Generalized Relations in Linguistics and Cognition

Bob Coecke, Fabrizio Genovese, Martha Lewis, and Dan Marsden[✉]

Department of Computer Science, University of Oxford, Oxford, UK
{bob.coecke,fabrizio.genovese,martha.lewis,
daniel.marsden}@cs.ox.ac.uk

Abstract. Categorical compositional models of natural language exploit grammatical structure to calculate the meaning of sentences from the meanings of individual words. This approach outperforms conventional techniques for some standard NLP tasks. More recently, similar compositional techniques have been applied to conceptual space models of cognition.

Compact closed categories, particularly the category of finite dimensional vector spaces, have been the most common setting for categorical compositional models. When addressing a new problem domain, such as conceptual space models of meaning, a key problem is finding a compact closed category that captures the features of interest.

We propose categories of generalized relations as source of new, practical models for cognition and NLP. We demonstrate using detailed examples that phenomena such as fuzziness, metrics, convexity, semantic ambiguity and meaning that varies with context can all be described by relational models. Crucially, by exploiting a technical framework described in previous work of the authors, we also show how we can combine multiple features into a single model, providing a flexible family of new categories for categorical compositional modelling.

1 Introduction

Distributional models of language describe the meaning of a word using co-occurrence statistics derived from corpus data. A central question with these models is how to combine meanings of individual words, in order to understand phrases and sentences. Categorical compositional models of natural language [15] address this problem, providing a principled approach to combining the meanings of words to form the meanings of sentences, by exploiting their grammatical structure. They also outperform conventional techniques for some standard NLP tasks [23,29].

Distributional models of language can be thought of as "process theories" [16] A process theory consists of a graphical language for reasoning about composite systems of abstract processes, and a categorical semantics modelling the application domain. A particularly important class of categorical models are the compact closed categories, which come equipped with an elegant graphical calculus. Process theoretic models built upon compact closed categories have been

successfully exploited in many application areas, including quantum computation [1], signal flow graphs [11], control theory [2], Markov processes [4], electrical circuits [3] and even linear algebra [43].

Recently [9], the categorical compositional approach to meaning has been applied to the conceptual space models of human cognition introduced in [21,22]. When addressing a new application domain, it is necessary to identify a compact closed category with mathematical structure compatible with the application phenomena of interest.

Amongst the compact closed categories the hypergraph categories [20] are a particularly well behaved class of practical interest. In [33] we presented a flexible parameterized mathematical framework for constructing hypergraph categories. We view this framework as a practical tool for building new models in a principled manner, by varying the parameter choices according to the needs of the application domain. These models are based upon generalizing the well understood notion of a binary relation, providing a concrete and intuitive setting for model development.

In the present work we demonstrate, via extensive examples, that categories of generalized relations present an attractive setting for constructing new models of language and cognition. We emphasize the intuitive interpretation of the models under construction, and their connections to concrete ideas in computation, NLP and further afield. These examples are structured as follows:

- In Sect. 3 we introduce relations with generalized truth values, and exploit them to model features such as distances, forces, connectivity and fuzziness. Relations with generalized truth values are well known in the mathematical community, but seem to have received little attention from the perspective of compositional semantics, with the recent exception of [19].
- In Sect. 4 we generalize relations in another direction, considering relations that respect algebraic structure. These relations can capture features such as convexity, which is important in conceptual spaces models [21,22]. In this case, we recover a model first used in [9], originally constructed in an ad-hoc manner using techniques from monad theory and the theory of regular categories. Importantly, we then show that we can combine generalized truth values with relations respecting algebraic structure, providing conceptual space models with access to distance measures.
- In Sect. 5 we view spans as generalized "proof aware" relations in which the apex of the span contains witnesses to relatedness between the domain and codomain. Spans can be extended to support generalized truth values, and to respect algebraic structure. Exploiting a combination of these features, we construct a new model of semantic ambiguity in conceptual space models of natural language, in which different proof witnesses allow us to vary how strongly different words are related, depending on how they are interpreted.
- The previous examples were essentially built upon the category of sets. Our techniques can be applied with different choices of ambient topos. In Sect. 6, as a practical example of this feature, we use presheaf toposes to build models

in which meanings can vary with context, such as the progress of time or states of the world.

All of our models are preorder enriched, providing a natural candidate for modelling semantic entailment or hyponymy [5,6]. Preorder enrichment also means we can consider internal monads within our various categories of relations. We emphasize the importance of these internal monads throughout our discussions. They provide access to important structured objects such as preorders, generalized metric spaces and ultrametric spaces, and similar well behaved relationships when we combine various modelling features.

2   Compositional Models of Meaning

The grammatical structure of natural language can be modelled using Lambek's pregroup grammars [31].

Definition 1. *A* **pregroup** *is a tuple* $(X, \cdot, 1, (-)^l, (-)^r, \leq)$ *where* $(X, \cdot, 1, \leq)$ *is a partially ordered monoid, or pomonoid, and* $(-)^r, (-)^l$ *are unary functions of type* $X \to X$ *such that for all* $x \in X$ *the following conditions hold,*

$$1 \leq x \cdot x^l \qquad x^l \cdot x \leq 1 \qquad 1 \leq x^r \cdot x \qquad x \cdot x^r \leq 1$$

We say that x **reduces** *to* x' *if* $x \leq x'$.

A grammar is typically described using the free pregroup over some set of basic types. For example, we may consider the free pregroup of the set $\{n, s\}$ where n and s are basic types for nouns and sentences respectively. More complex terms are then built up using the algebraic operations, for example the type of a transitive verb is $n^r s n^l$. We can calculate the type of a phrase by composing the types of the individual terms using the monoid multiplication. For example, the phrase "mice eat cheese" has type $n(n^r s n^l)n$. A composite term is then a well typed sentence if its type reduces to the sentence type. For example:

$$n(n^r s n^l)n = (nn^r)s(n^l n) \leq s(n^l n) \leq s$$

and so "mice eat cheese" is a well typed sentence. In this way, pregroups give us access to the *compositional* features of language.

On the other hand, distributional models [40] of the *meaning* of words in natural language are built using vector space models automatically derived from co-occurrence statistics in a large corpus of text. The key observation of the categorical compositional approach to natural language is that both pregroups and the category of finite dimensional real vector spaces carry the same categorical structure, that of an autonomous category.

Definition 2. *A monoidal category* $\mathcal{V}$ *has left/right duals if every object has an internal left/right adjoint when* $\mathcal{V}$ *is regarded as a one object bicategory. An* **autonomous category** *is a monoidal category in which every object has both left and right duals. A* **compact closed category** *is a symmetric monoidal category in which every object has right duals.*

A straightforward application oriented introduction to monoidal categories and compact closed categories can be found in [17].

This observation can be exploited to derive the meanings of sentences from the meanings of words. We fix a strong monoidal functor from a pregroup describing grammatical structure to the category of finite dimensional vector spaces. This functor maps type reductions to linear maps, allowing us to automatically derive the meaning of a sentence from its constituent parts. Clearly, this approach can be seen as an instance of functorial semantics. By varying the domain and preserved structure we can consider different categorial grammars [14]. By varying the codomain we can consider different models, as has been important in recent work broadening the scope to mathematical models of cognition [9,10]. When varying the category of meanings, it is desirable to remain within the domain of compact closed categories, in order to exploit connections with previous linguistic developments, and to retain access to their powerful graphical calculus.

The question then becomes: *How can we find or construct compact closed categories with desirable mathematical properties?* This is the question we explore in this paper. In fact, our constructions produce a subclass of compact closed categories, referred to as hypergraph categories [20,30], and so this is where we shall focus our attention.

Definition 3. *A* **hypergraph category** *is a symmetric monoidal category in which every object is equipped with a choice of special commutative Frobenius algebra, coherently with the monoidal structure.*

Details of the notion of a Frobenius algebra, and linguistic applications including modelling relative pronouns can be found in [38,39]. If I is the monoidal unit, we will occasionally refer to morphisms of types $I \to X$ and $X \to I$ as the **states** and **effects** of X. Morphisms of type $I \to I$ are referred to as **numbers**.

Example 1. The category **Rel** of sets and binary relations between them can be given the structure of a hypergraph category. The monoidal structure is given by forming Cartesian products of sets. A state of a set X is a subset of X and the numbers are the Boolean truth values. The Frobenius algebra is given by the copying relation $x \sim (x,x) : X \to X \times X$, the deletion relation $x \sim * : X \to I$, and their converses.

All the compact closed categories discussed in this paper will be hypergraph categories, generalizing Example 1 along different axes of variation.

3   Generalized Truth Values

A binary relation $R : A \to B$ between sets can be identified with a characteristic function of type $A \times B \to \{\top, \bot\}$ mapping the related pairs of elements to $\top$. It is fruitful to consider generalizing the codomain of such characteristic functions to a set Q, thought of as a collection of truth values. We can then consider functions of the form $A \times B \to Q$ as generalized relations, with truth values

in Q. In order for the corresponding binary relations to have satisfactory notions of identities and composition, the set Q must carry the structure of a quantale.

Definition 4 (Quantale). *A* **quantale** *is a join complete partial order Q with a monoid structure $(\otimes, k)$ satisfying the following distributivity axioms, for all $a, b \in Q$ and $A, B \subseteq Q$:*

$$a \otimes \left[\bigvee B\right] = \bigvee\{a \otimes b \mid b \in B\} \qquad \left[\bigvee A\right] \otimes b = \bigvee\{a \otimes b \mid a \in A\}$$

A quantale is said to be **commutative** *if its monoid structure is commutative.*

All the quantales encountered in this paper will be commutative. We introduce some examples of importance in later developments.

Example 2. The **Boolean quantale** is given by the two element complete Boolean algebra $\mathbf{B} = \{\top, \bot\}$, with the join and multiplication given by the join and meet in the Boolean algebra.

Example 3. The **Lawvere quantale L** is given by the chain $[0, \infty]$ of extended positive reals with the *reverse* ordering, hence minima in $[0, \infty]$ provide the joins of the quantale, and the monoid structure is given by addition.

Example 4. The quantale **F** has again the extended positive reals with reverse order as its partial order, but now with max as the monoid multiplication.

Example 5. The **interval quantale I** is given by the ordered interval $[0, 1]$ with minima as the monoid structure.

For a quantale Q, the Q-relations form a category $\mathbf{Rel}(Q)$ with composition and identities[1]

$$(S \circ R)(a, c) = \bigvee_b R(a, b) \otimes S(b, c) \qquad 1_A(a, b) = \bigvee\{k | a = b\}$$

If Q is a commutative quantale, $\mathbf{Rel}(Q)$ carries a symmetric monoidal structure, with the tensor product of objects given by the cartesian product of sets, and the action on relations given for $R : A \to C$ and $S : B \to D$ by

$$(R \otimes S)(a, b, c, d) = R(a, c) \otimes S(b, d)$$

The singleton set is the monoidal unit. A key observation from the perspective of this paper is:

Theorem 1. $\mathbf{Rel}(Q)$ *is compact closed with respect to this monoidal structure.*

Now that we have described how Q-relations compose, we can consider computational interpretations for our example choices of quantale.

---

[1] The slightly unusual formulation of identities is to avoid definition by cases. This means they can be interpreted in the internal language of an arbitrary topos.

Example 6. The relations over the Lawvere quantale **L** can be thought of as describing costs. The value $R(a, b)$ describes the cost of converting a into b. A cost of 0 means they are maximally related and can be freely inter-converted. A cost of ∞ indicates completely unrelated values, that cannot be converted between each other for finite cost. The value $(S \circ R)(a, c)$ describes the cheapest way of converting a into some b, and then converting that b into c, and adds the associated costs. If we perform two conversions in parallel $(R \otimes R')(a, a', b, b')$ describes the sum of the two individual conversion costs.

In this setting, we can think of a state $I \to A$ as giving a table of costs for acquiring the resources in A, and similarly an effect $A \to I$ is a table of costs for disposing of resources in A.

Example 7. The quantale **F** has the same underlying set as the Lawvere quantale, but its different algebraic structure leads to a very different interpretation. We think of $R(a, b)$ as the peak force required to move a to b. The value given by the composite $(S \circ R)(a, c)$ then describes optimum peak force we will require to move a to c. For example if we can convert a to b with one unit of force, and then move b to c for two units of force, then the peak force required is two units. An alternative procedure converting a to b' for zero units of cost, and then converting b' to c for 2.5 units of cost has a peak cost of 2.5 units, so we would prefer the first procedure to minimize our peak effort. Similarly, the truth value $(R \otimes R')(a, a', b, b')$ gives the peak force required to complete both conversions, assuming these costs are independently incurred.

As with Example 6, we can think of states and effects as tables of acquisition and elimination forces.

Example 8. We can interpret ordinary relations over the Boolean quantale as modelling connectivity. $R(a, b)$ tells us that a is connected to b, composition tells us that we can chain connections together, and the tensor product tells us that we can connect pairs of elements together using a pair of connections between their components. Generalizing to the interval quantale, we now think of $R(a, b)$ as a "connection strength" between a and b. The composite $(S \circ R)(a, c)$ gives the best connection quality that we can achieve in two steps via B. Similarly, the parallel composite $(R \otimes R')(a, a', b, b')$ gives a conservative judgment of the connection quality we can achieve simultaneously between both a and b and a' and b' as the lower of the two individual connection strengths. States describe the "transmission strength" with which signals enter the system from the environment, and effects describe the "reception quality" when consuming output signals.

Alternatively, we could view relations over **I** as fuzzy relations, with states and effects sets with fuzzy membership, and fuzzy predicates. Graded membership is widely used in cognitive science, for example in [8, 18, 24, 25, 37]. Concepts such as 'tall' have no crisp boundary and are better modelled using grades of membership. Although human concept use does not obey fuzzy logic [35], fuzzy relations may prove useful.

Rel(Q) is partial order-enriched if we order relations pointwise with respect to the underlying quantale order. It therefore makes sense to consider internal monads in **Rel**(Q) as interesting "structured objects". An internal monad on an object in a partially ordered category is an endomorphism R satisfying:

$$(R \circ R) \subseteq R, \qquad 1_A \subseteq R \tag{1}$$

Example 9. If we specialize condition (1) to **Rel**(**L**), it is equivalent to:

$$R(a, b) + R(b, c) \geq R(a, c), \qquad 0 = R(a, a)$$

We therefore consider these internal monads as describing *generalized metric spaces*. This observation is important in the field of monoidal topology [26].

As before, we can also interpret our internal monad as giving a well behaved collection of conversion costs between resources. Converting a resource to itself is free, and converting a resource via an intermediate state is at least as expensive as taking the direct route. Similarly, if we consider **Rel**(**F**) the conditions of (1) become:

$$\max(R(a, b), R(b, c)) \geq R(a, c), \qquad 0 = R(a, a)$$

and we can therefore see such internal monads as *generalized ultrametric spaces*. Again, the interpretation in terms of maximum force requirements extends to a sensible interpretation of these axioms.

Example 10. Internal monads in the category of ordinary relations are preorders on their underlying set. The generalization to the interval quantale then gives a fuzzy generalization of the notion of preorder. We can also apply our intuition in terms of connection strengths. Reflexivity tells us that every element can be perfectly connected to itself. Transitivity tell us that the optimal connection strength available is always at least as good as connecting via an intermediate node.

4   Incorporating Convexity

Up to this point, the domain and codomain of our relations have been sets. If we fix an algebraic structure (Σ, E) with set of operations Σ and equations between terms E, we can define a notion of binary relation between these algebras.

Definition 5. *An **algebraic Q-relation** of type $A \to B$ is an ordinary Q-relation R between the underlying sets, such that for each operation $\sigma \in \Sigma$ of arity n the following inequation holds in the quantale order:*

$$R(a_1, b_1) \otimes ... \otimes R(a_n, b_n) \leq R(\sigma(a_1, ..., a_n), \sigma(b_1, ..., b_n))$$

As shown in [33], algebraic Q-relations form a hypergraph category:

Theorem 2. *For commutative quantale Q and algebraic signature (Σ, E) there is a hypergraph category* **Rel**$_{(\Sigma, E)}(Q)$ *with objects (Σ, E)-algebras and morphisms algebraic Q-relations.*

In the conceptual spaces literature, convexity is conceptually important. In [9] this convexity was captured using relations between convex algebras. We refer to [9] and the extended paper [10] for explicit modelling of toy computations of composed concepts in this category.

These convex algebras can be described as the Eilenberg-Moore algebras of the finite distribution monad. They can in fact be presented by a family Σ_c of binary operations

$$+^p, \quad p \in (0,1)$$

satisfying suitable axioms. We can read $x +^p y$ as "choose x with probability p and y with probability $(1-p)$". By considering algebraic **B**-relations over this signature, we can construct a category isomorphic to the category **ConvexRel** of convex relations from [9]. By changing our quantale of truth values, we can go further than this.

Proposition 1. *In the category of convex* **L***-relations, the internal monads are generalized metric spaces satisfying the additional axioms for $p \in (0,1)$:*

$$R(a_1, b_1) + R(a_2, b_2) \geq R(a_1 +^p b_1, a_2 +^p b_2)$$

So internal monads in the category of convex relations over the Lawvere quantale are generalized metric spaces that interact well with formation of convex mixtures. The usual distance on $\mathbb{R}^n$ is an example of such a metric.

As shown in [33], every quantale homomorphism $h : Q_1 \to Q_2$ induces a strict monoidal functor of type $\mathbf{Rel}_{(\Sigma,E)}(Q_1) \to \mathbf{Rel}_{(\Sigma,E)}(Q_2)$. If the quantale morphism is injective, this functor is faithful. In particular, the mapping $\perp \mapsto \infty; \top \mapsto 0$ is an injective quantale homomorphism from the Boolean to the Lawvere quantale. This means we can find the ordinary Boolean binary relations as a monoidal subcategory of the category $\mathbf{Rel}(\mathbf{L})$. This presents some flexible modelling possibilities. If U and V are two subsets of a set X, they induce two states $U, V : I \to X$ in $\mathbf{Rel}(\mathbf{B})$. If we consider the number $V^\circ \circ U$, where R° denotes relational converse, it evaluates to true if and only if $U \cap V \neq \emptyset$.

Proposition 2. *If $U, V \subseteq X$ and d is an internal monad in $\mathbf{Rel}(\mathbf{L})$, the composite $V^\circ \circ d \circ U$ is the infimum of the distances between elements in U and V.*

This gives us the greatest lower bound on the distances between elements in U and V, providing a finer grain measure of similarity than can conventionally be achieved in relational models. We note that as distances are in general asymmetric, the number $U^\circ \circ d \circ V$ may give a different measure of similarity. Similarly, we can find the ordinary Boolean convex relations within the category of **L**-valued convex relations, presenting analogous opportunities for performing calculations with discrete convex relations, and then measuring their separation on a continuum of values.

Such asymmetric distance measures are of practical use in cognitive science applications. A fundamental concept in psychology is that of similarity, which can be used as the basis of concept formation. Similarity between objects or concepts can be explained by locating objects in some sort of conceptual or

feature space, and modelling similarity as a function of distance, for example in [42]. However, judgements of similarity are not necessarily symmetric [45]. In one study examining the similarity between pairs of countries, participants are asked to choose between statements 'Country A is similar to country B' or 'Country B is similar to country A'. In all cases, a majority of participants preferred the statement where the latter country was considered more prominent.

5   Proof Relevance

A **span** S of sets, between sets A and B, is a set X and a pair of functions $X \xrightarrow{p_1} A$ and $X \xrightarrow{p_2} B$. Paralleling the notation for relations, we will write

$$S_x(a, b) := x \in X \ \wedge \ p_1(x) = a \ \wedge \ p_2(x) = b$$

We can think of such a span as a *proof relevant relation* in which $S_x(a, b)$ tells us that x witnesses that a and b are related. In a computational linguistics or cognition application where relations may have been derived automatically from data in some way, we can exploit these proof witnesses to track evidence for our beliefs that certain relationships hold.

Sets and spans between them form a hypergraph category **Span** with composition given by pullback, and tensor product induced by a choice of products[2]. In fact, as we did for relations, we can extend these spans with algebraic structure and a choice of truth values in a partially ordered monoid. We no longer require full quantale structure on our truth values, as multiple proof witnesses mean we don't need to choose a single representative truth value when composing relations.

Definition 6. *For an algebraic signature (Σ, E) and pomonoid Q an* **algebraic Q-span** *of type $A \to B$ between (Σ, E)-algebras is a span $A \xleftarrow{p_1} X \xrightarrow{p_2} B$ between the underlying objects, with a* **characteristic morphism** $\chi : X \to Q$. *We require that the algebraic structure is respected in that for all $\sigma \in \Sigma$, with arity n:*

$$\bigwedge_{1 \leq i \leq n} (p_1(x_i) = a_i \wedge p_2(x_i) = b_i) \Rightarrow \bigotimes_{1 \leq i \leq n} \chi(x_i) \leq \chi(\sigma(x_1, ..., x_n))$$

Intuitively, these are intensional relations in which proof witnesses are weighted by a truth value, and the relations respect the algebraic structure. As shown in [33], algebraic Q-spans also form a hypergraph category:

Theorem 3. *For commutative pomonoid Q and algebraic signature (Σ, E) there is a hypergraph category* **Span**$_{(\Sigma, E)}(Q)$ *with objects (Σ, E)-algebras and morphisms algebraic Q-spans.*

---

[2] In fact, in order for composition to be associative, it is necessary to work with equivalence classes of spans. It is sufficient to consider representatives, and we do so to avoid distracting technicalities.

For algebraic Q span S we define

$$S_x^q(a, b) := x \in X \ \wedge \ p_1(x) = a \ \wedge \ p_2(x) = b \ \wedge \ \chi(x) = q$$

We then read $S_x^q(a, b)$ as telling us that x witnesses that a and b are related with strength q. In fact, we can order algebraic Q-spans in a manner similar to that for relations, but accounting for proof witnesses.

Definition 7. *For pomonoid Q, we define a preorder on algebraic Q-spans by setting $(X_1, f_1, g_1, \chi_1) \subseteq (X_2, f_2, g_2, \chi_2)$ if there is a **Set**-monomorphism $\varphi : X_1 \to X_2$ such that $f_1 = f_2 \circ \varphi$, $g_1 = g_2 \circ \varphi$ and $\forall x \,.\, \chi_1(x) \leq \chi_2(\varphi(x))$.*

The ordering accounts pointwise for strengths of relatedness in a natural way. The requirement that the function φ in Definition 7 is a monomorphism ensures that even if our truth values are trivial, we take account of the "number" of proof witnesses available.

As internal monads provided interesting objects in the setting of relations, we should consider them in the span setting as well.

Proposition 3. *An internal monad on A in **Span**(**L**) is an **L**-span $S : A \to A$ such that if $S_x^p(a_1, a_2)$ and $S_y^q(a_2, a_3)$ we can choose an element $\varphi(x, y)$ of the apex such that $S_{\varphi(x,y)}^r(a_1, a_3)$ and $p + q$ is greater than r in the usual ordering on the real numbers. Furthermore, we can do this in a way such that the assignment φ is injective.*

So internal **L**-span monads further generalize metric spaces to incorporate multiple possible distances, which we can think of as describing different paths between points. We now outline a new practical application of spans in models of language.

Example 11 (Semantic Ambiguity via Spans). In natural language, we often encounter ambiguous situations. For example the word "bank" can refer to either a "river bank" or a "financial bank". A compositional account of semantic ambiguity was presented in [36], using mathematical models of incomplete information from quantum theory. The techniques applied implicitly assume meanings are built upon a vector space model, to which we apply Selinger's CPM construction [41] to yield a new category of ambiguous meanings. The CPM construction can also be applied to categories of relations, but in this case it does not provide a satisfactory model of ambiguity [34].

An alternative approach to ambiguity in relational models is to use spans. We consider how the ambiguous word "bank" is related to the word "water"

– In the "river bank" context, we would expect a strong relationship
– In the "financial bank" context, we would expect a weaker relationship

By using spans rather than relations, we can introduce two different proof witnesses for the different contexts under consideration. By choosing our quantale of truth values to be the Lawvere quantale **L**, we can attach a different choice of distance to each of these choices. As we compose spans to describe the meanings of phrases and sentences, the proof witnesses will keep track of the different possible relationships in play.

6   Variable Contexts

Our definitions of algebraic Q-relations and algebraic Q-spans are constructive. This means that Theorems 2 and 3 continue to hold for any elementary topos, as proved in [33]. Standard sources on topos theory are [12,27,28,32]. We will write $\mathbf{Rel}^{\mathcal{E}}_{(\Sigma,E)}(Q)$ and $\mathbf{Span}^{\mathcal{E}}_{(\Sigma,E)}(Q)$ for the categories of spans and relations, to make the choice of topos $\mathcal{E}$ explicit. This generalization has practical implications if we move to different choices of background topos.

Definition 8. *Let $\mathcal{C}$ be a small category. A* **presheaf** *on $\mathcal{C}$ is a functor of type $\mathcal{C}^{op} \to \mathbf{Set}$. Presheaves and natural transformations between them form a topos, denoted $\mathbf{Set}^{\mathcal{C}^{op}}$. For presheaf X over a preorder, we will write X_i for the set in the image under X of element i of the preorder, and $X_{i,j}$ for the image of $j \leq i$ under X.*

Presheaves can be interpreted as *sets varying with context*. This is exactly the perspective we shall adopt in our examples. To exploit our generalized span construction, we need to describe internal pomonoids in presheaf categories.

Lemma 1. *A commutative partially ordered monoid in a presheaf category $\mathbf{Set}^{\mathcal{C}^{op}}$ is a presheaf Q such that for each $\mathcal{C}$-object x and $\mathcal{C}$-morphism f, $Q(x)$ is a commutative pomonoid and $Q(f)$ is a pomonoid morphism in $\mathbf{Set}$. See [28, D1.2.14].*

Example 12 (Temporal dependence). In Example 11 we modelled ambiguity using multiple proof witnesses to describe different interpretations of words. We now investigate the description of time dependent ambiguous relationships, by exploiting spans over presheaves. To do so, we consider presheaves over the partial order $\mathbb{N} = 0 \leftarrow 1 \leftarrow 2...$ having objects natural numbers. We view these presheaves as sets varying in time. We assume our notion of truth is fixed, and so we will consider $\mathbf{Span}^{\mathbf{Set}^{\mathbb{N}^{op}}}(\mathbf{L})$, where $\mathbf{L}$ is the constant presheaf on the pomonoid underlying the Lawvere quantale. An $\mathbf{L}$-span between presheaves X and Y then consists of natural transformations $p_1 : X \Rightarrow A$ and $p_2 : X \Rightarrow B$, and a characteristic natural transformation $\chi : X \Rightarrow \mathbf{L}$. We see naturality as a consistency condition between the relationships described by proof witnesses, as they move forward in time. As our pomonoid is constant, $\chi_i(x) = \chi_j(X_{i,j}(x))$, so the truth value associated with a proof witness is preserved through time. Intuitively, in this model, a steadily increasing collection of relationships hold over time.

Example 13 (Perspective Dependence). In Example 12, the truth object was fixed in all contexts. We now examine a brief example in which our notion of truth is context dependent. Consider two agents. Agent 0 has a binary view of the world, relationships either hold or they don't. Agent 1 has a richer view incorporating different strengths of relation in the unit interval. Consider presheaves on the category $\mathcal{C}$ with a single non-trivial arrow $0 \leftarrow 1$. We define an internal pomonoid Q with $Q(0) = \mathbf{B}$, $Q(1) = \mathbf{I}$ and $Q_{0,1}$ the canonical pomonoid morphism between the Boolean and interval quantales. Now if we consider a Q-span

between constant presheaves A and B with apex an arbitrary presheaf X, we can think of it as follows. Each element of X_0 relates two elements $a \in A$ and $b \in B$ with strength 0 or 1. The structure of X then forces that X_1 contains a witness relating those two elements with the same strength. As X_1 encodes the views of the more powerful agent, it may describe additional relationships, now with strengths weighted in the interval $[0, 1]$.

If we wish to consider algebraic Q-relations over an arbitrary topos things are more delicate since internal quantales cannot be defined pointwise. Nevertheless there are standard sources of internal commutative quantales, for example:

– If $\mathcal{C}$ is a groupoid and Q is a commutative quantale in **Set**, then Q can be lifted to an internal commutative quantale in $\mathbf{Set}^{\mathcal{C}^{op}}$.
– The subobject classifier Ω of a topos is an internal locale, and therefore an internal commutative quantale.

We conclude by establishing the relationship between our framework of generalized relations and the standard notion of the category of relations over a regular category. This will involve the internal locale given by the subobject classifier.

Definition 9. *A category $\mathcal{C}$ is* **regular** *if it is finitely complete, every kernel pair has a coequalizer and regular epimorphisms are stable under pullback.*

There is standard construction of a category of relations $\mathbf{Rel}(\mathcal{C})$ of a regular category $\mathcal{C}$, see for example [13]. For the category **Set** for example, this construction recovers exactly the usual category of binary relations. As we have been constructing categories of relations in this paper, it would be interesting to know how this relates to the relations of a regular category. Every topos is regular, and in fact for any algebraic theory (Σ, E), the category of internal (Σ, E)-algebras in a regular category [7], meaning we can consider the impact of algebraic structure. In fact, the resulting category of relations is equivalent to the one produced by our construction with the subobject classifier as the object of truth values.

Theorem 4. *Let $\mathcal{E}$ be a topos, Ω its subobject classifier and (Σ, E) an algebraic signature. The category $\mathbf{Rel}^{\mathcal{E}}_{(\Sigma, E)}(\Omega)$ resulting from the algebraic Q-relations construction is equivalent to the category of internal relations over the regular category of internal (Σ, E)-algebras in $\mathcal{E}$.*

In this way, we see that relations over suitable regular categories are a special case of our construction.

7 Conclusion

We have demonstrated that categories of generalized relations present a flexible modelling tool for categorical compositional models of natural language and cognition. We presented various potential models worthy of further investigation, capturing features such as fuzziness, distances, convexity, ambiguity and context sensitivity, and showed how these features can be used in combination within

a generic framework. One natural direction for further work would be empirical investigation of the compatibility of these theoretical models with concrete applications. Another one would be to investigate whether the techniques in [44] can be used to build models with either non-commutative or typed quantales, known as quantaloids.

References

1. Abramsky, S., Coecke, B.: A categorical semantics of quantum protocols. In: Proceedings of the 19th Annual IEEE Symposium on Logic in Computer Science, pp. 415–425. IEEE (2004)
2. Baez, J.C., Erbele, J.: Categories in control. Theory Appl. Categ. **30**(24), 836–881 (2015)
3. Baez, J.C., Fong, B.: A compositional framework for passive linear networks. arXiv preprint arXiv:1504.05625 (2015)
4. Baez, J.C., Fong, B., Pollard, B.S.: A compositional framework for Markov processes. J. Math. Phys. **57**(3), 033301 (2016)
5. Bankova, D., Coecke, B., Lewis, M., Marsden, D.: Graded entailment for compositional distributional semantics. arXiv preprint arXiv:1601.04908 (2015)
6. Bankova, D.: Comparing meaning in language and cognition - p-hypononymy, concept combination, asymmetric similarity. Master's thesis, University of Oxford (2015)
7. Barr, M.: Exact categories. Exact Categories and Categories of Sheaves. LNM, vol. 236, pp. 1–120. Springer, Heidelberg (1971). doi:10.1007/BFb0058580
8. Barsalou, L.W.: Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. J. Exp. Psychol. Learn. Mem. Cogn. **11**(4), 629 (1985)
9. Bolt, J., Coecke, B., Genovese, F., Lewis, M., Marsden, D., Piedeleu, R.: Interacting conceptual spaces. In: Kartsaklis, D., Lewis, M., Rimell, L. (eds.) Proceedings of the 2016 Workshop on Semantic Spaces at the Intersection of NLP, Physics and Cognitive Science, SLPCS@QPL 2016, Glasgow, Scotland, 11 June 2016. EPTCS, vol. 221, pp. 11–19 (2016). http://dx.doi.org/10.4204/EPTCS.221.2
10. Bolt, J., Coecke, B., Genovese, F., Lewis, M., Marsden, D., Piedeleu, R.: Interacting conceptual spaces I: Grammatical composition of concepts. arXiv preprint arXiv:1703.08314 (2017)
11. Bonchi, F., Sobocinski, P., Zanasi, F.: Full abstraction for signal flow graphs. ACM SIGPLAN Not. **50**(1), 515–526 (2015)
12. Borceux, F.: Handbook of Categorical Algebra: Volume 3, Categories of Sheaves. Cambridge University Press, Cambridge (1994)
13. Borceux, F.: Handbook of Categorical Algebra: Volume 2, Categories and Structures, vol. 2. Cambridge University Press, Cambridge (1994)
14. Coecke, B., Grefenstette, E., Sadrzadeh, M.: Lambek vs. Lambek: functorial vector space semantics and string diagrams for lambek calculus. Ann. Pure Appl. Logic **164**(11), 1079–1100 (2013)

15. Coecke, B., Sadrzadeh, M., Clark, S.: Mathematical foundations for distributed compositional model of meaning. Lambek festschrift. Linguist. Anal. **36**, 345–384 (2010)
16. Coecke, B., Kissinger, A.: Picturing Quantum Processes. A First Course in Quantum Theory and Diagrammatic Reasoning. Cambridge University Press (2017, forthcoming)
17. Coecke, B., Paquette, E.O.: Categories for the practising physicist. In: Coecke, B. (ed.) New Structures for Physics, pp. 173–286. Springer, Heidelberg (2010)
18. Dale, R., Kehoe, C., Spivey, M.J.: Graded motor responses in the time course of categorizing atypical exemplars. Mem. Cogn. **35**(1), 15–28 (2007)
19. Dostal, M., Sadrzadeh, M.: Many valued generalised quantifiers for natural language in the DisCoCat model. Technical report, Queen Mary University of London (2016)
20. Fong, B.: The algebra of open and interconnected systems. Ph.D. thesis, University of Oxford (2016)
21. Gärdenfors, P.: Conceptual Spaces: The Geometry of Thought. MIT Press, Cambridge (2004)
22. Gärdenfors, P.: The Geometry of Meaning: Semantics Based on Conceptual Spaces. MIT Press, Cambridge (2014)
23. Grefenstette, E., Sadrzadeh, M.: Experimental support for a categorical compositional distributional model of meaning. In: The 2014 Conference on Empirical Methods on Natural Language Processing, pp. 1394–1404 (2011). arXiv:1106.4058
24. Hampton, J.A.: Disjunction of natural concepts. Mem. Cogn. **16**(6), 579–591 (1988)
25. Hampton, J.A.: Overextension of conjunctive concepts: evidence for a unitary model of concept typicality and class inclusion. J. Exp. Psychol. Learn. Mem. Cogn. **14**(1), 12 (1988)
26. Hofmann, D., Seal, G.J., Tholen, W.: Monoidal Topology: A Categorical Approach to Order, Metric, and Topology, vol. 153. Cambridge University Press, Cambridge (2014)
27. Johnstone, P.T.: Sketches of an Elephant: A Topos Theory Compendium, vol. 1. Oxford University Press, Oxford (2002)
28. Johnstone, P.T.: Sketches of an Elephant: A Topos Theory Compendium, vol. 2. Oxford University Press, Oxford (2002)
29. Kartsaklis, D., Sadrzadeh, M.: Prior disambiguation of word tensors for constructing sentence vectors. In: The 2013 Conference on Empirical Methods on Natural Language Processing, pp. 1590–1601. ACL (2013)
30. Kissinger, A.: Finite matrices are complete for (dagger-)hypergraph categories. arXiv preprint arXiv:1406.5942 (2014)
31. Lambek, J.: Type grammar revisited. In: Lecomte, A., Lamarche, F., Perrier, G. (eds.) LACL 1997. LNCS, vol. 1582, pp. 1–27. Springer, Heidelberg (1999). doi:10.1007/3-540-48975-4_1
32. MacLane, S., Moerdijk, I.: Sheaves in Geometry and Logic: A First Introduction to Topos Theory. Springer Science & Business Media, Heidelberg (2012)
33. Marsden, D., Genovese, F.: Custom hypergraph categories via generalized relations. In: CALCO 2017 (2017, to appear)
34. Marsden, D.: A graph theoretic perspective on CPM(Rel). In: Heunen, C., Selinger, P., Vicary, J. (eds.) Proceedings 12th International Workshop on Quantum Physics and Logic, QPL 2015, Oxford, UK, 15–17 July 2015. EPTCS, vol. 195, pp. 273–284 (2015). http://dx.doi.org/10.4204/EPTCS.195.20
35. Osherson, D.N., Smith, E.E.: Gradedness and conceptual combination. Cognition **12**(3), 299–318 (1982)

36. Piedeleu, R., Kartsaklis, D., Coecke, B., Sadrzadeh, M.: Open system categorical quantum semantics in natural language processing. In: Moss, L.S., Sobocinski, P. (eds.) 6th Conference on Algebra and Coalgebra in Computer Science, CALCO 2015. LIPIcs, vol. 35, pp. 270–289. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik (2015)

37. Rosch, E., Mervis, C.B.: Family resemblances: studies in the internal structure of categories. Cogn. Psychol. **7**(4), 573–605 (1975)

38. Sadrzadeh, M., Clark, S., Coecke, B.: The Frobenius anatomy of word meanings I: subject and object relative pronouns. J. Logic Comput. **23**(6), ext044 (2013)

39. Sadrzadeh, M., Clark, S., Coecke, B.: The Frobenius anatomy of word meanings II: possessive relative pronouns. J. Logic Comput. **26**(2), exu027 (2014)

40. Schütze, H.: Automatic word sense discrimination. Comput. Linguist. **24**(1), 97–123 (1998)

41. Selinger, P.: Dagger compact closed categories and completely positive maps. Electron. Not. Theor. Comput. Sci. **170**, 139–163 (2007)

42. Shepard, R.N., et al.: Toward a universal law of generalization for psychological science. Science **237**(4820), 1317–1323 (1987)

43. Sobocinski, P.: Graphical linear algebra. Mathematical blog. https://graphicallinearalgebra.net/

44. Stubbe, I.: Categorical structures enriched in a quantaloid: categories and semicategories. Ph.D. thesis, Université Catholique de Louvain (2003)

45. Tversky, A.: Features of similarity. Psychol. Rev. **84**(4), 327 (1977)