_i\}_i$ has a commutative special Frobenius algebra over it, explicitly given as follows [4]:

$$\begin{aligned} \Delta &:: \overrightarrow{v}_i \mapsto \overrightarrow{v}_i \otimes \overrightarrow{v}_i & \iota &:: \overrightarrow{v}_i \mapsto 1 \\ \mu &:: \overrightarrow{v}_i \otimes \overrightarrow{v}_j \mapsto \delta_{ij} \overrightarrow{v}_i := \begin{cases} \overrightarrow{v}_i & i = j \\ 0 & i \neq j \end{cases} & \zeta &:: 1 \mapsto \sum_i \overrightarrow{v}_i \end{aligned} \quad (6)$$

For $\overrightarrow{u} \in V, \overline{w} \in V \otimes V$, we have that $\Delta(\overrightarrow{u}) \in V \otimes V$ is a diagonal matrix whose diagonal elements are the weights of \overrightarrow{u} , and $\mu(\overline{w}) \in V$ is a vector consisting only of the diagonal elements of \overline{w} .² For the purposes of this paper we refer to Δ map as the *copying* operation, and to μ map as the *merging* operation.

3 Linguistic uses of the Frobenius operators

The Frobenius operators in **FdVect** as given in Equation 6 adhere to intuitive interpretations that make them important additions to the underlying compact closed setting. In particular, while the standard ε -composition has a transformational effect,³ the μ map can be seen as an alternative form of composition that imposes equal contribution of the operands to the final result; linear-algebraically, while ε -composition is tensor contraction requiring one of the interacting words to be of a higher order than the other, μ -composition takes the form of element-wise multiplication between tensors of the same order:

$$\mu(\overrightarrow{x} \otimes \overrightarrow{y}) = \sum_{ij} \delta_{ij} x_i y_j \overrightarrow{v}_i = \sum_i x_i y_i \overrightarrow{v}_i = \overrightarrow{x} \odot \overrightarrow{y} \quad (7)$$

²In this paper the notation \overrightarrow{v} refers to a vector, while \overline{w} denotes a tensor of order > 1 .

³For example, an intransitive verb is a map $N \rightarrow S$, faithfully encoded as a matrix living in $N \otimes S$, that transforms the input noun into a sentence.

where \vec{v}_i is a basis vector and \odot denotes element-wise vector multiplication. From a linguistic perspective, there are many cases where such an interaction is desirable. Sadrzadeh et al. [12, 13], for example, work on nouns modified by relative clauses (e.g. “The man who likes Mary”) and present a construction for the relative pronoun that results in μ -composing the vector of the noun (‘man’) with the vector of the modifying verb phrase (‘likes Mary’). Similarly, Kartsaklis and Sadrzadeh [6] inject intonational information into a sentence vector by merging the vectors of the theme of the sentence (information that is known to both interlocutors in a conversation) and the rheme of the sentence (information that is new to the addressee). Finally, Coecke and Lewis [3] argue that μ -composition can approximate to a certain extent the meaning of non-compositional compounds, such as ‘pet fish’.

Technically, μ -composition has an intersective effect on the elements of the operands. This is more clear when one uses **Rel** (the category of sets and relations) as the co-domain of our syntax-to-semantics functor \mathcal{F} , instead of **FdVect**. Recall that elements in finite sets can be seen as basis vectors of free modules over the semi-ring of booleans. In this setting, nouns are represented by vectors corresponding to subsets of the universe of discourse, whose i -th component is 1 if the i -th element is included in the specific set and 0 otherwise. Furthermore, a verb becomes a relation, represented by an adjacency matrix in which the (i, j) -th component is 1 if the relation stands for the pair consisting of the i -th element of its domain and the j -th element of its codomain. It is clear that in these cases element-wise multiplication of the vectors/tensors corresponds exactly to the intersection of the involved sets or relations. The transition from the truth-theoretic setting of **Rel** to the real-valued **FdVect**, where a standard distributional model of semantics lives, results in a form of “quantitative” intersection between the components of the vectors that has been proved very effective in a number of standard NLP tasks (see, for example, [9]).

The Frobenius co-multiplication (the “copying” Δ map) has been also proved useful in linguistic applications. Kartsaklis et al. [7] use it in order to restore the functorial relationship between grammar and verb tensors of order lower than that dictated by their type. In a more conceptual use of this operation, the relative pronoun construction of Sadrzadeh et al. essentially copies the vector of the noun from its original position and allows the information to “flow” at the other side of the pronoun and interact with the vector of the verb phrase, providing a means of syntactic movement. The current paper builds on all the above intuitions, and in the following sections these ideas are applied to model coordination in language.

4 Coordination in CCDS

Coordination is perhaps the linguistic aspect in which the notions of merging and copying information find their most natural application. Two coordinated phrases or sentences can be seen as contributing equally to the final outcome; we would expect, for example, the vector of the sentence “John reads and Mary sleeps” to reflect equally the vectors of the two coordinated sentences, “John reads” and “Mary sleeps”. Furthermore, distributivity conditions suggest that parts of the coordinate structure should be copied and interact with each one of the conjuncts separately. For example, it is the case that:

- (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

where the symbol \models denotes entailment. Therefore, merging and copying can be seen as the key processes of coordination, and in this section we will show that an effective use of Frobenius operators in conjunction with the underlying compact closed structure of **FdVect** allows us to model a variety of coordination cases in categorical compositional distributional semantics.

For the analysis that follows we consider the usual ternary rule $X \text{ CONJ } X \rightarrow X$, which states that coordination always takes place between conjuncts of the same type and produces a result that again matches that specific type. In pregroup terms, this is achieved by assigning the type $x^r \cdot x \cdot x^l$ to the conjunction (where x can be an atomic or a compound type), which leads to the following generic derivation:

$$x \cdot (x^r \cdot x \cdot x^l) \cdot x \leq 1 \cdot x \cdot 1 = x \quad (8)$$

Taking $\vec{x}_1, \vec{x}_2 \in \mathcal{F}(x) = X$ to be the semantic representations of the two conjuncts and $\overline{\text{conj}}_X \in X^r \otimes X \otimes X^l$ the coordination tensor,⁴ we translate the above pregroup derivation to **FdVect** as follows:

$$\mathcal{F} \left[(\varepsilon_x^r \cdot 1_x \cdot \varepsilon_x^l) \circ (x \cdot x^r \cdot x \cdot x^l \cdot x) \right] = (\varepsilon_X^r \otimes 1_X \otimes \varepsilon_X^l) \circ (\vec{x}_1 \otimes \overline{\text{conj}}_X \otimes \vec{x}_2) \quad (9)$$

Our main concern is to find a way to transform the composition function $\varepsilon_X^r \otimes 1_X \otimes \varepsilon_X^l$ (which for the moment is solely expressed in terms of ε -maps) to a function that applies μ -composition between the two operands, thus enforces equal contribution of the conjuncts in the final result. The connection between the two ways of composing the conjuncts is made explicit in the following diagram:

$$\begin{array}{ccc} I \xrightarrow{\vec{x}_1 \otimes \vec{x}_2} X \otimes X & \xrightarrow{1_X \otimes \eta_X^r \otimes \eta_X^l \otimes 1_X} & X \otimes X^r \otimes X \otimes X \otimes X^l \otimes X \\ \mu_X \downarrow & & \downarrow 1_X \otimes 1_{X^r} \otimes \mu_X \otimes 1_{X^l} \otimes 1_X \\ X & \xleftarrow{\varepsilon_X^r \otimes 1_X \otimes \varepsilon_X^l} & X \otimes X^r \otimes X \otimes X^l \otimes X \end{array} \quad (10)$$

from which we derive:

$$\begin{aligned} \mu_X \circ (\vec{x}_1 \otimes \vec{x}_2) &= (\varepsilon_X^r \otimes 1_X \otimes \varepsilon_X^l) \circ (1_X \otimes 1_{X^r} \otimes \mu_X \otimes 1_{X^l} \otimes 1_X) \circ (1_X \otimes \eta_X^r \otimes \eta_X^l \otimes 1_X) \circ (\vec{x}_1 \otimes \vec{x}_2) \\ &= (\varepsilon_X^r \otimes 1_X \otimes \varepsilon_X^l) \circ \left(1_X \otimes \left[(1_{X^r} \otimes \mu_X \otimes 1_{X^l}) \circ (\eta_X^r \otimes \eta_X^l) \right] \otimes 1_X \right) \circ (\vec{x}_1 \otimes \vec{x}_2) \\ &= (\varepsilon_X^r \otimes 1_X \otimes \varepsilon_X^l) \circ \left(\vec{x}_1 \otimes \left[(1_{X^r} \otimes \mu_X \otimes 1_{X^l}) \circ (\eta_X^r \otimes \eta_X^l) \right] \otimes \vec{x}_2 \right) \end{aligned} \quad (11)$$

The morphism between the square brackets defines a state in $X^r \otimes X \otimes X^l$, and corresponds exactly to the semantic representation of the coordinator we need (see also Equation 9) in order to translate ε -composition to μ -composition. The coordinator morphism is stated explicitly below:

$$\overline{\text{conj}}_X : I \xrightarrow{\eta_X^r \otimes \eta_X^l} X^r \otimes X \otimes X \otimes X^l \xrightarrow{1_{X^r} \otimes \mu_X \otimes 1_{X^l}} X^r \otimes X \otimes X^l \quad (12)$$

In the following sections we will use the diagrammatic calculus of compact closed categories to demonstrate the application of Equation 12 in a number of coordination examples; the cases of atomic and compound types are treated separately.

4.1 Coordinating atomic types

We start with the simple case of coordinating conjuncts of atomic types, that is, conjuncts whose semantic representation is a vector (nouns) or it can be reduced to a vector by ε -composition (e.g. noun phrases and

⁴Although for a vector space V in **FdVect** it is always the case that $V \cong V^r \cong V^l$, in this paper the adjoints of vector spaces will be explicitly stated for additional clarity.

sentences). In the graphical language of compact closed categories, the composition of a coordination over noun phrases according to Equation 12 takes this form:

$$(13)$$

In this notation, a vector $\vec{v} \in V$ is represented as a state of V , i.e. a morphism $\vec{v} : I \rightarrow V$ (the triangle denotes that the domain is the monoidal unit). Similarly, tensors are states of tensor product spaces; e.g. the coordinator tensor above is a morphism $I \rightarrow N^r \otimes N \otimes N^l$. The cups (\cup) denote ε maps, and the caps (\cap) η maps; straight line segments denote identities. Composing morphisms amounts to connecting outputs to inputs, while the tensor product is juxtaposition. The dot node corresponds to the Frobenius multiplication.

As noted before, the morphism of Equation 12 provides an interface between standard ε -composition (tensor contraction) and μ -composition (element-wise vector multiplication), equipping the underlying compositional framework with an additional layer of flexibility. For the above example we have:

$$(\varepsilon_N^r \otimes 1_N \otimes \varepsilon_N^l) \circ (\overrightarrow{\text{apples}} \otimes \overrightarrow{\text{conj}_N} \otimes \overrightarrow{\text{oranges}}) = \mu(\overrightarrow{\text{apples}} \otimes \overrightarrow{\text{oranges}}) = \overrightarrow{\text{apples}} \odot \overrightarrow{\text{oranges}} \quad (14)$$

The type of a sentence coordinator is $s^r \cdot s \cdot s^l$, leading to a situation very similar to that of the noun phrase case:

$$(15)$$

Linear-algebraically this results in a combination of ε -composition that takes place within the context of each sentence, and μ -composition that merges the two sentences into a single one at the position of the coordinator:

$$(\overrightarrow{\text{men}}^\top \times \overrightarrow{\text{watch}} \times \overrightarrow{\text{football}}) \odot (\overrightarrow{\text{women}}^\top \times \overrightarrow{\text{knit}}) \quad (16)$$

4.2 Coordinating compound types

The simple cases addressed in Section 4.1 make use of the merging Frobenius operator, but they do not include any examples for which duplication of information is necessary. This requirement emerges when one moves to coordination over complex types, which is more interesting and involved. In order to understand how does this work, it would be instructive to examine the way in which the morphism of Equation 12 lifts to complex types. We will use as an example the case of verb phrase coordination. Recall that the pregroup type of a verb phrase is $n^r \cdot s$; that is, something that expects a noun (a subject) from the left in order to return a sentence. The semantic representation of a verb phrase in **FdVect** is a matrix living in $N^r \otimes S \cong N \otimes S$; note however that what follows can be directly generalized to tensors of higher order.

We start by expressing the maps on the compound objects in terms of the atomic objects. For an arbitrary compound object $U \otimes V$, recall that:

$$\begin{array}{ccc}
 I \xrightarrow{\eta_V^r} V^r \otimes V & & I \xrightarrow{\eta_U^l} U \otimes U^l \\
 \eta_{U \otimes V}^r \searrow & \downarrow 1_{V^r} \otimes \eta_U^r \otimes 1_V & \eta_{U \otimes V}^l \searrow & \downarrow 1_U \otimes \eta_V^l \otimes 1_{U^l} \\
 V^r \otimes U^r \otimes U \otimes V = & & U \otimes V \otimes V^l \otimes U^l = \\
 (U \otimes V)^r \otimes (U \otimes V) & & (U \otimes V) \otimes (U \otimes V)^l
 \end{array} \quad (17)$$

which is graphically depicted as follows:

$$\begin{array}{ccc}
 \begin{array}{c} \curvearrowright \\ (U \otimes V)^r \quad (U \otimes V) \end{array} & \mapsto & \begin{array}{c} \curvearrowright \quad \curvearrowright \\ V^r \quad U^r \quad U \quad V \end{array} & & \begin{array}{c} \curvearrowleft \\ (U \otimes V) \quad (U \otimes V)^l \end{array} & \mapsto & \begin{array}{c} \curvearrowleft \quad \curvearrowleft \\ U \quad V \quad V^l \quad U^l \end{array}
 \end{array} \quad (18)$$

Furthermore, it is also the case that the Frobenius operators coherently lift to compound objects; for the μ map, this lifting takes the following form:

$$\begin{array}{ccc}
 (U \otimes V) \otimes (U \otimes V) & \xrightarrow{1_U \otimes \sigma_{U,V} \otimes 1_V} & U \otimes U \otimes V \otimes V \\
 \mu_{U \otimes V} \searrow & & \downarrow \mu_U \otimes \mu_V \\
 & & U \otimes V
 \end{array} \quad (19)$$

and is graphically shown below:

$$\begin{array}{ccc}
 \begin{array}{c} \curvearrowright \\ (U \otimes V) \quad (U \otimes V) \\ \downarrow \\ (U \otimes V) \end{array} & \mapsto & \begin{array}{c} U \quad V \quad U \quad V \\ \downarrow \quad \downarrow \\ U \quad V \end{array}
 \end{array} \quad (20)$$

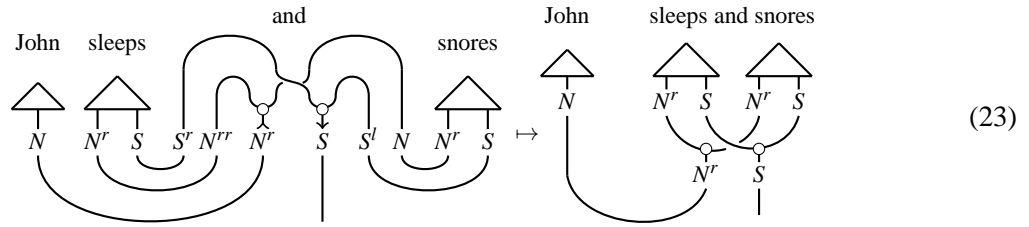
We can now apply the coordination morphism of Equation 12 to the compound type $N^r \otimes S$ and examine how does this translate in terms of the atomic types N^r and S . According to Equations 18 and 20 (and assuming the canonical extensions on identities) we have:

$$\begin{array}{ccc}
 \begin{array}{c} \curvearrowright \quad \curvearrowright \\ (N^r \otimes S)^r \quad (N^r \otimes S)^l \\ \downarrow \\ (N^r \otimes S) \end{array} & \mapsto & \begin{array}{c} \curvearrowright \quad \curvearrowright \\ S^r \quad N^{rr} \quad N^r \quad S \quad S^l \quad N \end{array}
 \end{array} \quad (21)$$

Symbolically, the morphism in its expanded form becomes:

$$\begin{aligned}
 \overline{\text{conj}}_{N^r \otimes S} &= (1_{(N^r \otimes S)^r} \otimes \mu_{N^r \otimes S} \otimes 1_{(N^r \otimes S)^l}) \circ (\eta_{N^r \otimes S}^r \otimes \eta_{N^r \otimes S}^l) \\
 &= (1_{S^r} \otimes 1_{N^{rr}} \otimes \mu_{N^r} \otimes \mu_S \otimes 1_{S^l} \otimes 1_N) \circ (1_{S^r} \otimes \eta_{N^r}^r \otimes \sigma_{N^r, S} \otimes \eta_S^l \otimes 1_N) \circ (\eta_S^r \otimes \eta_N^r)
 \end{aligned} \quad (22)$$

Putting the coordinator in context results in the following interaction:



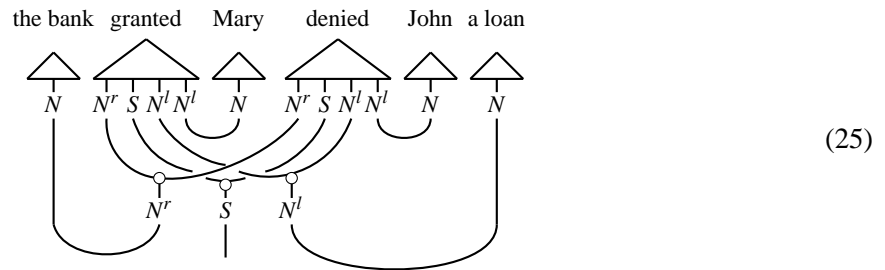
At the left-hand side of the diagram, we observe the following sequence of actions:

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 in order to return a single vector.

Thus the morphism of Equation 12 makes effective use of both notions of copying and merging in order to represent coordination between verb phrases in a way that makes linguistic sense. Note that the simplified diagram at the right-hand side of Equation 23 makes apparent a compact form for the coordinate structure "sleeps and snores". In fact, linear-algebraically this is nothing more than the Hadamard product between the matrices corresponding to the semantic representations of the two verb phrases. In the general case, the Hadamard product between two matrices $A = \sum_{ij} a_{ij} \vec{v}_i \otimes \vec{w}_j$ and $B = \sum_{ij} b_{ij} \vec{v}_i \otimes \vec{w}_j$ in $V \otimes W$ is given by $A \odot B = \sum_{ij} a_{ij} b_{ij} \vec{v}_i \otimes \vec{w}_j$. The element-wise merging of the two conjuncts is consistent with the way coordination is carried out over atomic types and makes possible the following convenient closed form for computing the meaning of the sentence:

$$\overrightarrow{John}^\top \times (\overrightarrow{sleep} \odot \overrightarrow{snores}) \quad (24)$$

Although the above exposition was based on compound types of order 2, the morphism of Equation 12 generalizes coherently to tensors of higher order, providing a first-class coordination object for categorical compositional distributional semantics. In the case of coordination between ditransitive verbs, for example, like in the sentence "The bank granted Mary but denied John a loan", the meaning of the sentence becomes:



Note that the above derivation accounts for copying *both* the subject and the direct object, as required by distributivity. This is translated to the following formula:

$$\overrightarrow{bank}^\top \times \left[(\overrightarrow{grant} \times \overrightarrow{Mary}) \odot (\overrightarrow{deny} \times \overrightarrow{John}) \right] \times \overrightarrow{loan} \quad (26)$$

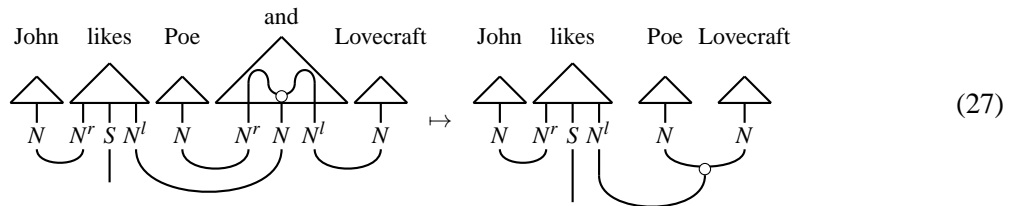
where \overrightarrow{grant} and \overrightarrow{deny} refer to tensors of order 4, and the \odot symbol denotes element-wise multiplication between tensors of order 3.

4.3 Non-standard forms of coordination

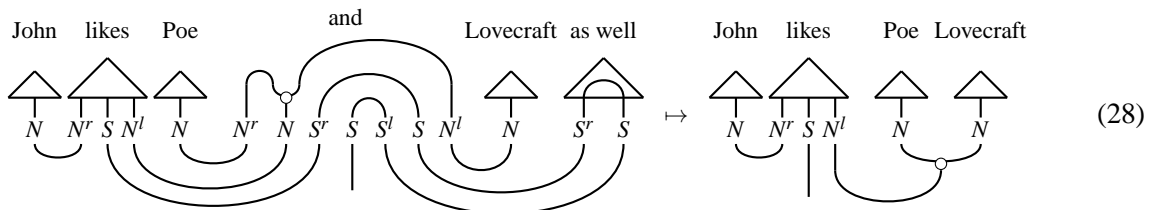
In this section we briefly examine the use of Frobenius operators in non-standard forms of coordination. Consider for example the following sentence:

(3) John likes Poe, and_[John likes] Lovecraft as well

Example 3 exhibits a form of ellipsis related to coordination known as *stripping*; we expect that the meaning of this sentence will be the same with that of the standard form “John likes Poe and Lovecraft”, which is computed below according to Equation 12:



As a special form of coordination, stripping can be addressed by exploiting a special-purpose tensor; we achieve the desired result by defining the internal structure of the coordinator as shown in the following diagram:



Note that the actual coordination step is still carried out by the morphism of Equation 12; however, additional wiring is necessary in order to allow the unhindered flow of information through the left-hand and right-hand parts of the sentence.

5 Conclusion and future work

This paper contributes to the ongoing research on categorical compositional distributional semantics by providing an account of coordination. While the presented ideas cover the most common cases of coordination, language is more complex than that; for example, coordination between text constituents of different syntactic categories (“John works evenings and on weekeends”) constitutes an interesting future direction of research. A proper distinction between conjunction and disjunction cases (which is not pursued in this paper) remains an open problem, as it implies the presence of an underlying logic; to this end, applications of quantum logic seem to provide a promising direction, as the work of Widdows [14] and Van Rijsbergen [11] shows. Finally, the Frobenius machinery and especially the copying operator provides a means for addressing other forms of ellipsis in language, such as verb phrase ellipsis or gapping, that we plan to explore in the future.

Acknowledgements

The author would like to thank the reviewers for their valuable comments, as well as Bob Coecke, Mehrnoosh Sadrzadeh and Matt Purver for useful suggestions and discussions. Support by the AFOSR grant “Algorithmic and Logical Aspects when Computing Meaning” is gratefully acknowledged.

References

- [1] Samson Abramsky & Bob Coecke (2004): *A Categorical Semantics of Quantum Protocols*. In: *19th Annual IEEE Symposium on Logic in Computer Science*, pp. 415–425, doi:10.1109/LICS.2004.1319636.
- [2] B. Coecke, M. Sadrzadeh & S. Clark (2010): *Mathematical Foundations for a Compositional Distributional Model of Meaning*. *Lambek Festschrift. Linguistic Analysis* 36, pp. 345–384.
- [3] Bob Coecke & Martha Lewis (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, pp. 179–192. Springer International Publishing, Cham, doi:10.1007/978-3-319-28675-4_14.
- [4] Bob Coecke & Dusko Pavlovic (2006): *Quantum Measurements Without Sums*. In Goong Chen, Louis Kauffman & Samuel J. Lomonaco, editors: *Mathematics of Quantum Computation and Quantum Technology*, pp. 559–596, doi:10.1201/9781584889007.ch16.
- [5] Dimitri Kartsaklis (2015): *Compositional Distributional Semantics with Compact Closed Categories and Frobenius Algebras*. Ph.D. thesis, University of Oxford.
- [6] Dimitri Kartsaklis & Mehrnoosh Sadrzadeh (2015): *A Frobenius Model of Information Structure in Categorical Compositional Distributional Semantics*. In: *Proceedings of the 14th Meeting on the Mathematics of Language (MoL 2015)*, Association for Computational Linguistics, Chicago, USA, pp. 62–74, doi:10.3115/v1/W15-2306.
- [7] Dimitri Kartsaklis, Mehrnoosh Sadrzadeh & Stephen Pulman (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*, The COLING 2012 Organizing Committee, Mumbai, India, pp. 549–558.
- [8] J. Lambek (2008): *From Word to Sentence*. Polimetrica, Milan.
- [9] Jeff Mitchell & Mirella Lapata (2010): *Composition in Distributional Models of Semantics*. *Cognitive Science* 34(8), pp. 1388–1439, doi:10.1111/j.1551-6709.2010.01106.x.
- [10] Robin Piedeleu, Dimitri Kartsaklis, Bob Coecke & Mehrnoosh Sadrzadeh (2015): *Open System Categorical Quantum Semantics in Natural Language Processing*. In: *Proceedings of the 6th Conference on Algebra and Coalgebra in Computer Science (CALCO)*, Nijmegen, Netherlands, doi:10.4230/LIPIcs.CALCO.2015.270.
- [11] Keith van Rijsbergen (2004): *The Geometry of Information Retrieval*. Cambridge University Press, doi:10.1017/CBO9780511543333.
- [12] M. Sadrzadeh, S. Clark & B. Coecke (2013): *The Frobenius anatomy of word meanings I: subject and object relative pronouns*. *Journal of Logic and Computation* Advance Access, doi:10.1093/logcom/ext044.
- [13] M. Sadrzadeh, S. Clark & B. Coecke (2014): *The Frobenius Anatomy of Word Meanings II: Possessive relative pronouns*. *Journal of Logic and Computation*, doi:10.1093/logcom/exu027.
- [14] D. Widdows (2003): *Orthogonal negation in vector spaces for modelling word-meanings and document retrieval*. In: *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics-Volume 1*, Association for Computational Linguistics, pp. 136–143, doi:10.3115/1075096.1075114.