Ambiguity and Incomplete Information in Categorical Models of Language

Dan Marsden
University of Oxford
daniel.marsden@cs.ox.ac.uk

We investigate notions of ambiguity and partial information in categorical distributional models of natural language. Probabilistic ambiguity has previously been studied in [27, 26, 16] using Selinger's CPM construction. This construction works well for models built upon vector spaces, as has been shown in quantum computational applications. Unfortunately, it doesn't seem to provide a satisfactory method for introducing mixing in other compact closed categories such as the category of sets and binary relations. We therefore lack a uniform strategy for extending a category to model imprecise linguistic information.

In this work we adopt a different approach. We analyze different forms of ambiguous and incomplete information, both with and without quantitative probabilistic data. Each scheme then corresponds to a suitable enrichment of the category in which we model language. We view different monads as encapsulating the informational behaviour of interest, by analogy with their use in modelling side effects in computation. Previous results of Jacobs then allow us to systematically construct suitable bases for enrichment.

We show that we can freely enrich arbitrary dagger compact closed categories in order to capture all the phenomena of interest, whilst retaining the important dagger compact closed structure. This allows us to construct a model with real convex combination of binary relations that makes non-trivial use of the scalars. Finally we relate our various different enrichments, showing that finite subconvex algebra enrichment covers all the effects under consideration.

1 Introduction

The categorical distributional approach to natural language processing [6] aims to construct the meaning of a sentence from its grammatical structure and the meanings of its parts. The grammatical structure of language can be described using pregroup grammars [20]. The distributional approach to natural language models the meanings of word as vectors of statistics in finite dimensional real vector spaces. Both the category of real vector spaces and linear maps, and pregroups are examples of monoidal categories in which objects have duals. This common structure is the key observation that allows us to functorially transfer grammatical structure to linear maps modelling meaning in a compositional manner.

Recent work in categorical models of language has investigated ambiguity, for example words with multiple meanings, [27, 26, 16]. These papers exploit analogies with quantum mechanics, using density matrices to model partial and ambiguous information. In order to do this, the category in which meanings are interpreted must change. Selinger's CPM construction [29] is exploited to construct compact closed categories in which ambiguity can be described. This construction takes a compact closed category $\mathscr{C}$ and produces a new compact closed category $\mathbf{CPM}(\mathscr{C})$. When applied to the category of finite dimensional Hilbert spaces and linear maps, **FdHilb**, the resulting category $\mathbf{CPM}(\mathbf{FdHilb})$ is equivalent to the category of finite dimensional Hilbert spaces and completely positive maps. In this way the construction takes the setting for pure state quantum mechanics and produces exactly the right setting for mixed state

Submitted to:
QPL 2016

quantum mechanics. It is therefore tempting to consider the CPM construction as a mixing device for compact closed categories. This perspective was adopted in [27] by starting with the category **Rel** of sets and binary relations and interpreting **CPM**(**Rel**) as a toy model of ambiguity. As argued in [24, 9], there are aspects of **CPM**(**Rel**) that conflict with its interpretation as a setting for mixing. Specifically:

- There are pure states that can be formed as a convex mixture of two mixed states
- There are convex combinations of distinct pure states that give pure states

It could be argued that this anomalous behaviour is due to the restricted nature of the scalars in **Rel**. This then points to another weakness of **CPM**(**Rel**) as a setting for ambiguity and mixing. We may wish to say that the word bank is 90% likely to mean a financial institution and 10% likely to refer to the boundary of a river. We can express this in **CPM**(**FdHilb**) as the scalars are sufficiently rich. In **CPM**(**Rel**) the best we can do is say that it's either one or the other, with all the quantitative information being lost.

In this paper we investigate some alternative models of ambiguity in the compact closed setting. We consider a variety of different interpretations of what it might mean to have ambiguous or limited information, and how these can be described mathematically. Specifically:

- In section 4 we describe constructions that model ambiguous and incomplete information in a non-quantitative manner. We show that compact closed categories can be freely extended so as to allow the modelling of incomplete information, ambiguity and a mixture of both phenomena.
- In section 5 we extend our constructions to describe quantitative ambiguous and incomplete information. Again we show that compact closed categories can be freely extended in order to model such features.
- Theorem 5.10 shows that these various informational notions fit into a hierarchy with the subdistribution monad at the top, capturing all the aspects we consider.

Our general perspective is to enrich the homsets of our categories in order to model the language features we are interested in. In order to do this systematically, we exploit some basic monad theory. Monads are commonly used to describe computational effects such as non-determinism, exceptions and continuations [25, 31]. In our case we instead view them as models of informational effects in natural language applications. Monads have previously been used in models of natural language, see for example [30]. Work of Jacobs [11] provides the connection between a certain class of monads and categories that provide good bases for enriched category theory. We aim to give explicit constructions of the categories of concrete interest throughout, rather than pursuing a policy of maximum abstraction. Our systematic approach means we should be able to incorporate additional informational features in a similar manner.

We assume familiarity with elementary category theory, and the notions of compact closed and dagger compact closed categories [1].

2 Monads

We will now outline the necessary background on monads required in later sections, and introduce the monads that will be of particular interest. The material in this section is standard, good sources for further background are [22, 2, 4, 2, 14]. In this paper we will only be interested in monads on the category **Set** of sets and total functions, although we will state some definitions more generally where it is cleaner to do so.

Definition 2.1 (Monad)**.** A **monad** on a category $\mathscr{C}$ is triple consisting of:

- An endofunctor $T : \mathscr{C} \to \mathscr{C}$

- A **unit** natural transformation $\eta : 1 \Rightarrow T$
- A **multiplication** natural transformation $\mu : TT \Rightarrow T$

such that the following three axioms hold:

$$\mu \circ (\eta * T) = T \qquad \mu \circ (T * \eta) = T \qquad \mu \circ (\mu * T) = \mu \circ (T * \mu)$$

We now introduce the monads of interest in this paper, and relate them to computational behaviour. Similarly to [28], we emphasize that monads are induced by algebraic operations modelling computational, or in our case informational, behaviour.

Definition 2.2 (Lift Monad)**.** The lift monad $((-)_\perp, \eta, \mu)$ is defined as follows:

- The functor component is given by the coproduct of functors $1 + \{\perp\} : \mathbf{Set} \to \mathbf{Set}$.
- The unit and multiplication are given componentwise by:

$$\eta_X(x) = x \qquad \mu_X(x) = \begin{cases} x \text{ if } x \in X \\ \perp \text{ otherwise} \end{cases}$$

The lift monad is commonly used to describe computations that can diverge.

Definition 2.3 (Powerset Monads)**.** The **finite powerset monad** P_ω has functor component the covariant finite powerset functor. The unit sends an element to the corresponding singleton, and the multiplication is given by taking unions. The **non-empty finite powerset monad** P_ω^+ arises in an analogous way by restricting the sets under consideration. The finite powerset monad is used to model finitely bounded non-determinism, and the non-empty finite powerset monad eliminates the possibility of divergence.

Definition 2.4 (Finite Distribution Monads)**.** The **finite distribution monad** has functor component:

$$\begin{aligned} D : \mathbf{Set} &\to \mathbf{Set} \\ X &\mapsto \{d : X \to [0,1] \mid d \text{ has finite support and } \sum_x d(x) = 1\} \\ f : X \to Y &\mapsto \lambda d\, y. \sum_{x \in f^{-1}\{y\}} d(x) \end{aligned} \tag{1}$$

The unit and multiplication are given componentwise by:

$$\eta_X(x) = \delta_x \qquad \mu_X(d)(x) = \sum_{e \in \mathrm{supp}(e)} d(e)e(x)$$

where δ_x is the Dirac delta function and $\mathrm{supp}(e)$ is the support of e.

The **finite subdistribution monad** (S, η, μ) has identical structure, except that we weaken the condition in equation (1) to:

$$\sum_x d(x) \leq 1$$

So our finite distributions are now sub-normalized rather than normalized to 1. Both the finite distribution and subdistribution monads are used to model probabilistic computations. Intuitively the subdistribution monad provides scope for diverging behaviour in the "missing" probability mass.

Remark 2.5. We adopt a convenient notational convention from [13] and write finite distributions as formal sums $\sum_i p_i |x_i\rangle$, where we abuse the physicists ket notation to indicate the sum is a formal construction. Using this notation, the unit of the (sub)distribution monad is the map $x \mapsto |x\rangle$ and multiplication is given by expanding out sums of sums in the usual manner.

Each monad can be canonically related to a certain category of algebras.

Definition 2.6 (Eilenberg-Moore Algebras)**.** Let (T,η,μ) be a monad on $\mathscr{C}$. An **Eilenberg-Moore algebra** [7] for T consists of an object A and a morphism $a : TA \to A$ satisfying the following axioms:

$$a \circ \eta_A = 1 \qquad\qquad a \circ \mu_A = a \circ Ta$$

A morphism of Eilenberg-Moore algebras of type $(A,a) \to (B,b)$ is a morphism in $\mathscr{C}$ such that:

$$h \circ a = b \circ Th$$

The category of Eilenberg-Moore algebras and their homomorphisms will be denoted $\mathrm{EM}(T)$.

Example 2.7. For the monads under consideration we note that:

- The Eilenberg-Moore category of the lift monad is equivalent to the category of pointed sets and functions that preserve the distinguished element, denoted $\mathbf{Set}_\bullet$.
- The Eilenberg-Moore category of the finite powerset monad is equivalent to the category of join semilattices and homomorphisms, denoted **JSLat**.
- The Eilenberg-Moore category of the non-empty finite powerset monad is equivalent to the category of affine join semilattices and homomorphisms, denoted **AJSLat**.
- The Eilenberg-Moore category of the finite distribution monad is the category of convex algebras and functions commuting with forming convex combinations, denoted **Convex**. This category has received a great deal of attention in for example [12, 15, 8].
- The Eilenberg-Moore category of the finite subdistribution monad is the category of subconvex algebras, that is algebras that can form subconvex combinations of elements in a coherent manner. The morphisms are functions that commute with forming subconvex combinations. We denote this category **Subconvex**.

We consider commutative monads on the category **Set**, and specialize their definition appropriately.

Definition 2.8 (Commutative Monad)**.** Let (T,η,μ) be a **Set** monad. There are a canonical **strength** and **costrength** natural transformations:

$$\begin{aligned} \mathrm{st}_{X,Y} &: X \times TY \to T(X \times Y) \\ (x,t) &\mapsto T(\lambda y.(x,y))(t) \end{aligned} \qquad\qquad \begin{aligned} \mathrm{st'}_{X,Y} &: TX \times Y \to T(X \times Y) \\ (t,y) &\mapsto T(\lambda x.(x,y))(t) \end{aligned}$$

The monad is said to be a **commutative monad** [19] if the following equation holds for all X,Y:

$$\mu_{X\times Y} \circ T(\mathrm{st'}_{X,Y}) \circ \mathrm{st}_{TX,Y} = \mu_{X\times Y} \circ T(\mathrm{st}_{X,Y}) \circ \mathrm{st'}_{X,TY}$$

This composite is then called the **double strength**, denoted dst.

Remark 2.9. Monads are intimately related to the topic of universal algebra. The Eilenberg-Moore algebras for a **Set** monad can be presented by operations and equations, if we permit infinitary operations. All the monads in this paper are in fact **finitary monads**, meaning they can presented by operations of finite arity. Let ϕ and ψ be operations of arities m and n respectively. These operations are said to commute with each other if the following equation holds:

$$\psi(\phi(x_{1,1},...,x_{1,m}),...,\phi(x_{n,1},...,x_{n,m})) = \phi(\psi(x_{1,1},...,x_{n,1}),...,\psi(x_{1,m},...,x_{n,m}))$$

If we unravel the definition of commutative monad, it says that all the operations in a presentation commute with each other. We can also phrase this as every operation being a homomorphism. More detailed discussion of connections to universal algebra and presentations can be found in [23].

Lemma 2.10. *Each of the lift, powerset, finite powerset, finite non-empty power, distribution and sub-distribution monads are commutative.*

Remark 2.11. It is interesting that all the notions of partial information and ambiguity considered in this paper give rise to commutative monads. Possibly we could regard this as showing these informational effects are independent of the order in which they are built up?

Clearly many monads are not commutative:

Example 2.12. The **list monad** has a functor component that sends a set to finite lists of its elements. The unit maps an element to the corresponding single element list and the multiplication concatenates lists of lists. The Eilenberg-Moore algebras of this monad are arbitrary, not necessarily commutative, monoids. Unsurprisingly, this monad is not commutative.

The following proposition captures the essential properties of Eilenberg-Moore categories of commutative monads that we will need, all in one place. The key symmetric monoidal closed structure is due to work of Jacobs [11], the other properties are well known.

Proposition 2.13. *Let* (T, η, μ) *be a commutative monad on* **Set**. *The category* $\mathrm{EM}(T)$*:*

- *Is a symmetric monoidal closed category.*
- *Has universal bimorphisms for the monoidal tensor.*
- *Has monoidal unit given by the free algebra* $(\{*\}, !)$.
- *The tensor product* $\mu_X \otimes \mu_Y$ *is isomorphic to* $\mu_{X \times Y}$.
- *Is complete.*
- *Is cocomplete.*

Proof. Completeness and cocompleteness of categories that are monadic over **Set** is standard. The category **Set** is a SMCC via its products and exponentials. **Set** is complete so we can use [11, lemma 5.3]. The additional properties of the monoidal structure come from [11, lemmas 5.1,5.2]. □

Remark 2.14. We avoid technical discussion of universal bimorphisms, details can be found in [18, 11]. The essential idea is to generalize the universal property of the tensor product of vector spaces, and their relationship to bilinear functions. So in the set theoretic case, homomorphisms out of our tensors will bijectively correspond to functions out of the cartesian product that are homomorphisms in each component separately.

3 Enriched Categories

An enriched category is a category in which the homsets have additional structure that interacts well with composition.

Example 3.1. The following are natural examples of enriched categories:

- As a trivial example, ordinary locally small categories are **Set**-enriched.
- A category is poset enriched if it's homsets have a poset structure and composition in monotone with respect to that structure. For example the category **Rel** is poset enriched.
- In the categorical quantum mechanics community, a category is said to have a **superposition rule** [10], or be a **process theory with sums** [5], if its homsets carry commutative monoid structure that is suitably compatible with composition (and possibly additional structure).

We have insufficient space for a detailed outline of the parts of enriched category theory we require, we refer the reader to [17, 4] for background. The informal discussion above should hopefully be sufficient to understand the discussions in later sections.

The idea of this paper is that complete and cocomplete categories with symmetric monoidal closed structure provide a very good base of enrichment for enriched category theory. If we select a commutative monad that captures the linguistic feature we are interested in, we can then consider categories enriched in the corresponding algebraic structure.

The universal bimorphism property of the monoidal structure of Eilenberg-Moore categories of commutative monads allows us provide concrete conditions for enrichment in our categories of interest:

Proposition 3.2. *A category $\mathscr{C}$:*

- *Is* **Set$_\bullet$**-*enriched if its homsets have pointed set structures such that:*

$$\bot \circ f = \bot \quad \textit{and} \quad f \circ \bot = \bot \tag{2}$$

- *Is* **AJSLat**-*enriched if its homsets have affine join semilattice structures such that:*

$$(f \vee g) \circ h = (f \circ h) \vee (g \circ h) \quad \textit{and} \quad f \circ (g \vee h) = (f \circ g) \vee (f \circ h) \tag{3}$$

- *Is* **JSLat**-*enriched if its homsets have join semilattice lattice structures such that both the equations of* (2) *and* (3) *hold.*
- *Is* **Convex**-*enriched if its homsets have convex algebra structures such that:*

$$(\sum_i p_i f_i) \circ g = \sum_i p_i (f_i \circ g) \quad \textit{and} \quad f \circ (\sum_i p_i g_i) = \sum_i p_i (f \circ g_i) \tag{4}$$

- *Is* **Subconvex**-*enriched if its homsets have subconvex algebra structures such that the equations* (4) *hold for all subconvex combinations.*

For a given cocomplete symmetric monoidal closed category $\mathscr{V}$, we can form the free $\mathscr{V}$-enriched category over an ordinary category. This construction will be exploited in the later sections, details can be found in [17, 4].

4 Unquantified Mixing

We begin by considering probably the simplest case, in which we have incomplete information. For example, I simply don't know the meaning of the word "logolepsy". In order to model this, we enrich our category in pointed sets, with the distinguished element denoting missing information.

Definition 4.1. For category $\mathscr{C}$ we define the category $\mathscr{C}_\bot$ as having:

- **Objects**: The same objects as $\mathscr{C}$
- **Morphisms**: We define $\mathscr{C}_\bot(A,B) = \mathscr{C}(A,B)_\bot$

Identities are as in $\mathscr{C}$. Composition is given as in $\mathscr{C}$, extended with the rules:

$$\bot \circ f = \bot \quad \text{and} \quad g \circ \bot = \bot$$

Proposition 4.2. *For a category $\mathscr{C}$, $\mathscr{C}_\bot$ is the free pointed set enriched category over $\mathscr{C}$.*

Theorem 4.3. *If $\mathscr{C}$ is a compact closed category then $\mathscr{C}_\perp$ is a compact closed category. The monoidal structure on morphisms extends that in $\mathscr{C}$ as in* (5).

$$f \otimes_\perp f' = \begin{cases} \perp \text{ if } f = \perp \text{ or } f' = \perp \\ f \otimes f' \text{ otherwise} \end{cases} \quad (5) \qquad f^{\dagger_\perp} = \begin{cases} \perp \text{ if } f = \perp \\ f^\dagger \text{ otherwise} \end{cases} \quad (6)$$

There is an identity and surjective on objects strict monoidal embedding $\mathscr{C} \to \mathscr{C}_\perp$.

If $\mathscr{C}$ is a dagger compact closed category then $\mathscr{C}_\perp$ is a dagger compact closed category with the dagger extending that of $\mathscr{C}$ as in (6).

Proof. We sketch the basic ideas. We must check that the extended definitions of the monoidal product and if necessary the dagger are functorial. We then wish to inherit the associator, left and right unitor and symmetry from the base. In order to do so we must check they remain natural with respect to the extended functor actions on morphisms. Cups and caps and other structure and axioms are then broadly speaking inherited from $\mathscr{C}$. □

Although $\mathscr{C}_\perp$ gives a new compact closed category, it is not particularly exciting. As soon as we compose or tensor with a $\perp$ element, the whole term becomes $\perp$. This is consistent with the intuition that if we have no idea about part of the information we require, we cannot know the whole either.

We took the opportunity to sketch the proof that $\mathscr{C}_\perp$ is compact closed as it is easiest to follow in this simple case. Later proofs of similar claims are analogous.

Remark 4.4. Although we have inherited some good properties from the base category in the construction 4.1, clearly not everything can be preserved. We are expanding the morphisms between each pair of objects, for example in $\mathbf{Rel}_\perp$ we now have three scalars, which we can interpret as true, false and unknown. This expansion will interfere with (co)limits from the base category. For example if we have a zero object in $\mathscr{C}$, that object will no longer be a zero object in $\mathscr{C}_\perp$ as there will be 2 morphisms of type $0 \to 0$.

Noting the similarity with the behaviour of the $\perp$ elements with that of zero morphisms in categories with zero objects, we note that:

Lemma 4.5. *Every category with a zero object is* $\mathbf{Set}_\bullet$*-enriched.*

Although the lift monad and $\mathbf{Set}_\bullet$-enrichment are extremely straightforward, we shall return to them later, in interaction with different types of ambiguity.

If the previous model captured incomplete information, we now move to consider ambiguity. In this case we intend situations where several things are possible, for example a bat is either a winged mammal or sporting equipment. In particular, unlike the previous model, we have complete information about the available possibilities, we are simply unaware of which one applies. If we don't have any sense of the relative likelihoods of the possibilities, we are just left with a non empty finite set of alternatives, and this points us in the direction of the monad P^+_ω.

Definition 4.6. For category $\mathscr{C}$ we define the category $\mathscr{C}_{P^+_\omega}$ as having:

- **Objects**: The same objects as $\mathscr{C}$
- **Morphisms**: We define $\mathscr{C}_{P^+_\omega}(A,B) = P^+_\omega(\mathscr{C}(A,B))$

We define composition as follows:

$$V \circ U = \{ v \circ u \mid v \in V, u \in U \}$$

Identities are then given by the singletons containing the identities from $\mathscr{C}$.

Proposition 4.7. *For a category $\mathscr{C}$, $\mathscr{C}_{P_\omega^+}$ is the free affine join semilattice enriched category over $\mathscr{C}$.*

Theorem 4.8. *If $\mathscr{C}$ is a compact closed category then $\mathscr{C}_{P_\omega^+}$ is a compact closed category. The action of the tensor on morphisms extends that of $\mathscr{C}$ as in* (7).

$$U \otimes_{P_\omega^+} U' = \{u \otimes u' \mid u \in U, u' \in U'\} \quad (7) \qquad\qquad U^{\dagger_{P_\omega^+}} = \{u^\dagger \mid u \in U\} \quad (8)$$

There is an identity and surjective on objects strict monoidal embedding $\mathscr{C} \to \mathscr{C}_{P_\omega^+}$.

If $\mathscr{C}$ is a dagger compact closed category then $\mathscr{C}_{P_\omega^+}$ is a dagger compact closed category with the dagger extending that of $\mathscr{C}$ as in (8).

As a final possibility, what if we wish to consider both ambiguous and incomplete information? It is then natural to consider the finite powerset monad as the source of our enrichment. The definition is almost identical to that of $\mathscr{C}_{P_\omega^+}$:

Definition 4.9. For category $\mathscr{C}$ we define the category $\mathscr{C}_{P_\omega}$ as having:

- **Objects**: The same objects as $\mathscr{C}$
- **Morphisms**: We define $\mathscr{C}_{P_\omega}(A,B) = P_\omega(\mathscr{C}(A,B))$

We define composition as follows:

$$V \circ U = \{v \circ u \mid v \in V, u \in U\}$$

Identities are then given by the singletons containing the identities from $\mathscr{C}$.

Proposition 4.10. *For a category $\mathscr{C}$, $\mathscr{C}_{P_\omega}$ is the free join semilattice enriched category over $\mathscr{C}$.*

Theorem 4.11. *If $\mathscr{C}$ is a compact closed category then $\mathscr{C}_{P_\omega}$ is a compact closed category. The action of the tensor on morphisms extends that of $\mathscr{C}$ as in* (9).

$$U \otimes_{P_\omega} U' = \{u \otimes u' \mid u \in U, u' \in U'\} \quad (9) \qquad\qquad U^{\dagger_{P_\omega}} = \{u^\dagger \mid u \in U\} \quad (10)$$

There is an identity and surjective on objects strict monoidal embedding $\mathscr{C} \to \mathscr{C}_{P_\omega}$.

If $\mathscr{C}$ is a dagger compact closed category then $\mathscr{C}_{P_\omega}$ is a dagger compact closed category with the dagger extending that of $\mathscr{C}$ as in (10).

Definition 4.12. We consider two sub-classes of monad:

- A monad is said to be **affine** [21, 11] if the component of its unit at the terminal object is an isomorphism.
- A monad is said to be **relevant** [11] if $\mathrm{dst} \circ \delta = T\delta$

Techniques for extracting the affine and relevant parts of a commutative monad can be found in [11].

Remark 4.13. Again we can think about the notion of affine monad in terms of presentations, as we did with commutativity. An operation is said to be idempotent if:

$$\psi(x, x, ..., x) = x$$

An algebraic theory is affine if all its operations are idempotent. In particular this means the theory can have no constants or non-trivial unary operations. It makes intuitive sense that descriptions of ambiguity should lead to affine algebraic theories. We cannot just conjure up elements out of thin air, and being confused between x and x should provide the same information as knowing x directly.

We now note a fundamental relationship between the three monads considered in this section.

Remark 4.14. As observed in [11], P_ω^+ and the lift monad are the affine and relevant parts of the finite powerset monad. In fact the finite powerset monad can be constructed from the non-empty finite powerset monad and the lift monad using a distributive law [3], and so in a mathematical sense it is precisely the description of incomplete information combined with non-quantitative ambiguity. A similar pattern will be repeated in the next section.

5 Quantified Mixing

We now move to the setting that has typically been considered in categorical models of mixing and ambiguity until now, probabilistic mixtures. Here we return to the situation where our state of knowledge is for example that the word "bank" suggests with 90% confidence a financial bank and 10% confidence a river bank. We now have quantitative information, and it should be possible to encode this information in our homsets.

Definition 5.1. For category $\mathscr{C}$ we define the category $\mathscr{C}_D$ as having:

- **Objects**: The same objects as $\mathscr{C}$
- **Morphisms**: We define $\mathscr{C}_D(A,B) = D(\mathscr{C}(A,B))$

Composition is given as follows:

$$\sum_j q_j|g_j\rangle \circ \sum_i p_i|f_i\rangle = \sum_{i,j} p_i q_j |g_j \circ f_i\rangle$$

Example 5.2. Describing mixing in **CPM**(**Rel**), as discussed in the introduction, was unsatisfactory as we could not encode quantitative data about our state of knowledge. The category $\mathbf{Rel}_D$ encodes a convex set of weights on its morphisms. For example the scalars in $\mathbf{Rel}_D$ correspond to the closed real interval $[0,1]$.

Proposition 5.3. *For category $\mathscr{C}$, $\mathscr{C}_D$ is the free convex algebra enriched category over $\mathscr{C}$.*

Theorem 5.4. *If $\mathscr{C}$ is a compact closed category then $\mathscr{C}_D$ is a compact closed category. The action of the monoidal structure on morphisms extends that of $\mathscr{C}$ as in* (11).

$$\sum_i p_i|f_i\rangle \otimes_D \sum_j q_j|g_j\rangle = \sum_{i,j} p_i q_j |f_i \otimes g_j\rangle \qquad (11)$$

$$(\sum_i p_i|f_i\rangle)^{\dagger_D} = \sum_i p_i|f^\dagger\rangle \qquad (12)$$

There is an identity and surjective on objects strict monoidal embedding $\mathscr{C} \to \mathscr{C}_D$.

If $\mathscr{C}$ is a dagger compact closed category then $\mathscr{C}_D$ is a dagger compact closed category with the dagger extending that of $\mathscr{C}$ as in (12).

Proposition 5.5. *For the finite subdistribution monad we have* [1]*:*

- *The finite distribution monad is the affine part of the subdistribution monad.*
- *The lift monad is the relevant part of the subdistribution monad.*
- *The finite subdistribution monad can be constructed using a distributive law combining the finite distribution monad and the lift monad.*

Remark 5.6. As we saw for unquantified ambiguity, quantified ambiguity is affine, so forming combinations is idempotent. Intuitively, quantified confusion between x and itself is the same as knowing x. Similarly to the unquantified case, we see that the subdistribution monad is exactly the result of combining quantified ambiguity and incomplete information.

Definition 5.7. For category $\mathscr{C}$ we define the category $\mathscr{C}_S$ as having:

- **Objects**: The same objects as $\mathscr{C}$
- **Morphisms**: We define $\mathscr{C}_S(A,B) = S(\mathscr{C}(A,B))$

[1] possibly these observations are well known or folklore, but I am unaware of a suitable prior reference

Composition is given as follows:

$$\sum_j q_j|g_j\rangle \circ \sum_i p_i|f_i\rangle = \sum_{i,j} p_i q_j |g_j \circ f_i\rangle$$

Proposition 5.8. *For category $\mathscr{C}$, $\mathscr{C}_S$ is the free subconvex algebra enriched category over $\mathscr{C}$.*

Theorem 5.9. *If $\mathscr{C}$ is a compact closed category then $\mathscr{C}_S$ is a compact closed category. The action of the monoidal structure on morphisms extends that of $\mathscr{C}$ as in* (13).

$$\sum_i p_i|f_i\rangle \otimes_S \sum_j q_j|g_j\rangle = \sum_{i,j} p_i q_j |f_i \otimes g_j\rangle \qquad (13)$$

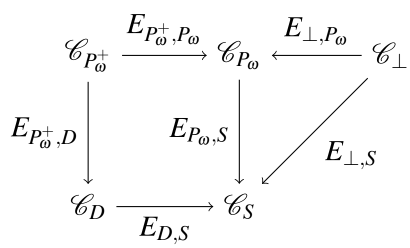
$$(\sum_i p_i|f_i\rangle)^{\dagger_S} = \sum_i p_i|f^{\dagger}\rangle \qquad (14)$$

There is an identity and surjective on objects strict monoidal embedding $\mathscr{C} \to \mathscr{C}_D$.

If $\mathscr{C}$ is a dagger compact closed category then $\mathscr{C}_S$ is a dagger compact closed category with the dagger extending that of $\mathscr{C}$ as in (14).

The various free models of ambiguity and incomplete information that we have constructed can be embedded into each other as follows:

Theorem 5.10. *For category $\mathscr{C}$ there are identity and surjective on objects embeddings:*



where:

$$E_{\perp,P_\omega}(f) = \begin{cases} \emptyset \text{ if } f = \perp \\ \{f\} \text{ otherwise} \end{cases}$$

$$E_{\perp,S}(f) = \begin{cases} \sum_\emptyset \text{ if } f = \perp \\ |f\rangle \text{ otherwise} \end{cases}$$

$$E_{P_\omega^+,P_\omega}(U) = U$$

$$E_{D,S}(\sum_i p_i f_i) = \sum_i p_i f_i$$

$$E_{P_\omega^+,D}(U) = \sum_{u \in U} \frac{1}{|U|}|u\rangle$$

$$E_{P_\omega,S}(U) = \begin{cases} \sum_\emptyset \text{ if } U = \emptyset \\ \sum_{u\in U} \frac{1}{|U|}|u\rangle \text{ otherwise} \end{cases}$$

where $\sum_\emptyset$ denotes the empty formal sum.

Theorem 5.10 shows that $\mathscr{C}_S$ is rich enough to describe all the effects we are interested in, and so serves as a "universal" model of partial information and ambiguity in the category composition distributional semantics framework.

6 Conclusion

In this paper we have considered freely extending the dagger compact closed categories used in categorical distributional models of meaning with sufficient algebraic structure to describe incomplete and ambiguous information. This was done in a systematic manner, constructing suitable bases for enrichment using monad theoretic principles. Our free models effectively record and combine the details of the information we lack. The data explicitly carried by this effectively syntactic construction suggests some algorithmic possibilities to be explored in later work.

Clearly, models other than the free models are of interest. For example, the category of Hilbert spaces and completely positive maps is subconvex algebra enriched, and **Rel** is join semilattice enriched. The category $\mathbf{Rel}_{\mathbf{Subconvex}}$ provides a category that allows non trivial mixing of relations with scalars in $[0,1]$. The author is unaware of other (non-free) models involving relations that allow non-trivial mixing with real scalars, and this remains an open question.

Acknowledgements

I would like to thank Martha Lewis, Bob Coecke, Samson Abramsky and Robin Piedeleu for feedback and discussions. This work was partially funded by the AFSOR grant "Algorithmic and Logical Aspects when Composing Meanings" and the FQXi grant "Categorical Compositional Physics".

References

[1] S. Abramsky & B. Coecke (2004): *A categorical semantics of quantum protocols*. In: *Logic in Computer Science, 2004. Proceedings of the 19th Annual IEEE Symposium on*, IEEE, pp. 415–425.

[2] M. Barr & C. Wells (2005): *Toposes, Triples and Theories*. *Reprints in Theory and Applications of Categories* 1, pp. 1–289.

[3] J. Beck (1969): *Distributive Laws*. In H. Appelgate, M. Barr, J. Beck, F.W. Lawvere, F. Linton, E. Manes, M. Tierney & F. Ulmer, editors: *Seminar on Triples and Categorical Homotopy Theory*, *Lecture Notes in Mathematics* 80, Springer, pp. 119–140.

[4] F. Borceux (1994): *Handbook of Categorical Algebra*. 2, Cambridge University Press.

[5] B. Coecke & A. Kissinger (2016): *Picturing Quantum Processes: A First Course in Quantum Theory and Diagrammatic Reasoning*. Book in preparation.

[6] B. Coecke, M. Sadrzadeh & S. Clark (2010): *Mathematical Foundations for Distributed Compositional Model of Meaning. Lambek Festschrift*. *Linguistic Analysis* 36, pp. 345–384.

[7] S. Eilenberg & J.C. Moore (1965): *Adjoint functors and triples*. *Illinois Journal of Mathematics* 9(3), pp. 381–398.

[8] T. Fritz (2009): *Convex Spaces I: Definition and Examples*. *arXiv preprint arXiv:0903.5522*.

[9] S. Gogioso (2015): *A Bestiary of Sets and Relations*. In Chris Heunen, Peter Selinger & Jamie Vicary, editors: Proceedings of the 12th International Workshop on *Quantum Physics and Logic*, Oxford, U.K., July 15-17, 2015, *Electronic Proceedings in Theoretical Computer Science* 195, Open Publishing Association, pp. 208–227, doi:10.4204/EPTCS.195.16.

[10] C. Heunen & J. Vicary (2016): *Categorical Quantum Mechanics: An Introduction*. Book in preparation.

[11] B. Jacobs (1994): *Semantics of Weakening and Contraction*. *Ann. Pure Appl. Logic* 69.

[12] B. Jacobs (2010): *Convexity, duality and effects*. In: *Theoretical Computer Science*, Springer, pp. 1–19.

[13] B. Jacobs (2011): *Coalgebraic walks, in quantum and Turing computation*. In: *Foundations of Software Science and Computational Structures*, Springer, pp. 12–26.

[14] B. Jacobs (2012): *Introduction to Coalgebra, Towards Mathematics of States and Observation*. Book in preparation.

[15] B. Jacobs & J. Mandemaker (2011): *The expectation monad*. *Quantum* 13, p. 10.

[16] D. Kartsaklis (2014): *Compositional Distributional Semantics with Compact Closed Categories and Frobenius Algebras*. Ph.D. thesis, University of Oxford.

[17] G.M. Kelly (2005): *Basic Concepts of Enriched Category Theory*. Reprints in the Theory and Applications of Category Theory.

[18] A. Kock (1971): *Bilinearity and cartesian closed monads*. Math. Scand. 29, pp. 161–174.

[19] A. Kock (1972): *Strong functors and monoidal monads*. Archiv der Mathematik 23(1), pp. 113–120.

[20] J. Lambek (1997): *Type grammar revisited*. In: Logical aspects of computational linguistics, Springer, pp. 1–27.

[21] H. Lindner (1979): *Affine parts of monads*. Archiv der Mathematik 33(1), pp. 437–443.

[22] S. MacLane (1998): *Categories for the Working Mathematician*. Graduate Texts in Mathematics 5, Springer.

[23] E.G. Manes (1976): *Algebraic Theories*. Graduate Texts in Mathematics 26, Springer.

[24] D. Marsden (2015): *A Graph Theoretic Perspective on CPM(Rel)*. In Chris Heunen, Peter Selinger & Jamie Vicary, editors: Proceedings of the 12th International Workshop on Quantum Physics and Logic, Oxford, U.K., July 15-17, 2015, Electronic Proceedings in Theoretical Computer Science 195, Open Publishing Association, pp. 273–284, doi:10.4204/EPTCS.195.20.

[25] E. Moggi (1991): *Notions of computation and monads*. Information and computation 93(1), pp. 55–92.

[26] R. Piedeleu (2014): *Ambiguity in Categorical Models of Meaning*. Master's thesis, University of Oxford.

[27] R. Piedeleu, D. Kartsaklis, B. Coecke & M. Sadrzadeh (2015): *Open System Categorical Quantum Semantics in Natural Language Processing*. In L. S. Moss & P. Sobocinski, editors: 6th Conference on Algebra and Coalgebra in Computer Science, CALCO 2015, June 24-26, 2015, Nijmegen, The Netherlands, LIPIcs 35, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, pp. 270–289.

[28] G. D. Plotkin & J. Power (2002): *Notions of Computation Determine Monads*. In M. Nielsen & U. Engberg, editors: Foundations of Software Science and Computation Structures, 5th International Conference, FOSSACS 2002. Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2002 Grenoble, France, April 8-12, 2002, Proceedings, Lecture Notes in Computer Science 2303, Springer, pp. 342–356.

[29] P. Selinger (2007): *Dagger compact closed categories and completely positive maps*. Electronic Notes in Theoretical Computer Science 170, pp. 139–163.

[30] C.C. Shan (2001): *Monads for Natural Language Semantics*. In K. Streignitz, editor: Proceedings of the ESSLLI-2001 Student Session, Institute for Logic, Language and Computation, University of Amsterdam, pp. 285–298.

[31] P. Wadler (1995): *Monads for functional programming*. In: Advanced Functional Programming, Springer, pp. 24–52.