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 [7]. 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 [6]. Compositional distributional models of meaning extend the vector space models from words to sentences. The categorical such models [4, 1] 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 [8, 9]. 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 [2].

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 [4]. 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 [10, 11]. 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:

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

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 $\mathcal{C}$:

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

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

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

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 naturak 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:

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

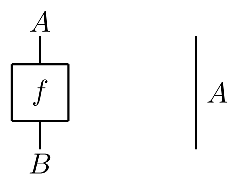
Every object in Rel has a Frobenius algebra over it given by the diagonal and codiagonal relations, as described below:

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

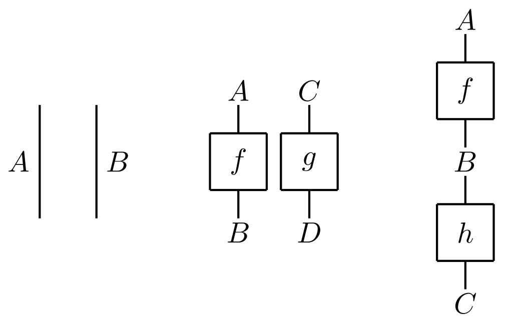
For the details of verifying that for each of the two examples above, the corresponding conditions hold see [3].

3 String Diagrams

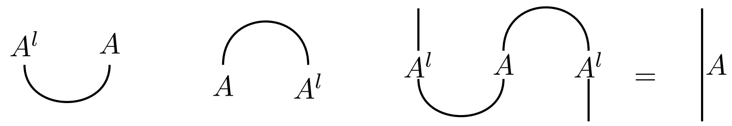
The framework of compact closed categories and Frobenius algebras comes with a complete diagrammatic calculus that visualises derivations, and which also simplifies the categorical and vector space computations. Morphisms are depicted by boxes and objects by lines, representing their identity morphisms. 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:

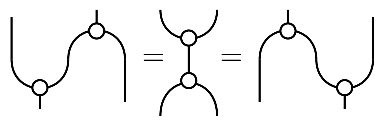


The ϵ maps are depicted by cups, η maps by caps, and yanking by their composition and straightening of the strings. For instance, the diagrams for $\epsilon^l \colon A^l \otimes A \to I$, $\eta \colon I \to A \otimes A^l$ and $(\epsilon^l \otimes 1_A) \circ (1_A \otimes \eta^l) = 1_A$ are as follows:

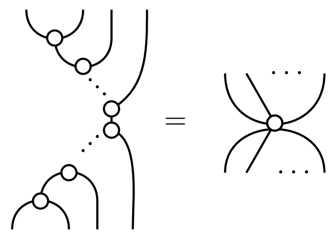


As for Frobenius algebras, the diagrams for the monoid and comonoid morphisms are as follows:

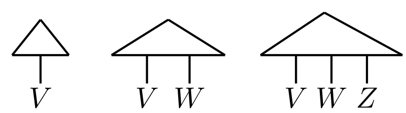
with the Frobenius condition being depicted as:



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’:

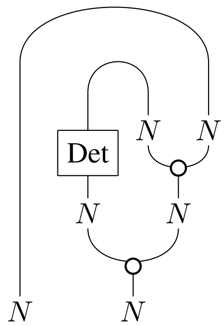


In the category FdVect, apart from spaces V, W, which are objects of the category, we also have vectors $\overrightarrow{v}, \overrightarrow{w}$. These are depicted by their representing morphisms and as triangles with a number of strings emanating from them. The number of strings of a triangle denote 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

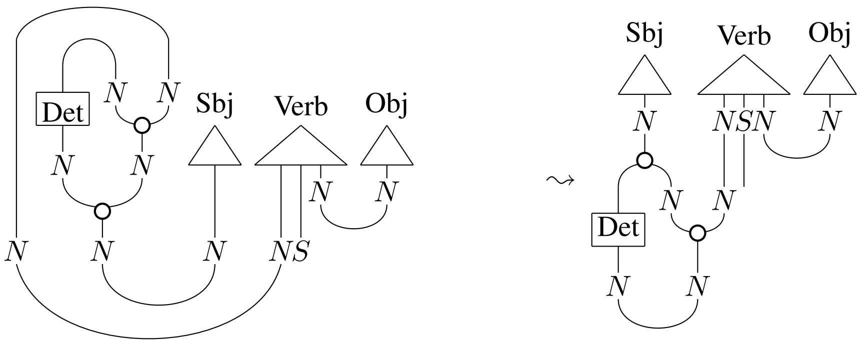
Following the terminology and notation of [2], given a phrase containing a quantifier followed by a noun, that is 'Q noun', we call 'Q' a determiner and the phrase 'Q noun' a quantified phrase.Hence, a quantified phrase is a noun phrase which is created by the application of a determiner to a noun phrase. 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 above is as follows:

$$(\epsilon_N \otimes 1_S) \circ (Det \otimes \mu_N \otimes 1_S) \circ (\delta_N \otimes 1_{N \otimes S})$$

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:

$$Det(\overrightarrow{n}_i) = \Phi\{\overrightarrow{w} \in N \mid \cos(\overrightarrow{n}_i, \overrightarrow{w}) = \alpha\}$$

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 the 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:

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

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 $\{s_1, s_2\}$. Suppose the linear expansion of $\overrightarrow{\text{Sbj}}$ in this space be $C_1\overrightarrow{n}_1 + C_2\overrightarrow{n}_2$ and the linear expansion of $\overrightarrow{\text{Verb}}$ be $C_{11}(\overrightarrow{n}_1 \otimes \overrightarrow{s}_1) + C_{12}(\overrightarrow{n}_1 \otimes \overrightarrow{s}_2) + C_{21}(\overrightarrow{n}_2 \otimes \overrightarrow{s}_1) + C_{22}(\overrightarrow{n}_2 \otimes \overrightarrow{s}_2)$. Suppose $\overrightarrow{n}_1$ is the word 'prison' and $\overrightarrow{n}_2$ is the word 'owner', then one will have the following vectors for the words 'cats' and 'murderer':

$$\overrightarrow{\text{cat}} = 0.8\ \overrightarrow{\text{owner}} + 0.2\ \overrightarrow{\text{prison}} \qquad \overrightarrow{\text{murderer}} = 0.1\ \overrightarrow{\text{owner}} + 0.9\ \overrightarrow{\text{prison}}$$

Hence, meaning of the phrase 'most cats' will be a word whose vectors is close to the word 'cat' in this space, for example 'kitten' or 'dog', whereas the meaning of 'few cats' will be a word whose vectors is far from the word 'cat', for example, 'murderer'. Meaning of the sentence 'most cats sneeze' will be close to the meaning of the sentence 'kittens sneeze', and meaning of the sentence 'few cats sneeze' will be close to the meaning of the sentence 'murderes sneeze'. In the first case, 'most cats' is represented by 'kittens' which shares most of the properties of 'cats', whereas in the second case, 'few cats' is represented by 'murderers' which shares very few properties with 'cats'.

6 Truth Theoretic Interpretation

For this part, we work in the category Rel of sets and relations. We take N to be the power set of a set of individuals. A common noun is then the powerset of the set of its individuals. For example, the set $\{\{m_1\}, \{m_2\}, \{m_1, m_2\}, \emptyset\}$ denotes the common noun 'men', for m_1, m_2 two male individuals and the set $\{\{c_1\}, \emptyset\}$ denotes the common noun 'cat', with one individual c_1. We take S to be the one dimensional space free over the singleton $\overrightarrow{1}$. The origin represents false and number one represents truth. A verb is the powerset of a relation (corresponding to its predicate). For an intransitive verb, this relation is on the set $N \times S$, where each relation corresponds to a subset of N, since we have $N \times S \cong N$. For a transitive verb, it is a relation on the set $N \times S \times N \cong N \times N$. Hence, an intransitive verb is the powerset of all individuals to which the verb applies. As an example, suppose the verb 'sneeze' applies to individuals m_1 and c_1, hence it is represented by $\{\{m_1\}, \{c_1\}, \{m_1, c_1\}, \emptyset\}$.

The map Det sends a subset of individuals to a set of its subsets exactly in the same way as defined by [2]. 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 non-empty subsets of individuals and so on. In our example, two(men) $= \{\{m_1, m_2\}\}$, whereas some(men) $= \{\{m_1\}, \{m_2\}, \{m_1, m_2\}\}$. For the case of 'cats', we have that two(cats) $= \{\emptyset\}$.

Following previous work [4], we represent the set of individuals as vector spaces free over (or spanned by) its elements, which in this setting are represented by basis vectors. Common nouns become sums of their elements in this space. The truth-theoretic meaning of the sentence "Det Sbj Verb" can then be obtained by computing the morphism developed in section 4. In the first step of the computation, we obtain the following:

$$(\delta_N \otimes 1_{N\otimes S})\Big(\overrightarrow{Sbj} \otimes \overrightarrow{Verb}\Big) = (\sum_i \overrightarrow{n}_i \otimes \overrightarrow{n}_i) \otimes (\sum_j \overrightarrow{n}_j \otimes \overrightarrow{1}) = (\sum_i \overrightarrow{n}_i \otimes \overrightarrow{n}_i) \otimes (\sum_j \overrightarrow{n}_j)$$

In the second step, we obtain the following (note that we are dropping the S morphisms because the S is now the unit of tensor):

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

The final step provides us with the following:

$$(\epsilon_N)\Big(Det(\sum_i \overrightarrow{n}_i) \otimes (\sum_i \sigma_{ij}\overrightarrow{n}_i)\Big) = \langle Det(\sum_i \overrightarrow{n}_i) \mid \sum_{ij} \sigma_{ij}\overrightarrow{n}_i\rangle$$

If the result is non-zero, the meaning of the sentence is true, else it is false. As an example, consider the meaning of 'some men sneeze', which becomes as follows:

$$\langle \{\{m_1\}, \{m_2\}, \{m_1, m_2\}\} \mid \{\{m_1\}, \{c_1\}, \{m_1, c_1\}, \emptyset\}\rangle = 1$$

For 'one cat sneezes', we have

$$\langle \{\{c_1\}\} \mid \{\{m_1\}, \{c_1\}, \{m_1, c_1\}, \emptyset\}\rangle = 1$$

Whereas for 'two cats sneeze' we have

$$\langle \{\{c_1, c_1\}\} \mid \{\{m_1\}, \{c_1\}, \{m_1, c_1\}, \emptyset\}\rangle = 0$$

We have not done the necessary expositions here, but as shown in [5, 3], the Frobenius μ map is the analog of set-theoretic intersection and the compact closed epsilon map is the analog of set-theoretic application. Using these it is not hard to show that this truth-theoretic interpretation of the compact closed semantics of quantified sentences provides us with the same meaning as their generalised quantifier semantics.

References

1. Baroni, M., Zamparelli, R.: Nouns are vectors, adjectives are matrices: Representing adjective-noun constructions in semantic space. In: Conference on Empirical Methods in Natural Language Processing (EMNLP-10). Cambridge, MA (2010)
2. Barwise, J., Cooper, R.: Generalized quantifiers and natural language. Linguistics and Philosophy 4(2), 159–219 (1981), `http://dx.doi.org/10.1007/BF00350139`
3. Coecke, B., Paquette, E.: Introducing categories to the practicing physicist. In: Coecke, B. (ed.) New Structures for Physics, Lecture Notes in Physics, vol. 813, pp. 167–271. Springer (2008)
4. Coecke, B., Sadrzadeh, M., Clark, S.: Mathematical Foundations for Distributed Compositional Model of Meaning. Lambek Festschrift. Linguistic Analysis 36, 345–384 (2010)
5. Coecke, B., Paquette, É.O., Pavlovic, D.: Classical and quantum structuralism. In: Gay, S., Mackie, I. (eds.) Semantical Techniques in Quantum Computation. pp. 29–69. Cambridge University Press (2009)
6. Curran, J.: From Distributional to Semantic Similarity. Ph.D. thesis, School of Informatics, University of Edinburgh (2004)
7. Firth, J.: A synopsis of linguistic theory 1930–1955. In: Studies in Linguistic Analysis (1957)
8. Grefenstette, E., Sadrzadeh, M.: Experimental support for a categorical compositional distributional model of meaning. In: Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP). pp. 1394–1404 (2011)
9. Kartsaklis, D., Sadrzadeh, M., Pulman, S.: A unified sentence space for categorical distributional-compositional semantics: Theory and experiments. In: Proceedings of 24th International Conference on Computational Linguistics (COLING 2012): Posters. pp. 549–558. The COLING 2012 Organizing Committee, Mumbai, India (December 2012)
10. Kelly, G., Laplaza, M.: Coherence for compact closed categories. Journal of Pure and Applied Algebra 19(0), 193 – 213 (1980), `http://www.sciencedirect.com/science/article/pii/0022404980901012`
11. Kock, A.: Strong functors and monoidal monads. Archiv der Mathematik 23, 113–120 (1972)