GENERALIZED RELATIONS FOR COMPOSITIONAL MODELS OF MEANING

FABRIZIO ROMANO GENOVESE

D.Phil Thesis

–

Department of Computer Science
Hertford College
University of Oxford

October 2017

Fabrizio Romano Genovese: *Generalized relations for compositional models of meaning*, D.Phil Thesis, © October 2017

SUPERVISORS:
Bob Coecke
Daniel Marsden

LOCATION:
Oxford

TIME FRAME:
October 2017

PUBLICATIONS

Part of the contents covered in this document (mainly Chapters 3, 5, 6 and 7) have been published by the author in:

[1] J. Bolt, B. Coecke, F. Genovese, M. Lewis, D. Marsden, and R. Piedeleu. "Interacting Conceptual Spaces." In: *Proceedings of the 2016 Workshop on Semantic Spaces at the Intersection of NLP, Physics and Cognitive Science, SLPCS@QPL 2016, Glasgow, Scotland, 11th June 2016.* Ed. by D. Kartsaklis, M. Lewis, and L. Rimell. Vol. 221. EPTCS. Aug. 2016, pp. 11–19. DOI: 10.4204/EPTCS.221.2. arXiv: 1608.01402.

[2] B. Coecke, F. Genovese, S. Gogioso, D. Marsden, and R. Piedeleu. "Uniqueness of composition in quantum theory and linguistics." In: *Electronic Proceedings in Theoretical Computer Science (Forthcoming)*. 2018. URL: http://qpl.science.ru.nl/papers/QPL%7B%5C_%7D2017%7B%5C_%7Dpaper%7B%5C_%7D62.pdf.

[3] B. Coecke, F. Genovese, M. Lewis, and D. Marsden. "Appications of Generalized Relations in Linguistics and Cognition." In: *Logic, Language, Information, and Computation: 24th International Workshop, WoLLIC 2017, London, UK, July 18-21, 2017, Proceedings*. Ed. by J. Kennedy and R. J. de Queiroz. Heidelberg: Springer, 2017, pp. 256–270. ISBN: 978-3-662-55386-2. DOI: 10.1007/978-3-662-55386-2_18. URL: http://link.springer.com/10.1007/978-3-662-55386-2%7B%5C_%7D18.

[4] B. Coecke, F. Genovese, M. Lewis, D. Marsden, and A. Toumi. "Generalized relations in linguistics and cognition." In: *Theoretical Computer Science (Forthcoming)* (2018).

[5] D. Marsden and F. Genovese. "Custom Hypergraph Categories via Generalized Relations." In: *7th Conference on Algebra and Coalgebra in Computer Science (CALCO 2017)*. Ed. by F. Bonchi and B. König; Leibniz International Proceedings in Informatics, 2017, 17:1–17:15. DOI: 10.4230/LIPIcs.CALCO.2017.17. arXiv: 1703.01204. URL: http://arxiv.org/abs/1703.01204.

In particular, the author contributions are as follows:

- In Chapters 1 and 2 the author put together the contents of various papers published in the field to give a coherent accounting of what has been done up to now;
- In Chapter 3 the author main contribution was in helping lying down the mathematical structure defining the category of convex sets and relations. This has been jointly done with Dan Marsden;

- All the proofs in Chapter 4 have been worked out by the author;
- In Chapter 5 the author contributed with conceptual ideas for some of the examples. Most of the examples have been worked out by Martha Lewis;
- In Chapter 6 the author contributed with ideas about how to define the category of generalized relations. All the proofs have been worked out by the author and Dan Marsden. In particular, the author proved all the results regarding varying the underlying algebras and toposes, both in the span and relational case. He also carried on some of the needed checks to prove that our definitions give us hypegraph categories;
- In Chapter 7 the author worked out the examples and proved the result relating our construction with the traditional relation construction in terms of subobjects in presheaf toposes;
- Chapters 8 and 9 are based on unpublished material and have been completely worked out by the author.

During his Dphil, the author has also published the following papers, not strictly related with the material covered here:

[1] J. Bolt, B. Coecke, F. Genovese, M. Lewis, D. Marsden, and R. Piedeleu. "Interacting Conceptual Spaces I: Grammatical Composition of Concepts." In: *arXiv preprint* (2017). arXiv: 1703.08314.

[2] S. Gogioso and F. Genovese. "Infinite-dimensional Categorical Quantum Mechanics." In: *Electronic Proceedings in Theoretical Computer Science* 236 (Jan. 2017). Ed. by R. Duncan and C. Heunen, pp. 51–69. ISSN: 2075-2180. DOI: 10.4204/EPTCS.236.4. arXiv: 1605.04305. URL: http://arxiv.org/abs/1605.04305%20http://dx.doi.org/10.4204/EPTCS.236.4.

[3] S. Gogioso and F. Genovese. "Towards Quantum Field Theory in Categorical Quantum Mechanics." In: *Electronic Proceedings in Theoretical Computer Science (Forthcoming)* (Mar. 2018). arXiv: 1703.09594. URL: http://arxiv.org/abs/1703.09594.

ACKNOWLEDGMENTS

Pectore ab imo.

— Virgil, Aeneid [125, Book VI, Line 55]

I want to thank Italia, Marina, Rino and Andrea, for having showed me the beauty of Maths. My parents, for having always believed in me and for having made a lot of sacrifices, emotional and economic, to allow me to complete this degree; Bob Coecke, for being the most chilled and friendly supervisor one could hope to have, ever. Dan Marsden, for having showed me the way in the crazy and nightmarish territory of category theory during these two long years.

I moreover want to thank all the friends that supported me for the time of my degree studies, all the colleagues in the Quantum Group, with which I have had the most enlightening discussions, and all the people that gave me the possibility of working closely with them (especially Joseph Bolt, Stefano Gogioso, Martha Lewis, Robin Piedeleu and Alexis Toumi, with which I proved many of the results listed in this document). I couldn't have achieved anything without you.

Finally, I tank the Air Force Office of Scientific Research (AFOSR) for having partially funded me with the grant "Algorithmic and Logical Aspects when Composing Meanings". You surely made my life (a bit) easier.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Part I

LITERATURE REVIEW

Here we review the key literature of everything that is relevant for what we are going to do in the second part. We start quite informally, and dive into technical details as we progress.

1 INTRODUCTION

Nil posse creari
de nilo.

— *Lucretius, De Rerum Natura [140, Book I, Lines 155-156]*

The struggle to model and formally represent natural language and meaning is probably as old as mathematical logic is [95, pp. 165-66, 192-95, 221-28, 248-50, and 654-66]. It is especially in the last decades, thanks to the improvements in computer science, that scientists and linguists saw for the first time a real opportunity to implement what was only a philosophical exercise into something practical. This brought new passion to the efforts to ask and solve questions regarding the formalization of language, a passion that stems from many different – and now independent – branches of study.

There are two different types of models that have been studied and exploited so far. The first ones account for a strong formalization of language structures by means of logical rules: We call these models "Compositional". This way of doing things draws inspiration from the work of Frege (to which the principle of compositionality is attributed) in logic [53, 66], and has been carried to natural language by pioneers like Montague [108–110] and Chomsky [34], that are among the first modern researchers who tried to mathematically formalize how language works. The compositional approach has multiple declensions, among which the most influential, at least with regard to what we want to do in this document, is due to Lambek. With the Lambek calculus [90], later restated and simplified in terms of pregroups [91], he gave an elegant account of how the grammar of natural language works. This is not the only way to axiomatize grammar in mathematical terms, but it is particularly appealing for our purposes since the deep connections with category theory that Lambek's approach has. Lambek's approach has then be used to develop techniques that take into account different grammatical constructions in many languages, even Latin [33]. It is out of doubt that this approach gladdens the mathematician, with its elegance and relative simplicity. Moreover, it can be formulated in categorical terms (pregroups are, de facto, categories) [39], and this is consistent with our intuition that what really matters, in language, is how the words compose with each other to form a sentence. This justifies the word "compositional" to define this class of models. Nevertheless, compositional approaches are also rather artificial, since they clash, from the very philosophical premises they start from, with the intrinsic empirical nature of language. People learn how to speak by trial and error, and not by memorizing a given set of axioms and rules; as a result language is irregular, and all the logical rules one can define to model it feel overimposed. Stating things in simpler terms, one can always find an example that does not fit a given compositional model and, as Lambek himself stated, "it is only fair to warn [. . .] that some authorities think that such difficulties are inherent in the present methods." [90, p. 154]

The killer feature of compositional models is that, since they are logic-oriented, defining a logic on the language we are studying is easier, see for instance [123]. Even in this case, though, this is not always straightforward: For ecample, if we can assign meaning to two sentences, say *John reads* and *John writes*, and we have some sort of operation to obtain a conjunction of these sentences, we would like *John reads and writes* to be assigned the same meaning of the logical conjunction between *John reads* and *John writes*. Hence, we would like words such as *and*, *or*, *not* to operate directly within the logical structure of our model. This is difficult though because these words, even if they operate in a logical fashion (at least according to our intuition), exhibit very different behaviors according to the context they are placed in. In our example, *John reads and writes* is the same as *John reads and John writes*, so the word *and* "slides through" the sentence; but *you and me make a good couple* is nowhere near the same as *you make a good couple and I make a good couple*, which makes no sense as a sentence. With *not* things are even harder: If we consider the sentence *he's not tall, he's short* then we have a first bit that is negated and a second bit that is intuitively compatible with this negation: *Short* is somehow the opposite of *tall*; but if we consider *he's not tall, he's a giant* then it is exactly the contrary that happens, and it is difficult to treat a sentence like this logically without avoiding paradoxes. Moreover, in both cases the grammatical role of *not* is the same, so we have no chance to understand how to interpret the word *not* just from the grammatical role it has in a sentence.

With these considerations in mind we highlighted the real problem of compositional models: They are very good at describing the structure of a sentence with regard to the grammatical roles words have in it and the way they compose with each other, but it is difficult to exploit this structure to assign a meaning to them: Meaning often acts in a way that disrespects sentence structure.

On the other end of the spectrum there are empirical models. The philosophy in this case is the opposite, we don't mind about any overimposed grammatical structure, we even disregard the grammar altogether, and we derive meaning of words studying the context in which they are used. These models are often called "distributional" and rely on the Wittgensteinian assumption that "meaning is use" [143]. The reason to study them is often driven by practical applications,

since they have been successful in tasks like thesaurus extraction [48, 72]. The killer feature here is that we can design algorithms that assign meaning to thousands of words in simple ways, essentially running statistical analyses on a huge corpus. The downside is that we do not have any hint about how the compositional structure works, and so it is much more difficult to derive the meaning of a sentence from its component words.
Clearly, defining a logic for distributional models is harder since we totally disregard compositionality, so we cannot straightforwardly plug in any intuition about how meanings should compose in our model.

People struggled a lot in the last decades to define a way to assign meanings to sentences that takes the best from both approaches. We still want to be able to automatically extract meanings of words from a corpus, but then we want the freedom to define rules to compose these meanings in a way that is compatible with the information we extracted.

One successful attempt to do this comes from the DisCoCat model, developed in [39, 44]. The core idea here is that distributional models of meaning often use a vector space formalism, and this formalism shares many categorical features with Lambek's pregroup approach. To be precise, vector spaces and linear morphisms can be formulated in categorical terms as compact closed categories. A pregroup, on the other hand, is a compact closed category too, so there is some structure these settings share. These similarities have been exploited to give distributional models of meaning a compositional structure coming from a pregroup.

One nice feature of this approach is that we can use the diagrammatic calculus developed in categorical quantum mechanics [2, 4, 43] in our setting, since this calculus ultimately relies on the fact that the category we are working in is compact closed. We have a clear way, then, to study and manipulate the means by which words compose in a sentence, and then the "magic" of compact closed categories gives us a way to reduce this composition to a definite meaning, in a purely graphical way.

DisCoCat is part of a much broader area of research that goes under the name of *process theory*. Process theories study the compositional behavior of processes abstractly, usually exhibiting some form of graphical calculus that is sound (sometimes also complete) with respect to some semantic category that represents the kind of processes we are interested in. This approach has been incredibly useful and has been employed in a lot of different research areas, such as linear algebra [138], control theory [12], Markov processes [14], signal flow graphs [25], natural language processing as we just said [44] and electrical circuits [13]. Techniques from this last application will in particular be useful later on in this document.

Lately [42] the compact closed prerequisite has been dropped, showing how similar results can be achieved requiring our categories just to be monoidal closed. We also somehow preserve the diagrammatic features, via the Baez-Stay diagrams developed in [15], that work well for monoidal closed categories[1].

At the end of the day, the mantra of categorical distributional semantics boils down to having two categories $\mathcal{G}$ and $\mathcal{S}$, representing grammar and semantics respectively, and a functor $F : \mathcal{G} \to \mathcal{S}$. Now, these categories have to share some structure, usually compact closedness, and the functor has to preserve this structure. What happens then is that if we have a sentence, we can parse the compositional structure the words in the sentence have, derive their reduction in $\mathcal{G}$ and then apply F to understand how to derive a meaning in $\mathcal{S}$. We immediately understand that we are not bound to vector spaces and pregroups to do this job: As long as we can find a categorical structure that is meaningful in terms of language composition (in our case compact closedness or monoidal closedness) then we can look through all the categories we know having that structure to find a suitable model for our grammar and our semantics. We then prove that a structure preserving functor F exists and we are back in the game. This simple observation motivated the start of a quest to find other categories that could substitute vector spaces and pregroups with the hope to outperform previous results. Alternative semantics proposed include density matrices [16], expressed categorically via the CPM construction [135]; double density matrices [10] (stating it practically this amounts to perform CPM construction on a vector space two times); relational semantics (used as a model of meaning in [39] exploiting categorical similarities with vector spaces already studied in [2, 5].[2]) and convex sets and relations [24], a model based on the work of Gärdenfors [70]. We will extensively talk about the last mentioned approach later. One of the main motivations for a change of semantics relies in the inherent difficulty to define a good logic in the vector spaces framework. We can define a logic for language in the vector spaces category as it has been done in [123], but the logic is not purely internal: It is built up by encoding things and this, from the categorical point of view, sounds a bit arbitrary, albeit many of the encodings are justified from an intuitive point of view. The researcher familiar with category theory wants to see the magic of categories at work, that is, how a logical structure that is compatible with language naturally arises from the category we choose, without any need of arbitrary choices: When a way to compose words is defined, a logic that describes how "difficult" words like *and* or *not* behave must be

1 At the moment there isn't, to my knowledge, a proof of completeness for these diagrammatic calculi, and it is safe to say that Baez-Stay diagrams are still not fully understood.

2 According to some researchers, this can be considered a categorical formalization of the Montague semantics [39, p. 4], but this idea is not universally accepted.

obtained for free from the compositional structure. When this does not happen, categorists conclude that this suitable logic must exists in some other category, and hence that we are maybe using the wrong one.

Back to the compositional distributional mantra "grammar/functor/semantics", it has to be said that things are not always so easy. For instance Preller shows [123, p. 119] that there isn't a functor from a pregroup to vector spaces that preserves compact closed structure and that maps a basic type in the pregroup to a vector space having dimension more than one. Luckily, if we take the freely generated compact closed category on a set of basic types and morphisms (here the generating morphisms represent a dictionary of words we use to generate the category) and we use it instead of the pregroup then our functor exists and everything works just as we expected. The formal construction of the freely generated compact closed category on a set of basic types and morphisms mentioned above can be found in [124].

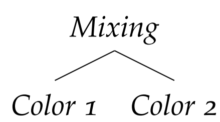
1.1 A COGNITIVE THEORY OF MEANING

We just mentioned Gärdenfors' work above, en passant. Actually, this work differs from everything else we summarised up to now in the fact that the adopted point of view changes again. Gärdenfors doesn't start from linguistic assumptions. Instead, he starts from cognitive considerations and only after tries to explain how language is perceived and understood. What he does is then to define a sort of "semantics of the mind", describing a model of how concepts interact in our heads. After doing this, he takes into account many of the most common grammatical types (pronouns, nouns, verbs, adjectives) and explains how they find a place in his framework. The main assumption is that human beings represent concepts in their minds by means of convex sets, products and domains, intermingling them in structures called "conceptual spaces". Essentially, domains represent fundamental properties a concept can have: For example, we can have a space domain, a time domain, a colour domain and so on. A concept is then a product of convex subsets of these domains. For instance, the concept of *ball* will have a colour connotation, since balls can be of almost any colour. This is formalized as a convex subset of the colour domain, that represents all the colours a ball can have according to the intuition of the person. *Ball* will also have a texture connotation, describing the fact that a ball can be very hard (like in the case of a Snooker ball), quite hard (a basket ball), quite soft (a sponge ball) or almost immaterial (a ball of gas, like the sun). Also these features will correspond to a subset in the texture domain. We can reiterate this reasoning again and again, obtaining a set of domains that describes all the properties our concepts have and that we want to account for. The product of these domains is a conceptual space. A concept will then be defined by a product of the convex subsets in every domain describing it. The convexity requirement is motivated by the assumption that if we can consider something having two different properties, we can consider that concept also to have all the possible mixings between the two: Using our example again, if we can think about a red ball and a yellow ball, Gärdenfors argues that we can think also about an orange ball or a ball having any other colour given by the composition of yellow and red.

The conceptual spaces approach gives elegant explanations about many linguistic phenomena, like the use of metaphors. The idea here is that we translate the meaning of a concept from one domain to another, and we exploit similarities between the two to describe our concept. So *black Wednesday*, referring to 16th September 1992, when the stock market in the UK went upside down, makes sense because we associate the concept *black*, in the domain of colours, with concepts like *darkness*, *sadness* and so on. Even without knowing the meaning of *black wednesday* one can then associate it to the idea of something very bad that happens (or happened or will happen) on a Wednesday. Note that these associations are most of the times non trivial: If *black Wednesday* means something bad, *black Friday* is instead perceived as a really happy thing, at least by people that love shopping.

Anyway, Gärdenfors' model makes many pretty good points and provides good suggestions about how to formalize meaning, even if it has its limitations. For instance, *piano key* means something black or white, but never gray, the point being exactly that piano keys are always painted in a way such that the distinction between the black keys and the white keys is clear. Gärdenfors' model struggles to account for this, since we cannot characterize *piano key* in the color domain as we would do with *ball*: The convexity requisite would force us to consider every shade of gray between white and black, that is the kind of thing we want to avoid with *piano key*.

One possible solution to this problem would be to note that the most pertinent characterization of *piano key* is not in terms of color, but of contrast. What is important about piano keys is not the fact that they can be white or black, but that they are always painted in a contrasting way such that the small keys (usually the black ones) are distinguishable from the big ones (usually white). The correct defining domain for *piano key* would then be the three element semilattice



Where the join of *Color 1* and *Color 2* can be interpreted as ambiguity (e.g. "either *Color 1* or *Color 2*"). The problem with this is that semilattices do not seem to have much in common with convexity. We will give a definition of convexity in Chapter 3 that will allow us to consider semilattices as legitimate domains and, in Chapter 6, we will recast our models to allow also for structures different than convexity. This, in principle, should offer some margin to solve this kind of problems but, even this being the case, it is clear how this solution sounds artificial. Moreover, it is evident how in many situations we cannot be naive in choosing our defining domains and we will have to come up with clever tricks and observations to characterize concepts in a way compatible with Gärdenfors' ideas.

Another limitation of Gärdenfors' model is that conceptual spaces are obtained as products of different domains, but Gärdenfors never states clearly what does he mean by "product". Most of the times he is referring to the cartesian product of convex sets, but this interpretation is not always the one to take to figure out what's happening in the examples he provides. Reading [70] it is clear he has in mind some kind of "parallel composition" that allows to put different domains side by side, but this composition is never clearly defined and the word "product" is used vaguely.

Moreover, convexity is often exploited beyond reason to justify some ideas. All in all, the model proposed by Gärdenfors contains clever ideas, but it is too vague to allow us to extract numbers out of it.

Nevertheless, the problems just highlighted are deemed to be solvable, since they are caused by a weak mathematical formulation of the theory and not by a conceptual inconsistency. For this reason we tried to give a first strong formalization of Gärdenfors' theory in categorical terms, defining the category of convex sets and convex-preserving relations [24], that will be reviewed in detail later in this document.

The most valuable idea in the conceptual spaces approach is that the term "grammar" has to be intended more as a "way to compose concepts" and not only as something related to language: In fact, the subdivision in domains works on a different level with respect to the type system commonly used when formalizing language grammars. For instance, the word *black* can be, according to Lambek, an adjective or a noun, and in these two instances it will have different types, since the type structure of pregroups relies on the type structure of language. For Gärdenfors, instead, these two words act on the very same domain, the domain of colours, and there is no real difference between *black* used as an adjective and *black* used as a noun, since they are both referring to the same concept in essence.

To assign meaning in terms of domains has, as we highlighted before, an intuitive explanation in cognitive terms: The word *meaning* ultimately refers to the way we represent concepts in our head, and moving from intuitive premises in defining a semantics just feels "right", whereas other approaches, like the vector spaces one, look hacky: No one questions their effectiveness, but the author deems as improbable that someone thinks about comparing vectors when asked how he intuitively perceives the idea of "meaning of a sentence". Effectiveness is, in fact, another thing to take into account, and probably a downside of Gärdenfors' approach (albeit we cannot be sure about it, since the formalization of meaning according to this theory is just in its first stage). Our concerns can here be split in two parts:

The first one pertains to the feasibility of defining a logic on our model. As in the vector spaces case, we want to be able to make inferences in our structure: We want to define a notion of entailment (for density matrices this has been done in [16, 18]), a notion of negation, conjunction and so on, such that these definitions are compatible with the language (we want to be able to explain how the word *and* acts by means of the logic we are going to define, for instance). This is even more strongly motivated if we think in cognitive terms; keeping it simple, note that the reason why the logical conjunction is almost always called "and" maybe relies on the intuitive idea that the word *and* makes conjunctions in natural language, and the same argument holds for many of the most used logical connectives. The author of this document spent a considerable amount of time in trying to find a way to implement a logic in our sets and convex relations approach, but the problem is non-trivial. It is likely, moreover, that our formalization of conceptual spaces by means of convex relations is not the right one: Much broader and powerful generalisations of the framework laid down in [24] have already been proposed [41, 104] and will be object of in depth study in this document. This uncertainty about the future of convex sets and convexity preserving relations as a semantics served as a motivation to look for a more high-level explanation of the logical connectives in terms of language, an explanation that does not explicitly rely on the category we are using as semantics. Again, the problem is non-trivial and needs to be further investigated.

The second concern has not been investigated yet, and it looks more difficult to account for. This is about the possibility of having an algorithm that is able to build conceptual spaces from a corpus of words. This possibility is heartily desired, because it is the only way to mark a transition between toy models and real models of meaning. If we do not have a way to assign meanings to concepts automatically, then it will be impossible to use the Gärdenfors' model for any practical application. We remember again that this is exactly what motivated the study of distributional models of meaning: Practical feasibility should always be considered when one tries to model cognition and language in mathematical terms.

With this brief introduction we hope to have given a satisfactory account of what has been done in the last years in the field. The synopsis of this document will then be as follows:

- In Chapter 2 we will provide a technical explanation of how the pregroup grammars work and what are the categorical properties we are interested in;

- In Chapter 3 we will introduce our first attempt to model a categorical semantics based on Conceptual Spaces, with detailed examples provided in Chapter 5. The mathematical properties of this semantics will be analyzed in depth in Chapter 4. We will show how our new model allows us to overcome some of the limitations of Gärdenfors' original one, giving us the possibility of extracting actual numbers out of our formalizations of words, sentences and the like;

- In Chapter 6 we will introduce a much broader generalization of our framework, that may well be considered the "last stop" of relational semantics for its generality. These generalized relations will contain the structures defined in Chapter 3 as a special case, and will allow us to perform additional operations that were impossible before. From an operative perspective, the most important result is probably the fact that we will now be able to talk about distances between concepts, as it will be extensively shown with examples in Chapter 7;

- In Chapters 8 and 9 we will adopt a radically different point of view, deeply criticizing our own work up to this point (every now and then a bit of self-deprecation does not harm) and using suggestions coming from exotic languages we will advocate for a radically different approach to the whole problem of language.

- In the conclusion we will stop boring the reader, trying to sum up what have we learned up to now.

2 LAMBEK CALCULUS AND PREGROUPS

Quare non ut intellegere possit
sed ne omnino possit non intellegere curandum.

— *Quintilianus, De Institutione Oratoria [100, Book VIII, 2.24]*

Pregroups have a much more complicated history than it looks. They have been, in fact, introduced by Lambek in 1999 [91], as a simplification of his work on what we call now Lambek calculus [90], formulated nearly forty years earlier. Both serve the same purpose, that is, defining a mathematical framework to perform formal analysis on language.

2.1 RESIDUATED MONOIDS

The original Lambek calculus was expressed in the form of a sequent calculus. What triggers the categorical reasoning is that this calculus (as shown in [29]) is sound and complete with respect to residuated monoids, since residuated monoids can be easily formalised in categorical terms. For this reason, we will give the definition of Lambek calculus directly in terms of residuated monoids as done in [42], avoiding sequents and reduction rules.

Definition 2.1.1. A partially ordered set $(L, \leqslant)$ is a set with a relation on it that is reflexive, antisymmetrical and transitive. A monoid $(L, \cdot, 1)$ is a set with a 2-ary operation on it, denoted with $\cdot$, and a 0-ary operation 1, called *unit*, such that $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ and $a \cdot 1 = a = 1 \cdot a$ for all $a, b, c \in L$.

Definition 2.1.2. A *partially ordered monoid*, denoted with $(L, \leqslant, \cdot, 1)$, is a set L with a partial order structure $(L, \leqslant)$ and a monoid structure $(L, \cdot, 1)$ on it, such that these structures are compatible with each other (multiplication preserves the order):

$$a \leqslant b \Rightarrow a \cdot c \leqslant b \cdot c, \qquad a \leqslant b \Rightarrow c \cdot a \leqslant c \cdot b \quad \forall a, b, c, \in L.$$

Definition 2.1.3. A *residuated monoid*, denoted with $(L, \leqslant, \cdot, 1, \multimap, \circ\!-)$ is a set L with a partially ordered monoid structure $(L, \leqslant, \cdot, 1)$ and two 2-ary operations $\multimap, \circ\!-$, called *right and left adjunctions of* $\cdot$, such that

$$a \leqslant b \multimap c \Leftrightarrow b \cdot a \leqslant c \Leftrightarrow b \leqslant c \circ\!- a \quad \forall a, b, c, \in L.$$

This structure encodes the compositionality of language as follows: Elements of the residuated monoid stand for grammatical types, and multiplication is concatenation of types, meaning that if a, b are types, then $a \cdot b$ is ab. The unit stands for the empty type and the order relation stands for reduction. The adjunctions represent function types (for instance, adjectives are usually modelled as elements $n \multimap n$, that is, things that carry the noun type to the noun type). This is further motivated by the observation, very easy to prove from the axioms, that $a \cdot (a \multimap b) \leqslant b$ (and similarly $(b \multimapinv a) \cdot a \leqslant b$); this intuitively means that if we plug something having type a next to something having type $a \multimap b$, this reduces to something having type b. This is exactly what happens in natural language: If we compose a noun to an adjective, we get a noun.

Definition 2.1.4. Given a set of words W and a set of basic grammar types T, a *Lambek type-dictionary* is a subset $D \subseteq W \times F(T)$, where $F(T)$ is the free residuated monoid generated by T [1].

The idea here is that to any word we associate one or more types (that is, algebraic terms of the residuated monoid). For example, if n stands for the noun type and $n \multimap n$ for the adjective one, then one could have $(\textit{black}, n)$ and $(\textit{black}, n \multimap n)$ as elements of D, representing the fact that the word *black* can be used as a noun or as an adjective in our grammar.

From this, we can finally define:

Definition 2.1.5. A *Lambek Grammar* is a pair $\langle D, S \rangle$, with D Lambek type-dictionary and S, subset of T, called *the set of basic grammatical types*.

To summarise, we choose some set of basic types T. In this T we include a subset that denotes the basic sentence types, like "declarative", "interrogative", "exclamative" and so on, call it S. We generate the residuated monoid on T, then we associate to every word one or more types that live into the monoid. Finally, we adopt the notation $\langle D, S \rangle$ to say "look, the ones on the left are our words, to every word corresponds a type. The sentence types are the ones on the right". This is important because

Definition 2.1.6. A string of words $w_1 \dots w_n$ is said to be *grammatical* if there is a sequence of types $t_1, \dots, t_n$ such that $(w_i, t_i) \in D$ for all $1 \leqslant i \leqslant n$, and $t_1 \cdot \dots \cdot t_n \leqslant s$ for some s in S. In this case we also say that the string of types $t_1 \cdot \dots \cdot t_n$ *reduces to* s. [2]

Now it should be clear what are we doing. We are interpreting order as reduction, and the idea then is: Take a string of words, consider all the words separately. Look in the dictionary for their types and compose them via multiplication. Then apply the residuated monoid rules to get something that is in S. If this is possible, then the sentence is grammatically correct. If this is not possible, then the sentence is grammatically incorrect.

1 the free residuated monoid construction can be found in [90].
2 By abuse of notation we can also say that a string of words $w_1 \dots w_n$ *reduces to* s.

2.2 PREGROUP GRAMMARS

Now we consider something a bit simpler. As we already mentioned, pregroups are a simplification of the original Lambek calculus. We would expect, then, to get a structure that is somehow easier to manage than residuated monoids. This is indeed true:

Definition 2.2.1. A *pregroup*, denoted with $(G, \leqslant, \cdot, 1, (-)^l, (-)^r)$ is a set L with a partially ordered monoid structure $(L, \leqslant, \cdot, 1)$ on it, and two 1-ary operations $(-)^l, (-)^r$ called *left and right adjunctions* respectively, such that

$$p \cdot p^r \leqslant 1 \leqslant p^r \cdot p \qquad p^l \cdot p \leqslant 1 \leqslant p \cdot p^l \qquad \forall p \in G.$$

The only true difference between a pregroup and a residuated monoid is that in the latter the adjunctions are 2-ary, so they take two arguments, while in the former they are 1-ary, so they only take one. This has quite big implications on the categorical properties these structures have. Note that every pregroup can be regarded as a residuated monoid defining $a \multimapinv b := a \cdot b^l$ and $a \multimap b := a^r \cdot b$. We can also do the opposite, and regard every residuated monoid as a pregroup putting $a^l = 1 \multimapinv a$ and $a^r = a \multimap 1$. The two definitions are not one the inverse of the other: In general if we take a residuated monoid, interpret it as a pregroup and then interpret the pregroup again as a residuated monoid what we get will be different from the residuated monoid we started with. [3]

Fiddling with the pregroup axioms one can prove that $1^r = 1^l = 1$ and that $(a^l)^r = (a^l)^r = a$ for all $a \in G$. Moreover, $(a \cdot b)^l = b^l \cdot a^l$ and $(a \cdot b)^r = b^r \cdot a^r$ for all $a, b \in G$. Be aware that adjoints are not nilpotent, they iterate: $((-)^r)^r \neq (-)^r$ and $((-)^l)^l \neq (-)^l$. What we said for residuated monoids also applies to pregroups:

Definition 2.2.2. Given a set of words W and a set of basic grammatical types T, a *Lambek (pregroup) type-dictionary* is a subset $D \subseteq W \times G(T)$, where $G(T)$ is the free pregroup [4] generated by T .

Definition 2.2.3. A *Lambek (pregroup) Grammar* is a pair $\langle D, S\rangle$, with D Lambek (pregroup) type-dictionary and S subset of T, the set of basic grammatical types.

Definition 2.2.4. A string of words $w_1 \dots w_n$ is said to be *grammatical* if there is a sequence of types $t_1, \dots, t_n$ such that $(w_i, t_i) \in D$ for all $1 \leqslant i \leqslant n$, and $t_1 \cdot \dots \cdot t_n \leqslant s$ for some s in S. In this case we also say that the string of types $t_1 \cdot \dots \cdot t_n$ *reduces to* s.

3 If you want to draw a universal algebraic comparison, note that every semigroup can be interpreted as a set, and every set can be interpreted as a semigroup taking union as operation and the empty set as unit. Nevertheless, interpreting a semigroup as a set and then interpreting the very same set as a semigroup gives us back an algebraic structure that is not isomorphic to the semigroup we started with: We clearly forgot the operations in the first passage and cannot recover them. For instance, we can never recover noncommutative semigroup structures.

4 The free pregroup construction can be found in [91].

2.2.1 *Examples*

Now let's show some examples. We will use the same examples in the pregroup case and in the residuated monoid case. Consider the sentence *John loves Mary*. The words *John* and *Mary* are nouns: If we agree to define with n the noun type in both our formalisms, then these words will both have type n. The word *loves* is a transitive verb, that is, something that takes an argument on the left, an argument on the right and gives back a well formed sentence. This is represented as $((n \multimap s) \multimapinv n)$ in residuated monoids and as $n^r \cdot s \cdot n^l$ in the pregroup setting. Then we have:

Pregroups	Words	Residuated Monoids
$n \cdot (n^r \cdot s \cdot n^l) \cdot n \leqslant$	*John loves Mary*	$n \cdot ((n \multimap s) \multimapinv n) \cdot n \leqslant$
$n \cdot (n^r \cdot s) \leqslant$	*John (loves Mary)*	$n \cdot (n \multimap s) \leqslant$
s	*(John loves Mary)*	s

Table 1: Reduction of *John loves Mary*.

In both cases the concatenation of words reduces to a sentence type, so the sentence is grammatical. Examples can be much more complicated: Take for instance the sentence *I was going quietly*, also used in [30]: Here we assign the following types:

Words	Types (pregroups)	Types (residuated monoids)
I	π_1	π_1
was	$\pi_1^r \cdot s_2 \cdot p_1^l$	$(\pi_1 \multimap s_2) \multimapinv p_1$
	$\pi_3^r \cdot s_2 \cdot p_1^l$	$(\pi_3 \multimap s_2) \multimapinv p_1$
	$\pi_1^r \cdot s_2 \cdot p_2^l$	$(\pi_1 \multimap s_2) \multimapinv p_2$
	$\pi_3^r \cdot s_2 \cdot p_2^l$	$(\pi_3 \multimap s_2) \multimapinv p_2$
going	$(p_2 \cdot i^l) \cdot i$	$(\pi_2 \multimapinv i) \cdot i$
quietly	$i^l \cdot i$	$(1 \multimapinv i) \cdot i$

Table 2: *I was going quietly*, type dictionary.

Type π_i stands for a pronoun: π_1 is the type of *I*, π_2 is the type of *you*, *we*, *they*, π_3 the one for *he*, *she*, *it*. The conjugation of verbs is obtained by modification of their infinitive form; we denote infinitive type as i (then *to go* has type i) and the participle type as p_i, where p_1 is present participle and p_2 the past participle. Then a conjugation to the participle form for a verb is obtained juxtaposing on the left the type $(p_i \cdot i^l)$ to the infinitive type. *Going* is then $(p_1 \cdot i^l) \cdot i$, and this is a consistent position, since $(p_1 \cdot i^l) \cdot i \leqslant p_1$, so *going* effectively reduces to a present participle. s_i is the type for a declarative sentence, where s_1 is a declarative sentence in the present and s_2 in the past. Conjugated forms of the verb *to be* used in auxiliary position are obtained as $\pi_i^r \cdot s_j \cdot p_k^l$, so *was* is $\pi_i^r \cdot s_2 \cdot p_k^l$: We have s_2 because the word gives us a sentence in the past, and π_i^r is taken with $i = 1$ or 3, because for second person pronouns the correct form is *were*. p_k can be p_1 or p_2 because we can juxtapose both a present or a past participle to the right. Finally, adverbs have form $i^l \cdot i$. For the residuated monoid the same considerations hold, but they have to be expressed in the residuated monoids language, so using $\multimap, \circ\!-$.

Some observations: One, these type choices look a bit arbitrary, but they are quite well motivated, see [92] for an extensive explanation. They are "arbitrary" in the sense that they don't come from mathematical analysis but from a diligent study of natural language (and this could not be different, type assignments always come from empirical observation). Two, they strictly depend on the language we are modeling, and are good only for English grammar. In other languages words will compose in different ways (especially the verbal systems vary a lot, think about how many languages express the construction of transitive verbs in the form "object - verb - subject"). Three, here you can clearly see how the same word can have many different type entries in the Lambek type-dictionary. Table 2 only accounts for a minimal part of it, since, for instance, the word *was* has many other grammatical entries in an extensive dictionary, think about *I was sad*: In this case *was* will probably have type $\pi_1^r \cdot s_2 \cdot n^l$. If we try to compose the words now:

Pregroups	Words
$(\pi_1) \cdot (\pi_1^r \cdot s_2 \cdot p_1^l) \cdot ((p_1 \cdot i^l) \cdot i) \cdot (i^l \cdot i) \leqslant$	*I was going quietly*
$(\pi_1) \cdot (\pi_1^r \cdot s_2 \cdot p_1^l) \cdot ((p_1 \cdot i^l) \cdot i) \cdot 1 =$	*I was going (quietly)*
$(\pi_1) \cdot (\pi_1^r \cdot s_2 \cdot p_1^l) \cdot ((p_1 \cdot i^l) \cdot i) \leqslant$	*I was going (quietly)*
$(\pi_1) \cdot (\pi_1^r \cdot s_2 \cdot p_1^l) \cdot (p_1) \leqslant$	*I was (going quietly)*
$(\pi_1) \cdot (\pi_1^r \cdot s_2) \leqslant$	*I (was going quietly)*
s_2	*(I was going quietly)*

Words	Residuated Monoids
I was going quietly	$\pi_1 \cdot ((\pi_1 \multimap s_2) \circ\!- p_1) \cdot ((p_1 \circ\!- i) \cdot i) \cdot ((1 \circ\!- i) \cdot i) \leqslant$
I was going (quietly)	$\pi_1 \cdot ((\pi_1 \multimap s_2) \circ\!- p_1) \cdot ((p_1 \circ\!- i) \cdot i) \cdot 1 =$
I was going (quietly)	$\pi_1 \cdot ((\pi_1 \multimap s_2) \circ\!- p_1) \cdot ((p_1 \circ\!- i) \cdot i) \leqslant$
I was (going quietly)	$\pi_1 \cdot ((\pi_1 \multimap s_2) \circ\!- p_1) \cdot (p_1) \leqslant$
I (was going quietly)	$\pi_1 \cdot (\pi_1 \multimap s_2) \leqslant$
(I was going quietly)	s_2

Table 3: Reduction of *I was going quietly*.

Clearly we chose one among the many type instances of *was*, precisely the only one that makes the string reduce to a sentence type: How do we choose the types? Well, we choose them all. We try all the possible combinations of types and try to reduce them; clearly some of them will reduce to a sentence type and some will not, we just discard the ones that don't because this means we considered the wrong grammar types for our words in those cases. If we have more than one choice of types that reduce to a sentence then we conclude that the original sentence can be read in different ways. Remembering what we said in the introduction about compositional distributional models, this means that different meanings could be assigned to our sentence. This does not contrast with our intuition, since there are multiple examples in natural language of sentences that can be interpreted in different ways.

Another question that may rise is: Do all reductions terminate? Clearly if a sequence of types reduces to a sentence type we will get there sooner or later, but what happens with an incorrectly typed sequence? If we have, say, *going I was quietly*, will we be able to say that this sentence is not grammatically correct or will our reduction rule run indefinitely without stopping? In its original paper [90] Lambek proved that his calculus based on residuated monoids is decidable. When he introduced pregroups [91] he also proved that the reduction procedure is decidable in the free pregroup (that is the one we use to build Lambek type-dictionaries and Lambek grammars), so you can reassure yourself that compositional approaches to natural language do not cause computational issues. We refer the reader interested in complexity and decidability problems related to pregroup grammars to [38, 114].

This was just a brief explanation of how pregroups and residuated monoids work. Keep in mind that in the original works by Lambek (and in many other that focus on the formal analysis of grammars using these structures) many more grammatical objects are taken into account. I don't dare to say that these models describe natural language grammar structure completely, but, at the moment, it is safe to affirm that they cover a huge part of the constructions that are legitimately meaningful in natural language (examples of how this can go nightmarishly wrong are given in Chapters 8 and 9).

2.3 MONOIDAL AND COMPACT CLOSED CATEGORIES

We briefly mentioned in the previous sections that pregroups and residuated monoids are interesting for categorical models of meaning because they can be formalized in categorical terms, and these categories have nice properties. Now that we made clear how pregroups and residuated monoids work we are ready to see their categorical formalization in detail. We will start with pregroups because it is simpler.

Definition 2.3.1. A *monoidal category* is a tuple $(C, \otimes, I, \alpha, \lambda, \rho)$, where

- C is a category (we denote as usual objects with $A, B, \ldots$ and arrows with $A \to B, X \to Y, \ldots$);
- $\otimes$, the *tensor*, is a functor $C \times C \to C$ where we write $A \otimes B$ for $\otimes(A, B)$;
- I, the *unit*, is an object of C.

We also require that the following natural isomorphisms exist:

- $\alpha_{A,B,C} : ((A \otimes B) \otimes C) \to (A \otimes (B \otimes C))$;
- $\rho_A : A \otimes I \to A$;
- $\lambda_A : I \otimes A \to A$.

These natural isomorphisms, moreover, must be such that any formal and well-typed diagram made up from $\otimes$, α, λ, ρ, α^{-1}, ρ^{-1}, λ^{-1}, and I commutes, where " formal" here means "not dependent on the structure of any particular monoidal category".

A monoidal category is moreover defined to be *symmetric* if there is also a natural isomorphism

$$\sigma_{A,B} : A \otimes B \to B \otimes A$$

such that any formal and well-typed diagram made up from $\otimes$, α, λ, ρ, σ, α^{-1}, ρ^{-1}, λ^{-1}, σ^{-1} and I commutes, where again " formal" means "not dependent on the structure. For a more satisfactory statement and discussion of the above definition, we redirect the reader to [97].

Monoidal have been introduced and studied decades ago, see for example [88]. For us, they are particularly useful (among other things) because they admit the following interpretation: The objects can be thought of as "system types"; a morphism $f : A \to B$ is then a process taking inputs of type A and giving outputs of type B. The object $A \otimes B$ represents the systems A and B in *parallel*; hence, a morphism $f \otimes g : A \otimes B \to C \otimes D$ is to be thought of as running the process $f : A \to C$ *whilst* running the process $g : B \to D$. The object I is thought of as the trivial system.

Example 2.3.2. The category **Rel** of sets and relations is monoidal. $\otimes$ is the Cartesian product and I is any singleton set $\{*\}$.

Example 2.3.3. The category $\text{FdVect}_{\mathbb{R}}$ of finite dimensional real vector spaces and linear maps is monoidal. The functor $\otimes$ is the tensor product, the trivial system I is the familiar one-dimensional real vector space $\mathbb{R}$.

Definition 2.3.4 (Dagger)**.** Given a category $\mathcal{C}$, a *dagger* on $\mathcal{C}$ is an *involutive functor* $\dagger : \mathcal{C}^{op} \to \mathcal{C}$, meaning that it is identity on objects and assigns to every morphism $f : A \to B$ a morphism $f^\dagger : B \to A$ such that

$$id_A = id_A^\dagger \qquad (g \circ f)^\dagger = f^\dagger \circ g^\dagger \qquad f^{\dagger\dagger} = f$$

For each couple of morphisms $f : A \to B, g : B \to C$ and any A, B, C.

Definition 2.3.5. A *compact closed category*[5] is a monoidal category such that, for every object A, there are two objects A^l, A^r and the following morphisms:

$$\epsilon^l : A^l \otimes A \to I \qquad \epsilon^r : A \otimes A^r \to I$$
$$\eta^r : I \to A^r \otimes A \qquad \eta^l : I \to A \otimes A^l$$

Such that the following equalities hold:

$$(id_A \otimes \epsilon^l) \circ (\eta^l \otimes id_A) = id_A \qquad (\epsilon^r \otimes id_A) \circ (id_A \otimes \eta^r) = id_A$$
$$(\epsilon^l \otimes id_{A^l}) \circ (id_{A^l} \otimes \eta^l) = id_{A^l} \qquad (id_{A^r} \otimes \epsilon^r) \circ (\eta^r \otimes id_{A^r}) = id_{A^r}$$

These equalities are usually called *snake equations* or *yanking equations*. We say that a category is *dagger compact closed* if it admits a dagger functor $(-)^\dagger$ and $\eta^l = \epsilon^{l\dagger}, \eta^r = \epsilon^{r\dagger}$.

The name "yanking equations" originated in the field of categorical quantum mechanics [3]. In [136] it was shown that there is a diagrammatic calculus that is sound and complete for compact closed categories. Briefly, this is how it works: Objects are labeled wires, and morphisms are given as nodes with input and output wires. Composing morphisms consists of connecting input and output wires, and the tensor product is formed by juxtaposition, as shown in Figure 1. By

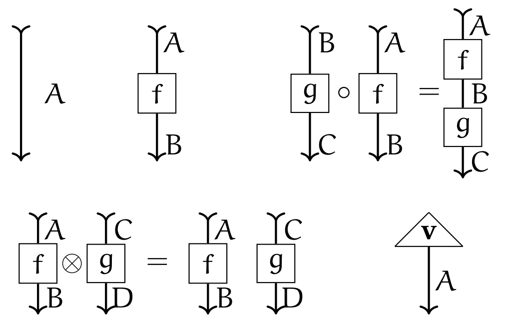


Figure 1: Monoidal graphical calculus.

convention, the wire for the monoidal unit is omitted. These diagrams hold in every monoidal category. For compact closed categories we

5 Note that this definition is more general than the one usually employed: Nearly always in our applications our category is symmetric and we have $A^l = A^* = A^r$. Hence we can just define $\epsilon : A \otimes A^* \to I, \eta^r : I \to A^* \otimes A$.

also have graphical rappresentations of the morphisms $\eta^l, \eta^r, \epsilon^l, \epsilon^r$, commonly called "caps" and "cups", see Figure 2. In this diagrammatic formalism the equalities given in the previous definition take the form shown in Figure 3, hence the name "snake equations". Note that, in many models, we will have $\eta^l = \eta^r$ and $\epsilon^l = \epsilon^r$. In this case the wires will be depicted as undirected, meaning that we will draw them simply as lines and not as arrows. Compact closed categories

Figure 2: Graphical compact structure.

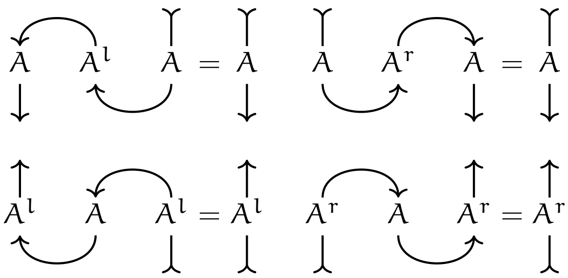


Figure 3: Snake equations.

lie at a convenient level of abstraction to work with. Many of the categories one would consider to use as either grammar or meaning environments have a compact closed structure.

Example 2.3.6. All objects in **Rel** are self-dual. The caps are given by

$$\epsilon^l_X = \epsilon^r_X := \epsilon_X : X \times X \to \{*\}$$
$$\epsilon_X(x, x', *) \Leftrightarrow x = x'$$

The associated cup η_X is just the converse of the above. Simple check shows that the snake equations hold.

Example 2.3.7. More generally, we can build a category of relations over any regular category, see for instance [26]. Moreover, the category of relations over a regular category is always compact closed, see for instance [77]. Since **Set** is regular, we immediately deduce the compact closedness of **Rel**, showed explicitly in the previous example, as a particular instance of this much broader result.

Example 2.3.8. **fdHilb** is the category of finite dimensional real inner product spaces. As in the case of $\text{FdVect}_{\mathbb{R}}$, $\otimes$ is the tensor product of vector spaces and I is the one-dimensional space $\mathbb{R}$. In defining cups and caps, we make use of the fact that if $\{v_i\}_i$ and $\{u_j\}_j$ are bases for vector spaces V and U respectively, then $\{v_i \otimes u_j\}_{i,j}$ is a basis for $V \otimes U$. Moreover, any linear map is fully determined by its action on a basis. Every finite-dimensional vector space is self-dual, and the cups and caps are given by

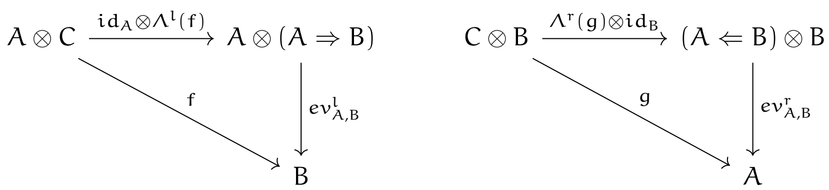
$$\epsilon^l_V = \epsilon^r_V := \epsilon_V : V \otimes V \to \mathbb{R}$$
$$\sum_{i,j} c_{i,j}\,(v_i \otimes v_j) \mapsto \sum_{i,j} c_{i,j}\langle v_i|v_j\rangle$$

$$\eta^r_V = \eta^r_V := \eta_V : \mathbb{R} \to V \otimes V$$
$$1 \mapsto \sum_i (v_i \otimes v_i)$$

Verifying that these maps satisfy the snake equations is again a simple check.

Compact closed categories can be used to define pregroups in categorical terms. We are not so lucky with residuated monoids and we need another definition, namely the one of "monoidal bi-closed category". Stating things in the most abhorrent and brutal way possible, a monoidal bi-closed category looks like a cartesian closed category where instead of the cartesian product you have a tensor. Let's make this precise:

Definition 2.3.9. A monoidal bi-closed category[6] is a monoidal category (the notation for objects, arrows, products and tensor unit will be the same as in the pregroup case) such that for every pair of objects A, B, there are two objects $A \Rightarrow B, A \Leftarrow B$, and two morphisms $ev^l_{A,B} : A \otimes (A \Rightarrow B) \to B$ and $ev^r_{A,B} : (A \Leftarrow B) \otimes B \to A$, called *left and right evaluations*, respectively. We also require that for every couple of morphisms $f : (A \otimes C) \to B$, $g : (C \otimes B) \to A$ there are two unique morphisms, denoted with $\Lambda^l(f) : C \to A \Rightarrow B, \Lambda^r(g) : C \to A \Leftarrow B$, that make these diagrams commute:

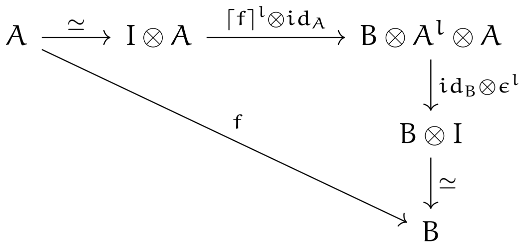


The word "bi-closed" comes from the fact that you have $A \Rightarrow B$ and $A \Leftarrow B$ for every object. If you only have one of the two then the category is called *right closed* or *left closed*, respectively.

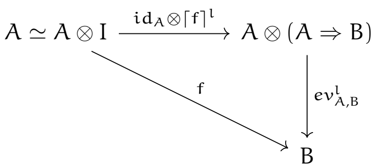
Note that in compact closed categories we can "bend the wires", that is, from $f : A \to B$ we can obtain $\ulcorner f \urcorner^l : I \to B \otimes A^l$ setting $\ulcorner f \urcorner^l := (f \otimes id_{A^l}) \circ \eta^l$ (here we are "bending on the right", we can "bend on

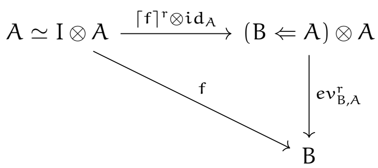
6 If you like a more compact definition, a monoidal category is bi-closed when all the functors $(-) \otimes A$, $A \otimes (-)$ have a right adjoint.

the left", that is $\ulcorner f \urcorner^r : I \to A^r \circ B$ setting $\ulcorner f \urcorner^r := (id_{A^r} \otimes f) \circ \eta^r)$. This makes sense because if we bend the wires again we come back to f via the yanking equations:



We can do something quite similar with monoidal bi-closed categories, via the evaluation and Λ arrows: From $f : A \to B$ we set $\ulcorner f \urcorner^l := \Lambda^l(f \circ \simeq) : I \to A \Rightarrow B$ (here $\simeq$ is the obvious isomorphism $A \otimes I \to A$, and we can obviously also do $\ulcorner f \urcorner^r := \Lambda^r(f \circ \simeq) : I \to B \Leftarrow A$.) Composing with the evaluation function we have:

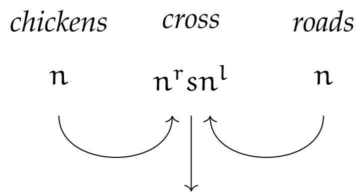




The morphisms $\ulcorner f \urcorner^l, \ulcorner f \urcorner^r$ are called, in both categories, *left and right name of* f, respectively.

2.4 THE CATEGORICAL POINT OF VIEW

Now, a pregroup can be presented as a compact closed category as follows: We interpret elements of a pregroup G as objects. It is common practice to write them in lower case when referring to a pregroup in the classical sense and to write them in upper case (following the category-theoretic convention) when the pregroup is presented as a category. So an element n of the pregroup will be interpreted as an element N of the category. Since partial orders are categories (precisely, categories in which there is at most one arrow between two objects), we interpret $n \leqslant m$ as $N \to M$. Reflexivity and transitivity of $\leqslant$ assures us that there is an identity arrow for every object and that composition behaves well. The tensor structure is given by the monoid fragment of the pregroup, so $m \cdot n$ gets interpreted into $M \otimes N$, and 1 is the tensor unit I. The morphisms $\eta^r, \eta^l, \epsilon^r, \epsilon^r$ are the interpretation of $1 \leqslant a^r \cdot a$, $1 \leqslant a \cdot a^l$, $a \cdot a^r \leqslant 1$, $a^l \cdot a \leqslant 1$, respectively. It is easy to check that yanking equations hold, take the first one for instance: $(\epsilon^r \otimes id_A) \circ (id_A \otimes \eta^r)$ is the interpretation of $a \cdot 1 \leqslant a \cdot a^r \cdot a \leqslant 1 \cdot a$, and since $1 \cdot a = a = a \cdot 1$, it follows $a \leqslant a$, that is interpreted as id_A. So we have $(\epsilon^r \otimes id_A) \circ (id_A \otimes \eta^r) = id_A$ as expected. The other yanking equations are verified in a similar way. Taking advantage of the graphical calculus defined above, now reduction has a straightforward graphical representation, for instance:

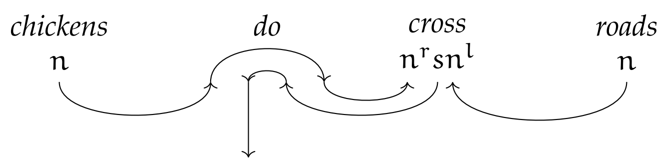


Residuated monoids are not compact closed categories, but they can be represented as monoidal bi-closed categories. We do essentially the same thing we did with pregroups, sending the partial order fragment into the morphism structure and the monoid fragment into the tensor one. Then we interpret $A \multimap B$ into $A \Rightarrow B$ and $a \multimapinv b$ into $A \Leftarrow B$. The residuated monoid axioms ensure that the equations involving evaluations still hold.

The names $\ulcorner f \urcorner^l$, $\ulcorner f \urcorner^r$ are incredibly important, since when we want to assign a meaning to our words, we apply the functor (the one from our grammar – pregroup or residuate monoid in our case – to the semantics) to words names. For instance, if we represent an adjective as $N \xrightarrow{f} N$, then we apply the functor to $\ulcorner f \urcorner^l$ or $\ulcorner f \urcorner^r$. What we get in the semantics is then a morphism of type $I \to T$, where T is the image of the codomain of $\ulcorner f \urcorner^l$ or $\ulcorner f \urcorner^r$ through our functor. So we end up with something of type $I \to T$, that is, something that "picks" an element of T: For instance, if your semantics is sets and relations over sets, then I is a singleton, so $I \to T$ is a relation from $\{*\}$ to T, hence a subset of T. If your semantics is vector spaces, then $I = K$, where K is the field the vector space is on, and then $I \to T$ is a linear map from K to the vector space T, that is, a 1-dimensional subspace of T. This ends up being a single vector if we work only with normalised elements (as we usually do), and so disregard all the scalars multiples.

2.4.1 *Beyond standard categorial grammar*

The compact closed structure is useful also to represent *functional words*. An example of functional word is the verb *do* used as an auxiliary. This particular case can indeed be accounted for using cup, caps and our graphical formalism [44]:



Unfortunately, the compact closed structure alone is not enough to represent other functional words, such as *relative pronouns*. To account for these words we need to introduce more structure, that will graphically be represented as *multi-wires*. We have the following sequence of definitions:[7]

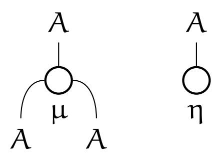
Definition 2.4.1 (Monoid)**.** Given a symmetric monoidal category $\mathcal{C}$ and an object A, a *monoid* on A is given by a couple of morphisms

$$(\eta : I \to A, \mu : A \otimes A \to A)$$

Such that the *associativity* and *unit* conditions hold:

$$\mu \circ (id_A \otimes \mu) = \mu \circ (\mu \otimes id_A) \qquad \mu \circ (id_A \otimes \eta) = id_A = \mu \circ (\eta \otimes id_A)$$

These structures are pictorially depicted as:



We moreover say that a monoid on A is *commutative* if

$$\mu \circ \sigma_{A,A} = \mu$$

where σ is the natural isomorphism representing the symmetry on $\mathcal{C}$.

The definition of monoid can promptly be dualized, as follows:

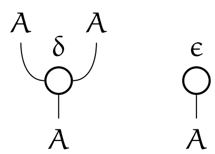
Definition 2.4.2 (Comonoid)**.** Given a symmetric monoidal category $\mathcal{C}$ and an object A, a *comonoid* on A is given by a couple of morphisms

$$(\epsilon : A \to I, \delta : A \to A \otimes A)$$

Such that the *associativity* and *unit* conditions hold:

$$(id_A \otimes \delta) \circ \delta = (\delta \otimes id_A) \circ \delta \qquad (id_A \otimes \epsilon) \circ \delta = id_A = (\epsilon \otimes id_A) \circ \delta$$

These structures are pictorially depicted as:



We moreover say that a comonoid on A is *commutative* if

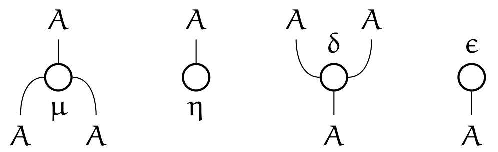
$$\sigma_{A,A} \circ \delta = \mu$$

where σ is the natural isomorphism representing the symmetry on $\mathcal{C}$.

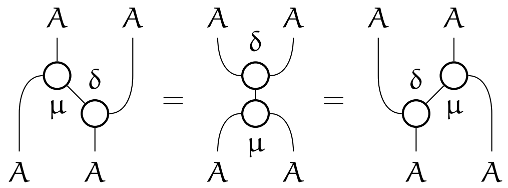
7 Since multi-wires will always be used in the undirected setting, we are only giving the relevant undirected definitions. This in particular means that the direction of the arrows will not be drawn in the figures that follow.

Finally, we give the definition that interests us the most, and that allows us to talk about multi-wires:

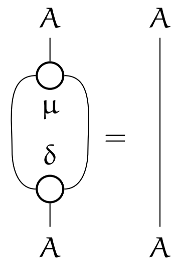
Definition 2.4.3 (Commutative Frobenius Algebras)**.** Given a symmetric monoidal category $\mathcal{C}$ and an object A, a *commutative Frobenius algebra* on A is given by a commutative monoid and comonoid structures:



Such that they satisfy the *Frobenius axiom*, graphically depicted as:



We say that the Frobenius algebra is *special* if the *special axiom* holds:



To conclude, we say that a Frobenius algebra is a *dagger Frobenius algebra* if $\mathcal{C}$ admits a dagger functor and the monoid multiplication and unit can be obtained applying the dagger functor to the comonoid comultiplication and counit, respectively.

Multi-wires, also called *spiders* [43] can be obtained "stacking up" (co)monoids (co)multiplications [43, 86], and can be thought of as representing the act of being connected. Graphically, spiders are just wires that can have more (or less) than two ends.

Multi-wires connect things just as wires do. As a consequence of the Frobenius axiom. the only thing that matters with multi-wires is if they are connected or not. Hence, graphical rule we adopt is that multi-wires fuse together, as depicted graphically in Figure 4. This definition will turn to be very useful in the following chapter and will be formalized further in definition 6.1.1, when hypergraph categories will be introduced.

Example 2.4.4. The category **Rel** admits multi-wires, defined as:

$$X \times \ldots \times X \to X \times \ldots \times X :: \{((x, \ldots, x), (x, \ldots, x)) \mid x \in X\}$$

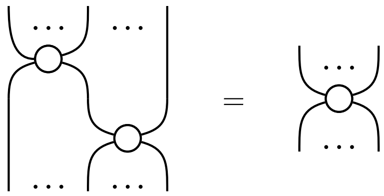


Figure 4: Spider fusion.

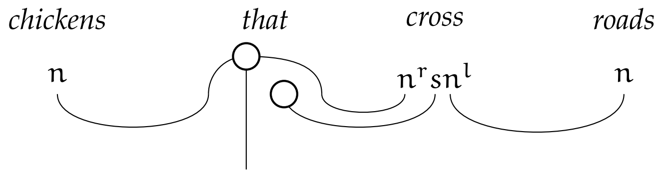
for every object X. It is easy to see that identities and the usual cups and caps can be obtained as particular instances of this definition.

Example 2.4.5. The category $\mathbf{fdHilb}_{\mathbb{R}}$ also admits multi-wires. Choosing any orthonormal basis $\{v_i\}_i$ for V, one can define them setting

$$V \otimes \ldots \otimes V \to V \otimes \ldots \otimes V :: v_{i_1} \otimes \ldots \otimes v_{i_n} \mapsto \delta_{i_1 \ldots i_n} v_{i_1} \otimes \ldots \otimes v_{i_1}$$

identities, cups and caps can again be obtained as special instances of this definition.

As promised, multi-wires help us when it comes to modeling relative pronouns [133, 134]:



Focusing on the reduction rules it is easy to see that what we get from this is a noun and not a sentence. The three-headed-multi-wire that models the word *that* acts conjoining the subject, namely *chickens*, with its verb, namely *cross*. In doing this it "stores a copy" of this conjunction and gets rid of the sentence type via the one-headed wire, giving us back a noun phrase.

The graphical calculus hereby exposed will be at the core of this entire document, and finding suitable categories on which it can be used will be one of our main tasks.

Part II

EXTENDING RELATIONS

Here we gradually extend the category of relations to accommodate our conceptual spaces formalism. We show the capabilities of our models with extensive examples. Finally, we draw general considerations about the links between grammar and semantics.

3 THE CATEGORY OF CONVEX SETS AND RELATIONS

Dum taxat, rerum magnarum parva potest res exemplare dare et vestigia notitiai.

— Lucretius, De Rerum Natura [140, Book II, Line 123]

Now that we have outlined our grammar, we focus on the semantics. The purpose of this chapter is to state in detail what has been done in [24], and to point out how the categorical treatment of convex relations allows for a broader formalization of Gärdenfors' work on conceptual spaces [70]. Moreover, we will show how categorical notion of convexity adopted here will allow us to consider as convex also structures that are not traditionally considered so, such as semilattices. The result of this is a quite expressive formalism with which we will model conceptual spaces.

Concept composition within conceptual spaces has been formalized in [8, 94, 129], for example. All these approaches focus on noun-noun composition rather than utilising any more complex structure, and the way in which nouns compose often focuses on correlations between attributes in concepts. Since then, Gärdenfors has started to formalize verb spaces, adjectives, and other linguistic structures [70]. However, he has not provided a systematic method for how to utilize grammatical structures within conceptual spaces. In this sense, the category-theoretic approach to concept composition we describe below will introduce a broader system for concept composition that can be applied to more general grammatical types.

In the distributional-categorical approach to natural language processing meanings are interpreted in categories of real vector spaces. For our intended cognitive application, we now introduce a category that emphasizes convex structure. The familiar definition of convex set is a subset of a vector space which is closed under forming convex combinations. In this document we lay down a different setting that includes convex subsets of vector spaces, but allows us to consider also further discrete examples.

We begin with some convenient notation. For a given set X we write $\sum_i p_i |x_i\rangle$ for a finite formal convex sum of elements of X, where $p_i \in \mathbb{R}^{\geqslant 0}$ and $\sum_i p_i = 1$. We moreover set $0|x\rangle = 0$ and $|x\rangle + 0 = |x\rangle$ for all $x \in X$. We then write $D(X)$ for the set of all such sums. Here we abuse the physicists ket notation to highlight that our sums are formal, following a convention introduced in [81]. Equivalently, these sums can be thought of as finite probability distributions on the elements of X.

Definition 3.0.1. A *convex algebra* is a set A together with a function $\alpha : D(A) \to A$ satisfying the following conditions

$$\alpha(|a\rangle) = a, \qquad \alpha\left(\sum_{i,j} p_i q_{i,j} |a_{i,j}\rangle\right) = \alpha\left(\sum_i p_i \left|\alpha\Big(\sum_j q_{i,j} |a_{i,j}\rangle\Big)\right\rangle\right) \tag{1}$$

Informally, α is a "mixing operation" that allows us to form convex combinations of elements, and the equations in (1) model the following good behavior

- Forming a convex combination of a single element a returns a as we would expect;
- Iterating forming convex combinations interacts as we would expect with flattening formal sums of sums.

It is worth to point out that the definition above can be elegantly formalized categorically using monads, as we will show in the next chapter. The categorical formalization of convex algebra will also allow us apply powerful machinery to study their properties.

3.0.1 *Examples*

We provide some examples of convex algebras.

Example 3.0.2. The closed real interval $[0,1]$ has an obvious convex algebra structure. Similarly, every real or complex vector space has a natural convex algebra structure using the underlying linear structure.

Example 3.0.3 (Simplices)**.** For any set X, the formal convex sums of elements of X themselves form the *free convex algebra* on X, which can also be seen as a simplex with vertices the elements of X. Mixtures are formed as follows

$$\sum_i p_i \left| \sum_j q_{i,j} |x_{i,j}\rangle \right\rangle \mapsto \sum_{i,j} p_i q_{i,j} |x_{i,j}\rangle$$

Example 3.0.4. The convex space of density matrices provides another example, with the convex structure given by the usual vector space structure on linear operators.

Example 3.0.5. For a set X, the functions of type $X \to [0,1]$ form a convex algebra pointwise, with mixing operation

$$\sum_i p_i |f_i\rangle \mapsto \left(\lambda x. \sum_i p_i f_i(x)\right)$$

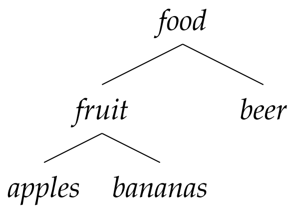
We can see this as a convex algebra of fuzzy sets.

Example 3.0.6 (Semilattices)**.** As a slightly less straightforward example, every join semilattice[1] has a convex algebra structure given by

$$\sum_i p_i |a_i\rangle = \bigvee_i a_i$$

Notice that here the scalars p_i are discarded and play no active role. These "discrete" types of convex algebras allow us to consider objects such as the Boolean truth values.

Example 3.0.7 (Trees)**.** Given a finite tree, perhaps describing some hierarchical structure, we can construct an affine semilattice in a natural way. For example, consider a limited universe of foods, consisting of bananas, apples, and beer. Given two members of the hierarchy, their join will be the lowest level of the hierarchy which is above them both. For instance, the join of *bananas* and *apples* would be *fruit*.



When α can be understood from the context, we abbreviate our notation for convex combinations by writing

$$\sum_i p_i a_i := \alpha\left(\sum_i p_i |a_i\rangle\right)$$

Putting this convention to good use, we define a *convex relation* of type $(A, \alpha) \to (B, \beta)$ as a binary relation $R : A \to B$ between the underlying sets that commutes with forming convex mixtures as follows

$$(\forall i.R(a_i, b_i)) \Rightarrow R\left(\sum_i p_i a_i, \sum_i p_i b_i\right)$$

We note that identity relations are convex, and convex relations are closed under relational composition and converse.

Example 3.0.8 (Homomorphisms)**.** If (A, α) and (B, β) are convex algebras, functions $f : A \to B$ satisfying

$$f\left(\sum_i p_i x_i\right) = \sum_i p_i f(x_i)$$

are convex relations. These functions are the *homomorphisms of convex algebras*. The identity function and constant functions are examples of convex homomorphisms.

1 Here we mean a partial order that has all finite *non-empty* joins. Many people call these structures "affine semilattices", while by "semilattice" they mean what is known in universal algebra as "complete semilattice". In this document we prefer the universal algebraic notation, so for us a "semilattice" will only have *finite non-empty* joins.

The singleton set $\{*\}$ has a unique convex algebra structure, denoted by I. Convex relations of the form $I \to (A, \alpha)$ correspond to *convex subsets*, that is, subsets of A closed under forming convex combinations.

Definition 3.0.9. We define the category **ConvexRel** as having convex algebras as objects and convex relations as morphisms, with composition and identities as for ordinary binary relations.

Given a pair of convex algebras (A, α) and (B, β) we can form a new convex algebra on the cartesian product $A \times B$, that will be denoted as $(A, \alpha) \otimes (B, \beta)$, with mixing operation

$$\sum_i p_i |(a_i, b_i)\rangle \mapsto \left(\sum_i p_i a_i, \sum_i p_i b_i \right)$$

This induces a symmetric monoidal structure on **ConvexRel**. In fact, the category **ConvexRel** has the necessary categorical structure for categorical compositional semantics:

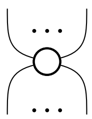
Theorem 3.0.10. *The category **ConvexRel** is a dagger compact closed category [2] The symmetric monoidal structure is given by the unit and monoidal product outlined above. Relational converse gives a dagger structure on **ConvexRel**. The cap is given by*

 $: I \to (A, \alpha) \otimes (A, \alpha) :: \{(*, (a, a)) \mid a \in A\}$

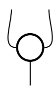
the cup by

 $: (A, \alpha) \otimes (A, \alpha) \to I :: \{((a, a), *) \mid a \in A\}$

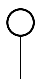
and more generally, the multi-wires by

 $: A \otimes \ldots \otimes A \to A \otimes \ldots \otimes A :: \{((a, \ldots, a), (a, \ldots, a)) \mid a \in A\}$

Every object (A, α) has a canonical commutative special dagger Frobenius structure [86], with copy

 $: (A, \alpha) \to (A, \alpha) \otimes (A, \alpha) :: \{(a, (a, a)) \mid a \in A\}$

and delete

 $: (A, \alpha) \to I :: \{(a, *) \mid a \in A\}$

2 We have given an elementary description of **ConvexRel**. More abstractly, it can be seen as the category of relations for the Eilenberg-Moore category of the finite distribution monad, see Chapter 4.

Proof. **ConvexRel** can be seen as the category of relations for the Eilenberg-Moore category of the finite distribution monad. Monads on sets are regular categories [26, Thm 4.3.5], and categories of relations over regular categories have a commutative special dagger Frobenius structure on every object [77, Thm 3.4, Ex. 3.5]. Compact closedness is trivially implied by the Frobenius structure, that induces a cup and a cap on every object (see for instance [43]). □

We note that the tensor product of **ConvexRel** is not a category theoretic product. For example, there are convex subsets of $[0, 1] \times [0, 1]$ such as

$$\{(x, x) \mid x \in [0, 1]\}$$

that cannot be written as the cartesian product of two convex subsets of $[0, 1]$. This behavior exhibits non-trivial *correlations* between the different components of the composite convex algebra: We have a genuine tensor.

4 CHARACTERIZATION OF CONVEX ALGEBRAS

Quibus vero natura tantum tribuit
sollertiae, acuminis, memoriae,
ut possint geometriam, astrologiam, musicen
ceterasque disciplinas penitus habere notas,
praetereunt officia architectorum
et efficiuntur mathematici.

— *Vitruvius, De Architectura [102, Book I, 1.17]*

In the last chapter we gave a definition of convex relations that works on a higher degree of generality with respect to the definition of convexity in terms of vector spaces we are used to: This is one of the benefits of using category theory: Things can be done in full generality. The purpose of this chapter is to analyze this definition in detail to understand its properties and limitations, so that we will be able to use it later on in a principled way. We will start recasting our definition using the formal language of monads, the we will study convex algebras from a universal algebraic perspective, getting some interesting cardinality results and understanding a bit better how convex algebras and semilattices are connected. Finally, we will study connections between our convex relations and betweenness relations, that constitute an interesting mathematical framework to describe what it means "to be between two things". This is interesting since such view is also the naïve interpretation we usually give to concepts as convexity, so studying the links between these different mathematical models seems natural. The results in this sections have all been proved by the author of this document.

4.1 ALGEBRAIC CHARACTERIZATION

In the previous chapter we denoted finite formal combinations on a set X as $\sum_i p_i |x_i\rangle$, and the set of all formal combinations on X as $D(X)$. Let's make this a bit more formal.

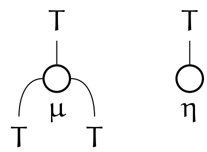
Definition 4.1.1 (Monad)**.** Given a category $\mathcal{C}$, an *endofunctor* is a functor $T : \mathcal{C} \to \mathcal{C}$. We denote with T^n the n-fold composition of T with itself. If there are natural transformations $\mu : T^2 \to T$ and $\eta : id_{\mathcal{C}} \to T$

(here $\mathrm{id}_{\mathcal{C}}$ is the identity functor on $\mathcal{C}$) such that the following diagrams commute

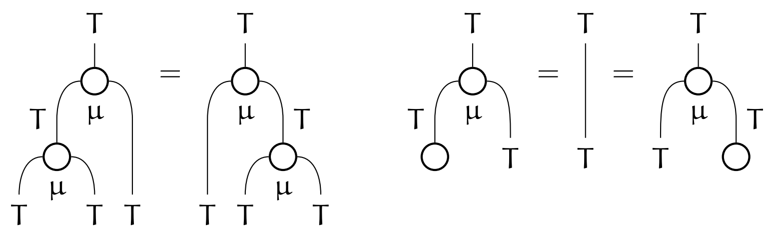
$$\begin{array}{ccc} T^3 & \xrightarrow{T\mu} & T^2 \\ {\scriptstyle \mu T}\downarrow & & \downarrow{\scriptstyle \mu} \\ T^2 & \xrightarrow[\mu]{} & T \end{array} \qquad \begin{array}{ccc} T & \xrightarrow{T\eta} & T^2 \\ {\scriptstyle \eta T}\downarrow & \diagdown\!\!\diagdown & \downarrow{\scriptstyle \mu} \\ T^2 & \xrightarrow[\mu]{} & T \end{array}$$

Then the triple (T, η, μ) is called a *monad* on $\mathcal{C}$.

Another interesting way to see monads is in terms of monoid objects. If you fix a category $\mathcal{C}$, then you can consider the category $[\mathcal{C}, \mathcal{C}]$ having endofunctors $\mathcal{C} \to \mathcal{C}$ as objects and natural transformations between them as morphisms. We can apply our diagrams in this setting: We represent an endofunctor T as a typed wire and a natural transformation as a process. The identity endofunctor $\mathrm{id}_{\mathcal{C}}$ is taken to be the trivial wire. Having a monad then means specifying two natural transformations



Such that they respect the following equations, graphically depicted:



But these are exactly the associative and unit law of a monoid: We can then say that a monad on $\mathcal{C}$ is just a monoid in the category of endofunctors on $\mathcal{C}$ and natural transformations between them.

Definition 4.1.2 (Eilenberg-Moore algebra)**.** Given a monad (T, η, μ) on a category $\mathcal{C}$, a *T-algebra* (A, α) is an object A together with a morphism $\alpha : TA \to A$ such that the following diagrams commute:

$$\begin{array}{ccc} T^2A & \xrightarrow{T\alpha} & TA \\ {\scriptstyle \mu_A}\downarrow & & \downarrow{\scriptstyle \alpha} \\ TA & \xrightarrow[\alpha]{} & A \end{array} \qquad \begin{array}{ccc} A & \xrightarrow{\eta_A} & TA \\ & \diagdown\!\!\diagdown & \downarrow{\scriptstyle \alpha} \\ & & T \end{array}$$

A morphism of T-algebras $(A, \alpha) \to (B, \beta)$ is a morphism $f : A \to B$ such that the following diagram commutes:

$$\begin{array}{ccc} TA & \xrightarrow{Tf} & TB \\ \alpha \big\downarrow & & \big\downarrow \beta \\ A & \xrightarrow[f]{} & B \end{array}$$

T-algebras and their morphisms form a category, called the *Eilenberg-Moore category of the monad* T.

It is well known that the function $D : X \to D(X)$ can be made functorial on **Set** and this defines a monad (D, η, μ) with identity η and multiplication μ

$$\eta : x \in X \to |x\rangle \in D(X),$$

$$\mu : \sum_i p_i \left| \left(\sum_j p_{j_i} \left| x_{j,i} \right\rangle \right) \right\rangle \in D^2(X) \to \sum_{i,j} p_i p_{j_1} \left| x_{j_i} \right\rangle \in D(X)$$

A convex algebra on X is then defined as an Eilenberg-Moore algebra for (D, η, μ). First of all, we want to show that the set X can be endowed with an uncountable set of algebraic operations having signature 2.

Definition 4.1.3. Let (X, α) be an Eilenberg-Moore algebra on the monad (D, η, μ). For each $p \in [0, 1]$, the *operation* $+_p : X \times X \to X$ is defined as

$$\forall x, y \in X, \quad +_p(x, y) \overset{def}{\Longleftrightarrow} \alpha(p\eta(x) + (1-p)\eta(y))$$

We will make almost exclusive use of the operation $+_p$ in infix notation, so we will always write $x +_p y$ instead of $+_p(x, y)$.

Proposition 4.1.4. $\langle X, \{+_p\}_{p \in [0,1]} \rangle$ *is an algebra. Moreover, it satisfies the equations*

ZERO COMBINATIONS $x +_0 y \simeq y$;

IDEMPOTENCY $x +_p x \simeq x$;

PARAMETRIZED COMMUTATIVITY $x +_p y \simeq y +_{1-p} x$;

PARAMETRIZED ASSOCIATIVITY $(x +_p y) +_q z \simeq x +_{pq} (y +_{\frac{q-pq}{1-pq}} z)$.

Note that the last three properties represent schemes *of equations, since they have to hold for every* p *(we cannot quantify on* p *because the* $+_p$ *are operations).*

When we regard $\langle X, \{+_p\}_{p\in[0,1]}\rangle$ as an algebra, $D(X)$ is the set of terms generated by the operations $+_p$: Just unfolding the definition of $+_p$, one can prove that

Corollary 4.1.5. *The following equality holds for every* $p_1\eta(x_1) + \cdots + p_n\eta(x_n) \in D(X)$*:*

$$\begin{aligned}\alpha(p_1\eta(x_1) + \ldots + p_n\eta(x_n)) &= \\ &= x_1 +_{f(1)} (x_2 +_{f(2)} (\ldots (x_{n-1} +_{f(n-1)} x_n) \ldots)\end{aligned}$$

Where $f(i)$ *is defined recursively as* $f(1) = p_1,\ f(i) = \frac{p_i}{(1-f(1))\ldots(1-f(i-1))}$.

This accounts for a completely algebraic representation of the generic convex combination on X and is useful to prove the "inverse" of Proposition 4.1.4, namely:

Proposition 4.1.6. *Let* $\langle X, \{+_p\}_{p\in[0,1]}\rangle$ *be an algebra where every* $+_p$ *has ariety two and respects zero combinations, idempotency, parametrized commutativity and parametrized associativity. Then* $\langle X, \{+_p\}_{p\in[0,1]}\rangle$ *is an Eilenberg-Moore algebra* (X, α) *with respect to the monad* (D, η, μ).

Proof. The generic element in $D(X)$ is in the form

$$\sum_{i=1}^{n} p_i \left|x_i\right\rangle = \sum_{i=1}^{n} p_i\eta(x_i) = p_1\eta(x_1) + \cdots + p_n\eta(x_n)$$

Define $\alpha : D(X) \to X$ as in Proposition 4.1.5. Parametrized commutativity and associativity ensure that α is well defined, meaning that it does not depend on how the convex combinations are presented: From the base case

$$\begin{aligned}\alpha(p\eta(x) + (1-p)\eta(y)) &= x +_p y \\ &= y +_{1-p} x \\ &= \alpha((1-p)\eta(y) + p\eta(x))\end{aligned}$$

Using parametrized associativity one can prove, by induction, that every permutation of sums in $D(X)$ goes through α to the same element in X. Idempotency allows us to prove the first requirement to be an Eilenberg-Moore algebra:

$$\alpha(\eta(x)) = \alpha(1\eta(x)) = \alpha(p\eta(x) + (1-p)\eta(x)) = x +_p x = x$$

It remains to prove that α behaves well with respect to to μ:

$$\alpha\left(\sum_{i,j} p_i q_{i,j}\eta(x_{i,j})\right) = \alpha\left(\sum_i p_i\eta\left(\alpha\left(\sum_j q_{i,j}\eta(x_{i,j})\right)\right)\right)$$

Applying the definition,

$$\alpha\left(\sum_{i,j} p_i q_{i,j}\eta(x_{i,j})\right) = x_{1,1} +_{f(1,1_1)} (\ldots (x_{n,m_n-1} +_{f(n,m_n-1)} x_{n,m_n}) \ldots)$$

With f defined as:

$$f(1,1_1) = p_1 q_{1,1},$$
$$f(i,j_i) = \frac{p_i q_{i,j}}{(1-f(1,1_1))\dots(1-f(1,m_1))(1-f(2,1_2))\dots(1-f(i,j_i-1))}$$

On the other hand,

$$\alpha\left(\sum_i p_i\eta\left(\alpha\left(\sum_j q_{i,j}\eta(x_{i,j})\right)\right)\right) =$$
$$= \alpha\left(\sum_i p_i\eta(x_{i,1} +_{g(i,1)} (\dots(x_{i,m_i-1} +_{g(i,m_i-1)} x_{i,m_i})\dots)\right)$$
$$= (x_{1,1} +_{g(1,1)} (\dots(x_{1,m_1-1} +_{g(1,m_1-1)} x_{i,m_1})\dots))+_{k(1)}$$
$$+_{k(1)} (\dots(x_{n,1} +_{g(n,1)} (\dots(x_{n,m_n-1} +_{g(n,m_n-1)} x_{n,m_n})\dots))\dots)$$

With g defined as $g(1,1) = q_{1,1}$, $g(i,j) = \frac{q_{i,j}}{(1-g(i,1))\dots(1-g(i,j-1))}$, and k defined as $k(1) = p_1$, $k(i) = \frac{k_1}{(1-k(1))\dots(1-k(i-1))}$. The equality is then obtained using parametrized associativity to rearrange parentheses in the second expression (the one in g and k) to match the disposition they have in the first (in f). □

Definition 4.1.7. We denote with **Convex** the category having convex algebras as objects and convex preserving functions (as in 3.0.8) as arrows.

Now consider $\mathcal{V}$, the class of algebras $\langle X, \{+_p\}_{p\in[0,1]}\rangle$ that satisfy zero combinations, idempotency, parametrized associativity and parametrized commutativity. Thanks to the Birkhoff theorem, this class is a variety, (that is, it is closed under direct products, subalgebras and homomorphic images). Connections between algebraic varieties and algebras of monads have being thoroughly investigated, see for example[80]. Albeit the abundance of general results, in our case we are able to prove some things directly: In fact, propositions 4.1.4 and 4.1.6 allow us to infer that

Theorem 4.1.8. ***Convex*** *can be regarded as the algebraic variety of Convex Algebras:* $\mathcal{V}$ *and* ***Convex*** *are isomorphic as categories.*

Proof. First of all we prove that every object in **Convex** can be uniquely sent to a convex algebra in $\mathcal{V}$, and vice-versa. Consider the function $f : \mathrm{obj}(\textbf{ConvexRel}) \to \mathcal{V}$ that carries every (X, α) in $\langle X, \{+_p\}_{p\in[0,1]}\rangle$ with operations defined as in 4.1.3. These objects are in $\mathcal{V}$ because of Proposition 4.1.4. Clearly different underlying sets in Eilenberg-Moore algebras give rise to different algebras. Now suppose $(X, \alpha) \neq (X, \beta)$. Then $\alpha(|x\rangle) \neq \beta(|x\rangle)$ for some x. This means $\alpha(p_1\eta(x_1) + \dots + p_n\eta(x_n)) \neq \beta(p_1\eta(x_1) + \dots + p_n\eta(x_n))$ for some convex combination $p_1\eta(x_1) + \dots + p_n\eta(x_n)$. But then, because of Proposition 4.1.5

$f((X,\alpha)), f((X,\beta))$ have different terms, so they are different algebras and f is injective.

On the other hand, consider $f' : \mathcal{V} \to \mathrm{obj}(\mathbf{ConvexRel})$ that associates to every$\langle X, \{+_p\}_{p\in[0,1]}\rangle$ an Eilenberg-Moore algebra (X,α) with α defined as in Proposition 4.1.5. Obviously algebras with different supports are different, and go to different objects in **ConvexRel**. It remains to prove that if $\langle X, \{+_p\}_{p\in[0,1]}\rangle \neq \langle X, \{+'_p\}_{p\in[0,1]}\rangle$ then also $f'(\langle X, \{+_p\}_{p\in[0,1]}\rangle) \neq f'(\langle X, \{+'_p\}_{p\in[0,1]}\rangle)$. Denote:

$$f'(\langle X, \{+_p\}_{p\in[0,1]}\rangle) \overset{def}{\iff} (X,\alpha)$$
$$f'(\langle X, \{+'_p\}_{p\in[0,1]}\rangle) \overset{def}{\iff} (X,\beta)$$

If $\langle X, \{+_p\}_{p\in[0,1]}\rangle \neq \langle X, \{+'_p\}_{p\in[0,1]}\rangle$ then at least two $+_p, +'_p$ are different functions, and so $x +_p y \neq x +'_p y$ for some x, y. But then $\alpha(p\eta(x) + (1-p)\eta(y)) \neq \beta(p\eta(x) + (1-p)\eta(y))$ and $(X,\alpha) \neq (X,\beta)$, as we wanted. To conclude the proof we just have to note how any morphism of **Convex** trivially preserves convex operations, and hence is a morphism of $\mathcal{V}$. The opposite is also trivially true, and we are done. □

We can now apply some nice universal algebraic tools to prove properties.

Proposition 4.1.9. *The monomorphisms in **Convex** are exactly the injective homomorphisms.*

Proposition 4.1.10. *In the category **Convex** the image of a convex morphism $h : (A,\alpha) \to (B,\beta)$ is a convex subalgebra of (B,β). Moreover, assume $h : (A,\alpha) \to (B,\beta)$ is a monomorphism in **Convex**. There is then a convex morphism $\mathrm{im}\, h \to (A,\alpha)$ given by inverse images.*

4.1.1 *Semilattices*

Definition 4.1.11. Let (X,α) be an object of **ConvexRel**. We say that α *disregards weights* if, for every $x, y \in X, p, q \in]0,1[$, it is

$$\alpha(p\,|x\rangle + (1-p)\,|y\rangle) = \alpha(q\,|x\rangle + (1-q)\,|y\rangle)$$

Proposition 4.1.12. (X,α) *disregards weights iff the corresponding algebra in $\mathcal{V}$ satisfies the equation $\forall x, y \in X, x +_p y \simeq x +_q y$ for every couple of p, q such that $p, q \neq 0, 1$.*

Call $\mathcal{V}_{\mathcal{L}}$ the class of algebras in $\mathcal{V}$ that disregard weights, that is, the algebras that satisfy the supplementary set of equations given in Proposition 4.1.12. Again, the existence of an equational definition is enough to apply apply Birkhoff theorem and state that $\mathcal{V}_{\mathcal{L}}$ is a variety. Then we have:

Lemma 4.1.13. *$\mathcal{V}_{\mathcal{L}}$ is equi-interpretable with the variety of semilattices.*

Proof. In the variety $\mathcal{V}_{\mathcal{L}}$ the following equations hold:

- $a +_p a \simeq a$;
- $a +_p b \simeq b +_{1-p} a \simeq b +_p a$ if $p \neq 0, 1$;
- $(a +_p b) +_p c \simeq a +_{pp} (b +_{\frac{p-pp}{1-pp}} c)$. If $p \neq 0, 1$ also $pp \neq 0, 1$ and hence $pp \neq p$, that means $\frac{p-pp}{1-pp} \neq 0, 1$. Weight disregarding implies $a +_{pp} (b +_{\frac{p-pp}{1-pp}} c) \simeq a +_p (b +_p c)$.

Interpreting elements into themselves and $a \vee b$ into $a +_p b$ for some $p \neq 0, 1$ ensures that the operation $+_p$ is idempotent, commutative and associative, so $\mathcal{V}_{\mathcal{L}}$ has a copy of the defining semilattice equations in itself, that is, semilattices are interpretable in $\mathcal{V}_{\mathcal{L}}$.

In the other direction, interpret elements into themselves, every term $a +_0 b$ into the term b and every term $a +_1 b$ into the term a. Hence zero combinations hold by definition. Interpret moreover every term $a +_p b$ with $p \neq 0, 1$ into $a \vee b$. Weights disregarding equations are then trivially satisfied since $a \vee b \simeq a \vee b$.

Idempotency holds because $a \vee a = a$ ensures $a +_p a = a$ for all $p \neq 0, 1$, while $a +_0 a = a$ and $a +_1 a = a$ trivially hold, so $a +_p a = a$ for all p.

Parametrized commutativity is similar, $b = a +_0 b$ and $b = b +_1 a$ trivially imply $a +_0 b = b +_1 a$, while $a +_p b = a +_{1-p} b$ holds because of the commutativity of $\vee$ along with the fact that if $p \neq 0, 1$ then $1 - p \neq 0, 1$.

For parametrized associativity, take $(a +_p b) +_q c$. If p, q are not 0 or 1 then $(a +_p b) +_q c$ gets interpreted into $(a \vee b) \vee c$. Moreover $p, q \neq 0, 1$ imply $pq, \frac{q-pq}{1-pq} \neq 0, 1$, then $a +_{pq} (b +_{\frac{q-pq}{1-pq}} c)$ goes into $a \vee (b \vee c)$ and the equation holds because of the associativity of $\vee$. The cases with p, q equal to 0 or 1 are trivial.

Now we have to prove that if we apply interpretations back and forth we go back to where we came from. Starting from lattices, elements trivially go back to themselves, and $a \vee b$ goes into $a +_p b$ that goes again into $a \vee b$ because $p \neq 0, 1$. From this we can inductively prove that every term in the variety of semilattices goes back to itself.

In the other direction the same argument holds for elements. The term $a +_0 b$ goes into b, b is interpreted into itself and $b \simeq a +_0 b$. For $a +_1 b$ it is sufficient to note that this is equal to $b +_0 a$. Finally $a +_p b$ with $p \neq 0, 1$ goes to $a \vee b$ that goes back to $a +_q b$ for some $q \neq 0, 1$. Then we have $a +_p b \simeq a +_q b$. □

This result states something that intuitively sounds like "If we squash all the $+_p$ into one operation, parametrized associativity and parametrized commutativity just become usual associativity and commutativity properties". Since idempotency, associativity and commutativity are the equations defining semilattices as algebraic structures, we can consider the algebras (X, α) that disregard weights as semilattices themselves.

In the algebraic formalisation of semilattices there is no real difference between meet and join (in contrast with what happens defining these structures order-theoretically): We can then interpret the operation $+_p$ as a meet or a join. Anyway, in this work the join notation is usually preferred, and we will stick to this convention when working with semilattices. The only exception is given by this observation: We would be tempted to consider the convex algebra Ω induced by the two element meet semilattice $\{\bot, \top\}$ with $\bot \leqslant \top$ as a subobject classifier in **Convex**, motivated by the following fact:

Proposition 4.1.14. *Let Ω be the convex algebra induced by the two element meet semilattice $\{\bot, \top\}$ with $\bot \leqslant \top$. For a finite set X, there is a bijective correspondence between **Convex**-morphisms of type $F(X) \to \Omega$ and sub-complexes of $F(X)$, where $F(X)$ denotes the* free convex algebra over X.

This is nevertheless not possible, since:

Counterexample 4.1.15. There are many more subalgebras of $F(X)$ than just the sub-complexes. It follows that Ω cannot serve as a subobject classifier, as there are insufficient morphisms $F(X) \to \Omega$ to serve as characteristic morphisms.

4.2 CARDINALITY CHARACTERIZATIONS

Now that we built up the basic algebraic machinery to look at Convex Algebras from a set-theoretic point of view, we can start pursuing a characterization of Convex Algebras in terms of cardinalities of their support sets.

This characterization will be worked out gradually, and will be heavily influenced by the topological properties of $\mathbb{R}$ that rule the relations among the $+_p$ operations through the choice of p. We will heavily rely on Lemma 4.1.8 to freely "confuse" a convex algebra (X, α) with its universal algebraic counterpart. The first thing we need is a little hack:

Definition 4.2.1. Let (X, α) be a convex algebra. α is *coherent* if, every time $a +_{p'} b = a +_p b$ with $p' \leqslant p$, then $\forall z \in [p', p]$ it is $a +_z b = a +_p b$.

Asking for coherence intuitively means that if the result of two different convex combinations between the same elements and with different weights gives back the same element, then every other convex combination "between" the two has to attain the same result.

This definition, albeit being something quite reasonable to ask if we want to interpret convex combinations as mixing of elements (as we normally do in the Conceptual Spaces framework), will be readily dropped: In fact, if we spoke about "hack" above it is because this useful property de facto comes for free:

Proposition 4.2.2. *Let* (X, α) *be a convex algebra.* (X, α) *is coherent.*

We can now prove some basic "cardinality facts". The proofs are tedious but can be easily worked out by cases.

Lemma 4.2.3. *Let* (X, α) *be a convex algebra. If* α *does not disregard weights, then* $|X| \neq 2$.

Lemma 4.2.4. *Let* (X, α) *be a convex algebra. If* α *doesn't disregard weights, then* $|X| \neq 3$.

These lemmas have a central role in the proof of the following theorem, responsible for the main characterization of convex algebras in terms of cardinality.

Theorem 4.2.5. *Let* (X, α) *be a convex algebra. If* α *does not disregard weights, then* X *is infinite.*

Proof. By Proposition 4.2.2 (X, α) is coherent. Since α does not disregard weights, there are $p, q \in]0, 1[$ such that $a +_p b \neq a +_q b$ for some $a, b \in X$. Suppose, without loss of generality, that $p < q$. Consider

$$\bar{p} := \sup\{p' \in [0, 1] | a +_{p'} b = a +_p b\}$$
$$\bar{q} := \inf\{q' \in [0, 1] | a +_{q'} b = a +_q b\}$$

These elements exists because $\mathbb{R}$, and hence $[0, 1]$, are complete.

Coherency implies that the segments of the real line $[p, \bar{p}[$ and $]\bar{q}, q]$ get sent by α to $a +_p b$, $a +_q b$, respectively. There are now three different possibilities:

CASE 1: $p < \bar{p} = \bar{q} < q$. If $a +_p b = a +_{\bar{p}} b$, then the algebra $\langle \{a +_p b, a +_q b\}, \{+_z\}_{z \in [0,1]} \rangle$ is a sub-algebra of $\langle A, \alpha \rangle$ that doesn't disregard weights having cardinality of the support equal to 2, contradicting Lemma 4.2.3. Same happens if $a +_q b = a +_{\bar{p}} b$. Now, if $a +_p b \neq a +_{\bar{p}} b \neq a +_q b$ then the algebra given by $\langle \{a +_p b, a +_q b, a +_{\bar{p}} b\}, \{+_z\}_{z \in [0,1]} \rangle$ is a sub-algebra of (A, α) that doesn't disregard weights having cardinality of the support equal to 3, contradicting Lemma 4.2.4.

CASE 2: $p = \bar{p} = \bar{q}$. First, there is no $p' < p$ such that $a +_{p'} b = a +_p b$, otherwise we could go back to the previous case and defining $\langle \{a +_{p'} b, a +_q b\}, \{+_z\}_{z \in [0,1]} \rangle$ would produce a contradiction. Since $0 < p$, there is a $0 < z < p$ such that $a +_z b \neq a +_p b$. Suppose that $a +_{z'} b = a +_z b$ for all $z' \in [z, p[$: We proceed as in the previous case, and the subalgebra of (X, α) given by $\langle \{a +_z b, a +_p b, a +_q b\}, \{+_k\}_{k \in [0,1]} \rangle$ does not disregard weights with cardinality of the support equal to 3 contradicting again Lemma 4.2.4. Then no such z can exist and we finally infer that $\forall z \in]0, p[, \exists z' \in [z, p[: a +_{z'} b \neq a +_z b$. The case $\bar{p} = \bar{q} = q$ is analogous.

CASE 3: $\bar{p} \neq \bar{q}$. Here, from the definition of $\bar{p}, \bar{q}$ we obtain that for all $z \in [\bar{p}, \bar{q}]$, it is $a +_p b \neq a +_z b \neq a +_q b$.

Now, if case 2 happens for some $a +_p b, a +_q b$ we are done. If not, then then there is a third point $a +_z b$ that is not equal to $a +_p b$ nor $a +_q b$, with $z \in]p, q[$. We can then repeat our reasoning with $a +_z b$ and $a +_p b$ and fall again into case 3, finding a new point $p < k < z < q$. Coherence implies all these $+_p, +_k, +_z, +_q$ send a, b into different elements. The reasoning can be reiterated by induction and clearly holds for every natural number, so X is infinite again. □

Note that the opposite implication is not true: Lattices disregard weights by definition and they can as well be infinite.

The author strongly believes that this result can be further improved, claiming that if α disregards weights the cardinality of X must at least be the same of $\mathbb{R}$ (meaning that if two non trivial convex combinations between elements are different, then every convex combination between them gets sent into a different element), but this is not really important: The relevant thing here is how Theorem 4.2.5 poses serious limitations on the choice of the weighting system.

Essentially what it states is that using the formalism of convex algebras there is no way to express the idea that two elements can have only $2, 3$ etc. fixed mixings: If one has to model how concepts, words, meanings or whatever interact between each other using convex algebras, given a couple of elements either one defines a mixing accounting for a plethora of different mixing instances, or he just defines mixing as a "platonic" concept (semilattice case) that is, one just expresses the fact that "a, b are somehow mixed together".

This, finally, means that the only way to account for finite different mixings in convex relations is to tweak directly with the categorical definition of **ConvexRel**, using a different semiring to form the monad of convex combinations (for instance, one could take a finite field $\mathbb{Z}_p$, with p prime, as semiring). This looks promising since changing the semiring structure for the distribution monad is functorial in the choice of semiring: As proved in [1], a homomorphism of semirings $R \to S$ induces a functor between the distribution monads built on R and S, respectively. This allows us to explore this direction of research in a controlled, parametrized way. Said investigations will be object of future work: For now, the author doesn't generally believe this limitation on the weight choice to be pathological with regard to the applications they ought to model (as in [24]), but its existence is good to know whenever someone may be wanting to use **ConvexRel** to model different phenomena.

4.3 BETWEENNESS RELATIONS

Since **ConvexRel** includes some objects extremely geometrical in nature (think about real vector spaces, canonical examples of euclidean geometry) it makes sense to ask if there is a common unifying geometrical ground for convex algebras inherited by the convex structure.

There is no obvious way to reason about a generic concept of metric for the elements of **ConvexRel**, hence the quest for a common geometrical framework between them has to start, for the sake of simplicity, with a definition of geometry that does not assume measure as a basic notion.

Betweenness relations constitute the defining concept we are looking for. Being the fulcrum of order geometry, these relations ought to model the idea of a point being between other two and provide a common framework for affine, euclidean, absolute, and hyperbolic geometry.
Many different versions of the axioms defining betweenness relations have been proposed across the XIX and the XX century by illustrious mathematicians, among others Pasch, Peano, Hilbert, Veblen [47]. Taking into account our last claim, we start clearing any ambiguity adopting the following definition of betweenness:

Definition 4.3.1. A *beetweenness relation* on X is a ternary relation B such that

1. $(a, b, c) \in B \Rightarrow (c, b, a) \in B$;
2. $(a, b, a) \in B \Rightarrow a = b$;
3. $\forall a, b, \in X, \exists c : (a, b, c) \in B$;
4. $\forall a, b, \in X, \exists c : (a, c, b) \in B$;
5. $(a, b, c), (b, a, c) \in B \Rightarrow a = b$;
6. $(a, b, c), (b, c, d) \in B \Rightarrow (a, b, d) \in B$;
7. $(a, b, d), (b, c, d) \in B \Rightarrow (a, b, c) \in B$.

As we already hinted, $(a, b, c) \in B$ can be interpreted as *the element* b *is between the elements* a *and* c.

We want to understand up to what point **ConvexRel** behaves well with respect to betweenness relations. Relying on the structure of convex algebras we can define a betweenness relation as:

Definition 4.3.2. Let (X, α) be a convex algebra. Define a ternary relation $B \subseteq X \times X \times X$ as follows:

$$(a, b, c) \in B \overset{def}{\iff} \exists p \in [0, 1] : a +_p c = b$$

B is called *the betweenness relation on* (X, α).

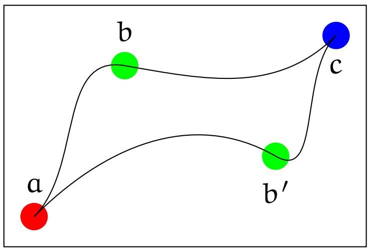


Figure 5: In a convex algebra situations like this never happen.

This relation is by definition compatible with the convex structure of (X, α) and could be roughly interpreted as "if b is a convex combination of a and c, then b is between a and c". As the name we chose suggests, these relations are very similar to the betweenness relations of Definition 4.3.1.

Lemma 4.3.3. *Let (X, α) be a convex algebra, and let B be the beetweennes relation on X. Then B satisfies axioms 1,2,3,4,5,7 of Definition 4.3.1.*

Counterexample 4.3.4. Convex algebras do not satisfy, in general, axiom 6. Take a semilattice with $b = c, b \neq a \neq d \neq a \vee d$. Then $a +_0 c = c = b$, $b +_1 d = b = c$ and hence $(a, b, c), (b, c, d) \in B$, but $(a, b, d) \notin B$.

A betweenness relation defined on (X, α) satisfies also other axioms, in particular:

Lemma 4.3.5. *Let (X, α) be a convex algebra and let B be the betweenness relation on X. Then B satisfies the property*

$$(a, b, c), (a, b', c) \in B \Rightarrow (a, b, b') \text{ or } (a, b', b) \in B$$

This means that B forces all the different convex combinations of the same couple of elements to be "in line", in the sense that situations as the one in Figure 5 never happen.

5 THINGS WE CAN DO

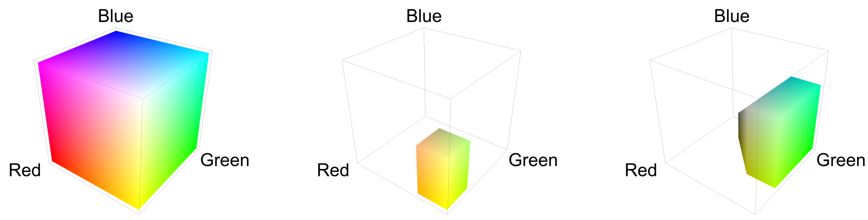
Plebeia ingenia magis exemplis
quam ratione capiuntur.

— *Macrobius, Saturnalia [99, Book VII, 4.4]*

Now we put the formalism developed up to this point to good use. First of all, we call an object of **ConvexRel** *conceptual space*. In our applications we will need two objects, namely N to represent the noun space and S to represent the sentence space. These will correspond to their analogues in the pregroup grammar. The reader familiar with conceptual space semantics will surely note how our choice of sentence space will be tailored on the specific examples we will give. This is because finding a satisfying, general definition for the sentence space should is a very hard problem, if not the hardest in distributive/conceptual space semantics. Many different ideas have been proposed: Gärdenfors, for instance, advocates for a model in terms of events [70]. Nevertheless, albeit the abundance of proposals there isn't, at least for now, a general agreement on what the most viable definition of sentence space should be. Worse than this, we are not even sure if this definition is context-dependent or if there is a general, broad way to define it that can encompass the specialized definitions one gives when examples are considered. Since framing this kind of issue with a satisfactory amount of detail would probably produce enough material for a thesis on its own, we preferred to provide simple, context dependent definitions of sentence spaces that will allow us to produce examples and to illustrate how the categorical formalism we developed up to now can be used. We can only hope that the aware linguist will forgive us.

This chapter is articulated as follows: In Section 5.1 we will describe in detail the noun space giving examples, then we will focus on the description of adjectives and verbs. This will give us enough grammatical types to make things interact together. In Section 5.2 we will study this interaction by means of other examples.

The classes of examples we will consider are *food and drink*, in which we focus on intrinsic properties of objects as taste, and *robot navigation*, in which we will give a representation of spatial movement. All the examples here have been worked out with the aid of coauthors (in particular Martha Lewis has to be credited for many of them).



(a) RGB colour cube (b) The property *yellow* (c) The property *green*

Figure 6: RGB colour space and *colour* properties.

5.1 ADJECTIVE AND VERB CONCEPTS

The noun space N is defined as a monoidal product:

$$N = N_{colour} \otimes N_{taste} \otimes \dots$$

where every component will represent an attribute we want to describe (such as colour, taste etc. as the subscripts suggest). A noun, being a state $I \to N$, will just be a convex subset of N. The sentence space is just a convex algebra, in which we will interpret events as individual points. We proceed giving some examples of nouns and sentences to further illustrate what we mean.

5.1.1 *Example: Food and drink*

We ought to model nouns corresponding to food and drinks. The properties we care about are *colour*, *taste* and *texture*, hence we choose to represent our noun space as a decomposition of domains N_{colour}, N_{taste}, $N_{texture}$, meaning that $N = N_{colour} \otimes N_{taste} \otimes N_{texture}$.

The domain N_{colour} is the RGB colour space (everyone familiar with image editing software should be confortable with this, see Figure 6). Its elements are triples $(R, G, B) \in [0, 1]^3$ where R, G, B represent the intensity of the colour red, green and blue respectively. Another possible choice would have been to use the HSV space, in which colours are defined in terms of hue and saturation. Representing convex combinations on the latter space is trickier though, and this is the main reason to prefer the RGB choice.
The domain N_{taste} is defined as the simplex having four fundamental tastes as its vertexes, namely *sweet*, *sour*, *bitter* and *salt* (Figure 7). N_{taste} can then be written as:

$$N_{taste} = \{\vec{t} | \vec{t} = \sum_{i \in I} w_i \vec{t}_i\}$$

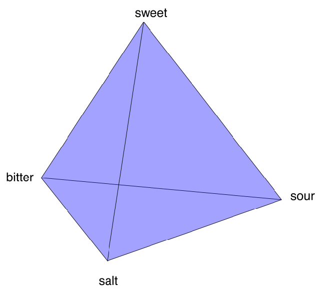
with $I = \{sweet, sour, bitter, salt\}$, $\sum_i w_i = 1$ and $\{\vec{t}_i\}$ the computational basis for the vector space $\mathbb{R}^4$.

$N_{texture}$ is the interval $[0,1]$ representing viscosity, where 0 stands for "completely liquid" and 1 for "completely solid".

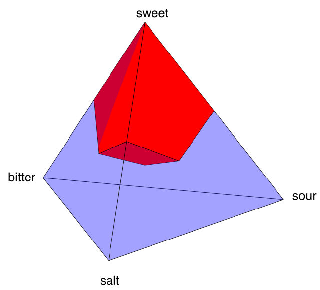
We call a convex subset of some domain *property*. Some examples of properties are defined below and illustrated in Figures 6 and 7.:

$$yellow = \{(R,G,B)|(R \geqslant 0.7),(G \geqslant 0.7),(B \leqslant 0.5)\}$$
$$green = \{(R,G,B)|(R \leqslant G),(B \leqslant G),(R \leqslant 0.7),(B \leqslant 0.7),(G \geqslant 0.3)\}$$
$$sweet = \{\vec{t}|t_{sweet} \geqslant t_l \text{ for } l \neq sweet\}$$

We can similarly define other properties we will use later in the same way, such as *sour* and *bitter*. We can use properties as components to



(a) The taste tetrahedron



(b) The region corresponding to $sweet = \{\vec{t}|t_{sweet} \geqslant t_l \text{ for } l \neq sweet\}$

Figure 7: The taste space and the property *sweet*.

define nouns. To do this in a pleasant way, we will denote the convex hull of a set A as $Cl(A)$.

$$banana = \{(R,G,B)|(0.9R \leqslant G \leqslant 1.5R),(R \geqslant 0.3),(B \leqslant 0.1)\}\times$$
$$\times\, Cl(\{sweet, 0.25sweet + 0.75bitter, 0.7sweet + 0.3sour\}) \times [0.2, 0.5]$$
$$apple = \{(R,G,B)|R - 0.7 \leqslant G \leqslant R + 0.7),(G \geqslant 1 - R),(B \leqslant 0.1)\}\times$$
$$\times\, [0.5,1] \times Cl(\{sweet, 0.75sweet + 0.25bitter, 0.3sweet + 0.7sour\}) \times [0.5, 0.8]$$
$$beer = \{(R,G,B)|(0.5R \leqslant G \leqslant R),(G \leqslant 1.5 - 0.8R),(B \leqslant 0.1)\}\times$$
$$\times\, Cl(\{bitter, 0.7sweet + 0.3bitter, 0.6sour + 0.4bitter\}) \times [0, 0.01]$$

The only part that deserves a bit of explanation in this definition is the taste one. Take *apple* as an example: Apples are never salty, hence the *salt* parameter is always zero in the convex combination defining the taste of *apple*. They can instead be very sweet, so *sweet* is chosen as one of the extremes defining the convex combination. Apples can moreover be a little bit, but not totally, bitter or sour, explaining why the points $0.75sweet + 0.25bitter$, $0.3sweet + 0.7bitter$ are chosen as extremal. The following pictorial definition is maybe worth one hundred words:

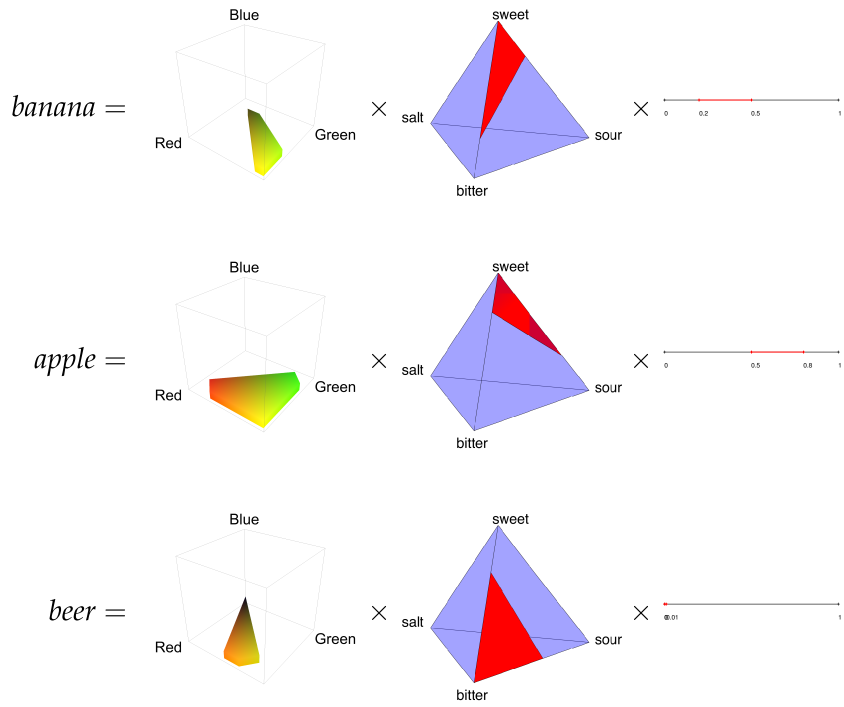


Figure 8: Pictorial description of *banana*, *apple*, *beer*.

What about the sentence space? We want it to model events related with drinking and eating. We give a discrete and very simple example, distinguishing only from positive or negative and surprising or unsurprising events. Our sentence space is then just $\mathbf{B} \times \mathbf{B}$, where $\mathbf{B}$ is the *boolean semilattice* $\{0, 1\}$, and the convex algebra structure is the canonical one on products of semilattices as in Example 3.0.6, that is in this case obtained taking the maximum element-wise. The elements of our sentence space S are then

$$\{0,0\} = \textit{negative, unsurprising} \qquad \{0,1\} = \textit{negative, surprising}$$
$$\{1,0\} = \textit{positive, unsurprising} \qquad \{1,1\} = \textit{positive, surprising}$$

5.1.2 *Adjectives*

In the pregroup grammar we set adjectives to have type $n \cdot n^l$. Since n, n^l both correspond to N in **ConvexRel**, we infer the that here the adjective type is $N \otimes N$. Using the name/coname construction we see that adjectives are convex relations $N \to N$, and hence subsets of $N \times N$, that is, convex combinations of ordered pairs.

In the following we will denote adjectives using the subscript $_{adj}$ to avoid ambiguities. We start with $yellow_{adj}$, that is just

$$yellow_{adj} = \{(\vec{x}, \vec{x}) | x_{colour} \in yellow\}$$

since it acts only on the colour domain. This definition is made simple from the fact that $yellow_{adj}$ does not care about the nature of the noun it is acting on. On the other hand, an adjective such as $soft_{adj}$ does: *Soft stone* and *soft sponge ball* will nowhere have the same level of "softness", that has to be adjusted according to the noun we are applying $soft_{adj}$ to. We can avoid problems opting for a definition by cases when possible: Restricting to *apple* and *banana* we set

$$\begin{aligned} soft_{adj} = \{(\vec{x}, \vec{x}) | \vec{x} \in banana \text{ and } x_{texture} \leqslant 0.35 \text{ or} \\ \vec{x} \in apples \text{ and } x_{texture} \leqslant 0.6\}. \end{aligned}$$

The above mentioned issues, namely the difficulty of classifying adjectives set-theoretically, are broadly analyzed in [85]. Going a bit more in depth, nouns are considered as one-place predicates, for instance setting $red = \{x | x \text{ is red}\}$ and $dog = \{x | x \text{ is a dog}\}$. Adjectives can then be classified as *intersective*, where the meaning of *adj noun* is just $adj \cap noun$; as *subsective*, where $adj\ noun \subseteq noun$; and as *privative*, meaning $adj\ noun \not\subseteq noun$.

Intersective adjectives are the simplest to describe: They act as set-theoretic intersections and this is clearly reflected by the way we defined them: *Yellow banana*, for instance, is just the intersection of *yellow* and *banana*, and we can see this as "all the bananas that happen to be yellow". We can exploit the Frobenius structure of Theorem 3.0.10 to describe intersective adjectives, compared with the general case, as the following picture shows:

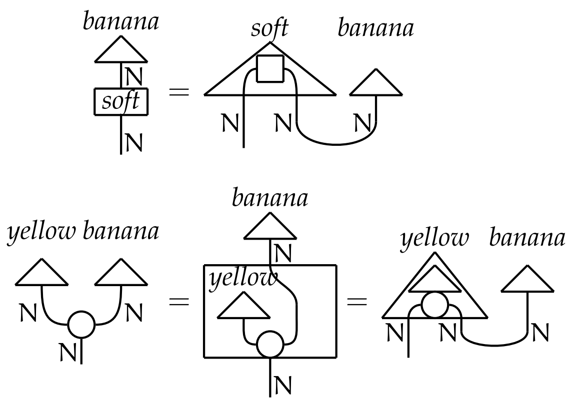


Figure 9: Pictorial description of adjective composition.

From this we infer that the internal structure of an intersective adjective is derived directly from a noun.

5.1.3 *Verbs*

As we saw in Section 2.2.1, we can set transitive verbs to have type $n^r \cdot s \cdot n^l$ in the pregroup formalism, that is sent to the object $N \otimes S \otimes N$ in **ConvexRel**. To define verbs, we use concept names as shorthand, where these can easily be calculated. For example,

$$\begin{aligned} \textit{green banana} = \{(R, G, B) | (R \leqslant G \leqslant 1.5R), (G \geqslant B), (0.3 \leqslant R \leqslant 0.7), \\ , (G \geqslant 0.3), (B \leqslant 0.1)\} \times \\ \times \, \mathrm{Cl}(\{\textit{sweet}, 0.25\textit{sweet} + 0.75\textit{bitter}, 0.7\textit{sweet} + 0.3\textit{sour}\}) \times \\ \times \, [0.2, 0.5] \end{aligned}$$

To fully specify a verb we should take into account all the nouns we could apply it to. This is obviously impractical, hence for expository purposes we restrict our nouns to *banana* and *beer*, exploiting the fact that having different textures they do not overlap. The adjective

$$\textit{taste} : I \to N \otimes S \otimes N$$

Is then defined as follows:

$$\begin{aligned} \textit{taste} = (\textit{green banana} \times \{(0,0)\} \times \textit{bitter}) \cup (\textit{green banana} \times \{(1,1)\} \times \textit{sweet}) \\ \cup (\textit{yellow banana} \times \{(1,0)\} \times \textit{sweet}) \\ \cup (\textit{beer} \times \{(0,1)\} \times \textit{sweet}) \cup (\textit{beer} \times \{(1,0)\} \times \textit{bitter}) \end{aligned}$$

5.1.4 *Example: Robot Navigation*

As we promised, now we model space-related concepts, using *robot navigation* as an example. Choices of noun and sentence spaces will have to be – unsurprisingly – radically different from the ones used in Subsection 5.1.1.

5.1.4.1 *Nouns*

We want to describe:

- Some objects, such as *armchair* and *ball*;
- A couple of robots, that we will call *Cathy* and *David*;
- Some places, such as *kitchen* and *living room*.

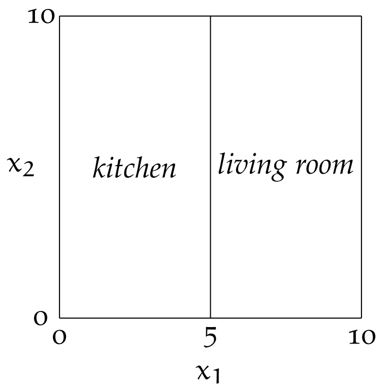
To avoid clutter we will call these nouns a, b, c, d, k, l, respectively. Che choose the noun space N to be decomposed in the following domains:

$$N_{location} \otimes N_{direction} \otimes N_{shape} \otimes N_{size} \otimes N_{colour} \otimes \ldots$$

where $N_{location}$ is just the real affine plane $\mathbb{AR}^2$. First we focus on *kitchen* and *living room*. What we really care about here is where these rooms are. We describe this property using convex subsets of the domain $N_{location}$ as follows:

$$p_{kitchen\ location} = \{(x_1, x_2)|x_1 \in [0,5], x_2 \in [0,10]\}$$
$$p_{living\ room\ location} = \{(x_1, x_2)|x_1 \in [5,10], x_2 \in [0,10]\}$$

$p_{kitchen\ location}, p_{living\ room\ location}$ have the convenient pictorial representation:



From here we can obtain the nouns *kitchen* and *living room* tensoring these properties together with other sets of characteristics defined in the shape, size domain etc. We won't describe these other characteristics in detail since they do not play any role in our examples.

$$kitchen = p_{kitchen\ location} \otimes p_{kitchen\ shape} \otimes p_{kitchen\ size} \otimes \dots$$
$$living\ room = p_{living\ room\ location} \otimes p_{living\ room\ shape} \otimes p_{living\ room\ size} \otimes \dots$$

We can define the other nouns in a similar way, as combinations of convex subsets of domains defining the noun space. We do not give any explicit description since all we require is that we are able to distinguish one noun from the other, and this is always true up to a sufficiently big number of domains taken into consideration.

5.1.4.2 *Verbs*

Before defining verbs we have to choose a suitable sentence space. The sentences we are interested in will have the form:

The ball is in the living room.
Cathy moves to the kitchen.

The characterizing feature of these sentences is that the subject (*the ball, Cathy*) is related to a complement (*living room, kitchen*) via a path through space and time. When the verb is *is in* this path is static, so it can be trivially represented as constant (space here is just one point); on the other hand in the case of *moves to* the path is non trivial, and we will represent it using subsets of the time and space domains. Our sentence space will then be defined as

$$S = N \otimes T \otimes N_{location}$$

Here the agent (that is, the subject of the verb) is represented by a noun in the noun space N, while the path the agent takes is described as a subset of time and location domains.

The time domain T is formalized according to our intuition as the real line $\mathbb{AR}^1$, interpreting 0 as "now", $t < 0$ as the past and $t > 0$ as the future.

Transitive verbs type is again as in Section 5.1.1, namely $N \otimes S \otimes N$. Plugging in our definition of sentence space this becomes:

$$N \otimes N \otimes T \otimes N_{location} \otimes N,$$

And can be thought of as sets of ordered tuples of the form:

$$(n_1, n_2, t, l, n_3)$$

Where n_i are points in the noun space, $t \in T$ and $l \in N_{location}$.

The verbs we will model are *is_in* and *moves_to*. If we want to be fussy, these verbs are not transitive, but intransitive verbs followed by a preposition: We are exploiting the fact that when an intransitive verb and a preposition are parsed together their complessive type is exactly the one defining a transitive verb. To see this, note that prepositions are usually modeled as having type $s^r \cdot s \cdot n^l$, hence the combination "intransitive verb + preposition"reduces as follows:

$$(n^r \cdot s)(s^r \cdot s \cdot n^l) \leqslant n^r \cdot s \cdot n^l$$

Motivating our choice. The verb *is_in* acts on any of the nouns a, b, c, d as subject, and any of the nouns k, l as object. As we said, this verb describes a static path, and accounts only for what is happening in the present (the fact that an agent *is in* some place means that the ageint is in some place now and doesn't tell us anything about where the agent was before or after), and hence we represent it as:

$$is\_in = \{(\vec{n}, \vec{n}, t_{now}, m_{location}, \vec{m}) | \vec{n} \in a \cup b \cup c \cup d, t_{now} = 0, \vec{m} \in k \cup l\} \quad (2)$$

The verb *moves_to*, instead, exhibits non-trivial dynamics. We can use the following observations:

- *moves_to* will take again any of the nouns a, b, c, d as subject, excluding k, l since we don't want to model *kitchen* and *living room* as capable to move;

- *moves_to* will take any of the nouns a, b, c, d, k, l as object, since we want to be able to give meaning to sentences like *David moves to the armchair*;

- By *moves_to*, we mean an action started in the past that is concluded now, at time 0. *Cathy moves to the kitchen* means that, from her position x at some point in the past t, Cathy's position x' is now in $p_{kitchen\ location}$.

- Since we are not interested in modeling quantum phenomena, we assume all the paths to be continuous: Things do not teleport instantaneously from one to another point in space.

We note that, in general, the meaning assigned to a sentence in this model will be a convex combination of paths. In the best case, this convex combination will consist of exactly one path, and this will be the maximum degree of specification we are able to attain. Using the previous considerations we define the verb *moves_to* as follows:

$$\begin{aligned} \textit{moves\_to} = \{(\vec{n}, \vec{n}, [t, 0], f([t, 0]), \vec{m}) \mid \vec{n} \in a \cup b \cup c \cup d, t < 0, \\ , f \text{ continuous}, f(t) \in \vec{n}_{location}, f(0) \in \vec{m}_{location}\} \end{aligned} \tag{3}$$

The constraints given in this definition ensure us that

- The agent cannot be k or l;
- The movement happened in the past;
- The movement desribes a path starting from the location of the agent and ending in the location of the object;
- The movement is continuous.

Note that at the moment the locations of all nouns are subsets of $\mathbb{AR}^2$, and are hence fixed in time. We saw this explicitly when we defined the locations of the nouns *kitchen* and *living room*. On the other hand, we would like the locations of *armchair*, *ball*, *Cathy* and *David* to be dynamic and prone to change over time. To take account of these issues we will probably need to extend our type system to a richer one, that will in fact be closer to that proposed by Gärdenfors in [70]. Definition of richer type structures is object of current research.

5.2 CONCEPTS IN INTERACTION

Now that we have defined enough types and given meaning to enough nouns, adjective and verbs, we want to show how interactions work in our categorical model of conceptual spaces. All we have to do is to apply the type reductions of the pregroup grammar within the conceptual spaces formalism.

5.2.1 *Sentences in the Food Space*

The application of $yellow_{adj}$ to *banana* works as follows.

$$\begin{aligned} yellow\ banana &= (1_N \times \epsilon_N)(yellow_{adj} \times banana) \\ &= (1_N \times \epsilon_N)\{(\vec{x}, \vec{x}) | x_{colour} \in yellow\} \\ &\quad \times (\{(R, G, B)|(0.9R \leqslant G \leqslant 1.5R), (R \geqslant 0.3), (B \leqslant 0.1)\} \\ &\quad\quad \times Cl(\{t_{sweet}, 0.25t_{sweet} + 0.75t_{bitter}, 0.7t_{sweet} + 0.3t_{sour}\}) \\ &\quad\quad\quad \times [0.2, 0.5]) \\ &= \{(R, G, B)|(0.9R \leqslant G \leqslant 1.5R), (R \geqslant 0.7), (G \geqslant 0.7), (B \leqslant 0.1)\} \\ &\quad \times Cl(\{t_{sweet}, 0.25t_{sweet} + 0.75t_{bitter}, 0.7t_{sweet} + 0.3t_{sour}\}) \\ &\quad\quad \times [0.2, 0.5] \end{aligned}$$

You may have noticed that in the last line above the colour propriety has changed. This reflects the intersective nature of the adjective $yellow_{adj}$: Restricting the possible choices of colour of the noun we started with to *yellow*. *Soft apple* is calculated in the same way, giving as result:

$$\begin{aligned} soft\ apple &= \{(R, G, B)|R - 0.7 \leqslant G \leqslant R + 0.7), (G \geqslant 1 - R), (B \leqslant 0.1)\} \\ &\quad \times Cl(\{t_{sweet}, 0.75t_{sweet} + 0.25t_{bitter}, 0.3t_{sweet} + 0.7t_{sour}\}) \times [0.4, 0.6] \end{aligned}$$

We now show interactions involving also verbs. We use the definition of *taste* that we gave in Subsection 5.1.3, finding out that even if sweet bananas are good:

$$\begin{aligned} bananas\ taste\ sweet &= (\epsilon_N \times 1_S \times \epsilon_N)(bananas \times taste \times sweet) \\ &= (\epsilon_N \times 1_S)(banana \\ &\quad \times (green\ banana \times \{(1, 1)\} \cup yellow\ banana \times \{(1, 0)\})) \\ &= \{(1, 1), (1, 0)\} = positive \end{aligned}$$

Sweet beer is, indeed, not[1]:

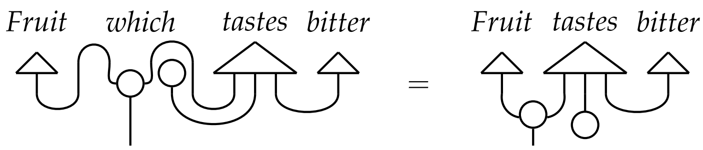
$$\begin{aligned} beer\ tastes\ sweet &= (\epsilon_N \times 1_S \times \epsilon_N)(beer \times taste \times sweet) \\ &= \{(0, 1)\} = negative\ and\ surprising \end{aligned}$$

5.2.2 *Relative Pronouns*

We already mentioned in Chapter 2, Subsection 2.4.1, that the existence of Frobenius algebras allows us to deal with relative pronouns. This is described in detail in [86]. By relative pronouns we mean

1 This is Bob's (my supervisor) idea. I honestly do like sweet beer.

words as "which" and "that". For example, we consider the noun phrase *Fruit which tastes bitter*. Its structure is displayed in Equation 4:



(4)

Figure 10: Pictorial description of pronoun composition.

In our toy model, we can compute that *Fruit which tastes bitter* $=$ *green banana*:

$$
\begin{aligned}
&\textit{Fruit which tastes bitter} = \\
&\quad = (\mu_N \times \iota_S \times \epsilon_N)(\mathrm{Cl}(\textit{bananas} \cup \textit{apples}) \times \textit{taste} \times \textit{bitter}) \\
&\quad = (\mu_N \times \iota_S)(\mathrm{Cl}(\textit{bananas} \cup \textit{apples}) \times (\textit{green banana} \times \{(0,0)\})) \\
&\quad = \mu_N(\mathrm{Cl}(\textit{bananas} \cup \textit{apples}) \times (\textit{green banana})) = \textit{green banana}
\end{aligned}
$$

Where μ_N is the converse of the Frobenius copy map on N and ι_S is the delete map on S from Theorem 3.0.10. This equality clearly comes from the fact that our choice of nouns, adjectives and verbs is indeed limited. In a bigger model we naturally expect the noun phrase *fruit which tastes bitter* to correspond to a plethora of possible "fruit - adjective" concepts.

5.2.3 *Sentences about Robot Movement*

Finally, we compute meanings of sentences involving robot movement. We use nouns and verbs defined in Subsection 5.1.4. We want to compute the meaning of the sentence *Cathy moves to the living room*. In order to do that, we need to assume that Cathy has a defined location.

$$
\begin{aligned}
&\textit{Cathy moves to the living room} \\
&= (\epsilon_N \otimes 1_s \otimes \epsilon_N)(C \otimes \textit{moves\_to} \otimes L) \\
&= (\epsilon_N \otimes 1_s \otimes)(C \otimes \{(\vec{n}, \vec{n}, [t, 0], f([t, 0]))| f(0) \in L_{location}\} \qquad (5) \\
&= \{C, [t, 0], f([t, 0])| f(t) \in C_{location}, f(0) \in L_{location}\}
\end{aligned}
$$

In line (5) further constraints apply to t and f as described in Equation (3), omitted here to avoid clutter. This calculation gives us a set of continuous line segments starting from Cathy's location at time t and ending in the living room at time 0.

It remains to check that this set of line segments is convex to be sure our model is correct. We postulated Cathy to have a location, but this is undetermined, meaning that Cathy can take any possible location while her other attributes remain static, hence the set of possible locations defining Cathy's position is convex. The set $[t, 0]$ with

$t < 0$ is clearly convex: If we take two of such time segments and define a pointwise convex mixture:

$$\begin{aligned} p[t_1, 0] + (1-p)[t_2, 0] &= [pt_1 + (1-p)t_2, p0 + (1-p)0] \\ &= [pt_1 + (1-p)t_2, 0] \end{aligned}$$

The condition that the start point is in the past and the end point is 0 is still satisfied. This can clearly be extended to arbitrary finite convex mixtures by induction.

We have moreover to prove that also the sets of locations $f([t_1, 0])$ is convex. We then consider two sets of locations $f_1([t_1, 0])$ and $f_2([t_2, 0])$ and transform the intervals $[t_1, 0]$ and $[t_2, 0]$ to $[-1, 0]$ by dividing through by $-t_i$, renaming the rescaled functions f'_i. We then form a pointwise convex combination:

$$pf'_1 + (1-p)f'_2 : [0, 1] \to N_{location}$$

By taking:

$$\begin{aligned} (pf'_1 + (1-p)f'_2)(\tau) &= pf'_1(\tau) + (1-p)f'_2(\tau) \\ &= pf_1((-t_1)\tau) + (1-p)f_2((-t_2)\tau) \end{aligned}$$

This convex combination is continuous because of composition of continuous functions. The constraints imposed on f_1, f_2, namely that $f_1(t_1), f_2(t_2) \in C_{location}$ and that $f_1(0), f_2(0) \in L_{location}$, are also preserved: This is because $C_{location}$ and $L_{location}$ are convex by definition, and hence

$$pf_1(t_1) + (1-p)f_2(t_2) \in C_{location}, \qquad pf_1(0) + (1-p)f_2(0) \in L_{location}$$

As we wanted.

In this chapter we have given an idea of how our model may be employed. In the next one we will generalize this formalism in a much broader way: This will allow us to incorporate a metric in it and to model things such as change of context and time-evolution.

6 ALGEBRAIC Q-RELATIONS AND Q-SPANS

Festina lente.

— *Augustus [68, Vita Divi Augusti, 25.4]*

As we mentioned at the end of the previous chapter, we want to redefine our **ConvexRel** in much broader terms to accomodate some features of interest, such as metrics and distances. We obviously want to do this without giving up on the compact closed structure, on one hand because it is the key feature to apply the pregroup grammar reductions to our semantic category, and on the other because it is ultimately what the slick and convenient graphical calculus we use relies on. The last reason motivates us to redefine **ConvexRel** without giving up on the dagger Frobenius structure either, that allows us to model "difficult" words like pronouns, as we have already seen. We start saying that we will be able to accomplish all this, but the construction hereby presented will require quite a lot of explanation. We start from general considerations of geometric nature before we move on with formal definitions.

A binary relation between two sets (that is, what we commonly refer to when we say "relation") can be thought as a function from its domain A to the powerset $\mathcal{P}(B)$ of its codomain B. What it does is to send an element a to the set $R(a)$ of elements in B that are related to A. Equivalently, we can see it as a *characteristic function*

$$A \times B \to \{0, 1\}$$

that sends a couple (a, b) to 1 if a and b are related, to 0 otherwise. The first idea that comes to mind is then to extend this definition considering generalized relations:

$$A \times B \to Q$$

where Q is a set of truth values. Since we want to define a category, it is paramount for us to require that composition of generalized relations – from now on called Q-relations – is still a generalized relation. To do this we need to impose some kind of structure on Q, namely requiring it to be a *quantale*. Quantales have been introduced in [117], and in particular the concept of quantale-valued relation is not new [6].

Definition 6.0.1 (Quantale)**.** A *quantale* is a join complete partial order $(Q, \bigvee)$ with a monoid structure $(Q, \odot, k)$ satisfying for all $a, b \in Q$ and $A, B \subseteq Q$ the following axioms:

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

We say that Q is *commutative* if its monoid structure is commutative.

Before going further we provide some examples of quantales that will be useful later.

Example 6.0.2. A *locale* $(A, \sup, \wedge)$ is a set A with finite meets ($\wedge$) and arbitrary joins (sup) such that meets distribute over joins. There are many equivalent names for these structures, such as *frame* or *complete Heyting algebra*, see [82] for more details. We can endow any locale with a commutative quantale structure setting

$$\bigvee A = \sup A$$
$$a_1 \odot a_2 = \min(a, b)$$
$$k = \top$$

- The chain $\{0, 1\}$ with its usual ordering is a complete distributive lattice, hence a commutative quantale. We will refer to this as the *Boolean quantale* **B**;
- The chain $[0, 1] \subseteq \mathbb{R}$ with its usual ordering is a complete distributive lattice, hence a commutative quantale. We will refer to this as the *interval quantale* **I**;
- The chain $[0, \infty]$ of extended positive reals with the *reverse* ordering is again a complete distributive lattice, hence a commutative quantale. We will refer to this using the notation **F**.

On the set of extended positive reals with reverse order (the same underlying set of **F**) we can define another commutative quantale structure that does not come from a locale, setting:[1]

$$\bigvee A = \inf A$$
$$a_1 \odot a_2 = a_1 + a_2$$
$$k = 0$$

We will call this structure the *Lawvere quantale* **C**. This structure will be of paramount importance in Chapter 7.

1 Here we agree that if a_1 or a_2 or both are ∞, then $a_1 + a_2 = \infty$.

When Q is a quantale we can compose two Q-relations

$$R : A \times B \to Q \qquad S : B \times C \to Q$$

Setting:

$$S \circ R : A \times C \to Q$$
$$(S \circ R)(a, b) := \bigvee_b R(a, b) \odot S(b, c)$$

A Q-relation can be thought of as a potentially infinite matrix of truth values. We will prove later than Q-relations form a category **Rel**(Q), with the composition given above. Moreover, when Q is commutative we will be able to define a symmetric monoidal structure on this category that makes **Rel**(Q) dagger compact closed.

These categories of generalized relations are used to investigate various topological notions, for example in [40, 79]. On the other hand, multi-valued relations have already been used in natural language processing, for instance in [52], and we are now looking to get the best from both approaches.

Summing up these considerations with the ones made in previous chapters, we now end up having two different generalizations of the relational construction:

- The first, defined in Chapter 3, incorporates some algebraic structure in the definition of relation, allowing us to talk about convexity;
- The second one generalizes the truth values. As we will see in Chapter 7, this will allow us to talk about metric and distances.

Clearly, this can't be the end of the story. Further questions naturally arise now that we have two different relational constructions at hand:

- Are these constructions related?
- Are these constructions instances of a more general one, that is, can we incorporate generalized truth values and convexity in the same setting?
- Can we generalize even further, obtaining other instances of compact closed categories?
- How do these parameters interact together? Will the corresponding categories be somehow related?

We will answer to these questions in this chapter. Our starting point will be the definition of *hypergraph category* [63], that captures some nice properties that a compact closed category can have. This definition originates from work on cospans and corelations [62, 63] that is part of a bigger research program in network theory started in [11], and all the constructions we will perform will be instances of hypergraph categories.

Remark 6.0.3. Be aware that every hypergraph category can be produced by means of decorated corelations: Albeit very general, this approach is rather abstract and moves from different premises. What we have in mind here is a construction that is parametrized by conceptual motivations that can be adjusted according to our modelling needs, as we did for **ConvexRel**.

Remark 6.0.4. We will make great use of topos theory, of which standard references are [27, 83, 84, 98]. Nevertheless this use will be most of the time "concealed": The reader with none or limited knowledge of topos theory should be able to read everything just interpreting our constructions as they were defined in **Set**. Note moreover that our relational and span constructions will always be *external*: We will never do internal category theory, and the internal language of a topos will be used just to manipulate diagrams in a convenient way. This will be stressed further later on, where ambiguity may arise. All in all, considering the use we will make of them, the reader could just forget that our categories are toposes: Our metalanguage is strictly classical and toposes, with respect to our application, can be thought as just "categories with a convenient language to manipulate usual diagrams".

We will proceed as follows:

- In Section 6.1 we formally introduce the definition of hypergraph category;
- In Sections 6.2 and 6.3 we will formally introduce algebraic Q-relations and spans, that are the main ingredients for our model. We will prove in Theorems 6.2.9, 6.3.2 and 6.3.7 that these constructions are paramterized instances of hypergraph categories;
- In Section 6.4 we will prove that our categories are order enriched. This key feature will allow us to introduce internal monads and metrics in Chapter 7;
- In Section 6.5 the interaction between algebraic Q-spans and relations will be studied;
- In Section 6.6 we will study how homomorphisms of truth values functorially induce functors between our constructions, preserving everything that matters;
- In Section 6.7 we will do the same with respect to the algebraic structure, providing connections with notions as resource sensitivity in the sense of linear logic;
- In Section 6.8 we will show that our constructions are also functorial in the choice of ambient topos, with the quantale structure being transferred along a logical functor. This amounts to

say that our construction is well behaved with respect to the variation of our universe of discourse;

- In Section 6.9 we will put everything together proving commutativity of all the induced functors, meaning that tweaks along different axes of variation behave well with respect to each other.

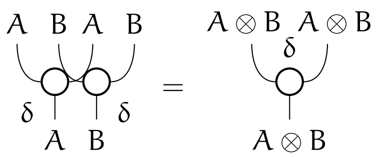
6.0.1 *Work we won't consider*

There is a lot of work on categories of relations out there. For instance, they can be characterized in terms of allegories [67]. This approach won't be considered since it does not behave well with respect to the graphical calculus, that we wanted to preserve. The concept of cartesian bicategory of [32] shares some similarity with hypergraph categories. The study of cartesian bicategories nevertheless aims towards an abstract direction of research, and is more interested in characterization rather than construction of models as we are. Another approach that is worth mentioning is the construction of categories with graphical calculi by means of PROPs [89, 96], that has recently been used to construct various categorical models with applications in control theory [25, 54, 144]. The way these methods work is starting with syntax and equations and building the free category that satisfies them. This approach works well when one has clear in mind what he wants to model and how the phenomena of interest behave compositionally. This is not our case: We instead emphasize the direct construction of models which can then be investigated for their suitability to a given application.

6.1 HYPERGRAPH CATEGORIES

We hit the ground running giving the definition of hypergraph category.

Definition 6.1.1 (Hypergraph Category)**.** A *hypergraph category* is a symmetric monoidal category (recall Definition 2.3.1) such that every object A has a commutative monoid structure and a cocommutative comonoid structure (Definitions 2.4.1 and 2.4.2). The comonoid has to satisfy the following equality, called *coherence condition*:



(6)

While the monoid structure to satisfy the coherence condition dual to the one above. Notice that in the diagram above we used the same

δ everywhere without specifying the object it acts on to avoid clutter. This will be customary in the following when context makes notation overload non-ambiguous.

Moreover, monoid and comonoid have to satisfy the Frobenius and the special axioms (Definition 2.4.3).

The definition above, to be brief, captures the property of having a symmetric monoidal category in which every object has a chosen special commutative Frobenius algebra, such that this choice is coherent with the tensor product.

Many examples of hypergraph categories are things we already know:

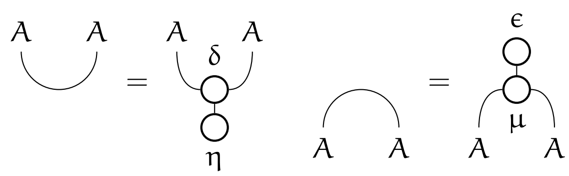
Example 6.1.2. **Rel** is an hypergraph category. The cocommutative comonoid is given by the relations:

$$\epsilon = \{(a, *) \mid a \in A\} \qquad \delta = \{(a, (a, a)) \mid a \in A\}$$

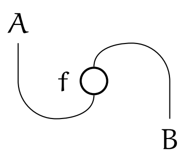
While the monoid is defined through the relational converse applied to the monoid structure.

But wait... Didn't we say that **Rel** was compact closed? It is indeed, and the whole point about hypergraph categories is exactly that they *always* induce a dagger compact closed structure in a rather pleasant way! To get aware of this the following observation is sufficient:

Proposition 6.1.3. *Every hypergraph category is a dagger compact closed category, with the cup and cap given by:*



Given a morphism $f : A \to B$*, its dagger is given by:*



Also called transpose *in the linear algebraic and quantum-mechanical jargon [43].*

6.2 RELATIONS

Now we make the intuition on Q-relations we summarized before precise. As we already said, Q-relations are morphisms between sets of type $A \times B \to Q$, where Q is a quantale representing the truth values our relation can assume. We already mentioned that composition of Q-relations is defined as:

$$(S \circ R)(a, c) = \bigvee_b R(a, b) \odot S(b, c)$$

Notice that if Q is the boolean quantale **B** this notion collapses to the usual one for binary relations on **Set**. We define the identity morphism for every set A as:

$$1_A(a_1, a_2) = \bigvee \{k \mid a_1 = a_2\}$$

This definition of identity is a bit cumbersome, but has one great advantage, that is, it is not defined by cases. Definitions by cases often make hidden use of the law of the excluded middle, and this would block a lot of interesting generalization: Observe that, albeit all these definitions collapse to the usual ones for relations on **Set** up to a wise choice of Q, they are all constructive; there is no hidden use of reductio ad absurdum or tertium non datur, and hence they make sense in the internal language of an arbitrary elementary topos[2]. With this we mean that we can work *semi-internally*, and use the internal language of a topos to characterize our generalized relations. This will allow us to prove a great deal of results in an easy way. Let's start with a definition:

Definition 6.2.1 (Q-relation)**.** Let $\mathcal{E}$ be a topos, and $(Q, \odot, k, \bigvee)$ an internal quantale. A Q-*relation* between $\mathcal{E}$ objects A and B is a $\mathcal{E}$-morphism of type:

$$A \times B \to Q$$

$\mathcal{E}$-objects and Q-relations between them form a category $\mathbf{Rel}(Q)$, with identities and composition as described above.

The next important step is to incorporate the algebraic structure in Definition 6.2.1. In this way we will obtain an environment in which the convex structure of conceptual spaces and the metric induced by the generalized truth values (yet to be defined) will coexist. First we formalize what we mean by algebraic structure.

Definition 6.2.2. An *algebraic signature* is a couple (Σ, E) were Σ is a set of functional symbols $\sigma_i : X^{n_i} \to X$ of finite arity n_i, called *operations*, over a formal set of variables X, and E is a set of statements between symbols of Σ in equational logic, that is, all the variables in every statement are non-free and universally quantified over X.

2 Often we will just say "topos" instead of "elementary topos". In this document no result will be proved using further properties that other kind of toposes, such as Grothendieck toposes, satisfy.

Definition 6.2.3. A set A can be endowed with an *algebraic structure* of type (Σ, E) if (Σ, E) can be *interpreted*[3] in A, meaning that for every function symbol $\sigma \in \Sigma$ there is a function $\sigma^A : A^n \to A$ such that n is equal to the arity of σ and the equations in E still hold for all elements in A substituting each σ with the correspondent σ^A. In this case we say that $\langle A, \{\sigma_i^A\}_{i:\sigma_i \in \Sigma}\rangle$ is an algebra of type (Σ, E).

Note that our definitions of signature and algebra make sense in any elementary topos[4]. It is clear then that we can again work *semi-inernally*, and manipulate our algebras (that are, indeed, external) using the same language we would use if we were to work in **Set**.

In the set-theoretic case we know that a relation respects the structure between algebras of the same signature if:

$$R(a_1, b_1) \wedge \cdots \wedge R(a_n, b_n) \Rightarrow R(\sigma(a_1, \dots, a_n), \sigma(b_1, \dots, b_n))$$

For every operation $\sigma \in \Sigma$ (regarding convex combinations as algebraic operations as we did in Chapter 4 it is easy to see that the definition of convex relations we used to formalize **ConvexRel** is just a particular instance of this more general one).

Clearly, since we are now working on using arbitrary quantales as sets of truth values, we want to replace the logical symbols popping up in the implication above with quantale operations. This leads us to the following definition:

Definition 6.2.4 (Algebraic Q-relation)**.** Let $\mathcal{E}$ be a topos, and let $(Q, \odot, k, \bigvee)$ an internal quantale. Let (Σ, E) be an algebraic variety in $\mathcal{E}$. An *algebraic* Q-relation between (Σ, E)-algebras A and B is a Q-relation between their underlying $\mathcal{E}$-objects such that for each $\sigma \in \Sigma$ the following axiom holds:

$$R(a_1, b_1) \odot \ldots \odot R(a_n, b_n) \leqslant R(\sigma(a_1, \dots, a_n), \sigma(b_1, \dots, b_n)) \qquad (7)$$

(Σ, E)-algebras and algebraic Q-relations form a category $\mathbf{Rel}_{(\Sigma,E)}(Q)$, with identities and composition as for their underlying Q-relations.

In Section 6.7 this algebraic structure will be studied in detail with particular focus on how it interacts with other choice parameters such as the truth values. For the moment we keep going with the definition and study the category we are interested in. Incorporating Definition 6.2.4 in our category we get the first, promising result:

3 Note that a set A can be endowed with many different algebraic structures, also of the same signature. For instance, if $|A| = 4$, the signature of groups can be interpreted in A in two different ways (obtaining the cyclic group of order 4 and the Klein 4 group, respectively.

4 Actually, our definition of algebra makes sense in any category with finite products: The interpretation of an algebraic signature is then given assigning morphisms $A^n \to A$ for every functional symbol of arity n and the equations are expressed as commutative diagrams. If our category is moreover an elementary topos, this definition coincides with the one given in terms of the internal language.

Proposition 6.2.5. *Let $\mathcal{E}$ be a topos, (Σ, E) a variety in $\mathcal{E}$, and $(Q, \odot, k, \bigvee)$ an internal commutative quantale. The category $\textbf{Rel}_{(\Sigma,E)}(Q)$ is a symmetric monoidal category. The symmetric monoidal structure is inherited from a choice of binary products and terminal object in $\mathcal{E}$.*

Proof. The proof that $\textbf{Rel}_{(\Sigma,E)}(Q)$ is a category follows from the unit and associative properties of quantale multiplication. We must first confirm that identities respect algebraic structure. For $\sigma \in \Sigma$:

$$\begin{aligned}
&\mathrm{id}_A(a_1, a_1') \odot \ldots \odot \mathrm{id}_A(a_n, a_n') = \\
&\qquad = \left[\bigvee\{k \mid a_1 = a_1'\}\right] \odot \ldots \odot \left[\bigvee\{k \mid a_n = a_n'\}\right] \\
&\qquad = \bigvee\{\underbrace{k \odot \ldots \odot k}_{n} \mid (a_1 = a_1') \wedge \cdots \wedge (a_n = a_n')\} \\
&\qquad \leqslant \bigvee\{k \mid \sigma(a_1, \ldots, a_n) = \sigma(a_1', \ldots, a_n')\} \\
&\qquad = \mathrm{id}_A(\sigma(a_1, \ldots, a_n), \sigma(a_1', \ldots, a_n'))
\end{aligned}$$

We do the same for composition:

$$\begin{aligned}
&(S \circ R)(a_1, c_1) \odot \ldots \odot (S \circ R)(a_n, c_n) = \\
&= \left[\bigvee_{b_1} R(a_1, b_1) \odot S(b_1, c_1)\right] \odot \ldots \\
&\qquad\qquad \ldots \odot \left[\bigvee_{b_n} R(a_n, b_n) \odot S(b_n, c_n)\right] \\
&= \bigvee_{b_1, \ldots, b_n} R(a_1, b_1) \odot S(b_1, c_1) \odot \ldots \odot R(a_n, b_n) \odot S(b_n, c_n) \\
&= \bigvee_{b_1, \ldots, b_n} [R(a_1, b_1) \odot \ldots \odot R(a_n, b_n)] \odot \\
&\qquad\qquad \odot [S(b_1, c_1) \odot \ldots \odot S(b_n, c_n)] \\
&\leqslant \bigvee_{b_1, \ldots, b_n} R(\sigma(a_1, \ldots, a_n), \sigma(b_1, \ldots, b_n)) \odot \\
&\qquad\qquad \odot S(\sigma(b_1, \ldots, b_n), \sigma(c_1, \ldots, c_n)) \\
&\leqslant \bigvee_{b} R(\sigma(a_1, \ldots, a_n), b) \odot S(b, \sigma(c_1, \ldots, c_n)) \\
&= (S \circ R)(\sigma(a_1, \ldots, a_n), \sigma(c_1, \ldots, c_n))
\end{aligned}$$

Now we check explicitly the associativity of composition. For relations R, S, T, we have

$$\begin{aligned}
(T \circ (S \circ R))(a, d) &= \bigvee_{c} (S \circ R)(a, c) \odot T(c, d) \\
&= \bigvee_{c} \left[\bigvee_{b} R(a, b) \odot S(b, c)\right] \odot T(c, d) \\
&= \bigvee_{b} R(a, b) \odot \left[\bigvee_{c} S(b, c) \odot T(c, d)\right]
\end{aligned}$$

$$
\begin{aligned}
&= \bigvee_b R(a,b) \odot (T \circ S)(b,d) \\
&= ((T \circ S) \circ R)(a,d)
\end{aligned}
$$

We also check the right identity law.

$$
\begin{aligned}
(R \circ \mathrm{id}_A)(a,b) &= \bigvee_{a'} \mathrm{id}_A(a,a') \odot R(a',b) \\
&= \bigvee_{a'} \left[\bigvee \{k \mid a = a'\}\right] \odot R(a',b) \\
&= \bigvee_{a'} \bigvee \{k \odot R(a',b) \mid a = a'\} \\
&= \bigvee_{a'} \{R(a,b)\} \\
&= R(a,b)
\end{aligned}
$$

The left identity law proof is similar. Next, we prove this category is symmetric monoidal. We define the monoidal unit to be the terminal algebra. On objects, the tensor is just products of algebras. We define the action on morphisms pointwise as:

$$
(R \otimes R')(a,a',b,b') = R(a,b) \odot R'(a',b')
$$

We first confirm $R \otimes R'$ respects the algebraic structure. For $\sigma \in \Sigma$ with arity n:

$$
\begin{aligned}
&(R \otimes R')(a_1,a_1',b_1,b_1') \odot \ldots \odot (R \otimes R')(a_n,a_n',b_n,b_n') = \\
&\quad = R(a_1,b_1) \odot R'(a_1',b_1') \odot \ldots \odot R(a_n,b_n) \odot R'(a_n',b_n') \\
&\quad = [R(a_1,b_1) \odot \ldots \odot R(a_n,b_n)] \odot \left[R'(a_1',b_1') \odot \ldots \odot R'(a_n',b_n')\right] \\
&\quad \leqslant R(\sigma(a_1,\ldots,a_n),\sigma(b_1,\ldots,b_n)) \odot R'(\sigma(a_1',\ldots,a_n'),\sigma(b_1',\ldots,b_n')) \\
&\quad = (R \otimes R')((\sigma(a_1,\ldots,a_n),\sigma(a_1',\ldots,a_n')), \\
&\qquad\qquad\qquad ,(\sigma(b_1,\ldots,b_n),\sigma(b_1',\ldots,b_n'))) \\
&\quad = (R \otimes R')(\sigma((a_1,a_1'),\ldots,(a_n,a_n')),\sigma((b_1,b_1'),\ldots,(b_n,b_n')))
\end{aligned}
$$

Then, we show that the tensor is functorial. Identities are preserved:

$$
\begin{aligned}
(1_{A_1} \otimes 1_{A_2})(a_1,a_2,a_1',a_2') &= 1_{A_1}(a_1,a_1') \odot 1_{A_2}(a_2,a_2') \\
&= \bigvee \{k \mid a_1 = a_1'\} \odot \bigvee \{k \mid a_2 = a_2'\} \\
&= \bigvee \{k \mid (a_1 = a_1') \wedge (a_2 = a_2')\} \\
&= \bigvee \{k \mid (a_1,a_2) = (a_1',a_2')\} \\
&= 1_{A_1 \otimes A_2}(a_1,a_2,a_1',a_2')
\end{aligned}
$$

For composition,

$$
\begin{aligned}
&[(S_1 \otimes S_2) \circ (R_1 \otimes R_2)](a_1, a_2, c_1, c_2) = \\
&\quad = \bigvee_{b_1, b_2} (R_1 \otimes R_2)(a_1, a_2, b_1, b_2) \odot (S_1 \otimes S_2)(b_1, b_2, c_1, c_2) \\
&\quad = \bigvee_{b_1, b_2} R_1(a_1, b_1) \odot R_2(a_2, b_2) \odot S_1(b_1, c_1) \odot S_2(b_2, c_2) \\
&\quad = \bigvee_{b_1, b_2} R_1(a_1, b_1) \odot S_1(b_1, c_1) \odot R_2(a_2, b_2) \odot S_2(b_2, c_2) \\
&\quad = \left[\bigvee_{b_1} R_1(a_1, b_1) \odot S_1(b_1, c_1)\right] \odot \left[\bigvee_{b_2} R_2(a_2, b_2) \odot S_2(b_2, c_2)\right] \\
&\quad = (S_1 \circ R_1)(a_1, c_1) \odot (S_2 \circ R_2)(a_2, c_2) \\
&\quad = [(S_1 \circ R_1) \otimes (S_2 \circ R_2)]\,(a_1, a_2, c_1, c_2)
\end{aligned}
$$

We consider $\mathcal{E}$ as a symmetric monoidal category with respect to our choice of binary products and terminal object. We then take the graphs (see Proposition 6.2.6 for the definition) of the corresponding left and right unitors, associator and symmetry as the corresponding structure in $\mathbf{Rel}(Q)$. We must confirm that these coherence morphisms are natural in their parameters. The proofs are all similar, we check the associator explicitly:

$$
\begin{aligned}
&R \otimes (S \otimes T) \circ \alpha_{A,B,C} = \bigvee_{x,y,z} \alpha_{A,B,C}(((a,b),c),(x,(y,z))) \odot \\
&\qquad\qquad \odot [R \otimes (S \otimes T)]\,((x,(y,z)),(a',(b',c'))) \\
&= \bigvee_{x,y,z} \left[\bigvee \{k \mid (a = x) \wedge (b = y) \wedge (c = z)\}\right] \odot \\
&\qquad\qquad \odot R(x,a') \odot S(y,b') \odot T(z,c') \\
&= \bigvee_{x,y,z} \bigvee \{R(x,a') \odot S(y,b') \odot T(z,c') \mid (a = x) \wedge (b = y) \wedge (c = z)\} \\
&= R(a,a') \odot S(b,b') \odot T(c,c') \\
&= \bigvee_{x,y,z} \bigvee \{R(a,x) \odot S(b,y) \odot T(c,z) \mid (x = a') \wedge (y = b') \wedge (z = c')\} \\
&= \bigvee_{x,y,z} R(a,x) \odot S(b,y) \odot T(c,z) \odot \left[\bigvee \{k \mid x = a' \wedge y = b' \wedge z = c'\}\right] \\
&= \bigvee_{x,y,z} [(R \odot S) \odot T]\,(((a,b),c),((x,y),z)) \otimes \\
&\qquad\qquad \otimes \alpha_{A',B',C'}(((x,y),z),(a',(b',c'))) \\
&= \alpha_{A',B',C'} \circ (R \otimes S) \otimes T
\end{aligned}
$$

These morphisms are isomorphisms by functoriality of graphs (see Proposition 6.2.6 for the proof). Their inverses are given by their converses, as established in the proof of Proposition 6.2.6.

Moreover, taking graphs commutes with our choice of products in $\mathcal{E}$ in the sense that:

$$(f \times g)_\circ = f_\circ \otimes g_\circ$$

To check this we reason as follows:

$$\begin{aligned}(f \times g)_\circ((a, a'), (b, b')) &= \bigvee\{k \mid (b, b') = (f \times g)(a, a')\} \\ &= \bigvee\{k \mid (b = f(a)) \wedge (b' = g(a'))\} \\ &= \left[\bigvee\{k \mid b = f(a)\}\right] \odot \left[\bigvee\{k \mid b' = g(a')\}\right] \\ &= f_\circ(a, b) \otimes g_\circ(a', b')\end{aligned}$$

This guarantees that the triangle and pentagon equations hold as the same equations hold for the cartesian monoidal structure in $\mathcal{E}$. The coherence conditions for symmetry follow similarly. □

As in the case of **ConvexRel** note that even if our tensor product is the cartesian product on objects, relations of type $A \otimes B \to A' \otimes B'$ cannot be separated into binary products $R \times R'$. Hence our definition gives us a genuine tensor product.

We clearly want to put a hypergraph structure on this newly defined symmetric monoidal category. To do this, we will need a couple of useful tools that will allow us to lift relevant stuff from the underlying topos $\mathcal{E}$. These are again generalizations of concepts from the standard **Set** case. One is taking the *converse* of a relation: Relations are, opposed to functions, symmetric in their arguments, and so we can always reverse their arguments. This can be promptly generalized to $\mathbf{Rel}_{(\Sigma,E)}(Q)$:

Proposition 6.2.6. *[Converse] Let $\mathcal{E}$ be a topos, (Σ, E) a variety in $\mathcal{E}$, and $(Q, \odot, k, \bigvee)$ an internal commutative quantale. There is an identity on objects strict symmetric monoidal functor:*

$$(-)^\circ : \boldsymbol{Rel}_{(\Sigma,E)}(Q)^{op} \to \boldsymbol{Rel}_{(\Sigma,E)}(Q)$$

Given by reversing arguments:

$$R^\circ(b, a) = R(a, b)$$

Moreover, in the standard case we know that functions are just a special kind of relation, and in fact we can lift any function $f : A \to B$ on **Set** to a morphism in **Rel** defining:

$$R_f := \{(a, b) \mid f(a) = b\}$$

This lifting is clearly functorial and can again be generalized to the category $\mathbf{Rel}_{(\Sigma,E)}(Q)$ as follows:

Proposition 6.2.7. *[Graph] Let $\mathcal{E}$ be a topos, (Σ, E) a variety in $\mathcal{E}$, and $(Q, \odot, k, \bigvee)$ an internal commutative quantale. There is an identity on objects strict symmetric monoidal functor:*

$$(-)_\circ : \boldsymbol{Alg}(\Sigma, E) \to \boldsymbol{Rel}_{(\Sigma,E)}(Q)$$

With action defined on morphism $f : A \to B$ by:

$$f_\circ(a, b) = \bigvee\{k \mid f(a) = b\}$$

The symmetric monoidal structure on $\boldsymbol{Alg}(\Sigma, E)$ is the finite product structure.

Now that we have the tools to lift morphisms from $\mathcal{E}$ to $\mathbf{Rel}_{(\Sigma,E)}(Q)$, we need to find something that is worth lifting. On every category with finite products (as $\mathcal{E}$ surely is being a topos) the diagonal defines a canonical comonoid structure as follows:

Proposition 6.2.8. *Let $\mathcal{E}$ be a category with finite products. Each object A carries a cocommutative comonoid structure via the canonical morphisms:*

$$! : A \to 1 \quad \textit{and} \quad \langle 1_A, 1_A \rangle : A \to A \times A$$

These morphisms satisfy the coherence condition (6) in Definition 6.1.1.

This is the perfect candidate to prove $\mathbf{Rel}_{(\Sigma,E)}(Q)$ is hypergraph, and in fact:

Theorem 6.2.9. *Let $\mathcal{E}$ be a topos, (Σ, E) a variety in $\mathcal{E}$, and $(Q, \odot, k, \bigvee)$ an internal commutative quantale. The category $\boldsymbol{Rel}_{(\Sigma,E)}(Q)$ is a hypergraph category. The cocommutative comonoid structure is given by the graphs of the canonical comonoids described in Proposition 6.2.8, and the monoid structure is given by their converses.*

Proof. For every object A of $\mathcal{E}$, call ϵ_A, δ_A the comultiplication and counit of Proposition 6.2.8, and η_A, μ_A their respective converses. The morphisms ϵ_A, δ_A have in the internal logic the explicit form:

$$\epsilon_A(a, x) = k \qquad \delta_A(a_1, (a_2, a_3)) = \bigvee\{k \mid a_1 = a_2 = a_3\}$$

Checking that η_A, μ_A form a monoid is straightforward from the definition of converse. With respect to this monoid/comonoid pair, we first confirm the special axiom:

$$\begin{aligned}
(\mu_A \circ \delta_A)(a_1,a_2) &= \bigvee_{(a,a')} \delta_A(a_1, (a, a')) \odot \mu_A((a, a'), a_2) \\
&= \bigvee_{(a,a')} \left[\bigvee\{k \mid a_1 = a = a'\}\right] \odot \left[\bigvee\{k \mid a = a' = a_2\}\right] \\
&= \bigvee_{(a,a')} \bigvee\{k \mid a_1 = a = a' = a_2\} \\
&= \bigvee\{k \mid a_1 = a_2\} \\
&= id_A(a_1, a_2)
\end{aligned}$$

Finally, we check the Frobenius axiom, omitting some stages where the expressions get long:

$$
\begin{aligned}
&((\mathrm{id}_A \otimes \delta_A) \circ (\mu_A \otimes \mathrm{id}_A))(a_1, a_2, a_3, a_4) = \\
&\quad = \bigvee_{x,y,z} \mathrm{id}_A(a_1, x) \odot \delta_A(a_2, (y, z)) \odot \mu_A((x, y), a_3) \odot \mathrm{id}_A(z, a_4) \\
&\quad = \bigvee_{x,y,z} \{k \mid x = y = z = a_1 = a_2 = a_3 = a_4\} \\
&\quad = \bigvee \{k \mid a_1 = a_2 = a_3 = a_4\} \\
&\quad = ((\delta_A \otimes \mathrm{id}_A) \circ (\mathrm{id}_A \otimes \mu_A))(a_1, a_2, a_3, a_4) \qquad \square
\end{aligned}
$$

Example 6.2.10. Let $(\emptyset, \emptyset)$ denote the signature with no operations or equations, to which we will sometimes refer as the *empty signature*. Algebras on the empty signature are just sets, and hence $\mathbf{Rel}_{(\emptyset,\emptyset)}(Q)$ is just $\mathbf{Rel}(Q)$. In particular, we have that **Rel**, $\mathbf{Rel}(\mathbf{B})$ and $\mathbf{Rel}_{(\emptyset,\emptyset)}(\mathbf{B})$ denote the same thing.

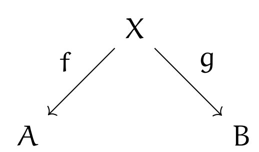
Example 6.2.11. We proved in Proposition 4.1.4 that every convex algebra can be algebraically presented as a set with infinite binary operations:

$$+^p \text{ where } p \in (0, 1)$$

satisfying some equations. These algebras form a variety equivalent to **Convex** as we proved in Theorem 4.1.8, and calling Convex the signature of this variety we can recast **ConvexRel** as $\mathbf{Rel}_{\mathrm{Convex}}(\mathbf{B})$.

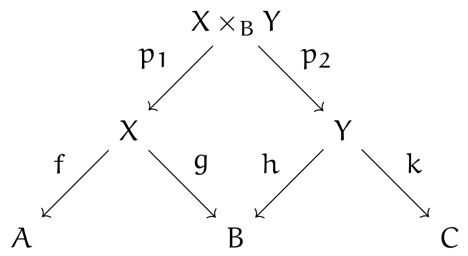
6.3 SPANS

$\mathbf{Rel}_{\mathrm{Convex}}(\mathbf{B})$ allows us to customize our relation choosing the logical universe we want to operate in (the underlying topos), the structure of truth values (the quantale) and the algebraic structure on our objects. In this setting, though, we can only say if two elements are related and with which strength. Elements can hence be related in one possible way, and we can introduce another, last degree of freedom generalizing this. Recall that a span of sets:



Represents the idea that elements $a \in A$ and $b \in B$ are related if $f(x) = a$ and $g(x) = b$. Clearly in this setting we can do more: If there is another x' such that $f(x') = a$ and $g(x') = b$, then a and b are related via the different elements x, x'. Elements of X then can be interpreted as *proof witnesses* of the relation between a and b. In a very similar fashion to what happens in computer science, where we choose to distinguish between different algorithms that do the same thing (and hence talk about complexity) or we don't (and hence talk about computability), here we are switching from a notion of "being related" to a notion of "being related in a specific way".

It is well known that isomorphism classes[5] of spans between sets form a category, where composition is defined by pullback:



In the category **Set** we have an explicit definition of pullback, given by

$$X \times_B Y = \{(x, y) \mid g(x) = h(y)\}$$

This confirms our intuition about spans as proof-relevant relations: Witnesses relating a, c are obtained considering couples of witnesses each relating a, b and b, c, respectively.

We would again try to leverage the similarities between toposes and sets to generalize the span construction to an arbitrary topos, but we have to pay attention to what we do: If we internalize the definition of span altogether, the fact that we are considering isomorphism classes of spans would put us in big trouble: Working with isomorphism classes of spans is only possible if our topos satisfies the axiom of choice and, as proven by Diaconescu [51], this automatically entails the law of the excluded middle, implying that we should restrict our working framework to classical toposes. Fortunately, we can avoid this issue by working again semi-internally: Our spans are not internalized in the internal logic of the topos: They are just ordinary, decorated spans and this allows us to freely consider isomorphism classes of spans while we work. Then, we can exploit the fact that each representative of an isomorphism class of spans can be described in pretty convenient terms using the internal language: For instance, the definition of pullback we recalled above works for an arbitrary topos, and this will make our proofs much easier. Note that to do this the proofs of independency from a particular representative in a given isomorphism class have to be carried out externally. All in all, to avoid making the understanding of the following results difficult, the reader that is fluent in topos theory should be aware that *we are using the internal logic to prove external facts*.

5 We need to consider isomorphism classes because pullbacks are associative only up to isomorphism, and hence we need to quotient by the isomorphism-induced equivalence relation on morphisms to obtain associativity with equality as the category theory axioms require.

We want to mimic what we did generalizing relations, and to do this we have to put some truth values on top of the span structure. Notice that now we don't need anymore a quantale structure for the truth values, indeed we can do with a much weaker one: We defined the composition of relations as:

$$(S \circ R)(a, c) = \bigvee_b R(a, b) \odot S(b, c)$$

Because we needed the quantale join exactly to "compress" all the possible $R(a, b) \odot S(b, c)$ we could obtain for different b: If we care only about being related or not, then we want only *one* truth value representing this relation, but which one should we pick? The quantale join $\bigvee$ solved the situation giving us a structural way to perform this picking operation. In the span case, though, this is not necessary since we *can* have different weights for different proof witnesses. Definition 6.2.1 can then be relaxed requiring Q to be just an internal monoid.

Definition 6.3.1 (Q-span)**.** Let $\mathcal{E}$ be a finitely complete category. Let moreover be $(Q, \odot, k)$ an internal monoid. A Q-*span* of type $A \to B$ is a quadruple (X, f, g, χ) where:

- $(X, f : X \to A, g : X \to B)$ is a span in $\mathcal{E}$;
- $\chi : X \to Q$ is a $\mathcal{E}$-morphism, referred to as the *characteristic morphism*.

We can compose Q-spans $(X, f, g, \chi), (Y, h, k, \xi)$ by composing their underlying spans by pullback, and taking the resulting characteristic morphism to be:

$$X \times_B Y \xrightarrow{\langle p_1, p_2 \rangle} X \times Y \xrightarrow{\chi \times \xi} Q \times Q \xrightarrow{\odot} Q$$

Where p_1 and p_2 are the pullback projections.

Generalizing the usual span case, we can introduce a *morphism of* Q-*spans* between two Q-spans of type $A \to B$

$$\alpha : (X_1, f_1, g_1, \chi_1) \to (X_2, f_2, g_2, \chi_2)$$

As a $\mathcal{E}$-morphism $\alpha : X_1 \to X_2$ such that:

$$f_1 = f_2 \circ \alpha \qquad g_1 = g_2 \circ \alpha \qquad \chi_1 = \chi_2 \circ \alpha$$

As we already mentioned before, we will work with isomorphism classes of spans. Everything is proven to be (see appendix, Section A) independent from the choice of representatives of these classes, and hence we can just ignore this fact in the notation to avoid clutter. This is common practice when considering categories of ordinary spans, and hence from now on by Q-span we will mean its corresponding isomorphism class. For every object A we can provide an identity Q-span $(A, 1, 1, \chi_k)$, where χ_k is the *constant morphism* defined by:

$$\chi_k = A \xrightarrow{!} 1 \xrightarrow{k} Q$$

With this notion of composition and identity, Q-spans form a category. If our internal monoid is moreover commutative, we get the following important result:

Theorem 6.3.2. *Let $\mathcal{E}$ be a finitely complete category, and $(Q, \odot, k)$ an internal commutative monoid. The category **Span**(Q) is a hypergraph category.*

We will not focus much on Theorem 6.3.2, that serves more as a proof of concept. In fact, all the details about the hypergraph structure for **Span**(Q) can be found as special cases in Proposition 6.3.4 and Theorem 6.3.7. What we really want to do is again to incorporate the algebraic structure in our definition. In this case, though, we need to keep track of the proof witnesses, slightly complicating our definition. We will, moreover, need to strengthen a bit the requirement that Q is a commutative monoid: We still do not need a quantale structure as in the relational case, but we need to introduce a partial ordering on the monoid structure to be able to compare different kind of weights.

Definition 6.3.3. Let $\mathcal{E}$ be a topos, (Σ, E) a variety in $\mathcal{E}$, and $(Q, \odot, k, \leqslant)$ an internal partially ordered commutative monoid. For (Σ, E)-algebras A and B, an *algebraic* Q-span is a quadruple (X, f, g, χ) which is a Q-span between the underlying $\mathcal{E}$-objects satisfying the following axiom. For every $\sigma \in \Sigma$ if:

$$f(x_1) = a_1 \wedge g(x_1) = b_1 \wedge \cdots \wedge f(x_n) = a_n \wedge g(x_n) = b_n$$

Then there exists x such that:

$$f(x) = \sigma(a_1, \ldots, a_n) \wedge g(x) = \sigma(b_1, \ldots, b_n)$$

And:

$$\chi(x_1) \odot \ldots \odot \chi(x_n) \leqslant \chi(x) \tag{8}$$

(Σ, E)-algebras and algebraic Q-spans form a category $\mathbf{Span}_{(\Sigma,E)}(Q)$ with identities and composition given as for the underlying Q-spans.

In this setting results similar to those for Q-relations proved in Section 6.2 hold. We have a symmetric monoidal category, a generalization of the graph and converse functors and, ultimately, using again the canonical comonoids of Proposition 6.2.8, an hypergraph structure.

Proposition 6.3.4. *Let $\mathcal{E}$ be a topos, (Σ, E) a variety in $\mathcal{E}$, and $(Q, \odot, k, \leqslant)$ an internal partially ordered commutative monoid. The category $\mathbf{Span}_{(\Sigma,E)}(Q)$ is then a symmetric monoidal category. The symmetric monoidal structure is inherited from the finite product structure in $\mathcal{E}$.*

Proof. Firstly, we check that $\mathbf{Span}_{(\Sigma,E)}(Q)$ is a category. We must confirm that the identity morphisms are algebraic. The required condition is immediate, as objects are closed under the algebraic operations, and the characteristic morphism is constant in the internal language.

Now we must confirm that composition is independent of representatives. Consider span isomorphisms:

$$\varphi : (X_1, f_1, g_1, \chi_1) \to (X_2, f_2, g_2, \chi_2)$$
$$\psi : (Y_1, h_1, k_1, \xi_1) \to (Y_2, h_2, k_2, \xi_2)$$

We must show an isomorphism between the Q-spans:

$$(X_1 \times_B Y_1, f_1 \circ p_1, k_1 \circ p_2, \odot \circ (\chi_1 \times \xi_1) \circ \langle p_1, p_2 \rangle)$$
$$(X_2 \times_B y_2, f_2 \circ p_1', k_2 \circ p_2', \odot \circ (\chi_2 \times \xi_2) \circ \langle p_1', p_2' \rangle)$$

We calculate, using properties of pullbacks:

$$f_2 \circ p_1' \circ \varphi \times_B \psi = f_2 \circ \varphi \circ p_1 = f_1 \circ p_1$$
$$g_2 \circ p_2' \circ \varphi \times_B \psi = g_2 \circ \psi \circ p_2 = g_1 \circ p_2$$

Also,

$$\begin{aligned} \odot \circ (\chi_2 \times \xi_2) \circ \langle p_1', p_2' \rangle \circ \varphi \times_B \psi &= \odot \circ (\chi_2 \times \xi_2) \circ (\varphi \times \psi) \circ \langle p_1, p_2 \rangle \\ &= \odot \circ (\chi_2 \circ \varphi, \xi_2 \circ \psi) \circ \langle p_1, p_2 \rangle \\ &= \odot \circ (\chi_1 \times \psi_1) \circ \langle p_1, p_2 \rangle \end{aligned}$$

And as $\varphi \times_B \psi$ has an inverse in $\mathcal{E}$ we can use Lemma A.0.1 to complete this part of the proof. Next, we must confirm that composition of algebraic Q-spans preserves the algebraic condition. Assume:

$$f(p_1(z_1)) = a_1 \wedge k(p_2(z_1)) = c_1 \wedge \dots$$
$$\dots \wedge f(p_1(z_n)) = a_n \wedge k(p_2(z_n)) = c_n$$

Then there exist $x_1, \dots, x_n$ and $y_1, \dots, y_n$ such that:

$$f(x_1) = a_1 \wedge g(x_1) = h(y_1) \wedge k(y_1) = c_1 \wedge \dots$$
$$\dots \wedge f(x_n) = a_n \wedge g(x_n) = h(y_n) \wedge k(y_n) = c_n$$

Therefore as the component spans are algebraic, there exist x and y such that:

$$f(x) = \sigma(a_1, \dots, a_n) \wedge g(x) = h(y) \wedge k(y) = \sigma(c_1, \dots, c_n)$$

And both:

$$\chi(x_1) \odot \dots \odot \chi(x_n) \leqslant \chi(x) \qquad \xi(y_1) \odot \dots \odot \xi(y_n) \leqslant \xi(y)$$

Therefore we have (x, y) in the apex of the composite span, with truth value $\chi(x) \odot \xi(y)$. By monotonicity of the monoid multiplication:

$$\chi(x_1) \odot \dots \odot \chi(x_n) \odot \xi(y_1) \odot \dots \odot \xi(y_n) \leqslant \chi(x) \odot \xi(y)$$

Finally, as the multiplication is commutative:

$$\chi(x_1) \odot \xi(y_1) \odot \ldots \odot \chi(x_n) \odot \xi(y_n) \leqslant \chi(x) \odot \xi(y)$$

Now we are ready to verify the axioms defining a category, namely identity laws and associativity of composition; we start with the left identity axiom. Firstly we note:

$$(B, id_B, id_B, \chi_k) \circ (X, f, g, \chi) = (X \times_B B, f \circ p_1, p_2, \odot \circ (\chi \times \chi_k) \circ \langle p_1, p_2 \rangle)$$

We claim p_1 is a Q-span morphism to the span (X, f, g, χ). The conditions for this being a span morphism are $f \circ p_1 = f \circ p_1$ and $g \circ p_1 = p_2$, and the second condition is obvious from the pullback square. Finally, we must confirm this also commutes with the characteristic functions.

$$\begin{aligned} \odot \circ (\chi \times \chi_k) \circ \langle p_1, p_2 \rangle &= p_1 \circ (\chi \times !) \circ \langle p_1, p_2 \rangle \\ &= p_1 \circ \langle \chi \circ p_1, ! \circ p_2 \rangle \\ &= \chi \circ p_1 \end{aligned}$$

We must prove that p_1 is an isomorphism, the required inverse being given by the universal property of pullbacks as $\langle 1, g \rangle$. Checking that this is an isomorphism follows from the universal property of pullbacks. We can then use Lemma A.0.1 to complete this part of the proof. The right identity axiom follows similarly.

We must then confirm associativity. We consider the composites:

$$\begin{aligned} L &= ((Z, l, m, \zeta) \circ (Y, h, k, \xi)) \circ (X, f, g, \chi) \\ R &= (Z, l, m, \zeta) \circ ((Y, h, k, \xi) \circ (X, f, g, \chi)) \end{aligned}$$

Via the usual proof for categories of ordinary spans,

$$\iota := \langle p_1 \circ p_1, \langle p_2 \circ p_1, p_2 \rangle \rangle : (X \times_B Y) \times_C Z \to X \times_B (Y \times_C Z)$$

Is an isomorphism of spans. It remains to show that this commutes with the characteristic morphisms. This is a space consuming exercise in tracking various canonical morphisms, and is most easily handled using the graphical calculus for a cartesian monoidal category. Details are omitted to avoid a long typesetting exercise for the diagrams. This concludes the proof that $\mathbf{Span}_{(\Sigma,E)}(Q)$ is a category.

It remains to prove that $\mathbf{Span}_{(\Sigma,E)}(Q)$ is symmetric monoidal. We define a functor $\otimes : \mathbf{Span}_{(\Sigma,E)}(Q) \times \mathbf{Span}_{(\Sigma,E)}(Q) \to \mathbf{Span}_{(\Sigma,E)}(Q)$ as follows:

$$\begin{aligned} A \otimes B &= A \times B \\ (X_1, f_1, g_1, \chi_1) \otimes (X_2, f_2, g_2, \chi_2) &= \\ &= (X_1 \times X_2, f_1 \times f_2, g_1 \times g_2, \odot \circ (\chi_1 \times \chi_2)) \end{aligned}$$

This functor respects the algebraic structure: for algebraic Q-spans (X, f, g, χ) and (X', f', g', χ'), if

$$(f \times f')(x_1, x_1') = (a_1, a_1') \wedge \cdots \wedge (f \times f')(x_n, x_n') = (a_n, a_n')$$

And

$$(g \times g')(x_1, x_1') = (b_1, b_1') \wedge \cdots \wedge (g \times g')(x_n, x_n') = (b_n, b_n')$$

Then:

$$f(x_1) = a_1 \wedge f'(x_1') = a_1' \wedge \cdots \wedge f(x_n) = a_n \wedge f'(x_n') = a_n'$$

And:

$$g(x_1) = b_1 \wedge g'(x_n') = b_1' \wedge \cdots \wedge g(x_n) = b_n \wedge g'(x_n') = b_n'$$

As the spans are algebraic, there exist x and x' such that:

$$\chi(x_1) \odot \ldots \odot \chi(x_n) \leqslant \chi(x) \qquad \chi'(x_1') \odot \ldots \odot \chi'(x_n') \leqslant \chi'(x')$$

By monotonicity and commutativity of the monoid multiplication, we then have:

$$\chi(x_1) \odot \chi'(x_1') \odot \ldots \odot \chi(x_n) \odot \chi'(x_n') \leqslant \chi(x) \odot \chi'(x')$$

Next step is to show that our functor respects equivalence classes of spans. Assume we have span isomorphisms:

$$\begin{aligned} \varphi &: (X_1, f_1, g_1, \chi_1) \to (X_1', f_1', g_1', \chi_1') \\ \psi &: (X_2, f_2, g_2, \chi_2) \to (X_2', f_2', g_2', \chi_2') \end{aligned}$$

The product $\varphi \times \psi$ gives an isomorphism of ordinary spans:

$$\begin{aligned} \varphi \times \psi : (X_1, f_1, g_1, \chi_1) \otimes (X_2, f_2, g_2, \chi_2) \to \\ \to (X_1', f_1', g_1', \chi_1') \otimes (X_2', f_2', g_2', \chi_2') \end{aligned}$$

It then remains to check this commutes with characteristic morphisms. We calculate:

$$\odot \circ (\chi_1' \times \chi_2') \circ (\varphi \times \psi) = \odot \circ \left[(\chi_1' \circ \varphi) \times (\chi_2' \circ \psi)\right] = \odot \circ (\chi_1 \times \chi_2)$$

And also this part is done. We now proceed to check functoriality. That our definition gives a functor as an operation on the underlying spans is well known. It remains to check the behaviour with respect to characteristic morphisms. For identity Q-spans, the resulting characteristic function is $\odot \circ (\chi_k \times \chi_k) = \chi_k$. For composition, we note there is an isomorphism of spans:

$$\begin{aligned} \langle\langle p_1 \circ p_1, p_1 \circ p_2\rangle, \langle p_2 \circ p_1, p_2 \circ p_2\rangle\rangle : (X \times_B Y) \times (X' \times_{B'} Y') \to \\ \to (X \times X') \times_{B \times B'} (Y \times Y') \end{aligned}$$

We must show that this commutes with the corresponding characteristic morphisms. The following unpleasant calculation establishes the required equality:

$$
\begin{aligned}
\odot \circ (\odot \times \odot) \circ \big[((\chi \times \xi) \circ \langle p_1, p_2\rangle) \times ((\chi' \times \xi')\langle p_1, p_2\rangle)\big] = \\
= \odot \circ (\odot \times \odot) \circ \big[(\chi \times \xi) \times (\chi' \times \xi')\big] \circ \\
\circ \langle\langle p_1 \circ p_1, p_2 \circ p_1\rangle, \langle p_1 \circ p_2, p_2 \circ p_2\rangle\rangle \\
= \odot \circ (\odot \times \odot) \circ \big[(\chi \times \chi') \times (\xi \times \xi')\big] \circ \\
\circ \langle\langle p_1 \circ p_1, p_1 \circ p_2\rangle, \langle p_2 \circ p_1, p_2 \circ p_2\rangle\rangle \circ \\
\circ \langle\langle p_1 \circ p_1, p_2 \circ p_1\rangle, \langle p_1 \circ p_2, p_2 \circ p_2\rangle\rangle \\
= \odot \circ (\odot \times \odot) \circ \big[(\chi \times \chi') \times (\xi \times \xi')\big] \circ \\
\circ \langle\langle p_1 \circ p_1, p_1 \circ p_2\rangle, \langle p_2 \circ p_1, p_2 \circ p_2\rangle\rangle
\end{aligned}
$$

Now we focus on proving that this functor gives us a monoidal structure: We take as our monoidal unit the terminal object in $\mathcal{E}$ and the tensor to be the functor we just defined. Since we want to proceed as in the relational case, lifting associators and unitors from the underlying topos $\mathcal{E}$ using the converse and graph functors defined in Proposition 6.3.5, we have to check that these functors commute with our tensor. For converse this is trivial. For graph, this amounts to check commutativity of the diagram:

$$
\begin{array}{ccc}
\mathbf{Span}_{(\Sigma,E)}(Q) \times \mathbf{Span}_{(\Sigma,E)}(Q) & \xrightarrow{\ \otimes\ } & \mathbf{Span}_{(\Sigma,E)}(Q) \\
{\scriptstyle (-)_\circ \times (-)_\circ} \uparrow & & \uparrow {\scriptstyle (-)_\circ} \\
\mathcal{E} \times \mathcal{E} & \xrightarrow[\ \times\ ]{} & \mathcal{E}
\end{array}
$$

On objects this is obvious, as all the functors involved act as the identity on objects. On morphisms we reason:

$$
\begin{aligned}
(f_1 \times f_2)_\circ &= (A_1 \times A_2, id_{A_1 \times A_2}, f_1 \times f_2, k) \\
&= (A_1, id_{A_1}, f, k) \otimes (A_2, id_{A_2}, f, k) \\
&= (f_1)_\circ \otimes (f_2)_\circ
\end{aligned}
$$

Therefore the graphs of the coherence morphisms in $\mathcal{E}$ lift to the category $\mathbf{Span}_{(\Sigma,E)}(Q)$. We must confirm that each of these remains natural. To do this, we first prove that the following equation holds for every couple of $\mathcal{E}$-morphisms l, h:

$$
l_\circ \circ (X, f, g, \chi) \circ {}_\circ h = (X, h \circ f, l \circ g, \chi)
$$

Where ${}_\circ h$ is an identity on objects functor[6] defined as:

$$\begin{aligned} {}_\circ(-) &: \mathcal{E}^{op} \to \mathbf{Span}_{(\Sigma,E)}(Q) \\ f &: A \to B \mapsto (A, f, id_A, \chi_k) \end{aligned}$$

We firstly consider the case of post composition with the graph of a morphism in the underlying category, that is, $l_\circ \circ (X, f, g, \chi)$. This composite is given by the pullback span:

$$(X \times_B B, p_1 \circ f, p_2 \circ l, \odot \circ (\chi \times \chi_k) \circ \langle p_1, p_2 \rangle)$$

We note that $p_1 \circ \langle id_X, g \rangle = id_X$ and

$$\langle id_X, g \rangle \circ p_1 = \langle p_1, g \circ p_1 \rangle = \langle p_1, p_2 \rangle = id_{X \times_B B}$$

And so p_1 and $\langle id_X, g \rangle$ witness an isomorphism in $\mathcal{E}$. We next confirm they give a span isomorphism. One of the conditions for p_1 to be a span morphism is trivial, for the other:

$$l \circ g \circ p_1 = l \circ id_B \circ p_2 = l \circ p_2$$

Finally, we must confirm that this commutes with the characteristic morphisms.

$$\odot \circ (\chi \times \chi_k) \circ \langle p_1, p_2 \rangle = \chi \circ p_1 \circ \langle p_1, p_2 \rangle = \chi \circ p_1$$

Now we note that:

$$\begin{aligned} (X, f, g, \chi) \circ {}_\circ h &= (X, f, g, \chi) \circ h_\circ^\circ \\ &= (h_\circ \circ (X, f, g, \chi)^\circ)^\circ \\ &= (h_\circ \circ (X, g, f, \chi))^\circ \\ &= (X, g, h \circ f, \chi)^\circ \\ &= (X, h \circ f, g, \chi) \end{aligned}$$

Combining these two observations then completes the proof that our equation holds. Back to naturality, applying what we just proved it is sufficient to show that the following are span isomorphisms:

$$\begin{aligned} &\lambda_X : (1 \times X, \lambda_A \circ (id_1 \times f), \lambda_B \circ (id_1 \times g), \odot \circ (\chi_k \times \chi)) \to (X, f, g, \chi) \\ &\rho_X : (X \times 1, \rho_A \circ (f \times id_1), \rho_B \circ (g \times id_1), \odot \circ (\chi \times \chi_k)) \to (X, f, g, \chi) \\ &\alpha_{X,Y,Z} : ((X_1 \times X_2) \times X_3, \alpha_{A_1,A_2,A_3}((f_1 \times f_2) \times f_3), \\ &\qquad , \alpha_{B_1,B_2,B_3} \circ ((g_1 \times g_2) \times g_3), \dots) \to (X_1 \times (X_2 \times X_3)) \\ &\sigma_{X,Y} : (X_1 \times X_2, \sigma_{A_1,A_2} \circ f_1 \times f_2, \sigma_{B_1,B_2} \circ g_1 \times g_2, , \odot \circ (\chi_1 \times \chi_2)) \to \\ &\qquad \to (X_2 \times X_1, f_2 \times f_2, g_2 \times g_1, \odot \circ (\chi_2 \times \chi_1)) \end{aligned}$$

And this is now just a straightforward (but very unpleasant) check. □

6 The proof of ${}_\circ(-)$ being a functor is completely analogous to the proof of $(-)_\circ$ being a functor found in Proposition 6.3.5. It is moreover easy to check that if Q is commutative then ${}_\circ(-) = (-)^\circ \circ (-)_\circ$.

Proposition 6.3.5. *[Converse] Let $\mathcal{E}$ be a topos, (Σ, E) a variety in $\mathcal{E}$, and $(Q, \odot, k, \leqslant)$ an internal partially ordered commutative monoid. There is an identity on objects strict symmetric monoidal functor:*

$$(-)^\circ : \boldsymbol{Span}_{(\Sigma,E)}(Q)^{op} \to \boldsymbol{Span}_{(\Sigma,E)}(Q)$$

Given by reversing the legs of the underlying span:

$$(X, f, g, \chi)^\circ = (X, g, f, \chi)$$

Proposition 6.3.6. *[Graph] Let $\mathcal{E}$ be a topos, and $(Q, \odot, k, \leqslant)$ an internal partially ordered commutative monoid. There is an identity on objects strict symmetric monoidal functor:*

$$(-)_\circ : \boldsymbol{Alg}(\Sigma, E) \to \boldsymbol{Span}_{(\Sigma,E)}(Q)$$

With the action on morphism $f : A \to B$ given by:

$$f_\circ = (A, id_A, f, \chi_k)$$

Theorem 6.3.7. *Let $\mathcal{E}$ be a topos, (Σ, E) a variety in $\mathcal{E}$, and $(Q, \odot, k, \leqslant)$ an internal partially ordered commutative monoid. The category $\boldsymbol{Span}_{(\Sigma,E)}(Q)$ is a hypergraph category. The cocommutative comonoid structure is given by the graphs of the canonical comonoids described in Proposition 6.2.8, and the monoid structure is given by their converses.*

In the following chapters we will see that spans are much easier to use than relations. Not only they are more "powerful", allowing us to choose proof witnesses, but more than anything else the requirement downgrading from commutative quantale to commutative partially ordered monoid makes things much easier when it comes to really exploit the structure of the topos we are working in.

Example 6.3.8. The span construction allows us to construct variations on the models we are already interested in. For example, using the same observations made in Example 6.2.11, we can construct a proof relevant version of **ConvexRel**. From a practical perspective, this presents the possibility of models in which we can describe the interaction of cognitive phenomena, providing quantitative evidence for any relationships that we conclude hold.

Example 6.3.9. Interesting instances of elementary toposes different than **Set** are presheaf toposes $\mathbf{Set}^{\mathcal{C}^{op}}$ for a small category $\mathcal{C}$. Presheaves can be thought of as "sets varying with time" and offer a natural choice to model relations and spans that are time or context sensitive. This idea will be explored in detail in the next chapter and constituted the main motivation to generalize our formalism to arbitrary logical universes.

6.4 ORDER ENRICHMENT

In this section we will lay the basics that will allow us to talk in detail about internal monads and distances in the next chapter. To talk about order enrichment we first have to specify what an enriched category is.

Intuitively, we say that a $\mathcal{C}$ is *enriched* over $\mathcal{V}$ (or that $\mathcal{C}$ is a $\mathcal{V}$-category) if, for every A, B objects of $\mathcal{C}$, $\mathrm{Hom}_{\mathcal{C}}[A, B]$ is an object of $\mathcal{V}$. Note that now since our homsets are not sets anymore (objects of $\mathcal{V}$ may not have an analogous of elements), we need to specify how composition and identity laws work in this setting. This is done requiring $\mathcal{V}$ to be monoidal, and specifying for each A object of $\mathcal{C}$ some $\mathcal{V}$-morphism $\mathrm{id}_A : I \to \mathrm{Hom}_{\mathcal{C}}[A, A]$ (here I is the monoidal unit of $\mathcal{V}$) and, for each A, B, C objects of $\mathcal{C}$, some morphism $c_{A,B,C} : \mathrm{Hom}_{\mathcal{C}}[B, C] \otimes \mathrm{Hom}_{\mathcal{C}}[A, B] \to \mathrm{Hom}_{\mathcal{C}}[A, C]$. id_A intuitively "picks" the identity morphism in $\mathrm{Hom}_{\mathcal{C}}[A, A]$, while $c_{A,B,C}$ intuitively represents morphism composition. Clearly the $\mathrm{id}_{(-)}$ and $c_{(-),(-),(-)}$ as defined above have to satisfy some conditions that mimic the usual unit and associativity laws for standard categories.

Connected to the notion of enriched category, there is the one of enriched functor. In fact, if $\mathcal{C}$ and $\mathcal{D}$ are $\mathcal{V}$-categories, having just a generic functor $F : \mathcal{C} \to \mathcal{D}$ doesn't say much about what happens to the enriched structure. An *enriched functor* $F : \mathcal{C} \to \mathcal{D}$ is then specified by the following data: To each object A of $\mathcal{C}$ we assign an object FA of $\mathcal{D}$, as in the "traditional" case. On morphisms, the definition is different: We require that for each $\mathrm{Hom}_{\mathcal{C}}[A, B]$ there is a $\mathcal{V}$-morphism $\mathrm{Hom}_{\mathcal{C}}[A, B] \to \mathrm{Hom}_{\mathcal{D}}[FA, FB]$ that respects conditions analogous to identity and composition preservation. If you want to know more about enriched categories and functors, see [87].

Example 6.4.1. $\mathbf{fdVect}_{\mathbb{R}}$ is enriched over itself: The linear applications between real vector spaces A and B form themselves a real vector space (if m, n are the dimensions of A, B respectively, $\mathrm{Hom}_{\mathbf{fdVect}_{\mathbb{R}}}[A, B]$ has dimension $m \times n$). Compatibility of the enrichment with composition in the base category is ensured by the usual laws of composition for linear applications.

Example 6.4.2. More generally, **fdHilb** is enriched over itself for the same reasons outlined in the previous example.

Example 6.4.3. **Rel** is enriched over the category **Pos**, since every couple of relations R, R' from A to B can be ordered pointwise and this implies that $\mathrm{Hom}_{\mathbf{Rel}}[A, B]$ forms a partial order compatible with relational composition.

Example 6.4.4. More in general, any symmetric monoidal closed category is canonically enriched over itself.

What we want to do now is enriching our categories of relations and spans over ordered sets. Order enriching is not new when it comes to models for language. It proved to be a relevant tool also in the study of ambiguity [121, 122] and lexical entailment [17], whose investigations bring awesome results when tackled from an order-theoretic perspective.

Example 6.4.3 is particularly relevant for us. It states that the category we started from is enriched over posets, and hence the question that spontaneously arises is: Does this hold also for algebraic Q-relations and spans? The answer is obviously a resounding *YES!*, otherwise we wouldn't be talking about this right now.

To prove that a category is enriched over posets (or on preorders) it is sufficient to prove that each homset is a poset (preorder, respectively) and that this poset (preorder) is compatible with composition (that is, composition must be monotone in both components). We honor our (recently established, for sure) tradition assessing the relational case first.

Definition 6.4.5. Let $\mathcal{E}$ be a topos and $(Q, \odot, k, \bigvee)$ an internal quantale. We define a partial order on Q-relations as follows:

$$R \subseteq R' \quad \text{iff} \quad \forall a, b. R(a, b) \leqslant R'(a, b)$$

Algebraic Q-relations are ordered similarly, by comparing their underlying Q-relations.

Theorem 6.4.6. *Let $\mathcal{E}$ be a topos, (Σ, E) a variety in $\mathcal{E}$, and $(Q, \odot, k, \bigvee)$ an internal commutative quantale. The category $\textbf{Rel}_{(\Sigma,E)}(Q)$ is a partially ordered (**Pos**-enriched) symmetric monoidal category.*

Proof. First of all we have to prove that our partial order is well defined. It is clearly reflexive. For transitivity, if $R \subseteq R' \quad \text{and} \quad R' \subseteq R''$, then we infer:

$$\vdash R(a, b) \leqslant R'(a, b) \quad \text{and} \quad \vdash R'(a, b) \leqslant R''(a, b)$$

Therefore internally:

$$\vdash (R(a, b) \leqslant R'(a, b)) \wedge (R'(a, b) \leqslant R''(a, b))$$

And so by transitivity of the order on the quantale:

$$\vdash R(a, b) \leqslant R''(a, b)$$

Finally, if $R \subseteq R'$ and $R' \subseteq R$, and so internally

$$\vdash (R(a, b) \leqslant R'(a, b)) \wedge (R'(a, b) \leqslant R(a, b))$$

By antisymmetry of the order on the internal quantale:

$$\vdash R(a, b) = R'(a, b)$$

And so:

$$\vdash \forall a, b. R(a, b) = R'(a, b)$$

Meaning externally $R = R'$.

Next, we must confirm that composition is monotone in both components. As the proofs are symmetrical, we only consider precomposition explicitly. Assume $R \subseteq R'$. We consider post-composing each of these relations with the relation S. Remembering that the quantale product preserves order, we calculate:

$$\begin{aligned}(S \circ R)(a, c) &= \bigvee_b R(a, b) \odot S(b, c) \\ &\leqslant \bigvee_b R'(a, b) \odot S(b, c) \\ &= (S \circ R')(a, c)\end{aligned}$$

Finally, we must confirm that the tensor on $\mathbf{Rel}_{(\Sigma,E)}(Q)$ is monotone in both arguments. Assume again $R \subseteq R'$. We calculate:

$$\begin{aligned}(R \otimes S)(a, b, c, d) &= R(a, b) \odot S(c, d) \\ &\leqslant R'(a, b) \odot S(c, d) \\ &= (R' \otimes S)(a, b, c, d)\end{aligned}$$ □

As usual in the span case proof witnesses must be accounted for explicitly when casting our definitions, even if what we do is conceptually analogous to the relational one.

Definition 6.4.7. Let $\mathcal{E}$ be a topos and $(Q, \odot, k, \leqslant)$ and internal partially ordered monoid, we define a preorder on Q-spans as follows:

$$(X_1, f_1, g_1, \chi_1) \subseteq (X_2, f_2, g_2, \chi_2)$$

If there is a $\mathcal{E}$-monomorphism $m : X_1 \to X_2$ such that:

$$f_1 = f_2 \circ m \qquad g_1 = g_2 \circ m \qquad \forall x. \chi_1(x) \leqslant \chi_2(m(x))$$

Algebraic Q-spans are ordered similarly, by comparing their underlying Q-spans.

Theorem 6.4.8. *Let $\mathcal{E}$ be a topos, (Σ, E) a variety in $\mathcal{E}$, and $(Q, \odot, k, \leqslant)$ an internal partially ordered commutative monoid. The category $\boldsymbol{Span}_{(\Sigma,E)}(Q)$ is a preordered (**Preord**-enriched) symmetric monoidal category.*

Proof. Firstly, we must confirm that this ordering is independent of choices of representatives for equivalence classes of spans.

Assume $(X_1, f_1, g_1, \chi_1) \subseteq (Y, h_1, k_1, \xi_1)$, and span isomorphisms:

$$\begin{aligned}i &: (X_1, f_1, g_1, \chi_1) \to (X_2, f_2, g_2, \chi_2) \\ j &: (Y_1, h_1, k_1, \xi_1) \to (Y_2, h_2, k_2, \xi_2)\end{aligned}$$

Let $m : (X_1, f_1, g_1, \chi_1) \to (Y, h_1, k_1, \xi_1)$ be the span morphism that is monic in $\mathcal{E}$ required by the assumed order structure. There is then a span morphism:

$$j \circ m \circ i^{-1} : (X_2, f_2, g_2, \chi_2) \to (Y_2, h_2, k_2, \xi_2)$$

Which is monic in $\mathcal{E}$ as monomorphisms are closed under composition. We then have:

$$\chi_2(x) = \chi_1(i^{-1}(x)) = \xi_1(m \circ i^{-1}(x)) = \xi_2(j \circ m \circ i^{-1}(x))$$

The relation $\subseteq$ is clearly reflexive via the identity Q-span morphism.

For transitivity, assume $(X, f, g, \chi) \subseteq (Y, h, k, \xi) \subseteq (Z, m, n, \zeta)$. Denote the required monomorphisms as $r : (X, f, g, \chi) \to (Y, h, k, \xi)$ and $s : (Y, h, k, \xi) \to (Z, m, n, \zeta)$, respectively. There is then an obvious span morphism $s \circ r$ that is a monomorphism in $\mathcal{E}$. We then have $\chi(x) \leqslant \xi(r(x))$ and so $\xi(r(x)) \leqslant \zeta(s \circ r(x))$. By transitivity of the quantale ordering the conclusion $\chi(x) \leqslant \zeta(s \circ r(x))$ follows. Then, we must confirm that composition is monotone in both components. As the proofs are symmetrical, we only consider precomposition explicitly.

Assume $(X_1, f_1, g_1, \chi_1) \subseteq (X_2, f_2, g_2, \chi_2)$ as witnessed by some $\mathcal{E}$ monomorphism $m : X_1 \to X_2$. We consider post-composing each of these spans with the span (Y, h, k, ξ). There is then a Q-span morphism:

$$m \times_B id_Y : \{(x_1, y) \mid f_1(x_1) = h(y)\} \to \{(x_2, y) \mid f_2(x_2) = h(y)\}$$

And the underlying morphism is a monomorphism in $\mathcal{E}$ by standard properties of pullbacks and monomorphisms. By monotonicity of the tensor:

$$\begin{aligned}
(\odot \circ (\chi_1 \times \xi) \circ \langle p_1, p_2 \rangle)(x, y) &= \chi_1(x) \odot \xi(y) \\
&\leqslant \chi_2(m(x)) \odot \xi(y) \\
&= (\odot \circ (\chi_2 \times \xi) \circ \langle p_1, p_2 \rangle)(m(x), y) \\
&= (\odot \circ (\chi_2 \times \xi) \circ \langle p_1, p_2 \rangle \circ (m \times_B id_Y))(x, y)
\end{aligned}$$

Finally, we must confirm that the tensor functor on $\mathbf{Span}_{(\Sigma,E)}(Q)$ is monotone in both arguments. Assume $(X_1, f_1, g_1, \chi_1) \subseteq (X_2, f_2, g_2, \chi_2)$ witnessed by $\mathcal{E}$ monomorphism $m : X_1 \to X_2$. There is then a span morphism:

$$m \times 1 : (X_1, f_1, g_1, \chi_1) \otimes (Y, h, k, \xi) \to (X_2, f_2, g_2, \chi_2) \otimes (Y, h, k, \xi)$$

And this is an $\mathcal{E}$ monomorphism by standard theory of products and monomorphisms. We then calculate:

$$\begin{aligned}
\odot \circ (\chi_1 \times \xi)(x, y) &= \chi_1(x) \odot \xi(y) \\
&\leqslant \chi_2(m(x)) \odot \xi(y) \\
&= (\odot \circ (\chi_2 \times \xi) \circ (m \times 1))(x, y) \qquad \square
\end{aligned}$$

We can say that a functor between categories $\mathcal{C}, \mathcal{C}'$ enriched over $\mathcal{D}$ is a *$\mathcal{D}$-functor* if it respects the structure of 2-cells. The general definition of this fact involves quite a bit of machinery, but in our case this just translates to the requirement that if $h \subseteq k$ for $h, k \in \mathrm{Hom}_{\mathcal{C}}[A, B]$, then $Fh \subseteq Fk$. We can use this notion to show that the converse functor preserves our orderings:

Proposition 6.4.9. *Let $\mathcal{E}$ be a topos, and (Σ, E) a variety in $\mathcal{E}$. Converses respect order structure, in that:*

- *If $(Q, \odot, k, \bigvee)$ is an internal quantale, the converse functor of Proposition 6.2.6 is a partially ordered functor;*
- *If $(Q, \odot, k, \leqslant)$ is an internal partially order monoid, the converse functor of Proposition 6.3.5 is a preordered functor.*

6.5 FROM SPANS TO RELATIONS

Now that we laid all our parameters into place, we can start studying the relationships between them. We start with the simplest one, relating proof relevance and provability.

Intuitively, the morphisms of $\mathbf{Span}_{(\Sigma,E)}(Q)$ distinguish between different ways to relate the same couple of elements. It would be reasonable to expect that a span can then be collapsed into a relation "compressing" all the weights assigned to proof witnesses of the relationship between a fixed couple of elements into only one. This "compressing" operation draws a perfect similarity with the role that the quantale join has in the definition of composition in $\mathbf{Rel}_{(\Sigma,E)}(Q)$: It skims through all the possible paths connecting two elements and picks the optimal one.

What we can infer from this is that to be able to collapse an algebraic Q-span into a Q-relation Q has to be again a quantale. This being the case, we can effectively prove algebraic Q-spans and Q-relations to be compatible. Even better, this compatibility preserves graphs and converses, de facto carrying the hypergraph structure of $\mathbf{Span}_{(\Sigma,E)}(Q)$ to the hypergraph structure of $\mathbf{Rel}_{(\Sigma,E)}(Q)$.

Theorem 6.5.1. *Let $\mathcal{E}$ be a topos, (Σ, E) a variety in $\mathcal{E}$ and $(Q, \odot, k, \bigvee)$ an internal commutative quantale. There is an identity on objects, strict symmetric monoidal **Preord**-functor with action on morphisms:*

$$V : \mathbf{Span}_{(\Sigma,E)}(Q) \to \mathbf{Rel}_{(\Sigma,E)}(Q)$$
$$V(X, f, g, \chi)(a, b) = \bigvee \{\chi(x) \mid f(x) = a \wedge g(x) = b\}$$

V *commutes with graphs and converses in that:*

$$\begin{array}{ccc} \textbf{\textit{Span}}_{(\Sigma,E)}(Q) & \xrightarrow{V} & \textbf{\textit{Rel}}_{(\Sigma,E)}(Q) \\ {\scriptstyle (-)_\circ} \nwarrow & & \nearrow {\scriptstyle (-)_\circ} \\ & \textbf{\textit{Alg}}(\Sigma, E) & \end{array}$$

$$\begin{array}{ccc} \textbf{\textit{Span}}_{(\Sigma,E)}(Q)^{op} & \xrightarrow{V^{op}} & \textbf{\textit{Rel}}_{(\Sigma,E)}(Q)^{op} \\ {\scriptstyle (-)^\circ}\downarrow & & \downarrow{\scriptstyle (-)^\circ} \\ \textbf{\textit{Span}}_{(\Sigma,E)}(Q) & \xrightarrow[V]{} & \textbf{\textit{Rel}}_{(\Sigma,E)}(Q) \end{array}$$

Proof. It is easy to check that this definition is independent of our choice of representatives. We moreover check preservation of identities:

$$\begin{aligned} V(A, id_A, id_A, \chi_k)(a_1, a_2) &= \bigvee\{\chi_k(a) \mid id_A(a) = a_1 \wedge id_A(a) = a_2\} \\ &= \bigvee\{k \mid a = a_1 \wedge a = a_2\} \\ &= \bigvee\{k \mid a_1 = a_2\} \\ &= id_A(a_1, a_2) \end{aligned}$$

That the functor commutes with the tensor is clear from the definition. That the coherence morphisms for the monoidal structures are preserved on the nose is clear as they were constructed using the graph constructions, and V preserves graphs. To see that V preserves preorders, we assume $(X_1, f_1, g_1, \chi_1) \subseteq (X_2, f_2, g_2, \chi_2)$, witnessed by a monomorphism $m : X_1 \to X_2$. We then calculate:

$$\begin{aligned} &\bigvee\{\chi_1(x) \mid f_1(x) = a \wedge g_1(x) = b\} \leqslant \\ &\qquad \leqslant \bigvee\{\chi_2(m(x)) \mid f_2(m(x)) = a \wedge g_2(m(x)) = b\} \\ &\qquad \leqslant \bigvee\{\chi_2(x') \mid f_2(x') = a \wedge g_2(x') = b\} \end{aligned}$$

Lastly, we check that V commutes with graph and converse. For graph, if f is a morphism $A \to B$ in $\mathcal{E}$, then $f^\circ = (A, id_A, f, \chi_k)$ and:

$$V(f_\circ)(a, b) = V(A, id_A, f, \chi_k)(a, b) = \bigvee\{\chi_k(x) \mid x = a \wedge f(x) = b\}$$

but, being $\chi_k(x) = k$ for every x, we can rewrite the right-hand-side as $\bigvee\{k \mid f(a) = b\}$, that is exactly the definition of graph in the category of Q-algebraic relations, and we are done. For converse,

$$\begin{aligned} V((X, f, g, \chi)^\circ)(b, a) &= V(X, g, f, \chi)(b, a) \\ &= \bigvee\{\chi(x) \mid g(x) = b \wedge f(x) = a\} \\ &= \bigvee\{\chi(x) \mid f(x) = a \wedge g(x) = b\} \\ &= V(X, f, g, \chi)(a, b) \\ &= V(X, f, g, \chi)^\circ(b, a) \end{aligned}$$

□

6.6 CHANGING TRUTH VALUES

We would expect that up to a wise choice of morphisms between our structures of truth values functorial relationships between models are established. This all goes through very smoothly, as we now elaborate.

Clearly it is natural to ask the morphisms between structures of truth values to preserve this structures. In the case of algebraic Q-relations, this amounts to consider internal quantale homomorphisms.

As the hypergraph structure is lifted from the underlying topos assigning truth values to monoids and comonoids in a trivial way by means of the graph functors, we would expect this structure to be also preserved. This is indeed the case, as we see:

Theorem 6.6.1. *Let $\mathcal{E}$ be a topos, (Σ, E) a variety in $\mathcal{E}$, and $h : Q_1 \to Q_2$ a morphism of internal commutative quantales. There is an identity on objects, strict symmetric monoidal **Pos**-functor h^*, with action on morphisms:*

$$\begin{aligned} h^* : \boldsymbol{Rel}_{(\Sigma,E)}(Q_1) &\to \boldsymbol{Rel}_{(\Sigma,E)}(Q_2) \\ A \times B \xrightarrow{R} Q_1 \quad &\mapsto \quad A \times B \xrightarrow{R} Q_1 \xrightarrow{h} Q_2 \end{aligned}$$

The assignment $h \mapsto h^*$ is functorial. Moreover, the functor h^* commutes with graphs and converses, that is, the following diagrams commute:

$$\begin{array}{ccc} \boldsymbol{Rel}_{(\Sigma,E)}(Q_1) & \xrightarrow{h^*} & \boldsymbol{Rel}_{(\Sigma,E)}(Q_2) \\ & {}_{(-)_\circ}\nwarrow \quad \nearrow{}_{(-)_\circ} & \\ & \boldsymbol{Alg}(\Sigma, E) & \end{array}$$

$$\begin{array}{ccc} \boldsymbol{Rel}_{(\Sigma,E)}(Q_1)^{op} & \xrightarrow{(h^*)^{op}} & \boldsymbol{Rel}_{(\Sigma,E)}(Q_2)^{op} \\ {\scriptstyle (-)^\circ}\downarrow & & \downarrow{\scriptstyle (-)^\circ} \\ \boldsymbol{Rel}_{(\Sigma,E)}(Q_1) & \xrightarrow[h^*]{} & \boldsymbol{Rel}_{(\Sigma,E)}(Q_2) \end{array}$$

Proof. Preservation of identities, compositions and tensors follows from the fact that h preserves quantale identities, products and joins. In particular, the preservation of quantale joins makes h also order-preserving, proving that h^* is a **Pos**-functor. The functoriality of assigments is trivial: Postcomposing with the identity function on a quantale gives back the same category we started from, and the composition of quantale homomorphisms is a quantale homomorphism.

Commutativity with graph and converse is straightforward, and we just prove the graph commutativity explicitly:

$$\begin{aligned} h^*(f^\circ)(a,b) &= h \circ f^\circ(a,b) \\ &= h\left(\bigvee\{k \mid f(a) = b\}\right) \\ &= \bigvee\{h(k) \mid f(a) = b\} \\ &= \bigvee\{k \mid f(a) = b\} \end{aligned}$$

Where the last equality follows from the fact that the quantale homomorpshism h preserves units. But the last line is exactly the definition of graph in $\mathbf{Rel}_{(\Sigma,E)}(Q_2)$, as we wanted. Preservation of graphs and converses clearly implies preservation of the hypergraph structure on the nose. □

Example 6.6.2. There is a quantale morphism $\mathbf{B} \to Q$ from the Boolean quantale to any other quantale Q, sending the top element to the top element and the bottom element to the bottom element. The induced functor identifies the ordinary binary relations as a subclass of morphisms in $\mathbf{Rel}(Q)$. This will in particular turn useful in the next chapter, where we will see how the Lavwere quantale $\mathbf{C}$ is used along with internal monads to induce a metric on our objects. The quantale homomorphism $\mathbf{B} \to \mathbf{C}$ will then allow us to "metrize" the classic category of sets and relations.

In the very same way we can consider $\mathbf{Rel}_{\text{Convex}}(\mathbf{B})$ and use the same trick to induce a metric on conceptual spaces.

In the case of the span construction we have a completely analogous result, with the difference that morphisms of partially ordered monoids are the appropriate notion of homomorphism to consider.

Theorem 6.6.3. *Let $\mathcal{E}$ be a topos, (Σ, E) a variety in $\mathcal{E}$, and $h : Q_1 \to Q_2$ a morphism of internal partially ordered commutative monoids. There is an identity on objects, strict symmetric monoidal **Preord**-functor h^*, with action on morphisms:*

$$\begin{aligned} h^* : \boldsymbol{Span}_{(\Sigma,E)}(Q_1) &\to \boldsymbol{Span}_{(\Sigma,E)}(Q_2) \\ (X, f, g, \chi) &\mapsto (X, f, g, h \circ \chi) \end{aligned}$$

The assignment $h \mapsto h^*$ is functorial. Moreover, the functor h^* commutes with graphs and converses, that is, the following diagrams commute:

$$\begin{array}{ccc} \boldsymbol{Span}_{(\Sigma,E)}(Q_1) & \xrightarrow{h^*} & \boldsymbol{Span}_{(\Sigma,E)}(Q_2) \\ {\scriptstyle (-)_\circ}\nwarrow & & \nearrow{\scriptstyle (-)_\circ} \\ & \boldsymbol{Alg}(\Sigma, E) & \end{array}$$

$$\begin{array}{ccc} \boldsymbol{Span}_{(\Sigma,E)}(Q_1)^{op} & \xrightarrow{(h^*)^{op}} & \boldsymbol{Span}_{(\Sigma,E)}(Q_2)^{op} \\ {\scriptstyle (-)^\circ}\downarrow & & \downarrow{\scriptstyle (-)^\circ} \\ \boldsymbol{Span}_{(\Sigma,E)}(Q_1) & \xrightarrow[h^*]{} & \boldsymbol{Span}_{(\Sigma,E)}(Q_2) \end{array}$$

Proof. Identical to the relational case. □

Example 6.6.4. For any commutative partially ordered monoid Q there is a partially ordered monoid morphism $1 \to Q$, induced by the monoid unit. Here, 1 is the terminal quantale. Therefore there is a strict symmetric monoidal functor

$$\mathbf{Span}_{(\Sigma,E)}(1) \to \mathbf{Span}_{(\Sigma,E)}(Q)$$

This example motivates our use of partially ordered monoids, rather than simply restricting to the quantales of interest in our primary applications, as the required morphism is not a quantale morphism.

6.7 ALGEBRAIC STRUCTURE

We now investigate the interaction between truth values and algebraic structure. Again, this will lead to functorial relationships between models, but the subject is more delicate than in the previous sections.

Definition 6.7.1. Recall Definition 6.2.2. A *term of type* Σ is a formal functional symbol $X^n \to X$ for some n (called *ariety of* τ) obtained by compostion of operation symbols in Σ, projections and diagonals.

We moreover say that a term τ over a finite set of variables is:

- *Linear* if it uses each variable exactly once (diagonals and projections are not part of the compositon);
- *Affine* if it uses each variable at most once (projections are allowed but diagonals are not);
- *Relevant* if it uses each variable at least once (diagonals are allowed but projections are not);
- *Cartesian* to emphasize that its use of variables is unrestricted (diagonals and projections both allowed).

Using compositonality of morphisms we can promptly extend Definition 6.2.3 to terms and talk of a term τ^A in an algebra A.

The concept of term is fundamental to define what interpretation of signatures is. Consider this example: We can say that any group is also a monoid, because we can "forget" the inverse operation and keep only the addition and unit, and associativity and unit law will still hold. The key of this consideration is the following:

- We can map every formal operation f in the signature of monoids to a *term* τ_f in the signature of groups. In our case this mapping is basically an immersion;
- We can use the equations in the signature of groups to prove that the equations in the signature of monoids hold for the terms τ that are in the image of the mapping defined above: The signature of groups intuitively *contains a copy* of the equations in the signature of monoids;
- Given a group G we can consider the term τ_f^G in G. $< G, \tau_{f_i}^G >$ then defines a monoid, as we wanted.

We can use the generalized version of this fact to define interpretations of signatures:

Definition 6.7.2 (Interpretation)**.** An *interpretation* of signature (Σ_1, E_1) in signature (Σ_2, E_2) is a mapping assigning each $\sigma \in \Sigma_1$ to a derived term of (Σ_2, E_2) of the same arity, such that the equations E_1 can be proved in equational logic from E_2. We say that an interpretation is *linear*, *affine*, *relevant* or *cartesian* if all the derived terms used in the interpretation are suitably restricted. It is standard that every interpretation i contravariantly induces a functor $\hat{i}$ between the categories of algebras, as clear from the previous considerations.

All the fuss about distinguishing between linear, affine, relevant and cartesian terms comes from the fact that Inequations (7) and (8) are only required to hold for the operations in our signature, hence do not extend to composition of operations involving projections and diagonals. Since interpretations of algebras use also these two types of morphisms we have to put restrictions on how our relations and spans are defined to consequently obtain a functorial relationship between the resulting categories of models.

Definition 6.7.3. Let $\mathcal{E}$ be a topos. If $(Q, \odot, k, \bigvee)$ is an internal quantale, we say that a Q-relation R is *affine* if $R(a_1, b_1) \odot R(a_2, b_2) \leqslant R(a_1, b_1)$ and *relevant* if $R(a, b) \leqslant R(a, b) \odot R(a, b)$. R is *cartesian* if it is both affine and linear. We say that R is *linear* to emphasize that no additional axioms are assumed to hold. Similarly, if $(Q, \odot, k, \leqslant)$ is an internal partially ordered monoid, we say that a Q-span (X, f, g, χ) is *affine* if $\chi(x_1) \odot \chi(x_2) \leqslant \chi(x_1)$, and *relevant* if $\chi(x) \leqslant \chi(x) \odot \chi(x)$. A Q-span is said to be *cartesian* if it is both affine and relevant, and *linear* if no additional axioms are assumed to hold.

Our terminology is derived from that sometimes used for variants of linear logic. If we view truth values as resources, the question is about when these resources can be "copied" or "deleted". The next proposition shows that if our truth values are well behaved, so are our morphisms.

Lemma 6.7.4. *Let* $(Q, \odot, k, \bigvee)$ *be an internal quantale. If for every* p, q *the inequality* $p \odot q \leqslant p$ *holds, then every relation is affine. If* $p \leqslant p \odot p$ *holds then every relation is relevant. Similarly, if* $(Q, \odot, k, \leqslant)$ *is an internal partially ordered monoid, if* $p \odot q \leqslant p$ *holds then every span is affine and if* $p \leqslant p \odot p$ *then every span is relevant.*

It is easy to check that if $p \odot q \leqslant p$ and $p \leqslant p \odot p$ both hold then $\odot$ is idempotent, making Q into a locale and viceversa. We then have the following direct corollary:

Corollary 6.7.5. *If* Q *is a locale all relations in* $\textbf{\textit{Rel}}_{(\Sigma,E)}(Q)$ *are cartesian.*

Restricting to one special kind of relations/spans among the ones defined above preserves the hypergraph structure:

Theorem 6.7.6. *Let* $\mathcal{E}$ *be a topos and* (Σ, E) *a variety in* $\mathcal{E}$*. If* $(Q, \odot, k, \bigvee)$ *is a commutative quantale, the affine, relevant and cartesian relations each form a sub-hypergraph category of* $\textbf{\textit{Rel}}_{(\Sigma,E)}(Q)$*. If* $(Q, \odot, k, \leqslant)$ *is a commutative partially ordered monoid, the affine, relevant and cartesian spans each form a sub-hypergraph category of* $\textbf{\textit{Span}}_{(\Sigma,E)}(Q)$*. In each case, the morphisms in the image of the graph functor are all cartesian.*

Proof. Just note that

- Composition of affine (relevant, cartesian) relations is again affine (relevant, cartesian);
- Similarly, composition of affine (relevant, cartesian) spans is again affine (relevant, cartesian);
- Tensor of affine (relevant, cartesian) relations is again affine (relevant, cartesian);
- Similarly, tensor of affine (relevant, cartesian) spans is again affine (relevant, cartesian);
- If a relation is constant with value the quantale unit, then it is cartesian, and hence, also affine and relevant;
- If a span with characteristic morphism χ is such that χ is constant with value the commutative partially ordered monoid unit, then the span is cartesian, and hence, also affine and relevant.

The first two observations, along with the fact that identies are constant in the span case and can only hit the quantale unit as a value in the relational case, tell us that restricting our categories of relations (spans) to affine (relevant, cartesian) ones gives us a subcategory. The fact that all the machinery used to define the hypergraph structure comes from taking the graphs of morphisms in the underlying topos $\mathcal{E}$, along with the last four observations, tells us that the hypergraph structure is preserved when we restrict to the affine (relevant, cartesian) version of our category, and this concludes the proof. □

Definition 6.7.7. We write $\mathbf{Rel}^{\mathrm{lin}}_{(\Sigma,E)}(Q)$, $\mathbf{Rel}^{\mathrm{aff}}_{(\Sigma,E)}(Q)$, $\mathbf{Rel}^{\mathrm{rel}}_{(\Sigma,E)}(Q)$ and $\mathbf{Rel}^{\mathrm{cart}}_{(\Sigma,E)}(Q)$ for the corresponding sub-hypergraph categories of cartesian relations described in Theorem 6.7.6. The definition extends straightforwardly to the span case.

Our restricted classes of relations respect the corresponding classes of terms.

Proposition 6.7.8. *Let $\mathcal{E}$ be a topos, (Σ, E) a variety in $\mathcal{E}$ and $(Q, \odot, k, \bigvee)$ an internal commutative quantale. For linear (affine, relevant, cartesian) algebraic Q-relation $R : A \to B$ the axiom:*

$$R(a_1, b_1) \odot \dots \odot R(a_n, b_n) \leqslant R(\tau(a_1, \dots, a_n), \tau(b_1, \dots, b_n))$$

Holds for every linear (affine, relevant, cartesian) n-ary derived operation τ.

Proof. Suppose R is a linear relation. We proceed by induction. By definition, if τ is any n-ary operation, then it is:

$$\bigodot_{k=1}^{n} R(a_k, b_k) \leqslant R(\tau(a_1, \dots, a_n), \tau(b_1, \dots, b_n))$$

Now let $\tau_1, \dots, \tau_n$ be operations of arities $k_1, \dots, k_n$, respectively. Being τ an operation, it is:

$$\begin{aligned}\bigodot_{i=1}^{n} R(\tau_i(a^i_1, \dots a^i_{k_i}), \tau_i(b^i_1, \dots b^i_{k_i})) \leqslant \\ \leqslant R(\tau(\tau_1(a^1_1, \dots a^1_{k_1}), \dots, \tau_n(a^n_1, \dots a^i_{k_n})), \\ , \tau(\tau_1(b^1_1, \dots b^1_{k_1}), \dots, \tau_n(b^n_1, \dots b^n_{k_n})))\end{aligned}$$

And combining with the same condition on the τ_i one obtains:

$$\begin{aligned}\bigodot_{i=1}^{n} \bigodot_{z=1}^{k_i} R(a^i_z, b^i_z) \leqslant R(\tau(\tau_1(a^1_1, \dots a^1_{k_1}), \dots, \tau_n(a^n_1, \dots a^i_{k_n})), \\ , \tau(\tau_1(b^1_1, \dots b^1_{k_1}), \dots, \tau_n(b^n_1, \dots b^n_{k_n})))\end{aligned}$$

This concludes the first part of the proof since every linear term can be written as a concatenation of operations.

Affine terms are obtained as compositions of operations and projections. It is thus sufficient to prove that the condition holds for affine relations if τ is a projection. Then, we can treat any n-ary projection as a generic operation and proceed as in the previous case. But the condition:

$$\bigodot_{k=1}^{n} R(a_k, b_k) \leqslant R(\pi(a_1, \dots, a_n), \pi(b_1, \dots, b_n))$$

Being the right hand side just $R(a_i, b_i)$, trivially holds when R is affine.

Relevant terms are obtained as compositions of operations and diagonals. Note that a diagonal δ is not a term when taken alone because it is not a morphism of the form $X^n \to X$ for some n. This means that if a term is built using diagonals, there is always at least one operation that is composed with the diagonal on the left. The proof is then very similar to the previous ones, with the additional step that if R is relevant, then the condition has to be proven to hold for every $\tau(x_1, \ldots, x_i, \delta(x), x_{i+1}, \ldots, x_n)$, where δ is the m-th diagonal and τ is any $(n+m)$-ary operation.

For cartesian terms it is sufficient to put all these observation together and proceed in the same way. □

Similarly, spans with sufficient structure respect the corresponding types of derived terms.

Proposition 6.7.9. *Let $\mathcal{E}$ be a topos, (Σ, E) a variety in $\mathcal{E}$, and $(Q, \odot, k, \leqslant)$ an internal partially ordered commutative monoid. For (Σ, E)-algebras A and B, and linear (affine, relevant, cartesian) algebraic Q-span (X, f, g, χ) and n-ary linear (affine, relevant, cartesian) term τ if:*

$$\bigwedge_i (f(x_i) = a_i \wedge g(x_i) = b_i)$$

Then there exists x such that:

$$f(x) = \tau(a_1, ..., a_n) \quad \textit{and} \quad g(x) = \tau(b_1, ..., b_n) \quad \textit{and} \quad \bigodot_i \chi(x_i) \leqslant \chi(x)$$

Proof. As in the relational case, we proceed by induction. Let (X, f, g, χ) be a span. By definition, if τ is any n-ary operation, then the axiom:

$$\bigwedge_{i=1}^{n} (f(x_i) = a_i \wedge g(x_i) = b_i) \implies \exists x : f(x) = \tau(a_1, \ldots, a_n) \wedge$$
$$\wedge g(x) = \tau(b_1, \ldots, b_n) \wedge \bigodot_{i=1}^{n} \chi(x_i) \leqslant \chi(x)$$

Is already satisfied. Now let $\tau, \tau_1, \ldots, \tau_n$ be operations of arities n, $k_1, \ldots, k_n$, respectively. Being τ an operation, it is:

$$\bigwedge_{i=1}^{n} (f(x^i) = \tau_i(a^i_1, \ldots a^i_{k_i}) \wedge g(x^i) = \tau_i(b^i_1, \ldots b^i_{k_i})) \implies$$
$$\implies \exists x : f(x) = \tau(\tau_1(a^1_1, \ldots a^1_{k_1}), \ldots, \tau_n(a^n_1, \ldots a^i_{k_n})) \wedge$$
$$\wedge g(x) = \tau(\tau_1(b^1_1, \ldots b^1_{k_1}), \ldots, \tau_n(b^n_1, \ldots b^i_{k_n})) \wedge$$
$$\wedge \bigodot_{i=1}^{n} \chi(x_i) \leqslant \chi(x)$$

And combining with the same condition on the τ_i one obtains:

$$\bigwedge_{i=1}^{n}\bigwedge_{j=1}^{k_i}(f(x_j^i) = a_j^i \wedge g(x_j^i) = b_j^i) \implies$$
$$\implies \exists x : f(x) = \tau(\tau_1(a_1^1, \dots a_{k_1}^1), \dots, \tau_n(a_1^n, \dots a_{k_n}^i))\wedge$$
$$\wedge g(x) = \tau(\tau_1(b_1^1, \dots b_{k_1}^1), \dots, \tau_n(b_1^n, \dots b_{k_n}^i))\wedge$$
$$\wedge \bigodot_{i=1}^{n}\bigodot_{j=1}^{k_i}\chi(x_j^i) \leqslant \chi(x)$$

This concludes the proof since every linear term can be written as a concatenation of operations.

For affine, relevant and cartesian terms the considerations done in the proof of Proposition 6.7.8 can easily be adapted to the span case. □

Finally, we can establish a contravariant functorial relationship between interpretations and functors between models.

Theorem 6.7.10. *Let $\mathcal{E}$ be a topos and $(Q, \odot, k, \bigvee)$ an internal commutative quantale. Let $i : (\Sigma_1, E_1) \to (\Sigma_2, E_2)$ be a linear interpretation of signatures. There is an identity on morphisms strict symmetric monoidal functor:*

$$i^* : \boldsymbol{Rel}^{\mathrm{lin}}_{(\Sigma_2,E_2)}(Q) \to \boldsymbol{Rel}^{\mathrm{lin}}_{(\Sigma_1,E_1)}(Q)$$

Sending each (Σ_2, E_2)-algebra to the corresponding (Σ_1, E_1)-algebra under the interpretation. The assignment $i \mapsto i^*$ extends to a contravariant functor. i^* commutes with graphs and converses, that is, the following diagrams commute:

$$\begin{array}{ccc} \boldsymbol{Rel}^{\mathrm{lin}}_{(\Sigma_2,E_2)}(Q) & \xrightarrow{i^*} & \boldsymbol{Rel}^{\mathrm{lin}}_{(\Sigma_1,E_1)}(Q) \\ {\scriptstyle (-)_\circ}\uparrow & & \uparrow{\scriptstyle (-)_\circ} \\ \boldsymbol{Alg}(\Sigma_2, E_2) & \xrightarrow[i^*]{} & \boldsymbol{Alg}(\Sigma_1, E_1) \end{array}$$

$$\begin{array}{ccc} \boldsymbol{Rel}^{\mathrm{lin}}_{(\Sigma_2,E_2)}(Q)^{op} & \xrightarrow{(i^*)^{op}} & \boldsymbol{Rel}^{\mathrm{lin}}_{(\Sigma_1,E_1)}(Q)^{op} \\ {\scriptstyle (-)^\circ}\downarrow & & \downarrow{\scriptstyle (-)^\circ} \\ \boldsymbol{Rel}^{\mathrm{lin}}_{(\Sigma_2,E_2)}(Q) & \xrightarrow[i^*]{} & \boldsymbol{Rel}^{\mathrm{lin}}_{(\Sigma_1,E_1)}(Q) \end{array}$$

The bottom functor in the first diagram is the obvious induced functor between categories of algebras. Similar results hold for affine, relevant and cartesian interpretations and relations.

Proof. An object of $\mathbf{Rel}^{\mathrm{lin}}_{(\Sigma_2,E_2)}(Q)$ is written as $\langle A, \sigma_j \rangle$, where A is an object of $\mathcal{E}$ and the σ_j are morphisms $A^n \to A$ in bijective correspondence with the operations in Σ_2, agreeing with them on arities, and such that they satisfy the equations in E_2 (these equations are just commutative diagrams between the above mentioned morphisms). The linear (affine, relevant, cartesian) interpretation i maps every operation in $\sigma'_k \in \Sigma_1$ to a linear (affine, relevant cartesian) term $i(\sigma'_k)$ on Σ_2, such that these terms satisfy the equations in E_1. This means that $\langle A, i(\sigma'_k) \rangle$ is an algebra of type (Σ_1, E_1).

The functor i^* then acts as follows: It sends every algebra $\langle A, \sigma_j \rangle$ to $\langle A, i(\sigma'_k) \rangle$, and it is identity on morphisms (the fact that morphisms of $\mathbf{Rel}^{\mathrm{lin}}_{(\Sigma_2,E_2)}(Q)$ are also morphisms of $\mathbf{Rel}^{\mathrm{lin}}_{(\Sigma_1,E_1)}(Q)$ is a direct consequence of Proposition 6.7.8). Functoriality then holds trivially being i^* identity on morphisms. Noting that i^* and $\hat{i}$ are identity on morphisms and they act on the same way on objects, whereas converse and graph are instead identity on objects, commutativity of i^* with converse and graphs holds trivially. □

A similar contravariant functorial relationship holds between interpretations and functors between span-based models.

Theorem 6.7.11. *Let $\mathcal{E}$ be a topos and $(Q, \odot, k, \leqslant)$ an internal partially ordered commutative monoid. Let $i : (\Sigma_1, E_1) \to (\Sigma_2, E_2)$ be a linear interpretation of signatures. There is an identity on morphisms strict monoidal functor:*

$$i^* : \boldsymbol{Span}^{\mathrm{lin}}_{(\Sigma_2,E_2)}(Q) \to \boldsymbol{Span}^{\mathrm{lin}}_{(\Sigma_1,E_1)}(Q)$$

Sending each (Σ_2, E_2)-algebra to the corresponding (Σ_1, E_1)-algebra under the interpretation. The assignment $i \mapsto i^*$ extends to a contravariant functor. i^* commutes with graphs and converses, that is, the following diagrams commute:

$$\begin{array}{ccc} \boldsymbol{Span}^{\mathrm{lin}}_{(\Sigma_2,E_2)}(Q) & \xrightarrow{i^*} & \boldsymbol{Span}^{\mathrm{lin}}_{(\Sigma_1,E_1)}(Q) \\ {\scriptstyle (-)_\circ}\uparrow & & \uparrow{\scriptstyle (-)_\circ} \\ \boldsymbol{Alg}(\Sigma_2, E_2) & \xrightarrow[i^*]{} & \boldsymbol{Alg}(\Sigma_1, E_1) \end{array}$$

$$\begin{array}{ccc} \boldsymbol{Span}^{\mathrm{lin}}_{(\Sigma_2,E_2)}(Q)^{op} & \xrightarrow{(i^*)^{op}} & \boldsymbol{Span}^{\mathrm{lin}}_{(\Sigma_1,E_1)}(Q)^{op} \\ {\scriptstyle (-)^\circ}\downarrow & & \downarrow{\scriptstyle (-)^\circ} \\ \boldsymbol{Span}^{\mathrm{lin}}_{(\Sigma_2,E_2)}(Q) & \xrightarrow[i^*]{} & \boldsymbol{Span}^{\mathrm{lin}}_{(\Sigma_1,E_1)}(Q) \end{array}$$

The bottom functor in the first diagram is the obvious induced functor between categories of algebras. Similar results hold for affine, relevant and cartesian interpretations and relations.

The extensional collapse functor of Section 6.5 also respects our different classes of spans and relations.

Proposition 6.7.12. *Let $\mathcal{E}$ be a topos, (Σ, E) a variety in $\mathcal{E}$ and $(Q, \odot, k, \bigvee)$ an internal commutative quantale. The functor of Theorem 6.5.1 maps linear (affine, relevant, cartesian) algebraic Q-spans to linear (affine, relevant, cartesian) algebraic Q-relations.*

Now we briefly discuss some examples.

Example 6.7.13. For any signature (Σ, E) there is a trivial *linear* interpretation $(\emptyset, \emptyset) \to (\Sigma, E)$. We therefore have, for every choice of internal quantale Q, strict symmetric monoidal forgetful functors:

$$\mathbf{Rel}_{(\Sigma,E)}(Q) \to \mathbf{Rel}_{(\emptyset,\emptyset)}(Q)$$
$$\mathbf{Span}_{(\Sigma,E)}(Q) \to \mathbf{Span}_{(\emptyset,\emptyset)}(Q)$$

Hence the category of Q-relations on sets is the *terminal object* in the category that has categories of algebraic Q-relations as objects and functors induced by algebraic interpretations between them as morphisms.

Example 6.7.14. We can present real vector spaces by a signature with a constant element representing the origin, and a family of binary mixing operations, indexed by the scalars involved, satisfying suitable equations. We denote this signature as Linear. It is easy to prove that there is an interpretation Convex $\to$ Linear. For any commutative quantale Q, this interpretation induces a functor:

$$\mathbf{Rel}_{\text{Linear}}(Q) \to \mathbf{Rel}_{\text{Convex}}(Q)$$

So, as we would expect, we can find the vector spaces in the convex algebras, in a manner respecting all the relevant categorical structure.

Example 6.7.15. An affine join semilattice is a set with an associative, commutative, idempotent binary operation. From an information theoretic perspective, we think of convex algebras as describing probabilistic ambiguity. Affine join semilattices can then be thought of as modelling unquantified ambiguity. As proved in Chapter 4, if we denote the signature for affine join semilattices as Affine there is an interpretation Convex $\to$ Affine inducing a functor:

$$\mathbf{Rel}_{\text{Affine}}(Q) \to \mathbf{Rel}_{\text{Convex}}(Q)$$

Relating these two different models of epistemic phenomena. This exhibits another interesting subcategory of **ConvexRel**.

6.8 CHANGING TOPOS

We now explore the last axis of variation, the topos structure. We would expect that, if $\mathcal{E}$ and $\mathcal{F}$ are elementary toposes, given a suitable functor $L : \mathcal{E} \to \mathcal{F}$ it would be possible to lift it to a functor between their respective relation and span constructions. Since the definitions of these categories make wide use of the internal language, it should not be surprising that by "suitable" we actually mean that L behaves well with respect to the logical properties of $\mathcal{E}, \mathcal{F}$.

Definition 6.8.1. Given toposes $\mathcal{E}, \mathcal{F}$, a functor $L : \mathcal{E} \to \mathcal{F}$ is called *logical*[7] if:

- L preserves products;
- L preserves exponentials;
- L preserves the subobject classifier.

Logical functors are the right functors to consider, since they preserve the validity of internal formulas: If $\models \phi$ in $\mathcal{E}$, then $\models L\phi$ in $\mathcal{F}$ for every formula ϕ written in the language of first order intuitionistic logic.

To make the following results more readable, we will have to slightly refine our notation, writing $\mathbf{Rel}^{\mathcal{E}}_{(\Sigma,E)}(Q)$ and $\mathbf{Span}^{\mathcal{E}}_{\Sigma,E}(Q)$ to explicitly indicate that the constructions are performed on topos $\mathcal{E}$. If $L : \mathcal{E} \to \mathcal{F}$ is a logical functor and Q is an internal quantale in $\mathcal{E}$, then the fact that L preserves models of first order intuitionistic theories implies that LQ is an internal quantale in $\mathcal{F}$. It makes sense, then, to consider how $\mathbf{Rel}^{\mathcal{E}}_{(\Sigma,E)}(Q)$ and $\mathbf{Rel}^{\mathcal{F}}_{(\Sigma,E)}(LQ)$ are related. The main result of the section is the following:

Theorem 6.8.2. *Let $\mathcal{E}, \mathcal{F}$ be toposes, and $L : \mathcal{E} \to \mathcal{F}$ be a logical functor. Let $(Q, \odot, k, \bigvee)$ be an internal commutative quantale in $\mathcal{E}$ and (Σ, E) be a signature. There is a symmetric monoidal functor:*

$$L^* : \boldsymbol{Rel}^{\mathcal{E}}_{(\Sigma,E)}(Q) \to \boldsymbol{Rel}^{\mathcal{F}}_{(\Sigma,E)}(LQ)$$

The assignment $L \mapsto L^*$ is functorial.

Proof. The proof heavily relies on the fact that logical morphisms preserve models of logical theories: We know that, if T is a logical theory, a logical functor $L : \mathcal{E} \to \mathcal{F}$ preserves every interpretation (that is, every model), of T in $\mathcal{E}$. This is because an interpretation of T in $\mathcal{E}$ assigns to every term and formula its correspondent in the Mitchell-Bérnabou internal language: Every type is interpreted in a product of objects, and every constant into a morphism of $\mathcal{E}$. The axioms correspond, finally, to commutative diagrams in $\mathcal{E}$. Since these diagrams involve only limits, exponentials and subobject classifiers, they are preserved by L up to isomorphism. This means that the image through L of objects and morphisms that constitute a model of T in $\mathcal{E}$ is a model of T in $\mathcal{F}$. The idea is then to state our definition of composition and identity of $\mathbf{Rel}^{\mathcal{E}}_{(\Sigma,E)}(Q)$ in terms of logical theories: In this case the composition and the identity of two algebra-preserving relations will be just a model of this theory in $\mathcal{E}$, and will hence be preserved by

7 Note that another definition of functor that may be worth consider here is the one of *geometric morphism*. We used logical functors because we are dealing with elementary toposes and we care about the preservation of intuitionistic first-order formulas. Geometric functors, instead, preserve geometric theories that are a completely different beast, and are widely used in the framework of Grothendiek toposes. Note moreover that every Grothendiek topos is an elementary topos, but the opposite is not true.

L. The images through L of our relations will then still satisfy our definition of composition in the internal language of $\mathcal{F}$, guaranteeing that $L(R \circ S)(a,c) = \bigvee_b \{LR(a,b) \odot LS(b,c)\} = (LR \circ LS)(a,c)$. From this, we can define $L^* : \mathbf{Rel}^{\mathcal{E}}_{(\Sigma,E)}(Q) \to \mathbf{Rel}^{\mathcal{F}}_{(\Sigma,E)}(LQ)$ as follows:

- On objects, $L^*(A) = L(A)$;
- On morphisms, denoting with κ the canonical isomorphism from $LA \times LB$ to $L(A \times B)$,

$$L^*(R) = LA \times LB \xrightarrow{\kappa} L(A \times B) \xrightarrow{LR} LQ$$

Now we have to state what our composition is in terms of logical theories. Given a signature (Σ, E), we can define a logical theory

$$T = (A, B, C, Q, \{\sigma_i^A\}_{\sigma_i \in \Sigma}, \{\sigma_i^B\}_{\sigma_i \in \Sigma}, \{\sigma_i^C\}_{\sigma_i \in \Sigma}, \\ \odot, \bigvee, k, id_B, R_{AB}, R_{BC}, R_{AC}),$$

where

- For a given $\sigma_i \in \Sigma$ of ariety n_i,
  - σ_i^A is a constant of type $A^{A^{n_i}}$;
  - σ_i^B is a constant of type $B^{B^{n_i}}$;
  - σ_i^C is a constant of type $C^{C^{n_i}}$;
- $\odot$ is a constant of type $Q^{Q \times Q}$;
- $\bigvee$ is a constant of type Q^{PQ};
- k is a constant of type Q;
- id_B is a constant of type $Q^{B \times B}$;
- R_{AB}, R_{BC}, R_{CD} are constants of type $Q^{A \times B}, Q^{B \times C}, Q^{A \times C}$, respectively.

We require this theory to satisfy the set of axioms :

$$\{\alpha_{id_B}, \{\alpha_A\}, \{\alpha_B\}, \{\alpha_C\}, \{\alpha_Q\}, \{\alpha_{R_{AB}}\}, \{\alpha_{R_{BC}}\}, \alpha_{comp}\}$$

Where:

- $\{\alpha_A\}$ is the set of axioms that makes $\langle A, \{\sigma_i^A\}_{\sigma_i \in \Sigma} \rangle$ into an algebra of type (Σ, E);
- $\{\alpha_B\}$ is the set of axioms that makes $\langle B, \{\sigma_i^B\}_{\sigma_i \in \Sigma} \rangle$ into an algebra of type (Σ, E);
- $\{\alpha_C\}$ is the set of axioms that makes $\langle C, \{\sigma_i^C\}_{\sigma_i \in \Sigma} \rangle$ into an algebra of type (Σ, E);

- $\{\alpha_Q\}$ is the set of axioms that makes $(Q, \otimes, k, \bigvee)$ into an internal quantale;
- α_{id_B} is the axiom $\forall b, b'.id_B(b, b') = \bigvee\{k \mid b = b'\}$;
- $\{\alpha_{R_{AB}}\}$ is the set of all the axioms, one for every $\sigma_i \in \Sigma$ of ariety n, of the form:

$$\forall_{a_1,\ldots,a_n,b_1\ldots,b_n} \bigvee \Bigg\{ R_{AB}(\sigma_i^A(a_1,\ldots,a_n), \sigma_i^B(b_1,\ldots,b_n)), \bigodot_{j=1}^{n} R(a_j, b_j) \Bigg\} = \\ = R_{AB}(\sigma_i^A(a_1,\ldots,a_n), \sigma_i^B(b_1,\ldots,b_n))$$

  (Note that in this setting to use the quantale order relation we have to write explicitly what it is. The axiom above is nothing but the algebra preservation axiom for R_{AB} written explicitly using the algebraic lattice structure);
- $\{\alpha_{R_{BC}}\}$ is the set of all the axioms, one for every $\sigma_i \in \Sigma$ of ariety n,of the form:

$$\forall_{b_1\ldots,b_n,c_1,\ldots,c_n} \bigvee \Bigg\{ R_{BC}(\sigma_i^B(b_1,\ldots,b_n), \sigma_i^C(c_1,\ldots,c_n)), \bigodot_{j=1}^{n} R(b_j, c_j) \Bigg\} = \\ = R_{BC}(\sigma_i^B(b_1,\ldots,b_n), \sigma_i^C(c_1,\ldots,c_n))$$

- Finally, α_{comp} is the axiom:

$$\forall a, c.R_{AC}(a, c) = \bigvee\{R_{AB}(a, b) \odot R_{BC}(b, c) \mid b \in B\}$$

An interpretation of this theory in the topos $\mathcal{E}$ then consists of three morphisms in $\mathcal{E}$

$$R_{AB} : A \times B \to Q \quad R_{BC} : B \times C \to Q \quad R_{AC} : A \times C \to Q$$

Where the sets of axioms $\{\alpha_A\}, \{\alpha_B\}, \{\alpha_C\}$ get interpreted into commutative diagrams ensuring that A, B, C are algebras of signature (Σ, E), respectively, while $\{\alpha_{R_{AB}}\}, \{\alpha_{R_{BC}}\}$ guarantee that R_{AB} and R_{BC} respect the usual algebraic condition. $\{\alpha_Q\}$ gets interpreted into diagrams ensuring that Q is an internal quantale and α_{comp} guarantees that R_{AC} is exactly the composition of relations R_{AB}, R_{BC} in $\mathbf{Rel}^{\mathcal{E}}_{(\Sigma,E)}(Q)$. □

As in the previous cases, graph and converse functors are preserved.

Proposition 6.8.3. *With the same assumptions, the induced functor* L^* *of Theorem 6.8.2 commutes with graphs and converses. That is, the following diagrams commute:*

$$\begin{array}{ccc} \textbf{\textit{Rel}}^{\mathcal{E}}_{(\Sigma,E)}(Q) & \xrightarrow{L^*} & \textbf{\textit{Rel}}^{\mathcal{F}}_{(\Sigma,E)}(LQ) \\ (-)_\circ \uparrow & & \uparrow (-)_\circ \\ \textbf{\textit{Alg}}^{\mathcal{E}}(\Sigma,E) & \xrightarrow[L]{} & \textbf{\textit{Alg}}^{\mathcal{F}}(\Sigma,E) \end{array}$$

$$\begin{array}{ccc} \textbf{\textit{Rel}}^{\mathcal{E}}_{(\Sigma,E)}(Q)^{op} & \xrightarrow{(L^*)^{op}} & \textbf{\textit{Rel}}^{\mathcal{F}}_{(\Sigma,E)}(LQ)^{op} \\ (-)^\circ \downarrow & & \downarrow (-)^\circ \\ \textbf{\textit{Rel}}^{\mathcal{E}}_{(\Sigma,E)}(Q) & \xrightarrow[L^*]{} & \textbf{\textit{Rel}}^{\mathcal{F}}_{(\Sigma,E)}(LQ) \end{array}$$

As with relations, morphisms between toposes extend functorially to morphisms between spans.

Theorem 6.8.4. *Let* $\mathcal{E}, \mathcal{F}$ *be toposes, and* $L : \mathcal{E} \to \mathcal{F}$ *be a logical functor. Let* $(Q, \odot, k, \leqslant)$ *be an internal partially ordered commutative monoid in* $\mathcal{E}$ *and* (Σ, E) *be a signature. There is a symmetric monoidal functor:*

$$L^* : \textbf{\textit{Span}}^{\mathcal{E}}_{(\Sigma,E)}(Q) \to \textbf{\textit{Span}}^{\mathcal{F}}_{(\Sigma,E)}(LQ)$$

The assignment $L \mapsto L^*$ *is functorial.*

Proof. Here the same considerations used to prove Theorem 6.8.2 hold. Given a signature (Σ, E), the logical theory we use is:

$$T = (X, A, B, Q, \{\sigma_i^A\}_{\sigma_i \in \Sigma}, \{\sigma_i^B\}_{\sigma_i \in \Sigma}, f, g, \chi, \odot, \leqslant, k)$$

Where:

- For a given $\sigma_i \in \Sigma$ of ariety n_i,
  - σ_i^A is a constant of type $A^{A^{n_i}}$;
  - σ_i^B is a constant of type $B^{B^{n_i}}$;
  - σ_i^C is a constant of type $C^{C^{n_i}}$;
- f, g, χ are constants of type X^A, X^B, X^Q", respectively;
- $\odot$ is a constant of type $Q^{Q \times Q}$;
- $\leqslant$ is a constant of type $\Omega^{Q \times Q}$;
- k is a constant of type Q;

We require this theory to satisfy the set of axioms:

$$\{\{\alpha_A\}, \{\alpha_B\}, \{\alpha_C\}, \{\alpha_Q\}, \{\alpha_X\}\}$$

- $\{\alpha_A\}$ is the set of axioms that makes $\langle A, \{\sigma_i^A\}_{\sigma_i \in \Sigma}\rangle$ into an algebra of type (Σ, E);
- $\{\alpha_B\}$ is the set of axioms that makes $\langle B, \{\sigma_i^B\}_{\sigma_i \in \Sigma}\rangle$ into an algebra of type (Σ, E);
- $\{\alpha_C\}$ is the set of axioms that makes $\langle C, \{\sigma_i^C\}_{\sigma_i \in \Sigma}\rangle$ into an algebra of type (Σ, E);
- $\{\alpha_Q\}$ is the set of axioms that makes $(Q, \odot, k, \leqslant)$ into an internal partially ordered monoid;
- $\{\alpha_X\}$ is the set of axioms, one for every $\sigma \in \Sigma$ of ariety n, of the form:

$$\begin{gathered}\forall_{x_1,\ldots,x_n} \exists x. f(x) = \sigma_i^A(f(x_1), \ldots, f(x_n) \wedge \\ \wedge \big(g(x) = \sigma_i^B(g(x_1), \ldots, g(x_n)\big) \wedge \\ \wedge \bigodot_{j=1}^{n} \chi(x_j) \leqslant \chi(x)\end{gathered}$$

A model of T in $\mathcal{E}$ is just a span that respects the algebraic structure and we know that L preserves this condition. L^* then agrees with L on objects and is defined as $(LX, Lf, Lg, L\chi)$ on the morphism (X, f, g, χ). For composition and identity we do not need to invoke any logical theory: The identity span is of the form $(X, 1_X, 1_X, \chi_k)$, and the span part is clearly preserved because functors preserve identities in general. The quantale part χ_k is the morphism $A \to 1 \to Q$ where the final arrow sends the terminal object to the quantale unit. Again, being L logical this is trivially preserved. For composition, note that the span part is composed via pullbacks and L preserves limits. For the quantale part we have, supposing (Z, h, k, ζ) to be the composite of (X, f, g, χ) and (Y, f', g', υ),

$$\begin{array}{ccccccc} LZ & \xrightarrow{L\langle p_1, p_2\rangle} & L(X \times Y) & \xrightarrow{L(\chi \times \upsilon)} & L(Q \times Q) & \xrightarrow{L(\odot)} & L(Q) \\ & \searrow^{\langle Lp_1, Lp_2\rangle} & \downarrow \text{iso} & & \text{iso} \downarrow & \nearrow_{\odot} & \\ & & LX \times LY & \xrightarrow[L\chi \times L\upsilon]{} & LQ \times LQ & & \end{array}$$

Where p_1, p_2 are the pullback projections. The top row is the image of ζ through L. The triangle on the left and the square on the center commute because L preserves limits, while the triangle on the right commutes because every partially ordered monoid is obviously a model of a theory, so the multiplication of Q gets carried in the multiplication of LQ. Moreover, since L trivially preserves isomorphisms, we can be sure that isomorphic spans get carried to isomorphic spans, hence our correspondence doesn't depend on the choice of representatives. □

The essential structure is again respected by the induced functors.

Proposition 6.8.5. *With the same assumptions, the induced functor* L^* *of Theorem 6.8.4 commutes with graphs and converses. That is, the following diagrams commute:*

$$\begin{array}{ccc} \boldsymbol{Span}^{\mathcal{E}}_{(\Sigma,E)}(Q) & \xrightarrow{L^*} & \boldsymbol{Span}^{\mathcal{F}}_{(\Sigma,E)}(LQ) \\ {\scriptstyle (-)_\circ}\uparrow & & \uparrow{\scriptstyle (-)_\circ} \\ \boldsymbol{Alg}^{\mathcal{E}}(\Sigma,E) & \xrightarrow[L]{} & \boldsymbol{Alg}^{\mathcal{F}}(\Sigma,E) \end{array}$$

$$\begin{array}{ccc} \boldsymbol{Span}^{\mathcal{E}}_{(\Sigma,E)}(Q)^{op} & \xrightarrow{(L^*)^{op}} & \boldsymbol{Span}^{\mathcal{F}}_{(\Sigma,E)}(LQ)^{op} \\ {\scriptstyle (-)^\circ}\downarrow & & \downarrow{\scriptstyle (-)^\circ} \\ \boldsymbol{Span}^{\mathcal{E}}_{(\Sigma,E)}(Q) & \xrightarrow[L^*]{} & \boldsymbol{Span}^{\mathcal{F}}_{(\Sigma,E)}(LQ) \end{array}$$

Example 6.8.6. Given any category $\mathcal{C}$ we can form a corresponding *presheaf category*, having representable functors from $\mathcal{C}$ to **Set** as objects and natural transformations as morphisms. Presheaves constitute one of the most important examples of toposes, and it makes sense to ask how Theorems 6.8.2, 6.8.4 behave in these circumstances.

In general, given arbitrary categories $\mathcal{C}, \mathcal{D}$ it is difficult to say when a functor $F : \mathcal{C} \to \mathcal{D}$ lifts to a logical functor between the corresponding presheaves. Nevertheless, the following result holds: If $\mathcal{C}, \mathcal{D}$ are groupoids (categories in which every arrow is an isomorphism), then any functor $F : \mathcal{C} \to \mathcal{D}$ lifts to a logical functor $\bar{F}$ between presheaves. This is because truth values in presheaf toposes are defined in terms of *sieves* (subfunctors of the homset functor) and these sieves trivialize when the only arrows at our disposal are isos. This in turn trivializes the structure of truth values in the presheaf itself, that ends up to be defined pointwise from **Set**.

Theorems 6.8.2, 6.8.4 then ensure that $\bar{F}$ can be lifted to the relational and span structures built on $\mathbf{Set}^{\mathcal{C}}$ and $\mathbf{Set}^{\mathcal{D}}$, respectively.

Example 6.8.7. If $\mathcal{E}$ is a topos, and $f : I \to J$ is a morphism of $\mathcal{E}$, then pulling back along f induces a logical functor $F : \mathcal{E}/J \to \mathcal{E}/I$. Theorem 6.8.2 guarantees the existence of a functor:

$$F^* : \mathbf{Rel}^{\mathcal{E}/J}_{(\Sigma,E)}(Q) \to \mathbf{Rel}^{\mathcal{E}/I}_{(\Sigma,E)}(FQ)$$

In particular, this means that there is always a functor:

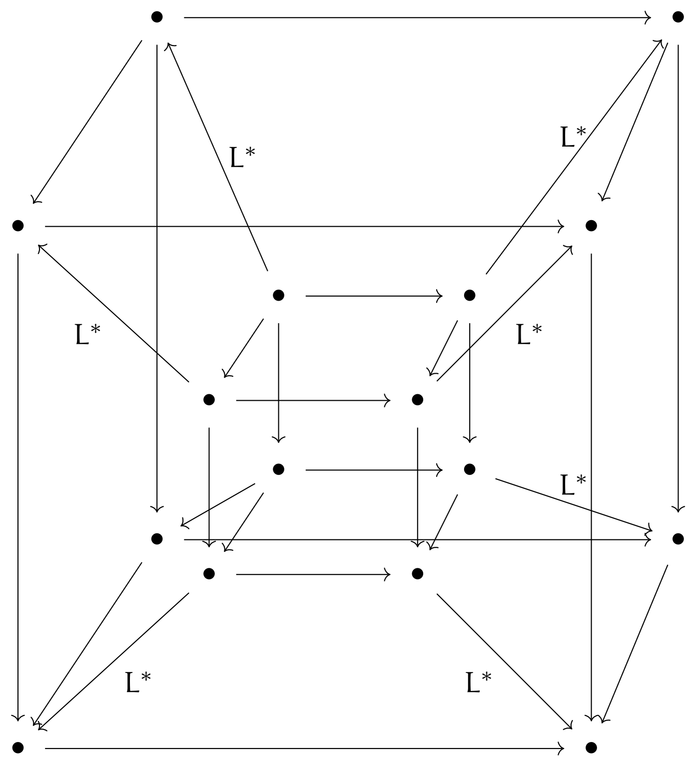
$$F^* : \mathbf{Rel}^{\mathcal{E}}_{(\Sigma,E)}(Q) \to \mathbf{Rel}^{\mathcal{E}/I}_{(\Sigma,E)}(FQ)$$

where $\mathcal{E}/I$ is any slice topos of $\mathcal{E}$.

6.9 INDEPENDENCE OF THE AXES OF VARIATION

Finally, we establish that our various induced functors between models are independent, in that they all commute with each other. Unfortunately, the commutativity of the functors induced by interpretations between algebras, order structure and quantale morphisms with L^* will hold only up to isomorphism. This depends intrinsicly on the definition of logical functor, that is, in turn, defined to preserve validity of formulas in the internal language only up to natural isomorphism.

Theorem 6.9.1. *Let $\mathcal{E}$ be a topos, $h : Q_1 \to Q_2$ a morphism of internal commutative quantales, $i : (\Sigma_1, E_1) \to (\Sigma_2, E_2)$ a linear interpretation and $L : \mathcal{E} \to \mathcal{F}$ a logical functor. For the induced functors of Theorems 6.6.1, 6.6.3, 6.7.10, 6.7.11, 6.8.2 and 6.8.4, the following diagram commutes (be aware that in the hypercube below commutative squares involving L^* only commute up to isomorphism. Other squares commute up to equality):*



Where the inner cube is:

$$
\begin{array}{ccccccc}
 & & \textbf{\textit{Span}}^{\text{lin},\mathcal{E}}_{(\Sigma_2,E_2)}(Q_1) & \xrightarrow{i^*} & \textbf{\textit{Span}}^{\text{lin},\mathcal{E}}_{(\Sigma_1,E_1)}(Q_1) & & \\
 & {\scriptstyle h^*}\swarrow & \downarrow & & {\scriptstyle h^*}\swarrow\quad\downarrow & & \\
\textbf{\textit{Span}}^{\text{lin},\mathcal{E}}_{(\Sigma_2,E_2)}(Q_2) & & \xrightarrow{i^*} & \textbf{\textit{Span}}^{\text{lin},\mathcal{E}}_{(\Sigma_1,E_1)}(Q_2) & & & \\
\downarrow & & \downarrow & \downarrow & \downarrow & & \\
 & & \textbf{\textit{Rel}}^{\text{lin},\mathcal{E}}_{(\Sigma_2,E_2)}(Q_1) & \xrightarrow{i^*} & \textbf{\textit{Rel}}^{\text{lin},\mathcal{E}}_{(\Sigma_1,E_1)}(Q_1) & & \\
\downarrow & \swarrow{\scriptstyle h^*} & & \downarrow & \swarrow{\scriptstyle h^*} & & \\
\textbf{\textit{Rel}}^{\text{lin},\mathcal{E}}_{(\Sigma_2,E_2)}(Q_2) & & \xrightarrow{i^*} & \textbf{\textit{Rel}}^{\text{lin},\mathcal{E}}_{(\Sigma_1,E_1)}(Q_2) & & &
\end{array}
$$

And the outer cube is:

$$
\begin{array}{ccccccc}
 & & \textbf{\textit{Span}}^{\text{lin},\mathcal{F}}_{(\Sigma_2,E_2)}(LQ_1) & \xrightarrow{i^*} & \textbf{\textit{Span}}^{\text{lin},\mathcal{F}}_{(\Sigma_1,E_1)}(LQ_1) & & \\
 & {\scriptstyle (Lh)^*}\swarrow & \downarrow & & {\scriptstyle (Lh)^*}\swarrow\quad\downarrow & & \\
\textbf{\textit{Span}}^{\text{lin},\mathcal{F}}_{(\Sigma_2,E_2)}(LQ_2) & & \xrightarrow{i^*} & \textbf{\textit{Span}}^{\text{lin},\mathcal{F}}_{(\Sigma_1,E_1)}(LQ_2) & & & \\
\downarrow & & \downarrow & \downarrow & \downarrow & & \\
 & & \textbf{\textit{Rel}}^{\text{lin},\mathcal{F}}_{(\Sigma_2,E_2)}(LQ_1) & \xrightarrow{i^*} & \textbf{\textit{Rel}}^{\text{lin},\mathcal{F}}_{(\Sigma_1,E_1)}(LQ_1) & & \\
\downarrow & \swarrow{\scriptstyle (Lh)^*} & & \downarrow & \swarrow{\scriptstyle (Lh)^*} & & \\
\textbf{\textit{Rel}}^{\text{lin},\mathcal{F}}_{(\Sigma_2,E_2)}(LQ_2) & & \xrightarrow{i^*} & \textbf{\textit{Rel}}^{\text{lin},\mathcal{F}}_{(\Sigma_1,E_1)}(LQ_2) & & &
\end{array}
$$

In both cases the vertical arrows are the functors of Theorem 6.5.1. Similar diagrams commute for affine, relevant and cartesian interpretations, relations and spans.

7 THINGS WE CAN DO - TAKE II

A disputandi subtilitate orationem
ad exempla traducimus.

— *Cicero, Tusculanae Disputationes [101, Book III, 56]*

Starting from our first conceptual spaces axiomatization of Chapter 3, in Chapter 6 we broadly extended our original model along four different axes of variation. We can now tweak the algebraic structure that was at the base of our definition of convexity, generalize the structure of truth values going beyond simple binary dichotomies and account for proof witnesses when establishing a relation between convex sets is not enough; lastly, we allowed for the use of different universes of discourse in the choice of an arbitrary topos other than **Set**. Remarkably, we were able to do all this saving the compact closed and hypergraph structure, meaning that we still have our beautiful graphical calculus made of cups, caps and multi-wires. Now we will elaborate on these properties further, and will focus in particular on three concepts:

- First of all, in Section 7.1 we will account for a construction we anticipated in Chapter 6. Namely, we will use the order enrichment to define internal monads on our models. We will show how internal monads will give us things that look like distances in the relational case and internal categories in the span case [131]. This is probably the biggest contribution of this chapter, since taking distances of concepts has always been one of the most usueful features of conceptual spaces models based on vector spaces. With internal monads, we are getting this feature in our model too.
- In Section 7.2 we will show how the span construction can be used to model semantic ambiguity, with different proof witnesses allowing us to vary how strongly different words are related, depending on how they are interpreted;
- In Section 7.3 we will switch to other universes of discourse, investigating which toposes could give us a base to model interesting phenomena. In particular, we will focus on presheaf toposes and we will show how, if the algebraic Q-relation construction is difficult to use because internal quantales are much less innocuous beasts than it may seem, on the other hand the

algebraic Q-span construction behaves particularly well, allowing us to formalize different arguments of interest in linguistics and cognition, such as concepts varying over time or different agents.

Since we already went through an awful lot of mathematics in the last chapter, we will now try to proceed in a more grounded way by means of examples, harvesting the fruits of our work.

7.1 METRIC AND DISTANCES (FINALLY)

Implementing metrics in a compact closed environment is tricky. The standard approach one could think of is to start from categories of metric spaces and considering relations over them. This approach is however destined to fail quite soon since these categories are not regular, and we cannot employ the results in [77] to infer compact closedness. Without any sugar, this means that the approach we used in the previous chapters to define **ConvexRel** is deemed to fail on metric spaces no matter what and that, in general, the "standard" categorified idea of "taking distances" is not compatible with our idea of categories admitting a diagrammatic calculus in terms of string diagrams. To solve this problem, we have to rely on different ways to categorify the notion of distance.

Our idea is to introduce something that looks like a metric structure on categories of relations relying on the the broadened definition of *truth values* that a relation can assume. This is directly compatible with the direction of our work in Chapter 6: To put a metric on $\mathbf{Rel}_{(\Sigma,E)}(Q)$ the last thing we need is the bicategorical notion of internal monad.

Consider a category $\mathcal{C}$ enriched over **Cat** (the category of categories and functors between them). In this setting, every homset of $\mathcal{C}$ is a category, and so we can talk about "elements" of the homsets (every morphism $A \to B$ in $\mathcal{C}$ is an object of $\mathrm{Hom}_{\mathcal{C}}[A, B]$). A category enriched over **Cat** is usually called *2-category*. We can further relax our definition, requiring our composition to be associative only up to isomorphism. What we get, then, is called a *bicategory* (for more information, see [21]).

Remark 7.1.1. In this setting objects of $\mathcal{C}$ are usually called 0-*cells*, while morphisms of $\mathcal{C}$ are called 1-*cells*. We will denote 0-cells with capital letters and 1-cells as $f : A \to B$, as we usually do. Now, since for each A, B $\mathrm{Hom}_{\mathcal{C}}[A, B]$ is itself a category, we can consider morphisms of $\mathrm{Hom}_{\mathcal{C}}[A, B]$, that are called 2-*cells*, and we will denote them as $h \Rightarrow k$.

When we have a bicategory, we can internalize the notion of monad, as follows: We can identify categories with 0-cells, functors with 1-cells and natural transformations with 2-cells. An internal monad on an object A is then just a 1-cell $d : A \to A$ together with 2-cells $\eta : id_A \Rightarrow d$, $\mu : (d \circ d) \Rightarrow d$ such as the usual monad diagrams hold. This prompts the following definition:

Definition 7.1.2. Given a bicategory $\mathcal{C}$ and an object A of $\mathcal{C}$, an *internal monad on* A is a monoid in the monoidal category $\mathrm{Hom}_{\mathcal{C}}[A, A]$ with composition as tensor and identity as tensor unit. More information about internal monads can be found in [139].

Now, note that both preordered sets and posets are special cases of categories, so we can apply the concept of internal monad to our setting. Being $\mathbf{Rel}_{(\Sigma,E)}(Q)$ enriched over posets (Theorem 6.4.6), we do not have much choice for our 2-cells: for every d either they exist, and in this case there is exactly one couple of arrows $\eta : id_A \Rightarrow d$ and $\mu : d \circ d \Rightarrow d$, for which the monad conditions will be satisfied trivially, or they do not, in which case d is not an internal monad. In our relational setting the internal monad condition then reads:

$$\forall a, b. \quad \bigvee\{k \mid a = b\} \leqslant d(a, b) \qquad \bigvee_x d(a, x) \odot d(x, b) \leqslant d(a, b) \quad (9)$$

Recall now the quantales given in Example 6.0.2:

- 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;
- The *Lawvere quantale* $\mathbf{C}$ 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;
- The *force quantale* $\mathbf{F}$ has again the extended positive reals with reverse order as its partial order, but now we use max as the monoid multiplication;
- The *interval quantale* $\mathbf{I}$ is given by the ordered interval $[0, 1]$ with suprema as joins and minima as the monoid multiplication.

If we specialize the conditions (9) to $\mathbf{Rel}_{(\Sigma,E)}(\mathbf{C})$ they are equivalent to:

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

Where the ordering, the $+$ and the 0 are the usual ones for real numbers. Using the Lawvere quantale, then, the condition of being an internal monad is equivalent to the condition of being a (non necessarily symmetric) *distance*. We therefore consider these internal monads as describing *generalized metric spaces*. This observation is important in the field of monoidal topology [79]. Similarly, if we consider $\mathbf{Rel}_{(\Sigma,E)}(\mathbf{F})$ the conditions (9) become:

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

and we can therefore see such internal monads as *generalized ultrametric spaces*.

So we see that, in some nice cases, internal monads end up being distances describing some (kind of) metric spaces. An interpretation of what these distances represent, that at the moment we still lack, will be inevitably linked to the interpretation we give to our truth values. In fact, all the quantales hereby considered happen to be the sensible choice to model a different class of phenomena:

Example 7.1.3. The relations over the Lawvere quantale **C** can be thought of as describing costs [93]. The value $R(a,b)$ describes the cost of converting a into b. A cost of 0 means a and b are maximally related and can be freely interconverted. A cost of ∞ indicates completely unrelated elements, that cannot be converted between each other for any finite cost. The value $(S \circ R)(a,c)$ describes the cheapest added cost of converting a into some b, and then converting that b into c. If we perform two conversions in parallel, $(R \odot R')(a,a',b,b')$ describes the sum of the two individual conversion costs.

With regard to the algebraic structure, consider the operations σ_i in the algebraic signature as means of going from *raw* materials or quantities $a_1, a_2, \dots a_n$ to a *processed* product $\sigma_i(a_1,\dots,a_n)$. The Inequation (7) in Definition 6.2.4 then states that the cost of converting a processed product into another obtained by the same means but different materials is smaller than the sum of the costs needed to convert raw materials between them (post-conversion is somehow more convenient than converting everything before).
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.

In the Lawvere quantale an internal monad d describes costs such that:

- Converting a into itself (e.g. by doing nothing) is free;
- Converting a into b directly is cheaper than converting a into b by means of an intermediate conversion to some other x.

Note that our definition of internal monad in a Lawvere quantale is compatible with the usual definition of distance. This observation is one of the key reasons why the Lawvere quantale is relevant in monoidal topology.

Proposition 7.1.4. *The usual distance on* $\mathbb{R}^n$ *defines an internal monad on* $\textbf{\textit{Rel}}_{(\emptyset,\emptyset)}(\mathbf{C})$, *where* $(\emptyset,\emptyset)$ *is the empty signature corresponding to just sets as already showed in Example 6.7.13. Hence, the usual euclidean distance defines an internal monad on the category of sets and* **C***-relations.*

Example 7.1.5. The quantale **F** has the same underlying set as the Lawvere quantale, but its different algebraic structure leads to another interpretation. We think of $R(a,b)$ as the peak force required to move a to b [93]. The value given by the composite $(S \circ R)(a, c)$ then describes the optimum peak force we will require to move a to c. To better explain this, imagine the following scenario: Suppose you want to move a to c and that you can perform this job in two different ways: You can either move a to b and then b to c; or you can move a to b' and then b' to c. The peak force required to perform each of these operations is as in the following table:

Movement	Peak force required (in some fixed units)
a to b	1
a to b'	1.5
b to c	1.7
b' to c	0.3

Table 4: Peak force costs for movements between a, b, b', c.

The total peak force required to move a to c through b is then 1.7, while moving a to c through b' requires a peak force of 1.5. Composition in $\mathbf{Rel}_{(\Sigma,E)}(\mathbf{F})$ returns the the preferred alternative if we want to minimize our peak effort, that would be the value 1.7 in the example above. Similarly, the truth value $(R \odot R')(a, a', b, b')$ gives the peak force required to complete both conversions, assuming these costs are independently incurred.

The Inequation (7) in Definition 6.2.4 can here be interpreted as follows: $\sigma_i \in \Sigma$ expresses some kind of relation $a_1, \dots, a_n$ satisfy (e.g. σ_i tells us all the a_j are tied one with the other and must be moved all together). Moving $\sigma_i(a_1, \dots, a_n)$ to $\sigma_i(b_1, \dots, b_n)$ then requires a peak force that is smaller than the sum of the peak forces required to move every a_j to b_j separately.

In this setting internal monads describe relations such that:

- The peak force required to move a to a (that is, to do nothing) is zero;
- The peak force needed to move a to b is less than moving a to x first and then x to b (a direct path is always the most convenient one).

As with Example 7.1.3, we can think of states and effects as tables of acquisition and elimination forces.

Example 7.1.6. 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. Inequation (7) in Definition 6.2.4 says that if elements are connected so are operations between them. Generalizing to the interval quantale, we now think of $R(a, b)$ as "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 $(R \odot R')(a, a', b, b')$ 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. With respect to the algebraic structure we have that the strength of the connection between operations performed between objects is at least as strong as the minimum connection strength between the objects themselves.

In this setting internal monads describe relations such that:

- a is always connected to itself with maximum strength;
- A direct connection between a and b is always stronger than connecting a and b my means of some other x: Bouncing the signal through intermediate stations may cause losses.

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: States and effects are sets with fuzzy membership, morphisms are fuzzy predicates. Graded membership is widely used in cognitive science, for example in [20, 49, 73, 74, 130]. 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 [119], fuzzy relations may prove useful.

In each of these cases, clearly, internal monads are also algebraic Q-relations, and hence respect the algebraic structure of the objects they act on.

We already proved in Chapter 6 that every quantale homomorphism $h : Q_1 \to Q_2$ induces a strict monoidal functor from $\mathbf{Rel}_{(\Sigma,E)}(Q_1)$ to $\mathbf{Rel}_{(\Sigma,E)}(Q_2)$ (Theorem 6.6.1). It is moreover easy to prove that:

Proposition 7.1.7. *If* $h : Q_1 \to Q_2$ *is an injective quantale morphism (in* ***Set****), the induced strict monoidal functor* $h^* : \boldsymbol{Rel}_{(\Sigma,E)}(Q_1) \to \boldsymbol{Rel}_{(\Sigma,E)}(Q)$ *is faithful.*

We put together this simple proposition with Example 6.6.2, specializing it to the case Q is the Lawvere quantale **C**.

Proposition 7.1.8. *There is a quantale homomorphism from the Boolean quantale* **B** *to the Lawvere quantale* **L** *given by:*

$$\bot \mapsto \infty \qquad \top \mapsto 0 \tag{10}$$

The quantale homomorphism in (10) is clearly injective, and hence the category of ordinary binary relations over sets can be seen as a subcategory of $\mathbf{Rel}_{(\emptyset,\emptyset)}(\mathbf{C})$.

Now remember that $\mathbf{Rel}_{(\emptyset,\emptyset)}(\mathbf{B})$ is just $\mathbf{Rel}$. A state $I \to X$ in $\mathbf{Rel}$ is clearly a *subset of* X. We can then consider two subsets U, V of X and express them as two states $U, V : I \to X$. Applying the converse and composing, $V^\circ \circ U$ is a relation $I \to I$. By definition, the monoidal unit comes from the underlying terminal object in **Set** and hence it is just the singleton set $\{*\}$. The relation $V^\circ \circ U$ then evaluates to 1 if U, V have non-empty intersection and 0 otherwise and acts as a rudimental definition of distance between sets or, to be more precise, as a "separation test". Seeing $\mathbf{Rel}$ as a subcategory of $\mathbf{Rel}_{(\emptyset,\emptyset)}(\mathbf{C})$ as we pointed out above, we can put our internal monads to good use hugely refining this consideration:

Proposition 7.1.9. *If* $U, V \subseteq X$ *and* d *is an internal monad in* ***Rel***$(\mathbf{C})$*, 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. Keep in mind that $V^\circ \circ d \circ U$ gives us the infimum of the distances between elements in U and V *according to* d, and since d is not required to be symmetric, $U^\circ \circ d \circ V$ may give a different measure of similarity. Asymmetric measures of distance are not at all a bad thing, instead they are quite relevant when it comes to applications in cognitive science. 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, modelling similarity as a function of distance, for example as in [137]. However, judgements of similarity are not necessarily symmetric [141]: 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.

Everything we said can be promptly generalized to algebraic **C**-relations, and we obtain:

Corollary 7.1.10. *If* $U, V \subseteq X$ *and* d *is an internal monad in* ***Rel***$_{(\Sigma,E)}(\mathbf{C})$*, the composite* $V^\circ \circ d \circ U$ *is the infimum of the distances between elements in* U *and* V*.*

In particular, if we see $\mathbf{Rel}_{\text{Convex}}(\mathbf{B})$ as a subcategory of $\mathbf{Rel}_{\text{Convex}}(\mathbf{C})$ we are finally able to measure distances between convex sets! The distances we pick can be rather exotic, but also the standard, familiar ones can be used. In fact,

Proposition 7.1.11. *The euclidean distance on* $\mathbb{R}^n$ *respects convexity.*

So one can measure the distance between convex sets also according to our naive geometrical intuition.

Internal monads are defined on objects, but what happens if our objects can be decomposed as a tensor of other objects in the category? Stating it differently, if we have an internal monad on every object X_i for a finite number of i, can these internal monads be lifted to an internal monad on $\bigotimes_i X_i$? This is particularly relevant for us since in Chapter 5 we defined the noun space as a big tensor of smaller domain spaces. We have an intuitive notion of distance for some of these domains, but we have no real clue of how distance between nouns could be measured. Luckily,

Proposition 7.1.12. *Let $X_1, \dots, X_n$ be objects of $\mathbf{Rel}_{(\Sigma,E)}(Q)$ such that we can choose an internal monad d_i for every X_i. Then $d_1 \otimes \dots \otimes d_n$ is an internal monad on $X_1 \otimes \dots \otimes X_n$.*

Example 7.1.13. Now we can use these results starting from the examples in Chapter 5. A noun in the food and drink example of sub-Section 5.1.1 was expressed as a subset of the tensor product $N_{colour} \otimes N_{taste} \otimes N_{texture}$. N_{colour} is just $[0,1]^3$, on which we can consider the restriction of the euclidean distance in $\mathbb{R}^3$, call it d_{colour}. Similarly, N_{taste} is just the 4-dimensional simplex, and we can take d_{taste} to be the restriction of the euclidean distance in $\mathbb{R}^4$. Finally, we take $d_{texture}$ to be the euclidean distance on $\mathbb{R}$ restricted to $[0,1]$.
Now recall that (Figure 8 may help):

$$\begin{aligned}
&\textit{banana} = \{(R,G,B) | (0.9R \leqslant G \leqslant 1.5R), (R \geqslant 0.3), (B \leqslant 0.1)\} \times \\
&\quad \times \mathrm{Cl}(\{\textit{sweet}, 0.25\textit{sweet} + 0.75\textit{bitter}, 0.7\textit{sweet} + 0.3\textit{sour}\}) \times [0.2, 0.5] \\
&\textit{apple} = \{(R,G,B) | R - 0.7 \leqslant G \leqslant R + 0.7), (G \geqslant 1 - R), (B \leqslant 0.1)\} \times \\
&\times [0.5, 1] \times \mathrm{Cl}(\{\textit{sweet}, 0.75\textit{sweet} + 0.25\textit{bitter}, 0.3\textit{sweet} + 0.7\textit{sour}\}) \times [0.5, 0.8] \\
&\textit{beer} = \{(R,G,B) | (0.5R \leqslant G \leqslant R), (G \leqslant 1.5 - 0.8R), (B \leqslant 0.1)\} \times \\
&\quad \times \mathrm{Cl}(\{\textit{bitter}, 0.7\textit{sweet} + 0.3\textit{bitter}, 0.6\textit{sour} + 0.4\textit{bitter}\}) \times [0, 0.01]
\end{aligned}$$

First things first, via the equivalence between the categories **ConvexRel** and $\mathbf{Rel}_{(\mathrm{Convex})}(\mathbf{B})$ and the injective quantale homomorphism $\mathbf{B} \to \mathbf{C}$ we consider **ConvexRel** as a subcategory of $\mathbf{Rel}_{(\mathrm{Convex})}(\mathbf{C})$.

d_{colour}, d_{taste} and $d_{texture}$ are internal monads of $\mathbf{Rel}_{(\mathrm{Convex})}(\mathbf{C})$ on their underlying objects by Proposition 7.1.10, and hence also the product $d_{colour} \otimes d_{taste} \otimes d_{texture}$ is an internal monad on $N_{colour} \otimes N_{taste} \otimes N_{texture}$ by Proposition 7.1.12. Identifying *banana*, *apple* and *beer* by their states, we can calculate now how distant these concepts are. Doing

$$\textit{banana}^{\circ} \circ (d_{colour} \otimes d_{taste} \otimes d_{texture}) \circ \textit{apple}$$

we get that the distance between *banana* and *apple* is 0: This is because both these nouns can be completely *yellow*, and hence by Proposition 7.1.9 the distance between *banana* and *apple* in N_{colour} is 0. Similarly, the distance between them in N_{taste} is again 0, because both contain t_{sweet} in the convex sets describing their tastes. Finally, both contain 0.5 in their texture domain, and hence their distance is 0 also in the *texture* domain. Since the tensor product in $\mathbf{Rel}_{(\text{Convex})}(\mathbf{C})$ is just addition of real numbers and $0+0+0=0$, the total distance between *banana* and *apple* is 0. On the contrary, doing

$$\textit{banana}^{\circ} \circ (d_{colour} \otimes d_{taste} \otimes d_{texture}) \circ \textit{beer}$$

has a different outcome: The distances between these two nouns are again 0 in the *colour* and *taste* domains, but the infimum of their distances in the *texture* domain is $0.5 - 0.01 = 0.49$. Hence the distance between *banana* and *beer* amounts to be $0 + 0 + 0.49 = 0.49$.

This measure of distance can look a bit bad, because it cannot really distinguish between *apple* and *banana*, that surely are different things. On the other hand, this could be taken as a measure of failure: The distance between *banana* and *apple* is 0 because:

- Both a banana and an apple can have the same yellowish colour. The texture of the colour may be different (spots, stains etc.), but we are only considering the colour itself;
- Both a banana and an apple can have the same texture: A particularly ripe apple can be very creamy as a particularly unripe banana can be quite hard;
- Apples and bananas taste differently, but it is very difficult to express this difference only in terms of the fundamental tastes (sweet, bitter, sour, salt). Again, the result is that according to our model a banana and an apple may eventually taste the same.

It is evident, then, that the real problem here is that our model is not taking into account some feature that really distinguishes *banana* and *apple*. On the other hand this is the case with *banana* and *beer*: The distance between these two nouns is nonzero and this is due to the evident difference in terms of consistence between them: No matter how ripe a banana can be, its texture will never be as liquid as beer is. When the distance between two nouns expressing intuitively different concepts is 0, we can then conclude that we need to take more parameters into account while casting our model.

This application of the euclidean distances prompts a final remark:

Remark 7.1.14. Given objects $X_1, \dots, X_n$ and internal monads on them $d_1, \dots, d_n$, the internal monad $d_1 \otimes \dots \otimes d_n$ on $X_1 \otimes \dots \otimes X_n$ may be different from what we expect because it heavily depends on the quantale of truth values we are considering. For instance, take $\mathbb{R}^2$: Seeing it as $\mathbb{R} \otimes \mathbb{R}$ and denoting with d^n the euclidean distance on $\mathbb{R}^n$, we can consider the internal monads d^2 and $d^1 \otimes d^1$ on $\mathbb{R}^2$. It is easy to check that $d^2 \neq d^1 \otimes d^1$ in $\mathbf{Rel}_{(\emptyset,\emptyset)}(\mathbf{C})$. In fact, $d^2(a,b) = \sqrt{(a_1-b_1)^2 + (a_2-b_2)^2}$, while $(d^1 \otimes d^1)(a,b) = (a_1 - b_1) + (a_2 - b_2)$.

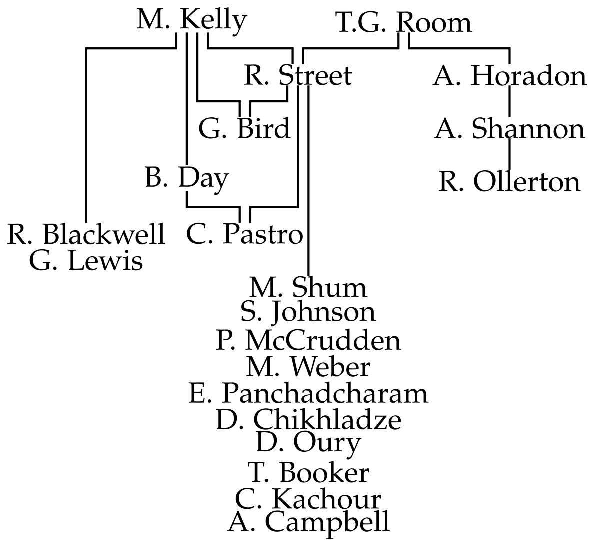


Figure 11: The mathematical family tree.

We now provide an extended example of the application of relations over generalized truth values.

Example 7.1.15 (Family Trees)**.** We will assume our universe of discourse to be the "mathematical family tree" in Figure 11, built using information about supervisor relationships freely available from the mathematics genealogy project [46]. A vertical line represents a Supervisor – PhD student relationship, with the supervisor in diagrammatically higher position. For example, T. G. Room is the supervisor of A. Horadon, Ross Street was supervised by both Room and Kelly, and Kelly supervised five different students. We will define two individuals to be "academic siblings" if they share one or more supervisors. For example M. Shun and S. Johnson are academic siblings.

What makes this family tree interesting is that there are relationships that rarely occur in ordinary genealogy trees. For instance, Bird is both an academic sibling and a student of Street. In a real family graph this would imply an unconventional relationship in which Street is both a parent and sibling of Bird. Such possibilities make the academic family tree an interesting set of relationships with nontrivial structure.

We will freely borrow terms from genealogy, saying for instance that Shum is the cousin of Shannon, or that Kelly is an ancestor of Weber. We set the following goals:

- We want to use a relational model to give meaning to sentences such as "Bird is a student of Kelly";
- If we define other genealogical relationships such as "grandparent", "cousin" or "ancestor" in the natural way, we expect these definitions to coincide with the ones obtained compositionally in our model. Ideally, "Kelly is a academic grandparent

of Shum" and "Kelly is a supervisor of a supervisor of Shum" should have the same meaning;

- We would like to express more complicated degrees of kinship, such as "Blackwell is the second-degree academic cousin of Horadon", again in a purely compositional way;
- We want this process of defining complicated relations from simpler ones to be scalable, such that it can be used on family trees of arbitrary size.

We model the compositional structure of these relationships using a very simple pregroup grammar, with only one basic type N denoting nouns. In particular, our sentence type will simply be the pregroup unit, meanings sentences will be interpreted as scalars [1] in our monoidal category. This is a rather heterodox choice: Usually, the sentence type is assumed to have non-trivial structure because we are interested in comparing the meaning of a rich space of potential sentences. In our setting however, it is not particularly interesting to compare the sentences "Ralph is the brother of Mary" and "John is the son of Mark". Instead, what we would really like is to measure *how true* the individual sentences are, ideally quantifying the degree of kinship between the people involved. We can achieve this by creatively varying our choice of truth values.

As sentence meanings are interpreted as numbers, they correspond to a single truth value. If we choose **B** as quantale for truth values, in the spirit of Montague there are only two possible choices, a sentence is either true of false. Things will get more exciting once we move to less conventional truth values, but we begin with some simple examples.

Taking **B** as our quantale, we define the following relation on sets pointwise in the obvious way:

$$C(x, y) = x \textit{ is the academic child of } y$$

$C(x, y)$ is $\top$ if x is a child of y, and $\bot$ otherwise. We can build many interesting academic relationships out of the child relation C, for example:

$$\begin{aligned}
S &= (C^\circ \circ C) \setminus 1_N && \textit{the sibling relationship} \\
P &= C^\circ && \textit{the parent relationship} \\
G &= P \circ P && \textit{the grandparent relationship} \\
K &= P \circ S \circ C && \textit{the cousin relationship}
\end{aligned}$$

Where 1_N is the usual identity relation on our nouns, that can be explicitly written as

$$1_N(x, y) = \begin{cases} \top \text{ if } x = y \\ \bot \text{ otherwise} \end{cases}$$

1 Recall scalars are morphisms of type $I \to I$.

We can interpret our various family tree relations as simple verbs, as illustrated in Figure 12. We drew the sentence space as a dashed

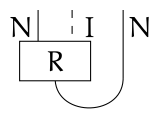
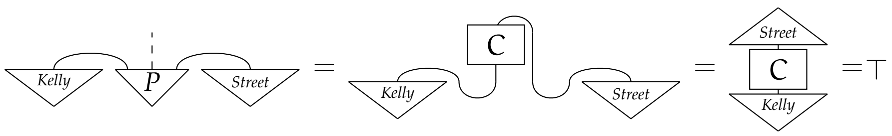


Figure 12: Relations are interpreted as simple verbs.

wire, as it is actually the monoidal unit and it would not normally be explicitly drawn according to our formalism. A simple graphical calculation establishes that "Kelly is a parent of Street", as follows:



Similar calculations show that "Shum is the cousin of Shannon", whereas "Shum is the cousin of Street" is false. More surprisingly, Pastro is his own cousin!

So far, so good. We now move to more expressive truth values that will allow us to quantify "how related" two individuals in the hierarchy are.

Definition 7.1.16. The *step quantale* **N** is given by the extended natural numbers $\mathbb{N} \cup \{\infty\}$ with the *reverse* ordering. Joins are minima and we take addition as the monoid multiplication. This can be seen as a discrete version of the Lawvere quantale **L**.

As expected, we now use $\mathbf{Rel}(\mathbf{N})$ as our semantics. In this case, we re-define C as follows:

$$C(x, y) = \begin{cases} 1 \text{ iff } x \text{ is directly below } y \\ \infty \text{ otherwise} \end{cases}$$

We then define the parent, grandparent and cousin relations as we did before. The sibling relation S is defined as $P \circ C$. It is easy to see how our truth values represent the degree of kinship between our individuals: A parent-child relation between x and y can assume value one or ∞, depending if it is satisfied or not according to our tree. The sibling relationship S can have value two or ∞: We are considering being a sibling as a more distant relationship than parenthood. Although slightly surprising at first sight, this observation makes sense from an heraldry perspective, where the parent-child relationship is considered to be stronger than that of siblings. If two individuals are cousins, the degree of kinship will be 4, and so on. The strongest degree of all, zero, can only be attained by the identity relation, corresponding to "being oneself". Note how in this framework an individual can be considered "their own sibling" but, in doing so, this relation will be satisfied only with value two, while considering an individual as "oneself" attains value zero.

The impact of using the truth values in the quantale **N** is most pronounced when we consider relations such as "ancestor" and "relative". In order to do so, we extend the notion of transitive closure to relations over a quantale. Firstly, we define for a relation $F : X \to X$:

$$F^1 = F \quad \text{and} \quad F^{n+1} = F \circ F^n$$

The transitive closure can then be defined as the relation:

$$\overline{F}(x,y) = \bigvee \{F^n(x,y) \mid n \geqslant 1\}$$

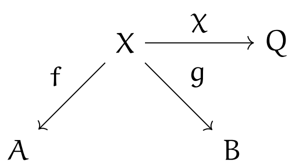
The ancestor relation A is the transitive closure of the child relation C. The value $A(x,y)$ is lowest number of child relation "steps" from x to y, returning ∞ when x is not an ancestor of y.
The relative relation R is slightly more complex, we define it using the transitive closure as follows:

$$R = \overline{P \cup C \cup 1_N}$$

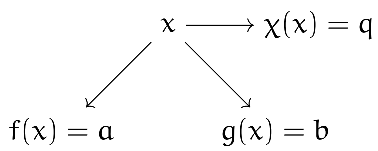
$R(x,y)$ is the shortest number of steps between x and y assuming that we can travel in in either direction, and that we can always reach ourselves in zero steps.

7.2 PROOF RELEVANCE

In Section 6.3 we interpreted spans as the proof-sensitive counterpart of relations: A span not only tells us if two elements are related, but also how related they are. Here we push further these investigations and, to do this, we need to refine our notation a bit. We denoted algebraic Q-spans as (X, f, g, χ), that stands for:



This notation is wonderful when we want to talk about spans in abstract terms to prove categorical properties, but it gets a bit in the way when we need direct access to elements, for example to identify "who-is-what" in practical applications as we want to do here. We solve this problem using the notation $S^q_x(a,b)$ to refer to:



that is, to say that elements a, b are related by x with strength q.

We moreover proved in Theorem 6.4.8 that also algebraic Q-spans are enriched over preorders, that are again categories. It makes sense then to ask what internal monads are in the algebraic Q-span case. We have the following results: Specializing the definition of internal monad to the span case, we get that an algebraic Q-span $(X, f, g, \chi) : A \to A$ is an internal monad if:

$$(X, id_X, id_X, \chi_k) \subseteq (X, f, g, \chi)$$
$$(X \times_A X, f \circ p_1, g \circ p_2, \odot \circ \chi \times \chi \circ \langle p_1, p_2 \rangle) \subseteq (X, f, g, \chi)$$

We recall that the order relation $\subseteq$ is defined via a $\mathcal{E}$-momorphism that in order induces a monomorphism of algebraic Q-spans. This is just a $\mathcal{E}$ morphism $A \to X$ for the first line and some $\varphi : X \times_A X \to X$ for the second.

Using our new notation we can see how this acts on elements directly:

$$S^k_x(a, a) \tag{11}$$
$$\text{if } S^p_x(a, b), S^q_y(b, c) \text{ then } \exists r : S^{r \leqslant p \odot q}_{\varphi(x,y)}(a, c)$$

As usual, now we try to understand what this means specializing the definition to our canonical choices of quantales[2]:

Example 7.2.1. An internal monad on A in **Span**(**C**) is an **C**-span $S : A \to A$ such that if $S^p_x(a, b)$ and $S^q_y(b, c)$ we can choose an element $\varphi(x, y)$ of the apex such that $S^r_{\varphi(x,y)}(a, c)$ 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.

Example 7.2.2. An internal monad on A in **Span**(**F**) is an **F**-span $S : A \to A$ such that if $S^p_x(a, b)$ and $S^q_y(b, c)$ we can choose an element $\varphi(x, y)$ of the apex such that $S^r_{\varphi(x,y)}(a, c)$ and $\max(p, q)$ 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 **C** and **F**-span monads further generalize metric and ultrametric spaces to incorporate multiple possible distances, which we can think of as describing different paths between points.

Example 7.2.3. Note that spans on the Boolean quantale do not make much sense if we interpret the truth values of **B** as "connected/not connected". A proof witness already tells us that a couple of elements are connected, so what sense could something like $S^{\perp}_x(a, b)$ have? Things start making much more sense when we consider **I**, interpreted as the fuzzy version of **B**.

2 Yes, we just need partially ordered monoids in the algebraic Q-spans case, but as a matter of fact the four quantales we consistently used as a case study happen to be the interesting ones also for spans. Clearly we do not rely on their completeness as join-semilattices since the quantale join is never used when working with spans.

An internal monad on A in $\mathbf{Span}(\mathbf{I})$ is an $\mathbf{I}$-span $S : A \to A$ such that if $S^p_x(a, b)$ and $S^q_y(b, c)$ we can choose an element $\varphi(x, y)$ of the apex such that $S^r_{\varphi(x,y)}(a, c)$ and $\min(p, q)$ is smaller 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.

We now outline a new practical application of spans in models of language.

Example 7.2.4 (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 [122], 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 [135] to yield a new category of ambiguous meanings. When it comes to relational semantics it makes sense to try to do the same, but unfortunately on **Rel** the CPM construction does not provide the needed results and hence is not a good way to treat semantic ambiguity [103].

Luckily enough, in the spans framework different proof witnesses can tell us not only how much, but also how words are connected. In our example, we can consider how the ambiguous word "bank" is related to the word "water", and provide different proof witnesses to address different contexts:

- One proof witness stands for the "river bank" context, where we would expect a strong relationship between "water" and "bank";
- Another proof witness stands for the "financial bank" context, and in this case we would expect the connection between "water" and "bank" to be much lower in intensity.

By choosing our quantale of truth values to be the Lawvere quantale $\mathbf{C}$, 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.

Example 7.2.5 (Proof Relevant Family Trees)**.** We return to the family tree Example 7.1.15, this time formalizing our semantics in $\mathbf{Span}(\mathbf{N})$. The intuition for such a span $(N \xleftarrow{f} X \xrightarrow{g} N, \chi)$ is that an element $x \in X$ witnesses a path from $f(x)$ to $g(x)$ of length $\chi(x)$. For example, we can introduce a span C describing the child relationship, admitting a path from a to b of length 1 if and only if a is a child of b. The parent span P is the converse of the child span, given by reversing its legs. A composite of two spans encodes composites of compatible paths and the sum of their corresponding lengths. The sibling span is the composite $P \circ C$, illustrated in Figure 13. If a and b are siblings,

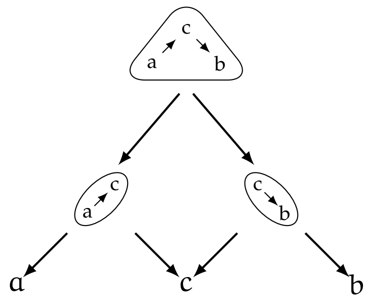


Figure 13: Interpretation of pullbacks as composition of paths.

they must have some common parent c, resulting in a length two path $a \to c \to b$, as illustrated in Figure 13. If the pair a and b have two different common parents, in contrast to the case of relations where this information is lost, the composite span will record two distinct paths between them.

Similarly, if we generalize the ancestor relation to a span, it would witness every possible way of relating two members of the family tree, and record the corresponding path length. In this way, we would explicitly record that Bird is related to Kelly in two distinct ways, directly in one step, and via Street in two steps.

Clearly, if one were to drop the definition of the span category in terms of isomorphism classes of spans and go for a bicategorical formalization, we would be able to distinguish between *actual paths*: In fact, more than reasoning with paths, at the moment we are reasoning with *kinds of paths*, that is, we consider two isomorphic paths to be the same. This amounts to say that we are grouping group witnesses in classes, and we are not considering them singularly. This more precise (and probably more satisfying) approach will surely be object of future work, also because the formalization of spans is naturally bicategorical, and mathematically more pleasing, albeit more involved: Categorically, modelling spans as bicategories just feels like the right thing to do.

7.3 VARIABLE CONTEXTS

Finally we further explore our last axis of variation, namely the choice of universe of discourse, represented by a topos. Following the convention already introduced in Chapter 6 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. There are, obviously, some toposes that are more interesting than others with respect to what we aim to model, **Set** being probably the most evident example. But are there other choice that it makes sense to consider? Well, if this were not the case all the hassle to formalize things in such general terms would have been for nothing. The first natural alternative to take into consideration is given by presheaf categories.

Definition 7.3.1. 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 $j \leqslant i$ under X.

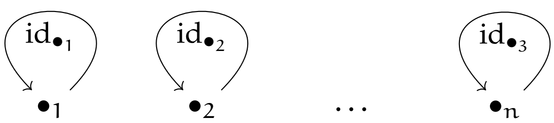
Presheaves are not only a logical, but also a conceptual direct generalization of sets since they can be interpreted as *sets varying with context*. This is exactly the perspective we shall adopt in our examples. It is useful to give a pictorial depiction of what presheaves are to make our intuition more precise.

Example 7.3.2. (Trivial presheaf) The simplest case we can think about is the one with $C = \mathbf{1}$, the category consisting of one object and one identity arrow. We will denote the only object of $\mathbf{1}$ with $\bullet$. $\mathbf{Set}^{\mathbf{1}^{op}}$ has then as objects functors $F : \mathbf{1} \to \mathbf{Set}$, that is, functors that pick one object, namely $F(\bullet)$, in $\mathbf{Set}$, along with their identity arrow. It is clear that objects of $\mathbf{Set}^{\mathbf{1}^{op}}$ are just sets. Morphisms are natural transformations between such functors, that is, arrows that make this diagram commute:

$$\begin{array}{ccc} F(\bullet) & \xrightarrow{\quad \tau_\bullet \quad} & G(\bullet) \\ \Big\| & & \Big\| \\ F(\bullet) & \xrightarrow{\quad \tau_\bullet \quad} & G(\bullet) \end{array}$$

So, since $F(\bullet), G(\bullet)$ are just sets and the commutative diagram is trivial, the only requirement we are asking for is that $\tau_\bullet$ has to be a morphism in $\mathbf{Set}$, hence a function. This means that $\mathbf{Set}^{\mathbf{1}^{op}}$ is isomorphic to $\mathbf{Set}$ as one would expect.

Example 7.3.3. (Multiple copies) This time, we pick $\mathcal{C}$ to be $\mathbf{1}^n$, the category with n objects in which the only morphisms are the identity arrows for every object.



As one would expect, a functor from $\mathbf{1}^{n^{op}}$ to $\mathbf{Set}$ just picks n different sets, hence we can identify it with a sequence $(X_1, \ldots, X_n)$. A natural transformation between two functors F, G is determined by functions

$\tau_{\bullet_1}, \ldots, \tau_{\bullet_n}$ such that the following diagram commutes, for every i and f:

$$\begin{array}{ccc} F(\bullet_i) & \xrightarrow{\tau_{\bullet_i}} & G(\bullet_i) \\ {\scriptstyle F(f)}\downarrow & & \downarrow{\scriptstyle G(f)} \\ F(\bullet_i) & \xrightarrow{\tau_{\bullet_i}} & G(\bullet_i) \end{array}$$

Note, however, that the only f allowed in the diagram above are identities, since $\mathbf{1}^n$ does not possess any other arrow. This means that, again, the only requirement for every $\tau_{\bullet_i}$ is to be a function. A natural transformation between F and G is then a set of functions such that $(X_1, \ldots X_n) \xrightarrow{\tau_{\bullet_1} \times \cdots \times \tau_{\bullet_n}} (Y_1, \ldots, Y_n)$, showing that $\mathbf{Set}^{\mathbf{1}^{n^{op}}}$ is just $\mathbf{Set}^n$. We represent this pictorially as in Figure 14. In this figure, we have three functors and two natural transformations between them. Every sequence of images through a functor (eg. $F(\bullet_1) \ldots F(\bullet_n)$) is an object in $\mathbf{Set}^{\mathbf{1}^{n^{op}}}$ and every natural transformation (τ, as an instance) is a morphism in $\mathbf{Set}^{\mathbf{1}^{n^{op}}}$.

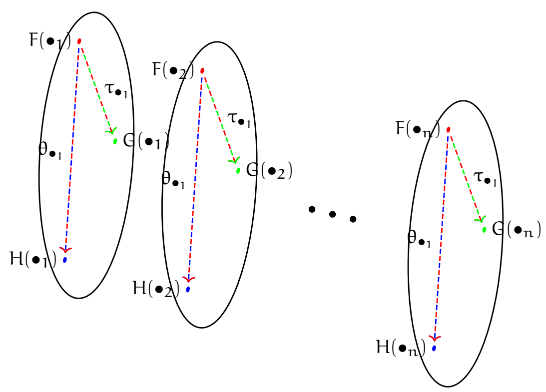


Figure 14: Morphisms in $\mathbf{Set}^{\mathbf{1}^{n^{op}}}$.

Example 7.3.4. The straightforward generalisation of what we did is considering presheaves on a little more complicated category, namely $\cdot \to \cdot$. This category has two objects, again called $\bullet_1$ and $\bullet_2$, and three morphisms, of which only one is non-trivial, the arrow $\bullet_1 \to \bullet_2$ that we will call f. The pictorial representation of a couple of morphisms in $\mathbf{Set}^{(\cdot \to \cdot)^{op}}$ is then just as in Figure 15, where the naturality of τ, σ guarantees that the parallelograms in the picture commute. We can interpret presheaves on $(\bullet_1 \to \bullet_2)$ as *sets varying over time*: $\bullet_2$ represents the state "before" while $\bullet_1$ the state "after".

Example 7.3.5. Generalizing a little more, we consider presheaves on $(\cdot \to \cdot \to \cdot)$. In this case the non trivial arrows are three, the one from $\bullet_1$ to $\bullet_2$ (call it f), the one from $\bullet_2$ to $\bullet_3$ (call it g) and their composition, see again Figure 15.

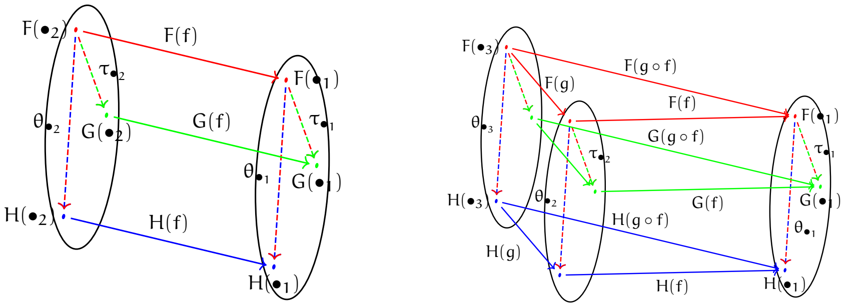


Figure 15: A couple of morphisms in $\mathbf{Set}^{(\cdot\to\cdot)^{op}}$ and in $\mathbf{Set}^{(\cdot\to\cdot\to\cdot)^{op}}$, respectively.

Now that we can visualize what a presheaf is we can go back to our models. To continue, though, a question must be answered: How do algebras, internal commutative partially ordered monoids and internal quantales look in a presheaf topos? Yes, from the point of view of the internal language they look exactly as they do in **Set**, but this is not what we need to put this stuff to good use! An object in a presheaf topos is a functor $\mathcal{C} \to \mathbf{Set}$, so what does it mean, for example, that a given presheaf is an internal quantale? How do we put a quantale structure on a functor?! The idea would be to select a presheaf that somehow represents something we are interested in, such as a truth values structure, and then prove that this is indeed an internal quantale/commutative partially ordered monoid in the internal language, allowing us to use it as we normally would in the **Set** case. Solving this problem amounts to connect the internal and external description of an object, and in general it is not an easy task: Working externally or in the internal logic separatedly is manageable, but working with both at the same time is a bit of a mess, since topos theory was in fact conceived with the idea that "if you have the internal language you work with it and forget about the external description of things".

The most important result we will use here can be found in [84, Theorem D1.2.14], and says that

Theorem 7.3.6. *If* $\mathbb{T}$ *is a* geometric theory, *then the categories:*

$$\mathrm{Mod}_{\mathbb{T}}([\mathcal{C}^{op}, \boldsymbol{Set}]) \qquad \textit{and} \qquad [\mathcal{C}^{op}, \mathrm{Mod}_{\mathbb{T}}(\boldsymbol{Set})]$$

Are isomorphic.

Where we used the alternative notation $[\mathcal{C}^{op}, \mathbf{Set}]$ for $\mathbf{Set}^{\mathcal{C}^{op}}$. What this theorem says is that if $\mathbb{T}$ is a "sufficiently well-behaved" theory then a model of the theory in the presheaf category can be seen as a presheaf on the models of $\mathbb{T}$ in **Set**. It doesn't really matter what a geometric theory is here, everything we have to know is that:

- Algebras on a finite signature (Σ, E) are geometric theories;
- Commutative partially ordered monoids are geometric theories;

- Unfortunately, quantales are NOT geometric theories.

At least from the first couple of points above we can get an useful corollary of Theorem 7.3.6:

Corollary 7.3.7. *A commutative partially ordered monoid in a presheaf category* $\mathbf{Set}^{\mathcal{C}^{op}}$ *is a presheaf* $Q : \mathcal{C}^{op} \to \mathbf{Set}$ *such that for each* $\mathcal{C}$*-object* x *and* $\mathcal{C}$*-morphism* f*,* $Q(x)$ *is a commutative partially ordered monoid and* $Q(f)$ *is a commutative partially ordered monoid morphism in* **Set***.*

Similarly, an internal (Σ, E)*-algebra in a presheaf category* $\mathbf{Set}^{\mathcal{C}^{op}}$ *is a presheaf* $A : \mathcal{C}^{op} \to \mathbf{Set}$ *such that for each* $\mathcal{C}$*-object* x *and* $\mathcal{C}$*-morphism* f*,* $A(x)$ *is a* (Σ, E)*-algebra and* $A(f)$ *is a* (Σ, E)*-algebra homomorphism in* **Set***.*

. . . and hence we can model these objects in the presheaf simply assembling them from **Set**. These ingredients are everything we need to talk about algebraic Q-spans, while the relational case, because of the quantale issues, will be trickier to study. For this reason we for once break our tradition and approach the span case first.

Example 7.3.8 (Temporal dependence)**.** In Example 7.2.4 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 $\mathbb{N} = 0 \leftarrow 1 \leftarrow 2 \leftarrow \dots$ 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 $\mathbf{Span}^{\mathbf{Set}^{\mathbb{N}^{op}}}_{(\emptyset,\emptyset)}(\mathbf{C})$, where $\mathbf{C}$ is the constant presheaf on the pomonoid underlying the Lawvere quantale. A $\mathbf{C}$-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 $\chi : X \Rightarrow \mathbf{C}$. 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 must be preserved through time. Intuitively, in this model, a steadily increasing collection of relationships hold over time: As time increases, we can do at most two things:

- Add proof witnesses, to which arbitrary truth values can be assigned (compatibly with our requirements for algebras);
- Collapse proof witnesses. We can collapse proof witnesses x, y at time t to a unique proof witness z at time $t + 1$ only if they have the same truth value p. To do this, though, we must pay attention to what happens to the elements connected by x, y. What is required to happen can be easily represented in pictures. Every square made of arrows in Figure 16 has to commute, since a span in our setting is given in terms of natural transformations and not just morphisms. The consequence of

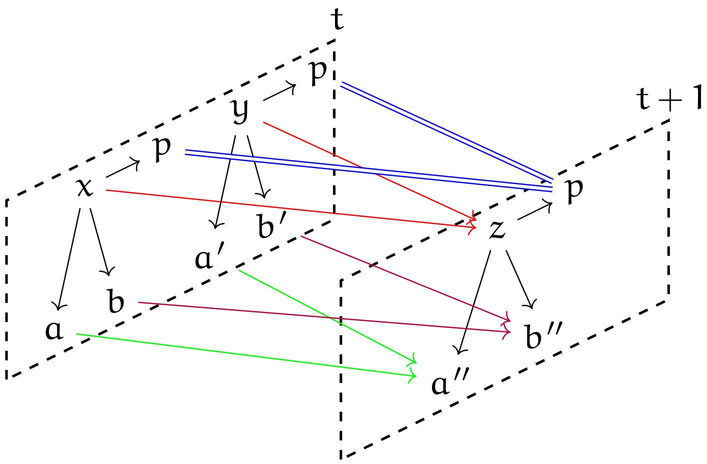


Figure 16: Collapsing proof witnesses. Every square of arrows has to commute.

this is that if we collapse a couple of proof witness we have to take care also of what happens to the elements they relate. In Figure 16 collapsing x and y forces us to collapse the couples a, a' and b, b' to common elements a'', b'' as time increases.

Example 7.3.9 (Perspective Dependence)**.** In Example 7.3.8, 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 $[0, 1]$.

Example 7.3.10 (Belief hierarchy)**.** The previous example can be also interpreted under the light of belief revision. Consider $\mathbf{Set}^{\mathcal{C}^{op}}$; as we already saw, if there is an arrow $t \to t'$ in $\mathcal{C}$, then given some arbitrary span the presence of a proof witness at stage t' implies the presence of a proof witness at stage t, in a way such that conditions that relate their respective truth values hold. On the contrary, at stage t we can add more proof witnesses in a way that is more or less independent from what happens at t'. If we use the span formalism to encode beliefs (that is, in $S^r_x(a, b)$ the parameter r represents how much an agent believes a, b are related by x), then we can use the presheaf model to formalize how much a given belief is important for a given agent. Having a, b related through x at stage t' is stronger that having them related at stage t, since changing the way they are related at stage t' (revising a belief) automatically implies that we have to revise also at t.

If we want 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 $\mathbf{Set}^{\mathcal{C}^{op}}$ (proof in A.0.3). A presheaf on a groupoid is somehow trivial: It describes contexts that can be arranged into disconnected blocks, while moving within a given block is a totally fluid operation, being all the morphisms invertible;
- The subobject classifier Ω of a topos is an internal locale, and therefore an internal commutative quantale. This is particularly interesting since Ω represents internal truth values structure of our underlying logic. Choosing Ω in our algebraic Q-relation construction then is like choosing the booleans to construct relations between sets: In the latter case, we are using "true" and "false" as the only truth values, according to our intuition in classical logic (on which set-theory is based). This seems to be indeed a natural choice, that directly generalizes to Ω in the arbitrary topos case: Using $\mathbf{Rel}^{\mathbf{Set}^{\mathcal{C}^{op}}}_{(\Sigma,E)}(\Omega)$ elements can be related only as Ω prescribes, that is, according to the notion of truth that comes with our choice of universe of discourse.

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. Given what we just said, it should be unsurprising that this relationship will involve the internal locale given by the subobject classifier.

Definition 7.3.11. 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 $\mathbf{Rel}(\mathcal{C})$ on a regular category $\mathcal{C}$, see for Example [26]. For the category **Set** for example, this construction recovers exactly the usual category of binary relations. As we have been constructing categories of relations up to this moment, it would be interesting to know how this relates to the relations on 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 [19], 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 7.3.12. *Let $\mathcal{E}$ be a topos, Ω its subobject classifier and (Σ, E) an algebraic signature. The category $\boldsymbol{Rel}^{\mathcal{E}}_{(\Sigma,E)}(\Omega)$ resulting from the algebraic Q-relations construction is isomorphic 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 "in the usual sense" are a special case of our construction.

Part III

DOWNSIDES

Here we analyze the downsides of our model, drawing considerations in particular from ancient and/or exotic languages.

8 COMPOSITIONALITY IN MIDDLE EGYPTIAN

Et manebant structis molibus litterae Aegyptiae
priorem opulentiam complexae.

— *Tacitus, Annales [126, Book II, 60]*

Since the moment we highlighted how, in categorical models of meaning, the functorial link between grammar and semantics works, pregroups have been often taken for granted, given that they model well many characteristic features of the English language (that is, of course, the language on which research efforts are usually focused). As we observed, less fortunate has been the semantic side of this paradigm, and the effort to characterize meaning in the most different ways, for instance trying to model traditionally "difficult" words as pronouns or prepositions, has often resulted in the definition of new compact closed categories to employ as semantics. This second consideration is, obviously, what prompted the writing of the last hundred-or-so pages.

If – as it is in our case – our main goal consists in having a better understanding of how the mind works, we should then embrace the cognitive point of view also with regard to the grammatical side of our framework. In particular it makes sense to ask, hence, if pregroups represent a good model for grammars of any language and if the way "wires are drawn in our heads" really follows pregroup-like rules.

In this chapter we want to highlight how this is not the case for languages with a high level of ambiguity that exhibit, nevertheless, a high level of compositionality. If compositionality makes us think that the categorical paradigm defined in the previous chapters could provide a good cognitive model of meaning, the continuous necessity to disambiguate statements depending on context makes the now familiar "pregroup $\to$ semantics" approach difficult to adopt, for at least a couple of reasons:

- At the moment, pregroup types represent grammatical entities motivated from linguistic considerations. Types are defined to represent nouns, adjectives and verbs, and these entities do not originate from cognitive considerations. We will not be concerned with this issue, utterly interesting per se, until the next chapter. We may even suppose, at this point, that this is not a real problem: One may argue, in fact, that even if our models for language are based on cognition, we are still trying to assign meanings to linguistic entities, and for this reason a relation of sort with grammar cannot be entirely superseeded by a cognitive approach.

- Even if we were to agree with this consideration, a very big problem remains, and is given by the fact that pregroup reductions are always linear: Given a string of types, we can only reduce adjacent elements in the string. This amounts to say that the wiring prescribed by pregroup reductions does not accomodate for wires that cross each other. If this a big but somehow fixable problem in English, it becomes much worse for highly non-linear languages, where having words referring to each other scattered around, requiring consistent wire crossing, is the norm. An application-oriented approach would not be shattered at all by this consideration: English is the most used language on the planet, hence the one that more than any other needs to be modeled; highly non-linear languages, being totally marginal for applications, do not constitute a problem on which one wants to spend time to solve it. But as the goal of this document is trying to model the way the mind works, we cannot be excused: What we are doing should, at least in principle and with good approximation, work for any language. The aim of this chapter is to show that this is not the case.

The language we will consider as a case study is Middle Egyptian. We will briefly describe some of its interesting features in the following section. Then, we will point out where the real problems arise in Section 8.2.

8.1 THE EGYPTIAN LANGUAGE

The Egyptian language, commonly known as *Hieroglyphic*[1], is an afroasiatic language used in ancient Egypt for more than four and a half millennia. Its history is roughly divided in four phases: Old Egyptian (3000 – 2135 B.C.E.), Middle Egyptian (divided itself into classical and post-classical, 2135 – 2000 and 2000-1300 B.C.E. respectively), Late Egyptian (1350 – 715 B.C.E.), Demotic (715 – 470 B.C.E.) and Coptic (from 3_{rd} to 16_{th} centuries C.E, still spoken in a revived version by the Coptic Church) [78]. In particular Middle Egyptian, the language we will focus on in this paper, is considered to be the "classical" Egyptian language, having defined a literary standard that lasted in the region until the third century C.E. For this reason, it is often considered the most interesting by Egyptologists.

Middle Egyptian used two writing scripts: Hieroglyphic, undoubtedly the most famous, was mainly used in stone inscriptions. *Hieratic*,

1 The word *Hieroglyphic* technically denotes one the scripts used to write Middle Egyptian, and not the language itself.

that will not be covered in this paper, was developed in parallel with Hieroglyphic and is closely related to it. This was a cursive writing system mainly used to write on papyrus, allowing scribes to save the large amount of time needed to draw complex logograms. It is believed that written Middle Egyptian resembled closely its spoken form [78].

The most well known characteristics of Egyptian are the use of *logograms* and the *rebus principle*. Logograms are symbols that represent words: For instance, 𓉐𓏤 (pronounced "pr") stands for "house", and is in fact a stylized drawing of a houseplant. This system allowed for an easy representation of simple concepts, but posed challenges when less concrete words, like prepositions or abstract nouns as "justice", "knowledge" etc. were to be written. The way Egyptian deals with this issue is called rebus principle. Since vowels were basically not written, sounding-alike logograms (that is, logograms having the same root consonants) were used in the written form. Providing some context, if 𓂋 (pronounced "r") was the logogram standing for "mouth", it could also be used to represent the preposition "to". Usually, a stroke was added to the logogram to distinguish the logographic interpretation from the phonetic one. In this case, 𓂋𓏤 stands for "mouth" and 𓂋 for the preposition "to".

The rebus principle, though, was often not enough to represent many concurring concepts having the same phonetic interpretation without ambiguity. For this reason, semantic determinatives where used. For example, if 𓂋𓂝 (pronounced "r‘"[2]) were the root consonants appearing both in the words "Day" and "Ra"(the Egyptian solar god), the two were written 𓂋𓂝𓇳 and 𓂋𓂝𓀭 (or equivalently 𓂋𓂝𓁛) respectively, where 𓇳 is the determinative for things related to the sun and 𓀭 the one used to denote gods and the like (𓁛 is the determinative explicitly related to the god Ra). If one is acquainted with the theory of conceptual spaces, it could be argued that conceptual spaces are somehow already embedded in the language itself in the form of determinatives, each one denoting in which category of though a group of phonemes has to be mapped.

Middle Egyptian is, moreover, a highly non-inflected language. It lacks cases and articles, possesses two genders (masculine and feminine), three persons (singular, plural and dual, albeit the latter one is quite rare to find in texts) and the verbal constructions are not many

2 There are very few cases in which we know the exact vowels composing the pronunciation of a word, mainly due to transliterations in other languages: This often happens with names of important people and places. Nearly always, though, we can be sure only about the consonants. When this happens, an "e" is conventionally added between consonants to make the pronunciation easier: For example, 𓄤𓆑𓂋, meaning "beautiful" and spelled "nfr", is conventionally pronounced "nefer". The real pronunciation could nevertheless have been radically different, as in "nafar". Recostructing pronunciation of words is a big area of research in the study of Egyptian language.

compared with the majority of European languages. For this reason, word order is of the utmost importance for inferring the correct meaning of sentences, in a very similar fashion to what happens in English with transitive verbs: "John eats the lunch" and "The lunch eats John" have quite different meanings. This behavior is though pushed to the extreme, as we will shortly see. The fundamental order of the Egyptian sentence[3] is

(iw) + verb + subect [+ modifiers] + object [+ modifiers] + complements + modifiers of the sentence

Here, the modifiers in square brackets refer to subject and object, respectively. These can be adjectives or even entire sentences, and behave like adverbs. Complements are considered modifiers of the verb itself, and go after the object. Finally, sentences related to the main one are considered to be modifiers of the sentence itself. The general rule is then that modifiers, let them be complements, specifications, adjectives et cetera, get always postponed to the word they modify. These consideration should highlight how Middle Egyptian is highly compositional and, up to some extent, "regular" in the way it works.

We briefly mentioned the lack of inflection. This is particularly evident from the absence of a specific construction for comparatives and superlatives: The formers are obtained with the use of the preposition , meaning "in comparison to":

(ʿ3 ṯsm·tn r ṯw)

This dog is bigger than you

Literally: *Big [is] this dog in comparison to you*

The latters are obtained recursively by means of a genitival construction. Something similar happens also in English in localized instances, as in "King of kings" when used to denote a great ruler. For example, if (pronounced "ʿ3") means "big" (the book roll glyph here is the determinative for abstract concepts), then "biggest" is just (ʿ3 ʿ3w), literally "big one of the big ones". Here the wave logogram between and its plural form, obtained using the plurality symbol given by three vertical strokes, denotes a genitival construction that gets translated with the preposition "of" in English. A more articulated example that provides some context is the following passage from the famous tale of the Eloquent Peasant [B53]:

(Imi-r pr wr, nb·i, wr n(y) wrw ...)

High Steward, my lord, greatest of the greats...

Literally: *High Steward, lord of mine, great one of [the] great ones...*

3 Here, we are in particular referring to the most diffused word order in statements of fact, to which subordinates are attached.

A very interesting feature of the language is that many constructions (the huge majority of them, to be honest) are *adverbial* in nature: Adjectives for example, even if intended to be modifiers of the noun, get stacked after it into an adverbial construction: In the previous example, "my lord" is an adverb modifying "high steward", and "this" is an adverb modifying "dog". Once one becomes familiar with this way of reasoning, the uselessness of relative pronouns becomes clear since relative sentences can be used as modifiers of the noun they refer to. Consider, for instance, the following example, that is a simplified version of a passage of the fourth story[6,26] in the Westcar papyrus:

(iw rh·n·i nds hmsi·f m Ḏd-Snfrw . . .)

I have learned of a commoner who lives in Djed-Snofru. . .

Literally: *[Statement of fact:] I learned* [*of a*] *commoner*, *lives he in Djed-Snofru. . .*

In the very same way, this system is also applied to whole phrases, where coordinate sentences are just modifiers of the principal one and get stacked after it. As a result, the relation between sentences is not explicit: One sentence just serves as a modifier for another. It has to be noted that, in fact, Egyptian generally lacked words as "and", "or" and the like. This, along with the high propensity to center-embedding and the lack of punctuation probably makes learning how to correctly distinguish one sentence from the other the most difficult task for the translator.

8.2 PROBLEMS

There are many problems arising when one tries to model Middle Egyptian using the pregroup framework. Nouns, adjectives and verbs roughly work as in English, so we will be quite brief about it: Nouns will typically have type n, adjectives will have type $n^r \cdot n$ since they are postponed to the word they modify as we mentioned in Section 8.1, while transitive and intransitive verbs will generally have types $s \cdot n^l \cdot n^l$ and $s \cdot n^l$ respectively, since they are found at the beginning of the sentence, in our case immediately following 𓇋𓅱.

The particle 𓇋𓅱 is tricky: This is usually employed to denote a statement of fact, and is found at the beginning of the sentence when it serves this purpose. According to Gardiner though [71] this particle is originally a verb, and can be modelled as $s_f \cdot s^l$ (it converts an unspecified sentence into a statement of fact).

This is, alas, where the "easy part" ends. The following observations are the ones making Middle Egyptian so problematic as a language:

- There is a nearly-total absence of sentence coordination. Words like *and*, *or*, *because* do exist, but are seldom used and, in fact, in translation one has often to provide them to make the translated sentence understandable. This highlights how for Egyptians correlation between sentences did not play a fundamental role to assign meaning: Saying "I won't go to the university (if, because, since, ...) I am ill" was perfectly acceptable. The total irrelevance of logical connectives in the language sounds incredible for western researchers, considering how endowing mathematical models with a decent interpretation of such words has always been considered a central and very hard task by the research community: Finding out that in some languages and cultures people can't be bothered at all even with representing such words comes as a shock! This "we can't be bothered at all" attitude will assume frightening magnitude when we will analyze peculiarities of the Pirahã language in Chapter 9. This problem is particularly felt in categorical models of meaning, where one doesn't just rely on grammar but wants to actually link grammatical meaning with a semantic meaning in some category.

- The previous observation is worsened by the fact that Hieroglyphic script lacked any form of punctuation, and as we mentioned before coordination must essentially be guessed. This makes the automatic assignment of meanings from a corpus of text incredibly hard, since the pregroup grammar won't have any way to understand by itself when a sentence is finished and another is starting. This is a common problem in many models of grammar, and it is not limited to pregroups;

- Sentence order is not stable, since pronouns tend to stick to the verb they refer to. For instance, even if the standard order is 𓇋𓅱 + verb + subect + object, this becomes 𓇋𓅱 + verb + object + subject when the object is a pronoun and the subject is not. This problem is particularly bad in models of grammar that rely on a correspondence between words and types, where the types depend on the position that the word has in the sentence. In pregroup grammar this is surely the case;

- Meaning-assignment in Egyptian is highly non-linear, and works backwards. Consider, for instance, the sentence in Figure 17. As it stands, the following translations are all deemed as acceptable:

  RELATIVE *The man that makes good things speaks.*

  FACT *The man says he makes good things.*

  FACT *The man said he made good things.*

  CAUSAL *The mans speaks because he makes good things.*

(iw dd̠ s ir·f nfr·t)

Literally: *[Statement of fact:] says [or speaks] [the] man makes he good [things]*

Figure 17: An ambiguous sentence in Middle Egyptian.

TEMPORAL *The man speaks while he makes good things.*

CONDITIONAL *The man that will make good things will speak.*

Here we can see three different levels of ambiguity acting at the same time: First there is a lexical ambiguity, coming from the fact that both the verbs "to speak" and "to say" are written in the same way. Second, we do not know if the personal pronoun refers to *the man* or stands as it is, referring to another person not mentioned in the sentence. Finally, the Egyptian verbal system doesn't allow us to place this sentence into a definite time: It could be in the past, in the present or in the future, since the tense system is a relative one: Tenses are expressed not in an absolute fashion, but in a relative one; it is possible to understand if some part of the sentence is happening before, at the same time or after some other part, but there is no way to collocate the whole sentence in a temporal frame if an explicit time reference is not given. For example, in the sentence:

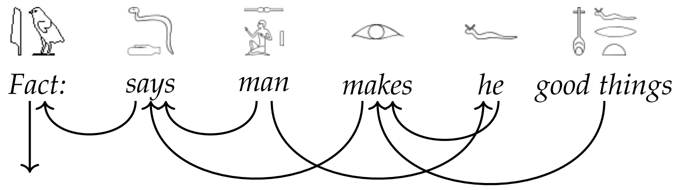
(iw dd̠ s ir·n·f nfr·t)

Literally: *[Statement of fact:] says [or speaks] [the]man makes [past] he good [things]*

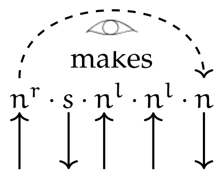
The verb "to make" has been put at the past tense using the suffix , and the sentence may be translated as *Having made good things, the man spoke*, as *The man speaks because he made good things*, or as *When the man will have made good things, he will speak*.

Middle Egyptian dealt with this high number of degrees of freedom by means of context disambiguation. When one starts reading, all the possible meanings of a sentence are assumed to be valid, and ordered from the most to the least likely. As reading progresses, some of these meanings are ruled out, until there is enough certainty to assign one. One has hence to continuously go backwards in the sentence to re-assign meanings to the previous parts. Pregroups, as context-free grammars in general, struggle quite a lot to cope with this level of ambiguity.

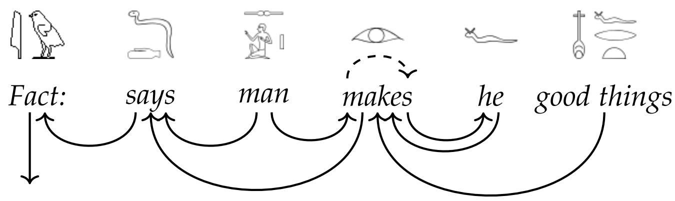
- Observing how the sentence structure is defined, it is clear that in the wiring that defines the grammatical dependencies there are a lot of cups that will necessarily have to cross with each other. If we assume than *the man* and *he* are the same person, then the wiring of the sentence in Figure 17 should look like this:



Unfortunately, as we said before, we are not allowed to form such combinations in the pregroup model, in which reduction rules force the cups not to intersect with each other. One way to solve the problem would be tweaking with the type structure of the word *makes*, changing it to $n^r \cdot s \cdot n^l \cdot n^l \cdot n$ in which the two outermost types are supposed to be connected by a cap in the semantics, as follows:



In this way the reduction is possible, and the only grammatical one would be the one below:



This tweak is nevertheless very unsatisfactory, because the quantity of different entries in the Lambek dictionary for even simple words would balloon quickly, having to deal with all the possible scenarios in which two words that must be connected are separated by one, two, ten words. Words like *makes* in our example should have one entry in the dictionary for every kind of context that may surround them, and this is clearly unfeasible.

At the moment, we have no real clue about how to deal with many of these problems in the pregroup setting. There have been endeavors to generalize the pregroup setting by means of logic modalities (see, for instance, [50, 111]). Nevertheless, the introduction of modalities could undermine the delicate functorial equilibrium that we have between grammar and semantics: Introducing modalities amounts to change the logical system we are working in, and hence also the categorical models of the system may have to change accordingly. This means that, while generalizing pregroups from a logical perspetive can seem sound from the grammatical point of view, we could end up loosing the functorial relationship between grammar and hypegraph categories, and hence our ability to reason diagrammatically. A similar argument can be given with respect to categorical grammars: Albeit many categorical generalizations of Lambek's work exist [112, 113, 115, 116], it is not immediately clear if these new category representing grammar are directly compatible with our semantics by means of a functorial relationship. Investigating these links will be object of future work.

We will try, then, to discuss the above-mentioned issues keeping in mind that the solution we seek has to be way less disruptive. We will focus, in particular, on the issue of personal pronouns making some observations, that are expressed below. Note that we won't try to solve this problem once and for all: what follows has to be interpreted more as a base for future work than as a definitive solution.

The crucial ability to distinguish if a pronoun such as *he* stands for its own or refers to some other noun in the sentence is mediated by the ability to request additional grammatical information when needed. In detail, we know that *he* is referring to a noun that is singular and masculine. Hence, when we see an occurence of *he* in a sentence we go back and we "explode" the nouns we encounter until we find one that meets our requirements (namely, it is masculine and singular). At this point we decide that *he* stands for that noun, connecting the two. We could try to model such behaviour with an algorithmic procedure as follows:

Instead of simple types, we will consider *towers* of types, that is, strings of finite length that we will typically write in a vertical fashion. The bottom type represents the basic grammatical one, while going up we have different levels of specification. For our example, we use towers of length two. A noun in this setting may look like this:

$$\begin{bmatrix} n_m^s \\ n \end{bmatrix}$$

The bottom type tells us that this is a noun, while the top type tells us that the noun is masculine singular. This may, for instance, represent *the man* in the sentence of Figure 17.

In our setting verbs will have bottom type $v \cdot n^l \cdot n^l$ and top type 1, since we do not need to specify any more information for the verb. The pronoun *he*, on the contrary, has type:

$$\begin{bmatrix} (n_m^s)^r \cdot n_m^s \\ n \end{bmatrix}$$

This means that the pronoun acts as a noun in the sentence, but it is referring to some other masculine singular noun that appears at some point at its left. The sentence in Figure 17 is then typed as follows:

$$\begin{bmatrix}1\\ s_f\cdot s^l\end{bmatrix}\cdot\begin{bmatrix}1\\ s\cdot s^l\cdot n^l\end{bmatrix}\cdot\begin{bmatrix}n^s_m\\ n\end{bmatrix}\cdot\begin{bmatrix}1\\ s\cdot n^l\cdot n^l\end{bmatrix}\cdot\begin{bmatrix}(n^s_m)^r\cdot n^s_m\\ n\end{bmatrix}\cdot\begin{bmatrix}n^p_f\\ n\end{bmatrix} \tag{12}$$

Where the verb (to say) has type $s\cdot s^l\cdot n^l$ because it takes another sentence as object. We are forced to type the verb in this way if we want to avoid inserting coordination words like "*that*", that as we pointed out before are not natural in Middle Egyptian. The reduction goes like this:

- We try to reduce the last row as we usually do with pregroups. This reduction (called *first reduction*) gives us the meaning where *he* stands on its own, and it is not connected to any other word in the sentence. We draw all the cups corresponding to this reduction (Figure 18, in red), obtaining a result – grammatical or not – in the standard sense (Figure 19a);

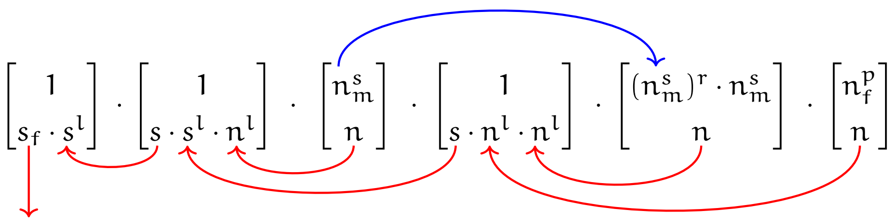
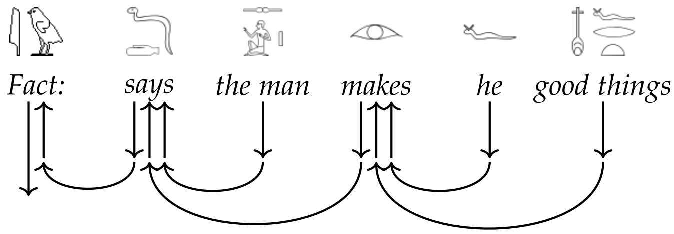


Figure 18: Reductions using towers of types. In blue, first reduction. In red, second reduction, drawn as caps for clarity.

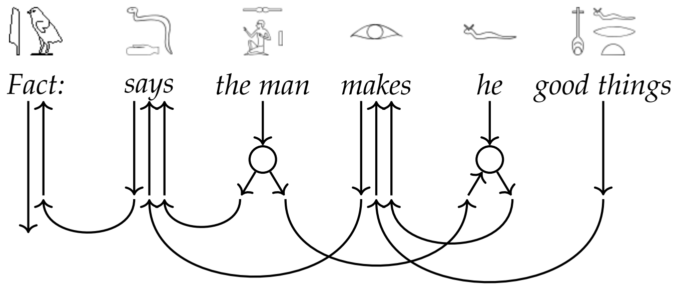
- Then we have a *second reduction*. We look at the top row (we "explode" the types): Every time a type with a $(-)^l$ appears, we look for the closest matching type on the right. When $(-)^r$ appears, we do the same on the left. If we can find one, we draw a cap between these words. In the case of Equation 12, we see that there is a $(n^s_m)^r$ in the type standing for *he* matching the n^s_m in the type standing for *the man*; We draw cups between these words (Figure 18, in blue);

- To finally obtain the correspondent second reduction, we have to merge the caps at the bottom level with the ones at the top one. This is done introducing oriented Frobenius structures (Figure 19b). The orientation is determined by the presence of adjunctions;

- The result of this is the wiring we need to use to assign our meanings in the semantics. This assignment is not well defined, since crossing wires are not allowed in pregroups. This means we are not working anymore in a pregroup category. If one wants to keep functoriality, then the wiring can be interpreted as existing in a suitable free category generated by the types showing up in the bottom row of the reduction, and the functor can be applied from there.



(a) Resulting first reduction.



(b) Resulting second reduction.

Figure 19: Resulting reductions using towers of types.

Note that if this procedure allows us to interpret personal pronouns in a meaningful way in the semantics. The interpretation of *he*, for instance, could be just the Frobenius comonoid counit. This represents maximum ambiguity, and it is consistent with the idea that a personal pronoun does not carry any meaning by itself, but it is more like a pointer referring to some other noun. In this case, applying the usual Frobenius laws, the spider under the word *he* in Figure 19b becomes just a cap, consistent with the interpretation that the meaning of *the man* gets copied and plugged in both sentences. On the other hand, the result of this semantic interpretation in Figure 19a is that the third slot of the verb *makes* gets fed with maximum ambiguity, consistent with the fact that, not knowing who this pronoun is referring to, we do not have any consistent way to assign meaning to it.

Moreover, note how the correctness of first and second reductions are not related. Is we substitute *she* to *he* in our sentence, then only the first reduction is correct, since *she* will have type $(n_f^s)^r \cdot n_f^s$ on the top row and will be looking for a feminine singular noun on its left, that does not appear in the sentence. So in this case only the first reduction is grammatical and we conclude that, without doubt, the pronoun is referring to a person not mentioned in the sentence.

Clearly this procedure is quite handwavy and not formally satisfying, but it is a first start to provide some notion of context in the grammar, context that pronouns need to act properly. A more formal way to approach the problem would be trying to employ techniques borrowed from contextuality to pregroup grammar, viz. recasting the theory of pregroup in a sheaf-theoretic way. Applications of sheaf theory to language are not new, see for instance [7]. Another – maybe easier – approach would be to forget the grammar altogether. After all, before feeding a sentence to our pregroup reductions we always suppose that there is some sort of algorithmic procedure able to parse the grammar of the sentence correctly, assigning the right types to words. In the same way, we could just suppose the wiring of a sentence to be simply given, maybe resorting to some machine-learning based algorithm.

The point being made here is that the pregroup grammar by itself is not at all essential, and its existence is motivated more on the ground of finding some common denominator with popular topics in linguistics that by actual model-theoretical needs. In the next chapter we will venture through a controversial road, arguing how the pregroup grammar – and the Chomskian paradigm in general – end up being counterproductive for what we want to do.

9 THE KATABASIS OF GRAMMAR

Difficile est longum subito deponere amorem.

— Catullus, Carmina [69, Carmen LXXVI]

In this chapter we will again elaborate on the problem of producing a cognitive-based model that works well for any language. We will, controversially, argue against the Chomskian approach to grammar and, more in general, against any kind of approach that wants to impose some fixed compositional grammar to our model "from above".

Noam Chomsky can arguably be considered the greatest linguist ever existed. His seminal work on generative and universal grammars [35] has massively shaped the last fifty years in the field of linguistics and linguistics-related sciences, such as natural language processing. Our compositional approach to meaning makes no exception, since in the way it is pursued it heavily relies on the pregroup formalism for the grammar part. While dealing with pregroups, we generate a set of strings from some basic types, and we have "grammatical rules" (the pregroup rules, in our case) to decide which strings are grammatical and which are not. As proved in [91] this procedure is moreover decidable, so it always terminates.

Links between pregroups and context-free grammars have been thoroughly investigated, and in fact pregroups are weakly equivalent to context-free grammars [31]. One of the obvious consequences of this equivalence is that, adopting the pregroup grammar, we automatically obtain that the language we are modelling is *recursive*, meaning that a sentence is able to refer to its own type.

For instance, the sentence *Alice saw that a rabbit was running* contains a subsentence, namely *a rabbit was running*: We are building a sentence using other sentences as basic blocs. In pregroups this kind of recursion is always allowed, since once we specify a sentence type s we can automatically consider triples of types like $s \cdot s^l \cdot s$, that are by definition grammatically correct. Clearly, iterating this we obtain strings like $s \cdot s^l \cdot s \cdot s^l \cdot \dots \cdot s \cdot s^l \cdot s$, of arbitrary finite length. These strings roughly correspond to sentences that may look like *He saw that she said that she noted ... that it fell*, and we could clearly go on forever according to our grammatical rules. As a consequence of this, we obtain that if a language is modeled by a pregroup grammar then the number of possible sentences that the language can produce is countable.

Recursion is a basic property shared all Chomsky and Montague-based generative grammars [34, 109, 120]. In particular, recursion is implied by one of the basic operations in the minimalist approach to generative syntax, called *Merge*, that roughly speaking is the process with which two syntactic objects are combined to form a new syntactic unit [36]. The concept of Merge is fundamental in the context of universal grammars [37] and, in particular, the fact that its presence it has been observed in (nearly, wait for the next section) all the languages known to men has prompted Chomsky to consider it as one of the fundamental features that distinguish human language from animal communication [76].

Before we go further, it is pertinent to give the reader a bit of context: Before the Chomskian revolution stormed the world of linguistics, the phenomenon of language acquisition was characterized by a behavioral approach: Researchers backing this theory believe that language is learned empirically, with a mechanism relying on trial and error and positive rewarding. All in all, this means that acquisition of language ultimately relies on the mere imitation of other speakers: It is the grammatical rules that are derived from the semantics, not viceversa [9].

Instead, according to the universal grammar (Chomskian) point of view things are exactly upside-down: The behavioral approach cannot work, disciples of Chomsky say. One of the most founded motivations they provide is what is called "Poverty of Stimulus" [127]. This argument goes more or less like this:

- In the context of language acquisition, we need to separate between positive and negative evidence. Positive evidence is defined as the set of grammatically correct sentences that the learner can experience listening or observing other speakers. Negative evidence is, instead, defined as the evidence that a learner has access to that proves how some sentences are grammatically incorrect. For instance, when a learner is corrected, negative evidence that the sentence uttered or written was grammatically incorrect is acquired;

- When they learn a language, children only get positive evidence: Native speakers, in fact, are just exposed to sentences that are grammatical, for example in the act of listening to their parents, while negative evidence is very limited. Nevertheless, they are able to learn the language grammar correctly;

- The conclusion is that there is a lack of negative stimulus that does not justify how people can learn what is grammatically correct and what is not: Positive rewarding by itself is not enough to acquire language, disproving the behaviorists theory.

The conclusion drawn is that human beings must have some innate capacity that allows them to learn language correctly. Chomskian researchers then argue that there must be some set of simple rules (like recursion) that is common to all human languages, hardwired in the human brain, that ultimately allows us to understand how to compose words and derive their meanings.

It is clear how the paradigms mentioned above work in a completely opposite way: If universal grammar completely justifies our "grammar $\rightarrow$ semantics" approach, the behavioral perspective backs up a "semantics $\rightarrow$ grammar" interpretation. These radically different points of view and all the debate originated defending or attacking both frameworks resulted in what is now informally known as "linguistics wars" [75, p. 105].

Needless to say, the approach used in categorical models of meaning is by all means Chomskian in spirit.

9.1 A lesson from Pirahã language

The Chomskian theory of language has become incredibly popular since its inception. Sure, many critiques have spawned during the decades, but nevertheless it remains one of the most important and central points of view when it comes to language acquisition. As we said, one of the biggest pieces of evidence backing up the theory of universal grammar was that some basic grammatical patterns, such as recursion, have been observed in any language known to man. But, as we are used to in mathematics, to disprove a statement like "This feature is common to *all* languages" we just need to find one counterexample.

Daniel Everett is a field linguist, particularly experienced in monolingual fieldwork. "Field linguist" roughly means "going into the real world and trying to understand how a given language works doing experiments with native speakers in their native environment", while "monolingual fieldwork" identifies those mightily skilled individuals that just dive into a new environment without having any common language to communicate with the natives (no translators, no nothing), with the goal of understanding how the language spoken by the people in the environment works. Clearly, this means "learning a language in the harshest conditions a man can imagine", and seeing how this works is damn interesting. Good sources of insight about this are (among others) a lecture given by Everett himself [61] or, if you fancy something more informal, this movie [142].

Anyway, Everett worked in particular with some tribes in the Brazilian amazon forest, and is known worldwide for his work on the Pirahã language, which he studied for more or less thirty years [45]. Pirahã is a language with some amazing features, that at the moment seem to be quite unique. Before going further, it has to be noted that Everett's work is considered to be highly controversial at the moment. This is not necessarily a bad thing, since with criticism comes also unusual attention for details: Much of his production about the Pirahã language is being audited and, hopefully, at some point we will be able to say if he was right or not. In any case, according to Everett, the most peculiar features of Pirahã are:

LACK OF COLORS Pirahã does not distinguish between colors, it just has words roughly corresponding to "light" and "dark"; Other colors are expressed by means of descriptive terms, such as *blood-like* for *red* [57];

LACK OF NUMBERS Pirahã has only two words to define the quantity of an object. In the beginning these words have been translated as the numbers *one* and *two* [55], but further research has shown how a better translation should interpret them as *few* and *fewer* [57]. To validate this, two experiments were made [65]: In the first, ten spools of thread were placed on a table, one at a time. Pirahã people used the word *one* when one spool was on the table, *two* when there were two and *two* again or *many* when there were more, backing up the hypothesis that these words mean *one* and *two*, respectively. In the second experiment, ten spools of thread were placed on a table, and subtracted one at a time. In this case the word for *one* began to be used when there were six (or four, depending on the test subject) spools left, providing evidence that the Pirahã words for numbers are not to be interpreted in an absolute sense, backing up the second hypothesis.

Interestingly, Pirahã[1] are aware of the fact that this is a limitation, and asked Everett to teach them to count because they were being constantly cheated in trade with other tribes. Everett tried to teach them for eight months, until they decided that it was too difficult and wanted to stop. No Pirahã was able to perform simple operations like "1+1" nor to count to ten after the training session [57, p. 11];

WHISTLING/HUMMING On the other hand, Pirahã has a remarkably complex system of tones and accents that allows them to whistle or hum any sentence [45]. Phonetic syllables have not to be pronounced to convey meaning. Whistling is used in particular while they hunt in the jungle;

LACK OF RECURSION This is the most controversial point. Everett argues that Pirahã, as a language, lacks not only recursion, but the capacity of embedding clauses altogether [57]. This can be seen as one of the reasons why counting is so difficult for them [58]. In his PhD thesis, Everett argued that recursion in the language was present [56], but after many more years of study he concluded that this was not the case, and his opinion depended on

1 "Pirahã" also denotes the Pirahã speakers.

prejudicial biases [58]. For instance, in Pirahã it would not be possible to say *My cousin's desk*, and the sentence would have to be split into *I have a cousin. This cousin has a desk*. The lack of embedding, if true, implies that Pirahã language is finite, meaning that we could compile a (very long) book with all the possible sentences appearing in the language.

These considerations – the last in particular – promped Everett to disbelieve the Chomskian approach to language, of which he was a disciple before [58, 107]. To further strengthen his argument, Everett also cites examples of how recursive behavior is frequently observed also in other species of the animal kindom (see, for instance [23]), and so it can't be what separates human language from animal communication. Recursivity, he then argues, is not a capacity that identifies uniquely human beings: It is just more developed in humans, and its expression in language is highly depending on the culture and the environment [58].

Everett's argument embittered many linguists (it has to be said that Chomsky reacted quite badly to Everett's research, dismissing it all together and calling him "a charlatain" [128]) and it is considered as highly controversial. Many recordings of Pirahã language are in the process of being reviewed by independent researchers, but definitive evidence proving or disproving Everett's conclusion still has to be found. What makes everything more difficult is that many Pirahã are now involved in initiatives sponsored by the Brazilian government that aim to teach Portuguese and maths to children. As a result, the Pirahã tribe will soon stop being monolingual and further field investigation will inevitably be falsified.

9.2 AN ALTERNATIVE THEORY OF LANGUAGE

One may think "Then what? Even if Everett is right, can we really say that one counterexample disproves a theory that works for more than six thousand languages?" Practically speaking, the answer is obviously "no". It is out of doubt that the Chomskian approach to language works well more than 99.983% of the times (that is, it works for nearly 5999 languages out of 6000). This, from an experimental point of view, means that Chomskian theory of language *is* a good theory. But the presence of even only one counterexample tells us that probably there is something more going on, that the theory may be incomplete, and that the matter surely requires further investigation. Anyway, this document is mainly concerned with mathematical models for language and not with linguistics itself: The reason why Everett's studies find a place in this pages is because they may indeed turn out very useful from a mathematical point of view.

In fact, what is really interesting for us is the alternative theory of language evolution that Everett proposes. First, it has to be noted that Everett's research backs up the idea that, if cognition shapes language, the opposite is also true. Within our framework this means that imposing a functor from the pregroup grammar to the semantics is missing how the semantics itself may be shaping the language grammar. In "Language: The Cultural Tool" [60] Everett advances the idea that language, as any other human tool, is created, shaped and employed to solve a problem, namely providing a mean of effective and efficient communication. Putting things in this perspective, the Pirahã never developed numbers because they are a tribe of hunter/-gatherers, so they never really needed it. Moreover, the absence of any cardinal or ordinal numbers in the language must be understood in cultural terms: One of Everett' sharpest observations is that Pirahã have the cognitive capability of counting, but choose not to do so. They could have, in fact, borrowed counting systems from other tribes, but they deliberately didn't because they believe their culture is complete as it is. Borrowing his own words,

> «The crucial thing is that the Pirahã have not borrowed any numbers – and they want to learn to count. They asked me to give them classes in Brazilian numbers, so for eight months I spent an hour every night trying to teach them how to count. And it never got anywhere, except for a few of the children. Some of the children learned to do reasonably well, but as soon as anybody started to perform well, they were sent away from the classes. It was just a fun time to eat popcorn and watch me write things on the board.» [58]

Similarly, they have the capability of expressing semantic recursion, but they do not need to implement it effectively in their language to convey recursive meanings. Chomskians argue that, seen in these terms, the lack of recursion does not falsify the theory of universal grammar [59, 118]: In fact, universal grammar is about the cognitive capabilities that allow human beings to learn language, and these capabilities may be there even if the language used is not recursive. This may be true, but if Everett is right then it becomes apparent how generative grammars do not tell us much about language itself, nor they are really useful to describe its structure: Culture and environment are the real ones that shape language as it is.

Embracing Everett's considerations, it is easy to realize that the approach "semantics $\rightarrow$ grammar" could be indeed very fruitful to help us understanding language. Moreover, if all the linguistics wars fuss is ultimately about how language is learned, it should be learning itself the central cognitive process we ought to model. Again, if one sides with Everett in this never-ending diatribe, this cognitive process should be modeled in a genuinely semantic way, along with a formalization of what "context" really means in the semantics. All in all,the real valueof Everett's proposal, from our point of view, is that it suggests a shift in perspective and a different direction of research in our quest to find suitable categorical models of meaning, probably based on modelling cognitive learning directly. This approach has never been tried in the framework of compositional models of meaning, and it will surely be one of the strongest things to push on in future work.

CONCLUSION

In this document, we started recalling what compositional models of meaning are, and reviewed part of what has been done in the field so far. We went on generalizing the concept of relation in many different ways, to accommodate desirable features like convexity or metrics, mainly by means of generalizing truth values in our models. We showed that this generalization is mathematically well-behaved, and we moreover proved how, with the majority of categories usually chosen to interpret word meaning, fixing a grammar also fixes the functorial interpretation of it in the semantics. What we never did, though, was to elaborate on a conceptual discussion that implicitly supports the models given. In practice, this means that we mathematically described a categorical construction that implements ideas from generative semantics and the cognitive theory of conceptual spaces, without really asking if these two theories were the best ones to actually model language cognitively. We were indeed justified to do so, since we are no linguists or cognitive scientists, and we felt we did not have enough technical background nor the confidence to write about something we do not really know well. All things considered, we did applied math, that is what computer scientists ultimately have to do.

In the last couple of chapters, though, we pointed out how many of the considerations we gave for granted before starting to draft our model do not represent well language nor meaning in its universality, and we did this in the most straightforward way a mathematician knows, by means of counterexamples. In Chapter 8 we used Middle Egyptian to show how the linearity of the pregroup grammar can be a problem to model meaning, while in Chapter 9 we questioned the approach “grammar $\rightarrow$ semantics” altogether.

After working for roughly two years on compositional models of meaning, I feel it is about time to express some personal opinions about what does work and what does not with our approach, in complete honesty. Again, I do not claim to be a linguist nor a cognitive behaviorist (since I am not), and the considerations that follow are mainly based on common sense and experience with compositional models.

Modelling meaning compositionally is amazing. The magic of seeing mathematical structures at work graphically gives us an insight of what is going on in our models that is in my opinion unrivaled. While many researchers in the field of natural language modeling and applied maths in general have to “divine” the parameters of their models empirically, often without having a real grip on what makes some parameters perform better than others and why this happens, we are ultimately able to rule this large amount of uncertainty out of our considerations by the means of a beautiful graphical calculus. This, undoubtedly, gives us the privilege of being concerned only with what really matters, ignoring all the infamous "background noise" that makes the understanding of certain phenomena even more complicated than what it already is.

But what is that really matters? In our effort to formalize meaning in cognitive terms, we often assume that our formalization effort must be directed to issues that are perceived as such in the western culture. This assumption relies on the unjustified belief that the centrality of these issues can be extended to other languages and considered universal: If for us it is important to specify and disambiguate meaning in a certain way, this should after all be true in general, we think. Alas, on this we are very wrong: The existence of cultures all around the world that do not put emphasis in the disambiguation of logical relationships between sentences, the existence of people that do not feel the need to express any numerical concept or that do not deem recursion as a strong desiderata when it comes about conveying meaning by means of composition of words should make us think. These people are humans too, and this implies that either our cognitive principles are culture-dependent, or most likely that we are still missing something fundamental in our picture.

One of the strongest examples supporting the Conceptual Spaces model, for instance, comes from the good representation that concepts related with colors have within it. But then again we are confronted with cultures that have no explicit representation of color at all, and what we should realize, ultimately, is that what really fits in our model is an idea of color that belongs to the western culture, and that is surely non-universal.

We should, all things considered, be much more careful on what we suppose to be true when we start formalizing meaning categorically, and realize that "what intuitively makes sense for us" is not often a universal denominator of the human species in its entirety.

In particular, the considerations expressed in the last couple of chapters led me to believe that imposing some grammatical structure "from above" does not help us if we want to model language in cognitive terms. We are acting out of prejudice, presuming how a given language must or must not be shaped, casting models out of our own arrogance and often forgetting what really matters: The culture, the environment, the semantics itself. We have to be humble, open minded, and more importantly we have to constantly confront our beliefs with the ones of people coming from environments that are radically different from ours.

As pointed out in Chapter 9, the effects of this are particularly evident in all the compositional models of meaning we reviewed and defined so far: All of them struggle when it comes to represent logical connectives (words like "and", for instance), since we are desperately looking for a set of rules to over-impose that incidentally happen to model these connectives well, instead of trying to extract these rules out from semantics and context. Moreover, and this is my strongest criticism to the way research in our field has been conducted so far, there is still no compositional model whatsoever that expresses crucial features as learning, nor any sensible idea about how to do it. It often looks like this problem is disregarded altogether, considered as "something we will start doing at some point" in all the research meetings and seminars I have ever attended to: Presently, what we do just stays still and doesn't show any kind of real evolution, nor dynamics. Now, language is obviously a dynamic process, constantly evolving, and trying to model it without allowing for structures, constructions and relations between words to be elastic, capable of being updated or changed within the model itself is an approach that, in my opinion, is doomed to fail.

It is my belief, then, that a true advancement in compositional models of meaning will be acquired only when the research will be directed to the effort – indeed titanic – of tracing back and modeling the cognitive roots of learning. The ability to learn is common to all human beings, is fundamental in the expression of intelligent behavior and, again, many questions should be asked about it, the most important being "is the way we learn culture-dependent or, instead, all the human beings process experiences and learn new things by means of a common cognitive process?"

The undeniable performance advancement in approaches like machine learning [22, 105, 106, 132] relies, in my opinion, exactly on the idea of abandoning any prejudice and focusing just on how to model learning. These approaches, though, follow a straight application-oriented perspective and do not care too much in understanding why for us "things work the way they do". In particular, there is no effort whatsoever in trying to find a compositional structure in the way learning happens. It has to be noted that some recent work in the categorification of machine learning techniques is now being carried out [64], but this categorical approach deals with a very low-level definition of learning: Namely, the endeavor is in modelling neural networks categorically. What we still lack is an *high-level categorical theory of learning*, that is, a compositional modelling of how the learning process happens cognitively (a proposal for a cognitive theory of learning can be found in [28], but it uses traditional machine learning techniques and it nor categorical nor compositional).

I do firmly believe – and this is the only presumption that I have left, I'd say – that learning for humans is indeed compositional: We have the innate ability of learning something, pack it into a block and then use this block to build bigger, more complicated processes. This is what we do, all of us, all the time, and mathematics, in its totality, is probably the most brilliant example of this: A tightly connected, endless chain of inferences in which every result is neatly generalized, boxed and stored away, only to be specialized and used again to prove something bigger later on. Long story short, the compositional structure behind learning, backed up with an adequate modeling of the environment in which we learn, is the key to explain the structure of human cognition itself.

With these considerations in mind, I hope to prompt researchers to invest energy in doing something that, surely, looks much more difficult than what are doing now, but that indeed remains the most fundamental problem yet to solve to unlock the secrets of human language.

Given such considerations, one may then be tempted to ask: "Has all this work been for nothing?" The answer is a resounding no, and for two different reasons. First of all, categories of relations have a very broad spectrum of applications, and may prove very useful in fields weakly or even not related at all with categorical models of meaning. Producing new mathematical objects is never a loss of time, and the ways to employ some mathematical structure to explain interesting phenomena are countless. All in all, from a purely technical perspective, we extended the applied mathematician's toolkit and this is always a good thing.

With respect to categorical models of meaning, it has to be noted that the last two chapters of this document are the result of a two year long journey in the development of the above-mentioned categories of generalized relations. Clearly, no one is crazy enough to dive into the formalization of a model based on principles he does not believe in, and in fact my trust in the "grammar $\rightarrow$ semantics" philosophy has not just abruptly vanished into thin air at some point. Instead, the last two years have been fundamental for me to gather incremental evidence that the Chomskian approach may not be the right one to follow when it comes to categorical models of meaning. I had the opportunity to study many subtleties of natural language and it has been precisely the endeavor to model those subtleties with compositional models of meaning to steer me towards the claims made in the last chapter. From this point of view, the experience developed in the writing of this document has been invaluable, and gave me a better insight of which direction of research I should pursue in the near future.

In particular, at the beginning of my doctorate, I was eager to employ math to model language and, as it is sadly quite common in the computer science and maths sector, I disregarded many considerations about the influence that culture and environment have on language as "philosophers' stuff". "We do actual numbers here", I thought again and again, and I was perfectly happy with my limited knowledge of linguistics and anthropology, all focused on formalizing language according to principles that "must necessarily be true".

Two years later, I feel like waking up from a long slumber, realizing that humanities matter, anthropology matters, and having a good understanding about how different cultures represent meaning is a fundamental step towards a good formalization of language in mathematical terms. In particular, strict collaboration with specialists coming from these fields will be invaluable to progress towards this effort, and finding a common language for communicating among these different disciplines will be a big challenge. The last part of this document, then, has not to be seen as a celebration of failure of our prefixed endeavor, but as an exhortation to extend one's area of competence to embrace a more interdisciplinary approach to compositional models of meaning.

It is not possible anymore, for the computer scientist, to proceed assuming that some very debated topics in linguistics have trivial solutions, exactly as it is not possible for the linguist to assume that the mathematical formalization of some linguistic phenomenon will be easy. This lack of communication can only produce models (on both sides) based on wrong assumptions, slowing the quest towards a complete, categorical model of meaning just for lack of communication, that is the most futile reason ever.

This is probably the most important message I wanted to share in this document, and the one with which I now conclude: Doing research with people having our very same background, or sharing with us the way with which they approach reality, that is easy. But, as Sophocles teaches us in his Antigone, real wisdom often comes from conflict, from animated discussions, from the uncomfortable feeling we experience when we face someone that looks at the world from a completely different point of view. It is only through this struggle that we will be able to broaden our knowledge enough to see what language is "from above", gathering enough insight to draft models that are not based on naive assumptions. So let us all come out of our comfort zones, and start working together without discarding decades of progress in fields we deemed "not important enough" for too long.

Part IV

APPENDIX

A
PROOFS

PROOFS FROM CHAPTER 4

Section 4.1

Proposition 4.1.4. $\langle X, \{+_p\}_{p\in[0,1]}\rangle$ *is an algebra. Moreover, it satisfies the equations*

ZERO COMBINATIONS $x +_0 y \simeq y$;

IDEMPOTENCY $x +_p x \simeq x$;

PARAMETRIZED COMMUTATIVITY $x +_p y \simeq y +_{1-p} x$;

PARAMETRIZED ASSOCIATIVITY $(x +_p y) +_q z \simeq x +_{pq} (y +_{\frac{q-pq}{1-pq}} z)$.

Note that the last three properties represent schemes *of equations, since they have to hold for every* p *(we cannot quantify on* p *because the* $+_p$ *are operations).*

Proof. $+_p$ carries elements of X to elements of X (since α has codomain in X) and is clearly functional, so every $+_p$ is an operation on X.

To verify that $\langle X, \{+_p\}_{p\in[0,1]}\rangle$ satisfies the given equations all we need to do is to apply properties of multiplication and addition of real numbers and follow the properties given by the monadic structure. For instance,

$$\begin{aligned} x +_p x &= \alpha(p\eta(x) + (1-p)\eta(x)) \\ &= \alpha(p\,|x\rangle + (1-p)\,|x\rangle) \\ &= \alpha(|x\rangle) \\ &= x \end{aligned}$$ □

Proposition 4.1.9. *The monomorphisms in* ***Convex*** *are exactly the injective homomorphisms.*

Proof. This is a standard property of categories monadic over **Set**. □

Proposition 4.1.10. *In the category* ***Convex*** *the image of a convex morphism* $h : (A, \alpha) \to (B, \beta)$ *is a convex subalgebra of* (B, β)*. Moreover, assume* $h : (A, \alpha) \to (B, \beta)$ *is a monomorphism in* ***Convex****. There is then a convex morphism* $\operatorname{im} h \to (A, \alpha)$ *given by inverse images.*

Proof. This is a direct consequence of the first isomorphism theorem, that holds in every algebraic variety, so in particular in $\mathcal{V}$. □

Section 4.1.1

Proposition 4.1.12. (X, α) *disregards weights iff the corresponding algebra in* $\mathcal{V}$ *satisfies the equation* $\forall x, y \in X, x +_p y \simeq x +_q y$ *for every couple of* p, q *such that* $p, q \neq 0, 1$.

Proof. Obvious from Definition 4.1.3 and Proposition 4.1.5. □

Proposition 4.1.14. *Let* Ω *be the convex algebra induced by the two element meet semilattice* $\{\bot, \top\}$ *with* $\bot \leqslant \top$. *For a finite set* X, *there is a bijective correspondence between* ***Convex***-*morphisms of type* $F(X) \to \Omega$ *and sub-complexes of* $F(X)$, *where* $F(X)$ *denotes the* free convex algebra over X.

Proof. Using the universal property of free algebras, there is a bijective correspondence between morphisms of type $X \to U\Omega$ and morphisms of type $F(X) \to \Omega$. We then observe that:

$$\sum_i p_i |x_i\rangle$$

is mapped to true iff each x_i is, as convex combinations in Ω are given by meets. □

Section 4.2

Proposition 4.2.2. *Let* (X, α) *be a convex algebra.* (X, α) *is coherent.*

Proof. Fix any two elements $a, b \in X$ and $p, p' \in [0, 1]$. Fiddling with eerie commutativity and associativity we can write, for all $k \in [0, 1]$:

$$\begin{aligned}
(a +_p b) +_k (a +_{p'} b) &= (b +_{1-p} a) +_k (a +_{p'} b) \\
&= b +_{(1-p)k} (a +_{\frac{k-(1-p)k}{1-(1-p)k}} (a +_{p'} b)) \\
&= b +_{(1-p)k} ((a +_{p'} b) +_{1-\frac{k-(1-p)k}{1-(1-p)k}} a) \\
&= b +_{(1-p)k} ((a +_{p'} b) +_{\frac{1-k}{1-(1-p)k}} a) \\
&= b +_{(1-p)k} ((b +_{1-p'} a) +_{\frac{1-k}{1-(1-p)k}} a) \\
&= b +_{(1-p)k} (b +_{\frac{(1-p')(1-k)}{1-(1-p)k}} a) \\
&= (b +_{\frac{(1-p')(1-k)}{1-(1-p)k}} a) +_{1-(1-p)k} b \\
&= (a +_{1-\frac{(1-p')(1-k)}{1-(1-p)k}} b) +_{1-(1-p)k} b \\
&= (a +_{\frac{p'-p'k+pk}{1-(1-p)k}} b) +_{1-(1-p)k} b \\
&= (a +_{\frac{kp+(1-k)p'}{1-(1-p)k}} b) +_{1-(1-p)k} b \\
&= a +_{kp+(1-k)p'} b
\end{aligned}$$

Now, suppose $(a +_p b) = (a +_{p'} b)$. Then by idempotency $(a +_p b) +_k (a +_{p'} b) = (a +_p b)$ and so $(a +_{kp+(1-k)p'} b) = (a +_p b)$. Now fix

$q \in [p', p]$. Since k varies over $[0,1]$ and $kp + (1-k)p'$ is a continuous function in k from $[0,1]$ to $[p', p]$, by intermediate value theorem it spans all the values in the interval $[p, p']$. Then there is a k such that $kp + (1-k)p' = q$. This means that:

$$(a +_q b) = (a +_{kp+(1-k)p'} b) = (a +_p b)$$

as we wanted. □

Lemma 4.2.3. *Let (X, α) be a convex algebra. If α does not disregard weights, then $|X| \neq 2$.*

Proof. Ex absurdo: By Proposition 4.2.2 (X, α) is coherent. Call a, b the elements of X. Since α does not disregard weights, there are two distinct convex combinations $a +_p b, a +_q b$ with $p \neq q$ such that $a +_p b \neq a +_q b$. Since $|X| = 2$, then it is either $a +_p b = b$ and $a +_q b = a$ or the opposite. Without loss of generality suppose it is the former possibility that happens. Coherence of α then forces the algebra to be defined as:

$$\begin{cases} a +_p b = b & \text{For each } p \leqslant x, \\ a +_p b = a & \text{For each } p > x. \end{cases}$$

For some $x \in]0,1[$. We will show such an x cannot exist. Suppose $p \leqslant x$. Then $a +_p b = b$. Since $a +_p b = b +_{1-p} a$, it is also $1 - p > x$, and hence $p \leqslant x < 1 - p$. Since these inequalities have to hold for every p, we can pick $p = x$, obtaining $x < 1 - x$. But we can also pick $p = 1 - x$, obtaining $1 - x < 1 - (1 - x) = x$. Clearly no real number satisfies both these inequalities, contradiction. □

Lemma 4.2.4. *Let (X, α) be a convex algebra. If α doesn't disregard weights, then $|X| \neq 3$.*

Proof. Ex absurdo: By Proposition 4.2.2 (X, α) is coherent. Call a, b, c the elements of X. Since α does not disregard weights, there are two distinct convex combinations of say a, b with $0 < p \neq q < 1$ such that $a +_p b \neq a +_q b$. First of all, for some $k \in [p, q]$ it has to be $a +_p b \neq a +_q b \neq a +_k b$, otherwise the operations of (X, α) restricted to the set $\{a +_p b, a +_q b\}$ would define a subalgebra with cardinality of the support 2, against Lemma 4.2.3. Since X has only three elements, coherence forces $a +_n b$ to be defined in one of the two following ways (the cases with a and b switched between each other are analogous):

$$a +_n b := \begin{cases} a & n > x_2; \\ c & x_1 \leqslant n \leqslant x_2; \\ b & n < x_1. \end{cases} \qquad a +_n b := \begin{cases} a & n \geqslant x_2; \\ c & x_1 < n < x_2; \\ b & n \leqslant x_1. \end{cases}$$

for some $x_1, x_2 \in]0,1[$. In both cases, eerie commutativity forces x_2 to be equal to $1 - x_1$, so:

$$a +_n b := \begin{cases} a & n > 1 - x_1; \\ c & x_1 \leqslant n \leqslant 1 - x_1; \\ b & n < x_1. \end{cases} \quad a +_n b := \begin{cases} a & n \geqslant 1 - x_1; \\ c & x_1 < n < 1 - x_1; \\ b & n \leqslant x_1. \end{cases}$$

Case on the right: The operations of (X, α) on $\{a, c\}$ are closed since, for each $k \in [0,1]$, $a +_k c = a +_k (a +_n b) = a +_{k+(1-k)n} b$ for some $x_1 \leqslant n \leqslant 1 - x_1$, hence $a +_k c$ is always equal to a or c. Then for $k < \frac{1-x_1-n}{1-n}$ we have $a +_k c = c$, while for $k \geqslant \frac{1-x_1-n}{1-n}$ it is $a +_k c = a$. By Lemma 4.2.3 this subalgebra has to disregard weights, implying $\frac{1-x_1-n}{1-n} = 1$ and hence $x_1 = 0$, contradicting $x_1 \in]0,1[$.

Case on the left: Applying the very same argument we obtain $a +_k c = a +_k (a +_{x_1} b) = a +_{k+(1-k)x_1} b$, then for $k \leqslant \frac{1-x_1-x_1}{1-x_1}$ we have $a +_k c = c$, while for $k > \frac{1-x_1-x_1}{1-x_1}$ it is $a +_k c = a$. Again by Lemma 4.2.3 it has to be $\frac{1-x_1-x_1}{1-x_1} = 0$, hence $x_1 = \frac{1}{2}$. But then:

$$\begin{aligned} b &= a +_{\frac{2}{5}} b = (a +_{\frac{3}{4}} b) +_{\frac{2}{5}} (a +_{\frac{1}{3}} b) = (b +_{\frac{1}{4}} a) +_{\frac{2}{5}} (a +_{\frac{1}{3}} b) \\ &= b +_{\frac{1}{10}} (a +_{\frac{1}{3}} (a +_{\frac{1}{3}} b)) = b +_{\frac{1}{10}} ((a +_{\frac{3}{5}} a) +_{\frac{5}{9}} b) = b +_{\frac{1}{10}} (a +_{\frac{5}{9}} b) \\ &= b +_{\frac{1}{10}} (b +_{\frac{4}{9}} a) = (b +_{\frac{1}{5}} b) +_{\frac{1}{2}} a = b +_{\frac{1}{2}} a = a +_{\frac{1}{2}} b \\ &= c \end{aligned}$$

Violating the hypothesis again. □

Section 4.3

Lemma 4.3.3. *Let (X, α) be a convex algebra, and let B be the beetweennes relation on X. Then B satisfies axioms 1,2,3,4,5,7 of Definition 4.3.1.*

Proof. We check that B satisfies the conditions case by case:

1. $(a, b, c) \in B \Rightarrow a +_p c = b \Rightarrow c +_{1-p} a = b \Rightarrow (c, b, a) \in B$;
2. $(a, b, a) \in B \Rightarrow a +_p a = b \Rightarrow a = b$;
3. $b = a +_0 b \Rightarrow (a, b, b) \in B$;
4. $b = a +_0 b \Rightarrow (a, b, b) \in B$;
5. Suppose $b = a +_p c, a = b +_q c$. If $q = 1$ then $a = b = b$, and same can be said if $p = 1$. Suppose then $p, q < 1$. Hence $a = ((a +_p c) +_q c) = a +_{pq} (c +_{\frac{q-pq}{1-pq}} c) = a +_{pq} c$. But then $b = a +_p c = a +_{\frac{p-pq}{1-pq}} (a +_{pq} c)$ and so $(a, b, a) \in B$. Apply 2 to get $a = b$;

6. Suppose $b = a +_p d, c = b +_q d$. Then $c = (a +_p d) +_q d = (a +_{pq} d)$. But then $b = a +_p d = a +_{\frac{p-pq}{1-pq}} (a +_{pq} d)$ and so $(a, b, c) \in B$. □

Lemma 4.3.5. *Let* (X, α) *be a convex algebra and let* B *be the betweenness relation on* X*. Then* B *satisfies the property*

$$(a, b, c), (a, b', c) \in B \Rightarrow (a, b, b') \text{ or } (a, b', b) \in B$$

Proof. To see this, take $b = a +_p c, b' = a +_q c$. If $p \geqslant q$ it follows that $b = a +_p c = a +_{\frac{p-q}{1-q}} (a +_q c)$ and $(a, b, b') \in B$. With the same reasoning we conclude $(a, b', b) \in B$ if $p \leqslant q$. □

PROOFS FROM CHAPTER 6

Section 6.2

Proposition 6.2.6. *[Converse] Let $\mathcal{E}$ be a topos, (Σ, E) a variety in $\mathcal{E}$, and $(Q, \odot, k, \bigvee)$ an internal commutative quantale. There is an identity on objects strict symmetric monoidal functor:*

$$(-)^\circ : \boldsymbol{Rel}_{(\Sigma,E)}(Q)^{op} \to \boldsymbol{Rel}_{(\Sigma,E)}(Q)$$

Given by reversing arguments:

$$R^\circ(b, a) = R(a, b)$$

Proof. We first check that taking the converse gives a well defined relation. For $\sigma \in \Sigma$ of arity n, we reason:

$$\begin{aligned} R^\circ(a_1, b_1) \odot \ldots \odot R^\circ(a_n, b_n) &= R(b_1, a_1) \odot \ldots \odot R(b_n, a_n) \\ &\leqslant R(\sigma(b_1, \ldots, b_n), \sigma(a_1, \ldots, a_n)) \\ &= R^\circ(\sigma(a_1, \ldots, a_n), \sigma(b_1, \ldots, b_n)) \end{aligned}$$

Next, we must confirm identities are preserved.

$$id_A^\circ(a_1, a_2) = id_A(a_2, a_1) = \bigvee\{k \mid a_2 = a_1\} = id_A(a_1, a_2)$$

We confirm also functoriality with respect to composition:

$$\begin{aligned} (R \circ S)^\circ(a, c) &= (R \circ S)(c, a) \\ &= \bigvee_b S(c, b) \odot R(b, a) \\ &= \bigvee_b S^\circ(b, c) \odot R^\circ(a, b) \\ &= \bigvee_b R^\circ(a, b) \odot S^\circ(b, c) \\ &= (S^\circ \circ R^\circ)(a, c) \end{aligned}$$

Finally, we must check that the converse distributes over tensors:

$$\begin{aligned} (R^\circ \otimes S^\circ)(b, b', a, a') &= R^\circ(b, a) \odot S^\circ(b', a') \\ &= R(a, b) \odot S(a', b') \\ &= (R \otimes S)(a, a', b, b') \\ &= (R \otimes S)^\circ(b, b', a, a') \end{aligned}$$ ◻

Proposition 6.2.7. *[Graph] Let $\mathcal{E}$ be a topos, (Σ, E) a variety in $\mathcal{E}$, and $(Q, \odot, k, \bigvee)$ an internal commutative quantale. There is an identity on objects strict symmetric monoidal functor:*

$$(-)_\circ : \boldsymbol{Alg}(\Sigma, E) \to \boldsymbol{Rel}_{(\Sigma,E)}(Q)$$

With action defined on morphism $f : A \to B$ *by:*

$$f_\circ(a, b) = \bigvee\{k \mid f(a) = b\}$$

The symmetric monoidal structure on ***Alg***(Σ, E) *is the finite product structure.*

Proof. First of all we have to check that the resulting relation respects the algebraic structure. For $\sigma \in \Sigma$:

$$\begin{aligned}
f_\circ(a_1, b_1) \odot \ldots \odot f_\circ(a_n, b_n) = \\
&= \left[\bigvee\{k \mid f(a_1) = b_1\}\right] \odot \ldots \odot \left[\bigvee\{k \mid f(a_n) = b_n\}\right] \\
&= \bigvee\{k \mid f(a_1) = b_1 \wedge \cdots \wedge f(a_n) = b_n\} \\
&\leqslant \bigvee\{k \mid \sigma(f(a_1), \ldots, f(a_n)) = \sigma(b_1, \ldots, b_n)\} \\
&= \bigvee\{k \mid f(\sigma(a_1, \ldots, a_n)) = \sigma(b_1, \ldots, b_n) \\
&= f_\circ(\sigma(a_1, \ldots, a_n), \sigma(b_1, \ldots, b_n))
\end{aligned}$$

Then we confirm this is functorial with respect to identities:

$$\mathrm{id}_{A\circ}(a_1, a_2) = \bigvee\{k \mid \mathrm{id}_A(a_1) = a_2\} = \bigvee\{k \mid a_1 = a_2\} = \mathrm{id}_A(a_1, a_2)$$

For functoriality with respect to composition,

$$\begin{aligned}
(g_\circ \circ f_\circ)(a, c) &= \bigvee_b f_\circ(a, b) \odot g_\circ(b, c) \\
&= \bigvee_b \left[\bigvee\{k \mid f(a) = b\}\right] \odot \left[\bigvee\{k \mid g(b) = c\}\right] \\
&= \bigvee\{k \mid g(f(a)) = c\} \\
&= (g \circ f)_\circ(a, c)
\end{aligned}$$

Finally we prove the preservation of the monoidal structure:

$$\begin{aligned}
(f \times g)_\circ((a, a'), (b, b')) &= \bigvee\{k \mid (b, b') \\
&= (f \times g)(a, a')\} \\
&= \bigvee\{k \mid b = f(a) \wedge b' = g(a')\} \\
&= \left[\bigvee\{k \mid b = f(a)\}\right] \odot \left[\bigvee\{k \mid b' = g(a')\}\right] \\
&= f_\circ(a, b) \odot g_\circ(a', b') \\
&= (f_\circ \otimes g_\circ)((a, a'), (b, b'))
\end{aligned}$$

We moreover prove a fact used in the proof of Proposition 6.2.5, that is, if f is an isomorphism in $\mathcal{E}$ then $(f^{-1})_\circ = (f_\circ)^\circ$. This is a simple check:

$$\begin{aligned}
(f^{-1})_\circ(b, a) &= \bigvee\{k \mid f^{-1}(b) = a\} \\
&= \bigvee\{k \mid f(a) = b\} \\
&= f_\circ(a, b) \\
&= (f_\circ)^\circ(b, a)
\end{aligned}$$

□

Section 6.3

Lemma A.0.1. *Let $\mathcal{E}$ be a finitely complete category, and $(Q, \odot, k)$ an internal monoid. If*

$$h : (X_1, f_1, g_1, \chi_1) \to (X_2, f_2, g_2, \chi_2)$$

Is a Q-span morphism with an inverse in $\mathcal{E}$, then it is an isomorphism.

Proof. We aim to show that h^{-1} is the required inverse as a Q-span morphism. We calculate:

$$f_1 \circ h^{-1} = f_2 \circ h \circ h^{-1} = f_2$$
$$g_1 \circ h^{-1} = g_2 \circ h \circ h^{-1} = g_2$$
$$\chi_1 \circ h^{-1} = \chi_2 \circ h \circ h^{-1} = \chi_2$$ □

Lemma A.0.2. *Let $\mathcal{E}$ be a topos, $(Q, \odot, k, \leqslant)$ an internal partially ordered commutative monoid, and (Σ, E) an algebraic signature. If (X_1, f_1, g_1, χ_1) is an algebraic Q-span, and (X_2, f_2, g_2, χ_2) is an isomorphic Q-span, then it is also an algebraic Q-span.*

Proof. For the assumptions in the question, with ι denoting the assumed isomorphism and $\sigma \in \Sigma$, if

$$f_2(x_1) = a_1 \wedge g_2(x_1) = b_1 \wedge \cdots \wedge f_2(x_n) = a_n \wedge g_2(x_n) = b_n$$

Then using our span isomorphism:

$$f_1(\iota^{-1}(x_1)) = a_1 \wedge g_1(\iota^{-1}(x_1)) = b_1 \wedge \ldots$$
$$\cdots \wedge f_1(\iota^{-1}(x_n)) = a_n \wedge g_1(\iota^{-1}(x_n)) = b_n$$

By assumption that the first span is algebraic, there exists x such that:

$$f_1(x) = \sigma(a_1, \ldots, a_n) \wedge g_1(x) = \sigma(b_1, \ldots, b_n) \wedge$$
$$\wedge \chi_1(\iota^{-1}(x_1)) \odot \ldots \odot \chi_1(\iota^{-1}(x_n)) \leqslant \chi_1(x)$$

Therefore, using our span isomorphism again:

$$f_2(\iota(x)) = \sigma(a_1, \ldots, a_n) \wedge g_2(\iota(x)) = \sigma(b_1, \ldots, b_n) \wedge$$
$$\wedge \chi_2(x_1) \odot \ldots \odot \chi_2(x_n) \leqslant \chi_2(\iota(x))$$ □

Theorem 6.3.2. *Let $\mathcal{E}$ be a finitely complete category, and $(Q, \odot, k)$ an internal commutative monoid. The category **Span**(Q) is a hypergraph category.*

Proof. Special case of Theorem 6.3.7. □

Proposition 6.3.5. *[Converse] Let $\mathcal{E}$ be a topos, (Σ, E) a variety in $\mathcal{E}$, and $(Q, \odot, k, \leqslant)$ an internal partially ordered commutative monoid. There is an identity on objects strict symmetric monoidal functor:*

$$(-)^\circ : \mathbf{Span}_{(\Sigma,E)}(Q)^{op} \to \mathbf{Span}_{(\Sigma,E)}(Q)$$

Given by reversing the legs of the underlying span:

$$(X, f, g, \chi)^\circ = (X, g, f, \chi)$$

Proof. First we prove that $(-)^\circ$ respects the algebraic structure: just observe that the condition for a Q-span to be algebraic is symmetrical in its domain and codomain, and therefore preserved. Now we prove it is a functor. It clearly preserves identities. We aim to show:

$$(X, f, g, \chi)^\circ \circ (Y, h, k, \xi)^\circ = ((Y, h, k, \xi) \circ (X, f, g, \chi))^\circ$$

These two spans are given by:

$$(X, f, g, \chi)^\circ \circ (Y, h, k, \xi)^\circ = (Y \times_B X, k \circ p_1, f \circ p_2, \odot \circ \langle \xi \circ p_1, \chi \circ p_2 \rangle)$$
$$((Y, h, k, \xi) \circ (X, f, g, \chi))^\circ = (X \times_B Y, k \circ p_2, f \circ p_1, \odot \circ \langle \chi \circ p_1, \xi \circ p_2 \rangle)$$

The morphisms:

$$\langle p_2, p_1 \rangle : X \times_B Y \to Y \times_B X \qquad \langle p_2, p_1 \rangle : Y \times_B X \to X \times_B Y$$

Witness an isomorphism in $\mathcal{E}$. We first confirm that this gives a span isomorphism. This follows trivially from elementary properties of pullbacks. Finally, we must prove that this commutes with characteristic morphisms. This makes essential use of commutativity of $\odot$:

$$\odot \circ \langle \xi \circ p_1, \chi \circ p_2 \rangle \circ \langle p_2, p_1 \rangle = \odot \circ \langle \xi \circ p_2, \chi \circ p_1 \rangle = \odot \circ \langle \chi \circ p_1, \xi \circ p_2 \rangle$$

Proving that the functor is strict symmetric is trivial. □

Proposition 6.3.6. *[Graph] Let $\mathcal{E}$ be a topos, and $(Q, \odot, k, \leqslant)$ an internal partially ordered commutative monoid. There is an identity on objects strict symmetric monoidal functor:*

$$(-)_\circ : \boldsymbol{Alg}(\Sigma, E) \to \boldsymbol{Span}_{(\Sigma, E)}(Q)$$

With the action on morphism $f : A \to B$ given by:

$$f_\circ = (A, id_A, f, \chi_k)$$

Proof. We have to check that the graph functor respects the algebraic structure. As the characteristic morphism is constant, we must simply confirm the existence of witnesses relating composite terms. If

$$f(a_1) = b_1 \wedge \cdots \wedge f(a_n) = b_n$$

Then, as f is a homomorphism:

$$f(\sigma(a_1, \ldots, a_n)) = \sigma(f(a_1), \ldots, f(a_n)) = \sigma(b_1, \ldots, b_n)$$

Now we prove it is a functor. Since this construction is well known for ordinary spans, we must just confirm the interaction with characteristic morphisms behaves appropriately. Firstly we note that:

$$(id_A)_\circ = (A, id_A, id_A, \chi_k)$$

And so identities are preserved. For composition, we have an ordinary span isomorphism:

$$\langle id_A, f\rangle : A \to A \times_B B$$

We must confirm this commutes with characteristic morphisms:

$$\odot \circ (\chi_k \times \chi_k) \circ \langle p_1, p_2\rangle \circ \langle id_A, f\rangle = \odot \circ \langle \chi_k, \chi_k \circ f\rangle = \odot \circ \langle \chi_k, \chi_k\rangle = \chi_k$$

Finally we have to prove that graph and converse are symmetric monoidal functors. For converse this is trivial. For graph, commutativity with tensor has been proved in Proposition 6.3.4. □

Theorem 6.3.7. *Let $\mathcal{E}$ be a topos, (Σ, E) a variety in $\mathcal{E}$, and $(Q, \odot, k, \leqslant)$ an internal partially ordered commutative monoid. The category $\textbf{Span}_{(\Sigma,E)}(Q)$ is a hypergraph category. The cocommutative comonoid structure is given by the graphs of the canonical comonoids described in Proposition 6.2.8, and the monoid structure is given by their converses.*

Proof. This is just a straightforward check, very similar in fashion to the ones already performed to prove Proposition 6.3.4. □

Section 6.4

Proposition 6.4.9. *Let $\mathcal{E}$ be a topos, and (Σ, E) a variety in $\mathcal{E}$. Converses respect order structure, in that:*

- *If $(Q, \odot, k, \bigvee)$ is an internal quantale, the converse functor of Proposition 6.2.6 is a partially ordered functor;*
- *If $(Q, \odot, k, \leqslant)$ is an internal partially order monoid, the converse functor of Proposition 6.3.5 is a preordered functor.*

Proof. For relations this is trivial since the definition of ordering is symmetric in the relation arguments. For span the same consideration holds, since the ordering is defined on the span heads and switching the span legs does not affect this definition. □

Section 6.7

Lemma 6.7.4. *Let $(Q, \odot, k, \bigvee)$ be an internal quantale. If for every p, q the inequality $p \odot q \leqslant p$ holds, then every relation is affine. If $p \leqslant p \odot p$ holds then every relation is relevant. Similarly, if $(Q, \odot, k, \leqslant)$ is an internal partially ordered monoid, if $p \odot q \leqslant p$ holds then every span is affine and if $p \leqslant p \odot p$ then every span is relevant.*

Proof. This is obvious noting that for every a, b, x, $R(a, b)$ and $\chi(x)$ are elements of Q. We only prove the relational affine case explicitly, all the rest being similar. For arbitrary a_1, a_2, b_1, b_2 consider the product $R(a_1, b_1) \odot R(a_2, b_2)$. Both $R(a_1, b_1)$ and $R(a_2, b_2)$ are elements of Q, hence, being Q affine, we can readily infer $R(a_1, b_1) \odot R(a_2, b_2) \leqslant R(a_1, b_1)$. Being the variables arbitrarily chosen, we universally quantify on them obtaining the axiom of being affine for R as a valid formula in the ambient topos $\mathcal{E}$. ◻

Theorem 6.7.11. *Let $\mathcal{E}$ be a topos and $(Q, \odot, k, \leqslant)$ an internal partially ordered commutative monoid. Let $i : (\Sigma_1, E_1) \to (\Sigma_2, E_2)$ be a linear interpretation of signatures. There is an identity on morphisms strict monoidal functor:*

$$i^* : \boldsymbol{Span}^{\text{lin}}_{(\Sigma_2,E_2)}(Q) \to \boldsymbol{Span}^{\text{lin}}_{(\Sigma_1,E_1)}(Q)$$

Sending each (Σ_2, E_2)-algebra to the corresponding (Σ_1, E_1)-algebra under the interpretation. The assignment $i \mapsto i^*$ extends to a contravariant functor. i^* commutes with graphs and converses, that is, the following diagrams commute:

$$\begin{array}{ccc} \boldsymbol{Span}^{\text{lin}}_{(\Sigma_2,E_2)}(Q) & \xrightarrow{\;i^*\;} & \boldsymbol{Span}^{\text{lin}}_{(\Sigma_1,E_1)}(Q) \\ {\scriptstyle (-)_\circ}\uparrow & & \uparrow{\scriptstyle (-)_\circ} \\ \boldsymbol{Alg}(\Sigma_2, E_2) & \xrightarrow[\;i^*\;]{} & \boldsymbol{Alg}(\Sigma_1, E_1) \end{array}$$

$$\begin{array}{ccc} \boldsymbol{Span}^{\text{lin}}_{(\Sigma_2,E_2)}(Q)^{op} & \xrightarrow{\;(i^*)^{op}\;} & \boldsymbol{Span}^{\text{lin}}_{(\Sigma_1,E_1)}(Q)^{op} \\ {\scriptstyle (-)^\circ}\downarrow & & \downarrow{\scriptstyle (-)^\circ} \\ \boldsymbol{Span}^{\text{lin}}_{(\Sigma_2,E_2)}(Q) & \xrightarrow[\;i^*\;]{} & \boldsymbol{Span}^{\text{lin}}_{(\Sigma_1,E_1)}(Q) \end{array}$$

The bottom functor in the first diagram is the obvious induced functor between categories of algebras. Similar results hold for affine, relevant and cartesian interpretations and relations.

Proof. i^* is defined as in the relational case, sending algebras of type (Σ_2, E_2) to their interpretations of type (Σ_1, E_1). Being it identity on morphisms by definition (the fact that morphisms of $\mathbf{Span}^{\text{lin}}_{(\Sigma_2,E_2)}(Q)$ are also morphisms of $\mathbf{Span}^{\text{lin}}_{(\Sigma_1,E_1)}(Q)$ is a direct consequence of Proposition 6.7.9) functoriality follows trivially. Noting that i^* and $\hat{i}$ are identity on morphisms and they act on the same way on objects, where converse and graph are instead identity on objects. Commutativity of i^* with converse and graphs then holds trivially. ◻

Section 6.8

Proposition 6.8.3. *With the same assumptions, the induced functor L^* of Theorem 6.8.2 commutes with graphs and converses. That is, the following diagrams commute:*

$$\begin{array}{ccc}
\boldsymbol{Rel}^{\mathcal{E}}_{(\Sigma,E)}(Q) & \xrightarrow{\;L^*\;} & \boldsymbol{Rel}^{\mathcal{F}}_{(\Sigma,E)}(LQ) \\
{\scriptstyle (-)_\circ}\uparrow & & \uparrow{\scriptstyle (-)_\circ} \\
\boldsymbol{Alg}^{\mathcal{E}}(\Sigma,E) & \xrightarrow[\;L\;]{} & \boldsymbol{Alg}^{\mathcal{F}}(\Sigma,E)
\end{array}$$

$$\begin{array}{ccc}
\boldsymbol{Rel}^{\mathcal{E}}_{(\Sigma,E)}(Q)^{op} & \xrightarrow{\;(L^*)^{op}\;} & \boldsymbol{Rel}^{\mathcal{F}}_{(\Sigma,E)}(LQ)^{op} \\
{\scriptstyle (-)^\circ}\downarrow & & \downarrow{\scriptstyle (-)^\circ} \\
\boldsymbol{Rel}^{\mathcal{E}}_{(\Sigma,E)}(Q) & \xrightarrow[\;L^*\;]{} & \boldsymbol{Rel}^{\mathcal{F}}_{(\Sigma,E)}(LQ)
\end{array}$$

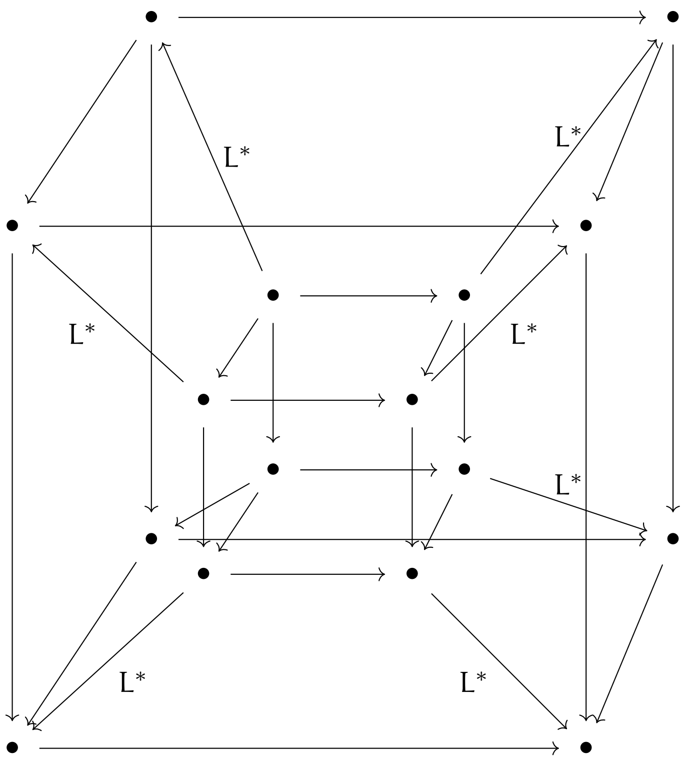
Proof. Start noting that the graph functor is identity on objects, so trivially $L^*(A)_\circ = L^*A = (L^*A)_\circ$ for every object A. For a morphism $f : A \to B$, in $\mathcal{E}$, consider the diagram:

$$\begin{array}{ccccccc}
LA \times LB & \overset{iso}{\text{—}} & L(A \times B) & \xrightarrow{\;L(f \times id^{\mathcal{E}}_B)\;} & L(B \times B) & \xrightarrow{\;L(id^{\mathbf{Rel}}_B)\;} & L(Q) \\
 & \diagdown\!\!\diagdown & \Big|\,{\scriptstyle iso} & & {\scriptstyle iso}\,\Big| & \nearrow{\scriptstyle id^{\mathbf{Rel}}_{LB}} & \\
 & & LA \times LB & \xrightarrow[\;Lf \times Lid^{\mathcal{E}}_B\;]{} & LB \times LB & &
\end{array}$$

Where $id^{\mathcal{E}}_B$ is the identity on B in $\mathcal{E}$ and $id^{\mathbf{Rel}}_B$ is the morphism of $\mathcal{E}$ corresponding to the formula that defines 1_B in $\mathbf{Rel}^{\mathcal{E}}_{\Sigma,E}(Q)$. The top row of the diagram is just $L^*(f)_\circ$, while the bottom one is $(L^*f)_\circ$. The left triangle commutes trivially, the center square commutes because L preserves products, the right triangle commutes because L preserves relational identities (previous proposition). Preservation of the converse follows trivially from the fact that any logical functor preserves products. □

Section 6.9

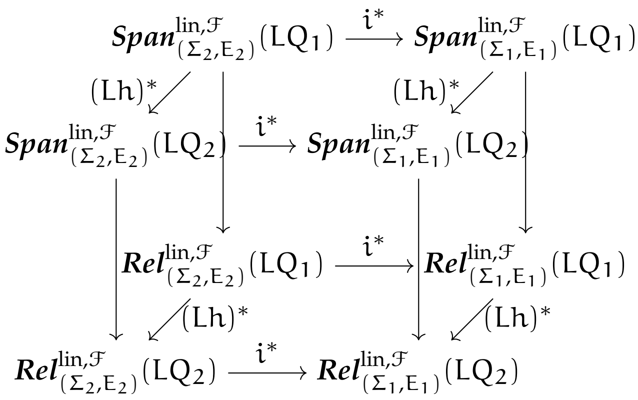
Theorem 6.9.1. *Let $\mathcal{E}$ be a topos, $h : Q_1 \to Q_2$ a morphism of internal commutative quantales, $i : (\Sigma_1, E_1) \to (\Sigma_2, E_2)$ a linear interpretation and $L : \mathcal{E} \to \mathcal{F}$ a logical functor. For the induced functors of Theorems 6.6.1, 6.6.3, 6.7.10, 6.7.11, 6.8.2 and 6.8.4, the following diagram commutes (be* *aware that in the hypercube below commutative squares involving* L^* *only commute up to isomorphism. Other squares commute up to equality):*



Where the inner cube is:

$$\begin{array}{ccccccc}
 & & \boldsymbol{Span}^{\mathrm{lin},\mathcal{E}}_{(\Sigma_2,E_2)}(Q_1) & \xrightarrow{i^*} & \boldsymbol{Span}^{\mathrm{lin},\mathcal{E}}_{(\Sigma_1,E_1)}(Q_1) & & \\
 & \swarrow^{h^*} & \downarrow & & & \swarrow^{h^*} & \downarrow \\
\boldsymbol{Span}^{\mathrm{lin},\mathcal{E}}_{(\Sigma_2,E_2)}(Q_2) & \xrightarrow{i^*} & & & \boldsymbol{Span}^{\mathrm{lin},\mathcal{E}}_{(\Sigma_1,E_1)}(Q_2) & & \\
\downarrow & & \boldsymbol{Rel}^{\mathrm{lin},\mathcal{E}}_{(\Sigma_2,E_2)}(Q_1) & \xrightarrow{i^*} & \downarrow & & \boldsymbol{Rel}^{\mathrm{lin},\mathcal{E}}_{(\Sigma_1,E_1)}(Q_1) \\
 & \swarrow^{h^*} & & & & \swarrow^{h^*} & \\
\boldsymbol{Rel}^{\mathrm{lin},\mathcal{E}}_{(\Sigma_2,E_2)}(Q_2) & \xrightarrow{i^*} & & & \boldsymbol{Rel}^{\mathrm{lin},\mathcal{E}}_{(\Sigma_1,E_1)}(Q_2) & &
\end{array}$$

And the outer cube is:



In both cases the vertical arrows are the functors of Theorem 6.5.1. Similar diagrams commute for affine, relevant and cartesian interpretations, relations and spans.

Proof. Here the notation $A \simeq B$ will denote that A and B are isomorphic. i^* trivially commutes with h^*, since the first is identity on morphisms and the second is identity on objects; for the very same reason, i^* commutes with V. V commutes with h^* because the former acts by postcomposition with a homomorphism of quantales, that commutes with joins and orders.

To show that $L^*i^* \simeq i^*L^*$, note that for morphisms this is trivial, being i^* the identity on them. Let then $\langle A, \sigma_j^A \rangle$ be an object of, say, $\mathbf{Rel}^{\mathcal{E}}_{(\Sigma_2,E_2)}(Q)$, and consider $L^*i^*\langle A, \sigma_j^A\rangle$. By definition this is equal to $L^*\langle A, i(\sigma^A)_{j'}\rangle$, where every $i(\sigma^A)_{j'}$ is a term derived from the σ_j^A, so a composition of σ_j^A (and eventually diagonals and projections, depending on the interpretation). Being L logical, operations of A get carried into operations of LA, hence $L^*\langle A, i(\sigma^A)_{j'}\rangle = \langle LA, Li(\sigma^A)_{j'}\rangle$ is an algebra of type (Σ_1, E_1) in $\mathbf{Rel}^{\mathcal{F}}_{(\Sigma_1,E_1)}(LQ)$. But, being $i(\sigma^A)_{j'}$ a composition of operations, projections and diagonals, and being L product preserving, it is $Li(\sigma^A)_{j'} \simeq i(\sigma^{LA})_{j'}$. Hence:

$$\begin{aligned} L^*i^*\langle A, \sigma_j^A\rangle &= \langle LA, Li(\sigma^A)_{j'}\rangle \\ &\simeq \langle LA, i(\sigma^{LA})_{j'}\rangle \\ &= i^*L^*\langle A, \sigma_j^A\rangle \end{aligned}$$

The proof is the same when L^* and i^* act on spans.

To prove that $L^*V \simeq V^*L^*$, consider the following logical theory:

$$T = (X, A, B, Q, \{\sigma_i^A\}_{\sigma_i \in \Sigma}, \{\sigma_i^B\}_{\sigma_i \in \Sigma}, f, g, \chi, \odot, \bigvee, k, R)$$

Where:

- $(X, A, B, Q, \{\sigma_i^A\}_{\sigma_i \in \Sigma}, \{\sigma_i^B\}_{\sigma_i \in \Sigma}, f, g, \chi, \odot, \bigvee, k)$ is the fragment that states that $(Q, \odot, k, \bigvee)$ is a quantale, that both $\langle A, \{\sigma_i^A\}_{\sigma_i \in \Sigma}\rangle$ and

$\langle B, \{\sigma_i^B\}_{\sigma_i \in \Sigma}\rangle$ are algebras of the required signature and that (X, f, g, χ) is an algebraic Q-span, with all the obvious axioms required to hold (see proof of Theorems 6.8.2 and 6.8.4 for details);

- R is a constant of type $Q^{A \times B}$ together with the axioms that say it is an algebraic Q-relation (again refer to the relational case in Theorem 6.8.2);
- The additional axiom:

$$R(a, b) = \bigvee \{\chi(m) \mid \exists m.(s(m) = a \wedge t(m) = b)\}$$

  Is satisfied.

This logical theory expresses exactly the fact that R is a relation coming from a span in the sense of the order functor, so if $R = V(X, s, t, \chi)$ then R and (X, s, t, χ) are a model for T. From this we get an isomorphism between LR and $V(LX, Lf, Lg, L\chi)$, and hence:

$$L^*V(X, f, g, \chi) = L^*R \simeq LR \simeq V(LX, Lf, Lg, L\chi)$$

Finally, to verify that $L^*h^* \simeq h^*L^*$, just note that it is possible to state what a quantale homomorphism is in terms of logical theories. This guarantees that if $h : Q_1 \to Q_2$ is a homomorphism of quantales, so is Lh. Everything then follows from the fact that h^* acts by postcomposition and L respects it. □

PROOFS FROM CHAPTER 7

Section 7.1

Proposition 7.1.7. *If* $h : Q_1 \to Q_2$ *is an injective quantale morphism (in* **Set***), the induced strict monoidal functor* $h^* : \textbf{\textit{Rel}}_{(\Sigma,E)}(Q_1) \to \textbf{\textit{Rel}}_{(\Sigma,E)}(Q)$ *is faithful.*

Proof. The action of the functor is:

$$R : A \times B \to Q_1 \;\mapsto\; h \circ R$$

Then:

$$h \circ R_1 = h \circ R_2 \Rightarrow R_1 = R_2$$

As h is monic. □

Proposition 7.1.8. *There is a quantale homomorphism from the Boolean quantale* **B** *to the Lawvere quantale* **L** *given by:*

$$\bot \mapsto \infty \qquad \top \mapsto 0 \tag{10}$$

Proof. Denote the function defined in the statement with h. We first note the monoid units of **B** and **L** are $\top$ and 0 respectively, so the monoid unit is preserved by h. For preservation of the multiplication, we check cases, noting we can skip one case by commutativity:

$$h(\bot \odot \bot) = h(\bot) = \infty = \infty + \infty = h(\bot) \odot h(\bot)$$

$$h(\bot \odot \top) = h(\bot) = \infty = 0 + \infty = h(\top) \odot h(\bot)$$

$$h(\top \odot \top) = h(\top) = 0 = 0 + 0 = h(\top) \odot h(\top)$$

Finally, we must check that joins are preserved. h preserves the bottom element, so empty joins are preserved. Singleton joins are trivial, so the only remaining case it to check the join of the whole of **B**. We have:

$$h\left(\bigvee\{\bot, \top\}\right) = h(\top) = 0 = \bigvee\{\infty, 0\} = \bigvee\{h(\bot), h(\top)\}$$

Where we recall that the ordering is reversed, so joins are infima. □

Proposition 7.1.9. *If* $U, V \subseteq X$ *and* d *is an internal monad in* $\textbf{\textit{Rel}}(\mathbf{C})$*, the composite* $V^\circ \circ d \circ U$ *is the infimum of the distances between elements in* U *and* V*.*

Proof. We have, by definition, that $U(*, a) = \bigvee\{k \mid a \in U\}$ and similarly $V^\circ(b, *) = \bigvee\{k \mid b \in V\}$. Again by definition we have:

$$\begin{aligned}(V^\circ \circ d \circ U)(*, *) &= \bigvee_{x,y} \left\{\bigvee\{k \mid x \in U\} \odot d(x,y) \odot \bigvee\{k \mid y \in U\}\right\} \\ &= \bigvee_{x,y} \{k \odot d(x,y) \odot k \mid (x \in U) \wedge (y \in V)\}\end{aligned}$$

And, remembering how $\bigvee$ and $\odot$ are defined in **C**,

$$\inf_{x,y}(0 + d(x,y) + 0 \mid (x \in U) \wedge (y \in V)) =$$
$$= \inf_{x,y}(d(x,y) \mid (x \in U) \wedge (y \in V)) \qquad \square$$

Proposition 7.1.11. *The euclidean distance on $\mathbb{R}^n$ respects convexity.*

Proof. The only thing to prove is compatibility with taking convex combinations, meaning that $d(x,y) \odot d(\tilde{x},\tilde{y}) \leqslant_{\mathbf{C}} d(x +^p \tilde{x}, y +^p \tilde{y})$. We rewrite the previous inequation in the equivalent form:

$$d(x +^p \tilde{x}, y +^p \tilde{y}) \leqslant d(x,y) + d(\tilde{x},\tilde{y})$$

Where addition and order are the usual ones in $\mathbb{R}$ (see Figure 20).

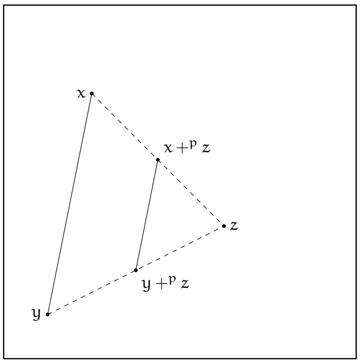


Figure 20: Convexity preservation of euclidean distance, geometrically explained.

First of all, we show that $d(x +^p z, y +^p z) \leqslant d(x,y)$ for every triple of points $x, y, z \in \mathbb{R}^n$ and $p \in [0,1]$. We unpack our definition of the operation $+^p$, obtaining:

$$\begin{aligned}
d(x +^p z, y +^p z) &= d(px - (1-p)z, py - (1-p)z) \\
&= \sqrt{\sum_{i=1}^{n}(px_i - (1-p)z_i - (py_i - (1-p)z_i))^2} \\
&= \sqrt{\sum_{i=1}^{n}(px_i - py_i)^2} \\
&= \sqrt{\sum_{i=1}^{n}p^2(x_i - y_i)^2} \\
&= p\sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \\
&\leqslant \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} = d(x,y)
\end{aligned}$$

Where the inequality follows from the fact that $p \in [0,1]$ (and hence $p^2 \in [0,1]$ as well).

From this, applying the triangular inequality for distances, we have:

$$\begin{aligned} d(x +^p \tilde{x}, y +^p \tilde{y}) &\leqslant d(x +^p \tilde{x}, y +^p \tilde{x}) + d(y +^p \tilde{x}, y +^p \tilde{y}) \\ &= d(x +^p \tilde{x}, y +^p \tilde{x}) + d(\tilde{x} +^{1-p} y, \tilde{y} +^{1-p} y) \\ &\leqslant d(x,y) + d(\tilde{x}, \tilde{y}) \end{aligned}$$

□

Proposition 7.1.12. *Let $X_1, \ldots, X_n$ be objects of $\mathbf{Rel}_{(\Sigma,E)}(Q)$ such that we can choose an internal monad d_i for every X_i. Then $d_1 \otimes \ldots \otimes d_n$ is an internal monad on $X_1 \otimes \ldots \otimes X_n$.*

Proof. Clearly, if $d_1, \ldots, d_n$ are internal monads in $\mathbf{Rel}_{(\Sigma,E)}(Q)$ they are all morphisms of $\mathbf{Rel}_{(\Sigma,E)}(Q)$, and so it is $d_1 \otimes \ldots \otimes d_n$. We just have to verify that $d_1 \otimes \ldots \otimes d_n$ respects the inequalities in (9). For identity, we have:

$$\begin{aligned} &id_{X_1 \otimes \ldots \otimes X_n}((a_1, \ldots a_n), (b_1, \ldots, b_n)) = \\ &\qquad = \bigvee\{k \mid (a_1, \ldots, a_n) = (b_1, \ldots, b_n)\} \\ &\qquad = \bigvee\{k \mid (a_1 = b_1) \wedge \cdots \wedge (a_n = b_n)\} \\ &\qquad = \bigvee\{k \mid (a_1 = b_1)\} \odot \ldots \odot \bigvee\{k \mid (a_n = b_n)\} \\ &\qquad \leqslant d_1(a_1, b_1) \odot \ldots \odot d_n(a_n, b_n) \\ &\qquad = (d_1 \otimes \ldots \otimes d_n)((a_1, \ldots a_n), (b_1, \ldots, b_n)) \end{aligned}$$

With regard to composition:

$$\begin{aligned} &((d_1 \otimes \ldots \otimes d_n) \circ (d_1 \otimes \ldots \otimes d_n))((a_1, \ldots a_n), (c_1, \ldots, c_n)) = \\ &\quad = \bigvee_{x_1, \ldots, x_n} \{(d_1 \otimes \ldots \otimes d_n)((a_1, \ldots a_n), (x_1, \ldots, x_n)) \odot \\ &\qquad\qquad \odot (d_1 \otimes \ldots \otimes d_n)((x_1, \ldots x_n), (c_1, \ldots, c_n))\} \\ &\quad = \bigvee_{x_1, \ldots, x_n} \{d_1(a_1, x_1) \odot \ldots \odot d_n(a_n, x_n) \odot \\ &\qquad\qquad \odot d_1(x_1, c_1) \odot \ldots \odot d_n(x_n, c_n) \\ &\quad = \bigvee_{x_1}\{d_1(a_1, x_1) \odot d_1(x_1, c_1)\} \odot \ldots \odot \bigvee_{x_n}\{d_n(a_n, x_n) \odot d_n(x_n, c_n)\} \\ &\quad \leqslant d_1(a_1, c_1) \odot \ldots \odot d_n(a_n, c_n) \\ &\quad = (d_1 \otimes \ldots \otimes d_n)((a_1, \ldots a_n), (c_1, \ldots, c_n)) \end{aligned}$$

Note that these proofs make sense in any topos, hence the result still holds for relations with underlying topos different from **Set**. □

Section 7.3

Theorem A.0.3. *Let $\mathcal{C}$ be a groupoid. Every quantale Q in **Set** can be canonically lifted to a quantale Q in $\mathbf{Set}^{\mathcal{C}^{op}}$.*

Proof. If $\mathcal{C}$ is a groupoid, then $\mathbf{Set}^{\mathcal{C}^{op}}$ inherits its elementary topos structure pointwise from **Set**. This in particular means that if

$$F : \mathcal{C} \to \mathcal{D}$$

is a groupoid functor, then $F^* : \mathbf{Set}^{\mathcal{D}^{op}} \to \mathbf{Set}^{\mathcal{C}^{op}}$ is logical. For every groupoid $\mathcal{C}$ there is a trivial groupoid homomorphism $F : \mathcal{C} \to \mathbf{1}$, where $\mathbf{1}$ is the category having one element and the identity morphism on it. Being $\mathbf{Set} = \mathbf{Set}^{\mathbf{1}}$ then we obtain a logical functor $F^* : \mathbf{Set} \to \mathbf{Set}^{\mathcal{C}^{op}}$. We moreover know that a logical functor preserves models: If $\mathbb{T}$ is a logical theory, and we denote with $\text{mod}_{\mathbb{T}}(\mathcal{E})$ the models of $\mathbb{T}$ in the topos $\mathcal{E}$, the statement above means that every time we have a logical functor $L : \mathcal{E} \to \mathcal{F}$, having $X \in \text{mod}_{\mathbb{T}}(\mathcal{E})$ implies $LX \in \text{mod}_{\mathbb{T}}(\mathcal{F})$. The equations defining a quantale clearly form a logical theory, call it $\mathbb{T}$. In our case this means that if Q is a model of $\mathbb{T}$ in **Set** (that is, a quantale in the set-theoretic sense), then $F^*(Q)$ is an internal quantale in $\mathbf{Set}^{\mathcal{C}^{op}}$, as we wanted. □

Lemma A.0.4. *Let $\mathcal{E}$ be a topos, and $t : 1 \to \Omega$ its subobject classifier. Then $\mathbf{Rel}^{\mathcal{E}}_{(\emptyset,\emptyset)}(\Omega)$ is equivalent to $\mathbf{Rel}(\mathcal{E})$, the category of relations on $\mathcal{E}$.*

Proof. First of all recall that $(\Omega, \sup, \wedge, t)$ is an internal locale in every topos, so we are legitimate to consider $\mathbf{Rel}^{\mathcal{E}}_{(\emptyset,\emptyset)}(\Omega)$. Moreover, $\mathbf{Rel}(\mathcal{E})$ has the same objects as $\mathcal{E}$ and, as morphisms $A \xrightarrow{R} B$, subobjects of $A \times B$. Being $\mathcal{E}$ a topos, there is a natural isomorphism:

$$\mathbf{Sub}(A \times C) \simeq \hom(A \times C, \Omega)$$

and we can then define a functor $F : \mathbf{Rel}(\mathcal{E}) \to \mathbf{Rel}^{\mathcal{E}}_{(\emptyset,\emptyset)}(\Omega)$ as being identity on objects and sending every morphism R to its classifier χ_R.

This correspondence is clearly bijective on objects and morphisms, so if we prove functoriality the isomorphism we want to prove will follow trivially. For composition, we have to prove $\chi_{S \circ R} = \chi_S \circ \chi_R$. The left hand side and the right hand side are the morphisms of type $(A \times C) \to \Omega$ defined, respectively, as:

$$(a, c) \xrightarrow{\quad \chi_{S \circ R} \quad} \exists b(\chi_R(a, b) \wedge \chi_S(b, c))$$

$$(a, c) \xrightarrow{\quad \chi_S \circ \chi_R \quad} \sup(\{\chi_R(a, b) \wedge \chi_S(b, c) | b \in B\})$$

We can prove these morphisms to be equal showing that they classify the same subobject, that is, proving they are equivalent as internal formulas. We know (see, as an instance, [27, p. 428]) that for every variable $\mathcal{P}$ of type Ω^{Ω} and every variable Z of type Ω,

$$Z = \sup \mathcal{P} \qquad \text{is equivalent to} \qquad \forall z (z \in Z \Leftrightarrow \exists T (T \in \mathcal{P} \wedge z \in T)).$$

Taking $\mathcal{P} = \{\chi_R(a, b) \wedge \chi_S(b, c) | b \in B\}$ and $Z = \sup(\mathcal{P})$, we substitute in the previous equivalence. On the left we obtain a tautology, namely

$\sup(\mathcal{P}) = \sup(\mathcal{P})$. This means that that what we obtain on the right hand side of the previous equivalence, that is:

$$\forall z(z \in \sup(\{\chi_R(a,b) \wedge \chi_S(b,c) | b \in B\}) \\ \Leftrightarrow \exists T(T \in \{\chi_R(a,b) \wedge \chi_S(b,c) | b \in B\} \wedge z \in T))$$

Is equivalent to a tautology, and so it is a validity in $\mathcal{E}$. Noting that the formula $T \in \{\chi_R(a,b) \wedge \chi_S(b,c) | b \in B\}$ is equivalent to the formula $\exists b'(T = \chi_R(a,b') \wedge \chi_S(b',c))$, we can focus on the right hand side of the equivalence and calculate:

$$\begin{aligned} &\exists T(T \in \{\chi_R(a,b) \wedge \chi_S(b,c) | b \in B\} \wedge z \in T) \Rightarrow \\ &\quad \Rightarrow \exists T(\exists b'(T = \chi_R(a,b') \wedge \chi_S(b',c)) \wedge z \in T) \\ &\quad \Rightarrow \exists T(\exists b'((T = \chi_R(a,b') \wedge \chi_S(b',c)) \wedge (z \in T))) \\ &\quad \Rightarrow \exists T \exists b'(z \in (\chi_R(a,b') \wedge \chi_S(b',c))) \\ &\quad \Rightarrow \exists b'(z \in (\chi_R(a,b') \wedge \chi_S(b',c))) \end{aligned}$$

And, in the opposite direction,

$$\begin{aligned} &\exists b'(z \in (\chi_R(a,b') \wedge \chi_S(b',c))) \Rightarrow \\ &\quad \Rightarrow \exists T(T = (\chi_R(a,b') \wedge \chi_S(b',c)) \wedge z \in T) \\ &\quad \Rightarrow \exists T(\exists b'(T = (\chi_R(a,b') \wedge \chi_S(b',c))) \wedge z \in T) \\ &\quad \Rightarrow \exists T(T \in \{\chi_R(a,b) \wedge \chi_S(b,c) | b \in B\} \wedge z \in T) \end{aligned}$$

Obtaining an equivalence. Substituting, we get the validity:

$$\forall z(z \in \sup(\{\chi_R(a,b) \wedge \chi_S(b,c) | b \in B\}) \Leftrightarrow \\ \Leftrightarrow \exists b'(z \in (\chi_R(a,b') \wedge \chi_S(b',c)))) \qquad (13)$$

Now, $\sup(\{\chi_R(a,b) \wedge \chi_S(b,c) | b \in B\})$, as well as $\chi_R(a,b') \wedge \chi_S(b',c)$, are morphisms of type $(A \times C) \to \Omega$ and, since the writing $a \in A$ is defined only for a of type $K \to X$ and A of type $K \to \Omega^X$ (for some K, X objects of $\mathcal{E}$), it is clear that z has to be a variable of type 1, that is, the unique morphism $(A \times C) \to 1$ in $\mathcal{E}$. Then the formula $z \in \sup(\{\chi_R(a,b) \wedge \chi_S(b,c) | b \in B\}$ is the bottom row of the diagram:

$$\begin{array}{ccccc} S & \longrightarrow & 1 & = & 1 \\ \downarrow & & \downarrow t & & \downarrow t \\ A \times C & \xrightarrow[\langle z, \sup(\{\chi_R(a,b) \wedge \chi_S(b,c) | b \in B\})\rangle]{} & 1 \times \Omega & \underset{\in}{=} & \Omega \end{array}$$

That is clearly equivalent to:

$$\begin{array}{ccc} S & \longrightarrow & 1 \\ \downarrow & & \downarrow t \\ A \times C & \xrightarrow[\sup(\{\chi_R(a,b) \wedge \chi_S(b,c) | b \in B\})]{} & 1 \times \Omega \end{array}$$

Hence the formulas in the bottom row classify the same morphism. We can do the same reasoning for $z \in (\chi_R(a,b') \wedge \chi_S(b',c))$ and $\chi_R(a,b') \wedge \chi_S(b',c)$ and substitute everything in Equation 13, obtaining:

$$\forall z(\sup(\{\chi_R(a,b) \wedge \chi_S(b,c) | b \in B\}) \Leftrightarrow \exists b'(\chi_R(a,b') \wedge \chi_S(b',c)))$$

Finally, we can drop the quantifier since the variable z does not appear anymore in its scope, and get:

$$\sup(\{\chi_R(a,b) \wedge \chi_S(b,c) | b \in B\}) \Leftrightarrow \exists b'(\chi_R(a,b') \wedge \chi_S(b',c))$$

Since this formula was equivalent to a validity, it is a validity itself. This means that it holds for all $(a,c) \in A \times C$, proving our functor respects compositionality. With regard to the identity, just note that, for $k : A \to B$ in $\mathbf{Rel}^{\mathcal{E}}_{(\emptyset,\emptyset)}(\Omega)$, $k = \chi_R$ for some $R : A \to B$ in $\mathbf{Rel}(\mathcal{E})$. Hence:

$$\begin{aligned} F(id_A) \circ k \circ F(id_B) &= F(id_A) \circ \chi_R \circ F(id_B) \\ &= F(id_A) \circ F(R) \circ F(id_B) \\ &= F(id_A \circ R \circ id_B) \\ &= F(R) = \chi_R = k \end{aligned}$$

And this concludes the proof. □

Theorem 7.3.12. *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 isomorphic to the category of internal relations over the regular category of internal (Σ, E)-algebras in $\mathcal{E}$.*

Proof. It is a consequence of [19, Thm. 5.11] that, being $\mathcal{E}$ regular, $\mathbf{Alg}_{(\Sigma,E)}(\mathcal{E})$ is regular too, so we can legitimately consider the category $\mathbf{Rel}(\mathbf{Alg}_{(\Sigma,E)}(\mathcal{E}))$. Note that we can apply [19, Thm. 5.11] because the categories of algebras we are considering are, as usual, not internal: The category $\mathbf{Rel}(\mathbf{Alg}_{(\Sigma,E)}(\mathcal{E}))$ is just the usual relation construction built on the regular category $\mathcal{E}$. Using the internal language of $\mathcal{E}$, we can describe morphisms of $\mathbf{Alg}_{(\Sigma,E)}(\mathcal{E})$ as just morphisms of $\mathcal{E}$ that preserve operations in Σ. Now, we define:

$$F : \mathbf{Rel}(\mathbf{Alg}_{(\Sigma,E)}(\mathcal{E})) \to \mathbf{Rel}^{\mathcal{E}}_{(\Sigma,E)}(\Omega)$$

To be identity on objects and as in Lemma A.0.4 on morphisms. If R is a morphism in $\mathbf{Rel}(\mathbf{Alg}_{(\Sigma,E)}(\mathcal{E}))$, it is by definition a subalgebra of some product, say $A \times B$. This amounts to ask R to be closed with respect to the operations, hence:

$$(a_i, b_i) \in R \Rightarrow (\sigma(a_1, \dots, a_n), \sigma(b_1, \dots, b_n)) \in R$$

For $a_1, \dots, a_n, b_1, \dots, b_n$ and any operation $\sigma \in \Sigma$. The consequent holds when all the antecedents are satisfied, so we can rewrite it using the internal language as:

$$\wedge_i((a_i, b_i) \in R) \implies (\sigma^A(a_1, \dots, a_n), \sigma^B(b_1, \dots, b_n)) \in R$$

From here and using the fact that Ω is an Heyting algebra we obtain the following chain of equivalences:

$$\bigwedge_i ((a_i, b_i) \in R(a_i, b_i)) \implies (\sigma^A(a_1, \dots, a_n), \sigma^B(b_1, \dots, b_n)) \in R$$
$$\chi(\bigwedge_i((a_i, b_i) \in R(a_i, b_i))) \leqslant \chi((\sigma^A(a_1, \dots, a_n), \sigma^B(b_1, \dots, b_n)) \in R)$$
$$\bigwedge_i \chi(((a_i, b_i) \in R(a_i, b_i)) \leqslant \chi_R(\sigma^A(a_1, \dots, a_n), \sigma^B(b_1, \dots, b_n))$$
$$\bigwedge_i \chi_R(a_i, b_i) \leqslant \chi_R(\sigma^A(a_1, \dots, a_n), \sigma^B(b_1, \dots, b_n))$$
$$\bigwedge_i F(R)(a_i, b_i) \leqslant F(R)(\sigma^A(a_1, \dots, a_n), \sigma^B(b_1, \dots, b_n))$$

Hence R is a morphism in $\mathbf{Rel}(\mathbf{Alg}_{(\Sigma,E)}(\mathcal{E}))$ if and only if its image through F is a morphism in $\mathbf{Rel}^{\mathcal{E}}_{(\Sigma,E)}(\Omega)$, proving that F is well defined. Since there is [19, Thm. 2.1] a functor from $\mathbf{Alg}_{(\Sigma,E)}(\mathcal{E})$ to $\mathcal{E}$ that preserves finite limits and epi-mono factorizations, this means that morphisms in $\mathbf{Rel}(\mathbf{Alg}_{(\Sigma,E)}(\mathcal{E}))$ compose exactly as morphisms in $\mathbf{Rel}(\mathcal{E})$, and the proof of functoriality is can be carried out exactly as in Lemma A.0.4. □

[1] S. Abramsky and A. Brandenburger. "The sheaf-theoretic structure of non-locality and contextuality." In: *New Journal of Physics* 13 (2011). ISSN: 13672630. DOI: 10.1088/1367-2630/13/11/113036. arXiv: 1102.0264 (cit. on p. 46).

[2] S. Abramsky and B. Coecke. "A Categorical Semantics of Quantum Protocols." In: *Logic in Computer Science, 2004. Proceedings of the 19th Annual IEEE Symposium on*. IEEE. IEEE, 2004, pp. 415–425. ISBN: 0-7695-2192-4. DOI: 10.1109/LICS.2004.1319636. arXiv: 0402130 [quant-ph]. URL: http://ieeexplore.ieee.org/ielx5/9221/29239/01319636.pdf?tp=%7B%5C&%7Darnumber=1319636%7B%5C&%7Disnumber=29239 (cit. on pp. 5, 6).

[3] S. Abramsky and B. Coecke. "Abstract Physical Traces." In: *Theory and Applications of Categories* 14.6 (2005). ISSN: 1201561X. arXiv: 0910.3144 (cit. on p. 20).

[4] S. Abramsky and B. Coecke. "Categorical Quantum Mechanics." In: *Handbook od Categorical Logic*. Ed. by K. Engesser, D. M. Gabbay, and D. Lehmann. Elsevier, 2009, pp. 261–323 (cit. on p. 5).

[5] S. Abramsky and R. Duncan. "A Categorical Quantum Logic." In: *Mathematical Structures in Computer Science* 16.03 (June 2006), p. 469. ISSN: 0960-1295. DOI: 10.1017/S0960129506005275. arXiv: 0512114 [quant-ph]. URL: http://www.journals.cambridge.org/abstract%7B%5C_%7DS0960129506005275 (cit. on p. 6).

[6] S. Abramsky and C. Heunen. "H*-Algebras and Nonunital Frobenius Algebras: First Steps in Infinite-Dimensional Categorical Quantum Mechanics." In: *Mathematical Foundations of Information Flow* 71 (2012), pp. 1–24. DOI: 10.1090/psapm/071/599. arXiv: 1011.6123 (cit. on p. 61).

[7] S. Abramsky and M. Sadrzadeh. "Semantic Unification: A Sheaf Theoretic Approach to Natural Language." In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 8222 (2014), pp. 1–13. ISSN: 16113349. DOI: 10.1007/978-3-642-54789-8_1. arXiv: 1403.3351 (cit. on p. 146).

[8] B. Adams and M. Raubal. "A Metric Conceptual Space Algebra." In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 5756 LNCS. 2009, pp. 51–68. ISBN: 364203831X.

DOI: 10.1007/978-3-642-03832-7_4. URL: http://link.springer.com/10.1007/978-3-642-03832-7%7B%5C_%7D4 (cit. on p. 31).

[9] B. Ambridge and E. V. M. Lieven. *Child Language Acquisition*. Cambridge: Cambridge University Press, 2011, p. 466. ISBN: 9780521768047 (cit. on p. 148).

[10] D. Ashoush. "Categorical Models of Meaning : Accommodating for Lexical Ambiguity and Entailment." MSc Thesis. University of Oxford, 2015 (cit. on p. 6).

[11] J. C. Baez. *Network Theory (Part 1)*. URL: https://johncarlosbaez.wordpress.com/2011/03/04/network-theory-part-1/ (visited on 07/16/2017) (cit. on p. 63).

[12] J. C. Baez and J. Erbele. "Categories in Control." In: *Theory and Applications of Categories* 30.24 (2015), pp. 836–881 (cit. on p. 5).

[13] J. C. Baez and B. Fong. "A Compositional Framework for Passive Linear Networks." In: *arXiv preprint arXiv:1504.05625* (2015) (cit. on p. 5).

[14] J. C. Baez, B. Fong, and B. S. Pollard. "A Compositional Framework for Markov Processes." In: *Journal of Mathematical Physics* 57.033301 (Mar. 2016). ISSN: 0022-2488. DOI: 10.1063/1.4941578. arXiv: 1508.06448. URL: http://aip.scitation.org/doi/10.1063/1.4941578 (cit. on p. 5).

[15] J. C. Baez and M. Stay. "Physics, Topology, Logic and Computation: A Rosetta Stone." In: *arXiv preprint* (Mar. 2009), pp. 1–73. arXiv: 0903.0340. URL: http://arxiv.org/abs/0903.0340 (cit. on p. 6).

[16] E. Balkır. "Using Density Matrices in a Compositional Distributional Model of Meaning." MSc Thesis. University of Oxford, 2014 (cit. on pp. 6, 10).

[17] D. Bankova. "Comparing Meaning in Language and Cognition: Hypononymy, Concept Combination, Asymmetric Similarity." PhD thesis. University of Oxford, 2015 (cit. on p. 85).

[18] D. Bankova, B. Coecke, M. Lewis, and D. Marsden. "Graded Entailment for Compositional Distributional Semantics." In: *arXiv preprint* (2016). arXiv: 1601.04908. URL: http://arxiv.org/abs/1601.04908 (cit. on p. 10).

[19] M. Barr. "Exact Categories." In: *Exact categories and categories of sheaves*. Springer, 1971, pp. 1–120 (cit. on pp. 131, 183, 184).

[20] L. W. Barsalou. "Ideals, Central Tendency, and Frequency of Instantiation as Determinants of Graded Structure in Categories." In: *Journal of experimental psychology: learning, memory, and cognition* 11.4 (1985), p. 629 (cit. on p. 114).

[21] J. Bénabou. *Introduction to Bicategories*. Tech. rep. Reports of the Midwest Category Seminar, 1967, pp. 1–77. DOI: 10.1007/BFb0074299. URL: http://link.springer.com/10.1007/BFb0074299 (cit. on p. 110).

[22] Y. Bengio, R. Ducharme, P. Vincent, and C. Jauvin. "A Neural Probabilistic Language Model." In: *Journal of Machine Learning Research* 3 (2003), pp. 1137–1155. URL: http://www.jmlr.org/papers/volume3/bengio03a/bengio03a.pdf (cit. on p. 157).

[23] O. Berger-Tal and S. Bar-David. "Recursive Movement Patterns: Review and Synthesis Across Species." In: *Ecosphere* 6.9 (2015), art149. ISSN: 2150-8925. DOI: 10.1890/ES15-00106.1. URL: http://www.esajournals.org/doi/full/10.1890/ES15-00106.1 (cit. on p. 151).

[24] J. Bolt, B. Coecke, F. Genovese, M. Lewis, D. Marsden, and R. Piedeleu. "Interacting Conceptual Spaces." In: *Proceedings of the 2016 Workshop on Semantic Spaces at the Intersection of NLP, Physics and Cognitive Science, SLPCS@QPL 2016, Glasgow, Scotland, 11th June 2016.* Ed. by D. Kartsaklis, M. Lewis, and L. Rimell. Vol. 221. EPTCS. Aug. 2016, pp. 11–19. DOI: 10.4204/EPTCS.221.2. arXiv: 1608.01402 (cit. on pp. 6, 9, 10, 31, 46).

[25] F. Bonchi, P. Sobociński, and F. Zanasi. "Full Abstraction for Signal Flow Graphs." In: *ACM SIGPLAN Notices* 50.1 (2015), pp. 515–526 (cit. on pp. 5, 65).

[26] F. Borceux. *Handbook of Categorical Algebra: Volume 2, Categories and Structures*. Cambridge: Cambridge University Press, 1994, p. 456. ISBN: 0-521-44179 (cit. on pp. 21, 35, 130).

[27] F. Borceux. *Handbook of Categorical Algebra: Volume 3, Categories of Sheaves*. Cambridge: Cambridge University Press, 1994, p. 540. ISBN: 0-521-44180-3 (cit. on pp. 64, 181).

[28] H. J. Briegel and G. De las Cuevas. "Projective Simulation For Artificial Intelligence." In: *Scientific reports* 2 (2012), p. 400. ISSN: 2045-2322. DOI: 10.1038/srep00400. arXiv: 1104.3787. URL: http://www.nature.com/srep/2012/120515/srep00400/full/srep00400.html (cit. on p. 157).

[29] W. Buszkowski. "Completeness Results for Lambek Syntactic Calculus." In: *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* 32.1-5 (1986), pp. 13–28. ISSN: 00443050. DOI: 10.1002/malq.19860320104. URL: http://doi.wiley.com/10.1002/malq.19860320104 (cit. on p. 13).

[30] W. Buszkowski. "Lambek Grammars Based on Pregroups." In: *Proceedings of the 4th International Conference on Logical Aspects of Computational Linguistics* (2001), pp. 95–109. ISSN: 16113349. URL: http://dl.acm.org/citation.cfm?id=645668.665052 (cit. on p. 16).

[31] W. Buszkowski and K. Moroz. “Pregroup Grammars and Context-free Grammars.” In: *Computational Algebraic Approaches to Natural Language*. 2008, pp. 1–22 (cit. on p. 147).

[32] A. Carboni and R. Walters. “Cartesian Bicategories I.” In: *Journal of Pure and Applied Algebra* 49.1-2 (Nov. 1987), pp. 11–32. ISSN: 00224049. DOI: 10.1016/0022-4049(87)90121-6. URL: http://linkinghub.elsevier.com/retrieve/pii/0022404987901216 (cit. on p. 65).

[33] C. Casadio and J. Lambek. “A Computational Algebraic Approach to Latin Grammar.” In: *Research on Language and Computation* 3.1 (Apr. 2005), pp. 45–60. ISSN: 15707075. DOI: 10.1007/s11168-005-1286-0. URL: http://link.springer.com/10.1007/s11168-005-1286-0 (cit. on p. 3).

[34] N. Chomsky. “Three Models for the Description of Language.” In: *IRE Transactions on Information Theory* 2.3 (Sept. 1956), pp. 113–124. ISSN: 0096-1000. DOI: 10.1109/TIT.1956.1056813. URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1056813 (cit. on pp. 3, 147).

[35] N. Chomsky. *Aspects of the Theory of Syntax*. Vol. 11. no 11. 1965, p. 251. ISBN: 0262530074. DOI: 10.1016/0732-118X(86)90008-5. arXiv: arXiv:1011.1669v3 (cit. on p. 147).

[36] N. Chomsky. “The Minimalist Program.” In: *Journal of Linguistics* 34.1 (1995), pp. 213–226. ISSN: 00222267. DOI: 10.1017/S0022226797006889. arXiv: arXiv:1011.1669v3 (cit. on p. 148).

[37] N. Chomsky. “Derivation by phase.” In: *Ken Hale: A life in language* 1998 (2001), pp. 1–52 (cit. on p. 148).

[38] K. Chvalovský and R. Horčík. “Full Lambek Calculus with Contraction is Undecidable.” In: *The Journal of Symbolic Logic* 81.02 (June 2016), pp. 524–540. ISSN: 0022-4812. DOI: 10.1017/jsl.2015.18. URL: http://www.journals.cambridge.org/abstract%7B%5C_%7DS0022481215000183 (cit. on p. 18).

[39] S. Clark, B. Coecke, and M. Sadrzadeh. “A Compositional Distributional Model of Meaning.” In: *Proceedings of the Second Quantum Interaction Symposium (QI-2008)* Schuetze 1998 (2008), pp. 133–140. arXiv: arXiv:1003.4394v1 (cit. on pp. 3, 5, 6).

[40] M. M. Clementino and W. Tholen. “Metric, Topology and Multicategory: A Common approach.” In: *Journal of Pure and Applied Algebra* 179.1 (2003), pp. 13–47 (cit. on p. 63).

[41] B. Coecke, F. Genovese, M. Lewis, and D. Marsden. “Appications of Generalized Relations in Linguistics and Cognition.” In: *Logic, Language, Information, and Computation: 24th International Workshop, WoLLIC 2017, London, UK, July 18-21, 2017, Proceedings*. Ed. by J. Kennedy and R. J. de Queiroz. Heidelberg: Springer, 2017, pp. 256–270. ISBN: 978-3-662-55386-2. DOI:

10.1007/978-3-662-55386-2_18. URL: http://link.springer.com/10.1007/978-3-662-55386-2%7B%5C_%7D18 (cit. on p. 10).

[42] B. Coecke, E. Grefenstette, and M. Sadrzadeh. "Lambek vs. Lambek: Functorial Vector Space Semantics and String Diagrams for Lambek Calculus." In: *Annals of Pure and Applied Logic* 164.11 (2013), pp. 1079–1100. ISSN: 01680072. DOI: 10.1016/j.apal.2013.05.009. arXiv: 1302.0393 (cit. on pp. 6, 13).

[43] B. Coecke and A. Kissinger. *Picturing Quantum Processes. A First Course in Quantum Theory and Diagrammatic Reasoning*. Cambridge: Cambridge University Press, 2017 (cit. on pp. 5, 26, 35, 66).

[44] B. Coecke, M. Sadrzadeh, and S. Clark. "Mathematical Foundations for a Compositional Distributional Model of Meaning." In: *Linguistic Analysis* 36.1-4 (2011), pp. 345–384. arXiv: arXiv:1003.4394v1. URL: http://www.cl.cam.ac.uk/%7B~%7Dsc609/pubs/lambek11.pdf%7B%5C%%7D5Cnhttp://arxiv.org/abs/1003.4394 (cit. on pp. 5, 24).

[45] J. Colapinto. *The Interpreter: Has a Remote Amazonian Tribe Upended our Understanding of Language?* New York, Apr. 2007. URL: http://www.newyorker.com/magazine/2007/04/16/the-interpreter-2 (cit. on pp. 149, 150).

[46] H. B. Coonce. *The Mathematics Genealogy Project*. URL: https://www.genealogy.math.ndsu.nodak.edu/index.php (visited on 07/16/2017) (cit. on p. 118).

[47] H. S. M. Coxeter. *Introduction to Geometry*. New York: John Wiley and Sons, Inc., 1969, p. 469. URL: http://plouffe.fr/simon/math/Wiley%20-%20Coxeter%20-%20Introduction%20To%20Geometry.pdf (cit. on p. 47).

[48] J. R. Curran. "From Distributional to Semantic Similarity." PhD Thesis. University of Edinburgh, 2003, p. 177. DOI: 10.1.1.10.6068. URL: http://www.era.lib.ed.ac.uk/handle/1842/563 (cit. on p. 5).

[49] R. Dale, C. Kehoe, and M. J. Spivey. "Graded Motor Responses in the Time Course of Categorizing Atypical Exemplars." In: *Memory & Cognition* 35.1 (2007), pp. 15–28 (cit. on p. 114).

[50] P. Dekker. *Scopes in Discourse*. 1999. DOI: 10.1.1.132.3841. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.132.3841%7B%5C&%7Drep=rep1%7B%5C&%7Dtype=pdf (cit. on p. 142).

[51] R. Diaconescu. "Axiom of Choice and Complementation." In: *Proceedings of the American Mathematical Society* 354.1 (1975), pp. 151–178. ISSN: 1088-6850. DOI: 10.1090/S0002-9947-01-02810-0. URL: http://www.ams.org/tran/2002-354-01/S0002-9947-01-02810-0/ (cit. on p. 75).

[52] M. Dostal and M. Sadrzadeh. *Many Valued Generalised Quantifiers for Natural Language in the DisCoCat Model*. Tech. rep. Queen Mary University of London, 2016 (cit. on p. 63).

[53] M. Dummett. *Frege: Philosophy of Language*. Cambridge, Massachusetts: Harvard University Press, 1993, p. 752. ISBN: 0674319311. DOI: 10.4324/9780203519615. URL: http://books.google.com/books?id=EYP7uCZIRQYC%7B%5C&%7Dpgis=1 (cit. on p. 3).

[54] J. Erbele. "Categories in Control: Applied PROPs." PhD thesis. University of California Riverside, 2016 (cit. on p. 65).

[55] D. L. Everett. "Pirahã." In: *Handbook of Amazonian Languages, Vol. 1*. Ed. by D. C. Derbshire and G. K. Pullum. Berlin: Mouton de Gruyter, 1986, p. 646. ISBN: 3-11-010257-9. URL: http://etnolinguistica.wdfiles.com/local--files/biblio%7B%5C%%7D3Apopjes-1986-canela/popjes%7B%5C_%7D1986%7B%5C_%7Dcanela.pdf (cit. on p. 150).

[56] D. L. Everett. "A Língua Pirahã e a Teoria da Sintaxe: Descrição, Perspectivas e Teoria." PhD Thesis. Universidade Estadual de Campinas, 1992, p. 400 (cit. on p. 150).

[57] D. L. Everett. "Cultural Constraints on Grammar and Cognition in Pirahã." In: *Current Anthropology* 46.4 (2005), pp. 621–646. ISSN: 0011-3204. DOI: 10.1086/431525. URL: http://www.journals.uchicago.edu/doi/abs/10.1086/431525%7B%5C%%7D5Cnhttp://www.journals.uchicago.edu/doi/10.1086/431525 (cit. on p. 150).

[58] D. L. Everett. *Recursion and Human Thought: Why the Pirahã don't have Numbers*. 2007. URL: https://www.edge.org/conversation/daniel%7B%5C_%7Dl%7B%5C_%7Deverett-recursion-and-human-thought (visited on 07/22/2017) (cit. on pp. 150–152).

[59] D. L. Everett. "Pirahã Culture and Grammar: A Response to Some Criticisms." In: *Language* 85.2 (2009), pp. 405–442. ISSN: 1535-0665. DOI: 10.1353/lan.0.0104. arXiv: arXiv:1011.1669v3. URL: http://muse.jhu.edu/content/crossref/journals/language/v085/85.2.everett.html (cit. on p. 152).

[60] D. L. Everett. *Language: The Cultural Tool*. London: Profile Books, 2012, p. 368. ISBN: 978-1846682674 (cit. on p. 152).

[61] D. L. Everett. *Monolingual Fieldwork Demonstration*. 2013. URL: https://www.youtube.com/watch?v=sYpWp7g7XWU (visited on 07/22/2017) (cit. on p. 149).

[62] B. Fong. "Decorated Cospans." In: *Theory and Applications of Categories* 30.33 (2015), pp. 1096–1120 (cit. on p. 63).

[63] B. Fong. "The Algebra of Open and Interconnected Systems." PhD thesis. University of Oxford, 2016 (cit. on p. 63).

[64] B. Fong, D. I. Spivak, and R. Tuyéras. "Backprop as Functor: A Compositional Perspective on Supervised Learning." In: (2017), pp. 1–33. arXiv: 1711.10455 (cit. on p. 157).

[65] M. C. Frank, D. L. Everett, E. Fedorenko, and E. Gibson. "Number as a Cognitive Technology: Evidence from Pirahã Language and Cognition." In: *Cognition* 108.3 (2008), pp. 819–824. ISSN: 00100277. DOI: 10.1016/j.cognition.2008.04.007 (cit. on p. 150).

[66] G. Frege. "Begriffsschrift." In: *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931*. Ed. by J. van Heijenoort. 4th. Cambridge, Massachusetts: Harvard University Press, 1967, pp. 1–82. ISBN: 0-674-32449-8 (cit. on p. 3).

[67] P. J. Freyd and A. Scedrov. *Categories, Allegories*. Vol. 39. Elsevier, 1990 (cit. on p. 65).

[68] Gaius Suetonius Tranquillus. *De Vitis Caesarum*. URL: http://thelatinlibrary.com/suet.html (visited on 07/27/2017) (cit. on p. 61).

[69] Gaius Valerius Catullus. *Carmina*. URL: http://thelatinlibrary.com/catullus.shtml (visited on 07/27/2017) (cit. on p. 147).

[70] P. Gärdenfors. *Geometry of Meaning: Semantics Based on Conceptual Spaces*. Cambridge, Massachusetts: The MIT Press, 2014. ISBN: 9780262026789 (cit. on pp. 6, 9, 31, 49, 57).

[71] A. H. Gardiner. *Egyptian Grammar*. 3rd ed. Oxford: Griffith Institute, 1957, p. 682. ISBN: 978-0-900416-35-4 (cit. on p. 139).

[72] G. Grefenstette. *Explorations in Automatic Thesaurus Discovery*. Boston: Springer, 1994. ISBN: 978-1-4613-6167-1. DOI: 10.1007/978-1-4615-2710-7. URL: http://link.springer.com/10.1007/978-1-4615-2710-7 (cit. on p. 5).

[73] J. A. Hampton. "Disjunction of Natural Concepts." In: *Memory & Cognition* 16.6 (1988), pp. 579–591 (cit. on p. 114).

[74] J. A. Hampton. "Overextension of Conjunctive Concepts: Evidence for a Unitary Model of Concept Typicality and Class Inclusion." In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14.1 (1988), p. 12 (cit. on p. 114).

[75] R. A. Harris. *The Linguistics Wars*. Oxford: Oxford University Press, 1993, p. 368. ISBN: 978-0195098341 (cit. on p. 149).

[76] M. D. Hauser, N. Chomsky, and W. T. S. Fitch. "The Faculty of Language: What is it, who has it, and how did it Evolve?" In: *Science* 298.5598 (2002), pp. 1569–79. ISSN: 1095-9203. DOI: 10.1126/science.298.5598.1569. URL: http://www.ncbi.nlm.nih.gov/pubmed/12446899 (cit. on p. 148).

[77] C. Heunen and S. Tull. "Categories of Relations as Models of Quantum Theory." In: *Proceedings of the 12th International Workshop on Quantum Physics and Logic, Oxford, U.K., July 15-17, 2015*. Ed. by C. Heunen, P. Selinger, and J. Vicary. Vol. 195. Electronic Proceedings in Theoretical Computer Science. Open Publishing Association, 2015, pp. 247–261. DOI: 10.4204 (cit. on pp. 21, 35, 110).

[78] J. E. Hoch. *Middle Egyptian Grammar*. Mississauga: Benben Publications, 1997, p. 307. ISBN: 0920168124 (cit. on pp. 136, 137).

[79] D. Hofmann, G. J. Seal, and W. Tholen. *Monoidal Topology: A Categorical Approach to Order, Metric, and Topology*. Vol. 153. Cambridge University Press, 2014 (cit. on pp. 63, 111).

[80] B. Jacobs. "Convexity, Duality and Effects." In: *IFIP Advances in Information and Communication Technology* 323 AICT (2010), pp. 1–19. ISSN: 18684238. DOI: 10.1007/978-3-642-15240-5_1 (cit. on p. 41).

[81] B. Jacobs. "Coalgebraic Walks, in Quantum and Turing Computation." In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 6604 LNCS (2011), pp. 12–26. ISSN: 03029743. DOI: 10.1007/978-3-642-19805-2_2 (cit. on p. 31).

[82] P. T. Johnstone. *Stone Spaces*. Vol. 3. Cambridge University Press, 1986 (cit. on p. 62).

[83] P. T. Johnstone. *Sketches of an Elephant: A Topos Theory Compendium*. Vol. I. Oxford: Oxford University Press, 2002, p. 1288. ISBN: 978-0198524960 (cit. on p. 64).

[84] P. T. Johnstone. *Sketches of an Elephant: A Topos Theory Compendium*. Vol. II. Oxford: Oxford University Press, 2002, p. 1288. ISBN: 978-0198524960 (cit. on pp. 64, 127).

[85] H. Kamp and B. Partee. "Prototype Theory and Compositionality." In: *Cognition* 57.2 (1995), pp. 129–191. ISSN: 00100277. DOI: 10.1016/0010-0277(94)00659-9 (cit. on p. 53).

[86] D. Kartsaklis, M. Sadrzadeh, S. Pulman, and B. Coecke. "Reasoning about Meaning in Natural Language with Compact Closed Categories and Frobenius Algebras." In: *Logic and Algebraic Structures in Quantum Computing* (Jan. 2014). Ed. by J. Chubb, A. Eskandarian, and V. Harizanov, pp. 199–222. DOI: 10.1017/CBO9781139519687.011. arXiv: 1401.5980. URL: http://ebooks.cambridge.org/ref/id/CBO9781139519687A017%20http://arxiv.org/abs/1401.5980 (cit. on pp. 26, 34, 58).

[87] G. M. Kelly. "Basic Concepts of Enriched Category Theory." In: *Theory and Applications of Categories* 64.10 (2005), pp. 1–136. ISSN: 1201561X. URL: http://books.google.com/books?hl=en%7B%5C&%7Dlr=%7B%5C&%7Did=MPE6AAAAIAAJ%7B%5C&%7Doi=fnd%7B%5C&%7Dpg=PA21%7B%5C&%7Ddq=Basic+concepts+of+enriched+category+theory%7B%5C&%7Dots=W2eGn5HFPa%7B%5C&%7Dsig=0qqnfbK8ejF16m4gUKnQab8QipQ (cit. on p. 84).

[88] G. M. Kelly and M. L. Laplaza. "Coherence for Compact Closed Categories." In: *Journal of Pure and Applied Algebra* 19 (1980), pp. 193–213. ISSN: 00224049. DOI: 10.1016/0022-4049(80)90101-2 (cit. on p. 19).

[89] S. Lack. "Composing PROPs." In: *Theory and Applications of Categories* 13.9 (2004), pp. 147–163 (cit. on p. 65).

[90] J. Lambek. "The Mathematics of Sentence Structure." In: *The American Mathematical Monthly* 65.3 (1958), pp. 154–170. ISSN: 00029890. DOI: 10.2307/2310058. URL: http://www.jstor.org/stable/2310058 (cit. on pp. 3, 4, 13, 14, 18).

[91] J. Lambek. "Type Grammar Revisited." In: *Lacl '97 (Lncs 1582)* (1999), pp. 1–27. ISSN: 16113349. DOI: 10.1007/3-540-48975-4_1 (cit. on pp. 3, 13, 15, 18, 147).

[92] J. Lambek. *From Word to Sentence: A Computational Algebraic Approach to Grammar*. Milano: Polimetrica, 2008. ISBN: ISBN 978-88-7699-117-2. URL: http://www.math.mcgill.ca/barr/lambek/pdffiles/2008lambek.pdf (cit. on p. 17).

[93] F. W. Lawvere. "Metric Spaces, Generalized Logic, and Closed Categories." In: *Rendiconti del Seminario Matematico e Fisico di Milano* 43.1 (1973), pp. 135–166. ISSN: 03707377. DOI: 10.1007/BF02924844 (cit. on pp. 112, 113).

[94] M. Lewis and J. Lawry. "Hierarchical Conceptual Spaces for Concept Combination." In: *Artificial Intelligence* 237.April (2016), pp. 204–227. ISSN: 00043702. DOI: 10.1016/j.artint.2016.04.008. URL: http://linkinghub.elsevier.com/retrieve/pii/S0004370216300492 (cit. on p. 31).

[95] L. E. Loemaker. *G.W. Leibniz: Philosophical Papers and Letters.* Synthese H. Dordrecht: D. Reidel, 1969, p. 736 (cit. on p. 3).

[96] S. Mac Lane. "Categorical Algebra." In: *Bulletin of the American Mathematical Society* 71.1 (1965), pp. 40–106 (cit. on p. 65).

[97] S. Mac Lane. *Categories for the Working Mathematician*. Ed. by S. Axler, F. W. Gehring, and K. Ribet. Vol. 5. Graduate Texts in Mathematics. New York: Springer, 1978, p. 262. ISBN: 978-1-4419-3123-8. DOI: 10.1007/978-1-4757-4721-8. URL: http://link.springer.com/10.1007/978-1-4757-4721-8 (cit. on p. 19).

[98] S. Mac Lane and I. Moerdijk. *Sheaves in Geometry and Logic, a First Introduction to Topos Theory*. Ed. by S. Axler, F. Gehring, and K. Ribet. Vol. 13. Universitext. New York: Springer, 1992, p. 630. ISBN: 978-0-387-97710-2. DOI: 10.1007/978-1-4612-0927-0. URL: http://link.springer.com/10.1007/978-1-4612-0927-0%7B%5C%%7D5Cnpapers3://publication/doi/10.1007/978-1-4612-0927-0 (cit. on p. 64).

[99] Macrobius Ambrosius Theodosius. *Saturnaliorum Libri Septem*. URL: http://penelope.uchicago.edu/Thayer/E/Roman/Texts/Macrobius/Saturnalia/home.html (visited on 07/27/2017) (cit. on p. 49).

[100] Marcus Fabius Quintilianus. *De Institutione Oratoria*. URL: http://thelatinlibrary.com/quintilian.html (visited on 07/27/2017) (cit. on p. 13).

[101] Marcus Tullius Cicero. *Tusculanae Disputationes*. URL: http://thelatinlibrary.com/cic.html (visited on 07/27/2017) (cit. on p. 109).

[102] Marcus Vitruvius Pollio. *De Architectura*. URL: http://thelatinlibrary.com/vitruvius.html (visited on 07/27/2017) (cit. on p. 37).

[103] D. Marsden. "A Graph Theoretic Perspective on CPM(Rel)." In: *Electronic Proceedings in Theoretical Computer Science*. EPTCS 195 (Nov. 2015). Ed. by C. Heunen, P. Selinger, and J. Vicary, pp. 273–284. ISSN: 2075-2180. DOI: 10.4204/EPTCS.195.20. URL: http://dx.doi.org/10.4204/EPTCS.195.20%20http://arxiv.org/abs/1504.07003v2 (cit. on p. 123).

[104] D. Marsden and F. Genovese. "Custom Hypergraph Categories via Generalized Relations." In: *7th Conference on Algebra and Coalgebra in Computer Science (CALCO 2017)*. Ed. by F. Bonchi and B. König; Leibniz International Proceedings in Informatics, 2017, 17:1–17:15. DOI: 10.4230/LIPIcs.CALCO.2017.17. arXiv: 1703.01204. URL: http://arxiv.org/abs/1703.01204 (cit. on p. 10).

[105] J. L. McClelland and D. E. Rumelhart. *Parallel Distributed Processing: Explorations in the Microstructure Cognition*. Vol. 2. Cambridge, Massachusetts: MIT Press, 1986, p. 632. ISBN: 978-0262631105. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.164.3258%7B%5C&%7Drep=rep1%7B%5C&%7Dtype=pdf (cit. on p. 157).

[106] J. L. McClelland and D. E. Rumelhart. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Vol. 1. Cambridge, Massachusetts: MIT Press, 1986, p. 567. ISBN: 978-0262680530. URL: http://www.osti.gov/energycitations/product.biblio.jsp?osti%7B%5C_%7Did=5838709 (cit. on p. 157).

[107] R. McCrum. *Daniel Everett: 'There is no Such Thing as Universal Grammar'*. Mar. 2012. URL: https://www.theguardian.com/technology/2012/mar/25/daniel-everett-human-language-piraha (cit. on p. 151).

[108] R. Montague. "English as a Formal Language." In: *Linguaggi nella Società e nella Tecnica*. Ed. by B. Visentini. Milano: Edizioni di Communita, 1970, pp. 188–221 (cit. on p. 3).

[109] R. Montague. "Universal Grammar." In: *Theoria* 36.3 (1970), pp. 373–398. ISSN: 17552567. DOI: 10.1111/j.1755-2567.1970.tb00434.x (cit. on pp. 3, 147).

[110] R. Montague. "The Proper Treatment of Quntification in Ordinary English." In: *Formal Semantics: The Essential Readings*. Ed. by P. H. Portner and B. H. Partee. Wiley, 2008, pp. 17–34. ISBN: 9780470758335. DOI: 10.1002/9780470758335.ch1 (cit. on p. 3).

[111] M. Moortgat. "Labelled Deduction in the Composition of Form and Meaning." In: *Logic, Language and Reasoning: Essays in Honour of Dov Gabbay*. Ed. by H. J. Ohlbach and U. Reyle. Dordrecht: Springer, 1999, pp. 377–399. DOI: 10.1007/978-94-011-4574-9_15. URL: http://www.springerlink.com/index/10.1007/978-94-011-4574-9%7B%5C_%7D15 (cit. on p. 142).

[112] M. Moortgat. "Symmetric Categorial Grammar." In: *Journal of Philosophical Logic* 38.6 (2009), pp. 681–710. ISSN: 00223611. DOI: 10.1007/s10992-009-9118-6. arXiv: 1008.0170 (cit. on p. 143).

[113] M. Moortgat. "Categorial Type Logics." In: *Handbook of Logic and Language*. Ed. by A. ter Meulen and J. van Benthem. 2nd. Amsterdam: Elsevier, 2011, pp. 95–179. ISBN: 9780444537263. DOI: 10.1016/B978-0-444-53726-3.00002-5 (cit. on p. 143).

[114] K. Moroz. "Parsing Pregroup Grammars in Polynomial Time." In: *2009 International Multiconference on Computer Science and Information Technology*. IEEE, Oct. 2009, pp. 257–264. ISBN: 978-1-4244-5314-6. DOI: 10.1109/IMCSIT.2009.5352719. URL: http:

//ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5352719 (cit. on p. 18).

[115] G. V. Morrill. *Type Logical Grammar*. Dordrecht: Springer Netherlands, 1994, p. 307. ISBN: 978-0-7923-3226-8. DOI: 10.1007/978-94-011-1042-6. URL: http://link.springer.com/10.1007/978-94-011-1042-6 (cit. on p. 143).

[116] G. V. Morrill. *Categorial Grammar: Logical Syntax, Semantics, and Processing*. Oxford: Oxford University Press, 2010, p. 254. ISBN: 9780199589869 (cit. on p. 143).

[117] C. J. Mulvey. "Observations." In: *Supplemento ai Rendiconti del Circolo Matematico di Palermo* Series II (1986), pp. 99–104 (cit. on p. 61).

[118] A. Nevins, D. Pesetsky, and C. Rodrigues. "Pirahã Exceptionality: A Reassessment." In: *Language* 85.2 (2009), pp. 355–404. ISSN: 1535-0665. DOI: 10.1353/lan.0.0107. URL: http://muse.jhu.edu/content/crossref/journals/language/v085/85.2.nevins.html (cit. on p. 152).

[119] D. N. Osherson and E. E. Smith. "Gradedness and Conceptual Combination." In: *Cognition* 12.3 (1982), pp. 299–318. ISSN: 00100277. DOI: 10.1016/0010-0277(82)90037-3 (cit. on p. 114).

[120] B. H. Partee and H. L. Hendriks. "Montague Grammar." In: *Handbook of Logic and Language*. Ed. by A. ter Meulen and J. van Benthem. 2nd. Amsterdam: Elsevier, 2011, pp. 3–94. ISBN: 9780444537263. DOI: 10.1016/B978-0-444-53726-3.00001-3 (cit. on p. 147).

[121] R. Piedeleu. "Ambiguity in Categorical Models of Meaning." MSc Thesis. University of Oxford, 2014 (cit. on p. 85).

[122] R. Piedeleu, D. Kartsaklis, B. Coecke, and M. Sadrzadeh. "Open System Categorical Quantum Semantics in Natural Language Processing." In: *6th Conference on Algebra and Coalgebra in Computer Science, CALCO 2015*. Ed. by L. S. Moss and P. Sobocinski. Vol. 35. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2015, pp. 270–289. DOI: 10.4230/LIPIcs.CALCO.2015.270 (cit. on pp. 85, 123).

[123] A. Preller. "From Logical to Distributional Models." In: *Electronic Proceedings in Theoretical Computer Science* 171.Qpl 2013 (2014), pp. 113–131. ISSN: 2075-2180. DOI: 10.4204/EPTCS.171.11. URL: http://eptcs.web.cse.unsw.edu.au/paper.cgi?QPLX.11 (cit. on pp. 4, 6, 7).

[124] A. Preller and J. Lambek. "Free Compact 2-Categories." In: *Mathematical Structures in Computer Science* 17.02 (Apr. 2007), p. 309. ISSN: 0960-1295. DOI: 10.1017/S0960129506005901 (cit. on p. 7).

[125] Publio Vergilius Maro. *Aeneis*. URL: http://www.thelatinlibrary.com/verg.html (visited on 07/27/2017) (cit. on p. v).

[126] Publius Cornelius Tacitus. *Annales*. URL: http://thelatinlibrary.com/tac.html (visited on 07/27/2017) (cit. on p. 135).

[127] G. K. Pullum and B. C. Scholz. "Empirical Assessment of Stimulus Poverty Arguments." In: *The Linguistic Review* 18.1-2 (2002), pp. 9–50. ISSN: 0167-6318. DOI: 10.1515/tlir.19.1-2.9. arXiv: arXiv:1011.1669v3 (cit. on p. 148).

[128] Redação. *"Ele Virou um Charlatão", diz Chomsky*. S. Paulo, Feb. 2009. URL: http://www1.folha.uol.com.br/fsp/ciencia/fe0102200904.htm (cit. on p. 151).

[129] J. T. Rickard, J. Aisbett, and G. Gibbon. "Reformulation of the Theory of Conceptual Spaces." In: *Information Sciences* 177.21 (2007), pp. 4539–4565. ISSN: 00200255. DOI: 10.1016/j.ins.2007.05.023 (cit. on p. 31).

[130] E. Rosch and C. B. Mervis. "Family Resemblances: Studies in the Internal Structure of Categories." In: *Cognitive psychology* 7.4 (1975), pp. 573–605 (cit. on p. 114).

[131] R. Rosebrugh and R. J. Wood. "Distributive Laws and Factorization." In: *Journal of Pure and Applied Algebra* 175.1 (2002), pp. 327–353 (cit. on p. 109).

[132] S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. 3rd. Essex, UK: Parentice Hall, 2009, p. 1152. ISBN: 0136042597. DOI: 10.1017/S0269888900007724. arXiv: 9809069v1 [arXiv:gr-qc]. URL: http://portal.acm.org/citation.cfm?id=1671238%7B%5C&%7Dcoll=DL%7B%5C&%7Ddl=GUIDE%7B%5C&%7DCFID=190864501%7B%5C&%7DCFTOKEN=29051579%7B%5C%%7D5Cnpapers2://publication/uuid/4B787E16-89F6-4FF7-A5E5-E59F3CFEFE88 (cit. on p. 157).

[133] M. Sadrzadeh, S. Clark, and B. Coecke. "The Frobenius Anatomy of Word Meanings I: Subject and Object Relative Pronouns." In: *Journal of Logic and Computation* 23.6 (2013), pp. 1293–1317 (cit. on p. 27).

[134] M. Sadrzadeh, S. Clark, and B. Coecke. "The Frobenius Anatomy of Word Meanings II: Possessive Relative Pronouns." In: *Journal of Logic and Computation* 26.2 (2014), pp. 785–815 (cit. on p. 27).

[135] P. Selinger. "Dagger Compact Closed Categories and Completely Positive Maps." In: *Electronic Notes in Theoretical Computer Science* 170.1 (Mar. 2007), pp. 139–163. ISSN: 15710661. DOI: 10.1016/j.entcs.2006.12.018. URL: http://linkinghub.

elsevier.com/retrieve/pii/S1571066107000606 (cit. on pp. 6, 123).

[136] P. Selinger. "A Survey of Graphical Languages for Monoidal Categories." In: *New structures for physics*. Vol. 813. Springer, 2010, pp. 289–355. ISBN: 9783642128202. DOI: 10.1007/978-3-642-12821-9_4. arXiv: 0908.3347 (cit. on p. 20).

[137] R. N. Shepard et al. "Toward a Universal Law of Generalization for Psychological Science." In: *Science* 237.4820 (1987), pp. 1317–1323 (cit. on p. 115).

[138] P. Sobociński. *Graphical Linear Algebra*. 2017. URL: https://graphicallinearalgebra.net/ (visited on 07/16/2017) (cit. on p. 5).

[139] R. Street. "The Formal Theory of Monads." In: *Journal of Pure and Applied Algebra* 2.2 (1972), pp. 149–168 (cit. on p. 111).

[140] Titus Lucretius Carus. *De Rerum Natura Libri Sex*. URL: http://thelatinlibrary.com/lucretius.html (visited on 07/27/2017) (cit. on pp. 3, 31).

[141] A. Tversky. "Features of Similarity." In: *Psychological review* 84.4 (1977), p. 327 (cit. on p. 115).

[142] D. Villeneuve. *Arrival*. 2016. URL: http://www.imdb.com/title/tt2543164/ (cit. on p. 149).

[143] L. Wittgenstein. *Philosophical Investigations*. Ed. by G. Anscombe, P. Hacker, and J. Schulte. 4th. Chicester, UK: Wiley-Blackwell, 2009, p. 592. ISBN: 1405159294 (cit. on p. 4).

[144] F. Zanasi. "Interacting Hopf Algebras: The Theory of Linear Systems." PhD thesis. École Normale Supérieure de Lyon, 2015 (cit. on p. 65).

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

https://bitbucket.org/amiede/classicthesis/

The LaTeX macros used are publicly available on *git*:

https://bitbucket.org/Nigredo/quantum-latex-macros

Stefano Gogioso and William Zeng have to be credited for having provided the first draft on which these macros were created.

Final Version as of June 27, 2018 (`classicthesis` version 4.2).