

VINCENT WANG-MAŚCIANICA

MY THESIS

Contents

1	Introduction	5			
1.0.1	Relation to conventional grammars	8			
1.1	What this thesis is not about	13			
2	String Diagrams	15			
2.1	A Partial History of String Diagrams	15			
2.1.1	Pre-formal Diagrams	15			
2.1.2	Convergent Evolution	16			
2.2	Process Theories	16			
2.2.1	What does it mean to copy and delete?	20			
2.2.2	What is an update?	21			
2.2.3	Spatial predicates	22			
2.2.4	Some basic properties of linguistic spatial relations	23			
2.3	Defining String Diagrams	23			
2.3.1	Symmetric Monoidal Categories	23			
2.3.2	1-object 4-categories	23			
2.3.3	PROPs	23			
3	Text Circuit Theory	25			
3.1	An everyday miracle of language	25			
3.1.1	Speaking grammars, listening grammars	25			
4	The category of continuous relations	29			
4.1	Continuous Relations	29			
4.2	Continuous Relations by examples	29			
4.3	The category TopRel	35			
4.4	Biproducts	35			
4.5	Symmetric Monoidal Closed structure	36			
4.6	TopRel diagrammatically	37			
4.6.1	Relations that are always continuous	37			
4.7	Populating space with shapes using sticky spiders	40			
4.7.1	When does an object have a spider (or something close to one)?	40			
4.8	Sticky spiders are collections of shapes in space	60			
4.9	Displacing shapes with symmetries	60			
4.10	Moving shapes continuously	60			
4.11	Configuration space of a sticky spider	60			
4.12	Topological models for a selection of concepts	60			
4.13	Modelling Lassos	60			
5	Things to do with formal depictions of language	71			
5.1	Modelling metaphor	71			
5.1.1	Orders, Temperature, Colour, Mood	72			
5.1.2	Complex conceptual structure	72			
5.1.3		73			

1
Introduction

THIS THESIS IS ABOUT A NEW WAY TO SEE LANGUAGE USING FORMAL DIAGRAMS.

THE FORMAL DIAGRAMS ARE COMPOSITIONAL BLUEPRINTS.

We know how grammar composes meanings in language.

We can quotient out grammar using these formal diagrams.

This turns big black boxes into composites of small black boxes.

This leaves smaller pieces for machines to learn representations for.

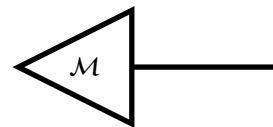
The compositional blueprints may be instantiated by classical or quantum computers.

We introduce the history, theory, and use of these diagrams in Chapter ??.

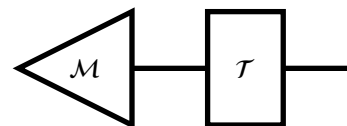
THE DIAGRAMS LOOK LIKE CIRCUITS.

Let's say that the meaning of text is how it updates a model.

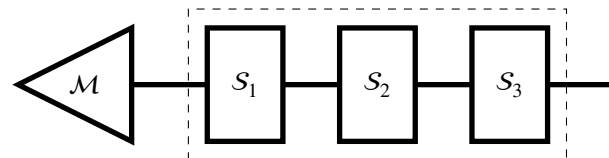
So we start with some model of the way things are.



Text updates that model.



Text is made of sentences.

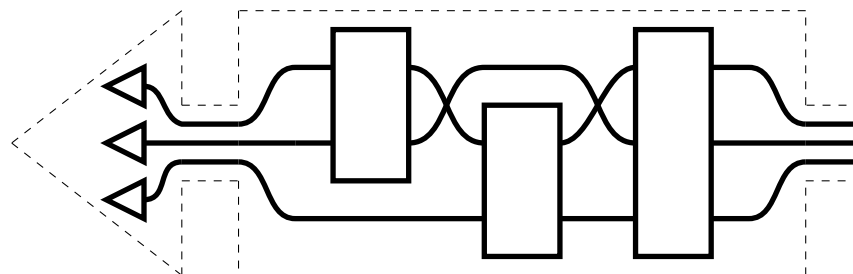


Let's say that *The meaning of a sentence is how it updates the meanings of its parts.*

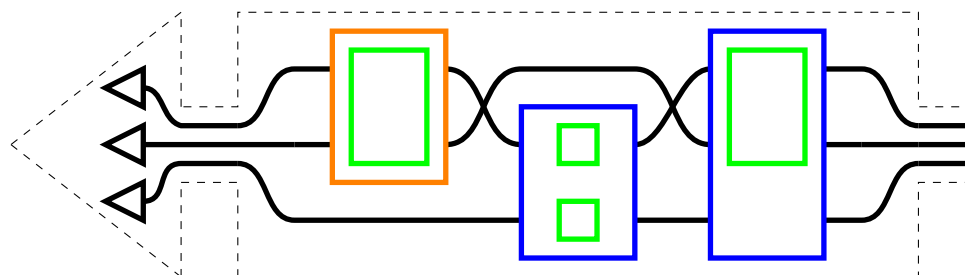
Let's say that the *parts* of a sentence are the nouns it contains or refers to.

Noun data is carried by wires.

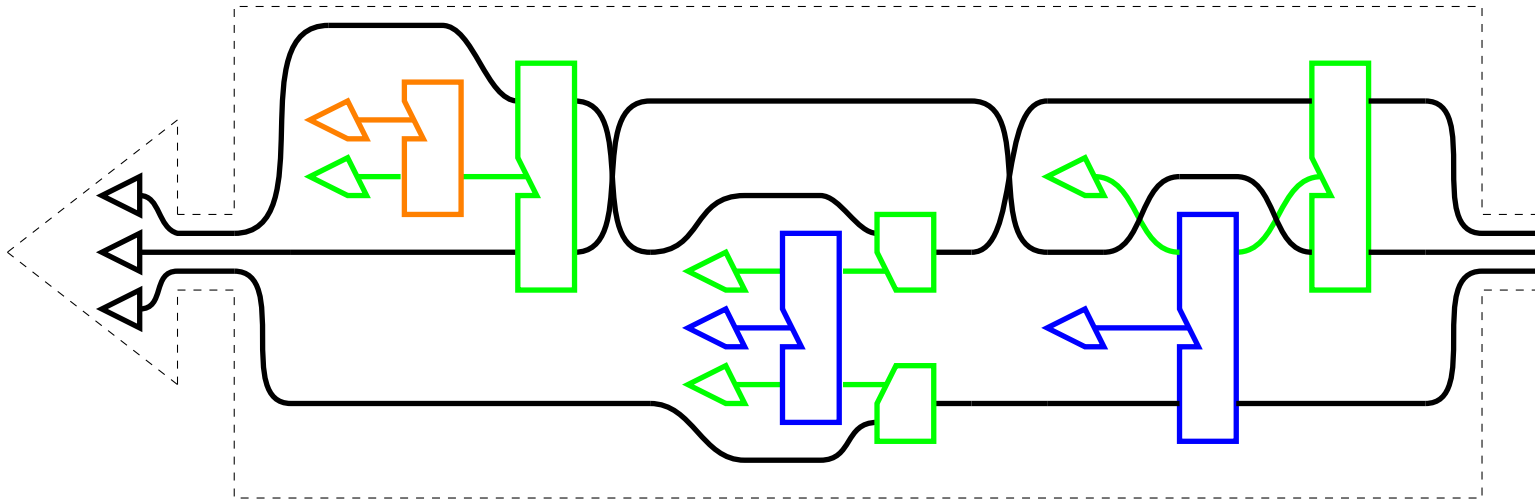
Collections of nouns are related by gates.



Gates can be related by higher order gates.



Higher order gates are gates that modify the parameters of other gates.



1.0.1 Relation to conventional grammars

DISCOURSE REPRESENTATION THEORY

Text is composed of sentences as circuits are composed from gates.

Text circuits are composed by connecting noun wires.

This can happen when two sentences refer to the same noun.

Alice likes Bob. Bob hates Claire.

placeholder

Or when a pronoun is used to refer to another noun. Alice likes the flowers that Bob gives Claire.

placeholder

So if you know how to resolve pronoun references, and you know how to represent sentences as gates acting on noun wires, you may represent text as circuits.

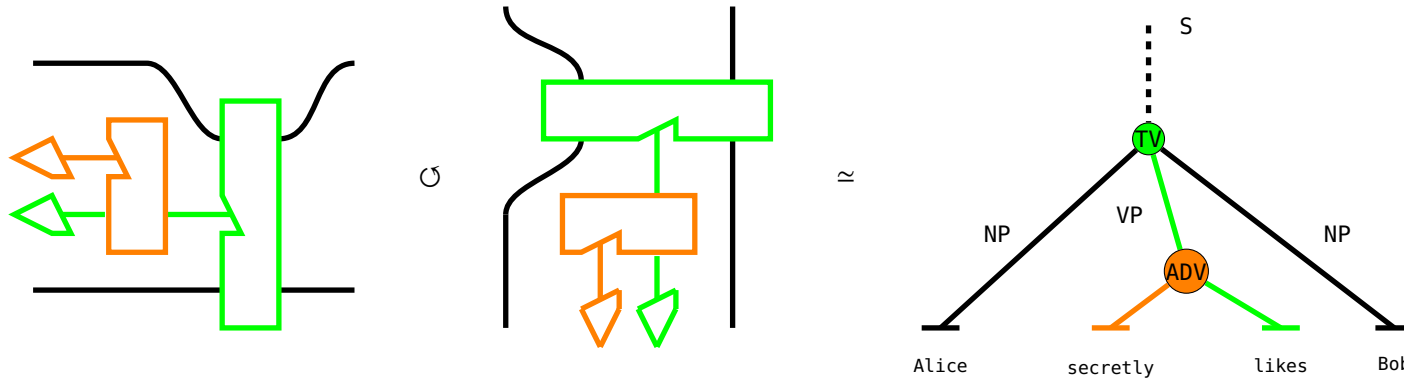
There are several ways to obtain gates from grammars for sentences.

CONTEXT-FREE GRAMMAR

In a context-free grammar, a sentence symbol S is expanded by replacing individual symbols with strings of symbols to obtain a well-formed sentence.

The shape of these expansions is a tree.

That tree gives the shape of gates in a sentence.



In Chapter ??, we characterise the generative grammar of text circuits in terms of a context-free grammar with additional structure.

TYPELOGICAL GRAMMAR

A typological grammar is like the inverse of a context-free grammar.

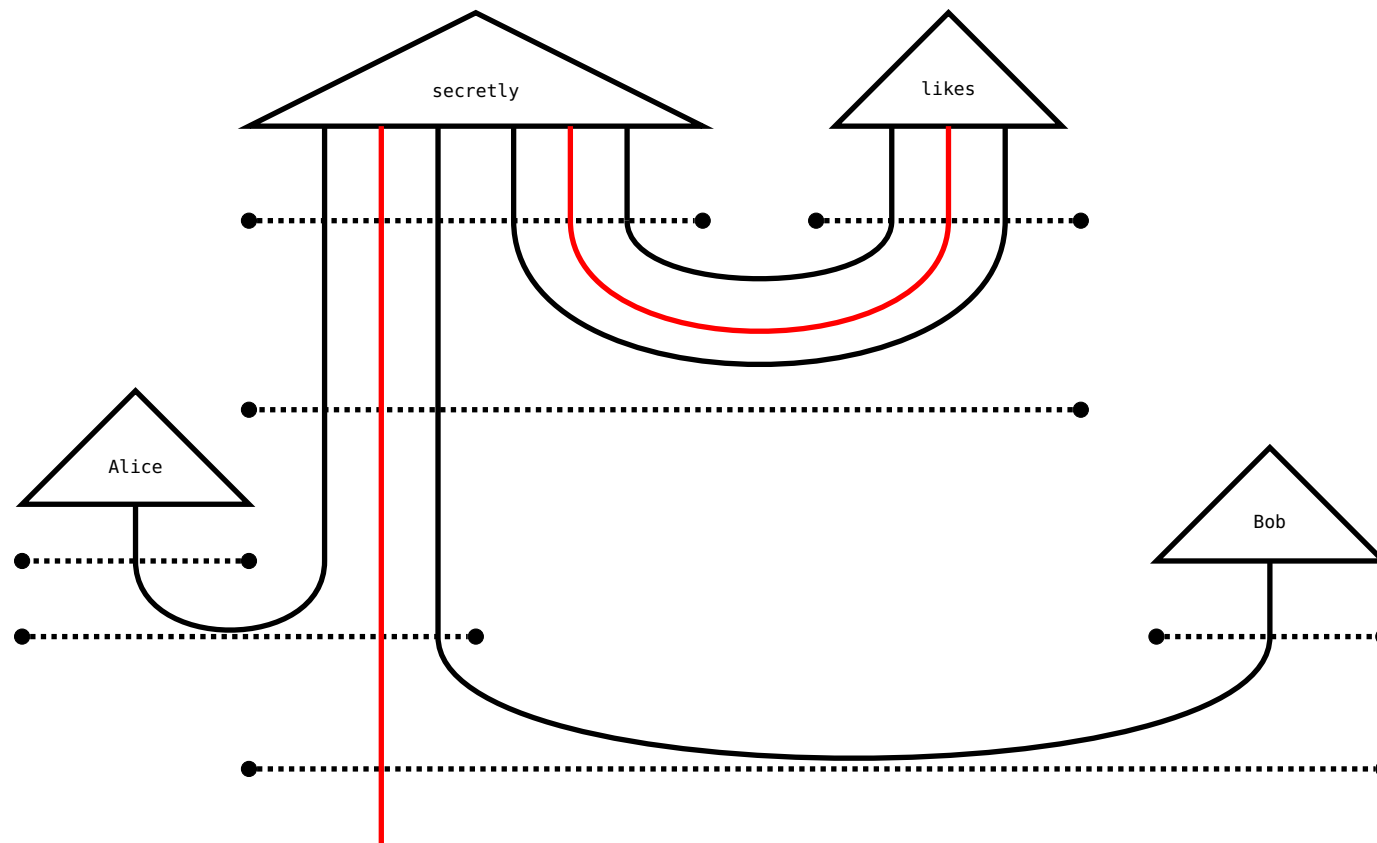
The latter is the grammar of the speaker: a full sentence is produced from S .

The former is the grammar of the listener: the sentence type s must be derived from the words of the sentence.

In a typological grammar such as a pregroup grammar, words are assigned types, and a logical proof witnesses the sentence type s :

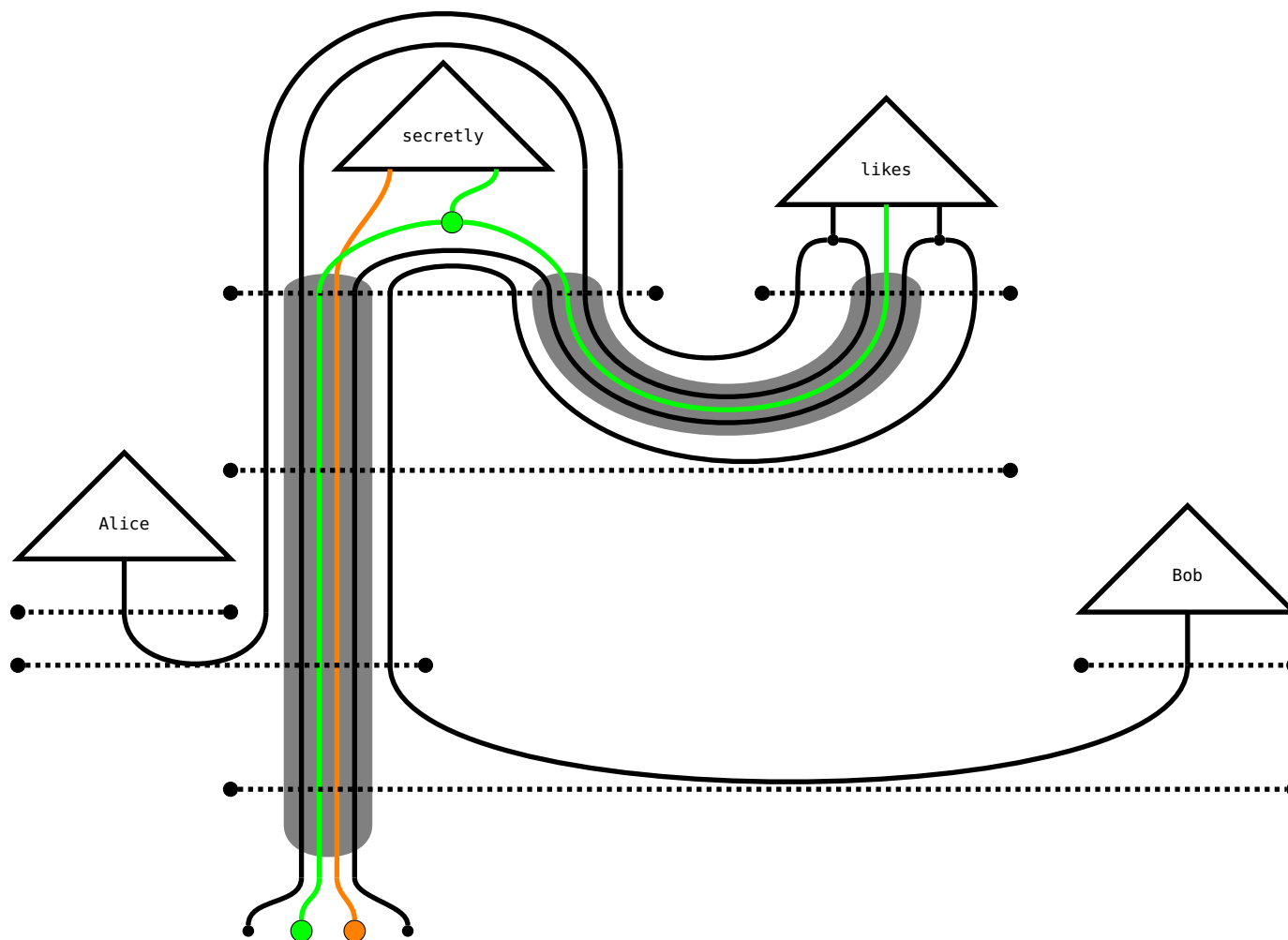
$$\begin{array}{c}
 \text{secretly} : (-^1 n \cdot s \cdot n^{-1}) \cdot (-^1 n \cdot s \cdot n^{-1})^{-1} \quad \text{likes} : ^{-1} n \cdot s \cdot n^{-1} \\
 \hline
 \text{Alice} : n \quad \text{secretly_likes} : ^{-1} n \cdot s \cdot n^{-1} \\
 \hline
 \text{Alice_secretly_likes} : s \cdot n^{-1} \quad [x^{-1} \cdot x \rightarrow 1] \quad \text{Bob} : n \\
 \hline
 \text{Alice_secretly_likes_Bob} : s \quad [x \cdot ^{-1} x \rightarrow 1]
 \end{array}$$

Such proofs of syntactic correctness can be re-expressed as diagrams:

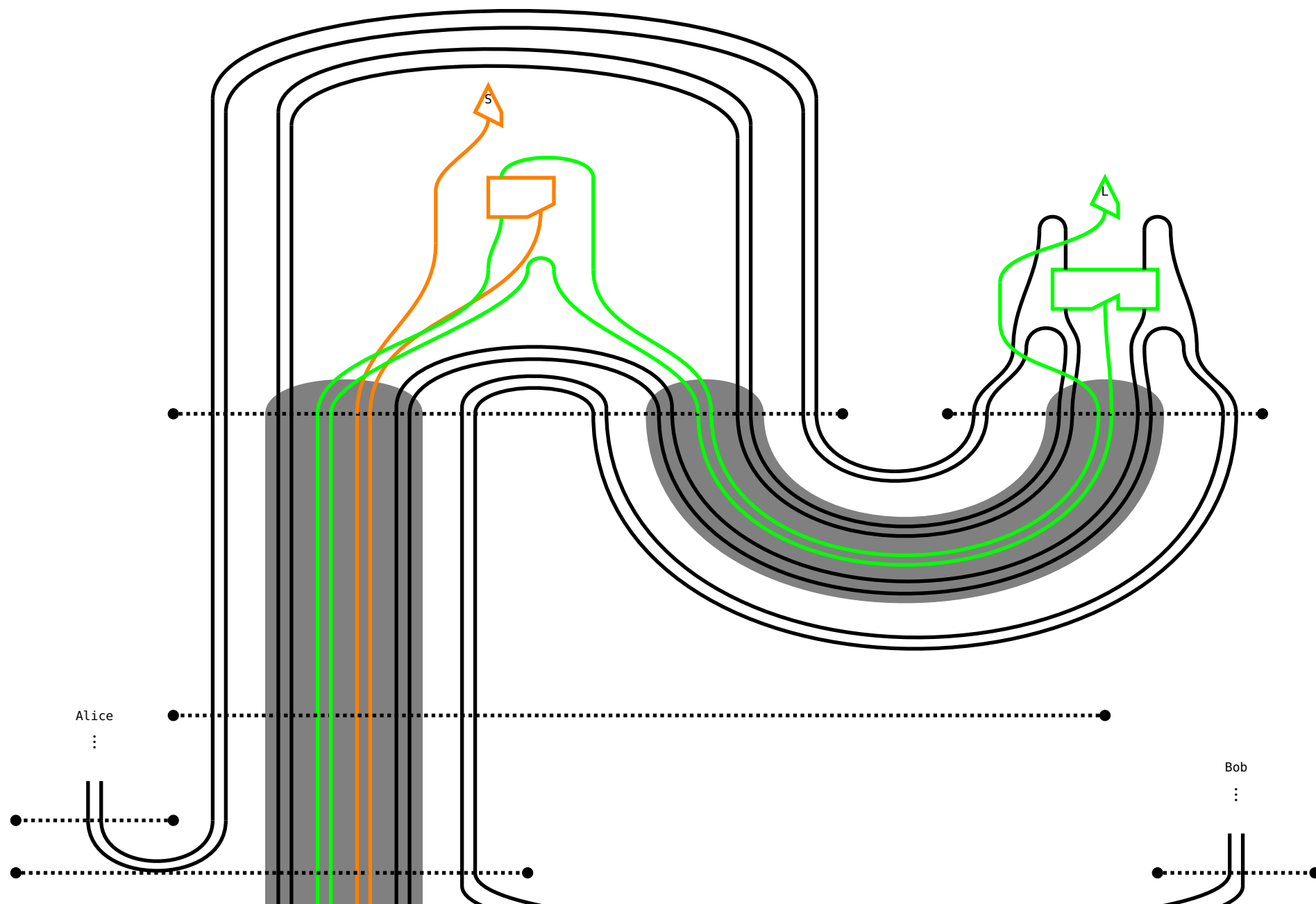


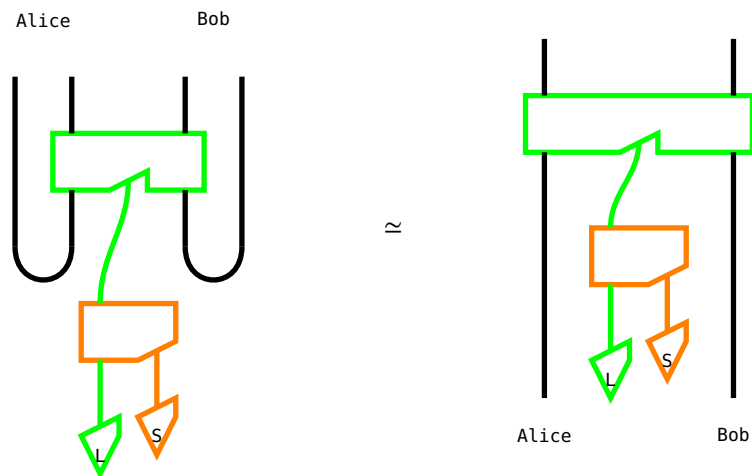
To obtain a circuit, we make the sentence type s dependent on the words that occur in the sentence, and we refine the structure on each word by introducing "internal wirings",

which make use of multiwires called spiders.



A certain choice of spider make these internal wiring diagrams topologically equivalent to the circuits we want.





In Chapter ??, we show how certain pregroup and dependency grammars may be equipped with internal wirings to obtain text circuits.

1.1 What this thesis is not about

ALTHOUGH THIS THESIS IS ABOUT LANGUAGE, IT IS NOT REALLY LINGUISTICS.

Linguistics, conventionally, is about understanding human language.

Formal linguistic theories rely on empirical observation of human language use.

Now there are not-human entities that use human language.

Now there are un-natural languages that humans use.

So I am interested in sketching a form of language that humans and computers can use.

This sketch is informed by, but not bound to, human language.

ALTHOUGH THIS THESIS IS MATHEMATICAL, IT IS NOT REALLY MATHEMATICS.

This thesis does not introduce significantly new mathematical constructions or relations.

So there are no new "know that"s.

This thesis does use relatively modern mathematics to approach an old problem.

The math is symmetric monoidal categories, and some weak n-categories.

The problem is depicting language.

So there is new "know how".

THIS THESIS IS COMPUTER SCIENCE, A LITTLE.

The mathematics used is, to a degree, implementation agnostic.

This thesis is only concerned with the "in principle", rather than the "in practice".

There will be no code demonstration nor machine learning experiments.

Computer Science : Programming :: Physics : Engineering

I will point out where experiments have been done.

Hypothetical procedures will be spelled out if needed.

The subject matter of this thesis touches on enough fields that I feel it necessary to get everyone on the same page.

The rest of this introduction will introduce the technical concepts and philosophical motivations that are the foundation for the rest of this work.

2

String Diagrams

2.1 A Partial History of String Diagrams

WE HAVE ALWAYS TRIED TO CAPTURE LANGUAGE IN DIAGRAMS.

Every so often, someone tries once again to draw the shape of language.

...

Now it's our turn.

The difference this time is formality; we are using *string diagrams*.

STRING DIAGRAMS ARE FORMAL REASONING AND REPRESENTATION SYSTEMS FOR MONOIDAL CATEGORIES. Chronologically, those adjectives occurred in this order: 1) visual representation, 2) visual reasoning, 3) formal.

2.1.1 Pre-formal Diagrams

Pre-formal diagrams are about visual expression. By the restrictions of writing technologies, the kind of visual representations we are interested in are two-dimensional. So visual means geometric in the plane. The ancient ancestors of modern string diagrams are systems for visual representation alone. For these ancestral diagrams, reasoning, if there is any, takes place 'elsewhere', not within and between diagrams.

Euclid: Geometry

Venn & Carroll: Syllogisms

Petri: Chemistry

2.1.2 *Convergent Evolution*

String diagrams arise naturally as diagrammatic representations for formal theories where reoccurring variables or bookkeeping indices for keep track of pairwise connections become too unwieldy for one-dimensional syntax.

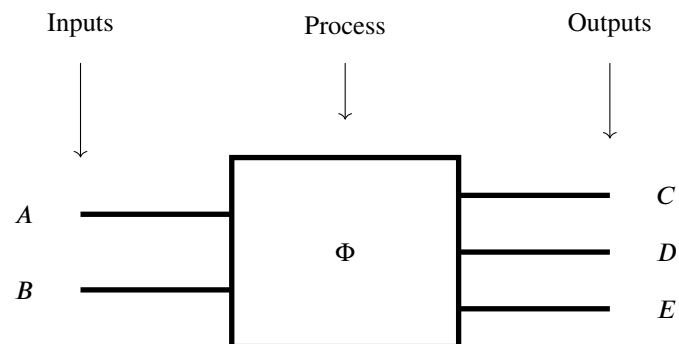
IN PHYSICS

IN LOGIC

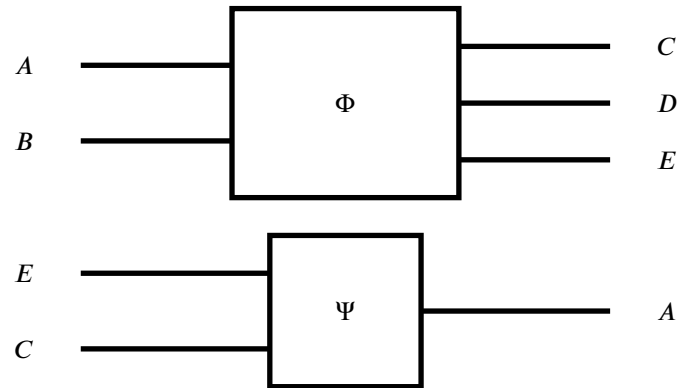
IN COMPUTER SCIENCE

2.2 *Process Theories*

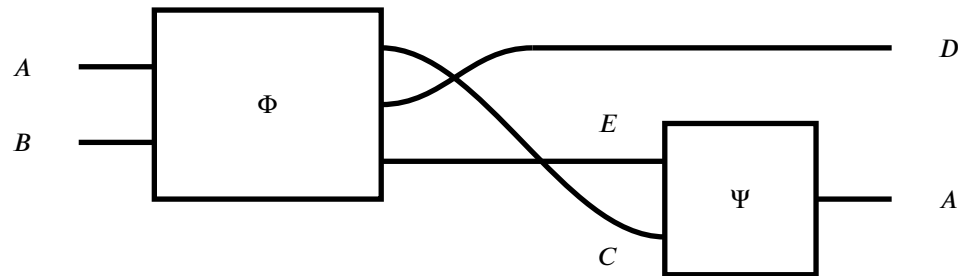
A *process* is something that transforms some number of input system types to some number of output system types. We depict systems as wires, labelled with their type, and processes as boxes. We read processes from left to right.



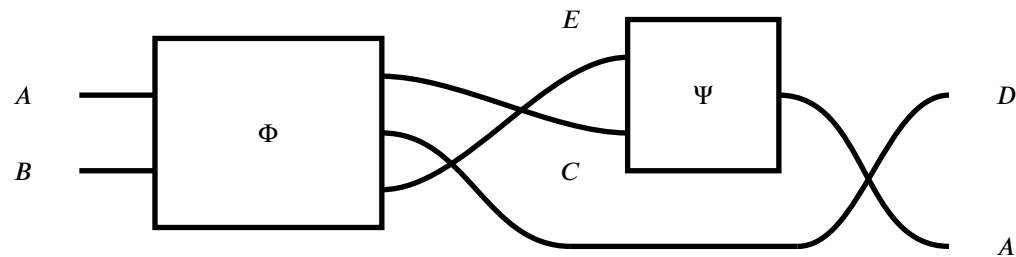
Processes may compose in parallel, which we depict as vertically stacking boxes.



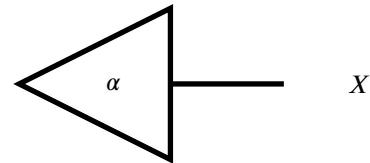
Processes may compose sequentially, which we depict as connecting wires of the same type from left to right.



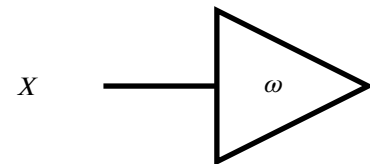
In these diagrams only input-output connectivity matters: so we may twist wires and slide boxes along wires to obtain different diagrams that still refer to the same process. So the diagram below is equal to the diagram above.



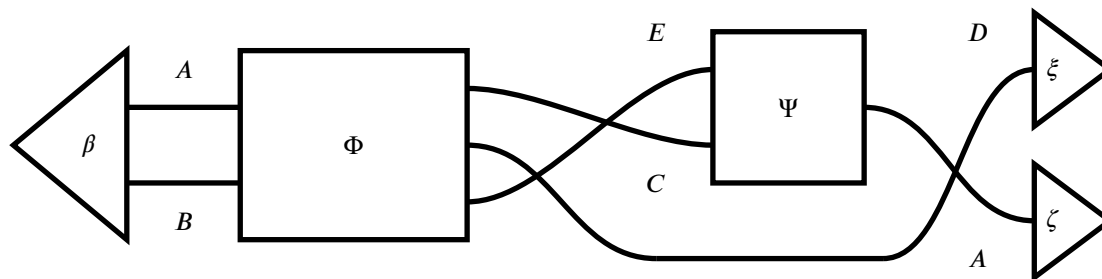
Some processes have no inputs; we call these *states*.



Some processes have no outputs; we call these *tests*.



A process with no inputs and no outputs is a *number*; the number tells us the outcome of applying tests to a composite of states modified by processes.



A process theory is given by the following data¹:

- A collection of systems
- A collection of processes along with their input and output systems
- A methodology to compose systems and processes sequentially and in parallel, and a specification of the unit of parallel composition.
- A collection of equations between composite processes

¹ Formally, process theories are symmetric monoidal categories []; see section ??.

Example 2.2.1 (Linear maps with direct sum). Systems are finite-dimensional vector spaces over R . Processes are linear maps, expressed as matrices with entries in R .

Sequential composition is matrix multiplication. Parallel composition of systems is the direct sum of vector spaces \oplus . The parallel composition of matrices \mathbf{A}, \mathbf{B} is the block-diagonal matrix

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix}$$

The unit of parallel composition is the singleton 0-dimensional vector space. States are row vectors. Tests are column vectors. The numbers are R .²

Example 2.2.2 (Sets and functions with cartesian product). Systems are sets A, B . Processes are functions between sets $f : A \rightarrow B$. Sequential composition is function composition. Parallel composition of systems is the cartesian product of sets: the set of ordered pairs of two sets.

$$A \otimes B = A \times B := \{(a, b) \mid a \in A, b \in B\}$$

The parallel composition $f \otimes g : A \times C \rightarrow B \times D$ of functions $f : A \rightarrow B$ and $g : C \rightarrow D$ is defined:

$$f \otimes g := (a, c) \mapsto (f(a), g(c))$$

The unit of parallel composition is the³ singleton set $\{\star\}$. States of a set A correspond to elements $a \in A$.⁴ Every system A has only one test $a \mapsto \star$.⁵ There is only one number.

Example 2.2.3 (Sets and relations with cartesian product). Systems are sets A, B . Processes are relations between sets $\Phi \subseteq A \times B$, which we may write in either direction $\Phi^* : A \rightharpoonup B$ or $\Phi_* : B \rightharpoonup A$.⁶ Sequential composition is relation composition:

$$A \xrightarrow{\Phi} B \xrightarrow{\Psi} C := \{(a, c) \mid a \in A, c \in C, \exists b \in B : (a, b) \in \Phi \wedge (b, c) \in \Psi\}$$

Parallel composition of systems is the cartesian product of sets. The parallel composition of relations $A \otimes C \xrightarrow{\Phi \otimes \Psi} B \otimes D$ of relations $A \xrightarrow{\Phi} B$ and $C \xrightarrow{\Psi} D$ is defined:

$$\Phi \otimes \Psi := \{((a, c), (b, d)) \mid (a, b) \in \Phi \wedge (c, d) \in \Psi\}$$

The unit of parallel composition is the singleton. States of a set A are subsets of A . Tests of a set A are also subsets of A .

² Usually the monoidal product is written with the symbol \otimes , which denotes hadamard product for linear maps. The process theory we have just described takes the direct sum \oplus to be the monoidal product. To avoid confusion we will use the linear algebraic notation when linear algebra is concerned.

³ There are many singletons, but this presents no problem for the later formal definition because they are all equivalent up to unique isomorphism.

⁴ We forgo the usual categorical definition of points from the terminal object in favour of generalised points from the monoidal perspective.

⁵ This is since the singleton is terminal in **Set**.

⁶ Relations between sets are equivalently matrices with entries from the boolean semiring. Relation composition is matrix multiplication with the boolean semiring. Φ^*, Φ_* are the transposes of one another.

2.2.1 What does it mean to copy and delete?

Now we discuss how we might define the properties and behaviour of processes by positing equations between diagrams. Let's begin simply with two intuitive processes *copy* and *delete*:



Example 2.2.4 (Linear maps). Consider a vector space \mathbf{V} , which we assume includes a choice of basis. The copy map is the rectangular matrix made of two identity matrices:

$$\Delta_{\mathbf{V}} : \mathbf{V} \rightarrow \mathbf{V} \oplus \mathbf{V} := \begin{bmatrix} \mathbf{1}_{\mathbf{V}} & \mathbf{1}_{\mathbf{V}} \end{bmatrix}$$

The delete map is the column vector of 1s:

$$\epsilon_{\mathbf{V}} : \mathbf{V} \rightarrow \mathbf{0} := \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

Example 2.2.5 (Sets and functions). Consider a set A . The copy function is defined:

$$\Delta_A : A \rightarrow A \times A := a \mapsto (a, a)$$

The delete function is defined:

$$\epsilon_A : A \rightarrow \{\star\} := a \mapsto \star$$

Example 2.2.6 (Sets and relations). Consider a set A . The copy relation is defined:

$$\Delta_A : A \rightharpoonup A \times A := \{(a, (a, a)) \mid a \in A\}$$

The delete relation is defined:

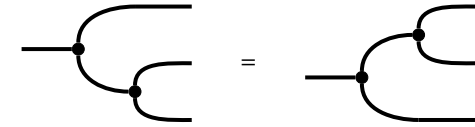
$$\epsilon_A : A \rightharpoonup \{\star\} := \{(a, \star) \mid a \in A\}$$

We may verify that, no matter the concrete interpretation of the diagram in terms of linear maps, functions or relations copy and delete satisfy the equations in the margin:

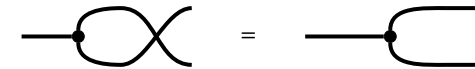
It is worth pausing here to think about how one might characterise the process of copying in words; it is challenging to do so for such an intuitive process. The diagrammatic equations in the margin, when translated into prose, provide an answer.

Coassociativity: says there is no difference between copying copies.

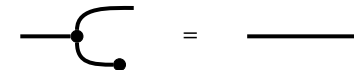
Formally, the following equations characterise a cocommutative comonoid internal to a monoidal category.



coassociativity



cocommutativity



counitality

(2.1)

Cocommutativity: says there is no difference between the outputs of a copy process.

Counitality: says that if a copy is made and one of the copies is deleted, the remaining copy is the same as the original.

Insofar as we think this is an acceptable characterisation of copying, rather than specify concretely what a copy and delete does for each system X we encounter, we can instead posit that so long as we have processes $\Delta_X : X \otimes X \rightarrow X$ and $\epsilon_X : X \rightarrow I$ that obey all the equational constraints above, Δ_X and ϵ_X are as good as a copy and delete.

Here is an unexpected consequence. Suppose we insist that *to copy* in principle also implies the ability to copy *anything* – arbitrary states. From Example 2.2.7 and Remark 2.2.8, we know that this demand is incompatible with certain process theories. In particular, this demand would constrain a process theory of sets and relations to a sub-theory of sets and functions. The moral here is that process theories are flexible enough to meet ontological needs. A classical computer scientist who works with perfectly copyable data and processes might demand universal copying along with Equations 2.1, whereas a quantum physicist who wishes to distinguish between copyable classical data and non-copyable quantum data might taxonomise copy and delete as a special case of a more generic quasi-copy and quasi-delete that only satisfies equations 2.1.⁷

2.2.2 What is an update?

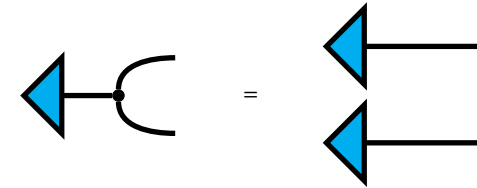
In the previous section we have seen how we can start with concrete examples of copying in distinct process theories, and obtain a generic characterisation of copying by finding diagrammatic equations copying satisfies in each concrete case. In this section, we show how to go in the opposite direction: we start by positing diagrammatic equations that characterise the operational behaviour of a particular process – such as *updating* – and it will turn out that any concrete process that satisfies the equational constraints we set out will *by our own definition* be an update.

Perhaps the most familiar setting for an update is a database. In a database, an **entry** often takes the form of pairs of **fields** and **values**. For example, where a database contains information about employees, a typical entry might look like:

< NAME:Jono Doe, AGE:69, JOB:CONTENT CREATOR, SALARY:\$420, ... >

There are all kinds of reasons one might wish to update the value of a field: Jono might legally change their name, a year might pass and Jono’s age must be incremented, Jono might be promoted or demoted or get a raise and so on. It was the concern of database theorists to formalise and axiomatise the notion of updating the value of a field – *independently of the specific programming language implementation of a database* – so that they had reasoning tools to ensure program correctness []. The problem is reducible to axiomatising a *rewrite*: we can think of updating

Example 2.2.7 (Not all states are copyable). Call a state *copyable* when it satisfies the following diagrammatic equation:

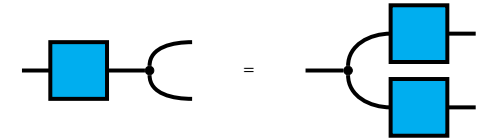


In the process theory of sets and functions, all states are copyable. Not all states are copyable in the process theories of sets and relations and of linear maps. For example, consider the two element set $B := \{0, 1\}$, and let $T : \{\star\} \rightarrow B := \{(\star, 0), (\star, 1)\} \simeq \{0, 1\}$. Consider the composite of T with the copy relation:

$$T; \Delta_B := \{(\star, (0, 0)), (\star, (1, 1))\} \simeq \{(0, 0), (1, 1)\}$$

This is a perfectly correlated bipartite state, and it is not equal to $\{0, 1\} \times \{0, 1\}$, so T is not copyable.

Remark 2.2.8. The copyability of states is a special case of a more general form of interaction with the copy relation:



A cyan map that satisfies this equation is said to be a homomorphism with respect to the commutative comonoid. In the process theory of relations, those relations that satisfy this equation are precisely the functions; in other words, this diagrammatic equation expresses *determinism*.

⁷ Quantum physicists *do* do this; see Dodo: []

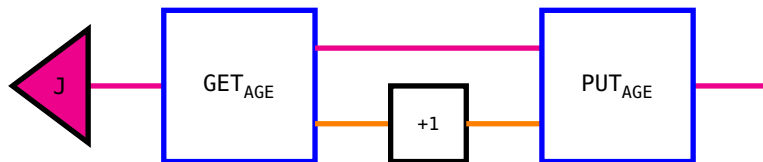
a value as first calculating the new value, then *putting* the new value in place of the old. Since often the new value depends in some way on the old value, we also need a procedure to *get* the current value.

That was a flash-prehistory of *bidirectional transformations* []. Following the monoidal generalisation of lenses in [], a rewrite as we have described above is specified by system diagrammatic equations in the margin, each of which we introduce in prose.

These diagrammatic equations do two things. First, they completely specify what it means to get and put values in a field in an implementation independent manner; it doesn't matter whether database entries are encoded as bitstrings, qubits, slips or paper or anything else, what matters is the interaction of get and put. Second, the diagrammatic equations give us the right to call our processes *get* and *put* in the first place: we define what it means to get and put by outlining the mutual interactions of get, put, copy, and delete. These two points are worth condensing and rephrasing:

A kind of process is determined by patterns of interaction with other kinds of processes.

Now we can diagrammatically depict the process of updating Jono's age, by *getting* Jono's *age value* from their *entry*, incrementing it by 1, and *putting* it back in.



2.2.3 Spatial predicates

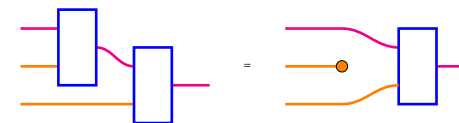
The following simple inference is what we will try to capture process-theoretically:

- Oxford is north of London
- Vincent is in Oxford
- Rocco is in London

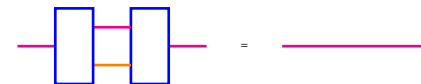
How might it follow that Rocco is south of Vincent?

One way we might approach such a problem computationally is to assign a global coordinate system, for instance interpreting 'north' and 'south' and 'is in' using longitude and latitude. Another coordinate system we might use is a locally flat map of England. The fact that either coordinate system would work is a sign that there is a degree of implementation-independence.

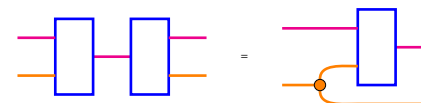
PutPut: Putting in one value and then a second is the same as deleting the first value and just putting in the second.



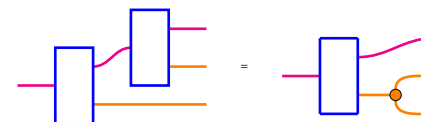
GetPut: Getting a value from a field and putting it back in is the same as not doing anything.



PutGet: Putting in a value and getting a value from the field is the same as first copying the value, putting in one copy and keeping the second.



GetGet: Getting a value from a field twice is the same as getting the value once and copying it.



This coordinate/implementation-independence is true of most spatial language: we specify locations only by relative positions to other landmarks or things in space, rather than by means of a coordinate system. This is necessarily so for the communication of spatial information between agents who may have very different reference frames.

So the process-theoretic modelling aim is to specify how relations between entities can be *updated* and *classified* without requiring individual spatial entities to intrinsically possess meaningful or determinate spatial location.

So far we have established how to update properties of individual entities. We can build on what we have so far by observing that a relation between two entities can be considered a property of the pair.

placeholder

Spatial relations obey certain compositional constraints, such as transitivity in the case of ‘north of’:

placeholder

Or the equivalence between ‘north of’ and ‘south of’ up to swapping the order of arguments:

There are other general constraints on spatial relations, such as order-independence: the order in which spatial relations are updated does not (at least for perfect reasoners) affect the resultant presentation. This is depicted diagrammatically as commuting gates:

placeholder

2.2.4 *Some basic properties of linguistic spatial relations*

2.3 *Defining String Diagrams*

2.3.1 *Symmetric Monoidal Categories*

2.3.2 *1-object 4-categories*

2.3.3 *PROPs*

3

Text Circuit Theory

3.1 An everyday miracle of language

SPEAKERS AND LISTENERS UNDERSTAND ONE ANOTHER. Even if we cautiously hedge this mundane statement to exclude pragmatics and context and only encompass small and boring fragments of factual language, as far as formal theories of syntax and semantics go, this is still a miracle. In this section I will first explain why this simple fact represents a gap in our understanding of language. Hopefully the urgency of filling the gap will be impressed, and then I will provide a mathematical structure that bridges this gap.

3.1.1 Speaking grammars, listening grammars

Here are some naïve observations on the nature of speaking and listening. Let's suppose that a speaker, Preube, wants to communicate a thought to Fondo. Just as a running example that does not affect the point, let's say we can gloss the thought as carrying the relations:

(alice, bob, flowers, like, give)

Preube and Fondo cooperate to achieve a minor miracle; Preube encodes his thoughts – a structure that doesn't look like a one-dimensional string of symbols – into a one-dimensional string of symbols, and Fondo does the reverse, turning a one-dimensional string of symbols into *a thought-structure like that of Preube's*. The two can do this for infinitely many thoughts, and new thoughts neither have encountered before.

placeholder

We assume Preube and Fondo speak the same language, so both know how words in their language correspond to

What I mean by this is just that both Preube and Fondo agree on the structure of entities and relations up to the words for those entities and relations. For example, Preube could ask Fondo comprehension questions such as WHO GAVE WHAT? TO WHO?, and if Fondo can always correctly answer – e.g. BOB GAVE FLOWERS. TO CLAIRE. – then both Preube and Fondo agree on the relational structure of the communicated thought to the extent permitted by language. It may still be that Preube and Fondo have radically different internal conceptions of what FLOWERS or GIVING or BEETLES IN BOXES are, but that is alright: we only care that the *interacting structure* of the thought-relations are the same, not their specific representations.

putative building blocks of thoughts, and how the order of words in sentences and special grammatical words affect the (de/re)construction procedures.

placeholder

The nature of the problem can be summarised as an asymmetry of information. The speaker knows the structure of a thought and has to make choices to turn that thought into text. The listener knows only the text, and must make choices to deduce the thought behind it. By this perspective, language is a shared and cooperative strategy to solve this (de/en)coding interaction. I will now outline the constraints imposed by this interaction, and then explain how context-free grammars and typological grammars partially model these constraints.

SPEAKERS CHOOSE. The speaker Preube must supply decisions about phrasing a thought in the process of speaking it. At some point at the beginning of an utterance, Preube has a thought but has not yet decided how to say it. Finding a particular phrasing requires choices to be made, because there are many ways to express even simple relational thoughts. For example, the relational content of our running example might be expressed as

ALICE LIKES THE FLOWERS BOB GIVES CLAIRE.

BOB GIVES CLAIRE FLOWERS. ALICE LIKES THE FLOWERS.

THE FLOWERS TO CLAIRE THAT BOB GIVES ARE LIKED BY ALICE.

Whether those decisions are made by committee or coinflips, those decisions represent information that must be supplied to Preube in the process of speaking a thought.

placeholder

For this reason, we consider context-free-grammars (and more generally, other string-rewrite systems) to be *Grammars of the speaker*. The start symbol S is incrementally expanded and determined by rule-applications that are selected by the speaker. The important aspect here is that the speaker has an initial state of information S that requires more information as input in order to arrive at the final sentence.

LISTENERS DEDUCE. The listener Fondo must supply decisions about which words are grammatically related, and how. Like right-of-way in driving, sometimes these decisions are settled by convention, for example, subject-verb-object order in English. Sometimes sophisticated decisions need to be made that override or are orthogonal to conventions, and we give two examples.

It is worth noting that in practice, neither grammar nor meaning strictly determines the other. Clearly there are cases where grammar supercedes: when Fondo hears *man bites dog*, despite his prior prejudices and associations about which animal is more likely to be biting, he knows that the *man* is doing the biting and the *dog* is getting bitten. Going the other way, there are many cases in which the meaning of a subphrase affects grammatical acceptability and structure.

Example 3.1.1 (Exclamations: how meaning affects grammar). The following examples from [Lakofflecture] illustrate how whether a phrase is an *exclamation* affects what kinds of grammatical constructions are acceptable. By this argument, to know whether something is an exclamation in context is an aspect of meaning, so we have cases where meaning determines grammar. Observe first that the following three phrases are all grammatically acceptable and mean the same thing.

nobody knows how many beers Bob drinks

who knows how many beers Bob drinks

God knows how many beers Bob drinks

The latter two are distinguished when *God knows* and *who knows* are exclamations. First, the modularity of grammar and meaning may not match when an exclamation is involved. For example, negating the blue text, we obtain:

somebody knows how many beers Bob drinks

who doesn't know how many beers Bob drinks

God doesn't know how many beers Bob drinks

The first two are acceptable, but mean different things; the latter means to say that everyone knows how many beers Bob drinks, which is stronger than the former. The last sentence is awkward: unlike in the first two cases, the quantified variable in the (gloss) $\dots \exists x_{\text{person}} \dots$ of *God knows* is lost, and what is left is a literal reading $\dots \neg \text{knows}(\text{God}, \dots) \dots$. Second, whether a sentence is grammatically acceptable may depend on whether an exclamation is involved. *God knows* and *who knows* can be shuffled into the sentence to behave as an intensifier as in:

Bob drinks *God knows* how many beers

Bob drinks *who knows* how many beers

But it is awkward to have:

Bob drank *nobody knows* how many beers

And it is not acceptable to have:

Bob drank *Alice knows* how many beers

Example 3.1.2 (Garden path sentences). So-called "garden path" sentences illustrate that listeners have to make choices to resolve lexical ambiguities. One such garden-path sentence is *The old man the boat*, where typically readers take *The old man* as a noun-phrase and *the boat* as another noun-phrase. We can sketch how the readers might think with a (failed) pregroup grammar derivation:

$$\frac{\frac{\text{the} : n \cdot n^{-1} \quad \text{old} : n \cdot n^{-1}}{\text{the_old} : n \cdot n^{-1}} \quad \text{man} : n \quad \frac{\text{the} : n \cdot n^{-1} \quad \text{boat} : n}{\text{the_boat} : n}}{\text{the_old_man} : n \quad \text{the_boat} : n}$$

Not a sentence!

So the reader has to backtrack, taking *The old* as a noun-phrase and *man* as the transitive verb. This yields a sentence as follows:

$$\frac{\frac{\text{the} : n \cdot n^{-1} \quad \text{old} : n}{\text{the_old} : n} \quad \text{man} :^{-1} n \cdot s \cdot n^{-1} \quad \frac{\text{the} : n \cdot n^{-1} \quad \text{boat} : n}{\text{the_old} : n}}{\frac{\text{the_old_man} : s \cdot n^{-1} \quad \text{the_old} : n}{\text{the_old_man_the_boat} : s}}$$

Garden-path sentences illustrate that listeners must make choices about what grammatical roles to assign words. We make these kinds of contextual decisions all the time with lexically ambiguous words or languages with many homophones; garden-path sentences are special in that they trick the default choices badly enough that the mental effort for correction is noticeable.

Example 3.1.3 (Ambiguous scoping). Consider the following sentence:

Everyone loves someone

The sentence is secretly (at least) two, corresponding to two possible parses. The usual reading is (glossed) $\forall x \exists y : \text{loves}(x, y)$. The odd reading is $\exists y \forall x : \text{loves}(x, y)$: a Raymond situation where there is a single person loved by everyone. We can sketch this difference formally using a simple combinatory categorial grammar.

$$\frac{\frac{\text{everyone} : (n \multimap s) \multimap s \quad \text{loves} : n \multimap s \multimap n}{\text{everyone_loves} : s \multimap n} \quad \text{someone} : (s \multimap n) \multimap s}{\text{everyone_loves_someone} : s}$$

$$\frac{\text{everyone} : (n \multimap s) \multimap s \quad \frac{\text{loves} : n \multimap s \multimap n \quad \text{someone} : (s \multimap n) \multimap s}{\text{loves_someone} : n \multimap s}}{\text{everyone_loves_someone} : s}$$

CCGs have functorial semantics in any cartesian closed category, such as one where morphisms are terms in the lambda calculus and composition is substitution []. So we might specify a semantics as follows:

$$\llbracket \text{everyone} \rrbracket = \lambda(\lambda x.V(x)).\forall x : V(x)$$

$$\llbracket \text{loves} \rrbracket = \lambda x \lambda y.\text{loves}(x, y)$$

$$\llbracket \text{someone} \rrbracket = \lambda(\lambda y.V(y)).\exists y : V(y)$$

Now we can plug-in these interpretations to obtain the two different meanings. We decorate with corners just to visually distinguish which bits are partial first-order logic.

$$\frac{\frac{\lambda(\lambda x.V(x)).\ulcorner \forall x : V(x) \urcorner : (n \multimap s) \multimap s \quad \lambda x \lambda y.\ulcorner \text{loves}(x, y) \urcorner : n \multimap s \multimap n}{\lambda y.\ulcorner \forall x : \text{loves}(x, y) \urcorner : s \multimap n} \quad \lambda(\lambda y.V(y)).\ulcorner \exists y : V(y) \urcorner : (s \multimap n) \multimap s}{\ulcorner \exists y \forall x : \text{loves}(x, y) \urcorner : s}$$

$$\frac{\lambda(\lambda x.V(x)).\ulcorner \forall x : V(x) \urcorner : (n \multimap s) \multimap s \quad \frac{\lambda x \lambda y.\ulcorner \text{loves}(x, y) \urcorner : n \multimap s \multimap n \quad \lambda(\lambda y.V(y)).\ulcorner \exists y : V(y) \urcorner : (s \multimap n) \multimap s}{\lambda x.\ulcorner \exists y : \text{loves}(x, y) \urcorner : n \multimap s}}{\ulcorner \forall x \exists y : \text{loves}(x, y) \urcorner : s}$$

4

The category of continuous relations

4.1 Continuous Relations

Recall that functions are relations, and the inverse image used in the definition of continuous maps is equivalent to the relational converse when functions are viewed as relations. So we can naïvely extend the notion of continuous maps to continuous relations between topological spaces.

Definition 4.1.4 (Continuous Relation). A continuous relation $R : (X, \tau) \rightarrow (Y, \sigma)$ is a relation $R : X \rightarrow Y$ such that

$$U \in \sigma \Rightarrow R^\dagger(U) \in \tau$$

where \dagger denotes the relational converse.

For shorthand, we denote the topology (X, τ) as X^τ . As special cases, we denote the discrete topology on X as X^\bullet , and the indiscrete topology X° .

4.2 Continuous Relations by examples

Let's consider three topological spaces and examine the continuous relations between them. This way we can build up intuitions, and prove some tool results in the process.

The **singleton space** consists of a single point which is both open and closed. We denote this space \bullet . Concretely, the underlying set and topology is

$$(\{\star\}, \{\{\star\}, \emptyset\})$$



Reminder 4.1.1 (Topological Space). A *topological space* is a pair (X, τ) , where X is a set, and $\tau \subseteq \mathcal{P}(X)$ are the *open sets* of X , such that: "nothing" and "everything" are open

$$\emptyset, X \in \tau$$

Arbitrary unions of opens are open

$$\{U_i : i \in I\} \subseteq \tau \Rightarrow \bigcup_{i \in I} U_i \in \tau$$

Finite intersections of opens are open $n \in \mathbb{N}$:

$$U_1, \dots, U_n \in \tau \Rightarrow \bigcap_{1 \leq i \leq n} U_i \in \tau$$

Reminder 4.1.2 (Relational Converse). Recall that a relation $R : S \rightarrow T$ is a subset $R \subseteq S \times T$.

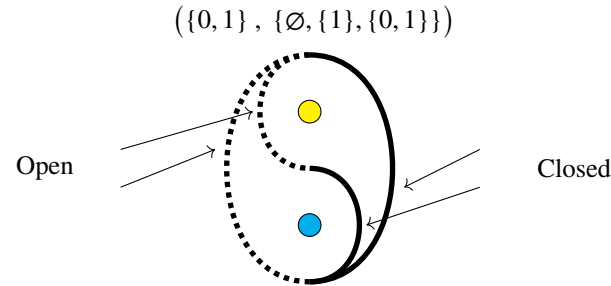
$$R^\dagger : T \rightarrow S := \{(t, s) : (s, t) \in R\}$$

Reminder 4.1.3 (Continuous Function). A function between sets $f : X \rightarrow Y$ is a continuous function between topologies $f : (X, \tau) \rightarrow (Y, \sigma)$ if

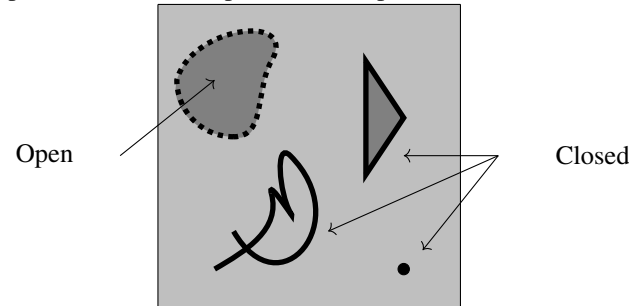
$$U \in \sigma \Rightarrow f^{-1}(U) \in \tau$$

where f^{-1} denotes the inverse image.

The **Sierpiński space** consists of two points, one of which (in yellow) is open, and the other (in cyan) is closed. We denote this space S . Concretely, the underlying set and topology is:



The **unit square** has $[0, 1] \times [0, 1]$ as its underlying set. Open sets are "blobs" painted with open balls. Points, lines, and bounded shapes are closed. We denote this space \blacksquare .



$\bullet \rightarrow \bullet$: There are two relations from the singleton to the singleton; the identity relation $\{(\bullet, \bullet)\}$, and the empty relation \emptyset . Both are topological.

$\bullet \rightarrow S$: There are four relations from the singleton to the Sierpiński space, corresponding to the subsets of S . All of them are topological.

$S \rightarrow \bullet$: There four candidate relations from the Sierpiński space to the singleton, but as we see in Example 4.2.1, not all of them are topological.

NOW WE NEED SOME ABSTRACTION. We cannot study the continuous relations between the singleton and the unit square case by case. We discover that continuous relations out of the singleton indicate arbitrary subsets, and that continuous relations into the singleton indicate arbitrary opens.

$\bullet \rightarrow \blacksquare$: Proposition 4.2.3 tells us that there are as many continuous relations from the singleton to the unit square as

Example 4.2.1 (A noncontinuous relation). The relation $\{(0, \bullet)\} \subset S \times \bullet$ is not a continuous relation: the preimage of the open set $\{\bullet\}$ under this relation is the non-open set $\{0\}$.

Terminology 4.2.2. Call a continuous relation $\bullet \rightarrow X^\tau$ a **state** of X^τ , and a continuous relation $X^\tau \rightarrow \bullet$ a **test** of X^τ .

Proposition 4.2.3. States $R : \bullet \rightarrow X^\tau$ correspond with subsets of X .

Proof. The preimage $R^\dagger(U)$ of a (non- \emptyset) open $U \in \tau$ is \star if $R(\star) \cap U$ is nonempty, and \emptyset otherwise. Both \star and \emptyset are open in $\{\star\}^\tau$. $R(\star)$ is free to specify any non- \emptyset subset of X . The empty relation handles \emptyset as

there are subsets of the unit square.

■ $\rightarrow \bullet$: Proposition 4.2.4 tells us that there are as many continuous relations from the unit square to the singleton as there are open sets of the unit square.

THERE ARE 16 CANDIDATE RELATIONS $S \rightarrow S$ TO CHECK. A case-by-case approach won't scale, so we could instead identify the building blocks of continuous relations with the same source and target space.

WHICH RELATIONS $X^\tau \rightarrow Y^\sigma$ ARE ALWAYS CONTINUOUS?

THE EMPTY RELATION IS ALWAYS CONTINUOUS.

Proposition 4.2.6. *Proof.* The preimage of the empty relation is always \emptyset , which is open by definition. \square

Reminder 4.2.5 (Empty relation). The **empty relation** $X \rightarrow Y$ relates nothing. It is defined:

$$\emptyset \subset X \times Y$$

FULL RELATIONS ARE ALWAYS CONTINUOUS

Proposition 4.2.8. *Proof.* The preimage of any subset of Y – open or not – under the full relation is the whole of X , which is open by definition. \square

Reminder 4.2.7 (Full relation). The **full relation** $X \rightarrow Y$ relates everything to everything. It is all of $X \times Y$.

FULL RELATIONS RESTRICTED TO OPEN SETS IN THE SOURCE ARE CONTINUOUS.

Proposition 4.2.9. Given an open $U \subseteq X^\tau$, and an arbitrary subset $K \subset Y^\sigma$, the relation $U \times K \subseteq X \times Y$ is open.

Proof. Consider an arbitrary open set $V \in \sigma$. Either V and K are disjoint, or they overlap. If they are disjoint, the preimage of V is \emptyset , which is open. If they overlap, the preimage of V is U , which is open. \square

CONTINUOUS FUNCTIONS ARE ALWAYS CONTINUOUS.

Proposition 4.2.10. If $f : X^\tau \rightarrow Y^\sigma$ is a continuous function, then it is also a continuous relation.

Proof. Functions are special cases of relations. The relational converse of a function viewed in this way is the same thing as the preimage. \square

THE IDENTITY RELATION IS ALWAYS CONTINUOUS. The identity relation is also the "trivial" continuous map from a space to itself, so this also follows from Proposition 4.2.10.

Proposition 4.2.12. *Proof.* The preimage of any open set under the identity relation is itself, which is open by assumption. \square

Reminder 4.2.11 (Identity relation). The **identity relation** $X \rightarrow X$ relates anything to itself. It is defined:

$$\{(x, x) : x \in X\} \subseteq X \times X$$

GIVEN TWO CONTINUOUS RELATIONS $R, S : X^\tau \rightarrow Y^\sigma$, HOW CAN WE COMBINE THEM?

Proposition 4.2.14. If $R, S : X^\tau \rightarrow Y^\sigma$ are continuous relations, so are $R \cap S$ and $R \cup S$.

Proof. Replace \square with either \cup or \cap . For any non- \emptyset open $U \in \sigma$:

$$(R \square S)^\dagger(U) = R^\dagger(U) \square S^\dagger(U)$$

As R, S are continuous relations, $R^\dagger(U), S^\dagger(U) \in \tau$, so $R^\dagger(U) \square S^\dagger(U) = (R \square S)^\dagger(U) \in \tau$. Thus $R \square S$ is also a continuous relation. \square

Corollary 4.2.15. Continuous relations $X^\tau \rightarrow Y^\sigma$ are closed under arbitrary union and finite intersection. Hence, continuous relations $X^\tau \rightarrow Y^\sigma$ form a topological space where each relation is an open set on the base space $X \times Y$, where the full relation $X \rightarrow Y$ is "everything", and the empty relation is "nothing".

A TOPOLOGICAL BASIS FOR SPACES OF CONTINUOUS RELATIONS

Definition 4.2.17 (Partial Functions). A **partial function** $X \rightarrow Y$ is a relation for which each $x \in X$ has at most a single element in its image. In particular, all functions are special cases of partial functions, as is the empty relation.

Lemma 4.2.18 (Partial functions are a \cap -ideal). The intersection $f \cap R$ of a partial function $f : X \rightarrow Y$ with any other relation $R : X \rightarrow Y$ is again a partial function.

Proof. Consider an arbitrary $x \in X$. $R(x) \cap f(x) \subseteq f(x)$, so the image of x under $f \cap R$ contains at most one element, since $f(x)$ contains at most one element. \square

Lemma 4.2.19 (Any single edge can be extended to a continuous partial function). Given any $(x, y) \in X \times Y$, there exists a continuous partial function $X^\tau \rightarrow Y^\tau$ that contains (x, y) .

Proof. Let $\mathcal{N}(x)$ denote some open neighbourhood of x with respect to the topology τ . Then $\{(z, y) : z \in \mathcal{N}(x)\}$ is a continuous partial function that contains (x, y) . \square

Proposition 4.2.20. Continuous partial functions form a topological basis for the space $(X \times Y)^{(\tau \multimap \sigma)}$ of continuous relations $X^\tau \rightarrow Y^\sigma$.

Proof. We will show that every continuous relation $R : X^\tau \rightarrow Y^\sigma$ arises as a union of partial functions. Denote the set of continuous partial functions \mathfrak{f} . We claim that:

$$R = \bigcup_{F \in \mathfrak{f}} (R \cap F)$$

Reminder 4.2.13 (Union, intersection, and ordering of relations).

Recall that relations $X \rightarrow Y$ can be viewed as subsets of $X \times Y$. So it makes sense to speak of the union and intersection of relations, and of partially ordering them by inclusion.

Reminder 4.2.16 (Topological Basis). $\mathfrak{b} \subseteq \tau$ is a basis of the topology τ if every $U \in \tau$ is expressible as a union of elements of \mathfrak{b} . Every topology has a basis (itself). Minimal bases are not necessarily unique.

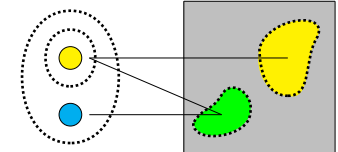


Figure 4.1: Regions of \blacksquare in the image of the yellow point alone will be coloured yellow, and regions in the image of both yellow and cyan will be coloured green:

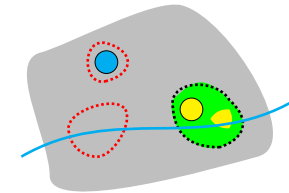
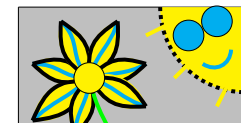


Figure 4.2: Regions in the image of the cyan point alone cannot be open sets by continuity, so they are either points or lines. Points and lines in cyan must be surrounded by an open region in either yellow or green, or else we violate continuity (open sets in red).



The \supseteq direction is evident, while the \subseteq direction follows from Lemma 4.2.19. By Lemma 4.2.18, every $R \cap F$ term is a partial function, and by Corollary 4.2.15, continuous. \square

$S \rightarrow S$: We can use Proposition 4.2.20 to write out the topological basis of continuous partial functions, from which we can take unions to find all the continuous relations, which we depict in Figure 4.5.

$S \rightarrow \blacksquare$: Now we use the colour convention of the points in S to "paint" continuous relations on the unit square "canvas", as in Figures 4.1 and 4.2. So each continuous relation is a painting, and we can characterise the paintings that correspond to continuous relations $S \rightarrow \blacksquare$ in words as follows: Cyan only in points and lines, and either contained in or at the boundary of yellow or green. Have as much yellow and green as you like.

$\blacksquare \rightarrow S$: The preimage of all of S must be an open set. So the painting cannot have stray lines or points outside of blobs. The preimage of yellow must be open, so the union of yellow and green in the painting cannot have stray lines or points outside of blobs. Point or line gaps within blobs are ok. Each connected blob can contain any colours in any shapes, subject to the constraint that if cyan appears anywhere, then either yellow or green must occur somewhere. Open blobs with no lines or points outside. Yellow and green considered alone is a painting made of blobs with no stray lines or points. If cyan appears anywhere, then either yellow or green have to appear somewhere.

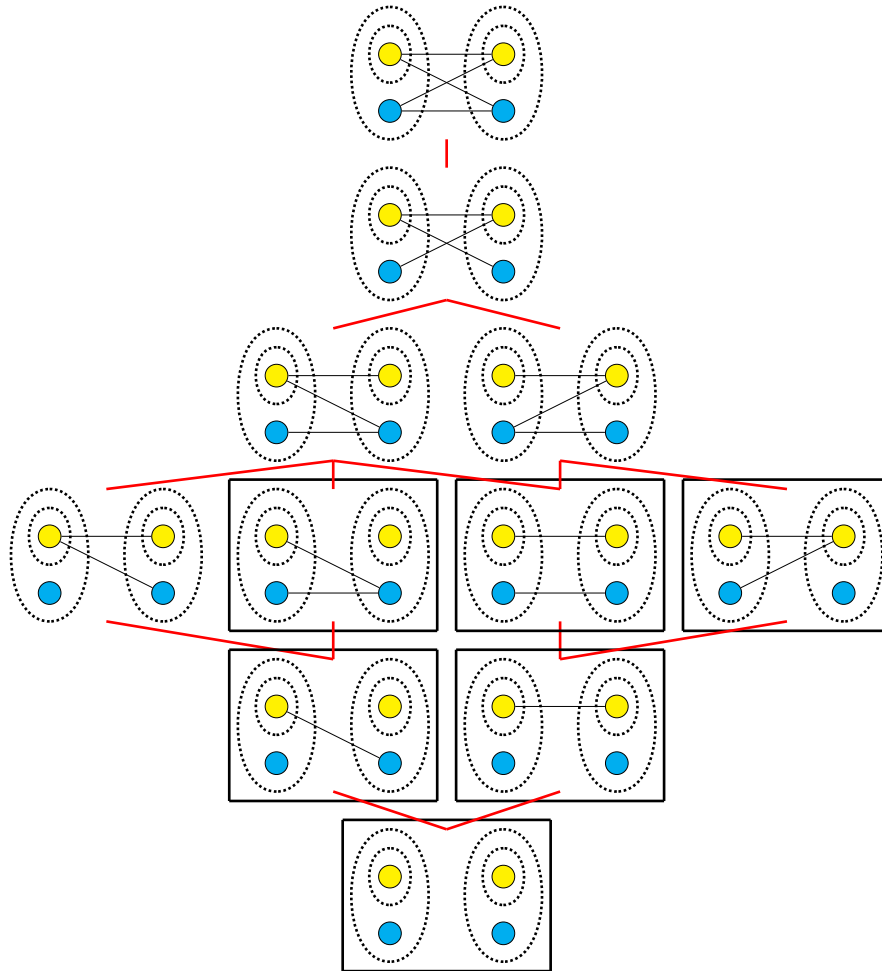


Figure 4.5: Hasse diagram of all continuous relations from the Sierpiński space to itself. Each relation is depicted left to right, and inclusion order is bottom-to-top. Relations that form the topological basis are boxed.

ONE MORE EXAMPLE FOR FUN: $[0, 1] \rightarrow \blacksquare$: We know how continuous functions from the unit line into the unit square look.

THEN WHAT ARE THE PARTIAL CONTINUOUS FUNCTIONS? If we understand these, we can obtain all continuous relations by arbitrary unions of the basis. Observe that the restriction of any continuous function to an open set in the source is a continuous partial function. The open sets of $[0, 1]$ are collections of open intervals, each of which is homeomorphic to $(0, 1)$, which is close enough to $[0, 1]$.

ANY PAINTING IS A CONTINUOUS RELATION $[0, 1] \rightarrow \blacksquare$. By colour-coding $[0, 1]$ and controlling brushstrokes, we can do quite a lot. Now we would like to develop the abstract machinery required to *formally* paint pictures with words.

4.3 The category *TopRel*

Proposition 4.3.1. continuous relations form a category **TopRel**.

Proof. IDENTITIES: Identity relations, which are always topological.

COMPOSITION: The normal composition of relations. We verify that the composite $X^\tau \xrightarrow{R} Y^\sigma \xrightarrow{S} Z^\theta$ of continuous relations is again continuous as follows:

$$U \in \theta \implies S^\dagger(U) \in \sigma \implies R^\dagger \circ S^\dagger(U) = (S \circ R)^\dagger \in \tau$$

ASSOCIATIVITY OF COMPOSITION: Inherited from **Rel**. □

4.4 Biproducts

We exhibit a free-forgetful adjunction between **Rel** and **TopRel**.

Definition 4.4.1 ($F: \mathbf{Rel} \rightarrow \mathbf{TopRel}$). We define the action of the functor F :

On objects $F(X) := X^\star$, (X with the discrete topology)

On morphisms $F(X \xrightarrow{R} Y) := X^\star \xrightarrow{R} Y^\star$. Recall that any relation between sets is continuous with respect to the discrete topology.



Figure 4.6: continuous functions $[0, 1] \rightarrow \blacksquare$ follow the naïve notion of continuity: "a line one can draw on paper without lifting the pen off the page".

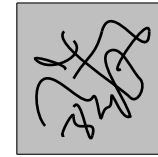


Figure 4.7: So a continuous partial function is "(countably) many (open-ended) lines, each of which one can draw on paper without lifting the pen off the page."

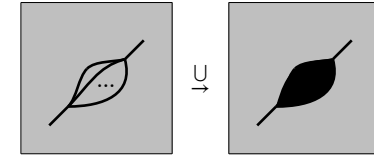


Figure 4.8: We can control the thickness of the brushstroke, by taking the union of (uncountably) many lines.



Figure 4.9: Assign the visible spectrum of light to $[0, 1]$. Colour open sets according to perceptual addition of light, computing brightness by normalising the measure of the open set.



Figure 4.10: Like it or not, a continuous relation $[0, 1] \rightarrow \blacksquare$: "The Starry Night", by Vincent van Gogh.

Evidently identities and associativity of composition are preserved.

Definition 4.4.2 ($U: \mathbf{TopRel} \rightarrow \mathbf{Rel}$). We define the action of the functor U :

On objects $U(X^\tau) := X$

On morphisms $U(X^\tau \xrightarrow{R} Y^\sigma) := X \xrightarrow{R} Y$

Evidently identities and associativity of composition are preserved.

Proposition 4.4.3 ($F \dashv U \dashv F$). *Proof.* By triangular identities.

The composite FU is precisely equal to the identity functor on \mathbf{Rel} . The unit natural transformation $1_{\mathbf{Rel}} \Rightarrow FU$ we take to be the identity morphisms.

$$\eta_X := \text{id}_X$$

The counit natural transformation $UF \Rightarrow 1_{\mathbf{TopRel}}$ we define:

$$\epsilon_{X^\tau} : X^\star \rightarrow X^\tau := \{(x, x) : x \in X\}$$

Now we verify the triangle identities...

placeholder

□

Corollary 4.4.4. \mathbf{TopRel} has a zero object and biproducts.

Proof.

□

4.5 Symmetric Monoidal Closed structure

Proposition 4.5.2. $(\mathbf{TopRel}, \{\star\}^\bullet, X^\tau \otimes Y^\sigma := (X \times Y)^{(\tau \times \sigma)})$ is a symmetric monoidal closed category.

Proof. TENSOR UNIT: The one-point space \bullet . Explicitly, $\{\star\}$ with topology $\{\emptyset, \{\star\}\}$.

TENSOR PRODUCT: For objects, $X^\tau \otimes Y^\sigma$ has base set $X \times Y$ equipped with the product topology $\tau \times \sigma$. For morphisms, $R \otimes S$ the product of relations. We show that the tensor of continuous relations is again a continuous relation. Take continuous relations $R : X^\tau \rightarrow Y^\sigma$, $S : A^\alpha \rightarrow B^\beta$, and let U be open in the product topology $(\sigma \times \beta)$.

Reminder 4.5.1 (Product Topology). We denote the product topology of X^τ and Y^σ as $(X \times Y)^{(\tau \times \sigma)}$. $\tau \times \sigma$ is the topology on $X \times Y$ generated by the basis $\{t \times s : t \in \mathfrak{b}_\tau, s \in \mathfrak{b}_\sigma\}$, where \mathfrak{b}_τ and \mathfrak{b}_σ are bases for τ and σ respectively.

Reminder 4.5.3 (Product of relations). For relations between sets $R : X \rightarrow Y, S : A \rightarrow B$, the product relation $R \times S : X \times A \rightarrow Y \times B$ is defined to be

$$\{((x, a), (y, b)) : (x, y) \in R, (a, b) \in S\}$$

We need to prove that $(R \times S)^\dagger(U) \in (\tau \times \alpha)$. We may express U as $\bigcup_{i \in I} y_i \times b_i$, where the y_i and b_i are in the bases \mathfrak{b}_σ and \mathfrak{b}_ρ respectively. Since for any relations we have that $R(A \cup B) = R(A) \cup R(B)$ and $(R \times S)^\dagger = R^\dagger \times S^\dagger$:

$$\begin{aligned} & (R \times S)^\dagger\left(\bigcup_{i \in I} y_i \times b_i\right) \\ &= \bigcup_{i \in I} (R \times S)^\dagger(y_i \times b_i) \\ &= \bigcup_{i \in I} (R^\dagger \times S^\dagger)(y_i \times b_i) \end{aligned}$$

Since each y_i is open and R is continuous, $R^\dagger(y_i) \in \tau$. Symmetrically, $S^\dagger(b_i) \in \alpha$. So each $(R^\dagger \times S^\dagger)(y_i \times b_i) \in (\tau \times \alpha)$. Topologies are closed under arbitrary union, so we are done.

UNITORS: The left unitors $\lambda_{X^\tau} : \bullet \times X^\tau \rightarrow X^\tau$ maps $(\star, x) \mapsto x$, and we reverse the direction of the mapping to obtain the inverse $\lambda_{X^\tau}^{-1}$. The construction is symmetric for the right unitors ρ_{X^τ} .

ASSOCIATORS:

The associators $\alpha_{X^\tau, Y^\sigma, Z^\rho}$

BRAIDS:

...

COHERENCES:

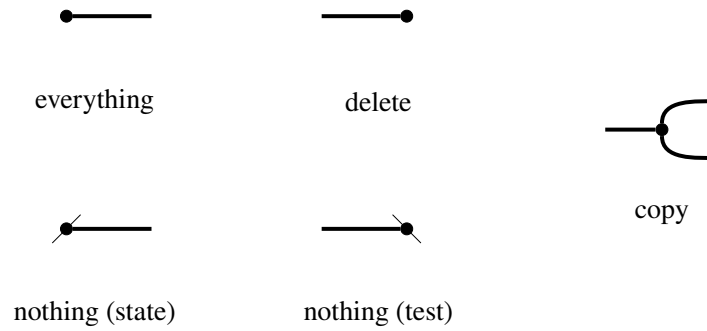
...

□

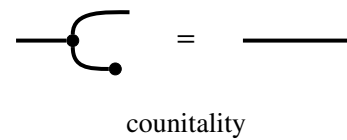
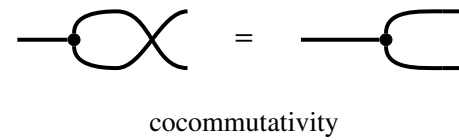
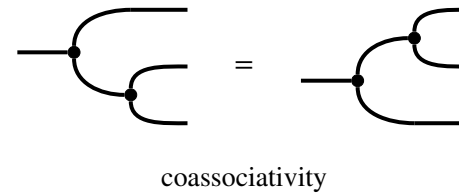
4.6 *TopRel* diagrammatically

4.6.1 Relations that are always continuous

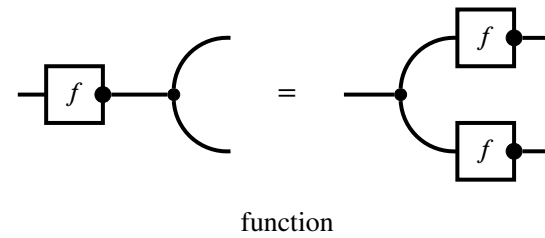
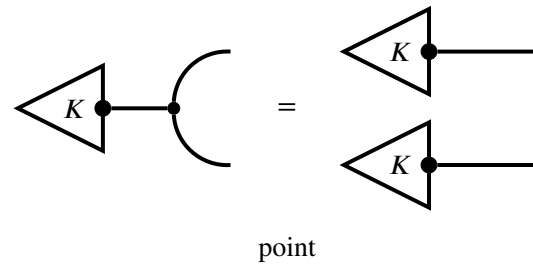
HERE ARE FIVE CONTINUOUS RELATIONS FOR ANY X^τ :



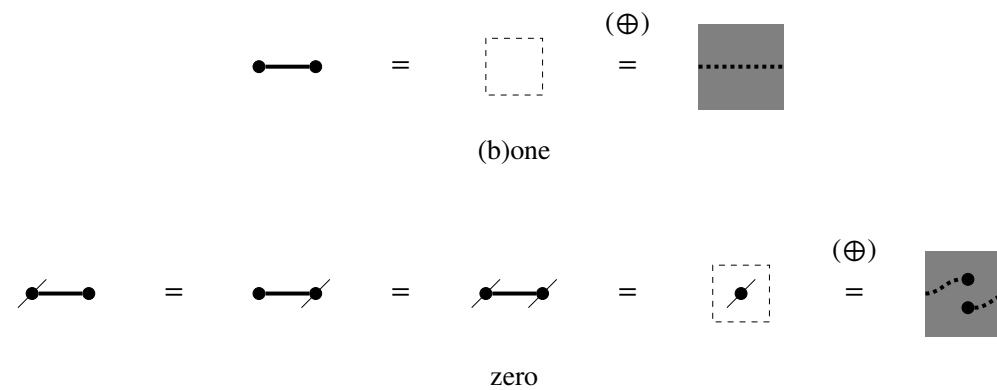
COPY AND DELETE OBEY THE FOLLOWING EQUALITIES:



THE COPY MAP CAN ALSO BE USED TO DISTINGUISH THE DETERMINISTIC MAPS – POINTS AND FUNCTIONS – WHICH WE NOTATE WITH AN EXTRA DOT.



EVERYTHING, DELETE, NOTHING-STATES AND NOTHING-TESTS COMBINE TO GIVE TWO NUMBERS, ONE AND ZERO. There are extra expressions in grey squares above: they anticipate the tape-diagrams we will later use to graphically express another monoidal product of **TopRel**, the direct sum \oplus .



ZERO SCALARS TURN ENTIRE DIAGRAMS INTO ZERO MORPHISMS. There is a zero-morphism for every input-output pair of objects in **TopRel**.

$$\begin{array}{c}
 \bullet \\
 \diagup \\
 X^\tau \text{ --- } \boxed{f} \text{ --- } Y^\sigma
 \end{array}
 \quad
 \forall X^\tau \forall Y^\sigma \forall f
 \quad
 =
 \quad
 \begin{array}{c}
 X^\tau \text{ --- } \bullet \quad \bullet \text{ --- } Y^\sigma \\
 \diagup \quad \diagup
 \end{array}$$

zero

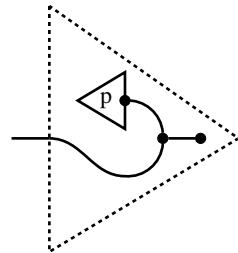
4.7 Populating space with shapes using sticky spiders

4.7.1 When does an object have a spider (or something close to one)?

Example 4.7.2 (The copy-compare spiders of **Rel** are not always continuous). The compare map for the Sierpiński space is not continuous, because the preimage of $\{0, 1\}$ is $\{(0, 0), (1, 1)\}$, which is not open in the product space of S with itself.

Proposition 4.7.3. The copy map is a spider iff the topology is discrete.

Proof. Discrete topologies inherit the usual copy-compare spiders from **Rel**, so we have to show that when the copy map is a spider, the underlying wire must have a discrete topology. Suppose that some wire has a spider, and construct the following open set using an arbitrary point p :



It will suffice to show that this open set is the singleton $\{p\}$ – when all singletons are open, the topology is discrete. As a lemma, using frobenius rules and the property of zero morphisms, we can show that comparing distinct points

Reminder 4.7.1 (copy-compare spiders of **Rel**). For a set X , the *copy* map $X \rightarrow X \times X$ is defined:

$$\{(x, (x, x)) : x \in X\}$$

the *compare* map $X \times X \rightarrow X$ is defined:

$$\{((x, x), x) : x \in X\}$$

These two maps are the (co)multiplications of special frobenius algebras. The (co)units are *delete*:

$$\{(x, \star) : x \in X\}$$

and *everything*:

$$\{(\star, x) : x \in X\}$$

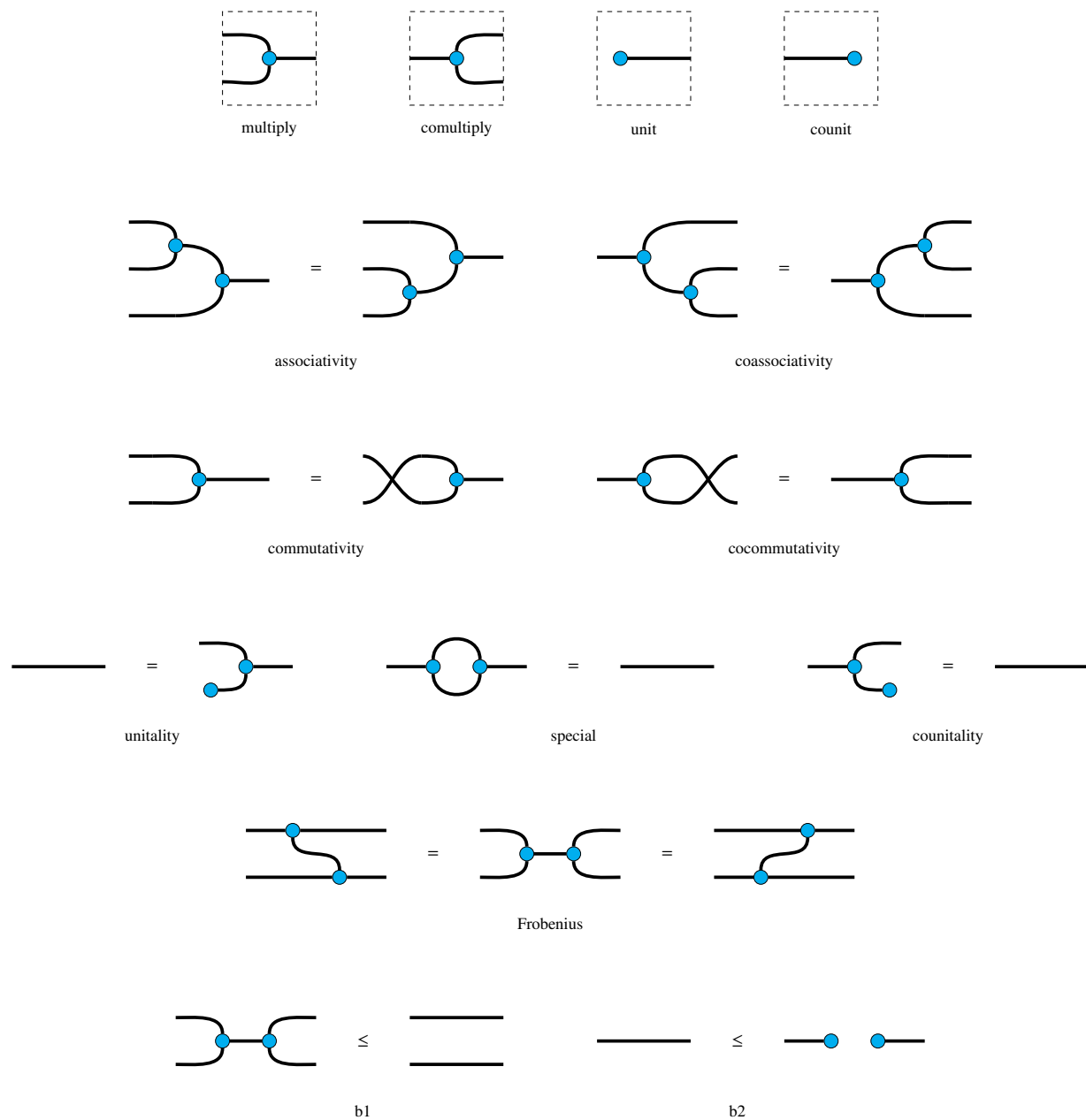
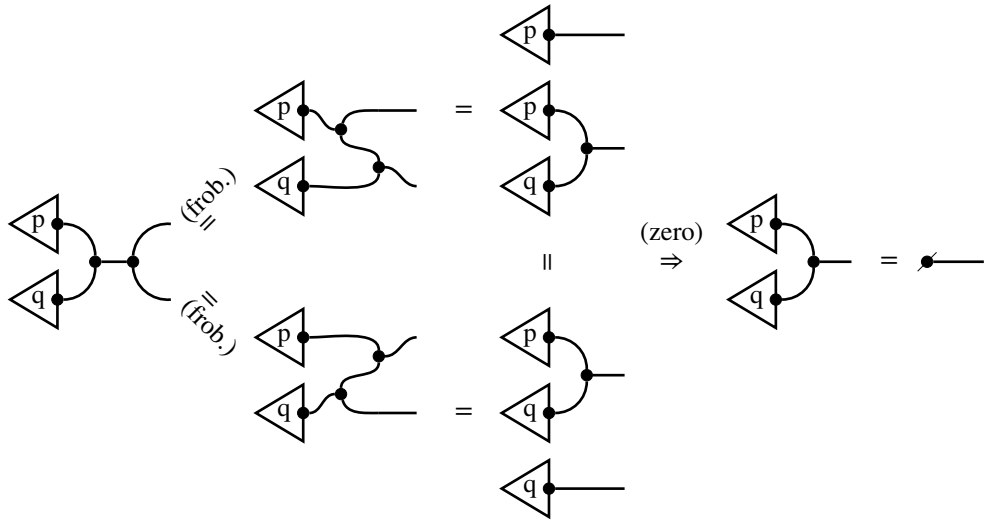
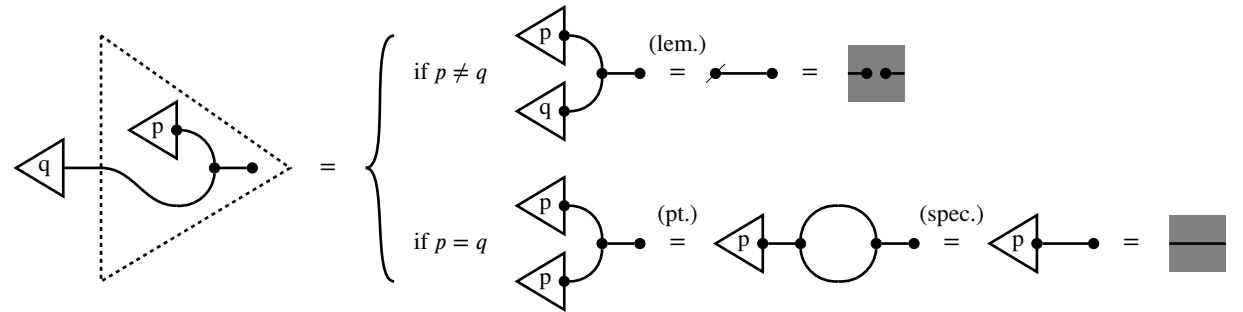


Figure 4.11: The generators (in dashed boxes) and relations that make a spider. When the spider satisfies in addition the three inequalities b1-3, we call it a **relation-spider**.

$p \neq q$ yields the \emptyset state.



The following case analysis shows that our open set only contains the point p .



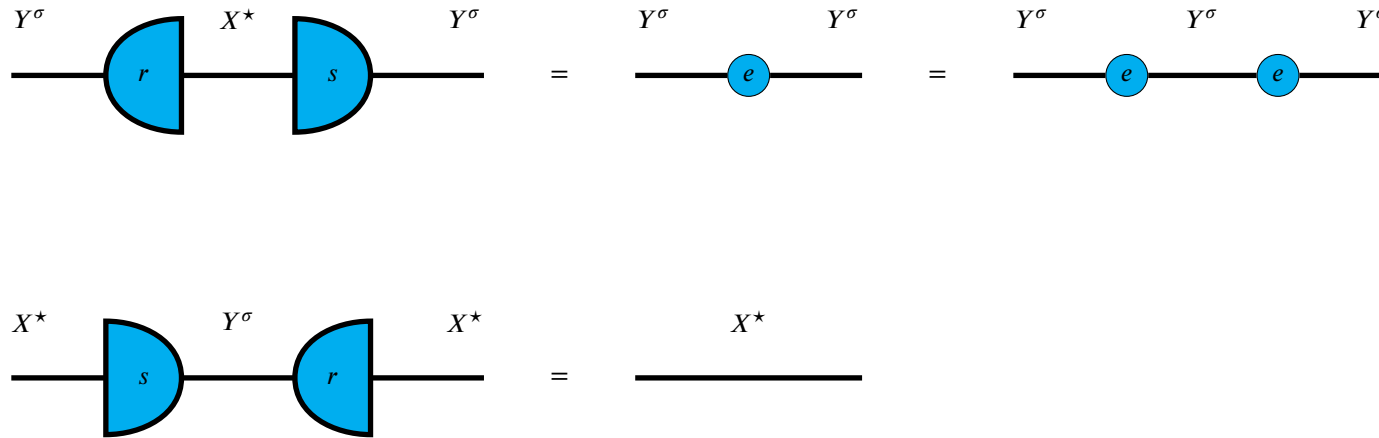
□

Reminder 4.7.4 (Split idempotents). An **idempotent** in a category is a map $e : A \rightarrow A$ such that

WE CAN USE SPLIT IDEMPOTENTS TO TRANSFORM COPY-SPIDERS FROM DISCRETE TOPOLOGIES TO ALMOST-SPIDERS ON OTHER SPACES. We can graphically express the behaviour of a split idempotent e as follows, where the semicircles for the section and retract r, s form a visual pun.

A **split idempotent** is an idempotent $e : A \rightarrow A$ along with a **retract** $r : A \rightarrow B$ and a **section** $s : B \rightarrow A$ such that:

$$\begin{array}{ccccc} A & \xrightarrow{e} & A & \xrightarrow{r} & B & \xrightarrow{s} & A \\ & & & & s & & r & & id \end{array}$$



AN IDEMPOTENT ON Y^σ THAT SPLITS THROUGH A DISCRETE TOPOLOGY X^\star DOES THESE THINGS:

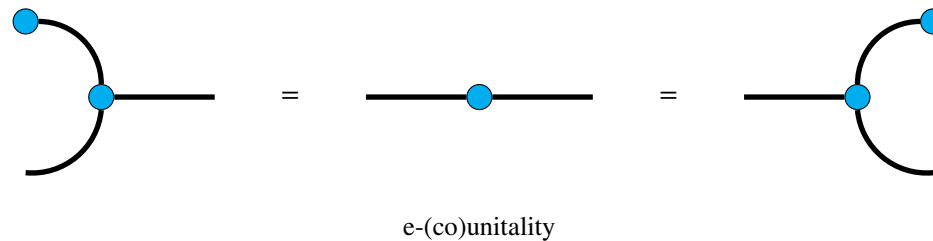
The section: picks a subset of $s(x) \subseteq Y$ for each point $x \in X$. In order to be a split idempotent, $s(x)$ must be distinct for distinct points x .

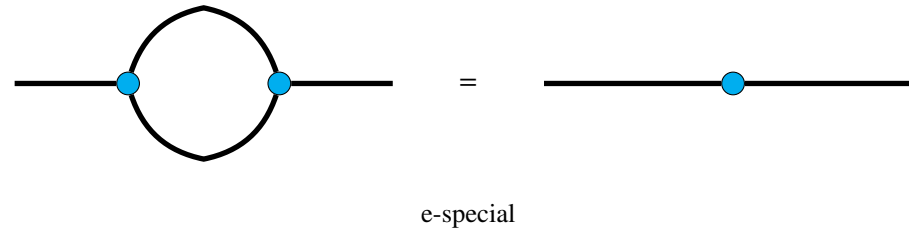
The retract: (reading backwards) picks an open set $r(x) \in \sigma$ for each point $x \in X$. In order to be a split idempotent, $r(x)$ only overlaps $s(x)$ out of all other selected shapes:

$$r(x) \cap s(x') = 1 \iff x = x'$$

The combined effect is that the shapes $s(x)$ become copiable elements of the sticky spider, just as the points are the copiable elements of a regular spider.

Definition 4.7.5 (Sticky spiders). A **sticky spider** (or just an e -spider, if we know that e is a split idempotent), is a spider *except* every identity wire on any side of an equation in Figure 4.11 is replaced by the idempotent e .

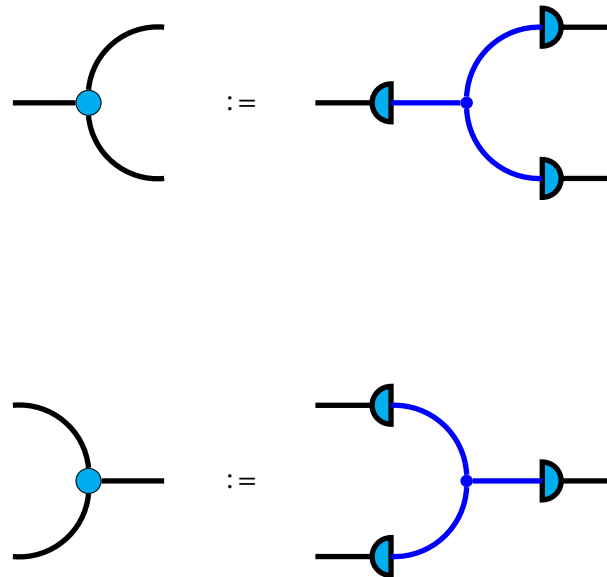




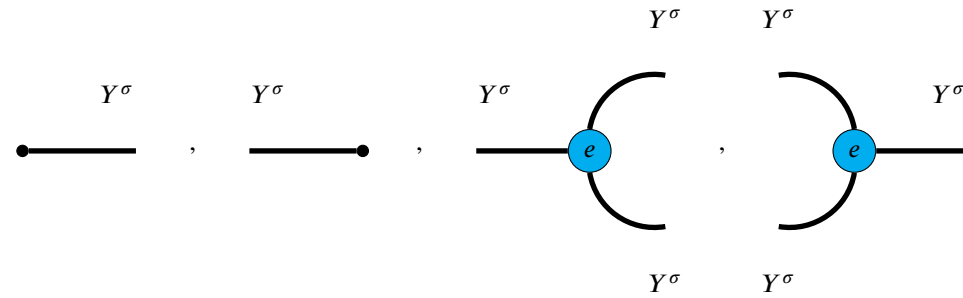
The desired graphical behaviour of a sticky spider is that one can still coalesce all connected spider-bodies together, but the 1-1 spider "sticks around" rather than disappearing as the identity. This is achieved by the following rules that cohere the idempotent e with the (co)unit and (co)multiplications.

placeholder

Construction 4.7.6 (Sticky spiders from split idempotents). Given an idempotent $e : Y^\sigma \rightarrow Y^\sigma$ that splits through a discrete topology X^\star , we construct a new (co)multiplication as follows:

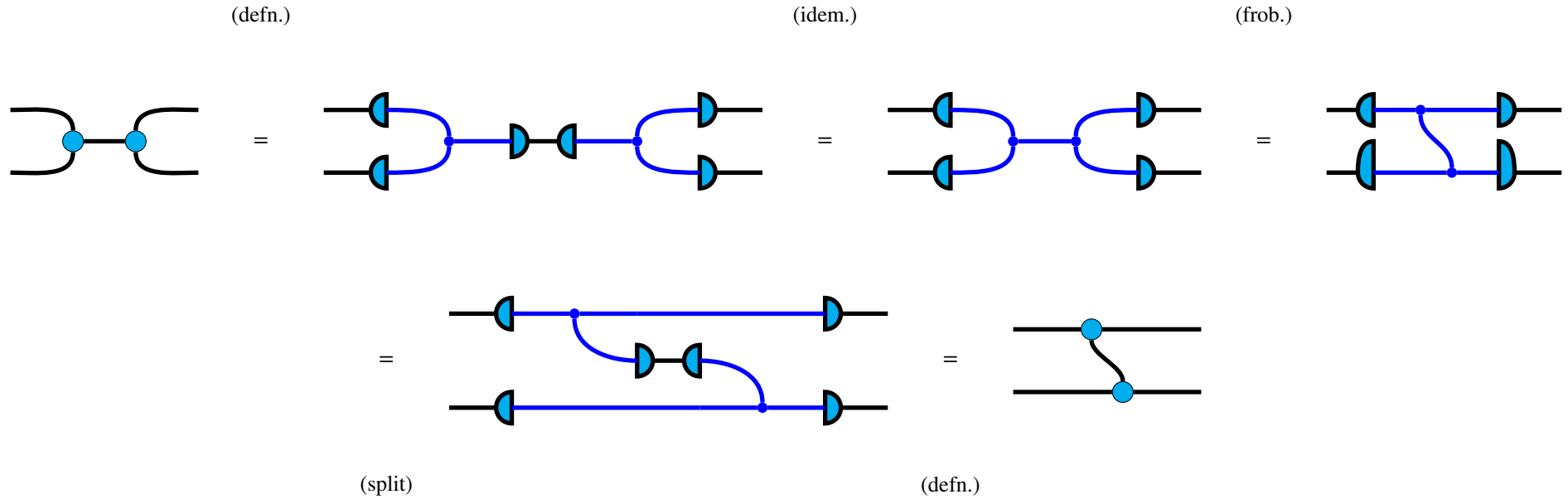


Proposition 4.7.7 (Every idempotent that splits through a discrete topology gives a sticky spider).



is a sticky spider

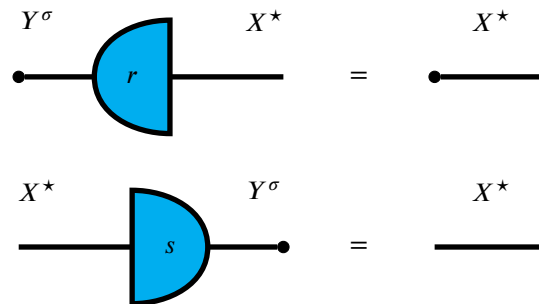
Proof. We can check that our construction satisfies the frobenius rules as follows. We only present one equality; the rest follow the same idea.



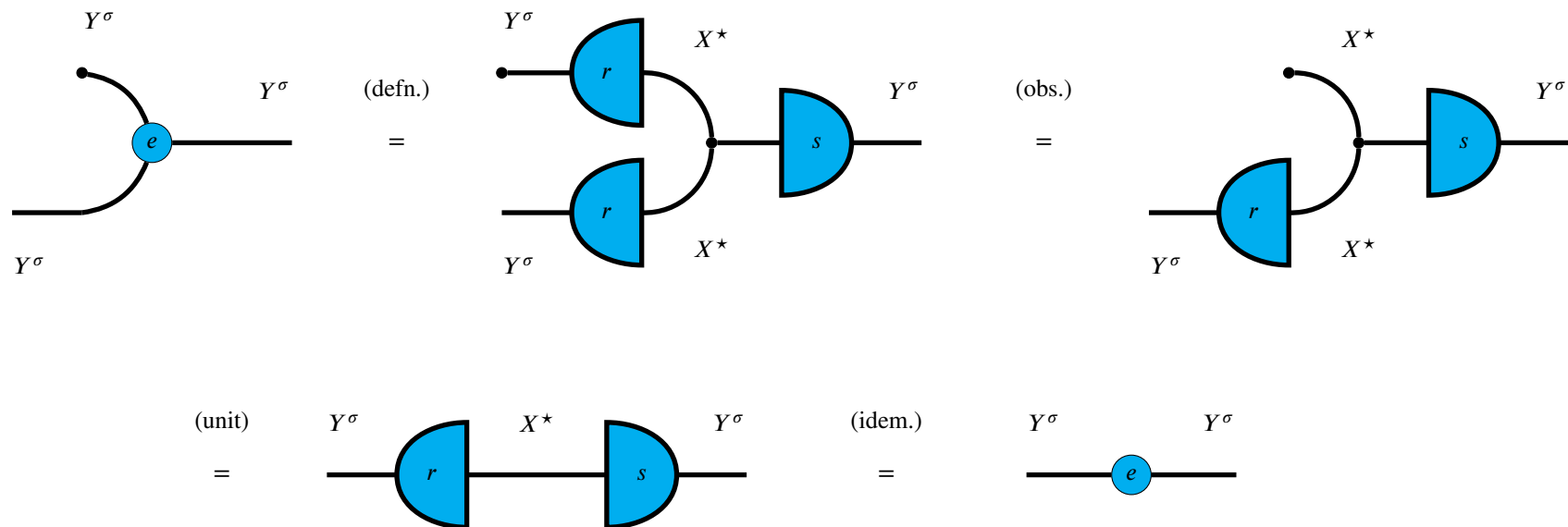
To verify the sticky spider rules, we first observe that since

$$X^\star \xrightarrow{s} Y^\sigma \xrightarrow{r} X^\star = X^\star \xrightarrow{id} X^\star$$

r must have all of X^\star in its image, and s must have all of X^\star in its preimage, so we have the following:



Now we show that e-unitality holds:

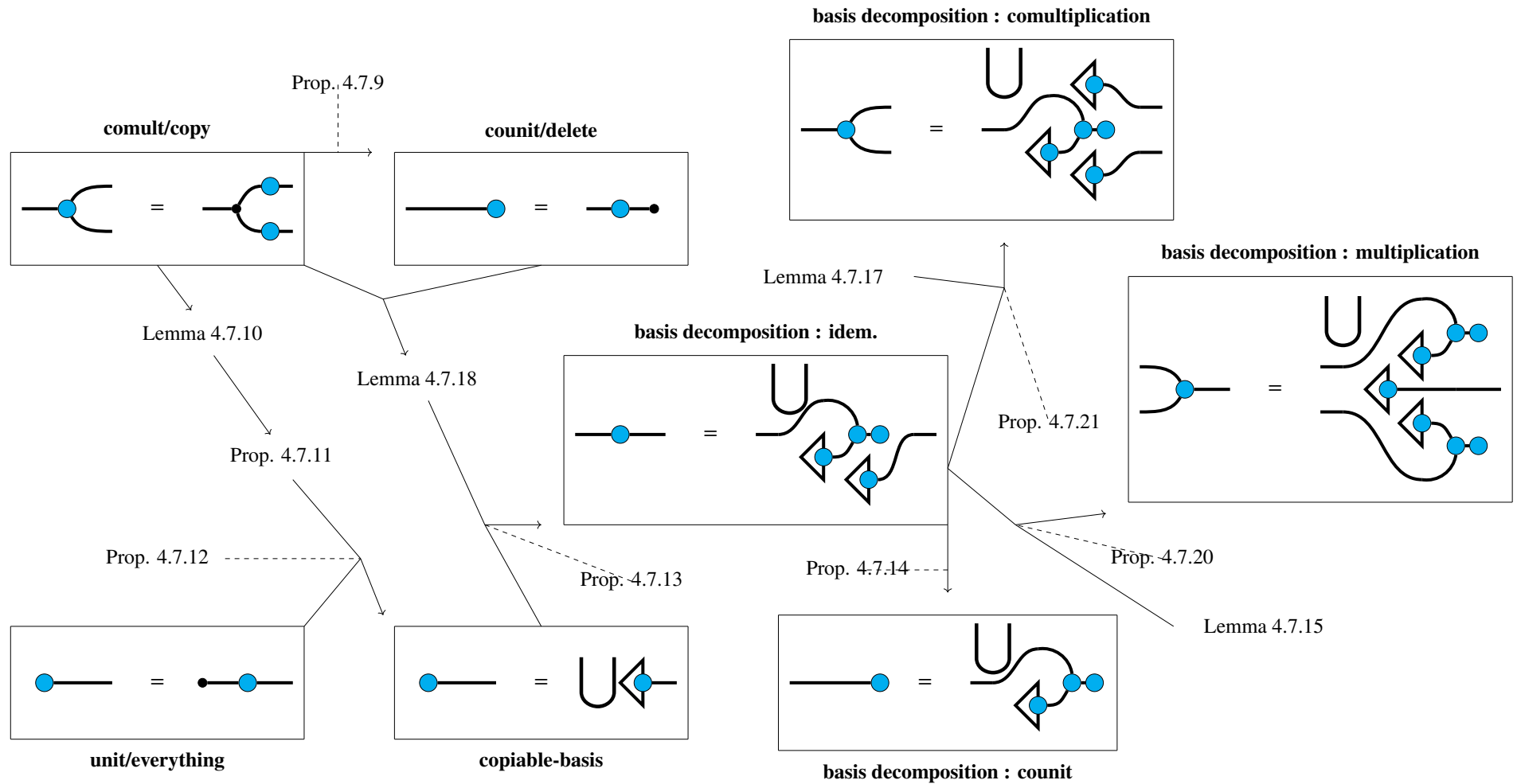


The proofs of e-counitality, and e-speciality follow similarly.

WE CAN PROVE A PARTIAL CONVERSE OF PROPOSITION 4.7.7: we can identify two diagrammatic equations that tell us precisely when a sticky spider has an idempotent that splits through some discrete topology.

Theorem 4.7.8. A sticky spider has an idempotent that splits through a discrete topology if and only if in addition to the sticky spider equalities, the following relations are also

satisfied.



Proposition 4.7.9 (comult/copy implies counit/delete).

$$\begin{array}{c} \text{---} \bullet \text{---} \\ \text{---} \end{array} = \begin{array}{c} \text{---} \bullet \text{---} \\ \text{---} \bullet \text{---} \end{array} \Rightarrow \text{---} \bullet = \text{---} \bullet \text{---}$$

Proof.

$$\begin{array}{c} \text{---} \bullet \text{---} \\ \text{---} \end{array} \stackrel{\text{(comult/copy)}}{=} \begin{array}{c} \text{---} \bullet \text{---} \\ \text{---} \bullet \text{---} \end{array} \stackrel{\text{(del)}}{\subseteq} \begin{array}{c} \text{---} \bullet \text{---} \\ \text{---} \bullet \text{---} \end{array} \\
 \stackrel{\cong}{\downarrow} \quad \quad \quad \stackrel{\text{(e-unit)}}{=} \quad \quad \quad \stackrel{\cong}{\downarrow} \quad \quad \quad \text{(copy-del)} \\
 \text{(del)} \quad \quad \quad \begin{array}{c} \text{---} \bullet \text{---} \\ \text{---} \bullet \text{---} \end{array} = \text{---} \bullet \text{---}$$

So:

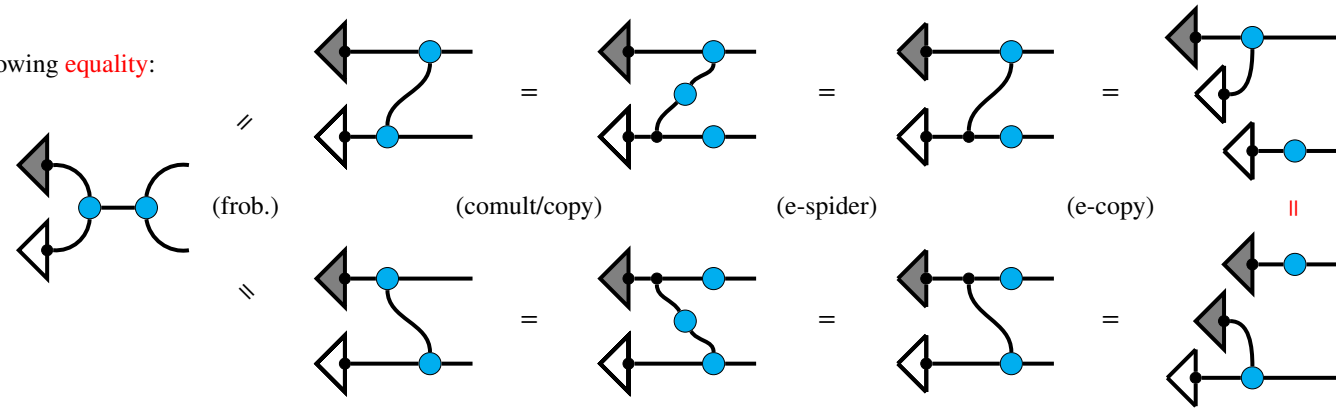
$$\text{---} \bullet \text{---} = \begin{array}{c} \text{---} \bullet \text{---} \\ \text{---} \bullet \text{---} \end{array}$$

So:

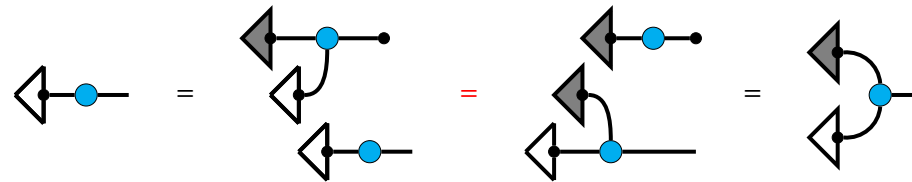
$$\text{---} \bullet \text{---} = \text{---} \bullet \text{---} \bullet = \begin{array}{c} \text{---} \bullet \text{---} \\ \text{---} \bullet \text{---} \end{array} = \text{---} \bullet \text{---}$$

□

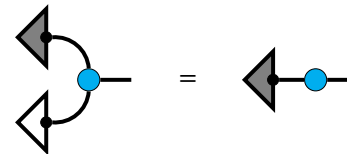
So we have the following **equality**:



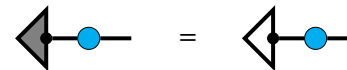
Which implies:



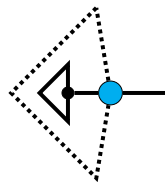
and symmetrically,



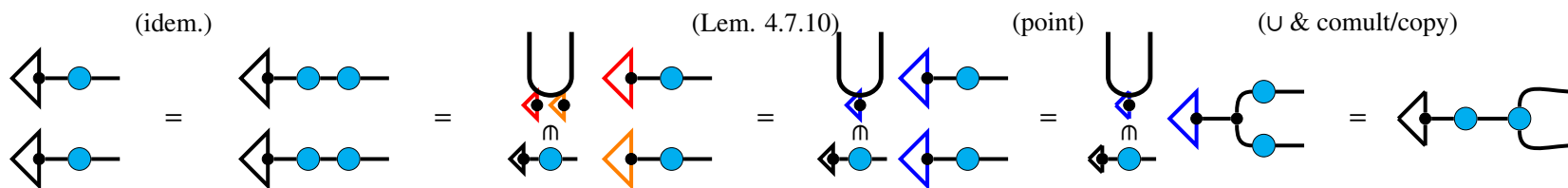
So we have the claim:



Proposition 4.7.11 (e of any point is e -copiable).

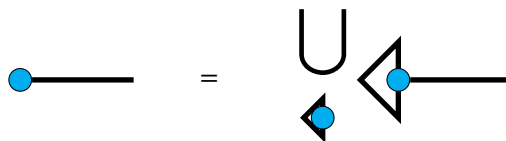


Proof.



□

Proposition 4.7.12 (The unit is the union of all e -copiables).



Proof.

The union of *all* e -copiables is a subset of the unit.

$$\begin{array}{ccccccc}
 \bigcup_{\forall} \triangleleft \bullet \text{---} & = & \bigcup_{\forall} \triangleleft \bullet \bullet \text{---} & \subseteq & \bullet \text{---} \bullet \text{---} & = & \bullet \text{---} \\
 & \text{(Lem. 4.7.18)} & & \text{(evr.)} & & \text{(unit/evr.)} &
 \end{array}$$

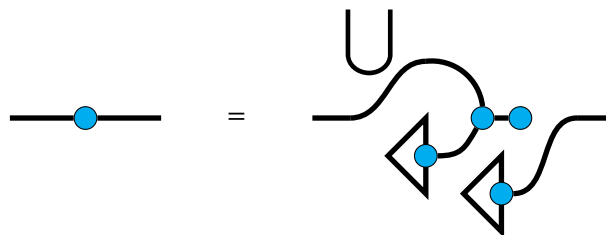
The unit is *some* union of e -copiables.

$$\begin{array}{ccccccc}
 \bullet \text{---} & = & \bullet \text{---} \bullet \text{---} & = & \begin{array}{c} \bigcup_{\forall} \triangleleft \bullet \text{---} \\ \in \bullet \text{---} \\ \triangleleft \bullet \text{---} \end{array} & = & \bigcup_{?} \triangleleft \bullet \text{---} \\
 & \text{(unit/evr.)} & & & & \text{(Prop. 4.7.11)} &
 \end{array}$$

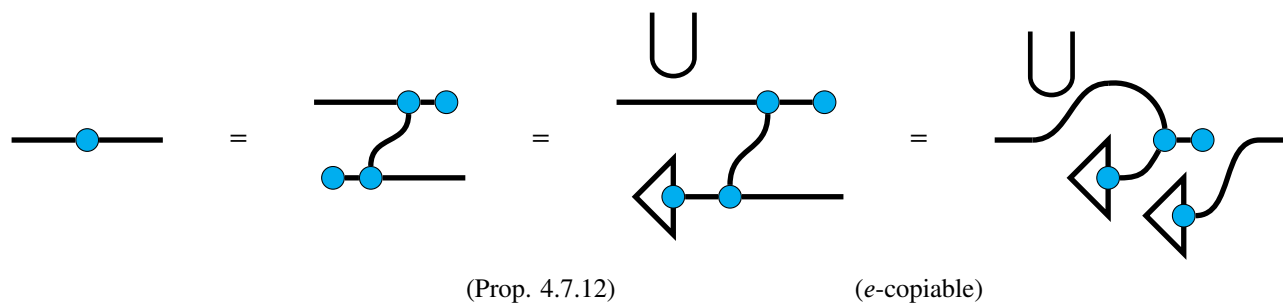
So the containment must be an equality.

$$\bullet \text{---} = \bigcup_{\forall} \triangleleft \bullet \text{---}$$

Proposition 4.7.13 (e -copiable decomposition of e).

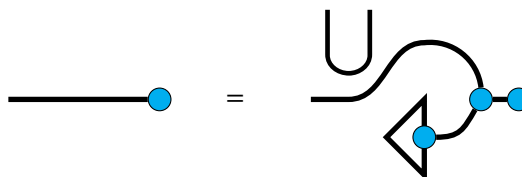


Proof.



□

Proposition 4.7.14 (e -copiable decomposition of counit).



Proof.

(Prop. 4.7.13)

The diagram shows a sequence of four expressions separated by equals signs. The first expression is a horizontal line with a blue dot. The second expression is a horizontal line with a blue dot followed by a black dot. The third expression is a complex diagram featuring a wavy line, a U-shape, and two triangles containing blue dots. The fourth expression is a similar diagram with a different arrangement of the triangles. The text "(Prop. 4.7.13)" is positioned above the third diagram.

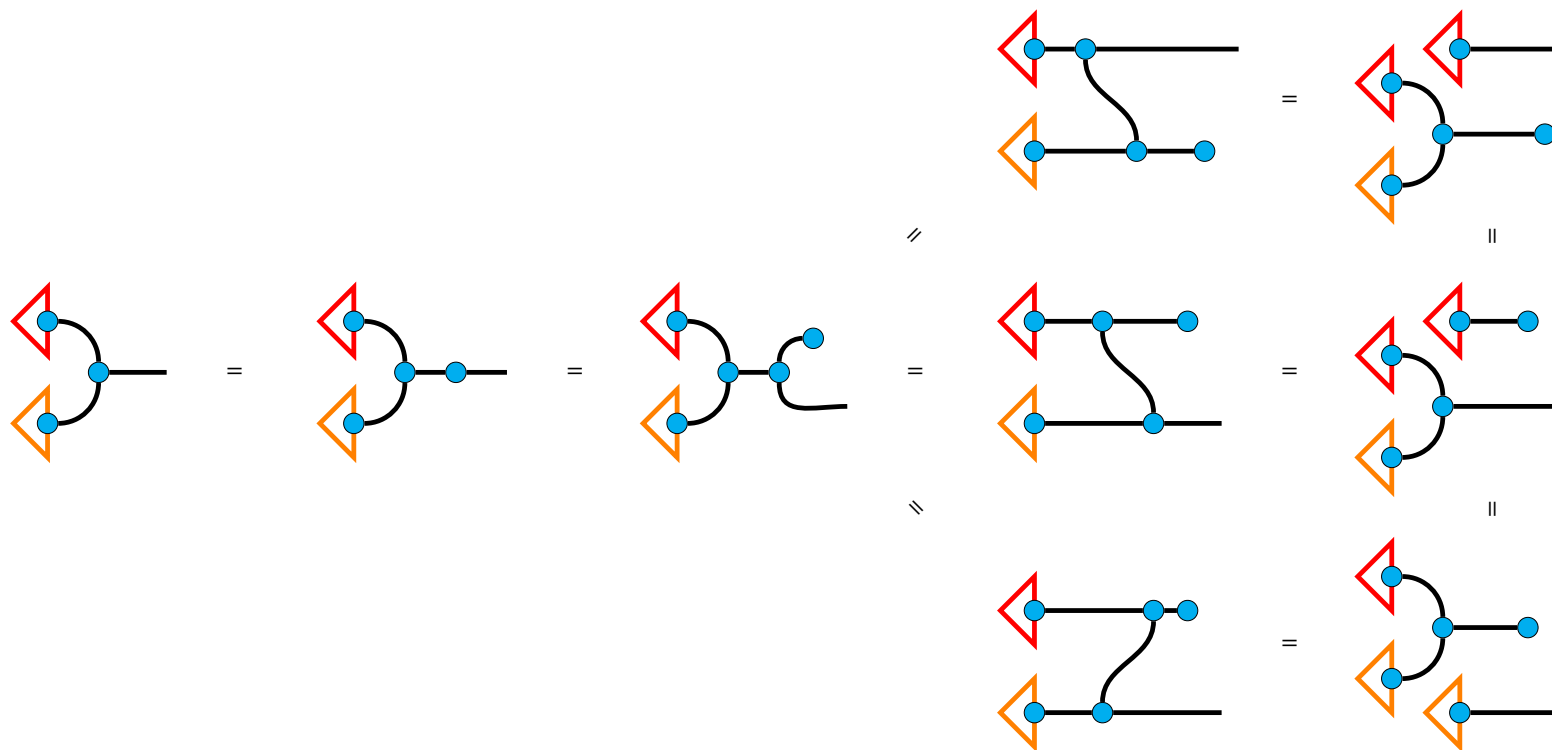
□

THE e -COPIABLE STATES REALLY DO BEHAVE LIKE AN ORTHONORMAL BASIS, AS THE FOLLOWING LEMMAS SHOW.

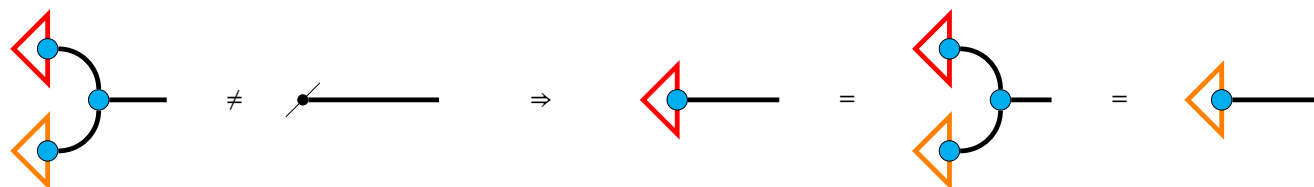
Lemma 4.7.15 (e -copiables are orthogonal under multiplication).

The diagram shows a red triangle and an orange triangle connected by a curved line, with a blue dot in the middle. This is followed by an equals sign and a large left curly brace. Inside the brace, there are two rows. The first row shows a black line with a diagonal slash, followed by "if" and a red triangle with a blue dot, followed by " \neq " and an orange triangle with a blue dot. The second row shows a red triangle with a blue dot, followed by "if" and a red triangle with a blue dot, followed by " $=$ " and an orange triangle with a blue dot.

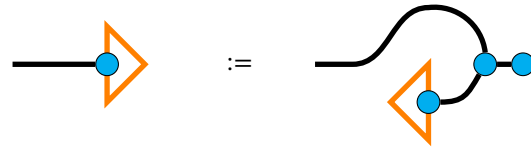
Proof.



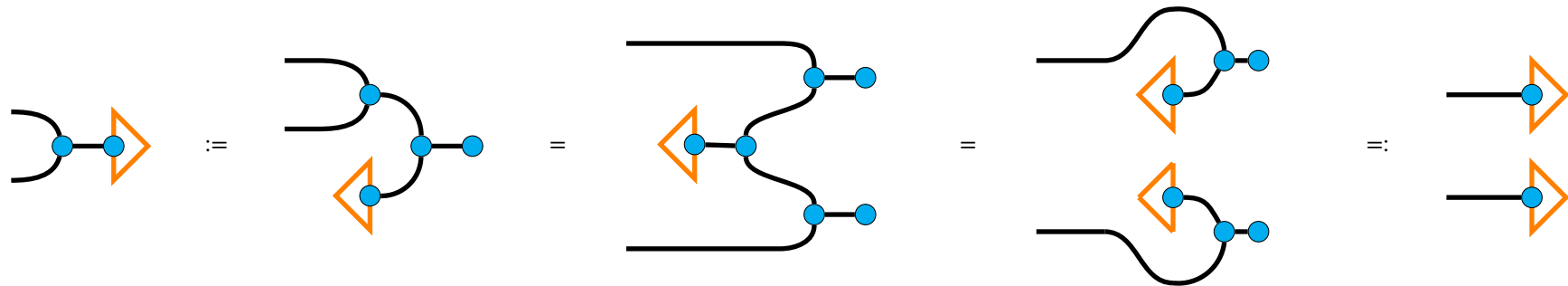
So:



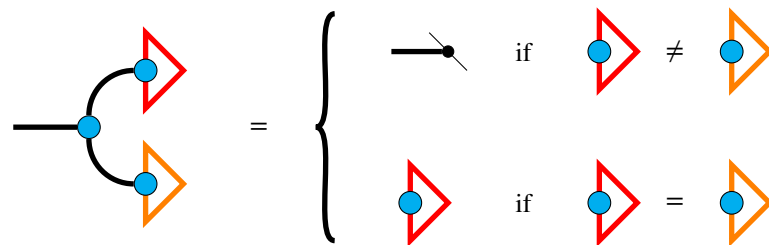
Convention 4.7.16 (Shorthand for the open set associated with an e -copiable). We introduce the following diagrammatic shorthand.



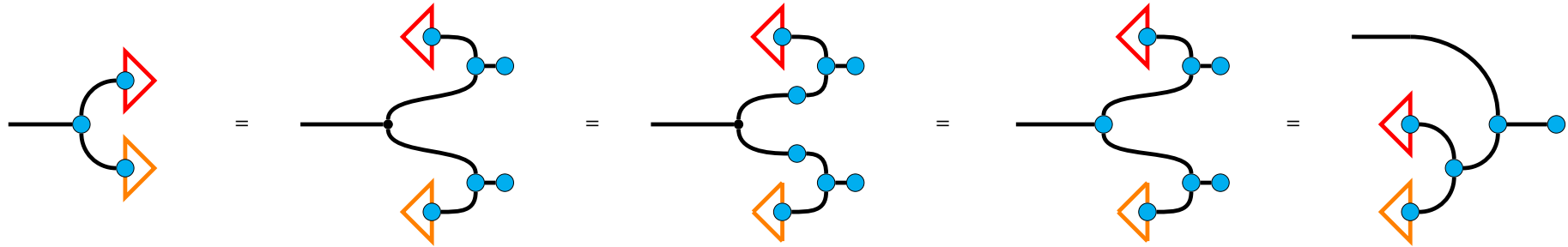
Including the coloured dot is justified, because these open sets are co-copiable with respect to the multiplication of the sticky spider.



Lemma 4.7.17 (Co-match).

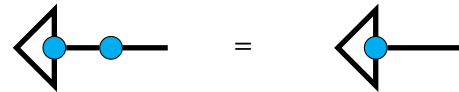


Proof.



The claim then follows by applying Lemma 4.7.15 to the final diagram. \square

Lemma 4.7.18 (e-copiables are e-fixpoints).

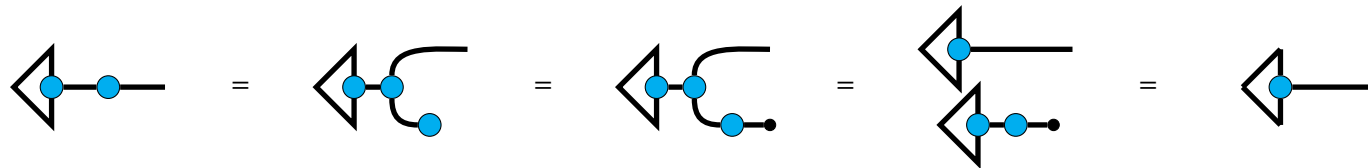


Proof.

(e-counit)

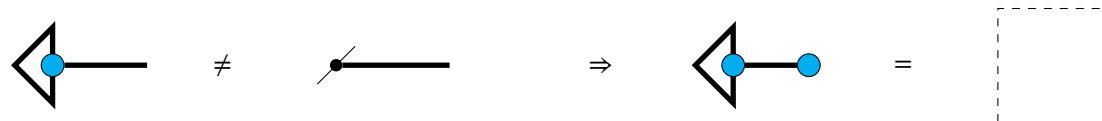
(coun/del)

(e-copy)



Observe that the final equation of the proof also holds when the initial e-copiable is the empty set. \square

Lemma 4.7.19 (e-copiables are normal).



Proof.

$$\begin{array}{ccccccc}
 & & \text{(coun/del)} & & \text{(Lem. 4.7.18)} & & \text{(Prem.)} \\
 \triangleleft \bullet \text{---} \bullet & = & \triangleleft \bullet \text{---} \bullet \text{---} \bullet & = & \triangleleft \bullet \text{---} \bullet & = & \square
 \end{array}$$

□

Proposition 4.7.20 (*e-copiable decomposition of multiplication*).

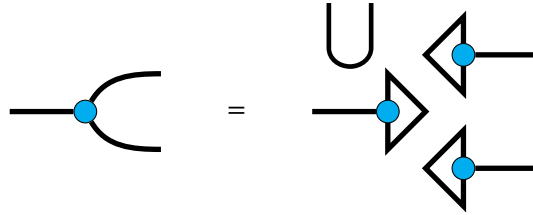
$$\text{multiplication} = \text{copiable decomposition}$$

Proof.

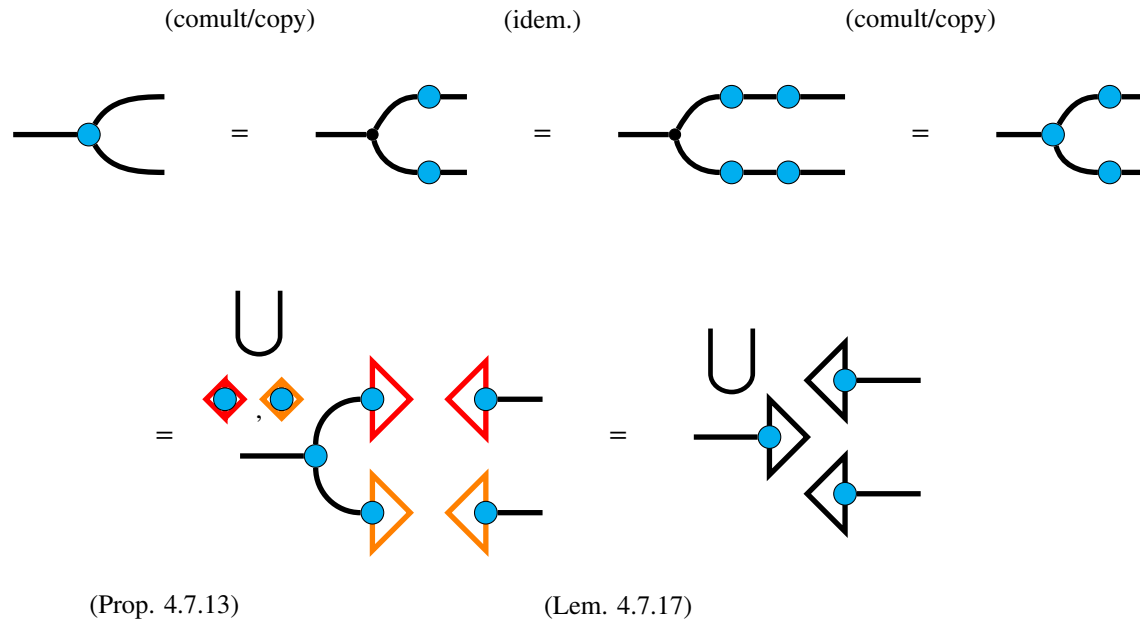
$$\begin{array}{ccccccc}
 \text{multiplication} & = & \text{e-spider} & = & \text{Prop. 4.7.13} & = & \text{Lem. 4.7.15} \\
 \text{multiplication} & = & \text{e-spider} & = & \text{Prop. 4.7.13} & = & \text{Lem. 4.7.15}
 \end{array}$$

□

Proposition 4.7.21 (*e*-copiable decomposition of comultiplication).



Proof.



□

NOW WE CAN PROVE THEOREM 4.7.8.

Proof. The key observation is that the *e*-copiable decomposition of the idempotent given by Proposition 4.7.13 is equivalent to a split idempotent though the set of *e*-copiables equipped with discrete topology.

placeholder

□

4.8 *Sticky spiders are collections of shapes in space*

...

4.9 *Displacing shapes with symmetries*

Definition 4.9.1 (Topological Group).

Definition 4.9.2 (Sticky spiders equivalent up to symmetric displacement).

4.10 *Moving shapes continuously*

4.11 *Configuration space of a sticky spider*

Definition 4.11.1 (The configuration space of a sticky spider).

Definition 4.11.2 (The "interaction" relation).

Definition 4.11.3 (Movement equivalence classes).

4.12 *Topological models for a selection of concepts*

Construction 4.12.1 (Putting-in and getting-out things from containers).

Construction 4.12.2 (Filling holes, fullness, and fitting).

Construction 4.12.3 (Sending by throwing and conduits).

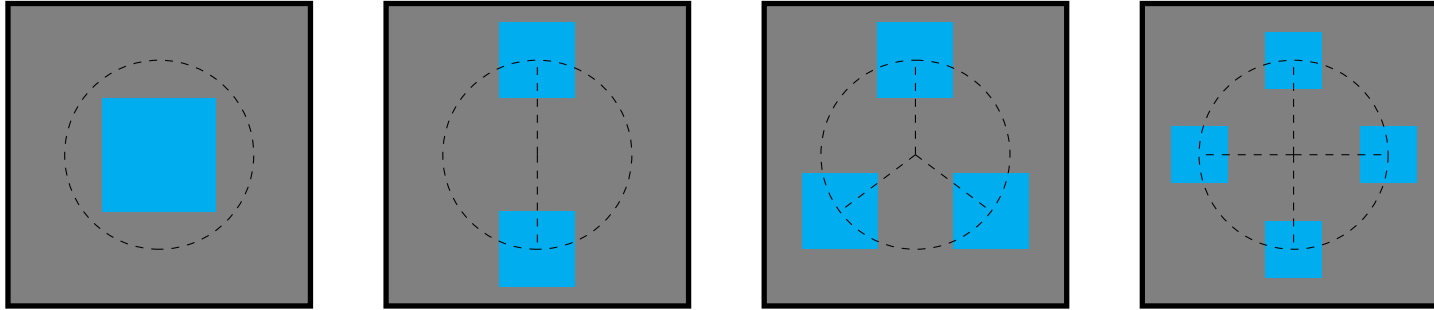
Construction 4.12.4 (Conduit Metaphor).

4.13 *Modelling Lassos*

Recall that Lassos – a graphical gadget that can encode arbitrary morphisms into a single wire – can be interpreted in a monoidal computer. Recall that monoidal computers require a universal object Ξ . Here we show how in **TopRel**, by taking $\Xi := \begin{smallmatrix} \blacksquare \\ \blacksquare \end{smallmatrix}$ the open unit square, we have a monoidal computer in **Rel** restricted to countable sets and the relations between them. We will make use of sticky spiders. We have to show that; $\begin{smallmatrix} \blacksquare \\ \blacksquare \end{smallmatrix}$ has a sticky-spider corresponding to every countable set; how there is a suitable notion of sticky-spider morphism to establish a correspondence with relations; what the continuous relations are on $\begin{smallmatrix} \blacksquare \\ \blacksquare \end{smallmatrix}$ that mimic various compositions of relations.

Proposition 4.13.1 ($((0, 1) \times (0, 1)$ splits through any countable set X). For any countable set X , the open unit square $\begin{smallmatrix} \blacksquare \\ \blacksquare \end{smallmatrix}$ has a sticky spider that splits through X^* .

Proof. The proof is by construction. We'll assume the sticky-spiders to be mereologies, so that cores and halos agree. Then we only have to highlight the copiable open sets. Take some circle and place axis-aligned open squares evenly along them, one for each element of X . The centres of the open squares lie on the circumference of the circle, and we may shrink each square as needed to fit all of them.



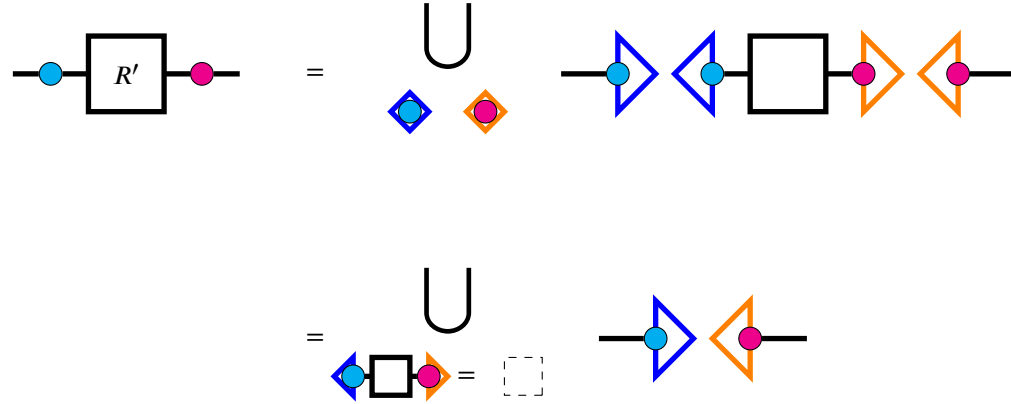
□

Definition 4.13.2 (Morphism of sticky spiders). A morphism between sticky spiders is any morphism that satisfies the following equation.

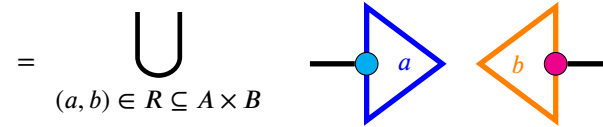
Proposition 4.13.3 (Morphisms of sticky spiders encode relations). For arbitrary split idempotents through A^\star and B^\star , the morphisms between the two resulting sticky spiders are in bijection with relations $R : A \rightarrow B$.

Proof.

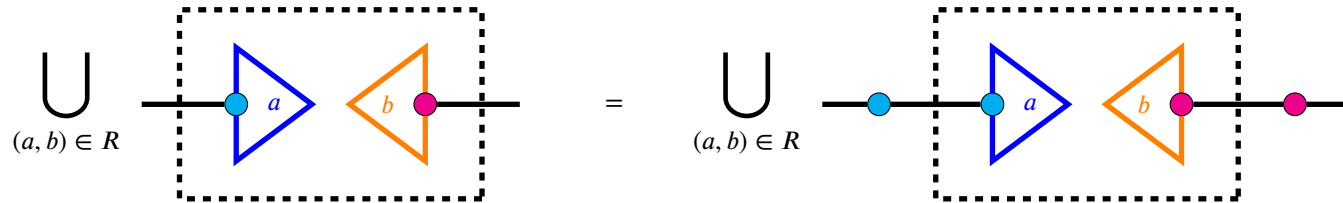
(\Leftarrow) : Every morphism of sticky spiders corresponds to a relation between sets.



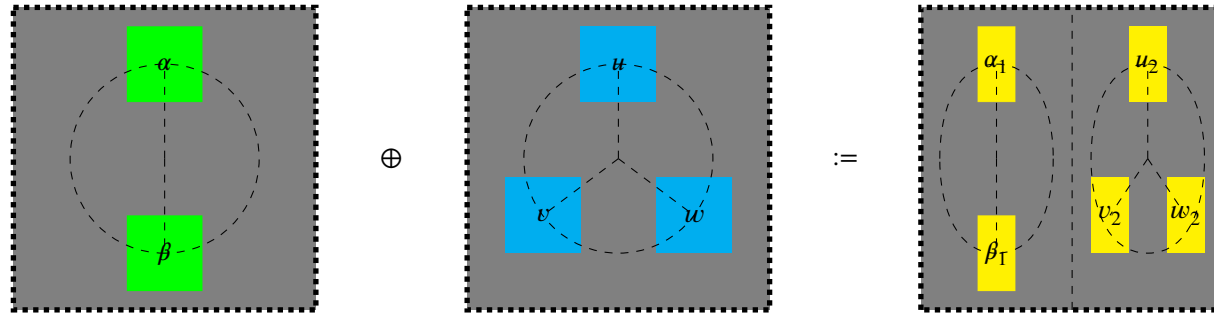
Since (co)copiables are distinct, we may uniquely reindex as:



(\Rightarrow) : By idempotence of (co)copiables, every relation $R \subseteq A \times B$ corresponds to a morphism of sticky spiders.



Construction 4.13.4 (Representing sets in their various guises within \mathbb{R}). We can represent the direct sum of two \mathbb{R} -representations of sets as follows.



The important bit of technology is the homeomorphism that losslessly squishes the whole unit square into one half of the unit square. The decompressions are partial continuous functions, with domain restricted to the appropriate half of the unit square.

$$\begin{array}{cccc}
 \text{---} \text{---} \text{---} & \text{---} \text{---} \text{---} & \text{---} \text{---} \text{---} & \text{---} \text{---} \text{---} \\
 (x, y) \mapsto (\frac{x}{2}, y) & (x, y) \mapsto (\frac{x+1}{2}, y) & (x, y)|_{x < \frac{1}{2}} \mapsto (2x, y) & (x, y)|_{x > \frac{1}{2}} \mapsto (2x - 1, y)
 \end{array}$$

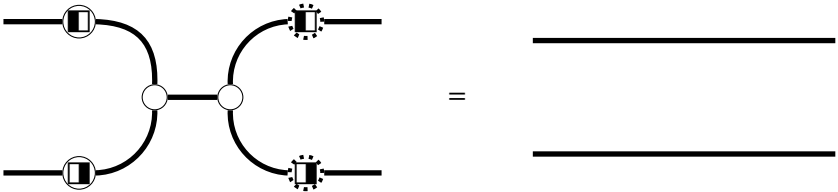
We express the ability of these relations to encode and decode the unit square in just either half by the following graphical equations.

$$\text{---} \text{---} \text{---} = \text{---} = \text{---} \text{---} \text{---}$$

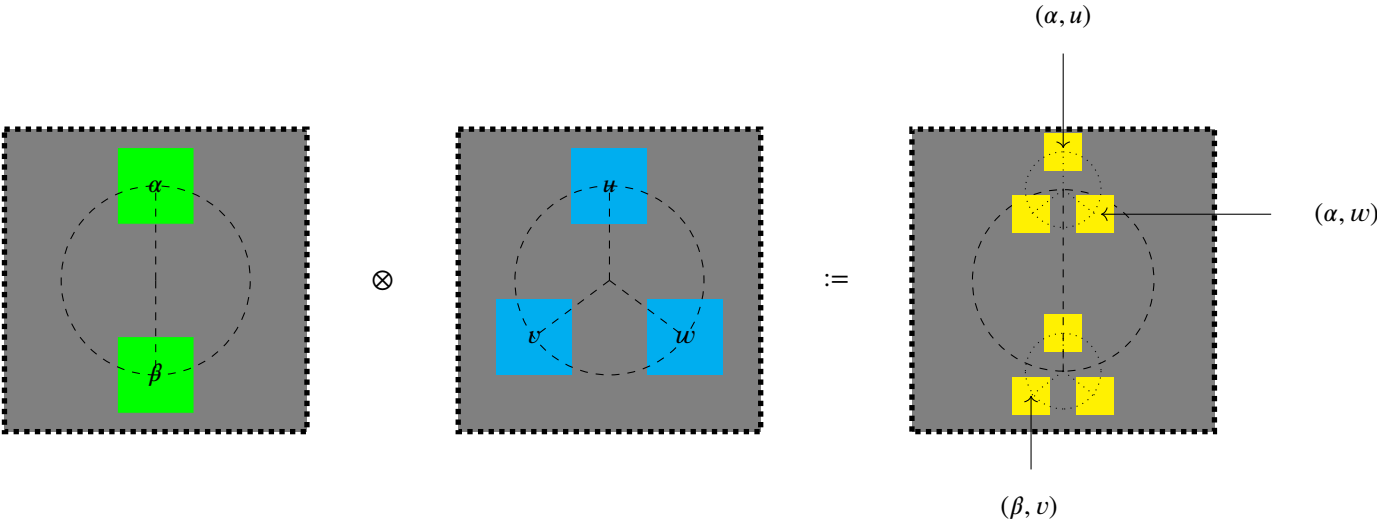
Now, to put the two halves together and to take them apart, we introduce the following two relations. In tandem with the squishing and stretching we have defined, these will behave just as the projections and injections for the direct-sum biproduct in **Rel**.

$$\begin{array}{ccc}
 \text{---} \text{---} \text{---} & := & \text{---} \text{---} \text{---} \cup \text{---} \text{---} \text{---} \\
 \text{---} \text{---} \text{---} & := & \text{---} \text{---} \text{---} \cup \text{---} \text{---} \text{---}
 \end{array}$$

The following equation tells us that we can take any two representations in \mathbb{B} , put them into a single copy of \mathbb{B} , and take them out again. Banach and Tarski would approve.

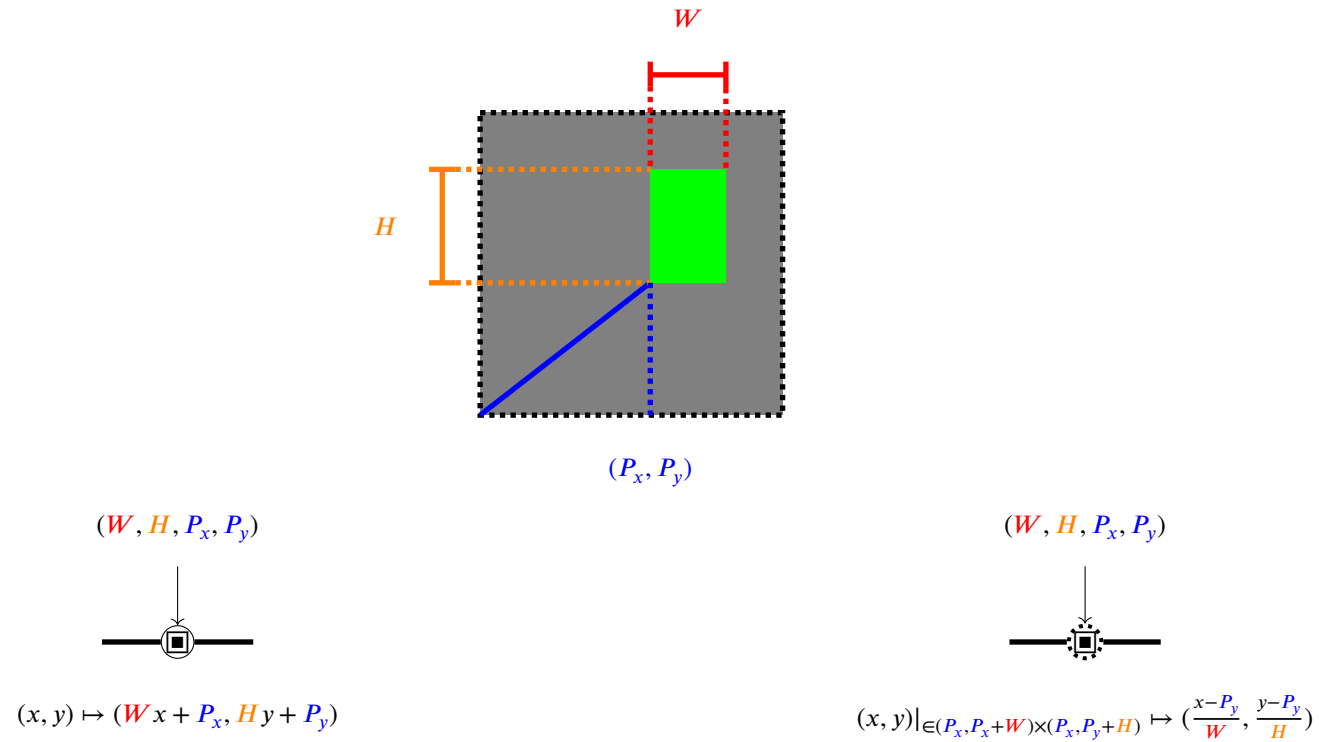


We encode the tensor product $A \otimes B$ of representations by placing copies of B in each of the open boxes of A .



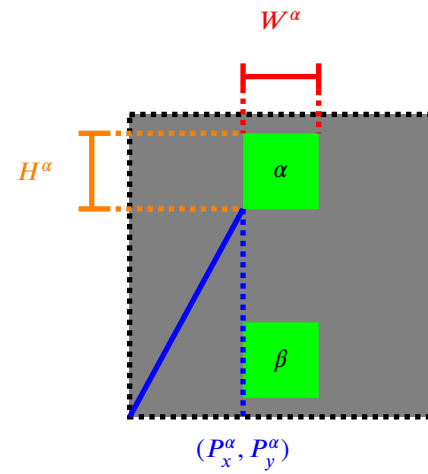
The important bit of technology here is a family of homeomorphisms of \mathbb{B} parameterised by axis-aligned open boxes. We depict the parameters outside the body of the homeomor-

phism for clarity. The squish is on the left, the stretch on the right.



Now, for every representation of a set in \mathbb{B} by a sticky-spider, where each element corresponds to an axis-aligned open box, we can associate each element with a squish-stretch

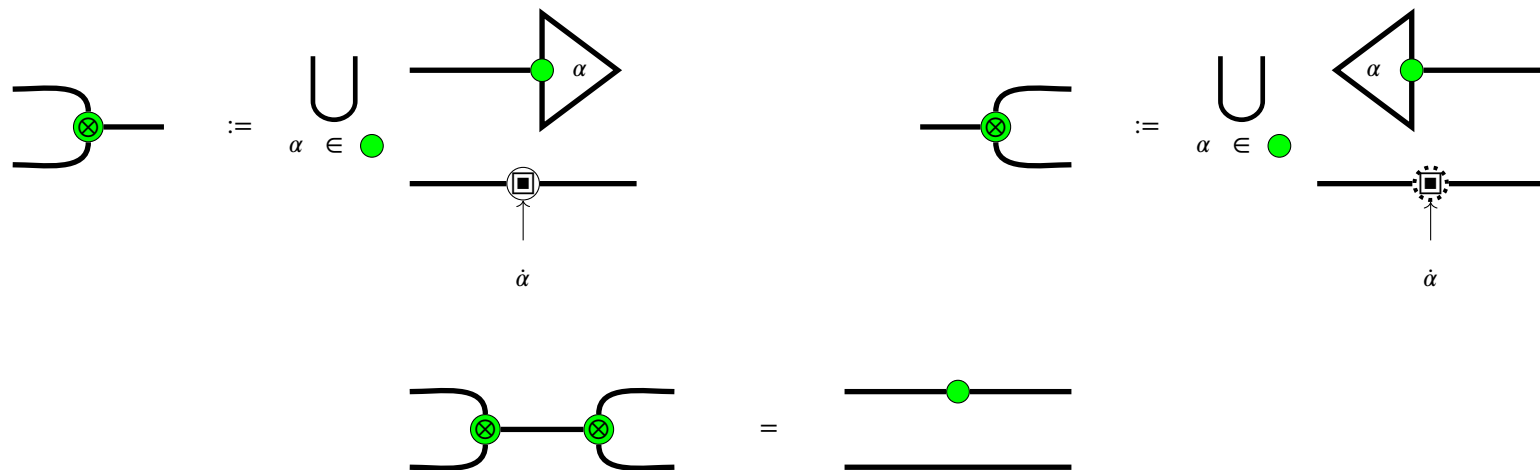
homeomorphism via the parameters of the open box, which we notate with a dot above the name of the element.



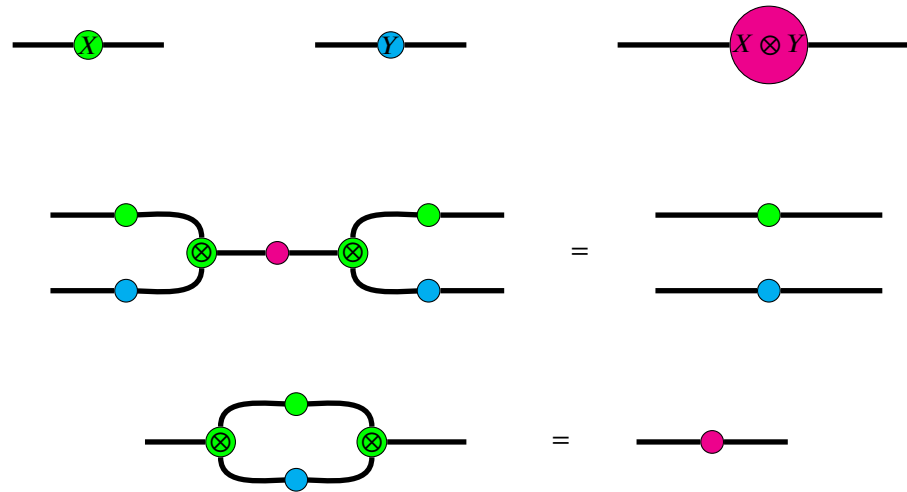
$$\dot{\alpha} = (W^\alpha, H^\alpha, P_x^\alpha, P_y^\alpha)$$

$$\dot{\beta} = (W^\beta, H^\beta, P_x^\beta, P_y^\beta)$$

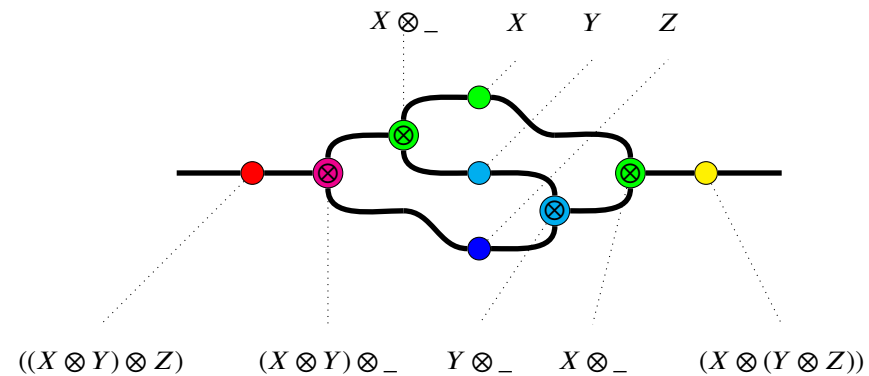
Now we can define the "tensor X on the left" relation $_ \rightarrow X \otimes _$ and its corresponding cotensor.



The tensor and cotensor behave as we expect from proof nets for monoidal categories.

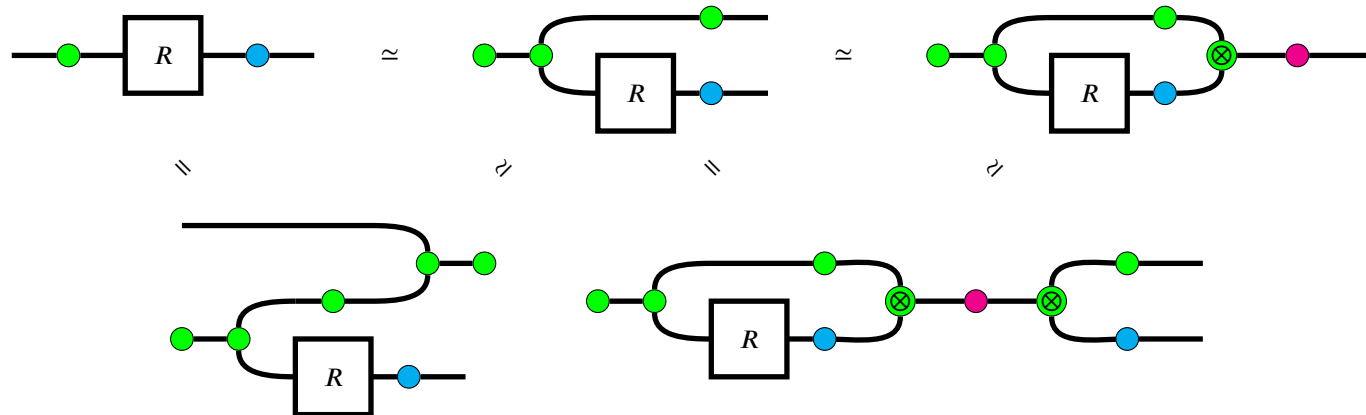


And by construction, the (co)tensors and (co)pluses interact as we expect, and they come with all the natural isomorphisms between representations we expect. For example, below we exhibit an explicit associator natural isomorphism between representations.

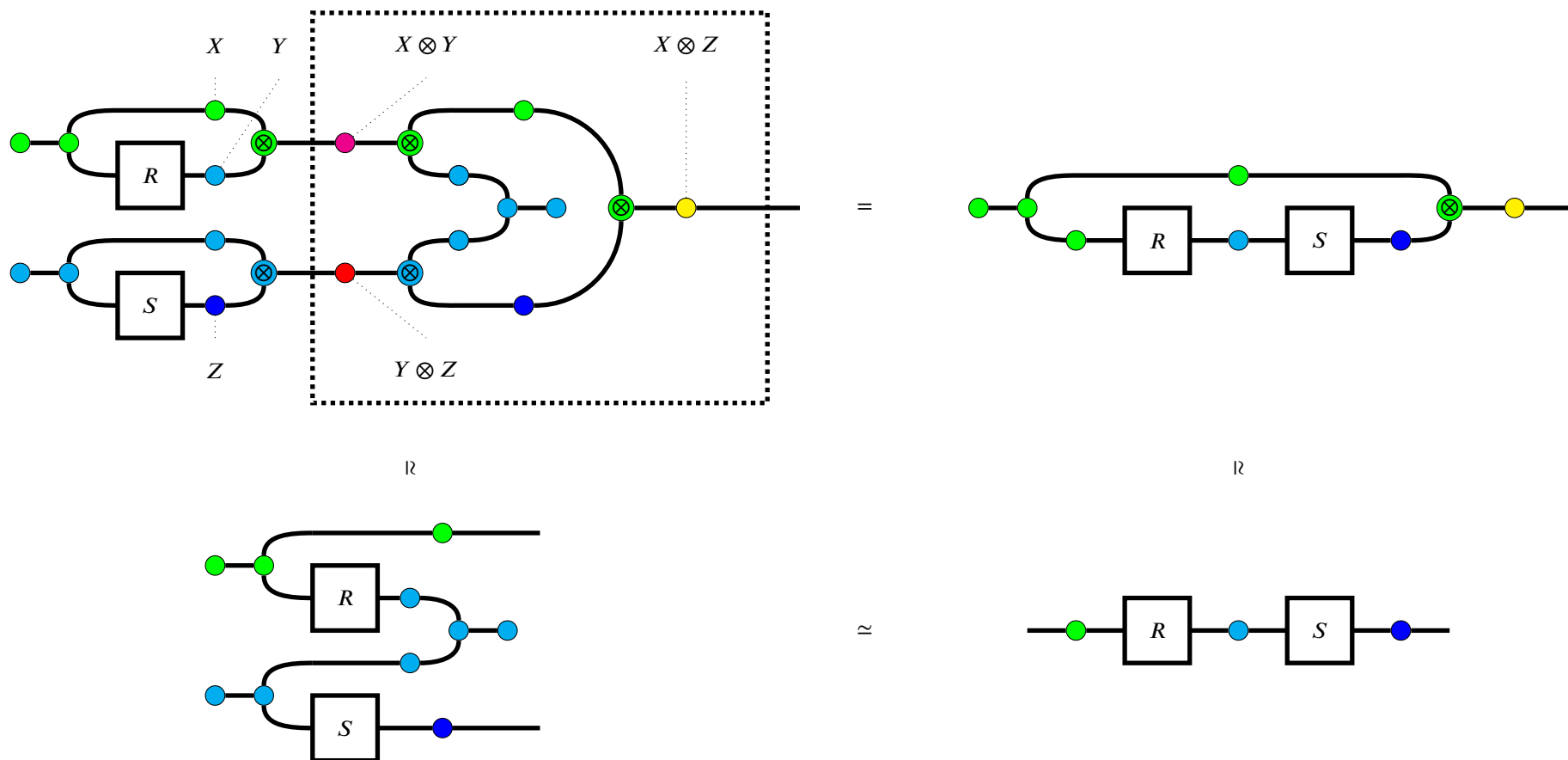


Construction 4.13.5 (Representing relations between sets and their composition within \mathbb{M}). With all the above, we can establish a special kind of process-state duality; relations as

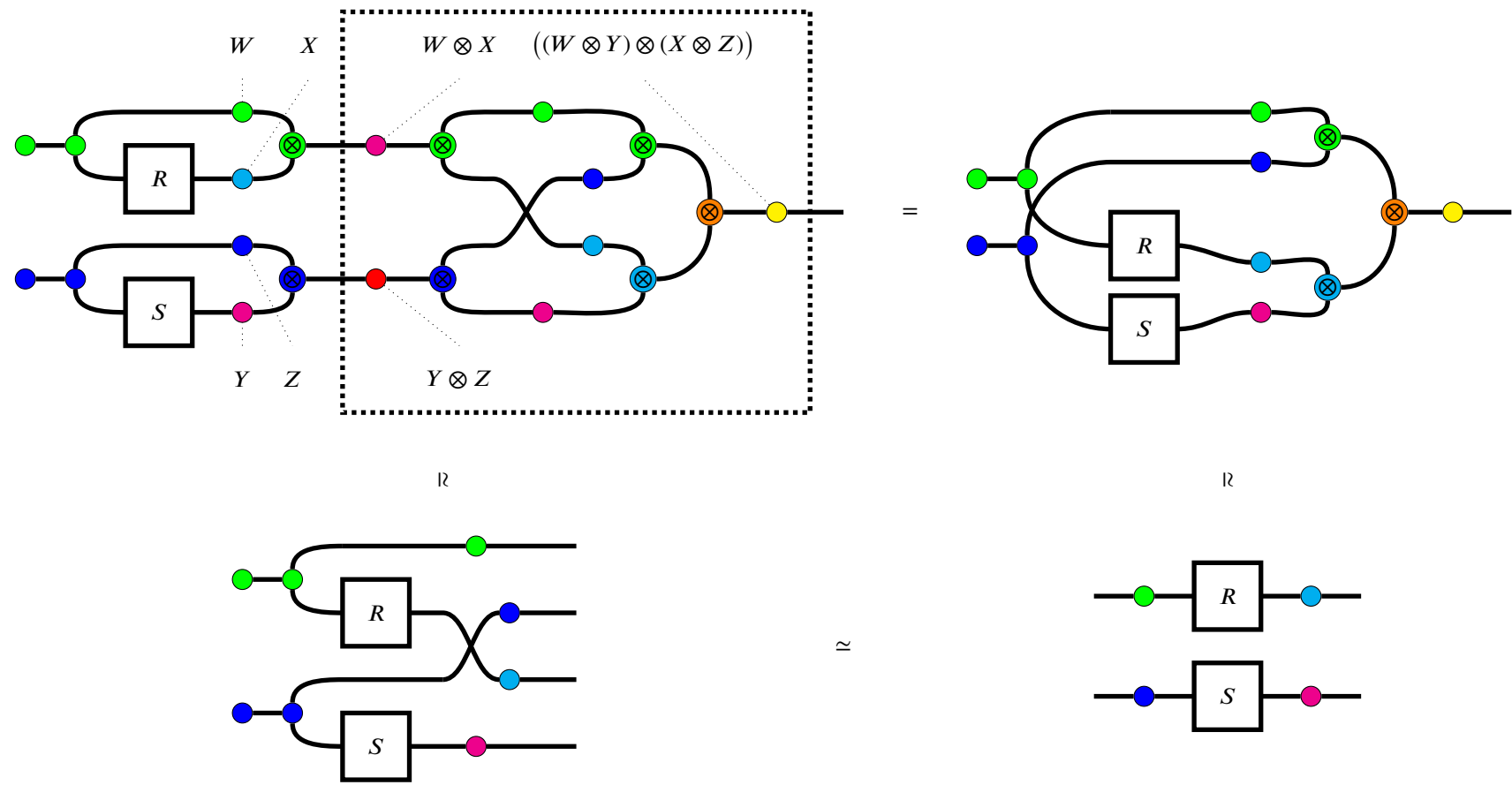
processes are isomorphic to states of \mathbb{M} , up to the representation scheme we have chosen.



Moreover, we have continuous relations that perform sequential composition of relations.



And we also know how to take the parallel composition of relations by tensors.



5

Things to do with formal depictions of language

5.1 Modelling metaphor

I will take a *metaphor* to be text where systematic language for one kind of concept is used to structure meanings for another kind of concept where literal meanings cannot apply. This may subsume some cases of what would otherwise be called *similes* or *analogies*.

THE IDEA: First, we observe that we can model certain kinds of analogies between conceptual spaces by considering structure-preserving maps between them. For example, Planck's law gives a partial continuous function from part of the positive reals measuring Kelvin to wavelengths of light, and the restriction of this mapping to the visible spectrum gives the so-called "colour temperature" framework used by (image-professionals?).

Second, we observe that we can use simple natural language to describe conceptual spaces, instead of geometric or topological models. Back to the example of colour temperature, instead of precise values in Kelvin, we may instead speak of landmark regions that represent both temperature and colour such as "candle", "incandescent", "daylight", which obey both temperature-relations (e.g. candle is a lower temperature than incandescent) and colour-relations (daylight is bluer than incandescent).

Third, we observe that we can describe complex conceptual schemes using simple natural language, and formally specify what it means for one conceptual scheme to structure another by describing structure-preserving maps between the text circuits of the descriptions. This will allow us to reason about and calculate with complex metaphors, such as "Time is Money".

THE MOTIVATION:

5.1.1 Orders, Temperature, Colour, Mood

5.1.2 Complex conceptual structure

TIME IS MONEY

"Do you have time to look at this?"

"Do you think this is worth the time?"

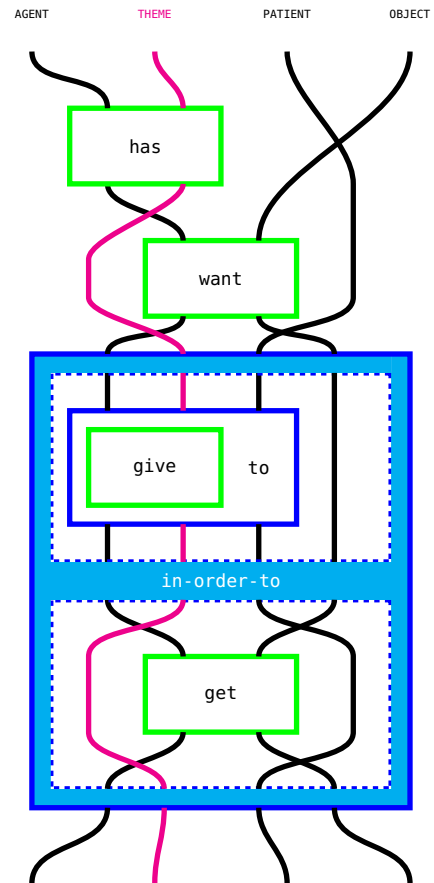
"What a waste of time."

I will work through an example of partially using the concept of Money to structure that of Time. Part of the concept of Money is that it can be *exchanged* for something. The concept of exchange can be glossed approximately as the following text, with variable noun-entries capitalised.

AGENT has THEME.

AGENT wants OBJECT.

AGENT gives THEME to PATIENT in-order-to get OBJECT.



The

5.1.3