A low-level treatment of quantifiers in categorical compositional distributional semantics
Extended Abstract

Ondřej Rypáček
University of Oxford
Dept. of Computer Science
ondrej@cs.ox.ac.uk

Mehrnoosh Sadrzadeh
Queen Mary University of London
School of Electronic Eng. and Computer Science
mehrs@eecs.qmul.ac.uk

Abstract. We show how one can formalise quantifiers in the categorical compositional distributional model of meaning. Our model is based on the generalised quantifier theory of Barwise and Cooper. We develop an abstract compact closed semantics and instantiate it in vector spaces and in relations. The former is an example for the distributional corpus-based models of language and the latter for the truth-theoretic ones.

1 Introduction

Vector space models of natural language are based on Firth's hypothesis that meanings of words can be deduced from the contexts in which they often occur [**?**]. One then fixes a context window, of for instance 5 words, and computes frequencies of how many times a words has occurred in this window with other words. These frequencies are often normalised to be better representatives of rare and very common words. These models have been applied to various language processing tasks, for instance thesauri construction [**?**]. Compositional distributional models of meaning extend the vector space models from words to sentences. The categorical such models [**?**,**?**] do so by taking into account the grammatical structure of sentences and the vectors of the words in there. These models have proven successful in practical natural language tasks such as disambiguation, term/definition classification and phrase similarity, for example see [**?**,**?**]. Nevertheless, it has been an open problem how to deal with meanings of logical words such as quantifiers and conjunctives. In this paper, we present preliminary work which aims to show how quantifiers can be deal with using the generalised quantifier theory of Barwise and Cooper [**?**].

According to generalised quantifier theory, the meaning of a sentence with a natural language quantifier Q such as 'Q Sbj Verb' is determined by first taking the intersection of the denotation of Sbj with the denotation of subjects of the Verb, then checking if the denotation of $Q(\text{Sbj})$ is an element of this set. The denotation of Q is specified separately, for example, for $Q = \exists$, it is the set of non-empty subsets of the universe, for $Q = 2$ it is the set of subsets of the universe that have exactly two elements and so on. As a result, and for example, the meaning of a sentence "some men sleep" becomes true if the set of men who sleep is non empty.

In what follows, we work in the categorical compositional distributional model of [**?**]. We first present a brief preliminary account of compact closed categories and Frobenius algebras over them and review how vector spaces and relations provide instances. Then, we develop a compact closed categorical semantic for quantifiers, in terms of diagrams and morphisms of compact closed categories. We present two concrete interpretations for this abstract setting: relations and vector spaces. The former is the basis for a truth-theoretic model and the latter works for a corpus-base model of language.

Our future work includes formalising this rather low-level treatment in the setting of categorical logic, where quantifiers are proven to be adjoints to substitution. Lack of much structure in vector spaces (and compact closed categories in general) and in particular lack of existence of pull-backs will be obvious obstacles. We also aim to experiment with this model on corpus-based datasets and tasks.

2 Preliminaries

This section briefly reviews compact closed categories and Frobenius algebras. For a formal presentation, see [**?**,**?**]. A compact closed category, $\mathcal{C}$, has objects A, B; morphisms $f\colon A \to B$; a monoidal tensor $A \otimes B$ that has a unit I; and for each object A two objects A^r and A^l together with the following morphisms:

$$A \otimes A^r \xrightarrow{\epsilon^r_A} I \xrightarrow{\eta^r_A} A^r \otimes A \qquad\qquad A^l \otimes A \xrightarrow{\epsilon^l_A} I \xrightarrow{\eta^l_A} A \otimes A^l$$

These morphisms satisfy the following equalities, sometimes referred to as the *yanking* equalities, where 1_A is the identity morphism on object A:

$$(1_A \otimes \epsilon^l_A) \circ (\eta^l_A \otimes 1_A) = 1_A \qquad\qquad (\epsilon^r_A \otimes 1_A) \circ (1_A \otimes \eta^r_A) = 1_A$$
$$(\epsilon^l_A \otimes 1_A) \circ (1_{A^l} \otimes \eta^l_A) = 1_{A^l} \qquad\qquad (1_{A^r} \otimes \epsilon^r_A) \circ (\eta^r_A \otimes 1_{A^r}) = 1_{A^r}$$

These express the fact the A^l and A^r are the left and right adjoints, respectively, of A in the 1-object bicategory whose 1-cells are objects of $\mathcal{C}$.

A Frobenius algebra in a monoidal category $(\mathcal{C}, \otimes, I)$ is a tuple $(X, \delta, \iota, \mu, \zeta)$ where, for X an object of $\mathcal{C}$, the triple (X, δ, ι) is an internal comonoid; i.e. the following are coassociative and counital morphisms of $\mathcal{C}$:

$$\delta\colon X \to X \otimes X \qquad\qquad\qquad \iota\colon X \to I$$

Moreover (X, μ, ζ) is an internal monoid; i.e. the following are associative and unital morphisms:

$$\mu\colon X \otimes X \to X \qquad\qquad\qquad \zeta\colon I \to X$$

And finally the δ and μ morphisms satisfy the following *Frobenius condition*:

$$(\mu \otimes 1_X) \circ (1_X \otimes \delta) \;=\; \delta \circ \mu \;=\; (1_X \otimes \mu) \circ (\delta \otimes 1_X)$$

Informally, the comultiplication δ dispatches the information contained in one object into two objects, and the multiplication μ unifies the information of two objects into one.

Finite Dimensional Vector Spaces. These structures together with linear maps form a compact closed category, which we refer to as FdVect. Finite dimensional vector spaces V, W are objects of this category; linear maps $f\colon V \to W$ are its morphisms with composition being the composition of linear maps. The tensor product $V \otimes W$ is the linear algebraic tensor product, whose unit is the scalar field of vector spaces; in our case this is the field of reals $\mathbb{R}$. Here, there is a natural isomorphism $V \otimes W \cong W \otimes V$. As a result of the symmetry of the tensor, the two adjoints reduce to one and we obtain the isomorphism $V^l \cong V^r \cong V^*$, where V^* is the dual space of V. When the basis vectors of the vector spaces are fixed, it is further the case that we have $V^* \cong V$.

Given a basis $\{r_i\}_i$ for a vector space V, the epsilon maps are given by the inner product extended by linearity; i.e. we have:

$$\epsilon^l = \epsilon^r \colon V \otimes V \to \mathbb{R} \quad \text{given by} \quad \sum_{ij} c_{ij}\; \psi_i \otimes \phi_j \quad\mapsto\quad \sum_{ij} c_{ij} \langle \psi_i \mid \phi_j \rangle$$

Similarly, eta maps are defined as follows:

$$\eta^l = \eta^r \colon \mathbb{R} \to V \otimes V \quad \text{given by} \quad 1 \;\mapsto\; \sum_i r_i \otimes r_i$$

Any vector space V with a fixed basis $\{\overrightarrow{v_i}\}_i$ has a Frobenius algebra over it, explicitly given as follows, where δ_{ij} is the Kronecker delta.

$$\begin{aligned}
\delta\colon V \to V \otimes V \quad & \text{given by} \quad \overrightarrow{v_i} \mapsto \overrightarrow{v_i} \otimes \overrightarrow{v_i} \\
\mu\colon V \otimes V \to V \quad & \text{given by} \quad \overrightarrow{v_i} \otimes \overrightarrow{v_j} \mapsto \delta_{ij} \overrightarrow{v_i} \\
\iota\colon V \to \mathbb{R} \quad & \text{given by} \quad \overrightarrow{v_i} \mapsto 1 \\
\zeta\colon \mathbb{R} \to V \quad & \text{given by} \quad 1 \mapsto \sum_i \overrightarrow{v_i}
\end{aligned}$$

Relations. Another important example of a compact closed category is Rel, the cateogry of sets and relations. Here, $\otimes$ is cartesian product with the singleton set as its unit $I = \{\star\}$, and * is identity on objects. Closure reduces to the fact that a relation between sets $A \times B$ and C is equivalently a relation between A and $B \times C$. Given a set S with elements $s_i, s_j \in S$, the epsilon and eta maps are given as follows:

$$\begin{aligned}
\epsilon^l = \epsilon^r\ : S \times S \to \{\star\} \quad & \text{given by} \quad \{((s_i, s_j), \star) \mid s_i, s_j \in S, s_i = s_j\} \\
\eta^l = \eta^r\ : \{\star\} \to S \times S \quad & \text{given by} \quad \{(\star, (s_i, s_j)) \mid s_i, s_j \in S, s_i = s_j\}
\end{aligned}$$

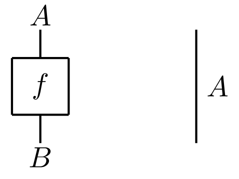
Every object in Rel has a Frobenius algebra over it given by the diagonal and codiagonal relations, as described below:

$$\begin{aligned}
\delta\ : S \to S \times S \quad & \text{given by} \quad \{(s_i, (s_j, s_k)) \mid s_i, s_j, s_k \in S, s_i = s_j = s_k\} \\
\mu\ : S \times S \to S \quad & \text{given by} \quad \{(s_i, s_j), s_k) \mid s_i, s_j, s_k \in S, s_i = s_j = s_k\} \\
\iota\ : S \to \{\star\} \quad & \text{given by} \quad \{(s_i, \star) \mid s_i \in S\} \\
\zeta\ : \{\star\} \to S \quad & \text{given by} \quad \{(\star, s_i) \mid s_i \in S\}
\end{aligned}$$

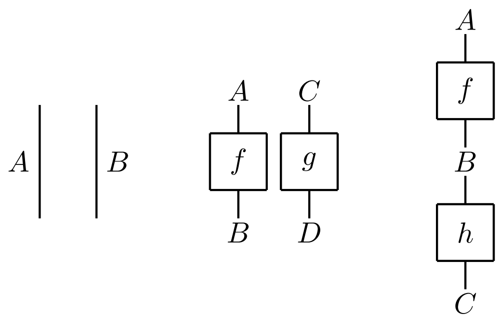
For the details of verifying that for each of the two examples above, the corresponding conditions hold see [**?**].

3 String Diagrams

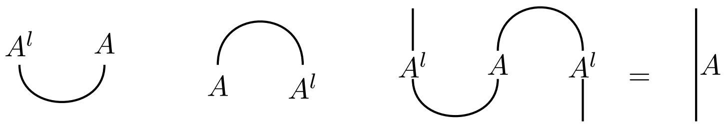
The framework of compact closed categories and Frobenius algebras comes with a complete diagrammatic calculus that visualises derivations, and which also simplifies the categorical and vector space computations. Morphisms are depicted by boxes and objects by lines, representing their identity morphisms. For instance a morphism $f\colon A \to B$, and an object A with the identity arrow $1_A\colon A \to A$, are depicted as follows:



The tensor products of the objects and morphisms are depicted by juxtaposing their diagrams side by side, whereas compositions of morphisms are depicted by putting one on top of the other; for instance the object $A \otimes B$, and the morphisms $f \otimes g$ and $f \circ h$, for $f\colon A \to B, g\colon C \to D$, and $h\colon B \to C$, are depicted as follows:

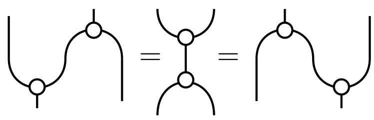


The ϵ maps are depicted by cups, η maps by caps, and yanking by their composition and straightening of the strings. For instance, the diagrams for $\epsilon^l \colon A^l \otimes A \to I$, $\eta \colon I \to A \otimes A^l$ and $(\epsilon^l \otimes 1_A) \circ (1_A \otimes \eta^l) = 1_A$ are as follows:

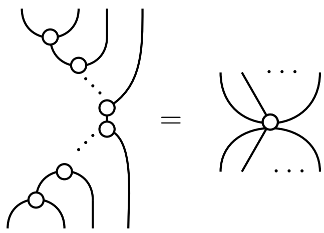


As for Frobenius algebras, the diagrams for the monoid and comonoid morphisms are as follows:

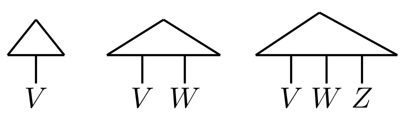
with the Frobenius condition being depicted as:



The defining axioms guarantee that any picture depicting a Frobenius computation can be reduced to a normal form that only depends on the number of input and output strings of the nodes, independent of the topology. These normal forms can be simplified to so-called 'spiders':

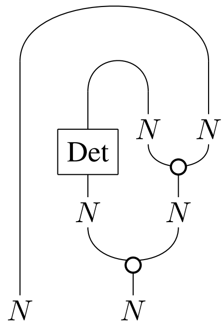


In the category FdVect, apart from spaces V, W, which are objects of the category, we also have vectors $\overrightarrow{v}, \overrightarrow{w}$. These are depicted by their representing morphisms and as triangles with a number of strings emanating from them. The number of strings of a triangle denote the tensor rank of the vector; for instance, the diagrams for $\overrightarrow{v} \in V$, $\overrightarrow{v'} \in V \otimes W$, and $\overrightarrow{v''} \in V \otimes W \otimes Z$ are as follows:



4 Diagrammatic Compact Closed Semantics

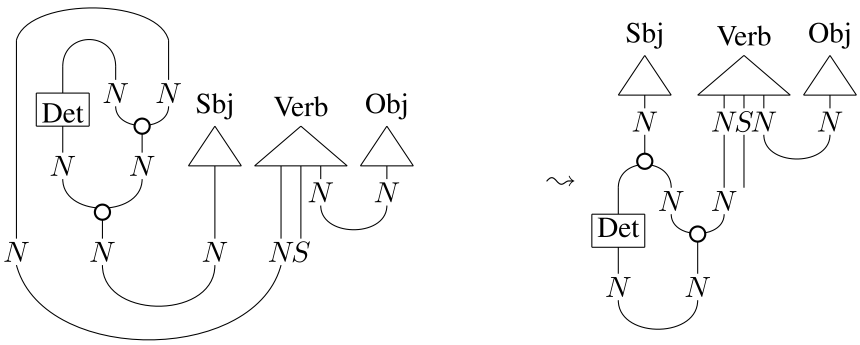
Following the terminology and notation of [?], given a phrase containing a quantifier followed by a noun, that is 'Q noun', we call 'Q' a determiner and the phrase 'Q noun' a quantified phrase.Hence, a quantified phrase is a noun phrase which is created by the application of a determiner to a noun phrase. We suggest the following diagrammatic semantics for a determiner Det:



It corresponds to the following compact closed categorical morphism:

$$(1_N \otimes \delta_N) \circ (1_N \otimes Det \otimes \mu_N) \circ (1_N \otimes \eta_N \otimes 1_N) \circ \eta_N$$

The meaning of the sentence with a quantified phrase in a subject position and its normalised form are as follows:



The symbolic representation of the normal form diagram is as follows:

$$(\epsilon_N \otimes 1_S) \circ (Det \otimes \mu_N \otimes 1_S) \circ (\delta_N \otimes 1_{N \otimes S})(\overrightarrow{\text{Sbj}} \otimes \overrightarrow{\text{Verb}} \otimes \overrightarrow{\text{Obj}})$$

The intuitive justification is that the determiner first makes a copy of the subject (via the Frobenius δ map), so now we have two copies of the subject. One of these is being unified with the subject argument of the verb (via the Frobenius μ map). In set-theoretic terms this is the intersection of the interpretations of subject and subjects-of-verb. The other copy is being inputted to the determiner map Det and will produce a modified noun based on the meaning of the determiner. The last step is the application of the unification to the output of Det. Set theoretically, this step will decide whether the intersection of the subject-of-verb and the noun belongs to the interpretation of the quantified noun.

The diagrams, morphisms, and intuitions for a quantified phrase in an object position are identical.

5 Vector Space Interpretation

In a concrete vector space model, built from a corpus using distributional methods, we assume that vector meaning of the subject is $\sum_i C_i \overrightarrow{n}_i \in N$ and the linear map corresponding to the verb is $\sum_{jk} C_{jk} \overrightarrow{n}_j \otimes \overrightarrow{s}_k \in N \otimes S$. For $\overrightarrow{n}_i$ a basis vector of N, we define the map Det as follows:

$$Det(\overrightarrow{n}_i) = \varPhi\{\overrightarrow{w} \in N \mid d(\overrightarrow{n}_i, \overrightarrow{w}) = \alpha\} \tag{1}$$

where we have:

- ϕ is a linear average function such as the arithmetic or weighted mean.
- α indicates how close $\overrightarrow{w}$ is to the $\overrightarrow{n}_i$ and depends on the quantifier expressed by Det.

The intuitive reading of the above is that Det of a word $\overrightarrow{n}_i$ is a linear combination, e.g. average, of all the words that are α-close to $\overrightarrow{n}_i$. In other words, the average of all the words whose distance from $\overrightarrow{n}_i$ is α. For instance, if Det is 'few', then α is a small number (closer to 0 than to 1), indicating that we are taking the average of vectors that are not so close to $\overrightarrow{n}_i$. If Det is 'most', then α will be a large number (closer to 1 than to 0), indicating that we are taking the average of vectors that are close to $\overrightarrow{n}_i$. The distance α can be learnt from a corpus using a relevant task. This will extend to any other (non-basis) word by linearity.

The underlying idea here is that the quantitative way of quantifying in set-theoretic models, which depends on the cardinality of the quantified sets, is now transformed into a geometric way of quantifying where the meaning of the quantified phrase depends on its geometric distance with other words. Hence, a quantified phrase such as 'few cats' returns a representative noun (obtained by taking the average of all such nouns) that is far from vector of 'cat' in the semantic space. This representative noun shares 'few' properties with 'cat'. A quantified phrase such as 'most cats' returns a representative noun that is close the the vector of 'cat' and stands for a noun that shares 'most' of the properties of 'cat'.

With this instantiation, the meaning of "Q Sbj Verb" is obtained by computing the following:

$$(\epsilon_N \otimes 1_S) \circ (Det \otimes \mu_N \otimes 1_S) \circ (\delta_N \otimes 1_{N\otimes S}) \left(\overrightarrow{\text{Sbj}} \otimes \overrightarrow{\text{Verb}}\right)$$

In the first step of computation we have:

$$(\delta_N \otimes 1_{N\otimes S})\Big(\sum_i C_i \overrightarrow{n}_i \otimes \sum_{jk} C_{jk} \overrightarrow{n}_j \otimes \overrightarrow{s}_k\Big) = (\sum_i C_i \overrightarrow{n}_i \otimes \overrightarrow{n}_i) \otimes (\sum_{jk} C_{jk} \overrightarrow{n}_j \otimes \overrightarrow{s}_k)$$

In the second step we obtain:

$$\begin{aligned}(Det \otimes \mu_N \otimes 1_S)\Big((\sum_i C_i \overrightarrow{n}_i \otimes \overrightarrow{n}_i) \otimes (\sum_{jk} C_{jk} \overrightarrow{n}_j \otimes \overrightarrow{s}_k)\Big) &= \sum_{ijk} C_i C_{jk} Det(\overrightarrow{n}_i) \otimes \mu(\overrightarrow{n}_i \otimes \overrightarrow{n}_j) \otimes \overrightarrow{s}_k \\ &= \sum_{ijk} C_i C_{jk} Det(\overrightarrow{n}_i) \otimes \delta_{ij} \overrightarrow{n}_i \otimes \overrightarrow{s}_k\end{aligned}$$

The final step is as follows:

$$(\epsilon_N \otimes 1_S)\Big(\sum_{ijk} C_i C_{jk} Det(\overrightarrow{n}_i) \otimes \delta_{ij} \overrightarrow{n}_i \otimes \overrightarrow{s}_k\Big) = \sum_{ijk} C_i C_{jk} \langle Det(\overrightarrow{n}_i) \mid \delta_{ij} \overrightarrow{n}_i \rangle \overrightarrow{s}_k$$

Example. As a distributional example, take N to be the two dimensional space with the basis $\{\overrightarrow{n}_1, \overrightarrow{n}_2\}$ and S be the two dimensional space with the basis $\{\overrightarrow{s}_1, \overrightarrow{s}_2\}$. Suppose the linear expansion of $\overrightarrow{\text{Sbj}}$ in this space is $C_1\overrightarrow{n}_1 + C_2\overrightarrow{n}_2$ and the linear expansion of $\overrightarrow{\text{Verb}}$ is $C_{11}(\overrightarrow{n}_1 \otimes \overrightarrow{s}_1) + C_{12}(\overrightarrow{n}_1 \otimes \overrightarrow{s}_2) + C_{21}(\overrightarrow{n}_2 \otimes \overrightarrow{s}_1) + C_{22}(\overrightarrow{n}_2 \otimes \overrightarrow{s}_2)$. Suppose further the following for the interpretation of the determiner:

$$Det(\overrightarrow{\text{Sbj}}) = Det(C_1\overrightarrow{n}_1 + C_2\overrightarrow{n}_2) = Det(C_1\overrightarrow{n}_1) + Det(C_2\overrightarrow{n}_2) = C'_1\overrightarrow{n}_1 + C'_2\overrightarrow{n}_2 \tag{2}$$

Then the result of the first step of the computation of a meaning vector for the sentence 'Q Sbj Verb' is:

$$(C_1\overrightarrow{n}_1 + C_2\overrightarrow{n}_2) \otimes (C_{11}\overrightarrow{n}_1 \otimes \overrightarrow{s}_1 + C_{12}\overrightarrow{n}_1 \otimes \overrightarrow{s}_2 + C_{21}\overrightarrow{n}_2 \otimes \overrightarrow{s}_1 + C_{22}\overrightarrow{n}_2 \otimes \overrightarrow{s}_2)$$

In the second step of computation we obtain:

$$C_1C_{11}Det(\overrightarrow{n}_1)\overrightarrow{n}_1\otimes\overrightarrow{s}_1 + C_1C_{12}Det(\overrightarrow{n}_1)\overrightarrow{n}_1\otimes\overrightarrow{s}_2 + C_2C_{21}Det(\overrightarrow{n}_2)\overrightarrow{n}_2\otimes\overrightarrow{s}_1 + C_2C_{22}Det(\overrightarrow{n}_2)\overrightarrow{n}_2\otimes\overrightarrow{s}_2$$

Since Det is a linear map, the above is equal to the following:

$$Det(C_1\overrightarrow{n}_1)(C_{11}\overrightarrow{n}_1 \otimes \overrightarrow{s}_1 + C_{12}\overrightarrow{n}_1 \otimes \overrightarrow{s}_2) + Det(C_2\overrightarrow{n}_2)(C_{21}\overrightarrow{n}_2 \otimes \overrightarrow{s}_1 + C_{22}\overrightarrow{n}_2 \otimes \overrightarrow{s}_2)$$

According to the expansion of the assumption in equation 2, the above is equivalent to the following:

$$(C'_1\overrightarrow{n}_1)(C_{11}\overrightarrow{n}_1 \otimes \overrightarrow{s}_1 + C_{12}\overrightarrow{n}_1 \otimes \overrightarrow{s}_2) + (C'_2\overrightarrow{n}_2)(C_{21}\overrightarrow{n}_2 \otimes \overrightarrow{s}_1 + C_{22}\overrightarrow{n}_2 \otimes \overrightarrow{s}_2)$$

Substituting this in the last step of the computation provide us with the following vector in S:

$$C'_1C_{11}\overrightarrow{s}_1 + C'_1C_{12}\overrightarrow{s}_2 + C'_2C_{21}\overrightarrow{s}_2 + C'_2C_{22}\overrightarrow{s}_2 \tag{3}$$

Small Corpus-Based Witness. A large scale experimentation for this model constitutes work in progress. For the sake of providing intuitions for the above symbolic constructions, we provide a couple of corpus-based witnesses here. In the distributional models, the most natural instantiation of the distance d in equation 1 is the co-occurrence distance. For a noun '*n*' and determiners '*few*' and '*most*', we define these as follows:

few(*n*) = $Avg\{$nouns that co-occurred with *n* few times$\}$
most(*n*) = $Avg\{$nouns that co-occurred with *n* most times$\}$

In this case, a sample query from the online *Reuter News Corpus*, with at most 100 outputs per query, provides the following instantiations:

few(*dogs*) = $Avg\{$bike, drum, snails$\}$
most(*dogs*) = $Avg\{$cats, pets,birds, puppies$\}$
few(*cats*) = $Avg\{$fluid, needle, care$\}$
most(*cats*) = $Avg\{$dogs, birds, rats, feces$\}$

A cosine-based similarity measure over this corpus results in the fact that any of the words in the 'most(*n*)' set are more similar to '*n*' than any of the words in the 'few(*n*)' set. This is indeed because the words in the former set are geometrically closer to '*n*' than the words in the latter set, since they have co-occurred with them more. This is the first advantage of our model over a distributional model, where words such as 'few' and 'most' are treated as noise and hence meanings of phrase such as 'few cats', 'most cats', and 'cats' become identical (and similarly for any other noun). Moreover, in our setting we can establish that 'most cats' and 'most dogs' have similar meanings, because of the over lap of their determiner sets. A larger corpus and a more thorough statistical analysis will let us achieve more, that for instance, 'few cats' and 'few dogs' also have similar meanings.

At the level of sentence meanings, compositional distributional models do not interpret determiners (e.g. see the model of [**?**]). As a result, meanings of sentences such as 'cats sleep', 'most cats sleep' and 'few cats sleep' will become identical; meanings of sentences 'most cats sleep' and 'few dogs snooze' become very close, since 'cats' and 'dogs' often occur in the same context and so do 'sleep' and 'snooze'. In our setting, equation 3 tells us that these sentences have different meanings, since their quantified subjects have different meanings. To see this, take $\overrightarrow{\text{cats}} = C_1\overrightarrow{n}_1 + C_2\overrightarrow{n}_2$, where as $few(\text{cats}) = C_1'\overrightarrow{n}_1 + C_2'\overrightarrow{n}_2$ and $most(\text{cats}) = C_1''\overrightarrow{n}_1 + C_2''\overrightarrow{n}_2$. Instantiating these in equation 3 provides us with the following three different vectors:

$$\begin{aligned}
\overrightarrow{\text{cats sleep}} &= C_1C_{11}\overrightarrow{s}_1 + C_1C_{12}\overrightarrow{s}_2 + C_2C_{21}\overrightarrow{s}_2 + C_2C_{22}\overrightarrow{s}_2 \\
\overrightarrow{\text{few cats sleep}} &= C_1'C_{11}\overrightarrow{s}_1 + C_1'C_{12}\overrightarrow{s}_2 + C_2'C_{21}\overrightarrow{s}_2 + C_2'C_{22}\overrightarrow{s}_2 \\
\overrightarrow{\text{most cats sleep}} &= C_1''C_{11}\overrightarrow{s}_1 + C_1''C_{12}\overrightarrow{s}_2 + C_2''C_{21}\overrightarrow{s}_2 + C_2''C_{22}\overrightarrow{s}_2
\end{aligned}$$

On the other hand, we have that 'most cats sleep' and 'most dogs snooze' have close meanings, one which is close to 'pets sleep'. This is because, their quantified subjects and their verbs have similar meanings, that is we have:

$$\begin{cases} most(\overrightarrow{\text{dogs}}) \sim most(\overrightarrow{\text{cats}}) \sim \overrightarrow{\text{pets}} \\ \overrightarrow{\text{snooze}} \sim \overrightarrow{\text{sleep}} \end{cases} \quad \Longrightarrow \quad \text{most cats sleep} \sim \text{most dogs snooze} \sim \text{pets sleep}$$

At the same time, 'few cats sleep' and 'most dogs snooze' have a less-close meaning, since their quantified subjects have different meanings, that is:

$$most(\overrightarrow{\text{dogs}}) \not\sim few(\overrightarrow{\text{cats}}) \implies \overrightarrow{\text{most dogs snooze}} \not\sim \overrightarrow{\text{few cats sleep}}$$

6 Truth Theoretic Interpretation

For this part, we work in the category Rel of sets and relations. We take U to be a universal reference set of individuals and take N to be the set of all subsets of U, denoted by $\mathcal{P}(U)$. A common noun is modelled by the set of all subsets of its individuals. We take S to be the singleton set $I = \{\star\}$, that is the unit of tensor product in Rel. A verb is the set of all subsets of a relation (corresponding to its predicate). For an intransitive verb, this relation is on the set $N \times S$; since we have $N \times S \cong N$, each relation corresponds to a subset of N. For a transitive verb, it is a relation on the set $N \times S \times N \cong N \times N$.

The map Det sends a subset of individuals to a set of its subsets exactly in the same way as defined by [**?**]. For example, for Det = 'two', the output is the set of subsets of individuals whose elements have cardinality exactly two; for Det = 'some', the output is the set of subsets of individuals and so on.

The truth-theoretic meaning of the sentence "Det Sbj Verb" is obtained by computing a simplified version of the morphism developed in section 4 in category Rel. The simplification is because we have $S = I$ and hence the morphisms that are applied to object S can be dropped.

Suppose we have:

$$\begin{aligned}
\overline{[\![\text{Sbj}]\!]} &= \{A_i \mid A_i \in \mathcal{P}([\![\text{Sbj}]\!])\} \\
\overline{[\![\text{Verb}]\!]} &= \{B_j \mid B_j \in \mathcal{P}([\![\text{Verb}]\!])\} \\
Det\left(\overline{[\![\text{Sbj}]\!]}\right) &= \{D_k \mid D_k \in Det([\![\text{Sbj}]\!])\}
\end{aligned}$$

where $[\![\text{Sbj}]\!] \subseteq U$ and $[\![\text{Verb}]\!] \subseteq U$ are the set-theoretic meanings of "Sbj" and "Verb", and $Det(S)$ is the same as in the generalised quantifier approach.

The compact closed meaning of a quantified sentence is computed in three steps as follows. In the first step, we obtain:

$$(\delta_N \otimes 1_N)\Big(\overline{[\![\text{Sbj}]\!]} \otimes \overline{[\![\text{Verb}]\!]}\Big) = \{(A_i, A_i) \mid A_i \in \mathcal{P}([\![\text{Sbj}]\!])\} \otimes \{B_j \mid B_j \in \mathcal{P}([\![\text{Verb}]\!])\}$$

In the second step, we obtain:

$$\begin{gathered}(Det \otimes \mu_N)\Big(\{(A_i, A_i) \mid A_i \in \mathcal{P}([\![\text{Sbj}]\!])\} \otimes \{B_j \mid B_j \in \mathcal{P}([\![\text{Verb}]\!])\}\Big) = \\ Det\Big(\{A_i \mid A_i \in \mathcal{P}([\![\text{Sbj}]\!])\}\Big) \otimes \{A_i \mid A_i = B_j, A_i \in \mathcal{P}([\![\text{Sbj}]\!]), B_j \in \mathcal{P}([\![\text{Verb}]\!])\} = \\ \{D_k \mid D_k \in Det([\![\text{Sbj}]\!])\} \otimes \{A_i \mid A_i = B_j, A_i \in \mathcal{P}([\![\text{Sbj}]\!]), B_j \in \mathcal{P}([\![\text{Verb}]\!])\}\end{gathered}$$

In the final step, we obtain:

$$\begin{gathered}\epsilon\Big(\{D_k \mid D_k \in Det([\![\text{Sbj}]\!])\} \otimes \{A_i \mid A_i = B_j, A_i \in \mathcal{P}([\![\text{Sbj}]\!]), B_j \in \mathcal{P}([\![\text{Verb}]\!])\}\Big) = \\ \{\star \mid D_k = A_i, D_k \in Det([\![\text{Sbj}]\!]), A_i = B_j, A_i \in \mathcal{P}([\![\text{Sbj}]\!]), B_j \in \mathcal{P}([\![\text{Verb}]\!])\}\end{gathered}$$

Recalling that, as shown in [**?**,**?**], the Frobenius μ map is the analog of set-theoretic intersection and the compact closed epsilon map is the analog of set-theoretic application, it is not hard to show that the truth-theoretic interpretation of the compact closed semantics of quantified sentences provides us with the same meaning as their generalised quantifier semantics. In this section we make this formal as follows.

Definition 1. *The compact closed meaning of the sentence "Q Sbj Verb" is true in Rel if and only if either $Det([\![Sbj]\!]) \neq \{\emptyset\}$ and we have*

$$\epsilon \circ (Det \otimes \mu) \circ (\sigma \otimes 1_N)\Big(\overline{[\![Sbj]\!]} \otimes \overline{[\![Verb]\!]}\Big) = \{\star\}$$

or $Det([\![Sbj]\!]) = \{\emptyset\}$, and we have the above as well as the following:

$$\mu\Big(\overline{[\![Sbj]\!]} \otimes \overline{[\![Verb]\!]}\Big) = \{\emptyset\}$$

Proposition 1. *The compact closed meaning of a quantified sentence computed in Rel is true if and only if its generalised quantifier meaning is true.*

Proof. For the right to left direction, suppose the generalised quantifier meaning of the sentence is true, that is $[\![\text{Sbj}]\!] \cap [\![\text{Verb}]\!] \in Det([\![\text{Sbj}]\!]$ and consider the case when $Det([\![Sbj]\!]) \neq \{\emptyset\}$. We have to show that $\epsilon \circ (Det \otimes \mu) \circ (\sigma \otimes 1_N)\Big(\overline{[\![\text{Sbj}]\!]} \otimes \overline{[\![\text{Verb}]\!]}\Big) = \{\star\}$. For this, we need to show that there is a set G equal to $D_k = A_i = B_j$ such that G is in $Det([\![\text{Sbj}]\!])$ and $\mathcal{P}([\![\text{Sbj}]\!])$ and $\mathcal{P}([\![\text{Verb}]\!])$. Take G to be $[\![\text{Sbj}]\!] \cap [\![\text{Verb}]\!]$. Then, it is a subset of $[\![\text{Sbj}]\!]$, hence an element of $\mathcal{P}([\![\text{Sbj}]\!])$, a subset of $[\![\text{Verb}]\!]$, hence an element of $\mathcal{P}([\![\text{Verb}]\!])$, and an element of $Det([\![\text{Sbj}]\!]$.

Now consider the case where $Det([\![Sbj]\!]) = \{\emptyset\}$, the above argument still holds and we have that $G = \emptyset$. It remains to check whether $\mu\Big(\overline{[\![\text{Sbj}]\!]} \otimes \overline{[\![\text{Verb}]\!]}\Big) = \{\emptyset\}$. This is indeed the case, since $D_k = A_i = B_j = \emptyset$, hence $\mu\Big(\overline{[\![\text{Sbj}]\!]} \otimes \overline{[\![\text{Verb}]\!]}\Big) = \{A_i \mid A_i = B_j, A_i \in \mathcal{P}([\![\text{Sbj}]\!]), B_j \in \mathcal{P}([\![\text{Verb}]\!])\} = \{\emptyset\}$.

For the left to right direction, suppose the compact closed meaning of the sentence is true in Rel, and consider the case when $Det([\![Sbj]\!]) \neq \{\emptyset\}$. Then we have subsets $D_k = A_i = B_j$ such that $D_k \in Det([\![Sbj]\!]), A_i \in \mathcal{P}([\![Sbj]\!])$, and $B_j \in \mathcal{P}([\![Verb]\!])$. Pick an arbitrary such subset, e.g. G, such that it equal to $D_k = A_i = B_j$; we have that $G \in \mathcal{P}([\![Sbj]\!])$, hence $G \subseteq [\![Sbj]\!]$; and that $G \in \mathcal{P}([\![Verb]\!])$, hence $G \subseteq [\![Verb]\!]$. It thus follows that $G \subseteq [\![Sbj]\!] \cap [\![Verb]\!]$. At the same time $G \in Det([\![Sbj]\!])$, hence the generalised quantifier meaning is also true.

For the case when $Det([\![Sbj]\!]) = \{\emptyset\}$, we still have the above, so the generalised quantifier meaning of the sentence remains true. In this case we moreover have that $\mu\Big(\overline{[\![\text{Sbj}]\!]} \otimes \overline{[\![\text{Verb}]\!]}\Big) = \{\emptyset\}$, which implies that $G = A_i = B_j = D_k = \emptyset$. If the second condition did not hold, we would have that $\mu\Big(\overline{[\![\text{Sbj}]\!]} \otimes \overline{[\![\text{Verb}]\!]}\Big) = \{A_i \mid A_i = B_j, A_i \in \mathcal{P}([\![\text{Sbj}]\!]), B_j \in \mathcal{P}([\![\text{Verb}]\!])\}$ is equal to a set of subsets, at least one of which is non empty, e.g. $\{X_1, X_2, \cdots\}$, where for X_w it holds that $X_w \neq \emptyset$, and $X_w \subseteq [\![Sbj]\!]$ and $X_w \subseteq [\![Verb]\!]$. It would then follows that $X_w = [\![Sbj]\!] \cap [\![Verb]\!] \neq \emptyset$; whereas $\text{Det}([\![Sbj]\!]) = \{\emptyset\}$, hence the generalised quantifier meaning of the sentence would become false.

Example. As a truth-theoretic example, suppose we have two male individuals m_1, m_2 and a cat individual c_1. Suppose further that the verb ‘sneeze’ applies to individuals m_1 and c_1. Hence, we have the following interpretations for the lemmas of words “man”, “cat”, and “sneeze”:

$$\overline{[\![\text{men}]\!]} = \{\emptyset, \{m_1\}, \{m_2\}, \{m_1, m_2\}\} \qquad \overline{[\![\text{cat}]\!]} = \{\emptyset, \{c_1\}\} \qquad \overline{[\![\text{sneeze}]\!]} = \{\emptyset, \{m_1\}, \{c_1\}, \{m_1, c_1\}\}$$

Consider the following quantified phrases and their interpretations:

$$Some\Big(\overline{[\![\text{men}]\!]}\Big) = \{\{m_1\}, \{m_2\}, \{m_1, m_2\}\} \qquad One\Big(\overline{[\![\text{man}]\!]}\Big) = \{\{m_1\}, \{m_2\}\} \qquad No\Big(\overline{[\![\text{men}]\!]}\Big) = \{\emptyset\}$$

In the first step of computation of the meaning of “some men sneeze”, we obtain:

$$(\delta_N \otimes 1_N)\Big(\overline{[\![\text{men}]\!]} \otimes \overline{[\![\text{sneeze}]\!]}\Big) =$$
$$\{(\{\emptyset\}, \{\emptyset\}), (\{m_1\}, \{m_1\}), (\{m_2\}, \{m_2\}), (\{m_1, m_2\}, \{m_1, m_2\})\} \otimes \{\emptyset, \{m_1\}, \{c_1\}, \{m_1, c_1\}\}$$

In the second step, we obtain:

$$\Big(Some \otimes \mu\Big)\Big(\{(\{\emptyset\}, \{\emptyset\}), (\{m_1\}, \{m_1\}), (\{m_2\}, \{m_2\}), (\{m_1, m_2\}, \{m_1, m_2\})\} \otimes \{\emptyset, \{m_1\}, \{c_1\}, \{m_1, c_1\}\}\Big) =$$
$$Some\Big(\{\{m_1\}, \{m_2\}, \{m_1, m_2\}\}\Big) \otimes \mu\Big(\{\emptyset, \{m_1\}, \{m_2\}, \{m_1, m_2\}\} \otimes \{\emptyset, \{m_1\}, \{c_1\}, \{m_1, c_1\}\}\Big) =$$
$$\{\{m_1\}, \{m_2\}, \{m_1, m_2\}\} \otimes \{\emptyset, \{m_1\}\}$$

In the last step, we obtain the following via the relation $\epsilon \colon N \times N \to \{\star\}$ being $\{((\{m_1\}, \{m_1\}), \star)\}$:

$$\epsilon\Big(\{\{m_1\}, \{m_2\}, \{m_1, m_2\}\} \otimes \{\emptyset, \{m_1\}\}\Big) = \{\star\}$$

Hence, the meaning of the sentence is true. For the sentence “One man sneezes”, the second and third steps of computation are as follows:

$$\epsilon\Big(One\Big(\{\{m_1\},\{m_2\},\{m_1,m_2\}\}\Big) \otimes \mu\Big(\{\emptyset,\{m_1\},\{m_2\},\{m_1,m_2\}\} \otimes \{\emptyset,\{m_1\},\{c_1\},\{m_1,c_1\}\}\Big)\Big) =$$
$$\epsilon\Big(\{\{m_1\},\{m_2\}\} \otimes \{\emptyset,\{m_1\}\}\Big) = \{\star\}$$

So the meaning of this sentence is also true (it has the same ϵ relation as the previous case). Now consider the case of the sentence "no man sneezes" in which case $No([\![man]\!]) = \emptyset$. In this case we obtain the following at the final step of computation

$$\epsilon\Big(No\Big(\{\{m_1\},\{m_2\},\{m_1,m_2\}\}\Big) \otimes \mu\Big(\{\emptyset,\{m_1\},\{m_2\},\{m_1,m_2\}\} \otimes \{\emptyset,\{m_1\},\{c_1\},\{m_1,c_1\}\}\Big)\Big) =$$
$$\epsilon\Big(\{\emptyset\} \otimes \{\emptyset,\{m_1\}\}\Big) = \{\star\}$$

However, this is not a true sentence, since

$$\mu(\overline{[\![men]\!]} \otimes \overline{[\![sneeze]\!]}) = \mu(\{\emptyset,\{m_1\}\} \otimes \{\emptyset,\{m_1\},\{c_1\},\{m_1,c_1\}\}) = \{\emptyset,\{m_1\}\} \neq \{\emptyset\}$$

If we had $\overline{[\![dog]\!]} = \{\emptyset,\{d_1\}\}$, then the compact closed meaning of the sentence "No dogs sneeze" would become true, since we would have

$$\mu(\overline{[\![dogs]\!]} \otimes \overline{[\![sneeze]\!]}) = \mu(\{\emptyset,\{d_1\}\} \otimes \{\emptyset,\{m_1\},\{c_1\},\{m_1,c_1\}\}) = \{\emptyset\}$$

and also that

$$\epsilon(No \otimes \mu)(\sigma \otimes 1)(\overline{[\![dogs]\!]} \otimes \overline{[\![sneeze]\!]}) = \epsilon(\{\emptyset\} \otimes \{\emptyset\}) = \{\star\}$$

Hence "no dogs sneeze" is true.

7 Truth-Theory in Vector Spaces

One can do the same calculations as in Rel in $FVect$ and obtain the same truth theoretic meanings in vector spaces. In this case, all we have to do is to model a set $N = \{n_1, n_2, \cdots\}$, by a vector space V_N spanned by N, that is $V_N = \{\overrightarrow{n}_i\}_i$. In the case of our model, $N = \mathcal{P}(U)$, for U a set of individuals, hence V_N is spanned by subsets of U. Denoting these subsets by U_i, we have $V_{\mathcal{P}(U)} = \{\overrightarrow{U}_i\}_i$. The one element set $\{\star\}$ is then modelled by the one dimensional vector space $\{\overrightarrow{1}\}$, which models the sentence space, that is we have $S = \{\overrightarrow{1}\} = V_{\{\star\}}$. The zero element set, that is the empty set, is modelled by the zero vector $\overrightarrow{0}$.

We demonstrate the computation for the truth theoretic meaning of "Q Sbj Verb" in the above vector space below. For the meanings of the words therein, we have:

$$\overrightarrow{\text{Sbj}} = \sum_i \overrightarrow{U}_i \qquad \text{for} \quad U_i \in \mathcal{P}([\![\text{Sbj}]\!])$$
$$\overrightarrow{\text{Verb}} = \sum_j \overrightarrow{U}_j \otimes \{\star\} \cong \sum_i \overrightarrow{U}_j \qquad \text{for} \quad U_j \in \mathcal{P}([\![\text{Verb}]\!])$$
$$Det\Big(\overrightarrow{\text{Sbj}}\Big) = \sum_k \overrightarrow{U}_k \qquad \text{for} \quad U_k \in Det([\![\text{Sbj}]\!])$$

The first step of the computation is as follows:

$$(\delta_N \otimes 1_N)\Big(\overrightarrow{\mathrm{Sbj}} \otimes \overrightarrow{\mathrm{V}erb}\Big) = (\delta_N \otimes 1_N)\Big(\sum_i \overrightarrow{U}_i \otimes \sum_j \overrightarrow{U}_j\Big) = (\sum_i \overrightarrow{U}_i \otimes \overrightarrow{U}_i) \otimes (\sum_j \overrightarrow{U}_j)$$

In the second step, we obtain:

$$(Det \otimes \mu_N)\Big(\sum_i \overrightarrow{U}_i \otimes \overrightarrow{U}_i) \otimes (\sum_j \overrightarrow{U}_j\Big) = Det(\sum_i \overrightarrow{U}_i) \otimes (\sum_i \sigma_{ij}\overrightarrow{U}_i) = \sum_k \overrightarrow{U}_k \otimes \sum_i \sigma_{ij}\overrightarrow{U}_i$$

The final step provides us with the following:

$$(\epsilon_N)\Big(\sum_k \overrightarrow{U}_k \otimes \sum_i \sigma_{ij}\overrightarrow{U}_i\Big) = \sum_{ijk}\langle \overrightarrow{U}_k \mid \sigma_{ij}\overrightarrow{U}_i\rangle$$

Definition 2. *The vector space meaning of the sentence "Q Sbj Verb" in $FVect$ is true if and only if either $Det(\overrightarrow{Sbj}) \neq \overrightarrow{0}$ and we have:*

$$\epsilon \circ (Det \otimes \mu) \circ (\sigma \otimes 1_N)\Big(\overrightarrow{Sbj} \otimes \overrightarrow{Verb}\Big) \geqslant \overrightarrow{1}$$

or $Det(\overrightarrow{Sbj}) = \overrightarrow{0}$ and we have the above as well as the following:

$$\mu\Big(\overrightarrow{Sbj} \otimes \overrightarrow{Verb}\Big) = \overrightarrow{0}$$

Proposition 2. *The compact closed meaning of a quantified sentence computed in $FVect$ is true if and only if its generalised quantifier meaning is true.*

Example. As an example, consider the meaning of "some men sneeze", in the first step of the computation we have:

$$(\delta \otimes 1_N)\Big(\overrightarrow{\mathrm{men}} \otimes \overrightarrow{\mathrm{s}neeze}\Big) = (\delta \otimes 1)\Big((\emptyset + \{m_1\} + \{m_2\} + \{m_1, m_2\}) \otimes (\emptyset + \{m_1\} + \{c_1\} + \{m_1, c_1\})\Big)$$
$$= (\emptyset \otimes \emptyset + \{m_1\} \otimes \{m_1\} + \{m_2\} \otimes \{m_2\} + \{m_1, m_2\} \otimes \{m_1, m_2\}) \otimes (\emptyset + \{m_1\} + \{c_1\} + \{m_1, c_1\})$$

In the second step, we apply $(Some \otimes \mu)$ to the above and obtain:

$$Some\Big(\emptyset + \{m_1\} + \{m_2\} + \{m_1, m_2\}\Big) \otimes \mu\Big((\emptyset + \{m_1\} + \{m_2\} + \{m_1, m_2\}) \otimes (\emptyset + \{m_1\} + \{c_1\} + \{m_1, c_1\})\Big)$$
$$= (\{m_1\} + \{m_2\} + \{m_1, m_2\}) \otimes (\emptyset + \{m_1\})$$

In the final step, we apply ϵ to the above and obtain:

$$\langle \{m_1\} + \{m_2\} + \{m_1, m_2\} \mid \emptyset + \{m_1\}\rangle = 1$$

For "one man sneezes" the hole computation is as follows:

$$\epsilon(One \otimes \mu)(\delta \otimes 1_N)\Big(\overrightarrow{\mathrm{men}} \otimes \overrightarrow{\mathrm{s}neeze}\Big) =$$
$$\epsilon(One \otimes \mu)(\emptyset \otimes \emptyset + \{m_1\} \otimes \{m_1\} + \{m_2\} \otimes \{m_2\} + \{m_1, m_2\} \otimes \{m_1, m_2\}) \otimes (\emptyset + \{m_1\} + \{c_1\} + \{m_1, c_1\})$$
$$= \epsilon\Big(One(\emptyset + \{m_1\} + \{m_2\} + \{m_1, m_2\}\Big) \otimes \mu\Big((\emptyset + \{m_1\} + \{m_2\} + \{m_1, m_2\}) \otimes (\emptyset + \{m_1\} + \{c_1\} + \{m_1, c_1\}$$
$$= \epsilon\Big((\{m_1\} + \{m_2\}) \otimes (\emptyset + \{m_1\})\Big)$$
$$= \langle \{m_1\} + \{m_2\} \mid \emptyset + \{m_1\}\rangle = 1$$

For "no man sneezes", we have that $Det(\overrightarrow{\mathrm{S}bj}) = 0$, hence for this to be true its $FVect$ meaning should be above 1 and also $\mu\Big(\overrightarrow{\mathrm{Sbj}} \otimes \overrightarrow{\mathrm{Verb}}\Big)$ should be the $\emptyset$ vector.

The first step of the computation of the $FVect$ meaning of the sentence is as before, in the second and third steps we obtain:

$$\epsilon\Big(No(\emptyset + \{m_1\} + \{m_2\} + \{m_1, m_2\}) \otimes \mu(\emptyset + \{m_1\} + \{m_2\} + \{m_1, m_2\}) \otimes (\emptyset + \{m_1\} + \{c_1\} + \{m_1, c_1\})\Big)$$
$$= \epsilon\Big(\{\emptyset\} \otimes (\emptyset + \{m_1\})\Big) = 1$$

But this sentence is not true, since we have

$$(1_N \otimes \mu) \circ (\sigma \otimes 1_N)(\overrightarrow{\mathrm{me}n} \otimes \overrightarrow{sneeze}) = \emptyset + \{m_1\} \neq \emptyset$$

Whereas, for "no dogs sneeze", where $\overrightarrow{\mathrm{d}ogs} = \emptyset + \{d_1\}$, we will have:

$$\mu\big(\overrightarrow{\mathrm{d}ogs} \otimes \overrightarrow{sneeze}\big) = \mu\big((\emptyset + \{d_1\}) \otimes (\emptyset + \{m_1\} + \{c_1\} + \{m_1, c_1\})\big) = \emptyset$$

And also that

$$\epsilon(No \otimes \mu)(\delta \otimes 1)(\overrightarrow{\mathrm{d}ogs} \otimes \overrightarrow{sneeze}) = \epsilon(\emptyset \otimes \emptyset) = 1$$

Hence "no dogs sneeze" is true.