UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI MATEMATICA "TULLIO LEVI–CIVITA"

CORSO DI LAUREA MAGISTRALE IN MATEMATICA

Internal mathematics for stochastic calculus: a tripos-theoretic approach

TESI DI LAUREA MAGISTRALE

RELATORE
Dott. Samuele Maschio

LAUREANDO
Dott. Matteo Capucci (matr. 1210253)

17 luglio 2020 **a.a. 2019–2020**

Contents

Introduction

This thesis was born as a tentative marriage of stochastic calculus and topos theory, and developed into a ramified investigation on the logical aspects of the elementary structures on which the former rests.

Work at the intersection between category theory and probability theory has been done for a long time [Lawvere 1962; Giry 1982], though we are currently witnessing surging interest in these questions, under the flagship of **categorical probability theory**. Much has been done regarding foundational aspects of probability theory, either by trying to develop a form of 'synthetic probability theory' [Simpson 2017; Fritz 2019] or by investigating relevant categorical structures for probability theory [Fritz and Perrone 2019; Simpson 2012; Heunen et al. 2017].

We tried instead to apply the technique of **internalization**. This method simplifies a mathematical theory by offsetting some of its complexity to the objects and arrows of a suitably constructed topos. In fact, a topos $\mathcal{E}$ looks very much 'set-like' from the internal point of view, i.e. if we adopt a suitable language interpreted in objects and arrows of $\mathcal{E}$. Everything one can express in higher-order intuitionistic logic, one can do it in a topos as if it was the usual world of sets.

In fact, if we fix a topos $\mathcal{E}$, there are now two levels of 'mathematics' relative to $\mathcal{E}$: the 'external' one (on par with $\mathcal{E}$ and expressed using the metalanguage), and the 'internal' one (expressed using the internal language of $\mathcal{E}$). An automatic procedure can convert any sentence or proof in the internal world into a sentence or proof in the external world. In doing so, all the complexity hidden in the objects and arrows of $\mathcal{E}$ resurfaces, and thus simple constructions and proofs are inflated to more convoluted ones, about more complex objects. Borrowing an example from [Blechschmidt 2017], a commutative ring defined in the sheaf topos of a topological site, once externalized becomes a sheaf of commutative rings on that site.

The idea is to take advantage of the opposite procedure: given a complex theory $\mathbb{T}$

in the external world, we could try to offset its complexity by constructing an ad-hoc topos $\mathcal{E}$ and a theory $\mathbb{T}'$ expressible in the language of $\mathcal{E}$ such that the externalization of $\mathbb{T}'$ is the theory $\mathbb{T}$ we started with. If we succeed to do this, then every (intuitionistically valid) proof we do in $\mathbb{T}'$ can be translated automatically to a proof of $\mathbb{T}$:

$$\begin{array}{ccc} \text{hard } P \text{ about } \mathbb{T} & \xrightarrow{\text{to } \mathcal{E}} & \text{simpler } P' \text{ about } \mathbb{T}' \\ \vdots & & \downarrow \\ \text{proof of } P \text{ in } \mathbb{T} & \xleftarrow[\text{to } \mathbf{Set}]{} & \text{proof of } P' \text{ in } \mathbb{T}' \end{array}$$

This idea is used, for example, in [Blechschmidt 2017] to investigate abelian sheaves and thus algebraic geometry, reducing many tedious proofs to simpler and more conceptual ones in commutative or even linear algebra.

In this thesis, we're trying to attack stochastic calculus with this technique. The question we're trying to answer is: **how much of stochastic calculus is 'stochastic' and how much is 'calculus'?** Can we simplify the theory by hiding some of the stochasticity in a suitable 'topos of stochastic sets'?

To understand how this could be done, let us briefly recall what the central object of stochastic calculus is: stochastic processes. In their most basic forms, they are collections of random variables $X = \{X_t : \underline{\Omega} \to \mathbb{R}^n\}_{t \in I}$, where I is a suitable set of 'times' (usually $\mathbb{N}$ or $\mathbb{R}^+$). Alternatively, they can be seen as mappings $X : \underline{\Omega} \to (\mathbb{R}^n)^I$ or $\underline{\Omega} \times I \to \mathbb{R}^n$. Intuitively, X should represent a 'random map' $I \to \mathbb{R}^n$. The picture is vastly complicated by the fact each X_t is often required to satisfy a different measurability condition, corresponding to the fact $\underline{\Omega}$ is not actually fixed but 'grows with time' (one says the process is **adapted**).

The first attempt we made at modelling this situation was to construct a category of random mappings 'with context' by using the machinery of probability monads, the Kleisli construction and *generalized lens categories*[1] [Spivak 2019]. A category of random mappings is what is proposed by [Lawvere 1962], as the Kleisli category of a suitable monad on the category of measurable spaces, namely the one assigning to a space its set of probability measures (this is discussed in more detail in Chapter 2).

However, this line of attack didn't prove successful, and it seemed hopeless to capture adapted processes through this lens.

Therefore we tried to use an old construction by Scott originally emerged in the setting of forcing in set theory [Scott 1967]. Essentially, he considered a Boolean-valued

[1]Generalized lens categories are basically a different point of view on the Grothendieck construction.

model of set theory where the Boolean algebra of truth values is the complete Boolean algebra one gets by quotienting a σ-field by its ideal of meager sets.

The idea of Scott can be recast in modern language using the categorical device of triposes. Triposes are a categorical gadget which encode intuitionistic higher-order theories as fibrations over a base category (whose objects will become the types of the language), satisfying some coherence properties. Intuitively, the fiber over an object X corresponds to the Heyting algebra of predicates over the type X.

Then, from a tripos, one can easily construct a topos by simply building sets in the logic encoded by it. In our particular case, the result is a topos (which we named after Scott) in which truth values are 'essential events' of the base space $\underline{\Omega}$. A proposition is globally true if it holds with probability one. The internal objects of natural and real numbers are natural and real valued random variables on $\underline{\Omega}$, respectively. To us, this seems a good start for an internalization procedure.

The benefit of the tripos approach is that the logical content of the construction is much more vivid. One directly encodes in a tripos what it means to have propositions valued in events of Ω, and this carries over to the associated topos. We consider this tool a powerful weapon in the arsenal of 'internalizing mathematicians', as they allow a rather explicit control of the logical properties of the target topos.

To get to the point at which internalization seems within arm's reach took us enough time to prevent us from actually doing any non-trivial internal mathematics. Nonetheless, we consider our work valuable for two reasons:

1. As already said, it laid the groundwork for further investigations in the internal mathematics of Scott topoi and their possible use in the conceptual simplification of probability theory and stochastic calculus. In particular, we are confident many arguments and 'basic' techniques in those areas can be given a compelling logical interpretation in Scott topoi, thereby providing cleaner and more conceptual proofs and constructions.

2. It uncovered (at least to the author) many links between stochastic calculus (and, more generally, the study of non-deterministic dynamical systems) and modal logic, which might be fruitful to both areas. Reflections upon this conclude this work, in Appendix A.

Contents

This master thesis consists of five chapters. The first two provide preliminary notions. In Chapter 1 we laid the groundwork for later chapters by providing basic facts about some algebraic structures for logic (mainly various flavours of lattices) and discussing the kind of formal languages we are going to employ later, when discussing triposes. Chapter 2 instead is a quick tour of measure theory and stochastic calculus, with some categorical remarks.

Chapter 3 is a brief introduction to tripos theory. We give a detailed account of the semantics of triposes, and their relationship with the topoi one constructs from them. It should provide the reader with 'off-the-shelf' tools to interpret theories in a tripos and its associated topos.

Chapter 4 discusses the so-called Scott triposes and their associated topoi. We give their construction and analyze some properties. Functoriality of the construction is discussed, and we end the chapter with some elementary examples of internalization, plus a tentative internalization of stochastic processes in a Scott topos.

Chapter 5 explores the logical features of filtered probability spaces. We observe there is a natural set of modalities one can adjoin to the internal language of the Scott topos of $\underline{\Omega}$ in order to be able to address the 'dynamic' aspect of its field of events. The chapter ends with a discussion of some ideas regarding how to use these modal operators to internalize measurability criteria.

Finally, Appendix A compares the modalities introduced in Chapter 5 to those appearing in known *process logics*, and conclude they have a very different meaning, which, however, we had no time to investigate further. We also discuss a possible link between the semantics of process logics and filtered probability spaces.

Acknowledgements

> O Captain! my Captain! our fearful trip is done,
> The ship has weather'd every rack, the prize we sought is won,
> The port is near, the bells I hear, the people all exulting
>
> Walt Whitman

By this time in my mathematical career, it has become clear mathematics is an essentially social endeavour. It is not possible to do mathematics alone, and it doesn't even make sense to do so: the very concept of proof entails having one human explaining something to another. This is true whether we adopt a Platonist view on mathematics, and it's even more true if we don't. Therefore it seems necessary to acknowledge here the contribution of my social entourage, whose influence on me is directly expressed in this work.

First and foremost, my family deserves the utmost gratitude for enabling me to undertake this long journey. They have been the wind in my sails, for longer than I can remember. Their sacrifice and love will inspire me for the rest of my life.

Second, my friends. They have been the crew escorting me in this perilous journey. Their sole presence eased the fatigue of it all, warmed my heart and filled my life with enthusiasm and joy, lifting the heavy weight mathematics sometimes can be.

Special mentions are due.

For the last year and a half, *Corollario* has been a major presence in my Paduan life: Nunzio, Martina, Maria, Isabella and many more than I can list here. From them I learned much more than singing (a skill which remains largely unknown to my body). I've learned the power of people united by passion. I've learned lightness of heart can be combined with the utmost commitment and passion. I will miss them dearly.

Davide, Francesco, Pietro, Valentina and the rest of my 'Paduan' friends were also immensely important for my time here. I remember the time spent with them as exciting and fulfilling, and conversations with them stimulating and enriching. I know this has been the beginning of a long-lasting friendship.

Not only my 'new' friends deserves an accolade: longtime friends have been the solid rock on which I could set ashore everytime I met them again, despite spending so much time far from them. All my friends from BVA, my flatmates, my ex university and high school classmates and many more (among them, I can't resist to name Linda, Paolo, Astrid, Sylwia). I apologize: the time and energy I devoted to them in the last two years wasn't always enough to fully enjoy their company and truly be each other's mates. What we had together, I will treasure for life.

Finally, I had the fortune to be welcomed in a great community this year, which makes me enthusiastic for my future. I have to thank Fosco, Fabrizio, and the rest of the people from ItaCa, as well as the wider ACT community. Your selfless engagement with my rookie excitement has been the America in which I didn't even hope to run into.

Last but not least, my supervisor, prof. Maschio, deserves all of my thankfulness. He gladly embarked with me in the last, foolish leg of my journey, helping me navigating an ocean I would have never tamed otherwise. His enthusiasm and trust in my work have been a great opportunity, and a major source of motivation. I hope our mathematical activity will cross again in the coming years.

There are many other people with whom I crossed paths in a fruitful way, many more than I can recall at the moment. Each of them brought me where I am now, therefore my gratitude goes to them as well.

Matteo Capucci
Padova, 2020-07-17

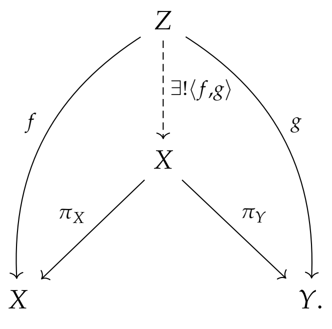
Notation

Category theory. Let $\mathbf{C}$ be a category. We always consider our categories to be locally small, but not always small. We use the notation $X : \mathbf{C}$ to say X is an object of $\mathbf{C}$. The set of morphisms from an object $X : \mathbf{C}$ and an object $Y : \mathbf{C}$ is denoted by $\mathbf{C}(X, Y)$. The identity morphism of X is denoted by 1_X. Composition of $X \xrightarrow{f} Y$ and $Y \xrightarrow{g} Z$ is denoted by $g \circ f$, or simply by juxtaposition (gf) if there's no risk of ambiguity.

If $\mathbf{C}$ has a terminal object, we denote it by 1 and we denote by $!_X$ the unique morphism $X \to 1$. Dually, we denote initial objects by 0 and we call $?_X$ the unique morphism $0 \to X$.

If $\mathbf{C}$ has finite products, we denote by Δ_X the **diagonal morphism** $X \rightarrowtail X \times X$, defined as the equalizer of the two projections $X \times X \rightrightarrows X$. If a subobject classifier is available in $\mathbf{C}$, the characteristic morphism of Δ_X will be denoted by δ_X.

If $\mathbf{C}$ has products, and $Z \xrightarrow{f} X$ and $Z \xrightarrow{g} Y$ are morphisms, we denote by $\langle f, g \rangle$ the unique morphism $Z \to X \times Y$ which makes the relevant diagram commute. Moreover, we use π_X and π_Y for the respective projections.



An analogous notation applies to products of more than two factors. If instead we have two morphisms $X \xrightarrow{f} Y$ and $Z \xrightarrow{g} W$, we denote by $f \times g$ the morphism $\langle f\pi_X, g\pi_Z \rangle$: $X \times Z \to Y \times W$, where π_X and π_Z are projections out of $X \times Z$ and the universal property invoked is the one of $Y \times W$.

If $\mathbf{C}$ has exponentials, then the exponential of Y by X is denoted by Y^X, and the evaluation map $Y^X \times X \to Y$ by ev_{Y^X}. The transpose $Z \to Y^X$ of a map $X \times Z \xrightarrow{f} Y$ is denoted by $\lambda_X f$.

If **C** has a subobject classifier, we denote it by $1 \xrightarrow{\mathsf{true}} \Omega$. The characteristic map of a subobject $U \overset{m}{\rightarrowtail} X$ will be denoted by χ_m or, if it's clear to which mono $U \to X$ we are referring to, simply χ_U.

Finally, we'll denote by よ (read 'yo') and ょ the Yoneda and co-Yoneda embeddings, respectively:

$$\mathbf{C} \xhookrightarrow{\;よ\;} [\mathbf{C}^{\mathsf{op}}, \mathbf{Set}] \qquad\qquad \mathbf{C}^{\mathsf{op}} \xhookrightarrow{\;ょ\;} [\mathbf{C}, \mathbf{Set}].$$

Logic. Throughout the thesis, we adopt classical Zermelo-Fraenkel set theory with axiom of choice (ZFC) as our metatheory.

When treating a type theory, we denote typing judgments using a colon: $t : X$. The context of a term or formula is put between square brackets at the end. If $t : Y\ [x : X]$ and $s : Z\ [y : Y]$ are terms, we denote the term of type Z obtained by substituting each occurence of y in s with t by $s[y/t] : Z\ [x : X]$.

Ordered sets. We'll often deal with totally ordered sets, such as $\mathbb{N}$ or $\mathbb{R}$. In that context, if $(I, \leq)$ is the total order and $t \in I$, then we employ the following notations:

$$t\!\uparrow = \{s \in I \mid t \leq s\}, \qquad t\!\upharpoonleft = t\!\uparrow \setminus \{t\},$$
$$t\!\downarrow = \{s \in I \mid s \leq t\}, \qquad t\!\downharpoonleft = t\!\downarrow \setminus \{t\}.$$

Miscellanea. Sometimes, to denote a structured set $(X, \text{structure})$, we underline it: $\underline{X}$, to distinguish it from the carrier set X.

Chapter 1

Preliminaries on logic

We recall some facts about the algebraic and logical structures we are going to use in the subsequent chapters.

1.1 Algebraic structures

Definition 1.1. An **Heyting algebra** is a poset $(\mathbb{H}, \leq)$ which admits all finite joins and meets, and moreover has a binary operation $\Rightarrow$ satisfying

$$a \wedge b \leq c \quad \underline{\text{iff}} \quad a \leq b \Rightarrow c.$$

This definition implies some good properties: the poset $\mathbb{H}$ is bounded, meaning there are a top ($\top$) and bottom ($\bot$) element, and it is a distributive lattice, meaning:

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c).$$

Using categorical lingo, a Heyting algebra is a posetal, finitely cocomplete and cartesian closed category. A Heyting algebra is **complete** if it is bicomplete in the categorical sense, hence if it admits arbitrary joins and meets (i.e. indexed by any set). Moreover, in this case infinite distributivity is automatic:

$$a \wedge \bigvee_{i \in I} b_i = \bigvee_{i \in I} (a \wedge b_i).$$

Logical meaning. A Heyting algebra is the algebraic incarnation of intuitionistic propositional logic. In fact the operators $\wedge$, $\vee$ and $\Rightarrow$ behave exactly as intuitionistic conjunction, disjunction and implication do, respectively, while the order structure captures entailment. Moreover, we can define

$$\neg a \equiv a \Rightarrow \bot$$

and see this too satisfies the required properties, including the possible failure of the law of excluded middle. In fact, in a Heyting algebra it is not necessarily the case that $\neg a \vee a = \top$, while $a \leq \neg\neg a$ and $a \wedge \neg a = \bot$ hold.

More rigorously speaking, one can show the set of propositions (up to logical equivalence, see later) of an intuitionistic propositional theory, ordered by logical entailment, is a Heyting algebra, whose meet, join and implication operator correspond to their logical equivalents. This algebra is called the Lindenbaum–Tarski algebra of the theory.

The equivalent algebraic structure for classical logic is now easy to guess:

Definition 1.2. A **Boolean algebra** is a Heyting algebra where

$$\neg\neg a = a.$$

In categorical lingo, a Boolean algebra is a $*$-autonomous Heyting algebra, whose dualizing object is $\bot$. Indeed, unpacking the additional axiom of Boolean algebras we get

$$a = (a \Rightarrow \bot) \Rightarrow \bot.$$

Internal Heyting algebras. We can formulate the definition of Heyting algebra in any sufficiently structured category, simply by requiring the axioms of Heyting algebras hold for generalized elements

Definition 1.3. Let **C** be a category with pullbacks. An **internal Heyting algebra** in **C** is an object $H : \mathbf{C}$ together with a subobject

$$O \overset{\leq}{\rightarrowtail} H \times H,$$

and three arrows

$$H \times H \xrightarrow{\wedge} H$$
$$H \times H \xrightarrow{\vee} H$$
$$H \times H \xrightarrow{\Rightarrow} H$$

such that, for any object $X : \mathbf{C}$,

$$(よ H(X), (よ \leq)_X)$$

is a Heyting algebra with respect to the operations $(よ\wedge)_X$, $(よ\vee)_X$, $(よ\Rightarrow)_X$.

Recall that $\text{よ}H(X)$ is the set of generalized element of H at stage X, hence the hom-set $\mathbf{C}(X, H)$. From this point of view, the definition of internal Heyting algebra prescribes the generalized elements of H to form a Heyting algebra at every stage of definition. Thus if $X \overset{a}{\to} H$ and $X \overset{b}{\to} H$ are generalized elements of $(H, \leq)$, we can say things like

$$a \wedge b \leq a, \quad b \leq a \vee b, \ldots$$

Prealgebras. As we've seen above, Lindenbaum–Tarski algebras become Heyting algebras only after quotienting by logical equivalence, otherwise entailment of propositions is not a partial order, but only a preorder. Nevertheless, we can turn to a slightly more lax version of Heyting and Boolean algebras:

Definition 1.4. A **Heyting prealgebra** is a preorder whose posetal reflection[1] is a Heyting algebra.

Likewise, we define Boolean prealgebras as those Heyting prealgebras whose posetal reflection is Boolean.

In a Heyting or Boolean prealgebra, all the objects singled out by an universal property are no more uniquely defined, but are defined only up to posetal equivalence. This fact brings the definition of join, meets and implication (and thus of top and bottom element as well) even more similar in spirit to their categorical generalizations. Unfortunately, this means definitions whose right hand side is an expression involving meets, joins or implications are not well-posed anymore, unless we assume there is a canonical choice of meet, join and implication between any two elements of the prealgebra. We'll employ this strategy silently in the rest of the thesis.

Morphisms. In order to pack up all these kinds of structures in full-blown categories, we must specify what are the morphisms between them.

Definition 1.5. A morphism of Heyting algebras is a monotone function preserving finite joins, finite meets and implication.

Boolean algebras have the same morphisms as Heyting algebras, as Boolean-ness is just a property, i.e. Boolean algebras have no additional structure for the morphisms to

[1] Given a preorder, its **posetal reflection** is the poset one gets by quotienting by **posetal equivalence**:

$$a \simeq b \quad \underline{\text{iff}} \quad a \leq b \text{ and } b \leq a.$$

preserve. This also means the category of Boolean algebras **Bool** is embedded in the category of Heyting algebras **Hey**.

Suitably relaxing the notion of 'preservation', we get also morphisms of Heyting prealgebras:

Definition 1.6. A morphism of Heyting prealgebras is a monotone function preserving joins, meets and implication *up to posetal equivalence*.

We denote the category of Heyting prealgebras as **PreHey**. Evidently, it sits inside the category **PreOrd** of preorders. Notice that **PreHey** and **Hey** also have a 2-categorical structure: a 2-cell from the 1-cell $\mathbb{H} \xrightarrow{f} \mathbb{H}'$ to the 1-cell $\mathbb{H} \xrightarrow{g} \mathbb{H}'$ is simply given by pointwise inequality:

$$f \leq g \;:\equiv\; f(a) \leq_{\mathbb{H}'} g(a), \text{ for all } a \in \mathbb{H}.$$

This allows us to speak of 2-categorical notions like equivalences and adjoint maps. The first are those maps such that $f \simeq g$ pointwise. The second are pairs

$$\mathbb{H} \underset{g}{\overset{f}{\rightleftarrows}} \mathbb{H}'$$

such that $1_{\mathbb{H}} \leq gf$ and $fg \leq 1_{\mathbb{H}'}$. In this case, we write $f \dashv g$.

Finally, as for any 2-category, it feels natural to consider suitably weakened versions of functors and natural transformations. A **pseudofunctor $\mathbf{C} \xrightarrow{F}$ PreOrd** is like a functor except functoriality is only satisfied up to equivalence, i.e. for all morphisms $A \xrightarrow{f} B \xrightarrow{g} C$ in $\mathbf{C}$, we have

$$F(g)F(f) \simeq F(gf), \quad F(1_A) \simeq 1_{F(A)}.$$

In the same spirit, a **pseudonatural transformation** is a transformation between parallel pseudofunctors satisfying naturality only up to equivalence.

1.2 Flavours of complete lattices

Locales, frames and complete Heyting algebras are three names for the same algebraic structure. Nonetheless, the difference in naming reflects a difference in the morphisms we consider between them [Picado and Pultr 2011, Section II.1]:

1. A morphism of frames is a (monotone) map preserving finite meets and arbitrary joins. Frames with these arrows form the category **Frm**.

2. A morphism of complete Heyting algebras is a morphism of frames which in addition preserves implication. Complete Heyting algebras with these arrows form the category **CHey**, which is a non-full subcategory of **Frm**.

3. A morphism of locales is a morphism of frames *in the opposite direction*, i.e. the category of locales **Loc** is formally $\mathbf{Frm}^{\mathrm{op}}$.

Complete Boolean algebras can also be seen as frames and locales. In particular, Boolean frames exhibit an interesting duality since if $\mathbb{B}$ is a Boolean frame, so is $\mathbb{B}^{\mathrm{op}}$. Thus a natural notion of morphism of Boolean frames is that of a morphism preserving both arbitrary meets and arbitrary joins. It turns out that the preservation of complements follows from these two, hence morphisms of Boolean frames and complete Boolean algebras coincide.

Moreover, complete Boolean algebras constitute a coreflective subcategory of **CHey**, where the coreflector is given by $\neg\neg$-*weakening*: it sends a Heyting algebra to its sub-Boolean algebra of fixed points of $\neg\neg$.

1.2.1 The adjoint functor theorem for posets

A pivotal theorem of category theory is Freyd's characterization of adjoint functors based on their continuity properties, i.e. their preservation of limits and colimits. Indeed, it is a known fact that left and right adjoints preserve colimits and limits, respectively, and *Freyd's Adjoint Functor Theorem* [Picado and Pultr 2011, Proposition 7.2.2, Appendix II] provides a sufficient condition for the converse to hold. The adjoint functor theorem specializes to a particularly elegant result in the case of posetal categories:

Theorem 1.7 (Adjoint Functor Theorem for posets)**.** *Let $f : \mathbb{A} \to \mathbb{B}$ be a morphism of posets, i.e. a monotone map between partially ordered sets. Assume $\mathbb{A}$ is complete and f preserves all small meets. Then f has a left adjoint, defined as*

$$f^!(b) = \bigwedge\{a \in \mathbb{A} \mid b \leq f(a)\}, \qquad \text{for all } b \in \mathbb{B}. \tag{1.2.1}$$

Dually, if $\mathbb{A}$ is cocomplete and f preserves all small joins, then it has a right adjoint, defined as

$$f_!(b) = \bigvee\{a \in \mathbb{A} \mid f(a) \leq b\}, \qquad \text{for all } b \in \mathbb{B}. \tag{1.2.2}$$

The Adjoint Functor Theorem for posets is especially relevant for frames (and thus locales). In fact morphisms of frames, which are cocomplete as posets, are of course monotone maps. Hence the theorem tells us that inside the (2-)category of posets, morphisms of frames always come with a right adjoint of the form (1.2.1). Another consequence is that every frame is an Heyting algebra, since distributivity of $a \wedge -$ means exactly such a map preserves small joins.

This is a useful fact also to make sense of morphisms of locales, which can be now be identified (at least morally) with these right adjoints.

In the case of complete Boolean algebras (*aka* Boolean frames), the situation is even richer, since a morphism of those structures preserves both meets and joins of arbitrary families, thereby we conclude it has both posetal adjoints.

1.2.2 Adjoint modalities from morphisms of Boolean frames

The following is partially informed by [Moerdijk and MacLane 1992, Chapter IX].

Every adjunction gives rise to a monad and a comonad. In the case of frames, this means that a morphism of frames $f : \mathbb{A} \to \mathbb{B}$ induces an endomorphism

$$ff_! =: \Box_f : \mathbb{B} \to \mathbb{B}. \tag{1.2.3}$$

having the following properties:

1. It is **intensive**:
$$\Box_f(b) \leq b, \qquad \text{for all } b \in \mathbb{B}.$$
2. It is **idempotent** (or **stable**):
$$\Box_f(b) \leq \Box_f\Box_f(b), \qquad \text{for all } b \in \mathbb{B}.$$
3. It is **left exact**, i.e. it preserves finite meets:
$$\Box_f(b \wedge b') \leq \Box_f(b) \wedge \Box_f(b'), \qquad \text{for all } b, b' \in \mathbb{B}.$$

The first two properties are just a general fact about adjunctions (they amount to say $\Box_f$ is a comonad). The last property is indeed peculiar of frames, since it is a consequence of the fact f preserves joins as morphism of frames.

When f is a morphism of Boolean frames it possess a left adjoint too, hence we also get a monad on $\mathbb{B}$, namely

$$ff^! =: \Diamond_f : \mathbb{B} \to \mathbb{B} \tag{1.2.4}$$

It enjoys dual properties to those of $\Box_f$, namely:

1. It is **extensive**:
$$b \leq \Diamond_f(b), \qquad \text{for all } b \in \mathbb{B}.$$

2. It is **idempotent** (or **stable**):

$$\Diamond_f \Diamond_f(b) \leq \Diamond_f(b), \qquad \text{for all } b \in \mathbb{B}.$$

3. It is **right exact**, i.e. it preserves finite joins:

$$\Diamond_f(b \vee b') \leq \Diamond_f(b) \vee \Diamond_f(b'), \qquad \text{for all } b, b' \in \mathbb{B}.$$

Moreover, since the two adjunctions from which $\Box_f$ and $\Diamond_f$ arise are on the opposite sides of the same morphism f, one gets

$$\Diamond_f \dashv \Box_f.$$

Indeed, for every $b, b' \in \mathbb{B}$:

$$\begin{aligned}
\Diamond_f(b) \leq b' \quad &\underline{\text{iff}} \quad f f^!(b) \leq b' \\
&\underline{\text{iff}} \quad f^!(b) \leq f_!(b') \\
&\underline{\text{iff}} \quad b \leq f f_!(b') \\
&\underline{\text{iff}} \quad b \leq \Box_f(b').
\end{aligned}$$

Another useful relation is duality

$$\neg\Diamond_f = \Box_f \neg, \tag{1.2.5}$$

which follows from the infinitary De Morgan's laws holding in a Boolean frame.

1.3 Typed languages and theories

We are now to describe exactly the syntax of the languages we'll be dealing with in the following pages. These are so-called **higher-order languages**, in the sense of [Jacobs 2005, Section 5.1], thus typed languages with a special type, the type of propositions Prop, whose terms correspond to formulae in a specific way.

First of all, with **language** we really mean *type theory*, meaning all the languages we'll deal with will be explicitly typed[2] and presented with meta-level statements called **judgments**:

[2] Just simply typed, hence no dependently typed or polymorphic.

judgment	applies to	usage	intended meaning
type	types	X type $[\vec{x}:\vec{\Gamma}]$	well-formed type in context
typing	terms	$t : Z\ [\vec{x}:\vec{\Gamma}]$	well-formed term in context
formula	formulae	$\varphi\ [\vec{x}:\vec{\Gamma}]$ formula	well-formed formula in context
equality	types	$(X$ is equal to $Y)$ type $[\vec{x}:\vec{\Gamma}]$	syntactically interchangeable
	terms	$(t$ is equal to $s) : Z\ [\vec{x}:\vec{\Gamma}]$	
	formulae	$(\varphi$ is equal to $\psi)$ formula $[\vec{x}:\vec{\Gamma}]$.	

Equality judgment is deemed reflexive, symmetric and transitive hence equality judgments arising from these properties are always tacitly assumed to hold. On the contrary, instances of other kinds of judgments are introduced by specific rules.

Definition 1.8. The specification of any language starts from its **vocabulary**, which is a collection of symbols:

1. A collection of **types** $X, Y, Z, \ldots$, for which the type judgment is deemed to hold.
2. A collection of **relation symbols** $R, S, T, \ldots$, each with a **signature** which is a finite, possibly empty sequence of types $(X_1, \ldots, X_n)$ whose length is called the **arity**. Relations of arity zero are **constants**.
3. A collection of **function symbols** $F, G, H, \ldots$, each with a signature $(X_1, \ldots, X_n)$ and a type Y.

As you can already appreciate from the description of judgments, we will also emphasize the dependence of terms and formulae on contexts, since they play a primary rôle in categorical logic:

Definition 1.9. A **context** is a finite, possibly empty, list[3] of distinct typed variables:

$$\vec{x}:\vec{\Gamma} \equiv x_1 : \Gamma_1, \ldots, x_n : \Gamma_n.$$

The positive integer n is the **length** of the context

Remark 1.9.1. Let us linger briefly on two meta-theoretic notation conventions we are going to use throughout the rest of the thesis.

[3] Being a list, order and multiplicity of its element matter. Being a list of typed variables, their names and their types are also an integral part of the datum of a context.

- The symbol $\equiv$ denotes 'definitional equality', hence defines the left hand side as a convenient placeholder for the right hand side.
- 'Vector notation' is a convention for writing down lists of variables, terms, types or formulae. We will employ it quite freely, hoping its intended meaning will always be unambiguous.

We stress a language[4] $\mathcal{L}$ is just a syntactical device, hence without further indications we are not able to speak of 'proofs' of a formula (or better, a sequent) nor it makes sense to speak of 'validity'.

If we imagine the theory $\mathbb{L}$ we are building out of $\mathcal{L}$ as embodied by the Lindenbaum–Tarski algebras[5] of its formulae under the different contexts available, then at this stage we only have the carrier set of the algebra.

To start talking about proofs, thus endowing the Lindenbaum–Tarski algebras with an entailment relation $\vdash$, one has to adopt a specific *proof calculus*, hence a flavour of *logic* (e.g. classical, intuitionistic, paraconsistent, linear, etc.). The combination of a language with a proof calculus[6] is a *theory*. Thus if we speak of, say, 'intuitionistic theories', we mean the datum of a language plus the logical rules of intuitionistic logic. We make only a requirement on the entailment: it must respect judgmental equality. In this way, judgmentally equal formulae are deemed equivalent (in the sense of the posetal reflection of $\vdash$).

Having a theory $\mathbb{L}$ at hand (e.g. the classical theory of groups), one can now start to look for structures behaving as prescribed by $\mathbb{L}$, i.e. *models* of $\mathbb{L}$ (e.g. specific examples of groups). This portends investigating the *semantics* of $\mathbb{L}$. This step requires one to specify an *interpretation* of the underlying language $\mathcal{L}$ of $\mathbb{L}$, which is a coherent assignment of 'meaning' (i.e. some part of the target model) to each syntactic expression of $\mathcal{L}$ (types, relation and function symbols, terms, formulae). Now one can speak of *validity* of a formula: a formula is valid when its meaning is 'true', ergo if its interpretation is the same as the interpretation of the true predicate. In other words, we are now confronting the theory with a given concrete model, which may or may not satsify a given formula of $\mathbb{L}$ (e.g. a group may or may not be Abelian). Observe the notion of truth, upon which

[4]In this sense, a language is sometimes called a **signature**, for instance Jacobs uses this terminology in [Jacobs 2005].

[5]This is partially true, as terms also need to be taken into account. We'll see in Example 3.2 a better way to wrap up a theory in an algebraic structure.

[6]Actually, 'proof calculi' and 'logics' do not entirely coincide, since different calculi might be logically equivalent (i.e. generate the same entailment relation on formulae). Nevertheless, the only meaningful way to distinguish a logic from another is by the derivations they allow, thence their proof calculi.

satisfiability and validity depend, is drawn from the metalevel[7], which is used to present and define the model.

Often at this stage it is of interest to restrict one's attention to the subclass of models satisfying a certain list of *axioms* of interest, which are simply formulae of $\mathbb{L}$ we'd like to be valid in the model (e.g. the commutativity axiom in the language of groups is modelled by Abelian groups).

Finally, an interpretation is *sound* if it is also logically coherent with $\mathbb{L}$, that is, if the entailment relation of $\mathbb{L}$ is preserved in passing to the meaning. Soundness is an especially important notion since it allows one to throw away the abstractness of $\mathbb{L}$ and work directly in the model, which can be considerably easier.

1.3.1 First- and higher-order languages

We are ready to describe the particular flavour of languages we are interested with.

Definition 1.10. A **first-order language** $\mathcal{L}$ is a typed language defined as follows. For any context $[\vec{x}:\vec{\Gamma}]$:

1. Terms in context are generated by the following rules:

$$\frac{}{x_i:\Gamma_i\,[\vec{x}:\vec{\Gamma}]} \quad \text{(for any } x_i:\Gamma_i \text{ appearing in the context)} \tag{1.3.1}$$

$$\frac{t_1:Z_1\,[\vec{x}:\vec{\Gamma}],\ldots,t_n:Z_n\,[\vec{x}:\vec{\Gamma}]}{F(t_1,\ldots,t_n):Y[\vec{x}:\Gamma]} \quad \text{(for any fun. symbol } F:\vec{Z}\to Y\text{)} \tag{1.3.2}$$

2. Formulae in context are generated by the following rules:

$$\frac{}{\mathsf{false}\,[\vec{x}:\vec{\Gamma}]\text{ formula}} \qquad \frac{}{\mathsf{true}\,[\vec{x}:\vec{\Gamma}]\text{ formula}} \tag{1.3.3}$$

$$\frac{t_1:Z_1\,[\vec{x}:\vec{\Gamma}],\;\ldots,\;t_n:Z_n\,[\vec{x}:\vec{\Gamma}]}{R(t_1,\ldots,t_n)\,[\vec{x}:\Gamma]\text{ formula}} \quad \text{(for any rel. symbol } R \text{ of signature } \vec{Z}\text{)} \tag{1.3.4}$$

$$\frac{\varphi\,[\vec{x}:\vec{\Gamma}]\text{ formula},\;\psi\,[\vec{x}:\vec{\Gamma}]\text{ formula}}{(\varphi\wedge\psi)\,[\vec{x}:\vec{\Gamma}]\text{ formula}} \qquad \frac{\varphi\,[\vec{x}:\vec{\Gamma}]\text{ formula},\;\psi\,[\vec{x}:\vec{\Gamma}]\text{ formula}}{(\varphi\vee\psi)\,[\vec{x}:\vec{\Gamma}]\text{ formula}} \tag{1.3.5}$$

[7] This 'definition of truth' is due to Tarski [Hodges 2018], and it's one of many possible alternatives. Another major form of semantics leans on games [Hodges and Väänänen 2019].

$$\frac{\varphi\,[\vec{x}:\vec{\Gamma}]\text{ formula},\ \psi\,[\vec{x}:\vec{\Gamma}]\text{ formula}}{(\varphi\to\psi)\,[\vec{x}:\vec{\Gamma}]\text{ formula}} \tag{1.3.6}$$

$$\frac{\varphi\,[\vec{x}:\vec{\Gamma},z:Z]\text{ formula}}{(\exists z:Z\,\varphi)\,[\vec{x}:\vec{\Gamma}]\text{ formula}} \qquad \frac{\varphi\,[\vec{x}:\vec{\Gamma},z:Z]\text{ formula}}{(\forall z:Z\,\varphi)\,[\vec{x}:\vec{\Gamma}]\text{ formula}} \tag{1.3.7}$$

Remark 1.10.1. We are going to use the following abbreviations:

$$(\neg\varphi)\,[\vec{x}:\vec{\Gamma}] \equiv (\varphi\to\mathsf{false})\,[\vec{x}:\vec{\Gamma}],$$
$$(\varphi\leftrightarrow\psi)\,[\vec{x}:\vec{\Gamma}] \equiv ((\varphi\to\psi)\wedge(\psi\to\varphi))\,[\vec{x}:\vec{\Gamma}].$$

Remark 1.10.2. Notice in a term or formula in context $\vec{x}:\Gamma$, every free variable must appears on the list $\vec{x}:\vec{\Gamma}$. Nonetheless, we have a certain leeway in the handling of contexts, reassuring us they only have a bookkeeping rôle in the syntax, as the following lemmata illustrates:

Lemma 1.11 (Substitution). *For any first-order language $\mathcal{L}$, the following are valid rules:*

$$\frac{t:Y\,[\vec{x}:\vec{\Gamma}],\ s:Z\,[y:Y,\vec{w}:\vec{\Theta}]}{(s[y/t]):Z\,[\vec{x}:\vec{\Gamma},\vec{w}:\vec{\Theta}]} \qquad \frac{t:Y\,[\vec{x}:\vec{\Gamma}],\ \varphi\,[y:Y,\vec{w}:\vec{\Theta}]\text{ formula}}{(\varphi[y/t])\,[\vec{x}:\vec{\Gamma},\vec{w}:\vec{\Theta}]\text{ formula}} \tag{1.3.8}$$

Proof. Induction on the construction of terms and formulae. □

Corollary 1.11.1 (Weakening). *Let $\vec{x}:\vec{\Gamma}$, $\vec{y}:\vec{\Delta}$ be contexts. Then the following are valid formation rules:*

$$\frac{t:Z\,[\vec{x}:\vec{\Gamma}]}{t:Z\,[\vec{x}:\vec{\Gamma},\vec{y}:\vec{\Delta}]} \qquad \frac{\varphi\,[\vec{x}:\vec{\Gamma}]\text{ formula}}{\varphi\,[\vec{x}:\vec{\Gamma},\vec{y}:\vec{\Delta}]\text{ formula}} \tag{1.3.9}$$

Corollary 1.11.2 (Permutation). *Let $\vec{x}:\vec{\Gamma}$ be a context of length n and σ a permutation of $\{1,...,n\}$ objects. Denote with $\vec{x}':\vec{\Gamma}'$ the context where we put $x_i':\Gamma_i' \equiv x_{\sigma(i)}:\Gamma_{\sigma(i)}$. Then the following are valid formation rules:*

$$\frac{t:Z\,[\vec{x}:\vec{\Gamma}]}{t:Z\,[\vec{x}':\vec{\Gamma}']} \qquad \frac{\varphi\,[\vec{x}:\vec{\Gamma}]\text{ formula}}{\varphi\,[\vec{x}':\vec{\Gamma}']\text{ formula}} \tag{1.3.10}$$

Definition 1.12. A **higher-order language** $\mathcal{L}$ is a first-order language augmented with:

1. For each list of types $\vec{\Delta}$, a type of **propositions over** $\vec{\Delta}$:

$$\frac{}{\vec{\Delta}\to\mathsf{Prop}\text{ type}\,[\vec{x}:\vec{\Gamma}]} \quad (\text{for every context } [\vec{x}:\vec{\Gamma}])$$

2. For each list of types $\vec{\Delta}$, a **membership** relation symbol $\in_{\vec{\Delta}}$ of signature $(\vec{\Delta},\vec{\Delta}\to\mathsf{Prop})$.

Membership relations are introduced in formulae through rule (1.3.4) for introducing relations, which specializes to:

$$\frac{\sigma : \vec{\Delta} \to \mathsf{Prop}\,[\vec{x} : \vec{\Gamma}],\ t_1 : \Delta_1\,[\vec{x} : \vec{\Gamma}], \dots t_n : \Delta_n\,[\vec{x} : \vec{\Gamma}]}{t_1, \dots, t_n \in_{\vec{\Delta}} \sigma\,[\vec{x} : \vec{\Gamma}] \text{ formula}} \tag{1.3.11}$$

Notice we are using infix notation to write membership relations. Moreover, we will often resort to the vector notation to denote the list of terms on the left of a membership symbol:

$$\vec{t} \in_{\vec{\Delta}} \sigma \equiv t_1, \dots, t_n \in_{\vec{\Delta}} \sigma.$$

Terms of type $\vec{\Delta} \to \mathsf{Prop}$ are introduced by the rule

$$\frac{\varphi\,[\vec{x} : \vec{\Gamma}, \vec{y} : \vec{\Delta}] \text{ formula}}{\{\vec{y} : \vec{\Delta} \mid \varphi\} : \vec{\Delta} \to \mathsf{Prop}\,[\vec{x} : \vec{\Gamma}]} \tag{1.3.12}$$

Remark 1.12.1. The language is higher-order since quantification over $\vec{\Gamma} \to \mathsf{Prop}$ simulates quantification over predicates over $\vec{\Gamma}$. As $\vec{\Gamma} \to \mathsf{Prop}$ is itself a list of types (of length 1), we can iterate this as many times as we want and quantify over predicates of arbitrary high order.

Remark 1.12.2. Observe we defined the types $\vec{\Gamma} \to \mathsf{Prop}$ ad-hoc, and thus are not proper function types. If we had assumed $\mathcal{L}$ to have function types, it would have been sufficient to postulate the presence of the type Prop (along with its membership relation $\in_1$) and then one could have defined

$$\vec{\Gamma} \to \mathsf{Prop} \equiv \Gamma_1 \to \cdots \to \Gamma_n \to \mathsf{Prop}.$$

The membership relation $\in_{\vec{\Gamma}}$ could have been then emulated by defining:

$$\vec{x} \in_{\vec{\Gamma}} \sigma\,[\vec{x} : \vec{\Gamma}] \equiv \in_1(\mathrm{ev}_{\vec{\Gamma} \to \mathsf{Prop}}(\sigma, \vec{x}))\,[\vec{x} : \vec{\Gamma}]$$

where $\mathrm{ev}_{\vec{\Gamma} \to \mathsf{Prop}}$ is the eliminator for the function type $\vec{\Gamma} \to \mathsf{Prop}$.

Definition 1.13. An **higher-order intuitionistic theory** $\mathbb{L}$ is a higher-order language $\mathcal{L}$ considered with the proof calculus of first-order intuitionistic predicate logic [Jacobs 2005, Figure 4.1] and the following rules

$$\frac{\Theta \vdash_{\vec{\Gamma}, \vec{\Delta}} \varphi\,[\vec{x} : \vec{\Gamma}, \vec{y} : \vec{\Delta}]}{\Theta \vdash_{\vec{\Gamma}, \vec{\Delta}} \vec{y} \in_{\vec{\Delta}} \{\vec{y} : \vec{\Delta} \mid \varphi\}\,[\vec{x} : \vec{\Gamma}, \vec{y} : \vec{\Delta}]}, \qquad \frac{\Theta \vdash_{\vec{\Gamma}, \vec{\Delta}} \vec{y} \in_{\vec{\Delta}} \{\vec{y} : \vec{\Delta} \mid \varphi\}\,[\vec{x} : \vec{\Gamma}, \vec{y} : \vec{\Delta}]}{\Theta \vdash_{\vec{\Gamma}, \vec{\Delta}} \varphi\,[\vec{x} : \vec{\Gamma}, \vec{y} : \vec{\Delta}]} \tag{1.3.13}$$

Remark 1.13.1. Rules (1.3.13) say that $\varphi\,[\vec{x} : \vec{\Gamma}, \vec{y} : \vec{\Delta}]$ and $\vec{y} \in_{\vec{\Delta}} \{\vec{y} : \vec{\Delta} \mid \varphi\}\,[\vec{x} : \vec{\Gamma}, \vec{y} : \vec{\Delta}]$ are logically equivalent. If we adopt the perspective by which functional types are special functional types, then both rules follows from β-conversion (indeed, this is how Jacobs deduces (1.3.13) [Jacobs 2005, p. 316]).

Chapter 2

Preliminaries on probability theory

2.1 Categorical preliminaries on measurability structures

For this section, we reference [Tao 2011].

Definition 2.1. A σ-algebra is a countably complete Boolean algebra, that is, a Boolean algebra which admits all countable joins.

Definition 2.2. A **σ-field** on a set X is a sub-σ-algebra of $\mathcal{P}(X)$.

A σ-field, being a sub-σ-algebra of $\mathcal{P}(X)$, is a Boolean algebra with respect to the set-theoretic operations. In particular, it's ordered by inclusion.

Definition 2.3. A **measurable space** is a pair $(X, \mathcal{E})$ where X is a set and $\mathcal{E}$ is a σ-field on X. The elements of $\mathcal{E}$ are called **measurable subsets** of X.

Definition 2.4. A **measurable map** is a map $(X, \mathcal{E}) \xrightarrow{f} (Y, \mathcal{F})$ between measurable spaces pulling measurable subsets of Y back to measurable subsets of X:

$$f^{-1}(A) \in \mathcal{F}, \qquad \text{for all } A \in \mathcal{E}.$$

Measurable spaces and maps gather in a category **Msbl** which is concrete, having a forgetful functor to **Set**. This functor has both adjoints: the left one equipping a set X with its **discrete σ-field** $\mathcal{P}(X)$ and the right one equipping a set X with **codiscrete σ-field** $\{\varnothing, X\}$.

Example 2.5. If (X, τ) is a topological space, then there's a standard way to turn it into a measurable space by equipping it with its **Borel σ-field** $\mathcal{B}(X)$, i.e. the σ-field generated by its open sets. It's easy to see a continuous map between topological spaces becomes measurable once both its domain and codomain are equipped with their Borel σ-fields. Hence $\mathcal{B}$ is a faithful functor **Top** $\to$ **Msbl**. It is not full as many discontinuous functions are nevertheless measurable, e.g. any monotone function $(\mathbb{R}, \mathcal{B}(\mathbb{R})) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is measurable regardless of its continuity.

Example 2.6. Famously, not every set of $\mathbb{R}$ can be measurable by the standard Lebesgue measure (see Example 2.20), i.e. the restriction of measures to proper subsets of $\mathcal{P}(X)$ is necessary. However, this depends on the axioms we choose for our set theory: every example of a non-measurable set (such as Vitali's) depends on the axiom of choice. There are models of ZF in which every subset of $\mathbb{R}$ is Lebesgue measurable [Solovay 1970]. Incidentally, the construction of such models uses the machinery of essential algebras we are going to use in Chapter 4

Lemma 2.7. *The category* **Msbl** *admits arbitrary products.*

Proof. Given a family of σ-fields $\{\mathcal{E}_i\}_{i \in I}$, we can define their product (in the categorical sense) by forcing the canonical projection maps to be measurable. In order to realize this, it suffices to define

$$\mathcal{E} = \bigotimes_{i \in I} \mathcal{E}_i := \langle \pi_i^{-1} A_i \mid i \in I,\ A_i \in \mathcal{E}_i \rangle$$

where $\langle - \rangle$ denotes the σ-field generated by the family of subsets in its argument. It coincides with the smallest σ-field containing it. Let now $\{(X_i, \mathcal{E}_i)\}_{i \in I}$ be a family of measurable spaces: its product is the pair $(X, \mathcal{E})$ where $X = \prod_{i \in I} X_i$ and $\mathcal{E}$ was defined above. □

Remark 2.7.1. Although **Msbl** admits arbitrary products, it's not cartesian closed [Aumann et al. 1961]. To overcome the latter obstacle [Heunen et al. 2017] recently introduced a category of so-called **quasi-Borel spaces**, in which real-valued measurable functions are axiomatized in lieu of measurable subsets.

Lemma 2.8. *Let G be a family of subsets of a set X. Then the σ-field $\langle G \rangle$ (the smallest σ-field containing G) coincides with the limit of the following sequence $(G_\alpha)_{\alpha \in \mathbf{ON}}$ defined by transfinite recursion:*

1. $G_0 := G$,

2. $G_{\alpha+1} = S(G_\alpha)$,

3. $G_\lambda := S(\bigcup_{\alpha<\lambda} G_\alpha)$ *for* λ *limit ordinal,*

where $S(\cdot)$ *denotes the family of all countable unions and complements of the sets in its argument.*

Proof. We start by noticing that the above transfinite sequence stabilizes at ω_1, since G_{ω_1} is stable under complements and countable unions. Indeed, if $E \in G_{\omega_1}$ then $E \in A_\alpha$ for some countable α and thus $X \setminus E \in A_{\alpha+1} \subset G_{\omega_1}$. Moreover, suppose $\{E_n\}_{n\in\mathbb{N}} \subseteq G_{\omega_1}$. For any E_n, there exists a countable ordinal α_n such that $E_n \in A_{\alpha_n}$. But now, as ω_1 is uncountable, $\alpha_\infty = \bigcup_{n\in\mathbb{N}} \alpha_n < \omega_1$, so that $\bigcup_{n\in\mathbb{N}} E_n \in A_{\alpha_\infty} \subset G_{\omega_1}$.

This property is actually equivalent to saying G_{ω_1} forms a σ-field itself. Hence we immediately get, by minimality, that $\langle G\rangle \subseteq G_{\omega_1}$. On the other hand, by the same stability property, since $G = G_0$ sits in $\langle G\rangle$ by definition, we conclude every A_α is contained in $\langle G\rangle$ as well. This proves the other containment, proving the claim. □

Lemma 2.9. *Let* $(X, \mathcal{E}_X)$, $(Y, \mathcal{E}_Y)$, $(Z, \mathcal{E}_Z)$ *be measurable spaces and* $\varphi : \underline{X \times Y} \to \underline{Z}$ *a measurable map. Then, for every* $y \in Y$ *and* $x \in X$*, its curryings* $\varphi(\cdot, y) : X \to Z$ *and* $\varphi(x, \cdot) : \underline{Y} \to \underline{Z}$ *are measurable maps too.*

Proof. By symmetry of the cartesian product, it suffices to consider the first currying $\varphi(\cdot, y) : X \to Z$. Let $A \in \mathcal{E}_Z$. Then

$$\varphi(\cdot, y)^{-1}A = \{x \in X \mid \varphi(x, y) \in A\} = \pi_X(\varphi^{-1}A).$$

Since φ is measurable, $B = \varphi^{-1}A$ is a measurable subset of $X \times Y$, i.e. it belongs to the product σ-field $\mathcal{E}_X \otimes \mathcal{E}_Y$. We now leverage Lemma 2.8 to prove $\pi_X B$ is measurable, arguing by transfinite induction. Indeed this is trivially true for any $B \in G_0 = \{C \times D \mid C \in \mathcal{E}_X,\ D \in \mathcal{E}_Y\}$, and since π_X is surjective, it commutes with $S(\cdot)$: images do always commute with unions and surjectivity of π_X gives commutativity with respect to complements. □

2.1.1 Probability spaces

As the name suggests, measurable spaces are just a platform for the gadget of a measure.

Definition 2.10. A **measure** on a measurable space $(X, \mathcal{E})$ is a function $\mu : \mathcal{E} \to [0, +\infty]$ such that $\mu(\varnothing) = 0$ and which is σ-additive: for every family of pairwise disjoint measurable sets $\{A_n\}_{n\in\mathbb{N}}$,

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n).$$

Remark 2.10.1. The (possibly infinite) sum at the right hand side of the identity defining σ-additivity is well-defined since the character (and, in case of convergence, the sum) of a series with positive terms doesn't depend on the order of summation.

Measurable spaces equipped with a measure are called **measure spaces**. They form a category as well, whose arrows are again measurable maps, that we denote by **Meas**. These maps transport measures covariantly: if $(X, \mathcal{E}) \xrightarrow{f} (Y, \mathcal{F})$ is a measurable map, and if μ is a measure on $\underline{X}$, then

$$f_*\mu := \mu \circ f^{-1}$$

is a measure on $\underline{Y}$ called the **pushforward measure**.

Meager subsets. Subsets of zero measure have a prominent role in measure theory.

Definition 2.11. A measurable subset of a measure space $(X, \mathcal{E}, \mu)$ is **meager** or **null** iff such is its measure. We denote the family of meager subsets by $\ker \mu$. If $\mathcal{P}(A) \subseteq \ker \mu$ for every $A \in \ker \mu$, then $\underline{X}$ is said to be **complete**.

Example 2.12. The Borel σ-field of $\mathbb{R}$ is not complete. Its completion is called **Lebesgue σ-field**.

Definition 2.13. A measurable map $(X, \mathcal{E}, \mu) \xrightarrow{f} (Y, \mathcal{F}, \nu)$ between measure spaces is **null-preserving** if it pulls meager subsets back to meager subsets:

$$f^{-1}(A) \in \ker \mu, \qquad \text{for all } A \in \ker \nu.$$

Null-preserving maps form a subcategory in both **Meas** and **Prob**, which we denote by $\mathbf{Meas}_0$ and $\mathbf{Prob}_0$, respectively. The first has been defined in [Wendt 1996], while the latter has been defined in [Adachi and Ryu 2016], although with arrows reversed.

Remark 2.13.1. Any measure space can be completed canonically by replacing its σ-field by the one generated by adjoining all subsets of meager subsets. Notice this completion procedure only involves the σ-field $\mathcal{E}$ and the measure μ, hence one often talks about *complete σ-fields*. This construction is left adjoint to the inclusion functor $\mathbf{CMsbl}_0 \hookrightarrow \mathbf{Msbl}_0$ of the full subcategory of complete measurable spaces and null-preserving maps, exhibiting it as a reflective subcategory of **Msbl**. Completion is *not* functorial on the whole **Msbl**, since adjoining new subsets to the codomain of a map may invalidate its measurability.

Definition 2.14. Let $(X, \mathcal{E})$ be a measurable space, and let μ, ν be two measures on $\underline{X}$. Then we say ν is absolutely continuous with respect to μ, written $\nu \ll \mu$, iff

$$\ker \mu \subseteq \ker \nu.$$

Example 2.15. If $(X, \mathcal{E}, \mu) \xrightarrow{f} (Y, \mathcal{F}, \nu)$ is a null-preserving map, then $f_*\mu$ is absolutely continuous with respect to ν. The converse is also true.

Finiteness conditions. We distinguish various kinds of such spaces:

Definition 2.16. A **probability space** is a measure space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{P}(\Omega) = 1$.

In this case, elements of $\mathcal{F}$ are called **events** and $\mathbb{P}$ is their **probability**.

Definition 2.17. A measure space is **σ-finite** if it can be expressed as the union of a countable family of sets of finite measure.

Construction of measures. To construct a measure can be a tricky business: σ-fields may contain very complicated sets, even if we generate them from 'nice' sets (like intervals in the case of $\mathbb{R}$). The following is thus a very important theorem is measure theory, as it allows one to extend an assignment on a sufficiently rich generating family to a full-blown measure.

Definition 2.18. A **premeasure** on a set X is a function $\text{pre}\mu : \mathcal{B} \to [0, +\infty]$ such that

1. $\mathcal{B} \subseteq \mathcal{P}(X)$ is a Boolean subalgebra, i.e. it's closed under finite unions and complements,
2. $\text{pre}\mu$ is σ-additive *whenever that makes sense*: if $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{B}$ are pairwise disjoint and such that $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{B}$,

$$\text{pre}\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \text{pre}\mu(A_n)$$

A measure $\mu : \mathcal{F} \to [0, +\infty]$ **extends** $\text{pre}\mu$ if $\mathcal{B} \subseteq \mathcal{F}$ and $\mu|_{\mathcal{B}} \equiv \text{pre}\mu$.

Theorem 2.19 (Hahn–Kolmogorov's Extension Theorem)**.** *Let X be a set, $\mathcal{B} \subseteq \mathcal{P}(X)$ a Boolean algebra. Then every premeasure* $\text{pre}\mu : \mathcal{B} \to [0, +\infty]$ *can be extended to a measure* $\mu : \mathcal{E} \to [0, +\infty]$, *where* $\mathcal{E} = \langle \mathcal{B} \rangle$.

Proof. See [Tao 2011, Theorem 1.7.8]. □

Remark 2.19.1. The theorem *doesn't guarantee uniqueness* of the extension, but it does so under the mild assumption that $(X, \mathcal{E}, \mu)$ is σ-finite. In that case, given a second extension μ' defined on the σ-field $\mathcal{E}'$, one can show μ and μ' agree on $\mathcal{E} \cap \mathcal{E}'$ [Tao 2011, Exercise 1.7.7]. Hence μ is the unique extension of preμ on $\mathcal{E}$ (although not necessarily the maximal possible one).

Example 2.20. The standard measure on $\mathbb{R}$, Lebesgue's measure $\lambda : \mathcal{B}(\mathbb{R}) \to [0, +\infty]$, is obtained by Hahn–Kolmogorov extension. The premeasure is first defined on closed intervals by pre$\lambda([a, b]) = b - a$, and then extended by imposing additivity to a full premeasure pre$\lambda : \mathcal{I} \to [0, +\infty]$ (where $\mathcal{I}$ is the Boolean algebra generated by closed intervals).

To us, the most useful application of Theorem 2.19 will be to define the product of measures.

Lemma 2.21. *If $(X, \mathcal{E}, \mu)$ and $(Y, \mathcal{F}, \nu)$ are σ-finite measure spaces, their product is uniquely given by*

$$(X \times Y, \mathcal{E} \otimes \mathcal{F}, \mu \times \nu)$$

where $\mu \times \nu$ is constructed by Hahn–Kolmogorov extension of the premeasure

$$(\mu \times \nu)(A \times B) = \mu(A)\,\nu(B), \qquad \text{for all } A \in \mathcal{E}, B \in \mathcal{F}.$$

Proof. See [Tao 2011, Proposition 1.7.11]. □

The σ-finiteness assumption cannot be relaxed in generality without giving up uniqueness of the product measure [Tao 2011, Remark 1.7.12]. Hence the category of measure spaces **Meas** does not admit every product, not even all the finite ones.

If we restrict our attention to the (sub)category of *probability* spaces **Prob** though, we get every product:

Definition 2.22. Let $\{(X_i, \mathcal{E}_i)\}_{i \in I}$ be a family of measurable spaces. A **cylindrical set** (or simply **cylinder**) is a subset $A \subseteq \prod_{i \in I} X_i$ such that $\pi_i A = X_i$ for all but finitely many $i \in I$.

Lemma 2.23. *Let $\{(X_i, \mathcal{E}_i, \mathbb{P}_i)\}_{i \in I}$ be a family of probability spaces. Then its product $(X, \mathcal{E}, \mathbb{P})$ exists, and $\mathbb{P}$ is defined on cylinders as*

$$\mathbb{P}(A) = \prod_{i \in I} \mathbb{P}_i(\pi_i A).$$

Proof. See [Cohn 2013, Proposition 10.6.1], [Saeki 1996]. □

Remark 2.23.1. Notice products of σ-finite and probability measures can be rephrased in term of pushforward measures: the measure μ is the smallest (domain-wise) such that $\pi_{i*}\mu = \mu_i$ for every $i \in I$, when it exists. This, in turn, amounts to the universal property of products.

Despite **Prob** having all products, it isn't complete: limits of systems of probability spaces do not always exist [Cohn 2013, Exercise 10.6.5]. The Kolmogorov Consistency Theorem gives sufficient conditions for a limit to exist [Cohn 2013, Theorem 10.6.2].

2.1.2 The Giry monad

The set of probability measures on a measurable space $(X, \mathcal{E})$, denoted by $G\underline{X}$ is itself a measurable space if equipped with the σ-field $G\mathcal{E}$ generated by evaluation maps:

$$\begin{aligned} \mathrm{ev}_A : G\underline{X} &\longrightarrow [0,1], \qquad A \in \mathcal{E} \\ \mathbb{P} &\longmapsto \mathbb{P}(A) \end{aligned}$$

The construction of this space is functorial. On maps, G acts by pushforward:

$$\begin{array}{ccc} \mathbf{Msbl} & \xrightarrow{G} & \mathbf{Msbl} \\ (X,\mathcal{E}) & & G\underline{X} \\ f\big\downarrow & \longmapsto & \big\downarrow f_* \\ (Y,\mathcal{F}) & & G\underline{Y} \end{array} \tag{2.1.1}$$

Proposition 2.24 ([Lawvere 1962]). *The endofunctor G is a monad. Its unit $\eta : 1 \to G$ sends points to Dirac measures:*

$$\eta_{\underline{X}}(x) = \delta_x, \qquad \textit{for all } x \in X.$$

Its multiplication $m : GG \to G$ is given by integration:

$$m(\mathfrak{P})(A) = \int_{G\underline{X}} \mathrm{ev}_A \, \mathrm{d}\mathfrak{P}, \qquad \textit{for all } \mathfrak{P} \in GG\underline{X} \textit{ and } A \in \mathcal{E}.$$

Monads of this kind are called **probability monads**. They generalize the monad of parts on **Set** to more structured valuations. In fact, variants of the Giry monad can be defined for Polish spaces [Giry 1982], quasi-Borel spaces [Heunen et al. 2017], topological spaces [Fritz, Perrone, and Rezagholi 2019], locales [Vickers 2011] and many other contexts [nLab authors 2020]. Distribution monads [Fritz 2009], which assign to a set its set of formal convex combinations of points, can be considered 'finite approximations' to probability monads, as described in [Fritz and Perrone 2019].

We will make use of the Kleisli category of the Giry monad, the category **Markov** of measurable spaces as objects and Markov kernels as arrows:

Definition 2.25. Let $(X, \mathcal{E})$, $(Y, \mathcal{F})$ be measurable spaces. A **Markov kernel** $\underline{X} \nrightarrow \underline{Y}$ is a map $K : \mathcal{F} \times X \to [0, 1]$ such that

1. for each $A \in \mathcal{F}$, $K(A, -) : X \to [0, 1]$ is an $\mathcal{E}$-measurable map,
2. for each $x \in X$, $K(-, x) : \mathcal{F} \to [0, 1]$ is a probability measure on $\underline{Y}$.

As noticed already by Lawvere [Lawvere 1962], Markov kernels can be thought as 'random maps': the probability measure $K(-, x)$ represents the probability that x lands in a given subset of Y.

Markov kernels compose by convolution: if $K : (S, \Sigma) \nrightarrow (T, \Xi)$ and $H : (T, \Xi) \nrightarrow (U, \Lambda)$, we have

$$HK(B, s) = \int_T H(B, -)\, \mathrm{d}K(-, s), \quad \text{for all } B \in \Lambda,\, s \in S.$$

Associativity holds by Fubini–Tonelli. Identities are given by Dirac kernels:

$$1_{(S,\Sigma)}(A, s) = \delta_s(A), \quad \text{for all } A \in \Sigma,\, s \in S.$$

2.2 Stochastic processes

This section is informed by [Caravenna 2011; Grimmett and Stirzaker 2001].

The central notion of a stochastic calculus is a **stochastic process** or, from now on, simply a **process**. This is misleading, however, since 'vanilla processes' are almost never used. As we will see, the real focus is actually on those processes which interact well with the second central notion of stochastic calculus, the **filtration**.

Throughout this section, we fix the following notation:

1. $\underline{\Omega} = (\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, i.e. $\mathcal{F}$ is a σ-field on the set Ω and $\mathbb{P}$ is a measure with total mass 1.
2. $\underline{E} = (E, \mathcal{E})$ is a measurable space, i.e. $\mathcal{E}$ is a σ-field on the set E. This space will usually be $\underline{\mathbb{R}} = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, which is the real numbers together with the σ-field generated by its open sets.
3. $\underline{I} = (I, 0, +, \leq)$ is a totally ordered commutative monoid (a **time monoid**), by convention a suborder of $(\mathbb{R}, \leq)$. Stochastic calculus is usually concerned with

two cases: $I = \mathbb{N}$ (**discrete time**) and $I = \mathbb{R}^+$ (**continuous time**). When used as a measurable space, I is assumed to be equipped with its Borel σ-field, i.e. the one generated by its downward closed subsets.

We call a **random variable** any real-valued measurable function. Otherwise, a measurable function $\underline{\Omega} \to \underline{E}$ will also be called a **random element** of $\underline{E}$.

Definition 2.26. A **stochastic process** X over $\underline{\Omega}$ is an $\underline{I}$-indexed collection of random elements of $\underline{E}$, i.e.

$$X = \{X_t : \underline{\Omega} \to \underline{E}\}_{t \in I}$$

From now on, if we refer to a process X we refer to a process as defined above.

Definition 2.27. Let X and Y be processes over the same space $\underline{\Omega}$, indexed by the same set I.

1. We say X is a **modification** of Y if

$$\mathbb{P}(X_t = Y_t) = 1, \qquad \text{for all } t \in I.$$

2. We say X is **indistinguishable** from Y if

$$\mathbb{P}(\forall t \in I,\ X_t = Y_t) = 1.$$

2.2.1 Filtrations

Definition 2.28. A **filtration** of $\underline{\Omega}$ is a family of sub-σ-fields $\mathcal{F}_\bullet = \{F_t\}_{t \in I}$ of $\mathcal{F}$ totally ordered by inclusion.

In the following, we denote by $\underline{\Omega}_t$ the probability space $(\Omega, \mathcal{F}_t, \mathbb{P})$.

Definition 2.29. A filtration $\{F_t\}_{t \in I}$ of a complete probability space $\underline{\Omega}$ is **complete** if each $\underline{\Omega}_t$ is complete.

This means $\ker \mathbb{P}$ is closed under taking subsets (completeness) and $\ker \mathbb{P} \subseteq \mathcal{F}_t$ for every $t \in I$.

Definition 2.30. A filtration $\{F_t\}_{t \in I}$ of a probability space $\underline{\Omega}$ is **right continuous** if

$$\mathcal{F}_t = \bigcap_{s > t} \mathcal{F}_s =: \mathcal{F}_{t^+} \tag{2.2.1}$$

A complete and right continuous filtration is called **augmented**. Any filtration can be augmented in the obvious way: first complete every $\mathcal{F}_t$ (as in Remark 2.13.1) and then make (2.2.1) into an assignment.

A filtration represents events we know about, as time passes. To give an example: our random base could be $\underline{\Omega} = (6 \times 6, 2^{6\times 6}, \#)$, representing the throw of two dice. Then a natural filtration for this space is:

$$\begin{aligned} \mathcal{F}_0 &= \{\varnothing, \Omega\} \\ \mathcal{F}_1 &= \{A \times B \mid A \subseteq 6, B = 6 \text{ or } B = \varnothing\} \\ \mathcal{F}_2 &= \mathcal{P}(\Omega). \end{aligned}$$

That is, before any dice is thrown we can only talk about anything or nothing happening, then after the first dice is thrown we can talk about anything regarding the first dice but still say nothing about the second, and finally we can talk about any possible event regarding the throw of both dice.

The absence of left continuity instead means we can have surprises from the future, which is usually a reasonable thing to expect (imagine a stock price suddenly jumping because of something unpredictable happening). Right continuity means these surprises come at definite moments in time.

Example 2.31. Consider the set $\Omega = 2^{\mathbb{N}}$ of infinite sequences of coin tosses. We construct its probability space structure by finite approximations. Hence for any $\Omega_n = 2^n$, we simply consider its full algebra of parts $\mathcal{P}(\Omega_n)$ together with the counting measure $\mathbb{P}_n$, so that any single sequence of launches has probability $1/2^n$. Define

$$\mathcal{F}_n := \{A \times 2^{\mathbb{N}} \mid A \in \mathcal{P}(\Omega_n)\}$$

Clearly each of these is a σ-field on Ω. Define $\mathcal{F}_\infty = \langle \mathcal{F}_n \mid n \in \mathbb{N} \rangle$. Moreover, let $\mathcal{B} = \bigcup_{n \in \mathbb{N}} \mathcal{F}_n$. This family is a Boolean algebra. On $\mathcal{B}$ we can define a premeasure pre$\mathbb{P}$ by using the probability measures on each $\underline{\Omega}_n$:

$$\text{pre}\mathbb{P}(A) := \mathbb{P}_n(A), \text{ where } n \text{ is the least } n \in \mathbb{N} \text{ such that } A \in \mathcal{F}_n.$$

Intuitively, a set $A \in \mathcal{B}$ 'contains only a finite number of finite sequences of launches', so that we can consider it a subset in $\mathcal{F}_n$ for $n =$ "length of the longest sequence appearing in A", and thus we can measure it using $\mathbb{P}_n$. By Hahn–Kolmogorov extension, we obtain a probability measure $\mathbb{P}$ defined on the whole $\mathcal{F}_\infty$ and agreeing with pre$\mathbb{P}$ on $\mathcal{B}$.

Eventually, we have built the space $(2^{\mathbb{N}}, \mathcal{F}_\infty, \mathbb{P})$ and a filtration $\mathcal{F}_\bullet = \{\mathcal{F}_n\}_{n \in \mathbb{N}}$.

Definition 2.32. The **natural filtration** associated to a stochastic process X is the filtration defined as

$$\mathcal{F}_t^X = \langle X_s \,|\, s \leq t \rangle$$

for every $t \in I$.

Definition 2.33. A filtration $\mathcal{G}_\bullet$ of $\underline{\Omega}$ is a **sub-filtration** of $\mathcal{F}_\bullet$ whenever

$$\mathcal{G}_t \subseteq \mathcal{F}_t, \qquad \text{for all } t \in I.$$

Clearly filtrations of $\underline{\Omega}$ with this order (which is just pointwise order) form a poset we denote by $\mathfrak{F}$. If $\mathcal{F}_\bullet$ is a sub-filtration of $\mathcal{G}_\bullet$, we'll also say $\mathcal{F}_\bullet$ is coarser than $\mathcal{G}_\bullet$, or that $\mathcal{G}_\bullet$ is finer than $\mathcal{F}_\bullet$.

2.2.2 Measurability properties

For this part, it is useful to recall $\underline{I}$ comes equipped with a canonical σ-field generated by its downward closed subsets.

Definition 2.34. A stochastic process X is **adapted** to the filtration $\mathcal{F}_\bullet$ if its natural filtration $\mathcal{F}_\bullet^X$ is a subfiltration of $\mathcal{F}_\bullet$.

The following are self-evident:

Proposition 2.35. *The natural filtration $\mathcal{F}_\bullet^X$ associated to a stochastic process X is the coarsest filtration in $\mathfrak{F}$ for which X is an adapted process.*

Proposition 2.36. *A stochastic process X is adapted to the filtration $\mathcal{F}_\bullet$ if and only if each X_t is $\mathcal{F}_t$-measurable.*

Definition 2.37. A stochastic process X is **measurable** if the function

$$X : \underline{\Omega} \times \underline{I} \to \underline{E},$$

defined by $X(\omega, t) = X_t(\omega)$, is a measurable function.

Remark 2.37.1. Obviously any such function gives rise to a stochastic process, since the components of a measurable map are measurable. The converse, however, isn't true, so the notion of measurable process is non-trivial.

Definition 2.38. A stochastic process X is **progressively measurable** if for every $t \in \underline{I}$, the function

$$X^{t\downarrow} : \underline{\Omega}_t \times t\!\downarrow \to \underline{E},$$

defined by $X^{t\downarrow}(\omega, t) = X_t(\omega)$, is a measurable function.

This last condition is the strongest of the three. Any progressively measurable process is measurable and adapted, while the converse is only true up to modification.

If E is a topological space, then we can speak of **continuity** of the process.

Definition 2.39. A (resp. right, left) **continuous process** is a process $\{X_t\}_{t\in I}$ such that, for every $\omega \in \Omega$, $X(\omega, -)$ is a (resp. right, left) continuous function $I \to E$ where I is considered equipped with its order topology.

Right continuous processes enjoy strong measurability properties: they are all measurable. Furthermore, if an adapted right continuous process is always progressively measurable. This second implication now holds on the nose, without resorting to a modification. Hence for right continuous processes,

progressive measurability <u>iff</u> adaptedness.

Finally, let us mention that this holds also in case the process is only almost surely right continuous, albeit only if we assume the given filtration to be standard.

This last setting is the most common in stochastic calculus, since it is a sweet spot between generality and good properties. Most importantly, it reflects the intuitive properties of the real-life situations stochastic calculus would like to model: a right continuous process is something varying continuously in time, up to unpredictable 'surprises' (the same intuition behind right continuous filtrations). **An adapted process is one whose 'value' at time t only depends on the information available up to that time.**

A good guiding application to keep in mind when trying to understand stochastic calculus notion is gambling or investment (the discrete and continuous time cases, respectively). In this setting:

1. the probability space $\underline{\Omega}$ models the context in which our gamble or investment occurs, such as the outcomes of a string of coin tosses or the varying prices in a stock market,
2. the filtration $\mathcal{F}_\bullet$ represents our growing knowledge of the events,
3. the stochastic process X is the outcome of such gamblings, or investment strategy, i.e. X_t is the amount of money we earned or lost up to time t.

Then right continuity of both $\mathcal{F}_\bullet$ and X are reasonable modelling assumptions, whilst the adaptedness of X means we cannot use information from the future to decide our gambling strategy.

2.2.3 Conditional expectation

Definition 2.40. Let $\mathcal{G}$ be a sub-σ-field of $\mathcal{F}$, V be a random variable over $\underline{\Omega}$. Then we define the **conditional expectation** of V on $\mathcal{G}$ to be the random variable $\mathbb{E}[V \mid \mathcal{G}]$ such that

$$\int_A \mathbb{E}[V \mid \mathcal{G}] \, \mathrm{d}\mathbb{P} = \int_A V \, \mathrm{d}\mathbb{P}, \qquad \forall A \in \mathcal{G}.$$

It is important to remark that $\mathbb{E}[V \mid \mathcal{G}]$ indeed exists and it is unique up to almost everywhere equality. This can be shown using Radon–Nikodym's theorem, which deserves to be stated here:

Theorem 2.41 (Radon–Nikodym)**.** *Let $(E, \mathcal{E}, \mu)$ be a σ-finite measure space. Let ν be a second σ-finite measure on $(E, \mathcal{E})$ such that $\nu \ll \mu$, i.e. such that ν is absolutely continuous with respect to μ. Then there exists a random variable*

$$\frac{\mathrm{d}\nu}{\mathrm{d}\mu} : \underline{E} \to \underline{\mathbb{R}},$$

*called the **Radon–Nikodym derivative** of ν with respect to μ, such that*

$$\nu(A) = \int_A \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \, \mathrm{d}\mu, \qquad \textit{for all } A \in \mathcal{E}.$$

Then

$$\mathbb{E}[V \mid \mathcal{G}] := \frac{\mathrm{d}\mu^V|_{\mathcal{G}}}{\mathrm{d}\mathbb{P}|_{\mathcal{G}}}$$

where μ^V is the measure $\mu^V(A) = \int_A V \, \mathrm{d}\mathbb{P}$. Intuitively, the conditional expectation of V can be considered a version of V 'averaged out' over $\mathcal{G}$.

Example 2.42. Let V be random variable, and consider any non-empty subset $A \in \mathcal{F}$ of $\underline{\Omega}$. It generates a sub-σ-field $\langle A \rangle = \{\varnothing, A, A^C, \Omega\}$. Then

$$\mathbb{E}[V \mid \langle A \rangle](\omega) = \begin{cases} \mathbb{E}[V \mathbf{1}_A] & \omega \in A \\ \mathbb{E}[V \mathbf{1}_{A^C}] & \omega \in A^C \end{cases}$$

In general, $\mathbb{E}[V \mid \mathcal{G}]$ is equal to $\mathbb{E}[V \mathbf{1}_A]$ on every atom $A \in \mathcal{G}$, i.e. every non-empty minimal subset of $\mathcal{G}$.

Theorem 2.43. *Let $\mathcal{G}$ be a sub-σ-field of $\mathcal{F}$, let V and W be random variables on $\underline{\Omega}$. Then the following hold almost surely:*

1. *The operator $\mathbb{E}[- \mid \mathcal{G}]$ is linear:*

$$\mathbb{E}[\alpha V + \beta W \mid \mathcal{G}] = \alpha \mathbb{E}[V \mid \mathcal{G}] + \beta \mathbb{E}[W \mid \mathcal{G}], \quad \forall \alpha, \beta \in \mathbb{R}.$$

2. *If V is $\mathcal{G}$-measurable,*

$$\mathbb{E}[V \mid \mathcal{G}] = V.$$

3. *If $V \perp \mathcal{G}$, in particular whenever $\mathcal{G} = \{\varnothing, \Omega\}$,*

$$\mathbb{E}[V \mid \mathcal{G}] = \mathbb{E}[V].$$

4. *If V is $\mathcal{G}$-measurable and VW is integrable,*

$$\mathbb{E}[VW \mid \mathcal{G}] = V\mathbb{E}[W \mid \mathcal{G}].$$

5. *If $\mathcal{H}$ is a sub-σ-field of $\mathcal{G}$,*

$$\mathbb{E}[\mathbb{E}[V \mid \mathcal{G}] \mid \mathcal{H}] = \mathbb{E}[V \mid \mathcal{H}].$$

In particular,

$$\mathbb{E}[\mathbb{E}[V \mid \mathcal{G}]] = \mathbb{E}[V].$$

2.2.4 Martingales

Definition 2.44. A **martingale** is a real-valued stochastic process X such that

1. X_t is integrable for every $t \in I$ and

2. for each $s \leq t$ in $\underline{I}$:

$$\mathbb{E}[X_t \mid \mathcal{F}_s] = X_s.$$

The first is a technical condition, which allows us to take expectations of the X_t. The second condition is called **martingality**. If in this condition we replace equality with $\leq$ or $\geq$ we get the definition of, respectively, **supermartingale** and **submartingale**.

Intuitively, a martingale is a process whose expected outcome doesn't change through time. In the stochastic-processes-as-gambling-strategies metaphor, a martingale is a gambling strategy which, on average, doesn't achieve a net gain.

Example 2.45. Consider the space of infinite sequences of coin tosses, as built in Example 2.31. Suppose we gamble on this random experiment, by betting a certain some of money on each toss. If we win, we are paid twice our bet. Define X to be the process representing our net gain when we gamble with the following betting strategy:

Bet 2^n € on the n-th toss being 'heads'.

The reason behind this choice is the inequality

$$1 + 2 + \ldots + 2^{n-1} < 2^n.$$

Therefore even if we lose the first $n-1$ bets, as soon as we win the n-th we compensate all of our losses so far. Explicitly, our stochastic process will be defined as

$$X_n = \sum_{i=1}^{n} 2^i \left(\mathbf{1}_{\omega_i = H} - \mathbf{1}_{\omega_i = T}\right)$$

where ω_i denotes the i-th item of the sequence $\omega \in 2^{\mathbb{N}}$.

We claim this process is a martingale. Start by noticing the events $\{\omega_i = H\}$ are all independent: for all $i \neq j$

$$0 = \mathbb{P}(\omega_{i_1} = H, \ldots, \omega_{i_n} = H) = \mathbb{P}(\omega_{i_1} = H) \cdots \mathbb{P}(\omega_{i_n} = H) = 0, .$$

In particular, the events $\{\omega_i = H\}$ for $i > m$ are independent from those with $i \leq m$, entailing $X_n - X_m \perp \mathcal{F}_m$. Moreover, any $\{\omega_i = H\}$ for $i \leq m$ is clearly $\mathcal{F}_m$ measurable. We then leverage Theorem 2.43 to compute the conditional expectation required for martingality: let $m < n \in \mathbb{N}$, then

$$\mathbb{E}[X_n \mid \mathcal{F}_m] = \mathbb{E}[X_m \mid \mathcal{F}_m] + \mathbb{E}[X_n - X_m \mid \mathcal{F}_m] = \underbrace{X_m}_{X_m \in \mathcal{F}_m} + \underbrace{\mathbb{E}[X_n - X_m]}_{X_n - X_m \perp \mathcal{F}_m} = X_m.$$

As the last computation suggests, martingales can also be characterized in terms of their increments:

Proposition 2.46. *An adapted, integrable process is a martingale if and only if, for each* $s \leq t \in \underline{I}$,

$$\mathbb{E}[X_t - X_s \mid \mathcal{F}_s] = 0.$$

Chapter 3

Tripos theory

The goal of this chapter is to describe the categorical semantics of higher-order intuitionistic theories, which are higher-order languages equipped with the proof calculus of intuitionistic logic. Although with a marginally different notation, such calculus is described in [Jacobs 2005, Figure 4.1].

3.1 Triposes

The scaffolding for the semantics we are going to use is fibrational, meaning we take quite literally the expression 'predicate over a context' by assigning a (Heyting) prealgebra of formulae to every 'context' of a category. This can be done in the obvious way, yielding a functor as in the definition below, or dually, yielding what would actually be a fibration of categories. As the two approaches are equivalent (thanks to the Grothendieck construction), we follow what we deem the more intuitive approach and we don't elaborate further on the matter.

Definition 3.1. Let **C** be a category with finite products. A **C-tripos** is a pseudofunctor

$$\mathbf{C}^{\mathsf{op}} \xrightarrow{\quad \mathsf{P} \quad} \mathbf{PreOrd}$$

such that:

1. It factors through **PreHey** $\hookrightarrow$ **PreOrd**, the category of Heyting prealgebras.
2. For every morphism $X \xrightarrow{f} Y$ in **C**, there's a string of adjoints:

   $$\exists_f \dashv f^* \dashv \forall_f$$

   where we set $f^* := \mathsf{P}f$.

3. The **Beck–Chevalley condition** holds: for every pullback square

$$\begin{array}{ccc} X & \xrightarrow{g} & Z \\ {\scriptstyle f}\downarrow & \lrcorner & \downarrow{\scriptstyle h} \\ Y & \xrightarrow{k} & W \end{array}$$

   the following commutes up to equivalence:

$$\begin{array}{ccc} \mathsf{P}X & \xrightarrow{\exists_g} & \mathsf{P}Z \\ {\scriptstyle f^*}\uparrow & & \uparrow{\scriptstyle h^*} \\ \mathsf{P}Y & \xrightarrow{\exists_k} & \mathsf{P}W\ . \end{array}$$

4. For every object $Y : \mathbf{C}$, there is an object $\pi(Y) : \mathbf{C}$ and an element $\in_Y \in \mathsf{P}(Y \times \pi(Y))$ having the following property: for every object $\Gamma : \mathbf{C}$ and element $\varphi \in \mathsf{P}(Y \times \Gamma)$, there is an arrow $\{\varphi\} : \Gamma \to \pi(Y)$ such that

$$(1_Y \times \{\varphi\})^*(\in_Y) \simeq \varphi. \tag{3.1.1}$$

Remark 3.1.1. The definition of tripos might appear rather obscure at first but, as anticipated, has a very lucid intuition behind: **a tripos embodies C with a certain higher-order predicate logic**.

The objects $\mathsf{P}X$ play the rôle of Lindenbaum–Tarski algebras of formulae over X. Indeed, it is customary to refer to elements of $\mathsf{P}X$ as P**-predicates**. The morphisms $f^* : \mathsf{P}Y \to \mathsf{P}X$ stand for substitution arrows, (ideally) sending a predicate $\varphi\,[y : Y]$ over Y to the predicate $\varphi[y/f(x)]\,[x : X]$ over X.

The adjoints $\exists_f$ and $\forall_f$ are, as one might have already guessed, simply existential and universal quantifiers, respectively (actually, they are slightly more general: the quantifiers we are accustomed to are adjoints to π_X^*, for a projection morphism $\pi_X : X \times Y \to X$).

The condition of Beck and Chevalley, as mysterious as it may look, is simply a necessary commutativity condition between existential quantification and substitution. Moreover, it follows from the axioms the dual Beck–Chevalley condition holds too: for a pullback like that in the definition we have

$$\forall_g\, f^* \simeq h^*\, \forall_k.$$

Finally, the fourth and most cryptic axiom can be interpreted in the following way: we want $\pi(Y)$ to play the rôle of the object of 'definable subobjects of Y', and $\in_Y$ to star as the **membership predicate**. Such a pair is called a **weak power object** of P for Y [Pasquali 2016]. Then the axiom prescribes the defining property of such a pair, by imposing that any predicate on Y, possibly dependent on a context given by Γ, actually defines a subobject of Y for every $x : \Gamma$. Indeed, if we imagine that a morphism f^* acts as a substitution, the axiom is saying:

"$y \in_Y \{\varphi\}(x)$ is logically equivalent to $\varphi(x, y)$".

Compare this to the rules for the type of propositions of a higher-order language (cf. (1.3.12) and (1.3.13)). In this sense, weak power objects 'reify' predicates.

Remark 3.1.2. Pseudofunctors satisfying only the first three axioms are called (**Lawvere**) **hyperdoctrines**. They're the analogous of triposes for first-order predicate logic. Indeed, they lack exactly the axiom which enables higher-order reasoning in the tripos. This means much of the following theory works for hyperdoctrines too, once one peels off the parts pertaining higher-order logic.

Remark 3.1.3. The requirement to have both adjoints to every map $f^* : \mathsf{P}Y \to \mathsf{P}X$ can actually be simplified, since one can build $\exists_f$ and $\forall_f$ with $\exists_{\pi_Y}$, $\forall_{\pi_Y}$ and $\exists_{\Delta_Y}$: for all $\varphi \in \mathsf{P}X$,

$$\begin{aligned} \exists_f(\varphi) &\simeq \exists_{\pi_Y}((f \times 1_Y)^*(\exists_{\Delta_Y}(\top_{\mathsf{P}(Y \times Y)})) \wedge \pi_X^*(\varphi)), \\ \forall_f(\varphi) &\simeq \forall_{\pi_Y}((f \times 1_Y)^*(\exists_{\Delta_Y}(\top_{\mathsf{P}(Y \times Y)})) \to \pi_X^*(\varphi)). \end{aligned} \tag{3.1.2}$$

We stress the importance of the predicate $\exists_{\Delta_Y}(\top_{\mathsf{P}(Y \times Y)})$. One might call this the **external equality predicate on** Y, and for this reason we'll denote it by $=_Y$, or simply $=$. In general, it holds that

$$\exists_{\Delta_Y}(\varphi) = \exists_{\Delta_Y}(\Delta_Y^* \pi_1^*(\varphi) \wedge \top_Y) = \pi_1^*(\varphi) \wedge \exists_{\Delta_Y}(\top_Y). \tag{3.1.3}$$

Example 3.2. It is well-known any *simply-typed language* $\mathcal{L}$ can be turned into a category with products $\mathbf{S}[\mathcal{L}]$, called its **syntactic category**, roughly as follows:

1. its objects are contexts considered up to 'α-equivalence', i.e. renaming of variables, hence lists of types;
2. morphisms $[\vec{x} : \vec{\Gamma}] \xrightarrow{t} [\vec{y} : \vec{\Delta}]$ are lists of terms

   $$\vec{t} : \vec{\Delta}\, [\vec{x} : \vec{\Gamma}] :\equiv t_1 : \Delta_1\, [\vec{x} : \vec{\Gamma}], \ldots, t_n : \Delta_n\, [\vec{x} : \vec{\Gamma}]$$

   where n is the length of $\vec{\Delta}$;

3. the composite of $[\vec{x}:\vec{\Gamma}] \xrightarrow{t} [\vec{y}:\vec{\Delta}]$ and $[\vec{y}:\vec{\Delta}] \xrightarrow{s} [\vec{z}:\vec{\Theta}]$ is given by substitution:

$$(\vec{s}:\vec{\Theta})\,[\vec{y}:\vec{\Delta}] \circ (\vec{t}:\vec{\Delta})\,[\vec{x}:\vec{\Gamma}] := (\vec{s}[\vec{y}/\vec{t}]:\vec{\Theta})\,[\vec{x}:\vec{\Gamma}];$$

4. identities are given by the same list of variables appearing in the context:

$$1_{[\vec{x}:\vec{\Gamma}]} = \vec{x}:\vec{\Gamma}\,[\vec{x}:\vec{\Gamma}];$$

5. and finally, binary products are given by concatenation of lists:

$$[\vec{x}:\vec{\Gamma}] \times [\vec{y}:\vec{\Delta}] = [\vec{x}:\vec{\Gamma}, \vec{y}:\vec{\Delta}]$$

   with projections simply given by the list of variables appearing in the target context; and the terminal object is the empty context [ ]. Of course, we suppose the variables appearing in $\vec{x}:\vec{\Gamma}$ and those appearing in $\vec{y}:\vec{\Delta}$ are distinct.

For a more detailed account of $\mathbf{S}[\mathcal{L}]$, see [Pitts 1989]. Suppose now $\mathcal{L}$ is actually higher-order (see Definition 1.12) with propositional equality, and make it into a theory $\mathbb{L}$ by choosing an intuitionistic proof calculus. Then the above motivation takes a literal meaning if we define the pseudofunctor:

$$\begin{array}{ccc} \mathbf{S}[\mathcal{L}]^{\mathsf{op}} & \xrightarrow{\mathsf{LT}_{\mathbb{L}}} & \mathbf{PreOrd} \\ [\vec{x}:\vec{\Gamma}] & & (\mathsf{LT}_{\mathbb{L}}(\vec{x}:\vec{\Gamma}), \vdash_{\vec{x}:\vec{\Gamma}}) \\ {\scriptstyle t}\downarrow & \longmapsto & \uparrow{\scriptstyle t^*} \\ [\vec{y}:\vec{\Delta}] & & (\mathsf{LT}_{\mathbb{L}}(\vec{y}:\vec{\Delta}), \vdash_{\vec{y}:\vec{\Delta}}) \end{array} \tag{3.1.4}$$

which is a tripos since

1. $\mathsf{LT}_{\mathbb{L}}(\vec{x}:\vec{\Gamma})$ is the Lindenbaum–Tarski algebra of propositions in context Γ, ordered by the entailment relation $\vdash_{\Gamma}$, thus a Heyting prealgebra as remarked before.

2. The morphism t^* induced by $\vec{t}:\vec{\Delta}\,[\vec{x}:\vec{\Gamma}]$ is defined by substitution:

$$t^*(\varphi\,[\vec{y}:\vec{\Delta}]) := (\varphi\,[\vec{y}/\vec{t}])\,[\vec{x}:\vec{\Gamma}].$$

3. The adjoints of t^* are obtained by using the actual quantifiers of $\mathcal{L}$:

$$\exists_t(\psi\,[\vec{x}:\vec{\Gamma}]) := \exists \vec{x}:\vec{\Gamma}\,(\vec{t}(\vec{x}) = \vec{y} \wedge \psi(\vec{x})),$$
$$\forall_t(\psi\,[\vec{x}:\vec{\Gamma}]) := \forall \vec{x}:\vec{\Gamma}\,(\vec{t}(\vec{x}) = \vec{y} \to \psi(\vec{x})).$$

   Notice equality too is crucial to their definition.

4. The weak power object of a context $[\vec{y}:\vec{\Delta}]$ is given by

$$\pi[\vec{y}:\vec{\Delta}] := [\sigma:\vec{\Delta} \to \mathsf{Prop}],$$

with membership predicate:

$$\in_{[\vec{y}:\vec{\Delta}]} := \vec{y} \in_{\vec{\Delta}} \sigma\, [\vec{y}:\vec{\Delta}, \sigma:\vec{\Delta} \to \mathsf{Prop}].$$

So every $\varphi\, [\vec{x}:\vec{\Gamma}, \vec{y}:\vec{\Delta}]$ in $\mathsf{LT}_{\mathbb{L}}(\vec{x}:\vec{\Gamma})$ determines a term

$$[\vec{x}:\vec{\Gamma}] \xrightarrow{\{\varphi\}} \pi[\vec{y}:\vec{\Delta}] := \{\vec{y}:\vec{\Delta} \mid \varphi\} : \vec{\Delta} \to \mathsf{Prop}\, [\vec{x}:\vec{\Gamma}]$$

so that

$$(1_{[\vec{y}:\vec{\Delta}]} \times \{\varphi\})^*(\in_{[\vec{y}:\vec{\Delta}]}) \simeq (\vec{y} \in_{\vec{\Delta}} \{\vec{y}:\vec{\Delta} \mid \varphi\})\, [\vec{y}:\vec{\Delta}, \vec{x}:\vec{\Gamma}] \simeq \varphi\, [\vec{y}:\vec{\Delta}, \vec{x}:\vec{\Gamma}]$$

as formulae by rule (1.3.13), and since judgmentally equal formulae are logically equivalent.

Remark 3.2.1. As noted in Remark 1.12.2, a higher-order language as defined in Definition 1.12 does not have all function types. This is easily seen to be equivalent to the fact that its syntactic category $\mathbf{S}[\mathcal{L}]$ does not admit every exponential (only exponentials $[\vec{x}:\vec{\Gamma}] \to [\sigma:\mathsf{Prop}]$ are present), hence it is not cartesian closed. This remark is of interest since the axiom of weak power objects seems to naturally impose cartesian closure on the base category of the tripos (and indeed, Pitts initially asked this in the definition of tripos), yet this is actually not true, as this example testifies.

Example 3.3. Let $\mathbb{H}$ be a complete Heyting algebra. A prominent example of tripos is the following:

$$\begin{array}{ccc} \mathbf{Set}^{\mathsf{op}} & \xrightarrow{\mathbb{H}^{(-)}} & \mathbf{PreOrd} \\ X & & (\mathbb{H}^X, \vdash_X) \\ f \downarrow & \longmapsto & \uparrow f^* \\ Y & & (\mathbb{H}^Y, \vdash_Y) \end{array} \tag{3.1.5}$$

The ordering $\vdash_X$ is the pointwise ordering of the maps $X \to \mathbb{H}$, f^* acts by precomposition, and the adjoints are given by

$$\exists f(\varphi)(y) = \bigvee\{\varphi(x) \mid f(x) = y\}$$
$$\forall f(\varphi)(y) = \bigwedge\{\varphi(x) \mid f(x) = y\}.$$

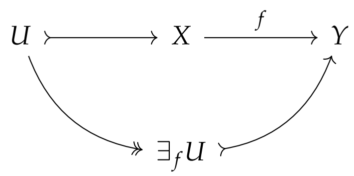
Notice completeness of $\mathbb{H}$ is essential for the definition of these. Finally, weak power objects $\pi(X)$ are simply $\mathbb{H}^X$, with membership predicate $\in_X$ given by evaluation, and representations $\{\varphi\}$ obtained by currying.

Example 3.4. If $\mathcal{E}$ is an elementary topos, the functor $\mathrm{Sub} : \mathcal{E}^{\mathrm{op}} \to \mathbf{PreOrd}$ is a tripos. This is expounded in [Moerdijk and MacLane 1992, Chapter IV]. In fact:

1. For any $X{:}\mathcal{E}$, $(\mathrm{Sub}\,X, \subseteq)$ is a Heyting algebra[1], where
   (a) $U \subseteq V$, for $U, V \in \mathrm{Sub}\,X$, means there is a monic $U \rightarrowtail V$ in $\mathcal{E}/X$.
   (b) The arrows $0 \overset{?_X}{\to} X$ and $X \overset{1_X}{\to} X$, both monic, are the bottom and top element, respectively.
   (c) The meet $U \cap V$ of two subobjects $U, V \in \mathrm{Sub}\,X$ is given by pulling back any of the two along the other.
   (d) The join $U \cup V$ of two subobjects $U, V \in \mathrm{Sub}\,X$ is given by the image of the universal morphism $U + V \to X$ (which exists since a topos admits epi-mono factorization).
   (e) Since $\mathrm{Sub}_{\mathcal{E}}\,X \cong \mathrm{Sub}_{\mathcal{E}/X}(1)$, and the exponential of two subterminal objects is again subterminal, we get an implication operator.

   Observe that, since Ω represents Sub, it follows Ω^X is an internal Heyting algebra (in the sense of Definition 1.3).

2. Morphisms $X \overset{f}{\to} Y$ lift by pullback to maps $\mathrm{Sub}\,Y \overset{f^*}{\to} \mathrm{Sub}\,X$. By direct calculation one shows these maps preserve the Heyting structure just described. Moreover, they have both adjoints:
   (a) The left adjoint $\mathrm{Sub}\,X \overset{\exists_f}{\to} \mathrm{Sub}\,Y$ is defined as the f-image functor:



   In [Moerdijk and MacLane 1992, Section IV.9] it is proved $\exists_f$ satisfy the Beck–Chevalley condition.
   (b) The right adjoint $\mathrm{Sub}\,X \overset{\forall_f}{\to} \mathrm{Sub}\,Y$ is harder to describe. First, one actually defines a right adjoint $\prod_f$ for $f^* : \mathcal{E}/Y \to \mathcal{E}/X$. In case $Y = 1$, $\prod_f(U \overset{h}{\to} X)$ is morally the subobject of U^X given by right inverses of h. One then generalizes to arbitrary Y by noticing $(\mathcal{E}/Y)/(f) \cong \mathcal{E}/X$. Then, since

[1] It should be first remarked Sub X is a set, and not a proper class. This follows from the fact $\mathcal{E}$ is locally small and Sub is represented by $\Omega_{\mathcal{E}}$.

$\prod_f$ is a right adjoint, it preserves monicity implying its restriction to $\operatorname{Sub} X \to \operatorname{Sub} Y$ is well-defined. This whole construction is carried out in detail [Moerdijk and MacLane 1992, Theorem I.9.4] and [Moerdijk and MacLane 1992, Proposition IV.9.3].

3. At last, weak power objects are obtained by taking $\pi(X)$ to be Ω^X and $\in_X$ to be the subobject $E_X \rightarrowtail X \times \Omega^X$ classified by ev_{Ω^X}. Then for any $\varphi : U \rightarrowtail \Gamma \times X$, define $\{\varphi\} : \Gamma \to \Omega^X$ to be the exponential transpose of the classifying map $\chi_\varphi : \Gamma \times X \to \Omega$ of φ. Finally, one has to prove the mono $\varphi' : U' \rightarrowtail X \times \Gamma$ one gets by pulling back $\in_X$ along $1_X \times \{\varphi\}$ is isomorphic to φ itself. Consider the following diagram:

$$\begin{array}{ccccc} U' & \longrightarrow & E_X & \longrightarrow & 1 \\ \downarrow{\scriptstyle \varphi'} & \lrcorner & \downarrow{\scriptstyle \in_X} & \lrcorner & \downarrow{\scriptstyle \mathsf{true}} \\ X \times \Gamma & \xrightarrow{1_X \times \{\varphi\}} & X \times \Omega^X & \xrightarrow{\mathrm{ev}_{\Omega^X}} & \Omega \end{array} \qquad \text{with } \chi_\varphi : X \times \Gamma \to \Omega \tag{3.1.6}$$

where both squares are pullbacks. By pasting property of pullbacks and since the bottom triangle commutes by definition of $\{\varphi\}$ as exponential transpose, we get $\varphi' \cong \varphi$.

3.2 Semantics in a tripos

In this section, we'll formalize the intuition that triposes are tailor-made for interpreting higher-order intuitionistic theories. The following is loosely informed by [Van Oosten 2008, Subsection 2.1.3] and [Pitts 1989].

Definition 3.5. Let $\mathcal{L}$ be a higher-order language as in Definition 1.12. An **interpretation** $\llbracket - \rrbracket$ of $\mathcal{L}$ in the **C**-tripos P assigns

1. to each type X, an object $\llbracket X \rrbracket$ of **C**, while a finite list of types $\vec{X}$ is canonically sent to
$$\llbracket \vec{X} \rrbracket := \llbracket X_1 \rrbracket \times \cdots \times \llbracket X_n \rrbracket,$$
and we recall the nullary product is the terminal object 1;

2. to each relation symbol R of signature $\vec{X}$, a P-predicate
$$\llbracket R \rrbracket \in \mathsf{P}\llbracket \vec{X} \rrbracket,$$

3. to each function symbol F of signature $\vec{X}$ and type Y, a morphism of $\mathbf{C}$:

$$[\![\vec{X}]\!] \xrightarrow{[\![F]\!]} [\![Y]\!].$$

We only require that for any propositional type $\vec{\Delta} \to \mathsf{Prop}$,

$$[\![\vec{\Delta} \to \mathsf{Prop}]\!] = \pi[\![\vec{\Delta}]\!], \qquad [\![\in_{\vec{\Delta}}]\!] = \in_{[\![\vec{\Delta}]\!]}.$$

We can then extend this assignment to terms and formulae of $\mathcal{L}$. For convenience, fix a context $\vec{x}:\vec{\Gamma}$ of length n. Each term $t:Y\,[\vec{x}:\vec{\Gamma}]$ will be sent to a morphism of $\mathbf{C}$:

$$[\![\vec{\Gamma}]\!] \xrightarrow{[\![t:Y[\vec{x}:\vec{\Gamma}]]\!]} [\![Y]\!],$$

while each formula $\varphi\,[\vec{x}:\vec{\Gamma}]$ becomes a P-predicate:

$$[\![\varphi\,[\vec{x}:\vec{\Gamma}]]\!] \in \mathsf{P}[\![\vec{\Gamma}]\!].$$

Interpretation of terms is inductively defined as follows:

1. for variables in context,

$$[\![x_i:\Gamma_i\,[\vec{x}:\vec{\Gamma}]]\!] := \pi_i,$$

   where $\pi_i : [\![\vec{\Gamma}]\!] \to [\![\Gamma_i]\!]$ is the projection morphism;

2. if F is a function symbol with signature $\vec{Z}$ with length m and type Y, and $t_1:Z_1\,[\vec{x}:\vec{\Gamma}], \ldots, t_m:Z_m\,[\vec{x}:\vec{\Gamma}]$ are terms in context,

$$[\![F(t_1,\ldots,t_m)\,[\vec{x}:\vec{\Gamma}]]\!] := [\![F]\!] \circ \langle [\![t_1:Z_1\,[\vec{x}:\vec{\Gamma}]]\!], \ldots, [\![t_m:Z_m\,[\vec{x}:\vec{\Gamma}]]\!] \rangle;$$

3. if $\varphi\,[\vec{x}:\vec{\Gamma}, \vec{y}:\vec{\Delta}]$ is a formula, then

$$[\![\{\vec{y}:\vec{\Delta} \mid \varphi\}:\vec{\Delta} \to \mathsf{Prop}\,[\vec{x}:\vec{\Gamma}]]\!] := \{[\![\varphi\,[\vec{x}:\vec{\Gamma}, \vec{y}:\vec{\Delta}]]\!]\} \tag{3.2.1}$$

   where $\{[\![\varphi\,[\vec{x}:\vec{\Gamma}, \vec{y}:\vec{\Delta}]]\!]\} : [\![\vec{\Gamma}]\!] \to \pi[\![\vec{\Delta}]\!]$ denotes the morphism associated to the P-predicate $[\![\varphi\,[\vec{x}:\vec{\Gamma}, \vec{y}:\vec{\Delta}]]\!]$ by the fourth tripos axiom.

Interpretation of formulae is inductively defined as well:

1. to the true and false formulae, we assign:

$$[\![\mathsf{false}\,[\vec{x}:\vec{\Gamma}]]\!] := \bot_{\mathsf{P}[\![\vec{\Gamma}]\!]},$$

$$[\![\mathsf{true}\,[\vec{x}:\vec{\Gamma}]]\!] := \top_{\mathsf{P}[\![\vec{\Gamma}]\!]}.$$

2. if R is a relation symbol of signature $\vec{Z}$ with length m, and $t_1 : Z_1\,[\vec{x}:\vec{\Gamma}]$, ..., $t_m : Z_m\,[\vec{x}:\vec{\Gamma}]$ are terms,

$$[\![\, R(t_1,\dots,t_m)\,[\vec{x}:\vec{\Gamma}]\,]\!] := \langle [\![\, t_1 : Z_1\,[\vec{x}:\vec{\Gamma}]\,]\!],\dots,[\![\, t_m : Z_m\,[\vec{x}:\vec{\Gamma}]\,]\!]\rangle^*[\![\, R\,]\!].$$

3. if $\varphi\,[\vec{x}:\vec{\Gamma}]$ and $\psi\,[\vec{x}:\vec{\Gamma}]$ are formulae,

$$[\![\,(\varphi\wedge\psi)\,[\vec{x}:\vec{\Gamma}]\,]\!] := [\![\,\varphi\,[\vec{x}:\vec{\Gamma}]\,]\!] \wedge [\![\,\psi\,[\vec{x}:\vec{\Gamma}]\,]\!],$$

$$[\![\,(\varphi\vee\psi)\,[\vec{x}:\vec{\Gamma}]\,]\!] := [\![\,\varphi\,[\vec{x}:\vec{\Gamma}]\,]\!] \vee [\![\,\psi\,[\vec{x}:\vec{\Gamma}]\,]\!],$$

$$[\![\,(\varphi\rightarrow\psi)\,[\vec{x}:\vec{\Gamma}]\,]\!] := [\![\,\varphi\,[\vec{x}:\vec{\Gamma}]\,]\!] \Rightarrow [\![\,\psi\,[\vec{x}:\vec{\Gamma}]\,]\!],$$

   where the operators appearing on the right side are those of the Heyting prealgebra $\mathsf{P}[\![\,\vec{\Gamma}\,]\!]$;

4. if $\varphi\,[\vec{x}:\vec{\Gamma}, z:Z]$ is a formula, then

$$[\![\,(\exists z:Z\,\varphi)\,[\vec{x}:\vec{\Gamma}]\,]\!] := \exists_{\pi_\Gamma}[\![\,\varphi\,[\vec{x}:\vec{\Gamma}, z:Z]\,]\!],$$

$$[\![\,(\forall z:Z\,\varphi)\,[\vec{x}:\vec{\Gamma}]\,]\!] := \forall_{\pi_\Gamma}[\![\,\varphi\,[\vec{x}:\vec{\Gamma}, z:Z]\,]\!],$$

   where $\pi_\Gamma : [\![\,\vec{\Gamma}\,]\!]\times[\![\,Z\,]\!]\rightarrow[\![\,\vec{\Gamma}\,]\!]$ is the projection morphism.

Following the same induction used to prove Lemma 1.11, Corollary 1.11.1 and Corollary 1.11.2, we can prove the following:

Lemma 3.6 (Semantics of substitution)**.** *Let $t : Y\,[\vec{x}:\vec{\Gamma}]$ and $s : Z[y:Y]$ be terms in context, and let $\varphi\,[y:Y]$ is a formula in context. Then*

$$[\![\,(s[y/t] : Z\,[\vec{x}:\vec{\Gamma}])\,]\!] = [\![\, s : Z[y:Y]\,]\!]\circ[\![\, t : Y\,[\vec{x}:\vec{\Gamma}]\,]\!],$$

$$[\![\,(\varphi\,[t/y])\,[\vec{x}:\vec{\Gamma}]\,]\!] = [\![\, t : Y\,[\vec{x}:\vec{\Gamma}]\,]\!]^*([\![\,\varphi\,[y:Y]\,]\!]).$$

Corollary 3.6.1 (Semantics of weakening)**.** *Let $t : Z\,[\vec{x}:\vec{\Gamma}]$ be a term in context $\varphi\,[\vec{x}:\vec{\Gamma}]$ a formula in context, and let $\vec{y}:\vec{\Delta}$ be another context. Then*

$$[\![\, t : Z[\vec{x}:\vec{\Gamma}, \vec{y}:\vec{\Delta}]\,]\!] = [\![\, t : Z\,[\vec{x}:\vec{\Gamma}]\,]\!]\circ\pi_\Gamma,$$

$$[\![\,\varphi\,[\vec{x}:\vec{\Gamma}, \vec{y}:\vec{\Delta}]\,]\!] = \pi_\Gamma^*[\![\,\varphi\,[\vec{x}:\vec{\Gamma}]\,]\!],$$

where $\pi_\Gamma : [\![\,\vec{\Gamma}\,]\!]\times[\![\,\vec{\Delta}\,]\!]\rightarrow[\![\,\vec{\Gamma}\,]\!]$ is the projection morphism.

Corollary 3.6.2 (Semantics of permutation)**.** *Let $\vec{x}:\vec{\Gamma}$ be a context of length n and σ a permutation of $\{1,...,n\}$. Denote with $\vec{x}':\vec{\Gamma}'$ the context define as*

$$x'_{\sigma(i)} : \Gamma'_{\sigma(i)} := x_i : \Gamma_i.$$

Moreover, let $t : Z\,[\vec{x} : \vec{\Gamma}]$ *be a term in context* $\varphi\,[\vec{x} : \vec{\Gamma}]$ *a formula in context. Then*

$$[\![\, t : Z[\vec{x}' : \vec{\Gamma}'] \,]\!] = [\![\, t : Z\,[\vec{x} : \vec{\Gamma}] \,]\!] \circ \langle \sigma \rangle,$$

$$[\![\, \varphi\,[\vec{x} : \vec{\Gamma}] \,]\!] = \langle \sigma \rangle^* [\![\, \varphi\,[\vec{x} : \vec{\Gamma}] \,]\!],$$

where $\langle \sigma \rangle : [\![\, \vec{\Gamma}' \,]\!] \to [\![\, \vec{\Gamma} \,]\!]$ *is the arrow* $\langle \pi_{\sigma(1)}, ..., \pi_{\sigma(n)} \rangle$.

Given an interpretation of $\mathcal{L}$, we can define validity for predicates:

Definition 3.7. Let $\mathcal{L}$ be a higher-order language and $[\![\, - \,]\!]$ an interpretation of $\mathcal{L}$ in the **C**-tripos P. Let $\varphi\,[\vec{x} : \vec{\Gamma}]$ be an $\mathcal{L}$-formula. Then we define

$$\mathsf{P} \vDash_{[\![-]\!]} \varphi\,[\vec{x} : \vec{\Gamma}] \quad \underline{\text{iff}} \quad \forall_! [\![\, \varphi\,[\vec{x} : \vec{\Gamma}] \,]\!] \simeq \top_{\mathsf{P}1}$$

and we read it "$\varphi\,[\vec{x} : \vec{\Gamma}]$ **is valid in** P **relative to** $[\![\, - \,]\!]$" or "P **satisfies** $\varphi\,[\vec{x} : \vec{\Gamma}]$ **relative to** $[\![\, - \,]\!]$".

Remark 3.7.1. The map $\forall_!$ is the universal closure operator, the right adjoint to $!^*$, where $! : [\![\, \vec{\Gamma} \,]\!] \to 1$ is the unique morphism existing by terminality of 1. In simpler terms, it quantifies away all the free variables occurring in φ:

$$\forall_! [\![\, \varphi\,[\vec{x} : \vec{\Gamma}] \,]\!] \simeq [\![\, \forall x_1 : X_1, \ldots x_n : X_n\ (\varphi)[\,] \,]\!].$$

It is equivalent to define

$$\mathsf{P} \vDash_{[\![-]\!]} \varphi\,[\vec{x} : \vec{\Gamma}] \quad \underline{\text{iff}} \quad [\![\, \varphi\,[\vec{x} : \vec{\Gamma}] \,]\!] \simeq \top_{\mathsf{P}[\![\vec{\Gamma}]\!]}.$$

The most important result of this section is the following:

Theorem 3.8 (Soundness Theorem Van Oosten 2008, Theorem 2.1.6)**.** *Let* $\mathcal{L}$ *be a typed relational language and* $[\![\, - \,]\!]$ *an interpretation of* $\mathcal{L}$ *in a tripos* P*. Let* φ *be an* $\mathcal{L}$*-sentence. If* φ *is provable in higher-order intuitionistic logic, then* $\mathsf{P} \models_{[\![-]\!]} \varphi$.

Corollary 3.8.1. *Let* $\mathcal{L}$ *be a typed relational language and* $[\![\, - \,]\!]$ *an interpretation of* $\mathcal{L}$ *in a tripos* P*. Let* $\varphi_1, \ldots, \varphi_n, \psi$ *be* $\mathcal{L}$*-sentences. Suppose* P *satisfies the* $\mathcal{L}$*-sentences* $\varphi_1, \varphi_2, \ldots, \varphi_n$ *relative to* $[\![\, - \,]\!]$*, and suppose* $\varphi_1, \varphi_2, \ldots \varphi_n$ *entail* ψ *in higher-order intuitionistic logic. Then* P *satisfies* ψ *relative to* $[\![\, - \,]\!]$.

A **C**-tripos can also be used syntactically, to 'put a logic on **C**'. In fact any **C**-tripos P defines a canonical higher-order language $\mathcal{L}[\mathsf{P}]$ (the **internal language of** P) whose types are objects of **C**, function symbols are morphisms of **C**, and weak power objects function as propositional types. This language has then an obvious canonical interpretation in P. This allows to manipulate the objects and arrows of **C** as if we were speaking of sets and functions. Moreover, the Soundness Theorem implies we are also able to carry out intuitionistic proofs in this language.

3.3 Tripos-to-topos construction

The fact $\mathcal{L}[\mathsf{P}]$ makes possible to 'pretend to work with sets' is remindful of what happens with topoi and their internal logic. Indeed, triposes and topoi are strongly linked together. In the words of Pitts [Pitts 2002]:

> The main use for triposes seems to occur when one has some non-standard notion of predicate and one wishes to see that it can be used to generate a topos.

Throughout this section, we'll use the language of a fixed **C**-tripos P, along with its canonical interpretation $[\![-]\!]$, with the goal of decluttering definitions and making more apparent the rationale behind them, which is to simply go and define a topos of 'sets according to P'.

Definition 3.9. The **category of partial equivalence relations** (**PERs**) $\mathbf{C}[\mathsf{P}]$ associated to P is defined in the following way:

1. An object is a pair $(X, \sim)$ where $X : \mathbf{C}$ and $\sim \in \mathsf{P}(X \times X)$ satisfies the axioms of a *partial equivalence relation* in the language of P (we abuse notation by using infix notation):

$$\begin{aligned} \text{(symmetric)} \quad & \mathsf{P} \models x \sim y \to y \sim x \, [x, y : X], \\ \text{(transitive)} \quad & \mathsf{P} \models x \sim y \land y \sim z \to x \sim z \, [x, y, z : X]. \end{aligned} \tag{3.3.1}$$

   This is called the **equality predicate** of $(X, \sim)$. Notice that by virtue of partiality, an equality predicate also specifies which 'elements' of X 'really belong' to $(X, \sim)$, namely those x in X for which $x \sim x$ holds; and the exact degree to which this holds specifies *how much* they belong to $(X, \sim)$.

2. A morphism $(X, \sim_X) \xrightarrow{F} (Y, \sim_Y)$ is an equivalence class of $F \in \mathsf{P}(X \times Y)$ which satisfy the axioms of a functional relation with respect to the equality predicates of its domain and codomain:

$$\begin{aligned} \text{(strict)} \quad & \mathsf{P} \models F(x, y) \to x \sim_X x \land y \sim_Y y \, [x : X, y : Y] \\ \text{(relational)} \quad & \mathsf{P} \models F(x, y) \land x \sim_X x' \land y \sim_Y y' \to F(x', y') \, [x, x' : X, y, y' : Y], \\ \text{(single valued)} \quad & \mathsf{P} \models F(x, y) \land F(x, y') \to y \sim_Y y' \, [x : X, y, y' : Y], \\ \text{(total)} \quad & \mathsf{P} \models x \sim_X x \to \exists y : Y \, F(x, y) \, [x : X]. \end{aligned} \tag{3.3.2}$$

   The equivalence is the posetal equivalence of $\mathsf{P}(X \times Y)$.

3. The identity of $(X, \sim)$ is the equivalence class of $\sim$ itself.
4. The composite of $(X, \sim_X) \xrightarrow{F} (Y, \sim_Y) \xrightarrow{G} (Z, \sim_Z)$ is given by

$$[\![\, \exists y : Y\, (F(x,y) \wedge G(y,z))\, [x : X, z : Z]\,]\!].$$

It is straightforward to show the above definition actually gives a category, by employing soundness of P with respect to intuitionistic logic (Theorem 3.8) to do the necessary computations.

Remark 3.9.1. Notice we didn't use weak power objects in the definition of $\mathbf{C}[\mathsf{P}]$, and that the definition only needs first-order constructs to work. In fact, any hyperdoctrine defines a category of partial equivalence relations.

Lemma 3.10. *Consider the category of PERs $\mathbf{C}[\mathsf{P}]$ arising from a tripos*[2]*, and let $(X, \sim)$ be an object. Then every subobject of $(X, \sim)$ can be presented by a strict relation on $(X, \sim)$, i.e. a predicate $\varphi \in \mathsf{P}X$ such that*

$$\begin{aligned} \textit{(strict)} \quad & \mathsf{P} \models \varphi(x) \to x \sim x\, [x : X] \\ \textit{(relational)} \quad & \mathsf{P} \models \varphi(x) \wedge x \sim x' \to \varphi(x')\, [x, x' : X]. \end{aligned} \tag{3.3.3}$$

Proof. The correspondence between subobjects and strict relations goes like this. First, to a mono $(Y, \approx) \overset{F}{\rightarrowtail} (X, \sim_X)$ we can associate the predicate

$$\varphi_F := \exists_{\pi_X} F = [\![\, \exists y : Y\, F(y,x)\, [x : X]\,]\!].$$

It is relational because F is and strict because F is. Conversely, if $\varphi \in \mathsf{P}X$ is strict on $(X, \sim)$ then we define

$$\sim_\varphi := \exists_{\Delta_X}(\varphi) = [\![\, x \sim x' \wedge \varphi(x)\, [x, x' : X]\,]\!]. \tag{3.3.4}$$

This determines the sought monomorphism into $(X, \sim)$:

$$(X, \sim_\varphi) \overset{\sim_\varphi}{\rightarrowtail} (X, \sim). \tag{3.3.5}$$

Finally, the two correspondences can be seen to be one inverse of the other rather easily: clearly

$$\exists_{\pi_X}(\sim_\varphi) = [\![\, \exists x' : X\, (x \sim x' \wedge \varphi(x))\, [x : X]\,]\!] = \varphi$$

while

$$(X, \sim_{\exists_{\pi_X} F}) = (X, [\![\, x \sim x' \wedge \exists y : Y\, F(y,x)\, [x, x' : X]\,]\!])$$

is isomorphic to $(Y, \approx)$ through F itself, considered as a morphism $(Y, \approx) \to (X, \sim_{\exists_{\pi_X} F})$. □

[2] Or any hyperdoctine, for that matter.

As anticipated, the internal logic of a tripos has enough expressive power to provide $\mathbf{C}[\mathsf{P}]$ with all the necessary structure to become a topos:

Theorem 3.11 (Pitts 1982)**.** *For a tripos* P*, the category* $\mathbf{C}[\mathsf{P}]$ *is an elementary topos, called the* ***Pitts' topos*** *of* P.

Sketch of proof. Indeed:

1. $\mathbf{C}[\mathsf{P}]$ **has finite limits**.

   (a) The terminal object is given by

   $$(1, \top_{\mathsf{P}(1\times 1)})$$

   where 1 is the terminal object of $\mathbf{C}$.

   (b) The product of $(X, \sim_X)$ and $(Y, \sim_Y)$ is $(X \times Y, \sim_{X\times Y})$, where

   $$\sim_{X\times Y} := [\![\, x \sim_X x' \wedge y \sim_Y y' \, [\langle x, y\rangle, \langle x', y'\rangle : X \times Y] \,]\!]$$

   together with the projection morphisms

   $$\pi_X := [\![\, x \sim_X x' \wedge y \sim_Y y \, [\langle x, y\rangle : X \times Y, x' : X] \,]\!],$$
   $$\pi_Y := [\![\, x \sim_X x \wedge y \sim_Y y' \, [\langle x, y\rangle : X \times Y, y' : Y] \,]\!].$$

   (c) The equalizer of $(X, \sim_X) \underset{F}{\overset{G}{\rightrightarrows}} (Y, \sim_Y)$ is $(X, \approx) \overset{E}{\rightarrowtail} (X, \sim_X)$, where

   $$E := \approx := [\![\, x \sim_X x' \wedge \exists y : Y \, (F(x, y) \wedge G(x', y)) \, [x, x' : X] \,]\!].$$

2. $\mathbf{C}[\mathsf{P}]$ **is cartesian closed**. The exponential $(Y, \sim_Y)^{(X, \sim_X)}$ is given by the object $(\pi(X \times Y), \sim)$, where

   $$\sim := [\![\, \mathsf{FR}(F) \wedge \forall x : X \, \forall y : Y \, (\langle x, y\rangle \in_{X\times Y} F \leftrightarrow \langle x, y\rangle \in_{X\times Y} G) \, [F, G : \pi(X \times Y)] \,]\!]. \tag{3.3.6}$$

   along with the evaluation map

   $$\mathrm{ev}_{Y^X} := [\![\, \langle x, y\rangle \in_{X\times Y} F \, [x : X, F : \pi(X \times Y), y : Y] \,]\!]. \tag{3.3.7}$$

   In the above definition, FR is a P-predicate expressing the fact F is a functional relation. It is basically the conjunction of the conditions (3.3.2).

3. $\mathbf{C}[\mathsf{P}]$ **has a subobject classifier**. It is the arrow $t : 1 \to (\Sigma, \Leftrightarrow)$, where $\Sigma = \pi(1)$,

   $$\Leftrightarrow := \pi_1^*(\in_1) \to \pi_2^*(\in_1) \ \wedge \ \pi_2^*(\in_1) \to \pi_1^*(\in_1), \tag{3.3.8}$$

and $\pi_1, \pi_2 : \Sigma \times \Sigma \to \Sigma$ are the two projections. The effect of such a map is really to give the logical double implication, as

$$\langle \{\varphi\}, \{\psi\} \rangle^*(\Leftrightarrow) \simeq \varphi \leftrightarrow \psi, \quad \text{for all } \varphi, \psi \in \mathsf{P}X.$$

To see this is really a subobject classifier, we can leverage Lemma 3.10 and show t can classify subobjects of type $(X, \sim_\varphi) \rightarrowtail (X, \sim)$: such a subobject is classified by the morphism $(X, \sim) \to (\Sigma, \Leftrightarrow)$ defined as

$$\chi_\varphi := [\![\, x \sim x \wedge (\{\varphi\}(x) \Leftrightarrow \sigma)\; [x : X, \sigma : \Sigma] \,]\!]. \tag{3.3.9}$$

□

Remark 3.11.1. We specified 'elementary' because not every topos obtained in this way is a Grothendieck[3] topos. The most celebrated of these is the **effective topos** $\mathcal{E}ff$, which can be obtained as the category of partial equivalence relations associated to a certain **Set**-tripos. More about this can be found in [Van Oosten 2008, Chapter 3].

Remark 3.11.2. If P were only an hyperdoctrine, the category $\mathbf{C}[\mathsf{P}]$ would only be a Heyting category, which is a category whose internal logic is first-order intuitionistic. In [Pitts 2002, Theorem 4.2], Pitts gives a necessary and sufficient condition on an hyperdoctrine P for its category of partial equivalence relations to be a topos:

Axiom (Comprehension Axiom, [Pitts 2002, Axiom 4.1]). *For all objects $Y : \mathbf{C}$ there is an object $\pi(Y) : \mathbf{C}$ and a P-predicate $\in_Y \in \mathsf{P}(Y \times \pi(Y))$ such that, for any $\Gamma : \mathbf{C}$ and P-predicate $\varphi \in \mathsf{P}(Y \times \Gamma)$,*

$$\mathsf{P} \models \forall x : \Gamma\, \exists \{\varphi\} : \pi(Y)\, \forall y : Y\, (x \in_Y \{\varphi\} \leftrightarrow \varphi(y, x)). \tag{3.3.10}$$

As noted in the article, the fourth axiom of triposes is but a Skolemization of this axiom, which means every tripos satisfies the Comprehension Axiom. On the other hand, the converse is not true! An example of such an hyperdoctrine is the **FinSet**-tripos of $\mathbb{B}$-valued (finite) sets, for an infinite complete Boolean algebra $\mathbb{B}$ [Pitts 2002, Example 4.8][4]. An hyperdoctrine satisfying the comprehension axiom is sometimes called a **moral tripos**, e.g. by Streicher [Streicher 2011], in contrast with **traditional triposes**.

[3] A topos is **Grothendieck** if it is equivalent to the topos of sheaves over a site, i.e. over a small category **C** equipped with a family of 'covering sieves' for each of its objects, which satisfy axioms inspired by the properties of open covers in a topological space.

[4] In short, for a finite set Y, $\pi(Y)$ is taken to be $\{\top, \bot\}^Y$ and $\in_Y$ is just function application, and then one

Example 3.12. Recall Example 3.3, where we defined the tripos of $\mathbb{H}$-valued sets (or simply $\mathbb{H}$-sets) for a complete Heyting algebra $\mathbb{H}$. Pitts' construction on the tripos of $\mathbb{H}$-sets yields a topos whose internal logic has $\mathbb{H}$ as its object of truth values, i.e. as subobject classifier. Since complete Heyting algebras are locales, every localic topos[5] arise from such a tripos. In particular:

Theorem 3.13. *The topos of* $\mathbb{H}$*-sets* $\mathbf{Set}[\mathbb{H}^{(-)}]$ *is equivalent to the topos of sheaves on* $\mathbb{H}$*, equipped with the topology of jointly epimorphic morphisms*[6].

Idea behind the proof. A detailed account of this proof can be found in [Johnstone 2002, Section C1.3], where it is proved the equivalence between sheaves on $\mathbb{H}$, local homeomorphism over $\mathbb{H}$ and $\mathbb{H}$-sets.

One starts defining the equivalence by turning a presheaf $\mathscr{F} : \mathbb{H}^{\mathrm{op}} \to \mathbf{Set}$ into an $\mathbb{H}$-set

$$\Theta(\mathscr{F}) := (\coprod_{h\in\mathbb{H}} \mathscr{F}(h), \sim_{\mathscr{F}}) \tag{3.3.11}$$

where

$$s \sim_{\mathscr{F}} t := \bigvee\{h \leq k \wedge \ell \mid s|_h \equiv t|_h\}, \qquad \text{for } s \in \mathscr{F}(k),\ t \in \mathscr{F}(\ell). \tag{3.3.12}$$

This definition is very close in spirit[7] to that of the étale space of $\mathscr{F}$, whereby one glues together all sections of $\mathscr{F}$ using the glueing instructions provided by $\mathscr{F}$ itself, in the form of its restriction morphisms[8]. Finally, maps $\mathscr{F} \xrightarrow{f} \mathscr{G}$ are sent to (with abuse of notation $\Theta(\mathscr{F}), \Theta(\mathscr{G})$ will denote the carrier sets):

$$\begin{aligned} \Theta(f) : \Theta(\mathscr{F}) \times \Theta(\mathscr{G}) &\longrightarrow \mathbb{H} \\ (s,t) &\longmapsto s \sim_{\mathscr{G}} f(s). \end{aligned} \tag{3.3.13}$$

Routine computations show Θ is actually a functor $\mathrm{Psh}\,\mathbb{H} \to \mathbf{Set}[\mathbb{H}^{(-)}]$. To conclude, one restricts Θ to sheaves.

shows CA holds by a clever argument exploiting Booleanity and then distributivity of B:

$$\begin{aligned} \top &= \bigwedge_{x\in\Gamma}\bigwedge_{y\in Y} \varphi(y,x) \vee \neg\varphi(y,x) \\ &= \bigwedge_{x\in\Gamma}\bigwedge_{y\in Y}\bigvee_{b\in\{\top,\bot\}} b \leftrightarrow \varphi(y,x) \\ &= \bigwedge_{x\in\Gamma}\bigvee_{\{\varphi\}\in\{\top,\bot\}^Y}\bigwedge_{y\in Y} s(y) \leftrightarrow \varphi(y,x). \end{aligned}$$

[5] A topos is **localic** if it is equivalent to the topos of sheaves on a locale. See [Moerdijk and MacLane 1992, Chapter 9] for an extensive treatment of this class of topoi.

[6] This means that a family of elements $\{h_i\}_{i\in I}$ of $\mathbb{H}$ covers an element $h \in \mathbb{H}$ iff $h_i \leq h$ for all $i \in I$, and $h \leq \bigvee_{i\in I} h_i$

Then one can show images of sheaves under Θ are *complete* $\mathbb{H}$-sets. Roughly, and $\mathbb{H}$-set $(X, \sim)$ is complete if one can extract a precise element from any of its singletons. We defer a more formal discussion on singletons and the object of singletons associated to an $\mathbb{H}$-set to the proof of Theorem 3.18. To our aim, it suffices to say completeness allows one to reason about singletons and then pass to elements straightforwardly.

For instance, if $(X, \sim_X)$ is complete, then one can define a presheaf:

$$\begin{array}{ccc} \mathbb{H}^{\mathsf{op}} & \xrightarrow{\Theta^{-1}(X,\sim_X)} & \mathbf{Set} \\ h & & \{x \in X \mid x \sim_X x = h\} \\ \rotatebox{90}{$\geq$} & \longmapsto & \uparrow -|_h \\ k & & \{x \in X \mid x \sim_X x = k\} \end{array} \tag{3.3.14}$$

Indeed, since a singleton is a subobject of $(X, \sim_X)$, it is represented by a function $X \to \mathbb{H}$. If $\{x\}$ is such a function, then $\{x\} \wedge h$ (for $h \in \mathbb{H}$) is a subobject and one can show it is again a singleton. Completeness allows us to say it is the singleton associated to an element of $(X, \sim_X)$, which we aptly name $x|_h$. This defines restrictions.

Likewise, one can use completeness to show a family of compatible sections $\{s_i \in \Theta^{-1}(X, \sim_X)(h_i)\}_{i \in I}$ glues to a common patching, since the function $x \mapsto \bigvee_{i \in I} s_i \sim_X x$ is a singleton and therefore comes from an element $s \in \bigvee_{i \in I} h_i$. This proves $\Theta^{-1}(X, \sim_X)$ is actually a sheaf.

Lastly, if $(X, \sim_X) \xrightarrow{f} (Y, \sim_Y)$ is a map between complete $\mathbb{H}$-sets, one can construct $\Theta^{-1}(f)$ simply by lifting f to singletons and then, by completeness, getting a map $X \to Y$ which can be shown to be such that $x \sim_X x = f(x) \sim_Y f(x)$ [Johnstone 2002, Lemma C1.3.8]. Therefore it defines a natural transformation $\Theta^{-1}(X, \sim_X) \to \Theta^{-1}(Y, \sim_Y)$ (f_h is f for each $h \in \mathbb{H}$, suitably restricted).

The final piece of the puzzle is given by the fact every $\mathbb{H}$-set is functorially isomorphic to a complete one, which is obtained as the object of its singletons [Johnstone 2002, Lemma C1.3.9]. Therefore the construction of Θ^{-1} on complete $\mathbb{H}$-sets suffices to define an inverse to Θ. To prove Θ and Θ^{-1} are inverses is a trivial computation. □

Example 3.14. We'll show later on that the Pitts' topos associated to the tripos of subobjects of an elementary topos $\mathcal{E}$ (Example 3.4) is equivalent to the $\mathcal{E}$ we started with.

[7] Indeed, as shown in [Johnstone 2002, p. 509], one can construct a local homeomorphism of locales using this exact idea. Hence the definition of $\Theta(F)$ is precisely the $\mathbb{H}$-set version of the étale space of F.

[8] The final picture recalls a lasagna, with large sheets of pasta on the bottom corresponding to global section and grated cheese on top corresponding to germs at points of $\mathbb{H}$.

3.3.1 Constant objects

In the tripos-to-topos construction, it is paramount to supply an equality predicate in defining the objects of $\mathbf{C}[\mathsf{P}]$, since this makes equality governed by the logic of P. However, 'external' equality is always there, and can be invoked if necessary. This gives a way to embed $\mathbf{C}$ in $\mathbf{C}[\mathsf{P}]$: define

$$\nabla_{\mathsf{P}}(X) := (X, =_X) \tag{3.3.15}$$

where $=_X$ stands for $\exists_{\Delta_X}(\top_{\mathsf{P}X})$ (see remark 3.1.3). The assignment ∇_{P} extends to morphisms: just send $X \xrightarrow{f} Y$ to

$$\nabla_{\mathsf{P}}(f) := \exists_{\langle 1_X, f\rangle}(\top_{\mathsf{P}X}) \simeq \mathsf{P}(f \times 1_Y)(=_Y) \simeq [\![\, f(x) = y\,[x\!:\!X, y\!:\!Y]\,]\!]. \tag{3.3.16}$$

Hence ∇_{P} becomes a functor $\mathbf{C} \to \mathbf{C}[\mathsf{P}]$, known as the functor of **constant objects**, which can also be shown to be left exact.

Its image is basically an internal copy of $\mathbf{C}$. However, in general $\mathbf{C}[\mathsf{P}]$ is much bigger. First of all, we only require $\mathbf{C}$ to have products, hence it's not reasonable to expect its copy to be a full-fledged topos. Moreover, since we equipped objects of $\mathbf{C}$ with synthetic equality predicates governed by the logic of P, $\mathbf{C}[\mathsf{P}]$ can be a far more exotic zoo.

Still, constant objects enjoy a privileged relationship with their host topos.

Proposition 3.15 (Van Oosten 2008, Proposition 2.4.3)**.** *Every object of* $\mathbf{C}[\mathsf{P}]$ *is the subquotient of a constant object.*

Proof. It is straightforward: for a given $(X, \sim)$, take $(X, =_{\Delta^*(\sim)})$ with the epimorphism given by $=_{\Delta^*(\sim)}$ itself. □

Intuitively, $(X, =_{\Delta^*(\sim)})$ is a 'discrete' version of X, where the possibly nuanced equality $\sim$ is replaced by an 'all-or-nothing' predicate. We keep the same elements however (in the sense of those $x \in X$ for which reflexivity of the equality predicates holds), they are just less 'cohesive'.

Finally, one can take Proposition 3.15 one step further and prove $\mathbf{C}[\mathsf{P}]$ is equivalent to the so-called *ex/reg completion* of the full subcategory $\mathbf{Ass}_{\mathbf{C}}(\mathsf{P})$ of subobjects of constant objects, which roughly means $\mathbf{C}[\mathsf{P}]$ is universal among the categories containing $\mathbf{Ass}_{\mathbf{C}}(\mathsf{P})$ and admitting all quotients of internal equivalence relations [Van Oosten 2008, Corollary 2.4.5].

3.4 Morphisms of triposes

There are a bunch of different notions of morphism between triposes. The most basic of them is that of pseudonatural transformation $\Phi : \mathsf{P} \to \mathsf{Q}$, where both P and Q are triposes on the same base $\mathbf{C}$. This is unsatisfactory, however, since morphisms usually preserve some amount of structure.

Hence one can start putting conditions on the pseudonatural transformation Φ. There are now two paths one can follow to come up with a sensible list of conditions. The first is to mimic morphisms of topoi. This makes sense since, as we've seen, triposes and topoi are cognate structures (this will become even more evident a posteriori). The second is to ask the morphisms to preserve structure, or, equivalently, to respect the kind of higher-order semantics one can do in triposes.

We'll follow both paths, and see how they actually cross.

3.4.1 Geometric morphisms

Definition 3.16. A **geometric morphism** of $\mathbf{C}$-triposes $\Phi : \mathsf{P} \to \mathsf{Q}$ is an adjoint pair $\Phi^+ \dashv \Phi_+$ of pseudonatural transformations

$$\mathsf{P} \underset{\Phi^+}{\overset{\Phi_+}{\rightleftarrows}} \mathsf{Q}$$

such that Φ^+ preserves finite meets (componentwise).

The definition, of course, echoes the homonymous notion of morphism of topoi, which we recall here:

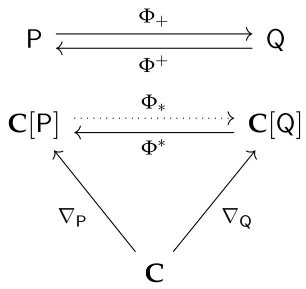
Definition 3.17. A **geometric morphism** of topoi $f : \mathcal{E} \to \mathcal{F}$ is an adjoint pair $f^* \dashv f_*$ of functors

$$\mathcal{E} \underset{f^*}{\overset{f_*}{\rightleftarrows}} \mathcal{F}$$

such that f^* is left exact (i.e. preserves finite limits).

In both cases, we'll call the left part of the adjunction **inverse image part** and the right one the **direct image part**.

Theorem 3.18 (Van Oosten 2008, Theorem 2.5.8(i))**.** *Every geometric morphism $\Phi : \mathsf{P} \to \mathsf{Q}$ of triposes induces a geometric morphism $\Phi : \mathbf{C}[\mathsf{P}] \to \mathbf{C}[\mathsf{Q}]$ between the corresponding Pitts* *topoi. Moreover, the inverse image part preserves constant objects:*



(3.4.1)

Sketch of proof. First of all, we construct the inverse image part of Φ. Notice that Φ^+, being a left adjoint commutes with left adjoints in **PreOrd**, in particular for every $f : X \to Y$ in $\mathbf{C}$, if $\exists^{\mathsf{P}}_f \dashv \mathsf{P}f$ and $\exists^{\mathsf{Q}}_f \dashv \mathsf{Q}f$, then by taking the left adjoint on both sides of the naturality condition of Φ for f, one gets

$$\Phi^+_Y \, \exists^{\mathsf{Q}}_f \simeq \exists^{\mathsf{P}}_f \, \Phi^+_X.$$

Now let $(X, \sim)$ be an object of $\mathbf{C}[\mathsf{Q}]$. Since Φ^+ preserves finite meets, $(X, \Phi^+_{X\times X}(\sim))$ is a well-defined object of $\mathbf{C}[\mathsf{P}]$, because $\Phi^+_{X\times X}(\sim)$ is again a partial equivalence relation. Moreover, Φ^+ preserves existential quantification, hence functional relations $F \in \mathsf{Q}(X \times Y)$ are sent to functional relations $\Phi^+_{X\times Y}(F) \in \mathsf{P}(X \times Y)$, meaning the latter is a well-defined arrow $(X, \Phi^+_{X\times X}(\sim)) \to (Y, \Phi^+_{Y\times Y}(\sim))$ in $\mathbf{C}[\mathsf{P}]$. Eventually, we defined a functor $\Phi^* : \mathbf{C}[\mathsf{Q}] \to \mathbf{C}[\mathsf{P}]$. It commutes with finite limits since those are defined in terms of meets and existential quantifiers alone. Finally, for the same reason one can show this functor preserves constant objects.

The construction of the direct image is more involved. The problem is, one cannot directly lift a functional relation F from $(X, \sim)$ to $(Y, \sim)$ to a functional relation from $(X, \Phi_+(\sim))$ to $(Y, \Phi_+(\sim))$: $\Phi_+(F)$ is not necessarily total anymore (Φ_+ doesn't commute on the nose with existential quantifiers).

To bypass this problem, one instead considers the 'object of singletons' $(\pi(X), \sim_S)$ associated to a given object $(X, \sim)$ of $\mathbf{C}[\mathsf{P}]$:

$$\begin{aligned} \mathsf{S}_X &:= [\![\, \exists x\!:\!X(x \sim x \wedge \forall x'\!:\!X(x' \in_X U \leftrightarrow x \sim x'))\,[U\!:\!\pi(X)]\,]\!], \\ \sim_S &:= \Leftrightarrow_{\mathsf{S}_X} := [\![\, \mathsf{S}_X(U) \wedge \forall x\!:\!X(x \in_X U \leftrightarrow x \in_X V)\,[U, V\!:\!\pi(X)]\,]\!]. \end{aligned} \tag{3.4.2}$$

The predicate S_X is the singleton predicate for $(X, \sim)$: it's satisfied by those subobjects whose elements are all equal (so-called **subsingletons**) and which are inhabited[9]. The

[9]Normally (in **Set**) subsingletons are either singletons or the empty set, but in a more general context non-singletons subsingletons can be non-empty yet not inhabited, hence the 'inhabitation condition' is important.

relation $\sim_S$ is defined as to pick the subobject of $\mathcal{P}(X,\sim)$ consisting of singleton subobjects of $(X,\sim)$, in the sense of P. Trivially, the object $(\pi(X),\sim_S)$ is isomorphic to $(X,\sim)$ via:

$$I_X := [\![\, x \sim x \wedge \forall x' : X(x' \in_X U \leftrightarrow x \sim x')\, [x : X, U : \pi(X)]\,]\!]$$

Now a functional relation $(\pi(X),\sim_S) \xrightarrow{F} (\pi(Y),\sim_S)$ gets sent to a total functional relation $(\pi(X),\Phi_+(\sim_S)) \xrightarrow{\Phi_+(F)} (\pi(Y),\Phi_+(\sim_S))$ [Van Oosten 2008, Lemma 2.5.7]. Thus we set

$$\Phi_*(X,\sim) := (\pi(X),\Phi_+(\sim_S))$$

and given an arrow $(X,\sim) \to (Y,\sim)$, we first make it into an arrow

$$(\pi(X),\sim_S) \to (\pi(Y),\sim_S)$$

by pre- and postcomposing it with the isomorphisms I_X and I_Y^{-1}, respectively; and then we apply Φ_+ to get an arrow $(\pi(X),\Phi_+(\sim)) \to (\pi(Y),\Phi_+(\sim_S))$.

At last, one shows functoriality of this construction and proves the adjunction $\Phi^* \dashv \Phi_*$ through routine computations. □

Remark 3.18.1. Van Oosten also proves the converse statement [Van Oosten 2008, Theorem 2.5.8(ii)]: every
geometric morphism $\mathbf{C}[\mathsf{P}] \to \mathbf{C}[\mathsf{Q}]$ whose inverse image part preserve constant objects is induced by an essentially unique geometric morphism of triposes $\mathsf{P} \to \mathsf{Q}$.

3.4.2 Logical morphisms

The morphisms we discuss here are naturally defined between triposes over (possibly) different bases, thus for the rest of the section fix two triposes $\mathsf{P} : \mathbf{C}^{\mathsf{op}} \to \mathbf{PreOrd}$ and $\mathsf{Q} : \mathbf{D}^{\mathsf{op}} \to \mathbf{PreOrd}$.

Definition 3.19. A **first-order logical morphism** of triposes is a pair (F,Φ) where

1. $F : \mathbf{C} \to \mathbf{D}$ is a functor preserving finite products,
2. $\Phi : \mathsf{P} \to \mathsf{Q}F$ is a pseudonatural transformation

such that

1. every component $\Phi_X : \mathsf{P}X \to \mathsf{Q}FX$ is a morphism of Heyting prealgebras,
2. for all morphisms $f : X \to Y$ in $\mathbf{C}$, Φ commutes with $\exists_f$ and $\forall_f$, i.e.

$$\Phi_Y \exists^{\mathsf{P}}_f \simeq \exists^{\mathsf{Q}}_{Ff} \Phi_X, \qquad \Phi_Y \forall^{\mathsf{P}}_f \simeq \forall^{\mathsf{Q}}_{Ff} \Phi_X.$$

Remark 3.19.1. In the proof of Theorem 3.18, we've managed to get commutation of Φ^+ and $\exists$ by exploiting the fact Φ^+ was a left adjoint. This technique shows that if $\Phi^+ \dashv \Phi_+$ is a geometric morphism of triposes, where Φ_+ is a morphism of Heyting prealgebras componentwise, and there exists a further right adjoint $\Phi_+ \dashv \Phi^!$, then Φ_+ is a first-order logical morphism.

Definition 3.20. An **higher-order logical morphism** of triposes $(F, \Phi) : \mathsf{P} \to \mathsf{Q}$ is a first-order logical morphism which preserves weak power objects, i.e. for every $Y : \mathbf{C}$:

$$\pi^{\mathsf{Q}}(FY) \cong F\pi^{\mathsf{P}}(Y), \qquad \in^{\mathsf{Q}}_{FY} \simeq \Phi_{Y \times \pi(Y)}(\in^{\mathsf{P}}_{Y}).$$

The major example of logical morphisms of triposes are given by interpretation, in the spirit of Lawvere's functorial semantics. In fact, recalling Definition 3.5 and Theorem 3.8, one readily realizes an interpretation $[\![-]\!]$ for a higher-order intuitionistic theory $\mathbb{T}$ in the tripos P coincides with the datum of a higher-order logical morphism

$$[\![-]\!] : \mathsf{LT}_{\mathbb{T}} \to \mathsf{P}.$$

This point of view is quite powerful. For example, we immediately see $\mathsf{P} \cong \mathsf{LT}_{\mathbb{L}[\mathsf{P}]}$, where $\mathbb{L}[\mathsf{P}]$ is the internal language of P equipped with higher-order intuitionistic logic (its **internal theory**): the isomorphism is witnessed by the canonical interpretation of $\mathcal{L}[\mathsf{P}]$ in P.

From the logical point of view, we see every higher-order intuitionistic theory $\mathbb{T}$ has a canonical interpretation $1_{\mathsf{LT}_{\mathbb{T}}}$ in itself, which can be identified with what it's traditionally called **generic** or **term model**. Moreover, we can now move interpretations along logical morphisms simply by composition:

Lemma 3.21 (First-order Translation Lemma)**.** *Let $(F, \Phi) : \mathsf{P} \to \mathsf{Q}$ be a first-order logical morphism between triposes. Suppose $[\![-]\!]$ is an interpretation of a first-order language $\mathcal{L}$ in P, and let $\varphi\, [\vec{x} : \vec{\Gamma}]$ be an $\mathcal{L}$-formula. Then there exists an interpretation $[\![-]\!]'$ of $\mathcal{L}$ in Q such that:*

$$\mathsf{P} \vDash_{[\![-]\!]} \varphi\, [\vec{x} : \vec{\Gamma}] \quad \underline{\textit{implies}} \quad \mathsf{Q} \vDash_{[\![-]\!]'} \varphi\, [\vec{x} : \vec{\Gamma}].$$

Proof. Let $[\![-]\!]'$ simply be $(F, \Phi) \circ [\![-]\!]$. □

Corollary 3.21.1 (Higher-order Translation Lemma)**.** *Let P and Q be $\mathbf{C}$-triposes, and let $\Phi : \mathsf{P} \to \mathsf{Q}$ be a higher-order logical morphism between them. Suppose $[\![-]\!]$ is an interpretation of a higher-order language $\mathcal{L}$ in P, and let $\varphi\, [\vec{x} : \vec{\Gamma}]$ be an $\mathcal{L}$-formula. Then there exists an interpretation $[\![-]\!]'$ of $\mathcal{L}$ in Q such that:*

$$\mathsf{P} \vDash_{[\![-]\!]} \varphi\, [\vec{x} : \vec{\Gamma}] \quad \underline{\textit{implies}} \quad \mathsf{Q} \vDash_{[\![-]\!]'} \varphi\, [\vec{x} : \vec{\Gamma}].$$

There is also a logical notion of morphisms between topoi:

Definition 3.22. A **logical morphism** of topoi $\Phi : \mathcal{E} \to \mathcal{F}$ is a left exact functor such that $\Phi(PX) \cong P(\Phi X)$ for each $X : \mathcal{E}$, where $P(-)$ denotes the power object operation.

Remark 3.22.1. Since a logical morphism is a left exact functor, $\Phi(\in_X^{\mathcal{E}}) \cong \in_{\Phi(X)}^{\mathcal{F}}$ for every $X : \mathcal{E}$ [Moerdijk and MacLane 1992, pp. 170–171]. In all respects, this means a logical morphism is a functor preserving the higher-order logical structure of topoi.

Remark 3.22.2. Contrary to the definitions of geometric morphisms for triposes and topoi, the definitions of (higher-order) logical morphism of triposes and topoi differ greatly in the requests. In fact, a(n higher-order) logical morphism of triposes, in addition to being 'left exact' and preserving power objects, needs to preserve all the Heyting structure plus the quantifiers, while a logical morphism of topoi doesn't have this requirement (or better, it is a consequence of the others). We strongly believe this state of affairs to be a mirage which would be dissolved by a more thorough investigation. In particular, we suspect the same ideas used to prove the second part of Theorem 3.18 might be repurposed to this aim.

As anticipated, here the two guiding principles behind the definition of morphisms of triposes meet: logical morphisms of triposes correspond to logical morphisms of topoi.

Theorem 3.23. *Every higher-order logical morphism $\Phi : \mathsf{P} \to \mathsf{Q}$ of triposes induces a logical morphism $\Phi^\circ : \mathbf{C}[\mathsf{P}] \to \mathbf{C}[\mathsf{Q}]$ between the corresponding Pitts topoi which commutes with constant objects:*

$$\mathsf{P} \xrightarrow{\Phi} \mathsf{Q}$$

$$\begin{array}{ccc} \mathbf{C}[\mathsf{P}] & \xrightarrow{\Phi^\circ} & \mathbf{C}[\mathsf{Q}] \\ & {\scriptstyle \nabla_{\mathsf{P}}} \nwarrow \quad \nearrow {\scriptstyle \nabla_{\mathsf{Q}}} & \\ & \mathbf{C} & \end{array} \tag{3.4.3}$$

Proof. Define

$$\begin{array}{ccc} \mathbf{C}[\mathsf{P}] & \xrightarrow{\Phi^\circ} & \mathbf{C}[\mathsf{Q}] \\ (X, \sim_X) & & (X, \Phi(\sim_X)) \\ {\scriptstyle F}\downarrow & \longmapsto & \downarrow{\scriptstyle \Phi_{X \times Y}(F)} \\ (Y, \sim_Y) & & (Y, \Phi(\sim_Y)) \end{array} \tag{3.4.4}$$

To see this is well-posed, we can apply the Higher-Order Translation Lemma to the canonical interpretation of $\mathcal{L}[\mathsf{P}]$ in P, to get a semantics in Q. This allows us to claim Φ° sends PERs to PERs, functional relations to functional relations, identities to identities, composites to composites, and finite limits to finite limits. Moreover, it commutes with constant objects since:

$$\Phi^\circ(X,=) = (X, \Phi(\exists^{\mathsf{P}}_{\Delta_X}(\top_{\mathsf{P}X}))) = (X, \exists^{\mathsf{Q}}_{\Delta_X}(\top_{\mathsf{Q}X})) = (X,=) : \mathbf{C}[Q].$$

It remains to check power objects are preserved. The power object of $(X,\sim)$ in $\mathbf{C}[\mathsf{P}]$ is given by (cf. (3.3.6) and (3.3.7))

$$(\pi^{\mathsf{P}}(1), \Leftrightarrow)^{(X,\sim_X)} = (\pi^{\mathsf{P}}(X), \sim), \qquad \in^{\mathbf{C}[\mathsf{P}]}_X := \mathrm{ev}_{(\pi^{\mathsf{P}}(1),\Leftrightarrow)^{(X,\sim_X)}} .$$

Now, since $\pi^{\mathsf{P}}(X) = \pi^{\mathsf{Q}}(X)$ by hypothesis and $\Phi(\sim)$ and $\Phi(\mathrm{ev}_{(\pi^{\mathsf{P}}(1),\Leftrightarrow)^{(X,\sim)}})$ continue to satisfy their defining properties in $\mathbf{C}[\mathsf{Q}]$ by virtue of the same Translation Lemma, the theorem stands proved. □

3.5 Semantics in a Pitts' topos

Every elementary topos $\mathcal{E}$ has an internal higher-order language with equality $\mathcal{L}[\mathcal{E}]$, sound with respect to higher-order intuitionistic logic, which is embodied by its tripos of subobjects $\mathrm{Sub}_{\mathcal{E}}$. Hence $\mathcal{L}[\mathcal{E}] = \mathcal{L}[\mathrm{Sub}_{\mathcal{E}}]$, and the same goes for semantics:

Definition 3.24. Let $\mathcal{L}$ be a higher-order language. An **interpretation $[\![-]\!]$ of $\mathcal{L}$ in a topos** $\mathcal{E}$ is an interpretation of $\mathcal{L}$ in the tripos $\mathrm{Sub}_{\mathcal{E}}$.

Thus, given an interpretation of $\mathcal{L}$ in $\mathcal{E}$, a formula $\varphi\,[\vec{x}:\vec{\Gamma}]$ is valid in $\mathcal{E}$ precisely when it is valid in its tripos of subobjects:

$$\mathcal{E} \vDash_{[\![-]\!]} \varphi\,[\vec{x}:\vec{\Gamma}] \quad \underline{\text{iff}} \quad \mathrm{Sub}_{\mathcal{E}} \vDash_{[\![-]\!]} \varphi\,[\vec{x}:\vec{\Gamma}] \quad \underline{\text{iff}} \quad \forall_!\,[\![\varphi\,[\vec{x}:\vec{\Gamma}]]\!] \simeq 1_1.$$

Remark 3.24.1. The notion of validity we just described is also known as **Kripke–Joyal semantics** [Moerdijk and MacLane 1992, Section VI.6]. In particular, when $\mathcal{L} = \mathcal{L}[\mathcal{E}]$, the canonical interpretation in $\mathcal{E}$ is denoted by $\{-\}$, to evoke the set-theoretic notation used for comprehension. Indeed the interpretation of a formula $\varphi\,[\vec{x}:\vec{\Gamma}]$ amounts to the maximal subobject of $\vec{\Gamma}$ on which it holds:

$$\{\varphi\,[\vec{x}:\vec{\Gamma}]\} \rightarrowtail \vec{\Gamma}.$$

Moreover, this subobject is classified by the morphism $\vec{\Gamma} \to \Omega$ given by the interpretation of

$$\{\vec{x}:\vec{\Gamma} \mid \varphi\} : \vec{\Gamma} \to \mathsf{Prop}\,[\vec{x}:\vec{\Gamma}].$$

Remark 3.24.2. If $U \xrightarrow{a} \Gamma$ is a generalized element of Γ, one can show that

$$\mathcal{E} \vDash_{[\![-]\!]} \varphi(a) \quad \underline{\text{iff}} \quad \begin{array}{ccc} \operatorname{im} a & \overset{\exists}{\dashrightarrow} & \{\varphi\,[\vec{x}:\vec{\Gamma}]\} \\ & \searrow & \downarrow \\ & & \Gamma \end{array} \tag{3.5.1}$$

or, equivalently, if $\chi_{\{\varphi\,[\vec{x}:\vec{\Gamma}]\}}a$ factors through $\mathsf{true} : 1 \rightarrowtail \Omega_{\mathcal{E}}$.

3.5.1 Relation between P and $\mathrm{Sub}_{\mathbf{C}[\mathsf{P}]}$

From a tripos P therefore, two languages arise:

$$\mathcal{L}[\mathsf{P}], \qquad \mathcal{L}[\mathbf{C}[\mathsf{P}]].$$

In the following, we are to establish a connection between these two languages and their semantics.

Theorem 3.25. *Every* **C***-tripos* P *admits a higher-order logical equivalence into* $\mathrm{Sub}_{\mathbf{C}[\mathsf{P}]}\nabla_{\mathsf{P}}$.

Proof. Let Q be $\mathrm{Sub}_{\mathbf{C}[\mathsf{P}]}\nabla_{\mathsf{P}}$. Since Q is a **C**-tripos too, the first component of the logical morphism (F, Φ) we are to build is just the identity on **C**. Hence we'll loosely refer to (F, Φ) as just Φ. Also, in the proof we'll make abundant use of the canonical interpretation $[\![-]\!]$ of $\mathcal{L}[\mathsf{P}]$.

Remember every subobject of $\nabla_{\mathsf{P}}(X)$ is given by a P-predicate $\varphi \in \mathsf{P}X$ (it follows from Lemma 3.10 and the fact every $\varphi \in \mathsf{P}X$ is $=$-strict). Therefore we can define

$$\Phi_X : \mathsf{P}X \longrightarrow \mathsf{Q}X$$
$$\varphi \longmapsto (X, =_\varphi) \rightarrowtail (X, =)$$

where $=$ denotes external equality, and $=_\varphi$ is defined as in (3.3.4). It's easy to see that this is a well-defined bijective correspondence with respect to both equivalence on $\mathsf{P}X$ and on $\mathsf{Q}X$. Moreover, although tedious to check, these maps are preserve all the structure of Heyting prealgebras. Naturality with respect to a **C**-morphism $f : X \to Y$ is expressed by the following square:

$$\begin{array}{ccc} \mathsf{P}Y & \xrightarrow{\mathsf{P}f} & \mathsf{P}X \\ {\scriptstyle \Phi_Y}\downarrow & & \downarrow{\scriptstyle \Phi_X} \\ \mathsf{Q}Y & \xrightarrow{\mathsf{Q}f} & \mathsf{Q}X \end{array}$$

If $\varphi \in \mathsf{P}Y$, we can chase it around the square to get that Φ is natural if and only if

$$(X, =_{\mathsf{P}f(\varphi)}) \rightarrowtail (X, =) \text{ is the pullback of } (Y, =_{\varphi}) \rightarrowtail (Y, =) \text{ along } \nabla_{\mathsf{P}}(f),$$

where we recall $\nabla_{\mathsf{P}}(f) = \mathsf{P}(f \times 1_Y)(=)$. But if $(X, \sim_{\text{pb}})$ is such a pullback, then we have

$$\begin{aligned}
[\![\, x \sim_{\text{pb}} x' \,]\!] &\simeq [\![\, x = x' \wedge \exists y\, (f(x) = y \wedge \varphi(y)) \,]\!] \\
&\simeq [\![\, x = x' \wedge \varphi(f(x)) \,]\!] \\
&\simeq [\![\, x =_{\mathsf{P}f(\varphi)} x' \,]\!].
\end{aligned}$$

Here we used the construction of inverse images in $\mathbf{C}[\mathsf{P}]$, which can be found in [Van Oosten 2008, p. 67].

We now turn to the proof of the 'logicality' of Φ. To show it commutes with quantifiers, we use the stronger property of having both adjoints, and this is true since Φ, we claim, is actually part of an adjoint equivalence. Hence we guess

$$\Phi^!_X((Y, \sim) \rightarrowtail (X, =)) := \Phi^+_X((Y, \sim) \rightarrowtail (X, =)) := [\![\, x \sim x \,]\!]$$

This would hold if

$$1_{\mathsf{Q}} \simeq \Phi\Phi^+, \qquad \Phi^+\Phi \simeq 1_{\mathsf{P}}$$

In fact, if $(Y, \sim) \rightarrowtail (X, =) \in \mathsf{Q}X$ and $\psi \in \mathsf{P}X$, this would mean

$$(Y, \sim) \rightarrowtail (X, =) \simeq (X, =_{\Phi^+_X((Y,\sim)\rightarrowtail(X,=))}) \rightarrowtail (X, =), \qquad \Phi^+_X((X, =_{\psi}) \rightarrowtail (X, =)) \simeq \psi$$

and this follows from the definition of Φ and Φ^+.

In conclusion, it remains to show weak power objects are preserved, meaning

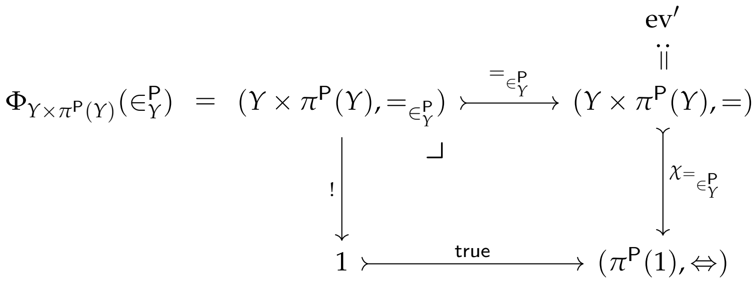
$$\pi^{\mathsf{P}}(Y) = \pi^{\mathsf{Q}}(Y), \qquad \in^{\mathsf{Q}}_Y = \Phi_{Y \times \pi^{\mathsf{P}}(Y)}(\in^{\mathsf{P}}_Y).$$

First of all, we recall from Example 3.4 that, for a tripos of subobjects, weak power objects are given by certain exponentials and evaluation maps. In our case however, we also need to pull the carrier back onto $\mathbf{C}$ through ∇_{P}:

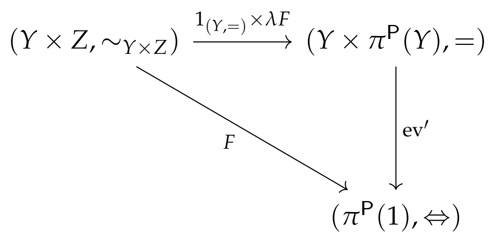
$$\pi^{\mathsf{Q}}(Y) := \nabla_{\mathsf{P}}^{-1}\left((\pi^{\mathsf{P}}(1), \Leftrightarrow)^{(Y,=)}\right),$$

$$\begin{array}{ccccc}
\in^{\mathsf{Q}}_Y & := & E_Y & \rightarrowtail & (\pi^{\mathsf{P}}(1), \Leftrightarrow)^{(Y,=)} \times (Y, =) \\
 & & \big\downarrow {\scriptstyle !} & \lrcorner & \big\downarrow {\scriptstyle \text{ev}} \\
 & & 1 & \xrightarrow{\ \mathsf{true}\ } & (\pi^{\mathsf{P}}(1), \Leftrightarrow)
\end{array}$$

Therefore, in order to prove our claim, it suffices to show $(\pi^{\mathsf{P}}(Y), =)$ satisfies the universal property of $(\pi^{\mathsf{P}}(1), \Leftrightarrow)^{(Y,=)}$, with evaluation map given by the classifying morphism of $=_{\in^{\mathsf{P}}_Y}$:



Thus let $(Z, \sim_Z) : \mathbf{C}[\mathsf{P}]$ and $(Y \times Z, \sim_{Y\times Z}) \xrightarrow{F} (\pi^{\mathsf{P}}(1), \Leftrightarrow)$. We want to show there exists a unique arrow $\lambda F : (Z, \sim_Z) \to (\pi^{\mathsf{P}}(Y), =)$ such that the following commutes:



To this aim, we define

$$\lambda F := [\![\, \forall y : Y\, \forall \tau : \pi^{\mathsf{P}}(1)\, (\mathrm{ev}'(\langle y, \sigma\rangle, \tau) \leftrightarrow F(\langle y, z\rangle, \tau))\, [z : Z, \sigma : \pi^{\mathsf{P}}(Y)]\,]\!].$$

One can check λF is a functional relation, by exploiting the fact ev′ and F are. More interestingly, λF satisfies the required factorization property (the following formulae are all in context $[\langle y, z\rangle : Y \times Z, \tau : \pi^{\mathsf{P}}(1)]$, which we omitted for better readability):

$$\begin{aligned}
&\mathrm{ev}' \circ (1_{(Y,=)} \times \lambda F) = \\
&= [\![\, \exists \langle y', \sigma\rangle : Y \times \pi^{\mathsf{P}}(Y)\, ((1_{(Y,=)} \times \lambda F)(\langle y, z\rangle, \langle y', \sigma\rangle) \wedge \mathrm{ev}'(\langle y', \sigma\rangle, \tau))\,]\!] \\
&= [\![\, \exists \sigma : \pi^{\mathsf{P}}(Y)\, (\lambda F(z, \sigma) \wedge \mathrm{ev}'(\langle y, \sigma\rangle, \tau))\,]\!] \\
&= [\![\, \exists \sigma : \pi^{\mathsf{P}}(Y)\, (\forall y' : Y\, \forall \tau' : \pi^{\mathsf{P}}(1)\, (\mathrm{ev}'(\langle y', \sigma\rangle, \tau') \leftrightarrow F(\langle y', z\rangle, \tau')) \wedge \mathrm{ev}'(\langle y, \sigma\rangle, \tau))\,]\!] \\
&= [\![\, \exists \sigma : \pi^{\mathsf{P}}(Y)\, (F(\langle y, z\rangle, \tau))\,]\!] \\
&= F.
\end{aligned}$$

Finally, suppose $\lambda' F$ has the same factorization property of λF. We want to show $\lambda' F$ is equivalent to λF, which amounts to proving (notice we only need one direction of implication, thanks to $\lambda' F$ and λF being functional relations):

$$\mathsf{P} \models \lambda' F(z, \sigma) \to \forall y : Y\, \forall \tau : \pi^{\mathsf{P}}(1)\, (\mathrm{ev}'(\langle y, \sigma\rangle, \tau) \leftrightarrow F(\langle y, z\rangle, \tau))\, [z : Z, \sigma : \pi^{\mathsf{P}}(Y)].$$

Informally, one can derive the double implication on the right as follows: from $\lambda' F(z, \sigma)$ and $\mathrm{ev}'(\langle y, z\rangle, \tau)$, we get $\exists \sigma : \pi^{\mathsf{P}}(Y)\, (\lambda F(z, \sigma) \wedge \mathrm{ev}'(\langle y, \sigma\rangle, \tau))$, which gives $F(\langle y, z\rangle, \tau)$ by hypothesis on $\lambda' F$. In the other direction, $F(\langle y, z\rangle, \tau)$ implies $\exists \sigma' : \pi^{\mathsf{P}}(Y)\, (\lambda F(z, \sigma') \wedge \mathrm{ev}'(\langle y, \sigma'\rangle, \tau))$, but since we have $\lambda' F(z, \sigma)$ and $\lambda' F$ is functional, it must be $\sigma = \sigma'$, from which we derive $\mathrm{ev}'(\langle y, \sigma\rangle, \tau)$ as desired. □

Corollary 3.25.1. *Suppose* $[\![-]\!]$ *is an interpretation of a higher-order language* $\mathcal{L}$ *in* P*, and let* $\varphi\, [\vec{x} : \vec{\Gamma}]$ *be an* $\mathcal{L}$*-formula. Then there exists an interpretation* $[\![-]\!]'$ *of* $\mathcal{L}$ *in* $\mathrm{Sub}_{\mathbf{C}[\mathsf{P}]} \nabla_{\mathsf{P}}$ *such that:*

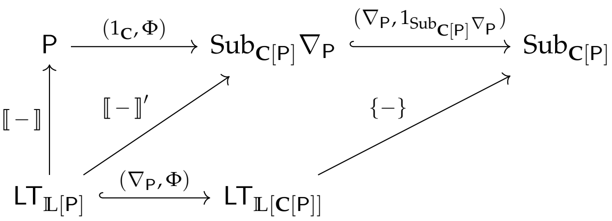
$$\mathsf{P} \vDash_{[\![-]\!]} \varphi\, [\vec{x} : \vec{\Gamma}] \quad \underline{\mathit{iff}} \quad \mathrm{Sub}_{\mathbf{C}[\mathsf{P}]} \nabla_{\mathsf{P}} \vDash_{[\![-]\!]'} \varphi\, [\vec{x} : \vec{\Gamma}].$$

Proof. This is the content of Corollary 3.21.1. □

Corollary 3.25.2. *Let* $\varphi\, [\vec{x} : \vec{\Gamma}]$ *be a well-formed formula in the language of* P*, let* $[\![-]\!]$ *be the canonical interpretation of* $\mathcal{L}[\mathsf{P}]$ *in* P*. Then* $\varphi\, [\vec{x} : \vec{\Gamma}]$ *is well-formed in the language of* $\mathbf{C}[\mathsf{P}]$ *(once we agree not to distinguish a type* $X : \mathbf{C}$ *from its representation* $\nabla_{\mathsf{P}}(X) : \mathbf{C}[\mathsf{P}]$*), the induced interpretation* $[\![-]\!]'$ *coincides with* $\mathbf{C}[\mathsf{P}]$*'s canonical one* $\{-\}$ *for these formulae, and*

$$\mathsf{P} \vDash_{[\![-]\!]} \varphi\, [\vec{x} : \vec{\Gamma}] \quad \underline{\mathit{iff}} \quad \mathbf{C}[\mathsf{P}] \vDash_{\{-\}} \varphi\, [\vec{x} : \vec{\Gamma}].$$

Proof. That formulae of $\mathcal{L}[\mathsf{P}]$ are also valid $\mathcal{L}[\mathbf{C}[\mathsf{P}]]$ formulae is straightforward. To say $[\![-]\!]'$ coincides with $\{-\}$ whenever both are defined is simply to ascertain the following diagram commutes, where hook arrows denote monic logical morphisms:



(3.5.2)

□

So far, we considered only the restriction of tripos $\mathrm{Sub}_{\mathbf{C}[\mathsf{P}]}$ to the subcategory of constant objects. Indeed, to seek a semantics from the full tripos $\mathrm{Sub}_{\mathbf{C}[\mathsf{P}]}$ to P is hopeless: $\mathbf{C}[\mathsf{P}]$ has just many more types and terms than $\mathbf{C}$. Nonetheless, $\mathbf{C}[\mathsf{P}]$ was built of of P-predicates, thus we can always express formulae in $\mathbf{C}[\mathsf{P}]$ using the language and logic of P (see [Van Oosten 2008, Section 2.3]).

For given a $\mathcal{L}[\mathbf{C}[\mathsf{P}]]$-formula $\varphi\, [x : (X, \sim_X)]$, consider its Kripke–Joyal interpretation, which realizes it as a subobject

$$\{\varphi\, [x : (X, \sim_X)]\} : (Y, \sim_Y) \rightarrowtail (X, \sim_X).$$

Suppose, without loss of generality, $\{\varphi\,[x : (X, \sim_X)]\}$ arises from a strict relation on $(X, \sim_X)$, which by abuse of notation we denote with the same name. Then that is the expression of $\varphi\,[x : (X, \sim_X)]$ in terms of P (recall Lemma 3.10):

$$R_{\{\varphi\,[x:(X,\sim_X)]\}} := \exists^{\mathsf{P}}_{\pi_X}\{\varphi\,[x : (X, \sim_X)]\} \in \mathsf{P}X \tag{3.5.3}$$

If $\psi\,[x : X]$ is an $\mathcal{L}[\mathsf{P}]$-formula such that

$$R_{\{\varphi\,[x:(X,\sim_X)]\}} \simeq [\![\, \psi\,[x : X]\,]\!], \tag{3.5.4}$$

then we'll write

$$\varphi\,[x : (X, \sim_X)] \rightsquigarrow \psi\,[x : X]. \tag{3.5.5}$$

To find such explicit syntactic representations, we now unfold (3.5.3) in various specific situations. In order to better understand the following, remember how we built $\mathbf{C}[\mathsf{P}]$ (cf. Definition 3.9), in particular have in mind (3.3.2) where we listed the axioms picking out morphisms $(X, \sim_X) \xrightarrow{F} (Y, \sim_Y)$ from predicates in $\mathsf{P}(X \times Y)$.

To start, one easily finds

$$\mathsf{true}\,[x : (X, \sim_X)] \rightsquigarrow x \sim_X x\,[x : X], \qquad \mathsf{false}\,[x : (X, \sim_X)] \rightsquigarrow \mathsf{false}\,[x : X]. \tag{3.5.6}$$

Then to translate formulae obtained as $S(t_1, \ldots, t_n)$ for a relation symbol[10] S of signature $(\vec{Z}, \vec{\sim}_Z)$, applied to terms (hence $\mathbf{C}[\mathsf{P}]$-arrows) $\vec{t} : (\vec{Z}, \vec{\sim}_Z)\,[x : (X, \sim_X)]$, one first recalls $\{S\}$ is a subobject of $(\prod_{i=1}^n Z_i, \sim_{\prod_{i=1}^n Z_i})$ which we can identify with the strict P-predicate by which it's represented. Then

$$\begin{aligned} &S(t_1, \ldots, t_n)\,[x : (X, \sim_X)] \\ &\qquad \rightsquigarrow \exists z : \prod_{i=1}^n Z_i\,(t_1(x, z_1) \wedge \cdots \wedge t_n(x, z_n) \wedge \{S\}(z))\,[x : X]. \end{aligned} \tag{3.5.7}$$

Disjunctions and conjunctions pass as-is:

$$\begin{aligned} \psi \wedge \eta\,[x : (X, \sim_X)] &\rightsquigarrow R_{\{\psi\,[x:(X,\sim_X)]\}} \wedge R_{\{\eta\,[x:(X,\sim_X)]\}}\,[x : X], \\ \psi \vee \eta\,[x : (X, \sim_X)] &\rightsquigarrow R_{\{\psi\,[x:(X,\sim_X)]\}} \vee R_{\{\eta\,[x:(X,\sim_X)]\}}\,[x : X]. \end{aligned} \tag{3.5.8}$$

Implication becomes

$$\psi \to \eta\,[x : (X, \sim_X)] \rightsquigarrow x \sim_X x \wedge R_{\{\psi\,[x:(X,\sim_X)]\}} \to R_{\{\eta\,[x:(X,\sim_X)]\}}\,[x : X]. \tag{3.5.9}$$

[10] The only relation symbols in $\mathcal{L}[\mathrm{Sub}_{\mathbf{C}[\mathsf{P}]}]$ are membership relations, which are translatable to $\mathcal{L}[\mathsf{P}]$.

Finally, quantifiers become

$$\begin{aligned}
&\exists z\!:\!(Z,\sim_Z)\,\psi(x,z)\,[x\!:\!(X,\sim_X)] \\
&\qquad\qquad \rightsquigarrow \exists z\!:\!Z\,R_{\{\psi\,[z:(Z,\sim_Z),\,x:(X,\sim_X)]\}}(z,x)\,[x\!:\!X], \\
&\forall z\!:\!(Z,\sim_Z)\,\psi(x,z)\,[x\!:\!(X,\sim_X)] \\
&\qquad\qquad \rightsquigarrow x \sim_X x \wedge \forall z\!:\!Z\,(z \sim_Z z \to R_{\{\psi\,[z:(Z,\sim_Z),\,x:(X,\sim_X)]\}}(z,x))\,[x\!:\!X].
\end{aligned} \tag{3.5.10}$$

Chapter 4

Scott triposes

In his article [Scott 1967], Scott explains a different point of view on Cohen's proof of independence of the continuum hypothesis. Following a discovery of Solovay, he notes Cohen's proof amounts to the construction of a 'Boolean valued' model of set theory. The logic of such models is exactly that of $\mathbf{Set}[\mathbb{B}^{(-)}]$ for the chosen complete Boolean algebra $\mathbb{B}$. In particular, Scott employs a Boolean algebra obtained from the σ-field of a probability space.

For the rest of the chapter, fix an arbitrary probability space $\underline{\Omega} = (\Omega, \mathcal{F}, \mathbb{P})$.

4.1 Essential algebras

Consider, on the σ-field $\mathcal{F}$, the order relation given by **essential inclusion**:

$$A \leq B \quad \underline{\text{iff}} \quad A \setminus B \in \ker \mathbb{P}.$$

We call $(\mathcal{F}, \leq)$ the **essential prealgebra** of $\underline{\Omega}$. Clearly, it inherits all the Boolean operations of $\mathcal{F}$, plus the countable joins. It is therefore a σ-Boolean prealgebra.

Define now $\mathcal{F} / \ker \mathbb{P}$ to be the posetal reflection of $(\mathcal{F}, \leq)$, i.e. the set of measurable subsets of $\underline{\Omega}$ considered up to **essential**[1] **equivalence**:

$$A \simeq B \quad \underline{\text{iff}} \quad A \leq B \text{ and } B \leq A \quad \underline{\text{iff}} \quad A \Delta B \in \ker \mathbb{P}.$$

We are to show $\mathcal{F} / \ker \mathbb{P}$ is a *complete* Boolean algebra, which we'll call the **essential algebra** of $\underline{\Omega}$. When $\mathbb{P}$ is unambiguously clear from the context, we denote $\mathcal{F} / \ker \mathbb{P}$ as $\mathbb{F}$.

[1] A proposition about measurable sets *holds essentially* if it holds up to a zero-measure modification of the sets under consideration, i.e. up to change of representatives in their essential equivalence classes.

Remark 4.0.1. One might notice the essential algebra $\mathbb{F}$ is complete if and only if the essential prealgebra $(\mathcal{F}, \leq)$ is complete, thus one might as well stick with the first instead of quotienting out (potentially precious) information by passing to the latter. What we'd lose in doing so, though, is the strictness of all definitions involving meets and joins, since these, in a prealgebra, are defined by universal properties. Hence they do not uniquely define an element of the prealgebra, but only pinpoint its posetal equivalence class.

Definition 4.1. A subset A of a lattice $(\mathbb{L}, \leq)$ is an **antichain** if $\perp \notin A$ and $a \wedge b = \perp$ for every pair of distinct $a, b \in A$. $\mathbb{L}$ satisfies the **countable chain condition** (**CCC**) if any antichain in $\mathbb{L}$ is at most countable.

Lemma 4.2 (Bell 2005, p. 150)**.** *A σ-algebra satisfying the CCC is complete.*

Proof. Let $(\mathbb{B}, \leq)$ be such a Boolean algebra. Suppose $\{x_i\}_{i \in I}$ is a family of elements indexed by an arbitrary set I. Denote by M the ideal generated by this family, that is, the set of elements of $\mathbb{B}$ covered by a finite join of elements of the family:

$$M = \{y \in \mathbb{B} \mid \exists x_{i_1}, \ldots, x_{i_n} (y \leq x_{i_1} \vee \cdots x_{i_n})\}.$$

Clearly, $\bigvee_i x_i = \bigvee M$, if they exist.

Consider now the family $\mathcal{A}$ of subsets of M which are antichains. The family is partially ordered by inclusion, and if $\mathcal{U}$ is a non-empty totally ordered subset of $\mathcal{A}$, then $U := \bigcup \mathcal{U}$ is an upper bound for $\mathcal{U}$. Therefore Zorn's lemma implies the existence of a maximal antichain F contained in M. By the CCC on $\mathbb{B}$, observe that F is at most countable, and therefore admits a supremum $\bigvee F$.

It remains to prove that $\bigvee M = \bigvee F$. Since $F \subseteq M$, $\bigvee F \leq \bigvee M$ trivally. Thus, we'll prove that if p is not an upper bound for M then it isn't an upper bound for F either. In fact, under this hypothesis, there exists an $m \in M$ such that $m \not\leq p$, which implies $m \wedge \neg p \neq \perp$. Obviously, $m \wedge \neg p \in M$ (because M is an ideal and $m \in M$). Then either $m \wedge \neg p \in F$, and since $m \wedge \neg p \not\leq p$, p is not an upper bound of F; or $m \wedge \neg p \notin F$, and by maximality of F there exists $q \in F$ such that $q \wedge m \wedge \neg p \neq \perp$, from which it follows that $q \wedge m \wedge \neg p \not\leq p$, so that p is not an upper bound of F. □

Lemma 4.3. *The essential algebra of a probability space satisfies the CCC.*

Proof. Let $\mathcal{A}$ be an antichain of $\mathbb{F}$. This means $\mathcal{A}$ is a family of essentially non-empty, essentially pairwise disjoint sets. Define

$$\mathcal{A}_n := \{[E] \in \mathcal{A} \mid \mathbb{P}(E) > 1/n\}$$

Since $\mathbb{P}$ is unitary, $\mathcal{A}_n$ has at most n elements. Thus $\mathcal{A} = \bigcup_{n\in\mathbb{N}} \mathcal{A}_n$ is at most countable. □

It follows that:

Theorem 4.4. *The essential algebra of a probability space is complete.*

We can thus define the central object of this chapter:

Definition 4.5. Given a probability space $\underline{\Omega} = (\Omega, \mathcal{F}, \mathbb{P})$, its associated **Scott tripos** is the **Set**-tripos $\mathbb{F}^{(-)}$. We denote its Pitts' topos (Definition 3.9) by

$$\mathbf{Scott}[\underline{\Omega}] := \mathbf{Set}[\mathbb{F}^{(-)}]$$

and we call it the **Scott topos** associated to the probability space $\underline{\Omega}$.

Is every complete Boolean algebras the essential algebra of some probability space? Any complete Boolean algebra $\mathbb{B}$ can be obtained by quotienting a σ-field by one of its σ-ideals. In particular, from $\mathbb{B}$ we can construct its *Stone space* and equip it with the σ-field of its clopen sets [Scott 1967]. Dividing this by the ideal of its meager sets yields an algebra isomorphic to $\mathbb{B}$. This seems to settle the question.

On the other hand, we defined essential algebras starting with spaces equipped with a probability measure, not just a σ-ideal. Hence one may ask: **can we endow every measurable space with a distinguished σ-ideal with a probability measure whose σ-ideal of null sets coincides with the one given?** This does not seem to be the case.

For example, let X be an infinite set considered with its discrete σ-field $\mathcal{P}(X)$. Suppose we fixed $\{\varnothing\}$ as σ-ideal and now we want to find a probability measure $\mathbb{P}$ on $\mathcal{P}(X)$ such that $\ker \mathbb{P} = \{\varnothing\}$. In that case, $\mathbb{P}(\{x\}) > 0$ for each $x \in X$, and a theorem of Ulam [Jech 2006, Theorem 10.1] says that if $\mathbb{P}$ is such a measure, then either

1. X has cardinality greater or equal to that of the least inaccessible cardinal

   or

2. the reals have cardinality greater or equal to that of the least weakly inaccessible cardinal.

In both case, we would have proven the existence of a weakly inaccessible cardinal, which is not possible in ZFC (at least, if it is consistent). Hence such a measure cannot be proven to exist in our chosen metatheory.

4.2 A bird-eye view of the construction

In the preceding paragraphs, we gave a construction starting from a probability space and yielding a topos, passing through a tripos. We now ponder on the functoriality of such constructions, and highlight some connections with the ideas in [Adachi and Ryu 2016] and [Jackson 2006].

The first remark we make is about the relationship between the category $\mathbf{Prob}_0$ (the category of probability spaces and null-preserving measurable maps, see Definition 2.13) and our construction. It is now clear that such a morphism has a compelling logical interpretation, since a null-preserving measurable map $(X, \Sigma_X, \mathbb{P}_X) \xrightarrow{f} (Y, \Sigma_Y, \mathbb{P}_Y)$ has all it's needed to induce a well-defined map between the essential algebras of X and Y:

$$f^{-1} : \Sigma_Y / \ker \mathbb{P}_Y \longrightarrow \Sigma_X / \ker \mathbb{P}_X.$$

Therefore, the category $\mathbf{Prob}_0$ can be thought as a category of probability spaces and 'logically relevant' morphisms between them, corroborating further the idea that such morphisms 'preserve knowledge' or 'information'.

Our construction recalls another instance (actually *the* instance) of such 'space-to-topos' construction, namely the construction of a topos of sheaves on a topological space. Similarly to our case, the construction begins by shifting the focus from the space X to its *frame* of open sets $\mathcal{O}(X)$. In doing this, continuous maps $X \to Y$ become maps of frames $\mathcal{O}(Y) \to \mathcal{O}(X)$ in the opposite direction. Hence if we want to replace X with $\mathcal{O}(X)$ (i.e. if we want to do topology 'without points'), we need to reverse maps again:

topological spaces	**Top**	$X \xrightarrow{f} Y$
frames	**Frm**	$\mathcal{O}(X) \xleftarrow{f^{-1}} \mathcal{O}(Y)$
locales	$\mathbf{Loc} := \mathbf{Frm}^{\mathrm{op}}$	$\mathcal{O}(X) \xrightarrow{f} \mathcal{O}(Y)$
topoi	**Topos**	$\mathrm{Sh}\, X \xrightarrow{f^* \dashv f_*} \mathrm{Sh}\, Y$

The category $\mathbf{Prob}_0$ is a natural candidate for the domain of a functor $\mathbf{Scott}[-]$ taking a probability space $\underline{\Omega}$ to its Scott topos $\mathbf{Scott}[\underline{\Omega}]$. However, this functor should factorize through various other functors, describing the steps of the construction. The first of these should be Ess:

$$\begin{array}{ccc} \mathbf{Prob}_0^{\mathrm{op}} & \xrightarrow{\text{Ess}} & \mathbf{CBool} \\ \\ (X, \Sigma_X, \mathbb{P}_X) & & \Sigma_X / \ker \mathbb{P}_X \\ f \big\downarrow & \longmapsto & \big\uparrow f^{-1} \\ (Y, \Sigma_Y, \mathbb{P}_Y) & & \Sigma_Y / \ker \mathbb{P}_Y \end{array} \tag{4.2.1}$$

The category **CBool** is the category of complete Boolean algebras and maps between them preserving all joins and negations. Then we compose this with the functor $(=)^{(-)}$:

$$\begin{array}{ccc} \mathbf{CBool}^{\mathrm{op}} & \xrightarrow{(=)^{(-)}} & \mathbf{Tripos}_{\mathbf{Set}} \\ \\ \mathbb{A} & & \mathbb{A}^{(-)} \\ f \big\downarrow & \longmapsto & \big\uparrow f^{+} \dashv f_{+} \\ \mathbb{B} & & \mathbb{B}^{(-)} \end{array} \tag{4.2.2}$$

where the adjoint pair of natural transformations $f^+ \dashv f_+$ is induced by the Adjoint Functor Theorem for posets (Theorem 1.7):

$$\begin{aligned} (f^+)_X : \mathbb{A}^X &\longrightarrow \mathbb{B}^X \\ \alpha &\longmapsto f \circ \alpha. \\ (f_+)_X : \mathbb{B}^X &\longrightarrow \mathbb{A}^X \\ \beta &\longmapsto \bigvee \{\alpha \in \mathbb{A}^X \mid f \circ \alpha \leq \beta\}. \end{aligned} \tag{4.2.3}$$

Clearly f^+ preserves meets since f is a map of **CBool**, although this can be strengthened further by observing it has a further left adjoint, namely

$$\begin{aligned} (f^!)_X : \mathbb{B}^X &\longrightarrow \mathbb{A}^X \\ \beta &\longmapsto \bigwedge \{\alpha \in \mathbb{A}^X \mid \beta \leq f \circ \alpha\}. \end{aligned} \tag{4.2.4}$$

By Remark 3.19.1, f^+ is a a first-order logical morphism of triposes (although not higher-order, since it doesn't preserve weak power objects in general).

Finally, Pitts' functor $\mathbf{Set}[-]$ gives us topoi

$$\begin{array}{ccc} \mathbf{Tripos}_{\mathbf{Set}} & \xrightarrow{\mathbf{Set}[-]} & \mathbf{Topos} \\ \\ \mathsf{P} & & \mathbf{Set}[\mathsf{P}] \\ f^{+} \dashv f_{+} \big\downarrow & \longmapsto & \big\downarrow f^{*} \dashv f_{*} \\ \mathsf{Q} & & \mathbf{Set}[\mathsf{Q}] \end{array} \tag{4.2.5}$$

where a geometric morphism of **Set**-triposes induces a geometric morphism of topoi by the construction of Theorem 3.18 (whose functoriality is trivial).

All in all, we arrived at the following factorization:

$$\mathbf{Prob}_0 \xrightarrow{\mathrm{Ess}^{\mathrm{op}}} \mathbf{CBool}^{\mathrm{op}} \xrightarrow{(=)^{(-)}} \mathbf{Tripos}_{\mathbf{Set}} \xrightarrow{\mathbf{Set}[-]} \mathbf{Topos} \qquad \text{with } \mathbf{Scott}[-] : \mathbf{Prob}_0 \to \mathbf{Topos} \tag{4.2.6}$$

Getting back to the analogy with topological spaces and their topoi of sheaves, we can represent the abstraction hierarchy of Scott's construction in the same vein:

probability spaces	$\mathbf{Prob}_0$	$\underline{\Omega} \xrightarrow{f} \underline{\Omega}'$
complete Boolean algebras	$\mathbf{CBool}$	$\mathrm{Ess}\,\underline{\Omega} \xleftarrow{f^{-1}} \mathrm{Ess}\,\underline{\Omega}'$
triposes	$\mathbf{Tripos}_{\mathbf{Set}}$	$\mathrm{Ess}\,\underline{\Omega}^{(-)} \xrightarrow{f^+ \dashv f_+} \mathrm{Ess}\,\underline{\Omega}'^{(-)}$
topoi	$\mathbf{Topos}$	$\mathbf{Scott}[\underline{\Omega}] \xrightarrow{f^* \dashv f_*} \mathbf{Scott}[\underline{\Omega}']$

The sheaf-theoretic side. In [Jackson 2006], Jackson shows there is a topology on the topos $\mathrm{Psh}\,\mathcal{F}$ whose sheaves form the topos $\mathrm{Sh}(\mathcal{F}/\ker\mathbb{P})$. This topos is equivalent to the Scott topos of $\underline{\Omega}$, by Theorem 3.13. In his work, Jackson formulates a 'sheaf-theoretic' version of measure theory, by generalizing many constructions and theorems to arbitrary locale and then expressing them in terms of sheaf theory.

Our work, although related and inspired by Jackson's, is going to differ since our goal is not to recast notions from stochastic calculus in tripos-theoretic language, but to relativize them in the peculiar setting of Scott topoi.

We also mention Wendt's [Wendt 1996], in which the topos $\mathrm{Sh}(\mathcal{F}/\ker\mathbb{P})$ is considered and some of its basic properties are analyzed. Wendt is trying to develop categorical tools to talk about fibrations in a measure-theoretic setting, in particular regarding Hilbert spaces.

4.3 Some facts about Scott topoi

We now outline some of the topos-theoretic and logical properties of Scott topoi. To avoid confusion between the probability space $\underline{\Omega} = (\Omega, \mathcal{F}, \mathbb{P})$ and the subobject classifier

$\Omega_{\mathbf{Scott}[\underline{\Omega}]}$ of $\mathbf{Scott}[\underline{\Omega}]$, we denote the latter by Σ, and its equality relation by $\Leftrightarrow$ (see proof of Theorem 3.11).

1. The localic reflection of $\mathbf{Scott}[\underline{\Omega}]$ is $\mathbb{F}$, i.e.

$$\mathrm{Sub}\, 1 = \mathbb{F} = \mathcal{F} / \ker \mathbb{P}.$$

   This is a general fact about localic topoi [Moerdijk and MacLane 1992, p. 490].

2. We punctualize Σ is *not* isomorphic to $\nabla\mathbb{F}$ in general. This because $\Leftrightarrow$ is a truly different equality predicate in most of the cases.

3. **It's Boolean**, meaning for every proposition φ in the language of $\mathbf{Scott}[\underline{\Omega}]$,

$$\mathbf{Scott}[\underline{\Omega}] \models \varphi \vee \neg\varphi.$$

   This is equivalent to the following, maybe surprising fact:

$$1 + 1 \cong \Sigma.$$

   The isomorphism is given by

$$\begin{aligned} F : \{0,1\} \times \mathbb{F} &\longrightarrow \mathbb{F} \\ (1, A) &\longmapsto A \\ (0, A) &\longmapsto \neg A \end{aligned}$$

   This means that, from the internal point of view, $\mathbf{Scott}[\underline{\Omega}]$ has only two logical values, i.e.

$$\mathbf{Scott}[\underline{\Omega}] \models \exists x{:}\Sigma\, \exists y{:}\Sigma\, \forall z{:}\Sigma\, (z \Leftrightarrow x \vee z \Leftrightarrow y).$$

   In fact

$$\begin{aligned} [\![\, \exists x{:}\Sigma\, \exists y{:}\Sigma\, \forall z{:}\Sigma\, (z \Leftrightarrow x \vee z \Leftrightarrow y) \,]\!] &= \bigvee_{x,y \in \Sigma} \bigwedge_{z \in \Sigma} (z \Leftrightarrow x \vee z \Leftrightarrow y) \\ &\supseteq \bigwedge_{z \in \Sigma} (z \Leftrightarrow \top \vee z \Leftrightarrow \bot) \\ &= \bigwedge_{z \in \Sigma} (z \vee \neg z) \\ &= \top. \end{aligned}$$

4. **It's *not* two-valued** in general, meaning Σ has more than two global elements. In fact, for any $A \in \mathbb{F}$, the following is a global element:

$$\begin{aligned} \ulcorner A \urcorner : 1 \times \Sigma &\longrightarrow \mathbb{F} \\ B &\longmapsto A \Leftrightarrow B \end{aligned} \tag{4.3.1}$$

Since any well-pointed topos (i.e. non-degenerate and generated by 1) is two-valued, this also shows **Scott**$[\underline{\Omega}]$ isn't well-pointed.

5. **It satisfies the axiom of choice**, in both the external and internal sense. The external axiom of choice (eAC) is satisfied by a topos $\mathcal{E}$ iff every surjective map $X \twoheadrightarrow Y$ admits a section, while the internal axiom of choice (iAC) is satisfied iff, for every object X and Y:

$$\mathcal{E} \models \forall f \in Y^X \, ((\forall y \in Y \, \exists x \in X \, f(x) = y) \to (\exists g \in X^Y \, \forall y \in Y \, f(g(y)) = y)).$$

   This is again the statement that every surjective map admits a section, although in the logic of $\mathcal{E}$. It's easy to see eAC entails iAC, while the opposite isn't true, since the internal logic of a topos can be weaker than the external one.

   Then **Scott**$[\underline{\Omega}]$ satisfies both versions of AC since it is Boolean and localic over **Set**, and these properties characterize topoi satisfying eAC [Johnstone 2002, Theorem D4.5.15].

4.4 Real numbers in a Scott topos

In a way, the Scott tripos of a probability space encodes the logic naturally attached to it. Therefore, the Scott topos is the natural environment to do mathematics dependent on such a stochastic context.

A confirmation of this is given by the mathematical practice of treating $\underline{\Omega}$ as an underlying unknown space whose only real purpose is to imbue random variables with randomness. Moreover, such random variables are informally treated as numbers, except they behave in an unusual way. That's usually where the analogy breaks down and mathematicians stop to pretend, resorting to probability and measure theory to properly handle such objects. However, the discrepancy between random variables and real numbers dissolves if we work inside the Scott topos of $\underline{\Omega}$.

4.4.1 Number systems in a topos

Since a topos has a rich internal language, one can carry out the usual definitions of $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$. The only missing ingredient is $\mathbb{N}$, which must be defined extrinsically.

Definition 4.6. A **natural numbers object** (**NNO** for short) in a topos $\mathcal{E}$ is an object

$\mathcal{N}_{\mathcal{E}} : \mathcal{E}$ equipped with two maps

$$\begin{array}{c} \mathcal{N}_{\mathcal{E}} \xrightarrow{\text{succ}} \mathcal{N}_{\mathcal{E}} \\ 1 \xrightarrow{0} \mathcal{N}_{\mathcal{E}} \end{array} \tag{4.4.1}$$

satisfying the following universal property: for every other $X : \mathcal{E}$ which comes equipped with maps $1 \overset{z}{\to} X \overset{s}{\to} X$, there is a unique morphism $\mathcal{N}_{\mathcal{E}} \overset{\varphi}{\to} X$ such that the following commutes:

$$\begin{array}{ccccc} & \overset{0}{\nearrow} & \mathcal{N}_{\mathcal{E}} & \xrightarrow{\text{succ}} & \mathcal{N}_{\mathcal{E}} \\ 1 & & \downarrow \varphi & & \downarrow \varphi \\ & \underset{z}{\searrow} & X & \xrightarrow{s} & X \end{array} \tag{4.4.2}$$

Remark 4.6.1. This definition is actually an induction schema in disguise. In fact, φ is exactly the function such that

$$\varphi(0) = z, \qquad \varphi(\mathsf{succ}\, n) = s(\varphi(n)) \quad \text{for all } n : \mathcal{N}_{\mathcal{E}}.$$

Remark 4.6.2. Not every (elementary) topos has a natural numbers object, for instance, **FinSet** does not have one. Anyway, all Grothendieck topoi, such as **Scott**$[\underline{\Omega}]$, have a NNO. It is obtained as the (externally) countable coproduct of 1:

$$\mathcal{N} \cong \coprod_{n \in \mathbb{N}} 1.$$

As anticipated, one might start from $\mathcal{N}$ to build an internal version of $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$. For $\mathbb{Z}$ and $\mathbb{Q}$, one simply carries out the same construction. For example, the **integer numbers object** $\mathcal{Z}$ can be defined as the quotient of $\mathcal{N}^2$ under the relation

$$\{n + n' = m + m' \; [n, n', m, m' : \mathcal{N}]\} \hookrightarrow \mathcal{N}^2 \times \mathcal{N}^2.$$

Since topoi admit quotients, this is a perfectly good definition for $\mathcal{Z}$. $\mathcal{Q}$ is then defined as the field of fractions of $\mathcal{Z}$.

Real numbers are subtler though, since there a bunch of constructions for them which are equivalent only under some assumptions. The two most important ones are Dedekind's and Cantor's constructions.

Dedekind defined the real numbers to be *cuts* in the rational numbers, i.e. pairs (L, R) of subsets of $\mathbb{Q}$ which should correspond to the set of rationals to the left and to the right of the definiendum. This is suitably axiomatized by requirements on the sets L and R. This construction of $\mathcal{R}$ is outlined in [Moerdijk and MacLane 1992, Section VI.8], and yields an object which is aptly named the **Dedekind real numbers object**.

Cantor instead defines the real numbers as equivalence classes of Cauchy sequences of rational numbers, under the equivalence which identifies sequences 'converging to the same point'. This version of $\mathcal{R}$ is the **Cauchy real numbers object**. To prove this construction is equivalent to the other one, one needs a choice principle. Hence, in a general topos, Cauchy real numbers are not the same as Dedekind real numbers.

However, in our case this is not a problem since we've seen $\mathbf{Scott}[\underline{\Omega}]$ is Boolean and satisfies the axiom of choice. Hence we will not distinguish between Dedekind or Cauchy real numbers, and we will simply refer to them as 'real numbers'.

4.4.2 Number systems in Scott topoi

Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Theorem 4.7. *The topos* $\mathbf{Scott}[\underline{\Omega}]$ *has a natural numbers object* $\mathcal{N}_{\underline{\Omega}}$ *given by*

$$\mathcal{N}_{\underline{\Omega}} = \nabla(\mathbb{N}).$$

Proof. See [Wendt 1996, Section 4]. □

Recall $\nabla : \mathbf{Set} \to \mathbf{Scott}[\underline{\Omega}]$ is the functor of constant objects, which takes a set X and endows it with an internalization $=$ of external equality for which:

$$[\![\, n = m \,]\!] = \begin{cases} \Omega & \text{if } n = m \\ \varnothing & \text{else.} \end{cases}$$

Theorem 4.8. *The topos* $\mathbf{Scott}[\underline{\Omega}]$ *has a (Dedekind) real numbers object* $\mathcal{R}_{\underline{\Omega}}$ *given by*

$$\mathcal{R}_{\underline{\Omega}} := \nabla_{\mathbb{F}(-)}(\mathbb{R}) = (\mathbb{R},\ =)$$

Proof. See [Wendt 1996, Section 4], where it is also proved these are isomorphic to Cauchy real numbers. □

Proposition 4.9. *Equip* $\mathbb{R}$ *with its Borel* σ*-field. Then the object*

$$\mathcal{R}'_{\underline{\Omega}} := (\mathbf{Msbl}(\underline{\Omega},\ \mathbb{R}),\ \{- = -\}), \tag{4.4.3}$$

is isomorphic to $\mathcal{R}_{\underline{\Omega}}$*; where* $\{X = Y\}$ *denotes the (equivalence class of the) measurable subset of* Ω *on which* X *and* Y *coincide.*

Proof. One can easily show the following is a well-defined arrow which is an isomorphism from $\mathcal{R}'_{\underline{\Omega}}$ to $\mathcal{R}_{\underline{\Omega}}$ in **Scott**[$\underline{\Omega}$]:

$$\begin{aligned} F : \mathbf{Msbl}(\underline{\Omega}, \mathbb{R}) \times \mathbb{R} &\longrightarrow \mathbb{F} \\ (X, x) &\longmapsto \{X = x\} \end{aligned} \tag{4.4.4}$$

In fact, recalling the identity of an object is given by its very equality predicate, we check:

$$\begin{aligned} F \circ F^{-1} &= [\![\, \exists X : \mathbf{Msbl}(\Omega, \mathbb{R})\, (F^{-1}(x, X) \wedge F(X, y))\, [x, y : \mathcal{R}_{\underline{\Omega}}] \,]\!] \\ &= (x, y) \mapsto \bigvee_{\Omega \xrightarrow{X} \mathbb{R}} \{X = x\} \wedge \{X = y\} \\ &= (x, y) \mapsto \{x = y\} \\ &= 1_{\mathcal{R}_{\underline{\Omega}}} \end{aligned}$$

$$\begin{aligned} F^{-1} \circ F &= [\![\, \exists x : \mathcal{R}_{\underline{\Omega}}\, (F(X, x) \wedge F^{-1}(x, Y))\, [X, Y : \mathbf{Msbl}(\Omega, \mathbb{R})] \,]\!] \\ &= (X, Y) \mapsto \bigvee_{x \in \mathbb{R}} \{X = x\} \wedge \{Y = x\} \\ &= (X, Y) \mapsto \{X = Y\} \\ &= 1_{(\mathbf{Msbl}(\Omega, \mathbb{R}),\, \{-=-\})}. \end{aligned}$$

□

Remark 4.9.1. Notice $\{X = Y\} = \Omega$ if and only if X and Y are almost-everywhere equal.

Remark 4.9.2. One can also easily check that the arithmetic operations induced on $\mathcal{R}^m_{\underline{\Omega}}$ by the isomorphism with $\mathcal{R}_{\underline{\Omega}}$ are determined by the relative componentwise operations.

Remark 4.9.3. The previous proof actually works for any algebra of measurable functions over $\underline{\Omega}$ containing all the constants and equipped with almost-everywhere equality. In particular, all the L^p spaces become isomorphic to $\mathcal{R}$.

We conclude with an application.

Definition 4.10. A sequence of random variables $\{X_n : \underline{\Omega} \to \mathbb{R}\}_{n \in \mathbb{N}}$ **converges almost surely** to the random variable $X : \underline{\Omega} \to \mathbb{R}$ if it converges pointwise almost everywhere:

$$X_n \underset{n}{\longrightarrow} X \text{ a.e.}$$

We can think of a sequence of random variables $\{X_n\}_{n \in \mathbb{N}}$ it as the function

$$\begin{aligned} X : \nabla\mathbb{N} \times \mathcal{R} &\longrightarrow \mathbb{F} \\ (n, Y) &\longmapsto \{X_n = Y\} \end{aligned}$$

which defines an element of $\mathcal{R}_{\underline{\Omega}}^{m\mathcal{N}_\Omega}$ which entirely belongs to it. Moreover, X is an element of $\mathcal{R}_{\underline{\Omega}}^m$ which belongs totally to it. After we have noticed these facts, the statement of the following theorem becomes clear.

Theorem 4.11. *A sequence of random variables on $\underline{\Omega}$ is almost-surely convergent if and only if it is a convergent sequence of real numbers in* **Scott**$[\underline{\Omega}]$.

Proof. Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of random variables over $\underline{\Omega}$ that converges pointwise to the random variable X. We have to show:

$$\mathbf{Scott}[\underline{\Omega}] \models \exists n : \mathcal{N}_{\underline{\Omega}} \,\forall n' : \mathcal{N}_{\underline{\Omega}} \,(n' > n \to |X_{n'} - X| < 1/m)\,[m : \mathcal{N}_{\underline{\Omega}}]. \tag{4.4.5}$$

We denote with $[\![-]\!]$ the canonical interpretation of predicates in the Scott tripos of $\underline{\Omega}$. Thus we have to show that for every $m \in \mathbb{N}$:

$$\bigvee_{n\in\mathbb{N}} \bigwedge_{n'\in\mathbb{N}} \left([\![n' \leq n]\!] \vee [\![|X_{n'} - X| < 1/m]\!]\right) \simeq \Omega$$

This is equivalent to requiring that

$$P\left(\bigcup_{n\in\mathbb{N}} \bigcap_{n'\in\mathbb{N}} \left(\{\omega \in \Omega |\, n' \leq n\} \cup \{\omega \in \Omega |\, |X_{n'}(\omega) - X(\omega)| < 1/m\}\right)\right) = 1. \tag{4.4.6}$$

If X_n converges to X almost surely, then there exists $\overline{n}$ such that for every $n' \geq \overline{n}$,

$$\mathbb{P}(\{\omega \in \Omega |\, |X_{n'}(\omega) - X(\omega)| < 1/m\}) = 1.$$

In particular, (4.4.6) holds.

Conversely, we can notice that (4.4.6) is equivalent to

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} \bigcap_{n'>n} \left(\{\omega \in \Omega |\, |X_{n'}(\omega) - X(\omega)| < 1/m\}\right)\right) = 1.$$

that is

$$\mathbb{P}\left(\{\omega \in \Omega |\, \exists n \in \mathbb{N}\, \forall n' \in \mathbb{N}(n' > n \to |X_{n'}(\omega) - X(\omega)| < 1/m))\}\right) = 1$$

If we consider the intersection of all the sets

$$\{\omega \in \Omega |\, \exists n \in \mathbb{N}\, \forall n' \in \mathbb{N}(n' > n \to |X_{n'}(\omega) - X(\omega)| < 1/m)\}$$

as m ranges in $\mathbb{N}$, we get a set of probability 1 which is the set

$$\{\omega \in \Omega |\, \forall m \in \mathbb{N}\, \exists n \in \mathbb{N}\, \forall n' \in \mathbb{N}(n' > n \to |X_{n'}(\omega) - X(\omega)| < 1/m)\}.$$

This means precisely that $\{X_n\}_{n\in\mathbb{N}}$ converges almost surely to X. □

4.4.3 A possible internalization of stochastic processes

One may now try to understand whether we arrived to a point at which we can start to internalize the theory of stochastic processes. Indeed, one way to see a (say, real-valued) stochastic process is as a collection of random variables over $\underline{\Omega}$ indexed by a time monoid I. This can be packaged into a function $I \to \mathbf{Msbl}(\underline{\Omega}, \mathbb{R})$. This suggests a stochastic process is a map $\nabla I \to \mathcal{R}_{\underline{\Omega}}$ in $\mathbf{Scott}[\underline{\Omega}]$, i.e. a real-valued function from the internal point of view.

Theorem 4.12. *Let $\{X_t : \underline{\Omega} \to \mathbb{R}\}_{t \in I}$ be a stochastic process. Then it defines a map $\nabla I \to \mathcal{R}_{\underline{\Omega}}$ in $\mathbf{Scott}[\underline{\Omega}]$, defined by*

$$\begin{aligned} X : I \times \mathbf{Msbl}(\underline{\Omega}, \mathbb{R}) &\longrightarrow \mathbb{F} \\ (t, Y) &\longmapsto \{X_t = Y\} \end{aligned} \tag{4.4.7}$$

Proof. We only have to show X is a well-defined map, hence verify conditions (3.3.2):

1. It is strict, since
$$X(t, Y) \leq [\![t = t]\!] \wedge \{Y = Y\} = \Omega.$$

2. It is relational, since
$$\begin{aligned} X(t, Y) \wedge [\![t = s]\!] \wedge \{Y = Y'\} &= \{X_t = Y\} \wedge [\![t = s]\!] \wedge \{Y = Y'\} \\ &\leq \{X_s = Y'\} \\ &= X(s, Y'). \end{aligned}$$

3. It is single valued, since
$$X(t, Y) \wedge X(t, Y') = \{X_t = Y\} \wedge \{X_t = Y'\} \leq \{Y = Y'\}.$$

4. It is total, since
$$\Omega = [\![t = t]\!] \leq \{X_t = X_t\} \leq \bigvee_{\underline{\Omega} \xrightarrow{Y} \mathbb{R}} \{X_t = Y\}.$$

□

Remark 4.12.1. An arrow $X : \nabla I \to \mathcal{R}_{\underline{\Omega}}$ determines a stochastic process in the external sense: one picks X_t to be the random variable $Y : \underline{\Omega} \to \mathbb{R}$ given by totality of X. In fact consider the interpretation of such clause in the canonical semantics of $\mathbf{Scott}[\underline{\Omega}]$: one gets

$$\Omega \leq \bigvee_{\underline{\Omega} \xrightarrow{Y} \mathbb{R}} X(t, Y).$$

This means Ω can be covered by a family of events over which '$X(t) = Y$', for Y possibly changing for each event in the covering. This cover, by virtue of the CCC property of $\mathbb{F}$, can be refined to a countable antichain, i.e. an essential, countable partition of Ω. Therefore we can patch together the definition of X_t over each of these pairwise disjoint sets, and get a global definition[2], as desired. Notice this process defines X_t in an essentially unique way.

Example 4.13. A common requirement for a process is almost-everywhere (right, left, or both) continuity, which means almost-all realizations of X_t are continuous.

Suppose then $I = \mathbb{R}^+$, so that $\nabla I = (\mathbb{R}^+, =) = \mathcal{R}^+$, as shown before. Then continuity trivially translates to the following requirement:

$$\mathbf{Scott}[\underline{\Omega}] \models \ulcorner X : I \to \mathcal{R}^+ \text{ is a continuous map} \urcorner.$$

Remark 4.13.1. As much as $\nabla\mathbb{R}$ is isomorphic to $(\mathbf{Msbl}(\underline{\Omega}, \mathbb{R}), \{- = -\})$, so ∇I is isomorphic to $(\mathbf{Msbl}(\underline{\Omega}, I), \{- = -\})$. Hence the function X can be defined not only on deterministic times, but also on 'random times' $\tau : \underline{\Omega} \to I$. One can compute these values by precomposing X with the isomorphism $(\mathbf{Msbl}(\underline{\Omega}, I), \{- = -\}) \xrightarrow{G} \nabla I$, which is analogous to the one given for real numbers (4.4.4) (we abuse notation by calling X both X as defined in (4.4.7) and its precomposition with G):

$$\begin{aligned} [\![\, X(\tau, Y) \,]\!] &= [\![\, \exists t \colon \nabla I \, (G(\tau, t) \wedge X(t, Y)) \, [\tau \colon \mathbf{Msbl}(\underline{\Omega}, I), Y \colon \mathbf{Msbl}(\underline{\Omega}, \mathbb{R})] \,]\!] \\ &= \bigvee_{t \in I} \{\tau = t \wedge X_t = Y\} \end{aligned}$$

Since X_t can be considered, as we observed above, the value of a function $I \xrightarrow{X} \mathbf{Msbl}(\underline{\Omega}, \mathbb{R})$ at time $t \in I$, there is no obstacle to consider the composite map $X_\tau := \underline{\Omega} \xrightarrow{\tau} I \xrightarrow{X} \mathbf{Msbl}(\underline{\Omega}, \mathbb{R})$. Therefore

$$[\![\, X(\tau, Y) \,]\!] = \{X_\tau = Y\}.$$

This is an half-victory, however, since we have already observed 'vanilla' stochastic processes are of little use. Adaptivity or stronger measurability properties are central to stochastic calculus, and to the very notion of 'time-dependent random process' itself.

In the next chapter, we embark on an investigation on some elementary additions to $\mathbf{Scott}[\underline{\Omega}]$ internal language which could allow to express measurability in an elementary manner.

[2]This way of extracting a witness of existence is akin to an **existence property** [Moschovakis 2018, Section 4.2] or to a **compactness property** [Blechschmidt 2017, Chapter II.7].

Chapter 5

Time in Scott topoi

5.1 Filtration modalities

Let $\underline{\Omega} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in I}, \mathbb{P})$ be a filtered probability space. We assume I has a top element ∞, so that $\mathcal{F}_\infty := \mathcal{F}$.

A filtration, as noted back in Chapter 2, represents growing observable knowledge about the world. An event $A \in \mathcal{F}_t$ is a certain observation one made before instant t.

Notice the bigger the event (inclusion-wise) the less information it contains. In fact, ideally, the best possible observation would allow one to pinpoint a specific outcome $\{\omega\} \in \mathcal{F}_t$. However, this is usually not possible, first of all because any realization at time t is oblivious of the future. Hence a completely unambiguous observation can only be made at $t = \infty$, when the experiment has ended. Until then, the best one can do is to come up with a decreasing sequence of events which, as time goes on and we acquire new information, narrows down on the final outcome:

$$A_0 \supseteq \cdots \supseteq A_t \supseteq \cdots \supseteq A_\infty. \tag{5.1.1}$$

On the other hand, the reverse problem is also interesting. Suppose the experiment has ended, and we are handed out 'final' observations. These amounts to an event B *at* ∞. Now, given a time $t \in I$, what does the information contained in B allow me to conclude about the status of the experiment at time t?

This might sound trivial at first, since $\mathcal{F}_\infty$ is omniscient about the outcomes of the experiment, yet it's not. If we identify events with the propositions on Ω which they reify, then it is clear that knowing an arbitrary proposition about the final state of a system (e.g. 'three heads were tossed') doesn't allow me to completely know what has been going on in the past states (e.g. 'how many heads were tossed after seven trials?').

We'll now see how the modalities introduced in Subsection 1.2.2 can be used to formalize this phenomenon. Indeed, a filtration of σ-fields carries over to the essential algebras associated to each of the $\mathcal{F}_t$. We denote them in blackboard bold. Hence we have a chain of inclusions:

$$\mathbb{F}_0 \hookrightarrow \dots \hookrightarrow \mathbb{F}_t \hookrightarrow \dots \hookrightarrow \mathbb{F}_\infty.$$

Remark 5.0.1. The essential algebra $\mathcal{F}_t$ is not $\mathcal{F}_t / \ker \mathbb{P}$, but $\mathcal{F}_t/(\ker \mathbb{P} \cap \mathcal{F}_t)$, unless $\mathcal{F}_\bullet$ is a complete filtration. Since meager subsets of $\mathcal{F}_t$ are still meager at later times, the inclusions are well-defined. Notice, however, that as time passes the equivalence relation $\simeq$ (essential equivalence) becomes strictly finer, i.e. if $A \leq B$ in $\mathbb{F}_t$ (from now on, also denoted as $A \leq_t B$) then $A \leq B$ at ∞, but the reverse is not always true. This has the intuitive meaning that, as times passes, we can exclude more and more 'possibilities' and hence our knowledge of what is impossible increases (since once something is known to be impossible, it remains such).

Let's focus, for now, on the inclusions $\mathbb{F}_t \hookrightarrow \mathbb{F}_\infty$. As observed in Subsection 1.2.2, they give rise to a pair of adjoint modal operators on $\mathbb{F}_\infty$:

$$\diamondsuit_t \dashv \Box_t : \mathbb{F}_\infty \to \mathbb{F}_\infty. \tag{5.1.2}$$

We can interpret them in the following way: if $B \in \mathbb{F}_\infty$ is an event at ∞, then

$$\begin{aligned} \diamondsuit_t B &= \bigwedge \{A \in \mathbb{F}_t \mid B \subseteq A\} \\ &= \text{possible state of the experiment at time } t \text{, given } B \text{ at time } \infty; \\ \Box_t B &= \bigvee \{A \in \mathbb{F}_t \mid A \subseteq B\} \\ &= \text{necessary state of the experiment at time } t \text{, given } B \text{ at time } \infty. \end{aligned} \tag{5.1.3}$$

The words 'necessary' and 'possible' have to be understood in the following way. The event $\Box_t B$ contains every outcome which must have happened at time t for B to happen later. Conversely, the event $\diamondsuit_t B$ contains every outcome which could have happened at time t if B subsequently happened. Therefore outcomes in $\diamondsuit_t B$ may not have happened later, but at time t they were not incompatible with the future realization of B; whereas outcomes in $\Box_t B$ all happened in the end, ending up in B.

Remark 5.0.2. By construction, notice an essential event $\Box_t A$ or $\diamondsuit_t A$ can be considered an essential event of $\mathbb{F}_t$ as well, since both modalities factor by the inclusion $\mathbb{F}_t \hookrightarrow \mathbb{F}_\infty$.

Example 5.1. Suppose in 2015 you sneaked up in a De Lorean car, retrofitted as a time machine, and traveled back to the year 1955 to give to your past self a sports almanac from the future you come from. In 1955, you now want to take advantage of your knowledge by betting on the sports events registered in the almanac you received. Unfortunately, a snarky kid (who, oddly enough, seems to come from 1985 himself), spoils the almanac so that outcomes of specific games are not intelligible anymore. The only information you have is in the cumulative statistics of each season. Hence:

$$\begin{aligned} B &= \text{information in the (spoilt) sports almanac,} \\ \Box_{1955} B &= \text{outcomes of 1955 and earlier games} \\ &\qquad \text{you can infer from the statistics in the almanac,} \\ \Diamond_{1955} B &= \text{outcomes of 1955 and earlier games} \\ &\qquad \text{you cannot exclude given the statistics in the almanac.} \end{aligned}$$

We can now read (5.1.1) in terms of the $\Diamond$ modalities: if one makes the best possible observations at each time t, the sequence of events narrowing down to the final event is given by

$$\Diamond_0 A \geq \cdots \geq \Diamond_t A \geq \cdots \geq \Diamond_\infty A = A. \tag{5.1.4}$$

The decreasing property of the sequence corresponds to the narrowing of the possibilities as the time passes and allows us to prune those outcomes which didn't realize.

Dually, the $\Box$ modality gives an increasing sequence of events representing growing definitive knowledge of the final event:

$$\Box_0 A \leq \cdots \leq \Box_t A \leq \cdots \leq \Box_\infty A = A. \tag{5.1.5}$$

Notice the distinction with the previous case: if $\Diamond$ narrows on the final event by pruning possibilities, $\Box$ grows a conservative estimate on the final event by collecting definitive evidence.

Identities. By elementary considerations on the definition of $\Diamond$ and $\Box$, we see that for every $s, t \in I$ one has

$$\begin{aligned} \Diamond_s \Diamond_t &= \Diamond_{s \wedge t}, \\ \Box_s \Box_t &= \Box_{s \wedge t}. \end{aligned} \tag{5.1.6}$$

where $s \wedge t$ denotes the smallest between s and t (since I is a total order). Mixed composites can also be reduced when $s \leq t$:

$$\begin{aligned} \Diamond_t \Box_s &= \Box_s, \\ \Box_t \Diamond_s &= \Diamond_s. \end{aligned} \tag{5.1.7}$$

This is just an instance of the obvious fact that both $\Diamond_t$ and $\Box_t$ fix the events in $\mathbb{F}_t$. Moreover, by (1.2.5), we can also relate the other two composites:

$$\neg\Box_s\Diamond_t = \Diamond_s\Box_t\neg \tag{5.1.8}$$

regardless of the relationship between $s, t \in I$ (of course if $s \geq t$, this generalized duality collapses to (1.2.5)).

5.2 In Scott topoi

Given a morphism of Boolean frames $\mathbb{A} \xrightarrow{f} \mathbb{B}$, what happens to their associated localic topoi?

We look at this situation in the framework of forcing triposes. First of all, notice $f^! \dashv f \dashv f_!$ induces a triplet of pseudonatural transformations:

$$\begin{aligned} f^!{}_X : \mathbb{B}^X &\longrightarrow \mathbb{A}^X \\ \beta &\longmapsto f^! \circ \beta, \\ f_X : \mathbb{A}^X &\longrightarrow \mathbb{B}^X \\ \alpha &\longmapsto f \circ \alpha, \\ f_{!X} : \mathbb{B}^X &\longrightarrow \mathbb{A}^X \\ \beta &\longmapsto f_! \circ \beta. \end{aligned} \tag{5.2.1}$$

The first, $f \dashv f_!$, has a left exact inverse part, therefore Theorem 3.18 gives us a geometric morphism between the Pitts' topoi associated to $\mathbb{A}^{(-)}$ and $\mathbb{B}^{(-)}$:

$$\mathbf{Set}[\mathbb{A}^{(-)}] \underset{f_*}{\overset{f^*}{\rightleftarrows}} \mathbf{Set}[\mathbb{B}^{(-)}] \tag{5.2.2}$$

However, $f^!$ is not left exact in general:

Example 5.2. Consider the morphism $\mathcal{P}(2) \xrightarrow{f} \mathcal{P}(\mathbb{N})$ induced by the assignment

$$\top \mapsto \{\text{even numbers}\}, \quad \bot \mapsto \{\text{odd numbers}\}.$$

Then consider

$$B_1 = \{\text{even numbers}\} \cup \{1\}, \quad B_2 = \{\text{odd numbers}\} \cup \{0\}.$$

Clearly, $f^!(B_1 \cap B_2) = \{0, 1\}$, but $f^!(B_1) \cap f^!(B_2) = \mathcal{P}(\mathbb{N}) \cap \mathcal{P}(\mathbb{N}) = \mathcal{P}(\mathbb{N})$.

This means $\mathbf{Set}[\mathbb{B}^{(-)}] \xrightarrow{f} \mathbf{Set}[\mathbb{A}^{(-)}]$ is an essential geometric morphism, but in general not an open one (i.e. $f^! \dashv f$ doesn't lift to a geometric morphism $\mathbf{Set}[\mathbb{A}^{(-)}] \to \mathbf{Set}[\mathbb{B}^{(-)}]$)

Nonetheless, one can bring both modalities to the internal language of $\mathbf{Set}[\mathbb{B}^{(-)}]$ by defining them fiberwise, and introducing new rules in the language:

$$\begin{aligned} [\![\Box_f \varphi\,[\vec{x}:\vec{\Gamma}]]\!] &:= \Box_f^{[\![\Gamma]\!]} [\![\varphi\,[\vec{x}:\vec{\Gamma}]]\!], \\ [\![\Diamond_f \varphi\,[\vec{x}:\vec{\Gamma}]]\!] &:= \Diamond_f^{[\![\Gamma]\!]} [\![\varphi\,[\vec{x}:\vec{\Gamma}]]\!] \end{aligned} \tag{5.2.3}$$

where $\Box_f^{[\![\Gamma]\!]} := f_X f_{!X}$ and $\Diamond_f^{[\![\Gamma]\!]} := f_X f^!{}_X$.

Proposition 5.3. *A morphism of complete Boolean algebras $\mathbb{A} \xrightarrow{f} \mathbb{B}$ induces two arrows $(\mathbb{B}, \Leftrightarrow) \to (\mathbb{B}, \Leftrightarrow)$ in the topos $\mathbf{Set}[\mathbb{B}^{(-)}]$, namely*

$$\begin{aligned} \Box_f : \mathbb{B} \times \mathbb{B} &\longrightarrow \mathbb{B} \\ (b, b') &\longmapsto \Box_f(b) \Leftrightarrow b' \\ \Diamond_f : \mathbb{B} \times \mathbb{B} &\longrightarrow \mathbb{B} \\ (b, b') &\longmapsto \Diamond_f(b) \Leftrightarrow b' \end{aligned} \tag{5.2.4}$$

Proof. By naturality of $\Box_f^X$ and $\Diamond_f^X$ (as transformations $\mathrm{Sub}_{\mathbf{Set}[\mathbb{B}^{(-)}]} \to \mathrm{Sub}_{\mathbf{Set}[\mathbb{B}^{(-)}]}$) and Yoneda Lemma. □

Remark 5.3.1. Notice $\Box_f$ (and, analogously, $\Diamond_f$) is surjective:

$$\bigvee_{\sigma \in \Sigma} \Box_f(\tau) \Leftrightarrow \sigma \geq (\Box_f(\varnothing) \Leftrightarrow \sigma) \vee (\Box_f(\top_{\mathbb{B}}) \Leftrightarrow \sigma) = \neg\sigma \vee \sigma = \top_{\mathbb{B}}.$$

This might seem weird at first, but it isn't. Compare this to the following, 'classical' fact: the map $2\lfloor -/2 \rfloor : \mathbb{N} \to \mathbb{N}$ is surjective. This doesn't mean it is trivial. For example, one can characterize even numbers as its fixed points.

The topos-theoretic approach to such adjoint pairs of modalities is described and thoroughly investigated by Reyes in [Reyes et al. 1991] and by Reyes and Zolfaghari in [Reyes and Zolfaghari 1991]. In these works, the authors recognize how certain geometric morphisms of topos $\mathcal{E} \xrightarrow{\Gamma} \mathcal{S}$ give rise to an adjoint pair of modal operators $\Diamond \dashv \Box$ (which they call MAO, as in 'modal adjoint operators') which is analogous to ours.

They interpret the geometric morphism Γ to exhibit $\mathcal{E}$ as a topos of 'variable sets' over a topos of 'constant sets' $\mathcal{S}$. The intuition behind their work is then to use the 'variability' of $\mathcal{E}$ to represent possible worlds, in the Kripkean sense.

Call Δ the inverse part of Γ. They induce an obvious pair of morphisms:

$$\mathrm{Sub}_{\mathcal{S}}(X) \underset{\Gamma_X}{\overset{\Delta_X}{\rightleftarrows}} \mathrm{Sub}_{\mathcal{E}}(\Delta X).$$

The first is defined by simply applying Δ to any mono $U \rightarrowtail X$, while the second is defined on $V \rightarrowtail \Delta X$ as the pullback of $\Gamma(V) \rightarrowtail \Gamma\Delta(X)$ along the unit $\eta_X : X \to \Gamma\Delta(X)$ (or, alternatively, as the image of the transpose map $\Gamma(V) \to X$).

To make Δ_X well-defined, Δ has to preserve all meets (this is 'assumption $(*)$' in [Reyes and Zolfaghari 1991]). As we observed, this is equivalent to have a left adjoint Π_X to Δ_X. In the end, $\Pi_X \dashv \Delta_X \dashv \Gamma_X$ is equivalent to the adjoint string $f^!{}_X \dashv f_X \dashv f_{!X}$ we found independently, as described by [Reyes and Zolfaghari 1991, Proposition 1.1].

From this string of adjoints, by naturality, we get an induced triple of adjoints in $\mathcal{S}$[1]:

$$\Omega_{\mathcal{S}} \xrightarrow{\delta} \Gamma(\Omega_{\mathcal{E}}), \qquad \pi, \gamma : \Gamma(\Omega_{\mathcal{E}}) \to \Omega_{\mathcal{S}}.$$

Unfortunately, in our case this approach does not yield the desired modalities, since $\Box_f$ and $\Diamond_f$ are modalities on $\mathbb{B}$, and thus should appear as operators on $\Omega_{\mathcal{E}}$. We were not able to 'transpose' Reyes' formalism to yield a similar theory on $\mathcal{E}$.

We find the reason for this in the different character of our modalities with respect to the ones investigated by Reyes. While his are alethic modalities, questioning the validity of a proposition in the accessible variations (pictured to be $\mathcal{E}$, accessed through $\Delta \dashv \Gamma$) of a fixed world (picture to be $\mathcal{S}$), our modalities question the available knowledge of a process which already took place.

This is very clear when f is instantiated to be the inclusion $\mathbb{F}_t \hookrightarrow \mathbb{F}_\infty$ which gives rise to the modalities $\Box_t \dashv \Diamond_t$ we outlined at the beginning of the chapter. Those modal operators are used to speak about 'accessed worlds' and not 'accessible worlds', as the events they produce have already happened.

5.2.1 Internal measurable criteria for stochastic processes

Return now to Scott topoi. If $\underline{\Omega}$ is filtered by $\{\mathcal{F}_t\}_{t\in I}$, the above construction provides us with a large arsenal of modalities for the internal language of $\mathbf{Scott}[\underline{\Omega}] = \mathbf{Set}[\mathbb{F}_\infty^{(-)}]$.

Can we use them to express filtration-related properties of a process X**?** We've seen in Theorem 4.12 that a process $\{X_t : \underline{\Omega} \to \mathbb{R}\}_{t\in I}$ gives a map $X : \nabla I \to \mathcal{R}$ in

[1] In the general case Π is only lax natural, so that π is lax too. In our case however, Π is natural and π is strict. This is addressed by [Reyes and Zolfaghari 1991, Example 1].

Scott$[\underline{\Omega}]$, confirming the conceptual view that processes ought to be functions, in some sense. What we couldn't do was to impose measurability conditions on X, such as adaptivity, which is the most important among the ones we saw back in Chapter 2.

The above question remains largely open, since we didn't have the time to delve into this issue thoroughly enough. Therefore, we collect here some ideas, comments and open questions concerning this matter.

Adaptivity. Recall a process $\{X_t : \underline{\Omega} \to \mathbb{R}\}_{t\in I}$ is adapted if each X_t is $\mathcal{F}_t$-measurable; hence if its value in the present doesn't depend on future events. One could approach the formalization of such condition using the modalities arising from $\{\mathcal{F}_t\}_{t\in I}$. In fact to say X_t is $\mathcal{F}_t$-measurable is equivalent to say $\Box_t$ and $\Diamond_t$ fix every event $X_t^{-1}A$, $A \in \mathcal{B}(\mathbb{R})$. Since $\mathcal{B}(\mathbb{R})$ is generated by half-lines $(-\infty, a]$, we could as well express this condition as

$$\{X_t \leq a\} \leq \Box_t\{X_t \leq a\}, \qquad \text{for all } a \in \mathbb{R}.$$

Now this looks like something which could be internalized rather easily. In fact $\{X_t \leq a\}$ is the truth value of the proposition $X(t) \leq a$ in **Scott**$[\underline{\Omega}]$. Nonetheless, the naïve attempt may be a failure. Suppose we define (using **Scott**$[\underline{\Omega}]$'s language directly, not its tripos') the following **internal adaptivity predicate**:

$$\forall t : \nabla I \, \forall a : \mathcal{R}_{\underline{\Omega}} \, (X(t) \leq a \to \Box_t(X(t) \leq a)) \, [X : \mathcal{R}_{\underline{\Omega}}^{\nabla I}]. \tag{5.2.5}$$

This says $X(t) \leq a$ has to be a fixed point of $\Box_t$. We've seen in Remark 5.3.1 that even if $\Box_t$ is surjective, it need not to be trivial, and its fixed points are indeed not trivial: they correspond to those essential events in $\mathbb{F}_t$.

Question 1. *Indeed, an interesting thing to observe is the data of a filtration $\{\mathcal{F}_t\}_{t\in I}$ and of a family of adjoint modalities $\{\Diamond_t \dashv \Box_t\}_{t\in I}$ are somewhat equivalent.* ***What is the exact extent to which this is true?***

The predicate (5.2.5) has two main defects we have to iron out. The first is that $\mathcal{R}_\Omega$ includes also non-constant reals, as we've seen in the previous chapter. Hence the measurability condition $X(t) \leq a$ might not be right. The second is we use t both as a variable and as index for $\Box_{(-)}$. Such indexes are 'external', while t is an internal symbol, so the predicate written above is ill-formed. One may circumvent this by defining a map like

$$\begin{aligned} \Box : I \times \Sigma^\Sigma &\longrightarrow \mathbb{F}_\infty \\ (t, \ell) &\longmapsto \bigwedge_{A,B\in\mathbb{F}_\infty} ((\Box_t(A) \Leftrightarrow B) \leftrightarrow \ell(A,B)). \end{aligned} \tag{5.2.6}$$

A second way is to define, externally, a family of predicates, indexed by I and by $\mathbb{R}$ (thereby externalizing the two universal quantifications). Then one may express (5.2.5) by taking the infimum of all of them. This, however, is less appetizing as it is not a completely internal construction.

Question 2. *Is* (5.2.5) *a good candidate for internalizing the concept of adaptivity?*

Filtration modalities for stochastic processes. Another line of investigation is given by the distinction between $\Diamond_t$ and $\Box_t$. Indeed, we wrote $\Box_t$ in (5.2.5) where $\Diamond_t$ would have been equally effective, since both have the same fixed points.

Question 3. *Are there any concepts, in the theory of stochastic processes, that one can represent with one but not the other? More generally, what is the significance of their meaning, as explored at the beginning of this chapter, for stochastic calculus and non-deterministic dynamical systems in general?*

We embarked on an investigation of this question in Appendix A.

Appendices

Appendix A

Process logic and stochastic processes

We want now to explore the logical potential of the set of modalities associated to a stochastic filtration that we introduce in Chapter 5. To do so, we first begin by reviewing the principal logics developed to speak about 'processes' in 'time'. These can be distinguished in two kinds: temporal and dynamic logics.

Temporal logics are *endogenous* logics, which means they fix a specific 'system' and reason about its behaviour only. In particular, their syntax does not allow one to write a description of the system. Dynamic logic instead is *exogenous*, meaning one can speak about the system itself in the language. Moreover one can describe their composition and thus create new ones: complex systems are built out of 'atomic ones', which are the only one whose semantics has to be explicitly given. Hopefully, this distinction will become clearer as we describe both kinds of logic.

Remark A.0.1. We stress both CTL^* and PDL, which are, respectively the temporal and dynamic logics we are about to expound, are subsystems of the **modal μ-calculus**. This system is remarkably simpler compared to those two, having two fixpoint operators at its disposal. See [Venema 2008] for an elementary introduction to the μ-calculus, through the lens of process logic, and [Emerson 1997] for a discussion of how CTL^* and PDL can be embedded in the modal μ-calculus.

Let us spend some words on the meaning of 'process', 'time', 'system', 'state' and 'transition'. These are very intuitive words, inextricably related to each other, whose meaning has many different hues.

Ultimately, a 'system' can be anything capable of change. In our matters, systems only have internal properties and exhibit a one-way external behaviour[1]. Then a 'process' is the unfolding of the changes a system undergoes throughout 'time'. 'Time' is an ordered collection of instants. At each instant of 'time', a 'system' can be in a variety of 'states'. Thus the change of a 'system' can be broken up in 'transitions' from a 'state' to another.

One may dispense with a 'system' altogether, and directly speak of 'processes', adopting a **behavioural** approach instead. This is more of a philosophical choice than a mathematical one, although such a choice makes a difference in ease of application, depending on the context. The counterpart to 'behavioural' is 'ontological'. The 'behavioural' versus 'ontological' attitudes are reflected in the 'exogenous' versus 'endogenous' characteristics of dynamic and temporal logic, respectively[2].

In our context, processes are certainly more central. Indeed, a filtration usually encodes the unfolding of a process[3] whose 'generating system' can be unknown. Nevertheless, to better compare and understand the set of modalities we found we are going to give a canonical construction of a filtered probability spaces from the specification of a system. Such construction will also help in defining and understanding the semantics of the logic we are going to develop.

A.1 Transition systems

As customary for modal logics, the semantics of the logical systems we are going to see is done through Kripke frames [Venema 2008, Chapter 1]. However, in the process logic literature these are often called 'transition systems' (e.g. Huth and Ryan 2004; Stirling 2001; Pnueli 1977).

[1] In other words, they do not interact with other systems or with the special system usually called 'environment'. This is the topic of the growing literature on **open systems**, which is the core and probably the first motivation behind applied category theory [Fong 2016; Ghani et al. 2018; Coecke and Duncan 2011; Baez and Pollard 2017; Clark, Coecke, and Sadrzadeh 2008; Myers 2020].

[2] The distinction between these two approaches is the subject of the notion *bisimilarity*. Roughly, two systems are bisimilar if they have the same behavioural properties. This is formalized as the ability of simulating each other. See [Venema 2008, Section 1.3] for some results relating bisimilarity and intrinsic equivalence for modal logics.

[3] We are not referring to *stochastic processes* in the mathematical sense, which instead represent observable properties of an underlying system. For instance, the stock market is a system, which generates a process (e.g. the prices of a given portfolio), which in turn has observable properties such as the value I extract from it with an investment strategy, economical indicators, aggregate information, etc. These are stochastic processes in the traditional sense.

Definition A.1. A **transition system** is a set of **states** S with a **transition relation** $\rightsquigarrow$, which is a total relation on S.

Remark A.1.1. Usually S is a set of 'possible worlds' and $\rightsquigarrow$ is called an 'accessibility relation'.

Definition A.2. An **(admissible) path** in a transition system is an infinite sequence of transitions

$$\pi = s_0 \rightsquigarrow s_1 \rightsquigarrow s_2 \rightsquigarrow \cdots .$$

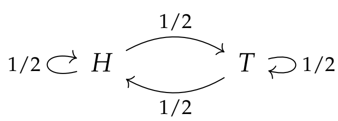
To denote the suffix of π starting from the i-th entry (included), we write π^i. For instance, $\pi^0 = \pi$.

We are going to generalize these devices to include more expressive transition relations and continuous time. In the process, we will find ourselves surprisingly close to core topics in stochastic calculus, such as Markov chains and stochastic differential equations.

Definition A.3. A transition system $(S, \rightsquigarrow)$ is **deterministic** if for every state $s \in S$, there exists a unique $s' \in S$ such that $s \rightsquigarrow s'$, i.e. if $\rightsquigarrow$ is single-valued.

Deterministic transition systems are equivalent to state machines. Non-deterministic transition systems can be equipped with a *probability labelling* and become what is known as a **time-homogeneous Markov chain**.

Example A.4. A sequence of coin flips is modelled by the transition system $S = \{H, T\}$, $\rightsquigarrow = S \times S$, labelled by probabilities in this way:



Intuitively speaking, a probability labelling should be an assignment $\mathbb{P} : S \times S \to [0, 1]$ of 'transition probabilities' to a (non-deterministic) transition system $(S, \rightsquigarrow)$ such that

1. if $\mathbb{P}(s, s') > 0$, then $s \rightsquigarrow s'$ holds (we call this an **allowed transition**),
2. for every state $s \in S$, $\sum_{s' \in S} \mathbb{P}(s, s') = 1$.

When S is a finite set, such a $\mathbb{P}$ amounts to a so-called right stochastic matrix, i.e. a matrix having non-negative entries and in which each row sums to 1. In the infinite (uncountable) case, we have to be more clever with our definitions since a plain map

$S \times S \to [0, 1]$ is not good enough for what we are going to do with it downstream in the construction.

Definition A.5. A **probability labelling** for a transition system $(S, \rightsquigarrow)$ is given by:

1. a measurable structure on S, thus a σ-field Σ on S;
2. a Markov kernel $S \nrightarrow S$.

Notice this structure collapses to the right stochastic matrices we discussed earlier when S is finite and $\Sigma = \mathcal{P}(S)$, simply because a probability measure on a discrete space is completely determined by its value on singletons, thus elements of S. Since this still holds for countable discrete spaces, we see how the general definition is only needed when S is an uncountable set, like a continuum of states.

Example A.6. Let $S = \mathbb{R}^n$ and $\Sigma = \mathcal{B}(\mathbb{R}^n)$. We are to describe a deterministic continuous dynamical system (represented by a smooth flow on S) with a non-deterministic perturbation as a probability labelled transition system on S.

Thus consider a flow $\Phi = \{\Phi_t\}_{t\in\mathbb{R}}$, which is a one-parameter group of diffeomorphisms of $\mathbb{R}^n$. This determines a foliation of $\mathbb{R}^n$ into flow lines, which are streams along which points *flow* according to the action of Φ (a point $x_0 \in \mathbb{R}^n$ moves along the curve parametrized by $t \mapsto \phi_t(x_0)$). This can be described as a transition system by defining $\rightsquigarrow_\Phi$ to be[4]

$$x \rightsquigarrow_\Phi y \quad \underline{\text{iff}} \quad y = \Phi_1(x).$$

Notice we are discretizing the flow. This is actually a limit of our description, which is inherently discrete. A slight improvement is to parametrize the relation on the step size $t > 0$:

$$x \rightsquigarrow_{\Phi,t} y \quad \underline{\text{iff}} \quad y = \Phi_t(x).$$

We'll discuss later how to overcome this obstacle and work directly with the action Φ.

We now extend this description with a stochastic noise. We want to add a small probability of jumping from one flow line to another, as if a particle moving along Φ were subject to random nudges making it wiggle in the current, while still moving with the flow macroscopically. Observe we cannot keep $(\mathbb{R}^n, \rightsquigarrow_\Phi)$ and just label its transitions with probabilities, since, as of now, transitions across flow lines are not permitted. We really need to 'perturb' the transitions, in the sense of mathematical physics, that is

[4] This is a 'directed' version of the equivalence relation partitioning $\mathbb{R}^n$ into the orbits of the action of Φ.

define a new relation $\rightsquigarrow$ which is obtained by overlaying a genuinely new disturbance to the transitions we already have in place. Informally:

$$\rightsquigarrow = \rightsquigarrow_{\Phi} + \varepsilon \rightsquigarrow_{\text{noise}}$$

Formally, this must be specified by a family $\{K_t\}_{t\in\mathbb{R}^+}$ of Markov kernels describing the exact dynamics of the perturbation. The description of such a kernel comes from the study of **Itô diffusions**, which are stochastic processes obeying a stochastic differential equation (SDE) of a particular form, and can be rather hard. We do not delve into details, but the interested reader can consult [Grimmett and Stirzaker 2001, Chapter 13].

What matters to us is that under suitable hypothesis on the flow Φ (which is called the 'drift' of the diffusion process) one can solve the equation and get an $\mathbb{R}^n$-valued process[5] $(X_t)_{t\in\mathbb{R}^+}$ which describes the evolution of a particle under the situation we setup. From this solution, one extracts the kernels by setting

$$\begin{aligned} K_t : \mathcal{B}(\mathbb{R}^n) \times \mathbb{R}^n &\longrightarrow [0,1] \\ (A, x) &\longmapsto \mathbb{P}(X_t \in A \mid X_0 = x). \end{aligned} \tag{A.1.1}$$

Here $\mathbb{P}$ is the probability measure on the auxiliary stochastic base one introduced to solve the aforementioned SDE. It is a fictitious intermediary since one only needs X_t law (which is, by definition, the measure pushed forward by X_t), and there's a direct way to get it through the Feynman–Kac formula. See [Caravenna 2011, Section 7.3].

The previous example gives us a couple of hints/ideas about what we developed so far:

1. A transition relation is not powerful enough to capture continuous transitions. We should employ flows instead, *aka* monoidal actions on the state space. However, this comes at the cost of losing the ability of specify non-deterministic systems, which brings us to the next point:

2. Markov kernels are much more expressive than transition relations. With a Markov kernel we have complete freedom over the determinism of the transitions *plus* a way to exactly specify how much and where the non-determinism lies.

Hence we give a more encompassing definition:

Definition A.7. A **Markovian action** of a monoid $(M, 1, \cdot)$ on the measurable space (S, Σ) is a monoidal map $\Phi : M \to \mathbf{Markov}(S, S)$.

Definition A.8. A **stochastic transition system** $(\underline{S}, \underline{I}, K)$ is a probability space of states $\underline{S} = (S, \Sigma, \mu)$ together with a Markovian action K of a time monoid $\underline{I} = (I, 0, +, \leq)$ on S.

The monoid I is usually $\mathbb{N}$ or $\mathbb{R}^+$. Its order relation will be useful later on to define a filtration. The action $K : M \to \mathbf{Markov}(S, S)$ is called the **transition kernel**.

Remark A.8.1. A finer formalization would distinguish the time monoid I, which indexes the time shifts, from the *observed time fragment* $[a, b] = a\uparrow \cap b\downarrow \subseteq I$, which indexes the trajectories.

Remark A.8.2. This notion of transition system is strictly related to time-homogeneous Markov processes, since every such process determines a stochastic transition system and viceversa. In fact, from both a time-homogeneous Markov process and a stochastic transition system we can extract a transition kernel. For the first, such a kernel can be shown to be a Markovian action, a fact usually ascribed to the so-called *Chapman–Kolmogorov equation*, which is exactly the statement $K_t K_s = K_{t+s}$. A minimal framing for these notions can be found in [Caravenna 2011, Chapter 8].

Example A.9. Example A.6 can be now recast in more conceptual terms. As before, the state space is $\mathbb{R}^n$ with its Borel σ-field. The Markovian action is then given by the sequence of kernels (A.1.1). As remarked above, the fact such kernels form a Markovian action is known as the Chapman–Kolmogorov equation. We only need to show K_0 is the identity kernel. But this is trivial given the definition of K_t:

$$K_0(A, x) = \mathbb{P}(X_0 \in A \mid X_0 = x) = \delta_x(A).$$

Example A.10. Consider the discrete transition system of Example A.4. There $S = \{H, T\}$ and $\Sigma = \mathcal{P}(S)$. The Markovian action is given by

$$\begin{aligned} K_t : \mathcal{P}(S) \times S &\longrightarrow [0, 1] \\ (A, s) &\longmapsto \left(\frac{1}{2}\delta_H + \frac{1}{2}\delta_T\right)(A). \end{aligned}$$

Example A.11. Let $(S, \rightsquigarrow)$ be a transition system, with S finite. Equip S with the uniform probability measure, and set $I = \mathbb{N}$. Then the relation $\rightsquigarrow$ can be encoded by the following Markovian action:

$$\begin{aligned} K_1 : \mathcal{P}(S) \times S &\longrightarrow [0, 1] \\ (B, s) &\longmapsto \frac{1}{|\{s' \in S \mid s \rightsquigarrow s'\}|} \sum_{s \rightsquigarrow s'} \delta_{s'}(B). \end{aligned}$$

$$\begin{aligned} K_n : \mathcal{P}(S) \times S &\longrightarrow [0, 1] \\ (B, s) &\longmapsto K_1^n(B, s). \end{aligned}$$

Notice K_1 is well-defined since $\rightsquigarrow$ is total, i.e. the set $\{s' \in S \mid s \rightsquigarrow s'\}$ is never empty. Its definition assigns a uniform probability of transitioning from a given s to any other state s' to which we can transition. K_n is then automatically defined since $(\mathbb{N}, 0, +)$ is generated by 1.

Remark A.11.1. The previous construction can be easily generalized to the case S is countable, by dropping the uniformity of the measures on S and each measure $K_1(-, s)$.

A.1.1 The filtered space of trajectories

We now aim to describe a general construction to turn a stochastic transition system $(\underline{S}, \underline{I}, K)$ into a filtered probability space.

The filtered probability space we are going to produce is the space of trajectories of the system $(\underline{S}, \underline{I}, K)$, hence 'possible paths' starting from a state $s_0 \in S$ and following the stochastic action of K. In practice, we are generalizing and formalizing the way we constructed, the filtered probability space of the coin flips experiment (Example 2.31).

Construction of the space of trajectories. Trajectories of a system live in the product space $(S^I, \Sigma^{\otimes I})$. It can be easily equipped with a probability measure thanks to Hahn–Kolmogorov's extension (Theorem 2.19), which allows us to define the probability measure by extension of an assignment the cylinders of S^I, which are those subsets of the form

$$C = \{\omega \in S^I \mid \omega_{t_1} \in A_1, \omega_{t_1+t_2} \in A_2, \dots, \omega_{t_1+\dots+t_k} \in A_k\},$$

where $k \in \mathbb{N}$, $t_1 < \dots < t_k \in I$ and $A_1, \dots, A_k \in \Sigma$. These can be assigned a probability rather naturally:

$$\mathbb{P}(C) := \prod_{i=1}^{k} \int_{A_{i-1}} K_{t_i}(A_i, -) \, \mathrm{d}\mu \tag{A.1.2}$$

where we put $A_0 = S$. In fact, $K_{t_i}(A_i, s)$ is the probability of transitioning from s to A_i within the next t_i units of time. Integrating this over A_{i-1} we get the probability of transitioning from a state in A_{i-1} to a state in A_i. Multiplying all this probability we get the probability of evolving through the given sets of states in the given amounts of time.

The initial likelihood of finding the system in an initial state belonging to A_0 is given by the probability distribution of μ. K then moves the probability mass according to the possible evolutions of the stochastic system. Notice, moreover, that if $t_1 = 0$ then the first term of the product is given by

$$\int_{A_0} K_0(A_1, -) \, \mathrm{d}\mu = \int_S \delta_s(A_1) \, \mathrm{d}s = \mu(A_1).$$

Definition A.12. The **space of trajectories** of a stochastic transition system $(\underline{S}, \underline{I}, K)$ is the probability space carried by S^I together with the product σ-field $\Sigma^{\otimes I}$ and the probability measure induced by Carathéodory extension of the assignment (A.1.2) to the whole $\Sigma^{\otimes I}$.

The space of trajectories is naturally filtered by the order of I, in the following way:

Definition A.13. The **canonical filtration** of the space of trajectories of a stochastic transition system $(\underline{S}, \underline{I}, K)$ is

$$\mathcal{F}_t := \{A \times S^{t\uparrow} \mid A \in \Sigma^{\otimes t\downarrow}\}.$$

If I lacks a top element, we forcibly define $\mathcal{F}_\infty := \Sigma^{\otimes I}$.

Admissible trajectories. The space S^I contains a lot of trajectories the system never experience. That piece of information, which is contained in K, is reflected in the probability measure. In fact, if the system at a state $s \in A \in \Sigma$ cannot transition in time t to a state $s' \in B \in \Sigma$, then

$$\mathbb{P}(\{\omega \in S^I \mid \omega_0 \in A, \omega_t \in B\}) = \int_A K_t(B, -)\, \mathrm{d}\mu = 0. \tag{A.1.3}$$

Nonetheless, it's handy to define exactly what it means for a specific trajectory $\omega \in S^I$ to be admissible by the system. In order to do that, we need to assume (S, Σ) is topological, i.e. $\Sigma = \mathcal{B}(S)$ for some topological structure on S. This is already true when S is finite, or countable, and equipped with its discrete σ-field, which is the Borel σ-field of its discrete topology. When S is uncountable, then it's usually (in meaningful situations at least) a piece of a manifold, in which case the topological structure was there even before the measurable structure. Therefore, we do not deem such a requirement to be too taxing.

In return, we can say the possible transitions (in time t) from a state $s \in S$ are given by the *support* of $K_t(-, s)$:

Definition A.14 (Ambrosio, Gigli, and Savaré 2008)**.** The **support** of a measure μ on the topological measurable space $(X, \mathcal{B}(X))$ is the set

$$\operatorname{supp} \mu = \overline{\{x \in X \mid \text{every open neighbourhood of } x \text{ has positive measure}\}}.$$

The support of a measure is the set on which its 'mass' is completely concentrated. For example, a Dirac delta on a point x_0 is supported on the singleton $\{x_0\}$. In the case of a probability measure, sets outside the support can be considered as *really* impossible events, as opposed to mere sets of zero probability *in* the support which are possible but 'extremely diluted'.

Example A.15. When constructing a Markov action from a classical transition system $(S, \rightsquigarrow)$ (Example A.11), $\operatorname{supp} K_1(-, s)$ contains only those states reachable from s in a single transition:

$$\operatorname{supp} K_1(-, s) = \{s' \in S \mid s \rightsquigarrow s'\}.$$

The support $\operatorname{supp} K_n(-, s)$ contains only those states reachable from s in exactly n transitions.

Example A.16. In Example A.6, the supports are trivial: $\operatorname{supp} K_t(-, x) = \mathbb{R}^n$ for any $t \in I$, $x \in \mathbb{R}^n$. This because the probability mass is spread all over $\mathbb{R}^n$, albeit lumped around $\Phi_t x$. Tweaking the diffusion model used to construct K one might as well get non-trivial supports, e.g. by cutting the 'tails' of a Gaussian distribution (and rescaling suitably).

Thus we can give the following definition:

Definition A.17. A transition from state $s \in S$ to state $s' \in S$ is **admissible in time** t if

$$s' \in \operatorname{supp} K_t(-, s). \tag{A.1.4}$$

Remark A.17.1. The concept of 'admissible in time t' makes it possible to define a family $\{\rightsquigarrow_t^K\}_{t \in I}$ of transitions 'discretizing' K: for $s, s' \in \mathbf{S}$, we define

$$s \rightsquigarrow_t^K s' \quad \underline{\text{iff}} \quad \text{it is admissible to transition from } s \text{ to } s' \text{ in time } t.$$

This family of discretizations is exactly what we encountered when approaching the diffusion model in Example A.6.

Finally, an admissible trajectory is one in which every transition is possible:

Definition A.18. A trajectory $\omega \in S^I$ is **admissible** if

$$\omega_t \rightsquigarrow_s^K \omega_{t+s}, \qquad \text{for all } s, t \in I. \tag{A.1.5}$$

By the observation we made at the beginning of the paragraph (computation (A.1.3)), we conclude the set of inadmissible transitions is meager in $\underline{S}^I$.

'Nice' transition systems. The filtered spaces associated to a stochastic transition system $(\underline{S}, \underline{I}, K)$ have wildly different properties depending on the cardinality of S and I. In this regard, a particularly well-behaved class of stochastic transition systems is given by the following:

Definition A.19. A stochastic transition system $(\underline{S}, \underline{I}, K)$ is **discrete** if S is a countable set and $\underline{I} = (\mathbb{N}, 0, +, \leq)$.

They are well-behaved by virtue of a 'separability' property: at every finite time, we can distinguish between trajectories even up to essential equivalence, i.e. a 'finite trajectory' has not zero measure. We now formalize this idea:

Definition A.20. The **truncation operator** at $t \in I$ is the map

$$\operatorname{tr}_t \omega = \pi_{t\downarrow}^{-1} \pi_{t\downarrow}(\omega), \qquad \omega \in \Omega. \tag{A.1.6}$$

Its effect is to take a definite trajectory and 'forget about the future'. Hence it yields the set of every possible trajectory prolonging ω after t. Truncation is very similar to $\diamondsuit_t$, however, $\diamondsuit_t$ cannot be used on single trajectories as in most of the cases the singleton $\{\omega\}$ is meager for $\mathbb{P}$.

Definition A.21. A stochastic transition system $(\underline{S}, \underline{I}, K)$ **has the singleton property** if, for any *admissible* trajectory $\omega \in S^I$,

$$\mathbb{P}(\operatorname{tr}_t \omega) > 0, \qquad \text{for all } t \in I.$$

Proposition A.22. *Let $(\underline{S}, \underline{\mathbb{N}}, K)$ be a discrete stochastic transition system. Suppose, moreover, that the singletons of S have positive measure with respect of the probability measure of $\underline{S}$. Then $(\underline{S}, \underline{\mathbb{N}}, K)$ has the singleton property.*

Remark A.22.1 (Future work). The generalized transition systems we defined above are begging to be recast in the framework of open dynamical systems [Myers 2020]. The improvement we expect from switching to such a language is twofold. First, we'd get a more conceptual description, which better highlights the modelling assumptions we are making and the role played by the different structures involved. Second, we expect to achieve a more versatile definition, which might encompass time-inhomogeneous Markov chains (a task that looks candidly within hand reach) and perhaps non-Markovian processes as well.

A.1.2 Stopping times for trajectories

Let $\underline{\Omega} = (\Omega, \mathcal{F}_\infty, \{\mathcal{F}_t\}_{t \in I}, \mathbb{P})$ be the space of trajectories of the stochastic transition system $(\underline{S}, \underline{I}, K)$.

Definition A.23. A **stopping time** on $\underline{\Omega}$ is a random variable $\tau : \Omega \to I$ such that

$$\{\tau \leq t\} \in \mathcal{F}_t, \qquad \text{for all } t \in I.$$

Stopping times are an important tool in stochastic calculus since they provide a kind of 'feedback loop' between a process and the probability space on which it is defined. This is due to the fact that stopping times are actually 'times of realizations' for certain propositions. The fact they're 'adapted' means one can assess, at each time $t \in I$, whether the event controlled by τ has happened yet or not.

Moreover, stopping times behave very similar to 'deterministic' times (which, by the way, are themselves stopping time).

Proposition A.24 (Facts about stopping times)**.** *Let τ, σ be stopping times on $\underline{\Omega}$.*

1. *The random variables $\tau + \sigma$, $\tau \wedge \sigma$, $\tau \vee \sigma$ are stopping times too.*
2. *The following is a σ-field:*

$$\mathcal{F}_\tau = \{A \in \mathcal{F}_\infty \mid A \cap \{\tau \leq t\} \in \mathcal{F}_t, \textit{ for all } t \in I\}$$

3. *If $\tau \leq \sigma$ almost everywhere, then $\mathcal{F}_\tau \subseteq \mathcal{F}_\sigma$.*
4. *If X is a progressively measurable process on $\underline{\Omega}$ and τ is almost surely finite, then X_τ is a well-defined[6] $\mathcal{F}_\tau$-measurable random variable.*

Proof. See [Caravenna 2011, p. 64]. □

A common way to define stopping times is indeed to look for the realization of a proposition φ. This proposition can be thought of as an event in $\mathcal{F}_\infty$, hence as an element of $\mathbb{F}_\infty$. As mentioned above, an event at infinity gives two sequences of sets, namely $\{\Box_t \varphi\}_{t\in I}$ and $\{\Diamond_t \varphi\}_{t\in I}$, which are 'adapted' in the sense that $\Box_t \varphi, \Diamond_t \varphi \in \mathcal{F}_t$ for every $t \in I$. Moreover, they are monotone (one increasing, one decreasing). We are going to develop some machinery to express facts about the realization of φ. We need some care to make such definitions because the set φ is only defined up to essential equivalence.

Proposition A.25. *Suppose $(\underline{S}, \underline{I}, K)$ has the singleton property, let $\varphi \in \mathbb{F}_\infty$. Then the following are stopping times:*

$$\begin{aligned} \tau_\varphi(\omega) &:= \inf\{t \in I \mid \mathrm{tr}_t\, \omega \leq_t \Box_t \varphi\}, \\ \tau^\varphi(\omega) &:= \sup\{t \in I \mid \mathrm{tr}_t\, \omega \leq_t \Diamond_t \varphi\}. \end{aligned} \tag{A.1.7}$$

Proof. Since $(\underline{S}, \underline{I}, K)$ has the singleton property, every $\mathrm{tr}_t\, \omega$ is an event of positive probability, thus the definitions of τ^φ and τ_φ are non-trivial.

[6] Up to the null set on which τ is infinite, where we could nevertheless set $X_\tau = C$ for some arbitrary constant C in the codomain.

Let now $s \in I$. Since $\{\Box_t \varphi\}_{t \in I}$ is an increasing sequence of sets, the following is straightforward:

$$\{\tau_\varphi \leq s\} = \{\omega \in \Omega \mid \operatorname{tr}_t \omega \leq_t \Box_t \varphi \text{ for some } t \leq s\} \leq_s \Box_s \varphi \in \mathbb{F}_s.$$

To prove the other essential inclusion (which then proves $\{\tau_\varphi \leq s\} \in \mathcal{F}_s$), we notice the following fact: if $\omega \in A \in \mathcal{F}_s$, then $\operatorname{tr}_s \omega \leq_s A$. Therefore suppose $A \in \mathcal{F}_s$ and $A \leq_s \Box_s \varphi$. For every $\omega \in A$, one has $\operatorname{tr}_s \omega \leq_s \Box_s \varphi$ by transitive property, hence $A \leq_s \{\tau_\varphi \leq s\}$. We conclude $\Box_s \varphi \simeq_s \{\tau_\varphi \leq s\}$, thereby proving the claim.

Conversely, the sequence $\{\Diamond_t \varphi\}_{t \in I}$ is decreasing, but since $\mathcal{F}_s$ is closed under complements, the following proves the claim nevertheless:

$$\{\tau^\varphi > s\} = \{\omega \in \Omega \mid \operatorname{tr}_t \omega \leq_t \Diamond_t \varphi \text{ for all } t \leq s\} \leq_s \Diamond_s \varphi \in \mathbb{F}_s.$$

Again, suppose $A \in \mathcal{F}_s$ and $A \leq_s \Diamond_s \varphi$, then for every $\omega \in A$ one has $\operatorname{tr}_s \omega \leq_s \Diamond_s \varphi$, hence $\operatorname{tr}_s \omega \leq_s \Diamond_s \varphi$. It follows $A \leq_s \Diamond_s \varphi$, thus $\{\tau^\varphi > s\} \simeq \Diamond_s \varphi$ and the proof is done. □

Remark A.25.1. In the proof we tacitly used the following fact: if $A \in \mathcal{F}_s$, $B \in \mathcal{F}_t$ ($t \leq s$) and $A \simeq_t B$, then we are able to conclude $A \in \mathcal{F}_t$. Indeed, $A \simeq_t B$ means $A \Delta B \in \ker \mathbb{P} \cap \mathcal{F}_t$, thus A differs from an event of $\mathcal{F}_t$ by a meager set of $\mathcal{F}_t$, which is again in $\mathcal{F}_t$, and therefore $A \in \mathcal{F}_t$.

Remark A.25.2. The stopping time τ_φ represents the earliest time at which we can say φ has happened. Dually, τ^φ tells the latest time at which we cannot say φ hasn't happened, or, in other words, the earliest time at which we have a disproof of φ.

The other two combinations of inf / sup and $\Box/\Diamond$ are degenerate: take, for instance

$$\sigma_\varphi(\omega) := \inf\{t \in I \mid \operatorname{tr}_t \omega \leq_t \Diamond_t \varphi\}.$$

This random variable is either 0 or ∞. In fact, since $\{\Diamond_t \varphi\}_{t \in I}$ is a decreasing sequence, if $\operatorname{tr}_t \omega \leq_t \Diamond_t \varphi$ for some t then actually $\operatorname{tr}_s \omega \leq_t \Diamond_s \varphi$ for every $s < t$, hence $\sigma_\varphi(\omega) = 0$. Conversely, if $\operatorname{tr}_t \omega \not\leq_t \Diamond_t \varphi$ for every $t \in I$, then by definition $\sigma_\varphi(\omega) = \infty$. Analogous considerations hold for

$$\sigma^\varphi(\omega) := \sup\{t \in I \mid \operatorname{tr}_t \omega \leq_t \Box_t \varphi\}.$$

In general, even one among τ_φ and τ^φ could be degenerate, when the event φ cannot be 'decided' with certainty in finite time, or, dually, when $\neg\varphi$ can't:

Example A.26. Consider the random experiment of the infinite sequence of coin flips, as formalized in Example 2.31. Let

$$\varphi = \{\text{we witness less than } k \text{ heads}\},$$
$$\psi = \{\text{we witness more than } k \text{ heads}\}.$$

The event φ cannot be assessed in a finite amount of time, since at any given time we cannot know how many heads we are going to witness in the future. Hence $\tau_\varphi \equiv \infty$, while τ^φ is the time at which we witness the $(k+1)$-th head in the sequence. Conversely, for the event ψ, τ^ψ is trivial (again, constantly ∞), since at no point the sequence we can exclude that sequence will contain more than k heads. On the other hand, τ_ψ is non-trivial and gives the time at which we witness the k-th head in the sequence.

Example A.27. Consider a Turing machine with a specific program preloaded. Obviously, it can be considered as a (discrete) transition system, hence it enjoys the singleton property. Then statements φ about the program which are positively undecidable (i.e. cannot be confirmed in finite time) give a degenerate τ^φ, whereas negatively undecidable ones (i.e. cannot be refuted in finite time, such as halting) give a degenerate τ_φ.

Systems without the singleton property. Even if the singleton property is crucial to be able to define τ^φ and τ_φ, it turns out we can still define the associated σ-fields. The pivotal observation is that

$$\{\tau_\varphi \le t\} \simeq_t \Box_t \varphi, \qquad \{\tau^\varphi \le t\} \simeq_t \Diamond_t \varphi.$$

Therefore, given $\varphi \in \mathbb{F}_\infty$ for the space of trajectories of any stochastic transition system $(\underline{S}, \underline{I}, K)$, the following definitions still make sense:

$$\begin{aligned} \mathcal{F}_{\tau_\varphi} &= \{A \in \mathcal{F}_\infty \mid (A \wedge \Box_t \varphi) \cap \mathcal{F}_t \neq \varnothing, \text{for all } t \in I\}, \\ \mathcal{F}_{\tau^\varphi} &= \{A \in \mathcal{F}_\infty \mid (A \wedge \Diamond_t \varphi) \cap \mathcal{F}_t \neq \varnothing, \text{for all } t \in I\}. \end{aligned} \tag{A.1.8}$$

The fact $A \wedge \Box_t \varphi$ doesn't determine a specific event in $\mathcal{F}_t$ isn't an issue, by what we observed in Remark A.25.1.

Modalities associated to stopping times. All in all, the interest we have in stopping times is that they provide a very expressive set of modalities, induced by the inclusion

$\mathcal{F}_\tau \subseteq \mathcal{F}_\infty$:

$$\begin{aligned}\Diamond_\varphi \psi &:= \Diamond_{\tau_\varphi} \psi = \psi \text{ hasn't been falsified yet by the time } \varphi \text{ is proved,}\\ \Box_\varphi \psi &:= \Box_{\tau_\varphi} \psi = \psi \text{ has already happened by the time } \varphi \text{ is proved,}\\ \Diamond^\varphi \psi &:= \Diamond_{\tau^\varphi} \psi = \psi \text{ hasn't been falsified yet by the time } \varphi \text{ is falsified,}\\ \Box^\varphi \psi &:= \Box_{\tau^\varphi} \psi = \psi \text{ has already happened by the time } \varphi \text{ is falsified.}\end{aligned} \tag{A.1.9}$$

Some evident identities:

Corollary A.27.1. *Let $\varphi, \psi \in \mathbb{F}_\infty$. Then we have*

$$\tau_{\varphi\cup\psi} = \tau_\varphi \wedge \tau_\psi, \qquad \tau^{\varphi\cap\psi} = \tau^\varphi \vee \tau^\psi,$$

from which it follows that

$$\Box_{\varphi\cup\psi} = \Box_\varphi \Box_\psi, \qquad \Diamond_{\varphi\cup\psi} = \Diamond_\varphi \Diamond_\psi. \tag{A.1.10}$$

A.2 A review of two process logics

A.2.1 CTL*: Full Computational Tree Logic

Temporal logics were born to help formalizing (and thus check) the behaviour of programs, especially concurrent ones, in an article of [Pnueli 1977][7]. Their syntaxes come equipped with a set of modal operators to express facts about the temporal realization of formulae about a process unfolding in time.

The two most common incarnation of temporal logic are LTL (*Linear Time Logic*) and CTL (*Computational Tree Logic*). Both are subsystems of CTL*, for which we'll discuss the semantics. All three flavours of temporal logic are modal logics defined on top of propositional logic, so they have a common '**atemporal** fragment'.

Atemporal propositions are built inductively by conjunction, disjunction and implication[8], starting from $\bot$ (false), $\top$ (true) and a set of **atomic propositional constants** P. The constants in P stand for specific properties of the system we are actually going to describe, hence is a 'parameter' of the logic (this is what it's meant by endogenous).

Example A.28. Consider, as a system, the semaphores controlling a road intersection. Atomic propositions here might be 'Semaphore 1 is green', 'Semaphore 2 is red', etc.

[7] Though it was far from the first time logic and time were discussed together. Even Aristotle concerned himself with the problems of putting logical truth and time together. See [Goranko and Rumberg 2020].

[8] Negation is defined by abbreviation: $\neg\varphi \equiv \varphi \to \bot$.

Propositions in the atemporal fragment can be thought to refer to the state of the system at the beginning of its evolution (in mathematical physics lingo, the *initial conditions* of the system).

What puts the 'temporal' in 'temporal logic' is a set of modalities which explicitly reference time and evolution. Here's where LTL, CTL and CTL* differ.

The main difference between LTL and CTL lies in the ability of quantifying over paths. In LTL, propositions refer to a single path (or to all paths, for what matters), whereas CTL and CTL* have quantifiers explicitly referring to multiple possible paths: that's why CTL and CTL* are called *branching time* logics, as opposed to LTL which is a *linear time* logic.

Therefore CTL and CTL* are suitable for non-deterministic systems. On the other hand, model checking for CTL(*) is much more difficult—computationally speaking—so one might want to stick with LTL as much as possible when describing a real-life scenario. See [Huth and Ryan 2004, Section 3.6], [Huth and Ryan 2004, Remark 3.19] for some insights on the pros and cons of each flavour.

Since CTL* encompasses both LTL and CTL, and strictly so, we are going to describe its set of modalities and its semantics, and then we'll point out which fragments correspond to LTL and CTL. For this part we refer to [Emerson and Halpern 1983], where CTL* was first introduced.

Formulae of CTL* are best introduced in a mutually recursive fashion by defining two 'kinds' of formulae:

1. *State formulae* are all atemporal formulae and all quantified path formulae: if φ is a path formula, then

$\mathsf{A}\varphi$	**for all paths**	φ holds in every path starting from the current state.
$\mathsf{E}\varphi$	**for some paths**	φ holds in at least one path starting from the current state

   are state formulae

2. *Path formulae* are all state formulae, conjunctions, disjunctions and implications of path formulae and, for φ, ψ path formulae:

$\mathsf{X}\varphi$	**next**	φ will be true in the next state.
$\psi\mathsf{U}\varphi$	**until**	ψ is true until φ becomes true, and this happens sometime in the future.

are path formulae.

Notice the separation between path and state formulae is really just a way to present them, since in the end all formulae are path formulae. Also the distinction is conceptual, and deemed important by the inventors of CTL^* [Emerson and Halpern 1983, p. 131]. That said, we stressed this fact since the semantics of these two classes of formulae is apparently different, while in reality everything can be expressed in terms of paths.

In addition to the temporal modalities we have defined, it is customary to add the following, which are defined in terms of the first ones[9]:

$\mathsf{F}\varphi \equiv \top\mathsf{U}\varphi$	**future**	φ will be true sometime in the future.
$\mathsf{G}\varphi \equiv \neg\mathsf{F}\neg\varphi$	**globally**	φ will always be true from now on.
$\psi\mathsf{W}\varphi \equiv \psi\mathsf{U}\varphi \vee \mathsf{G}\psi$	**weak until**	ψ is true until φ becomes true.
$\psi\mathsf{R}\varphi \equiv \neg(\neg\psi\mathsf{U}\neg\varphi)$	**release**	φ remains true up to the moment ψ becomes true, and this happens sometime in the future.

We then deduce the following useful identities [Huth and Ryan 2004, p. 184], [Reynolds 2001]. Here $\varphi = \psi$ means φ and ψ are equisatisfiable (in the sense of the $\models$ relation defined later):

$$\begin{aligned} \neg\mathsf{G}\varphi &= \mathsf{F}\neg\varphi, & \neg\mathsf{X}\varphi &= \mathsf{X}\neg\varphi, \\ \mathsf{G}(\varphi\wedge\psi) &= \mathsf{G}\varphi\wedge\mathsf{G}\psi, & \mathsf{F}(\varphi\vee\psi) &= \mathsf{F}\varphi\vee\mathsf{F}\psi, \\ \mathsf{F}\varphi &= \top\mathsf{U}\varphi, & \mathsf{G}\varphi &= \bot\mathsf{R}\varphi, \\ \psi\mathsf{U}\varphi &= \psi\mathsf{W}\varphi\wedge\mathsf{F}\psi, & \psi\mathsf{W}\varphi &= \psi\mathsf{U}\varphi\vee\mathsf{G}\varphi, \\ \psi\mathsf{W}\varphi &= \varphi\mathsf{R}(\varphi\vee\psi), & \psi\mathsf{R}\varphi &= \varphi\mathsf{W}(\varphi\wedge\psi). \end{aligned} \tag{A.2.1}$$

In light of these relations, it should be evident the choice of the 'primary' modalities (we chose X, F, U) and the 'derived' ones is not unique.

Finally, a formula in CTL^* is an LTL formula if it doesn't use the E modality; whereas it is a CTL formula is any path formula which is generated from state formulae using U and X only (hence no top level logical connectives, or negations, and U and X are applied just to state formulae).

[9] A pedant *sed* obligatory sidenote: in formulae, modal operators bind stronger than logical connectives, and unary operators bind stronger than binary ones.

A.2.1.1 Classical semantics

Definition A.29 (Reynolds 2001). A **model** for CTL^* is a transition system $(S, \rightsquigarrow)$ together with a labelling function $L : P \to \mathcal{P}(S)$, where P is the set of system-specific atomic propositions we mentioned earlier. One defines inductively what it means for a path $\pi = s_0 \rightsquigarrow s_1 \rightsquigarrow \cdots$ to satisfy a CTL^* formula:

1. $M, \pi \models \top$,

2. $M, \pi \not\models \bot$,

3. for each atomic proposition $p \in P$, $M, \pi \models p$ iff $s_0 \in L(p)$,

4. $M, \pi \models \varphi \wedge \psi$ iff $M, \pi \models \varphi$ and $M, \pi \models \psi$,

5. $M, \pi \models \varphi \vee \psi$ iff $M, \pi \models \varphi$ or $M, \pi \models \psi$,

6. $M, \pi \models \varphi \to \psi$ iff $M, \pi \models \psi$ whenever $M, \pi \models \varphi$,

7. $M, \pi \models \mathsf{A}\varphi$ iff for every path $\pi' = s_0 \rightsquigarrow \cdots$, where s_0 is the first state of π, $M, \pi' \models \varphi$,

8. $M, \pi \models \mathsf{E}\varphi$ iff there exists a path $\pi' = s_0 \rightsquigarrow \cdots$, where s_0 is the first state of π, such that $M, \pi' \models \varphi$,

9. $M, \pi \models \psi \mathsf{U} \varphi$ iff there is some $i \geq 1$ such that $M, \pi^i \models \psi$ and for all $j = 1, \ldots, i-1$ we have $M, \pi^j \models \varphi$,

10. $M, \pi \models \mathsf{X}\varphi$ iff $M, \pi^1 \models \varphi$.

A.2.2 PDL: Propositional Dynamic Logic

Propositional Dynamic Logic (PDL), is the exogenous companion of temporal logic. It is more similar to an abstract programming language than a logical system, since it codifies in its syntax all the necessary constructs to build the systems it talks about, namely (computer) programs. See [Harel, Kozen, and Tiuryn 2001] for a full elementary treatment of PDL.

In PDL there are two kinds of syntactic objects, formulae and programs. Both are defined by mutual recursion starting from a set of atomic propositions P (including $\top$, $\bot$) and a set of atomic programs A. If α and β are programs and φ, ψ are formulae, then

$\varphi \to \psi$	**implication**	
$\perp$	**false**	
$\langle\alpha\rangle\varphi$	**program possibility**	if α halts when started in the current state, then it may do so in a state where φ holds.

are formulae and

$\alpha;\beta$	**sequential composition**	execute α from the current state, then when it halts execute β from the last state of α.
$\alpha \cup \beta$	**choice**	execute either of α or β.
α^*	**iteration**	execute α an arbitrary but finite number of times.
$\varphi?$	**test**	do nothing if φ holds, hang if it doesn't.

are programs.

Moreover, the following abbreviations are introduced

$\mathsf{skip} := 1?$	**skip**	do nothing.
$\mathsf{fail} := 0?$	**fail**	hang (never halt).
$[\alpha]\varphi := \neg\langle\alpha\rangle\neg\varphi$	**program necessity**	if α halts when started in the current state, then it does so in a state in which φ holds.
$\mathsf{if}\ \varphi\ \mathsf{then}\ \alpha\ \mathsf{else}\ \beta := \varphi?;\alpha \bigcup \neg\varphi?;\beta$	**conditional choice**	if φ holds execute α, else β.
$\mathsf{while}\ \alpha\ \mathsf{do}\ \beta := (\varphi?;\alpha)^*;\neg\varphi?$	**conditional iteration**	iterate α as long as φ holds.

Remark A.29.1. We adopt the convention under which unary operators bind tighter than binary ones and ; binds tighter than $\cup$. Moreover, in light of the semantics we are about to give, both ; and $\cup$ are deemed associative.

A.2.2.1 Classical semantics

The semantics of PDL we are going to lay down is denotational, meaning we are going to construct some mathematical objects which represent the intended effect of programs.

Formally, the semantics of PDL uses again a family of labelled transition systems

$$(S, \{\rightsquigarrow_a\}_{a \in A}, L)$$

parametrized by the atomic programs, each analogous to what we have defined for temporal logics.

Then formulae are interpreted as subsets of S (corresponding to 'those states on which they hold'), while programs are interpreted as relations in S (corresponding to the input-output relations they encode). The interpretation will be **compositional**, meaning the interpretation of compound formulae and programs is built from that of their components.

In other words, the interpretation is going to be inductively defined on the structure of both programs and formulae. Thus the family $\{\rightsquigarrow_a\}_{a \in A}$ and the labelling $L : P \to \mathcal{P}(S)$ provide the 'base cases' of such induction.

Remark A.29.2. Below, $\circ$ denotes relational composition: if $R \subseteq H \times G$ and $T \subseteq G \times K$, then $T \circ R \subseteq H \times K$ is defined as

$$h\ T \circ R\ k \quad \underline{\text{iff}} \quad \text{there exists } g \in G \text{ such that } hRg \text{ and } gTk.$$

Notice a subset $R \subseteq H$ is equivalent to a relation $R \subseteq H \times 1$. That's how one should interpret the composition of a relation and a subset in the following definition.

Definition A.30 (Harel, Kozen, and Tiuryn 2001, p. 116)**.** A model of PDL (with set of atomic propositions P and set of atomic programs A) is a family of labelled transition systems $M = (S, \{\rightsquigarrow_a\}_{a \in A}, L)$. The interpretation of formulae and programs in M is defined inductively:

1. $\|\bot\|_M := \varnothing$,
2. $\|p\|_M := L(p)$ for any atomic formula $p \in P$,
3. $\|a\|_M := \rightsquigarrow_a$ for any atomic program $a \in A$,
4. $\|\varphi \to \psi\|_M = \|\varphi\|_M^{\complement} \cup \|\psi\|_M$,
5. $\|\langle\alpha\rangle\varphi\|_M := \|\alpha\|_M \circ \|\varphi\|_M$,
6. $\|\alpha;\beta\|_M := \|\alpha\|_M \circ \|\beta\|_M$,
7. $\|\alpha \cup \beta\|_M := \|\alpha\|_M \cup \|\beta\|_M$,
8. $\|\alpha^*\|_M := \bigcup_{n \in \mathbb{N}} \|\alpha\|_M^{\circ n}$,

9. $\|\varphi?\|_M := \Delta_S\|\varphi\|_M$.

Remark A.30.1. The interpretation of $\|\alpha^*\|_M$ is also known as the reflexive, transitive closure of $\|\alpha\|_M$, while the interpretation of $\|\varphi?\|_M$ is more explicitly given as $\{(s,s) \mid s \in \|\varphi\|_M\}$.

Remark A.30.2. Even if all the atomic programs are deterministic, the interpretations of $\cup$ and $.^*$ are non-deterministic relations. In fact,

$s\ \|\alpha \cup \beta\|_M\ s'$ <u>iff</u> executing α or β on s gets us to s',

$s\ \|\alpha^*\|_M\ s'$ <u>iff</u> we can execute α on s some finite number of times and get to s'.

Remark A.30.3. We remark the semantics of a non-halting program is $\varnothing$. In particular, if α is an atomic program which doesn't halt on a state $s \in S$, then $s\,\|\alpha\|_M\, s'$ is not defined for any state s'. Moreover, relational composition has $\varnothing$ as absorbing element. This means two things: first of all, if α does not terminate on state s, so does $\alpha;\beta$ for any other program β. Secondly, when interpreting a program containing either $\cup$ or $.^*$ (the 'indeterminate' operators), the semantics automatically prunes away non-halting execution paths in favour of halting ones.

Example A.31. To exemplify the last sentence, we analyze the meaning of $\mathsf{while}\ \varphi\ \mathsf{do}\ \alpha$:

$$\begin{aligned}
\|\mathsf{while}\ \varphi\ \mathsf{do}\ \alpha\|_M &= \|(\varphi?;\alpha)^*;\neg\varphi?\|_M \\
&= \|(\varphi?;\alpha)^*\|_M \circ \|\neg\varphi?\|_M \\
&= \left(\bigcup_{n\in\mathbb{N}} (\|\varphi?\|_M \circ \|\alpha\|_M)^{\circ n}\right) \circ \|\varphi?\|_M^C \\
&= \left(\bigcup_{n\in\mathbb{N}} (\Delta_S\|\varphi\|_M \circ \|\alpha\|_M)^{\circ n}\right) \circ \Delta_S\|\varphi\|_M^C.
\end{aligned}$$

We do not unwind the definition beyond this point, pain it would get too unwieldy to understand. Let's analyze what we got.

The first term is the reflexive, transitive closure of $\Delta_S\|\varphi\|_M \circ \|\alpha\|_M$. This relation represent the program that fails on a state $s \in S$ if φ isn't true for s, or if α doesn't terminate on s. Hence to repeat this program n times means to alternatively check φ against a state, then execute α, then check φ again on the result, and so on. Starting from a state s, there are two ways by which this can fail: either (1) φ stops being valid after a finite number of iterations, or (2) after a finite number of iterations I get to a state on which α doesn't terminate.

We can avoid (1) by iterating less. The number of iterations is arbitrary, which denotationally correspond to the fact that if iterating too much on s brings me to a non-halting state then simply this state won't be related to s in the transitive closure of $\Delta_S \|\varphi\|_M \circ \|\alpha\|_M$. Let's look at the second fragment. This doesn't terminate on the states which satisfy φ. Hence, composed with the previous, prunes out all the 'execution paths' in which we didn't iterate enough times and so we executed $\neg\varphi$? while φ was still true. Eventually, the program will terminate on a state s iff α halts on s (avoiding (2)), and the resulting state will be the one we get by repeating α the least number of times before φ stops being valid (possibly 0).

This is precisely the denotational semantics of a while loop.

A.3 Semantics in stochastic transition systems?

We were not able to find a semantics of CTL^* in a stochastic transition system $(\underline{S}, \underline{I}, K)$, if not 'trivially', i.e. by merely reproducing the semantics in the discrete case.

In particular we were not able to use the filtration modalities to define the temporal operator of CTL^*. Perhaps the only exception is given by A and E, which appear to be captured by $\Box_0$ and $\Diamond_0$, respectively.

The modalities (A.1.9) arising from 'realizing times' of a proposition ψ were a promising candidate to capture the meaning of W or U. Unfortunately, this is yet again not the case. In fact suppose we defined

$$\|\psi \mathsf{W} \varphi\| := \Box^{\|\psi\|} \|\varphi\|,$$

as suggested by the 'interpretation' given to $\Box^\psi$ in (A.1.9). Then, by the identities (A.2.1), one would obtain

$$\begin{aligned}
\|\psi \mathsf{R} \varphi\| &= \|\varphi \mathsf{W} (\varphi \wedge \psi)\| = \Box^{\|\varphi\|} \|\varphi \wedge \psi\|, \\
\|\psi \mathsf{U} \varphi\| &= \|\neg(\neg\psi \mathsf{R} \neg\varphi)\| = \Diamond^{\|\neg\varphi\|} \|\varphi \vee \psi\|, \\
\|\mathsf{F}\varphi\| &= \|\top \mathsf{U} \varphi\| = \Diamond^{\|\neg\varphi\|} \|\top\|, \\
\|\mathsf{G}\varphi\| &= \|\bot \mathsf{R} \varphi\| = \Box^{\|\varphi\|} \|\bot\|.
\end{aligned}$$

However, the only reasonable definitions for $\|\bot\|$ and $\|\top\|$ are $\varnothing$ and S^I, respectively, which would entail $\mathsf{F}\varphi$ and $\mathsf{G}\varphi$ are always interpreted to S^I and $\varnothing$, respectively.

The crux of the problem seems to be in the fact the modal framework of the modalities doesn't allow one to start evaluating a proposition at an arbitrary suffix of a path, as required by F, for instance. The proposition is either true from the beginning or never true.

The semantics for PDL is disappointing too. By defining $sSIK = (S, \{\rightsquigarrow_t^K\}_{t\in I}, \Sigma \hookrightarrow \mathcal{P}(S))$, one gets a a model of $\mathsf{PDL}(\Sigma, I)$, i.e. PDL where the set of atomic propositions is Σ and the set of atomic programs is I.

We recall the transitions $\{\rightsquigarrow_t^K\}_{t\in I}$ have been defined in Remark A.17.1 as the 'discrete' version of K. It's actually nice to observe the programs $t; s$ and $t + s$ have the same interpretation in the model, by virtue of the fact K is a monoidal action, hence

$$\|t; s\|_{sSIK} = \rightsquigarrow_t^K \circ \rightsquigarrow_s^K = \rightsquigarrow_{t+s}^K = \|t + s\|_{sSIK}.$$

On the other hand, the modalities $[t]$ and $\langle t\rangle$ do not correspond to the filtration modalities $\Box_t$ and $\Diamond_t$: this is easy to see since

$$[t][s]\varphi = [t; s]\varphi = [t + s]\varphi$$

while

$$\Box_t\Box_s\varphi = \Box_{t\wedge s}\varphi.$$

as observed in (5.1.6).

Maybe this is obvious, since we are interpreting PDL on states, while the filtration modalities act on paths. Indeed, one might define the following **shift relations**:

$$\omega +_t^K \omega' \quad \underline{\text{iff}} \quad \text{for every } s \in I, \omega'_{s+t} = \omega_s$$

Then the model $pSIK = (S^I, \{+_t^K\}_{t\in I}, \Sigma^{\otimes I} \hookrightarrow \mathcal{P}(S^I))$ is a model of $\mathsf{PDL}(\Sigma^{\otimes I}, I)$, i.e. PDL where the set of atomic propositions is $\Sigma^{\otimes I}$ and the set of atomic programs is I. As before, one has $t; s = t + s$ as programs, and thus the filtration modalities do not show up.

These results are disappointing, yet some future research questions remain open:

Question 4.

1. *Are stochastic transition systems suitable for a generalized Kripke semantics nevertheless? Perhaps in their 'open form' (Remark A.22.1).*
2. *Are there systems in process logics which capture the semantics of the modal operators $\Diamond_t \dashv \Box_t$? We have the feeling one may have to look at either epistemic logic [Rendsvig and Symons 2019], which is used to speak about the knowledge and beliefs of various agents; or define a new polymodal logic[10] altogether.*
3. *Could we need to generalize semantics even more? For example, probabilistic semantics might be relevant, see [Feldman and Harel 1984].*

[10] Indeed, this fragment of μ-calculus is trivially captured by filtration modalities, see [Venema 2008].

4. *Can process logics provide an original perspective and useful insights in the study of stochastic calculus, or, more generally, stochastic dynamical systems?*

Errata corrige

Theorem 4.8 and Theorem 4.9 are wrong as stated. Theorem 4.8 should read

Theorem. *The topos* **Scott**$[\underline{\Omega}]$ *has a (Dedekind) real numbers object* $\mathcal{R}_{\underline{\Omega}}$ *given by*

$$\mathcal{R}_{\underline{\Omega}} := (\mathbf{Msbl}(\underline{\Omega}, \mathbb{R}),\ \{- = -\})$$

which is indeed what's proven in the given reference (namely [Wendt 1996, Section 4]), although Wendt states the theorem for sheaves. To get this version one has to pass $\mathcal{R}_{\underline{\Omega}}$ through the functors defined in Theorem 3.13.

Theorem 4.9 is just wrong, unfortunately. The problem lies in the totality of the relation F defined in Equation 4.4.4, which isn't guaranteed in general (see below for additional assumptions which restore this property). To see why, notice that totality of F means that for any given measurable $X : \underline{\Omega} \to \mathbb{R}$, one has

$$\Omega \leq \bigvee_{x \in \mathbb{R}} \{X = x\} \tag{5.3.1}$$

This is hardly the case if $\underline{\Omega}$ is not discrete and X is continuous, though. For the sake of concreteness, consider $\underline{\Omega} = ([0,1], \mathcal{B}([0,1]), \mathrm{d}x)$ and the random variable $X(\omega) = \omega$. Then clearly each event $\{X = x\}$ is going to have zero measure, hence it's the bottom of the algebra $\mathbb{F}$. It follows that the supremum in (5.3.1) is the bottom, and not the top as desired.

Theorem 4.9 can be modified in many ways in order to make it true:

1. Assume that Ω is discrete. In that case each event $\{X = x\}$ has positive measure whenever x actually belongs to the image of X.

2. Replace $\mathbb{R}$ with $\mathbb{N}$, $\mathbb{Z}$, or $\mathbb{Q}$. In these cases, the image of any random variable X is at most countable, therefore the preimages $\{X = x\}$ form a countable cover of Ω in $\mathbb{F}$, i.e. the family $\{X = x\}$ doesn't collapse anymore.

3. Replace $\mathbf{Msbl}(\underline{\Omega}, \mathbb{R})$ with its subobject of simple functions, i.e. those measurable functions $\underline{\Omega} \to \mathbb{R}$ whose image is at most countable. Notably, such functions are dense in $\mathbf{Msbl}(\underline{\Omega}, \mathbb{R})$.

Bibliography

Adachi, Takanori and Yoshihiro Ryu (2016). *A Category of Probability Spaces*.

Ambrosio, Luigi, Nicola Gigli, and Giuseppe Savaré (2008). *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media.

Aumann, Robert J. et al. (1961). "Borel structures for function spaces". In: *Illinois Journal of Mathematics* 5.4, pp. 614–630.

Baez, John C and Blake S Pollard (2017). "A compositional framework for reaction networks". In: *Reviews in Mathematical Physics* 29.09, p. 1750028.

Bell, J.L. (2005). *Set Theory: Boolean-Valued Models and Independence Proofs*. 3rd. Oxford University Press.

Blechschmidt, Ingo (2017). "Using the internal language of toposes in algebraic geometry". PhD thesis. University of Augsburg. URL: https://rawgit.com/iblech/internal-methods/master/notes.pdf.

Caravenna, Francesco (2011). *Moto Browniano e analisi stocastica*. Version 3.3. URL: http://staff.matapp.unimib.it/~fcaraven/did1112/ps/dispense-3.3-single.pdf.

Clark, Stephen, Bob Coecke, and Mehrnoosh Sadrzadeh (2008). "A compositional distributional model of meaning". In: *Proceedings of the Second Quantum Interaction Symposium (QI-2008)*. Oxford, pp. 133–140.

Coecke, Bob and Ross Duncan (2011). "Interacting quantum observables: categorical algebra and diagrammatics". In: *New Journal of Physics* 13.4, p. 043016.

Cohn, Donald L. (2013). *Measure Theory: Second Edition*. 2nd ed. Birkhäuser Advanced Texts Basler Lehrbücher. Birkhäuser Basel.

Emerson, E. Allen (1997). "Model checking and the μ-calculus". In: *DIMACS series in discrete mathematics* 31, pp. 185–214.

Emerson, E. Allen and Joseph Y. Halpern (1983). ""Sometimes" and "not never" revisited: on branching versus linear time (preliminary report)". In: *Proceedings of the 10th ACM*

SIGACT-SIGPLAN symposium on Principles of programming languages, pp. 127–140. URL: https://dl.acm.org/doi/10.1145/567067.567081.

Feldman, Yishai A. and David Harel (1984). "A probabilistic dynamic logic". In: *Journal of Computer and System Sciences* 28.2, pp. 193–215. URL: https://www.sciencedirect.com/science/article/pii/0022000084900655.

Fong, Brendan (2016). "The algebra of open and interconnected systems". In: *arXiv preprint arXiv:1609.05382*.

Fritz, Tobias (2009). "Convex spaces I: Definitions and examples". In: URL: https://arxiv.org/abs/0903.5522.

— (2019). "A synthetic approach to Markov kernels, conditional independence, and theorems on sufficient statistics". In: *arXiv preprint arXiv:1908.07021*.

Fritz, Tobias and Paolo Perrone (2019). "A probability monad as the colimit of spaces of finite samples". In: *Theory and Applications of Categories* 34. URL: http://arxiv.org/abs/1712.05363.

Fritz, Tobias, Paolo Perrone, and Sharwin Rezagholi (2019). "Probability, valuations, hyperspace: Three monads on **Top** and the support as a morphism". In: *arXiv preprint arXiv:1910.03752*. URL: https://arxiv.org/abs/1910.03752.

Ghani, Neil et al. (2018). "Compositional game theory". In: *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science*, pp. 472–481.

Giry, Michele (1982). "A categorical approach to probability theory". In: *Categorical aspects of topology and analysis*. Springer, pp. 68–85.

Goranko, Valentin and Antje Rumberg (2020). "Temporal Logic". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2020. Metaphysics Research Lab, Stanford University.

Grimmett, Geoffrey R. and David R. Stirzaker (2001). *Probability and random processes*. Third edition. Oxford University Press.

Harel, David, Dexter Kozen, and Jerzy Tiuryn (2001). "Dynamic logic". In: *Handbook of philosophical logic*. Springer, pp. 99–217.

Heunen, Chris et al. (2017). "A convenient category for higher-order probability theory". In: *2017 32nd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. IEEE, pp. 1–12.

Hodges, Wilfrid (2018). "Tarski's Truth Definitions". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Fall 2018. Metaphysics Research Lab, Stanford University.

Hodges, Wilfrid and Jouko Väänänen (2019). "Logic and Games". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Fall 2019. Metaphysics Research Lab, Stanford University.

Huth, Michael and Mark Ryan (2004). *Logic in Computer Science: Modelling and reasoning about systems*. Cambridge university press.

Jackson, Matthew Tobias (2006). "A sheaf theoretic approach to measure theory". PhD thesis. University of Pittsburgh.

Jacobs, Bart (2005). *Categorical logic and type theory*. Studies in Logic and the Foundations of Mathematics vol. 141. Elsevier.

Jech, Thomas (2006). *Set Theory*. The Third Millennium Edition, revised and expanded. Academic Press.

Johnstone, Peter T. (2002). *Sketches of an elephant: A topos theory compendium*. Vol. 2. Oxford University Press.

Lawvere, Francis William (1962). "The category of probabilistic mappings-with applications to stochastic processes, statistics, and pattern recognition". In: URL: https://github.com/mattearnshaw/lawvere/blob/master/pdfs/1962-the-category-of-probabilistic-mappings.pdf.

Moerdijk, Ieke and Saunders MacLane (1992). *Sheaves in geometry and logic*. Springer.

Moschovakis, Joan (2018). "Intuitionistic Logic". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2018. Metaphysics Research Lab, Stanford University.

Myers, David Jaz (2020). "Double Categories of Open Dynamical Systems". In: *arXiv preprint arXiv:2005.05956*. URL: https://arxiv.org/abs/2005.05956.

nLab authors (2020). *monads of probability, measures, and valuations*. URL: https://ncatlab.org/nlab/show/monads+of+probability%2C+measures%2C+and+valuations (visited on).

Pasquali, Fabio (2016). "Remarks on the tripos to topos construction: Comprehension, extensionality, quotients and functional-completeness". In: *Applied Categorical Structures* 24.2, pp. 105–119.

Picado, Jorge and Aleš Pultr (2011). *Frames and Locales: topology without points*. Springer Science & Business Media.

Pitts, Andrew M. (1982). "The theory of triposes". PhD thesis. University of Cambridge. URL: http://www.cl.cam.ac.uk/~amp12/papers/thet/thet.pdf.

— (1989). "Notes on categorical logic". In: *Graduate Lecture Course, Cambridge University Computer Laboratory*.

Pitts, Andrew M. (2002). "Tripos theory in retrospect". In: *Mathematical structures in computer science* 12.3, pp. 265–279. URL: https://www.sciencedirect.com/science/article/pii/S1571066104001070.

Pnueli, Amir (1977). "The temporal logic of programs". In: *18th Annual Symposium on Foundations of Computer Science (sfcs 1977)*. IEEE, pp. 46–57. URL: https://ieeexplore.ieee.org/abstract/document/4567924/.

Rendsvig, Rasmus and John Symons (2019). "Epistemic Logic". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2019. Metaphysics Research Lab, Stanford University.

Reyes, Gonzalo E et al. (1991). "A topos-theoretic approach to reference and modality." In: *Notre Dame Journal of Formal Logic* 32.3, pp. 359–391. URL: https://projecteuclid.org/euclid.ndjfl/1093635834.

Reyes, Gonzalo E and Houman Zolfaghari (1991). "Topos-theoretic approaches to modality". In: *Category Theory*. Springer, pp. 359–378.

Reynolds, Mark (2001). "An axiomatization of full computation tree logic". In: *Journal of Symbolic Logic*, pp. 1011–1057.

Saeki, Sadahiro (1996). "A proof of the existence of infinite product probability measures". In: *The American mathematical monthly* 103.8, pp. 682–683.

Scott, Dana (1967). "A Proof of the Independence of the Continuum Hypothesis". In: *Mathematical systems theory* 1.2, pp. 89–111. URL: https://www.karlin.mff.cuni.cz/~krajicek/scott67.pdf.

Simpson, Alex (2012). "Measure, randomness and sublocales". In: *Ann. Pure Appl. Logic* 163.11, pp. 1642–1659.

— (2017). "Probability sheaves and the Giry monad". In: *7th Conference on Algebra and Coalgebra in Computer Science (CALCO 2017)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

Solovay, Robert M. (1970). "A model of set-theory in which every set of reals is Lebesgue measurable". In: *Annals of Mathematics*, pp. 1–56. URL: https://www.jstor.org/stable/1970696.

Spivak, David I. (2019). "Generalized Lens Categories via functors $\mathbf{C}^{\mathrm{op}} \to \mathbf{Cat}$". In: *arXiv preprint arXiv:1908.02202*.

Stirling, Colin (2001). *Modal and Temporal Properties of Processes*. 1st ed. Texts in Computer Science. Springer-Verlag New York.

Streicher, Thomas (2011). *Fibered View of Geometric Morphisms*. URL: https://www2.mathematik.tu-darmstadt.de/~streicher/FIBR/fvgm.pdf (visited on 05/04/2020).

Tao, Terence (2011). "An introduction to measure theory". In: URL: https://terrytao.wordpress.com/books/an-introduction-to-measure-theory/.

Van Oosten, Jaap (2008). *Realizability: an introduction to its categorical side*. Elsevier.

Venema, Yde (2008). *Lectures on Modal μ-calculus*. URL: https://web.archive.org/web/20160305061922/http://philo.ruc.edu.cn/logic/reading/Yde/mu0612.pdf (visited on 07/08/2020).

Vickers, Steven L. (2011). "A monad of valuation locales". In: URL: https://www.cs.bham.ac.uk/~sjv/Riesz.pdf.

Wendt, Michael A. (1996). "Measurable Hilbert sheaves". In: *Journal of the Australian Mathematical Society* 61.2, pp. 189–215.