Compositionnalité des monades par lois de distributivité faibles

On the compositionality of monads via weak distributive laws

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 573, Interfaces : matériaux, systèmes, usages
Spécialité de doctorat : Informatique
Unité de recherche : Université Paris-Saclay, CentraleSupélec, Mathématiques et Informatique pour la Complexité et les Systèmes, 91190, Gif-sur-Yvette, France
Référent : CentraleSupélec

Thèse présentée et soutenue à Gif-sur-Yvette, le 19 octobre 2021, par

Alexandre GOY

Composition du Jury

Christine TASSON Professeure, Sorbonne Université	Rapporteure & Examinatrice
Daniele VARACCA Professeur, Université Paris-Est Créteil	Rapporteur & Examinateur
Jean GOUBAULT-LARRECQ Professeur, ENS Paris-Saclay	Examinateur
Gordon PLOTKIN Professeur, University of Edinburgh	Examinateur
Damien POUS Directeur de recherche CNRS, ENS Lyon	Examinateur
Ana SOKOLOVA Maîtresse de conférences, Universität Salzburg	Examinatrice

Direction de la thèse

Marc AIGUIER Professeur, CentraleSupélec	Directeur de thèse
Daniela PETRIŞAN Maîtresse de conférences, Université de Paris	Co-encadrante

Thèse de doctorat

NNT : 2021UPAST080

Remerciements

Une thèse, c'est plus dur qu'on ne le pense, et même en entendant cette phrase à répétition, on ne s'en rend compte que quand on y est. La chose aurait été impossible sans la présence précieuse d'une foultitude de personnes dont j'espère oublier ici le plus petit nombre possible.

Dans un cours d'informatique théorique égaré au milieu d'un cursus d'ingénieur dont le sens m'échappait, un orateur passionné fascine les étudiants. Son nom est celui de mon directeur de thèse : Marc Aiguier, auquel vont mes premiers remerciements. Après m'avoir conseillé la perle qu'est le master LMFI, après m'avoir aidé à trouver un stage palpitant aux Pays-Bas, c'était une suite logique pour moi de poursuivre une thèse sous ta direction. Je n'ai finalement eu la pleine mesure de tes qualités humaines qu'au beau milieu de cette thèse, où tu as été incroyablement compréhensif envers mon projet de réorientation de sujet. Pour cette grande liberté et cette immense confiance que tu m'as accordée, Marc, merci.

Au pied d'un avion grec, ramenant à Paris les participants de ma toute première conférence, une jeune chercheuse me reconnaît et lance la conversation. Son nom est celui de ma co-encadrante : Daniela Petrişan. Daniela, je ne pourrai te remercier assez pour tout ce que tu as fait pour moi pendant cette thèse – depuis le moment où, à Nice, tu as accepté d'encadrer cet étudiant en questionnement, jusqu'au bout des relectures approfondies de mon manuscrit, en passant par les encouragements en coulisse dans les salles virtuelles de l'ère du confinement. Ton incroyable clairvoyance sur les sujets de recherche qui nous passionnent a bien sûr été une précieuse alliée tout au long de ces années. Sans tes intuitions qui ont fait germer en moi tant d'idées, ce manuscrit s'arrêterait ici.

Mais - peut-être au grand dam du lecteur - nous sommes loin d'en avoir fini, et je tiens dès à présent à remercier chaleureusement les membres du jury. Merci à mes rapporteurs, eux que j'ai croisés au fil du doctorat, Christine Tasson et Daniele Varacca, dont les retours détaillés ont grandement contribué à améliorer ce texte. Merci à eux et à Jean Goubault-Larrecq, Gordon Plotkin, Damien Pous et Ana Sokolova pour m'honorer de leur participation au jury. Tant de chercheurs chevronnés dont j'admire le travail réunis pour discuter du mien, c'est vertigineux.

Bien d'autres ont influencé directement le travail qui va suivre. Merci à Isabelle Bloch, pour m'avoir encadré avant la réorientation du sujet, et m'avoir aidé dès que cela était possible - en particulier en m'incitant à aller à Nice, où tout s'est éclairci. Merci à Arthur, à qui je dois toujours une bière, à Andreas, dont le franc-parler a été crucial, à Arnaud et Fabienne, à Christophe, à Leonor, pour leur point de vue apaisé et lucide. Merci à Alexis Saurin, non seulement pour avoir été un responsable de master dévoué, mais aussi pour cette discussion autour d'un bol de nouilles. Scientifiquement, je dois beaucoup à Jurriaan Rot, qui m'a fait découvrir le monde merveilleux des coalgèbres en encadrant mon stage à Nijmegen. Il se posait alors la question de l'existence d'une double monade des parties – deux ans plus tard, c'est en tombant par hasard sur l'article de Klin et Salamanca qui répondait par la négative que je me suis intéressé de nouveau à ces sujets. Merci à Bartek Klin, Paolo Perrone, Thomas Colcombet, Axel Osmond, Ralph Sarkis ainsi que le Topos Reading Group pour les inspirantes discussions qui ont fini par laisser une empreinte ici où là. Merci à Maaike Zwart pour avoir osé consacrer sa thèse à des résultats d'impossibilité, et à laquelle le titre du présent document est un clin d'œil.

Le travail qui est sous vos yeux clôt un chapitre commencé il y a bien longtemps : celui des études. Je voudrais rendre hommage aux trois professeurs non encore cités qui ont façonné le scientifique que je suis. Le premier, Jean-François Ahyerre, a su éveiller en classe de première mon goût pour la science en me faisant envisager chaque question sans réponse comme un défi. La seconde, Carine Sort, a déchaîné un intérêt brûlant pour les mathématiques qui ne s'est pas consumé depuis l'époque du lycée. Le troisième, Christophe Cuny, m'a initié avec une bienveillance intarissable aux arcanes de la recherche au sein du Parcours Recherche de ce qui était alors l'Ecole Centrale Paris. Eux et d'autres que je n'oublierai jamais - je pense en particulier à Aline Girard, François Mattei, Jean Nougayrède, Clément de Seguins Pazzis, Erick Herbin, Pauline Lafitte, François Le Maître et bien évidemment François Lussier - tous à leur façon ont gravé en moi le désir de transmettre.

Cette soif d'enseigner n'a pas été en reste durant ces trois ans. Ce fut un plaisir de créer, de dispenser et de gérer avec Pascale le Gall et Marc Aiguier ces cours d'informatique (plus ou moins) théorique qui imprimeront encore longtemps, je l'espère, leur marque sur le cursus ingénieur de CentraleSupélec. Je tiens d'ailleurs à adresser des remerciements appuyés à tous les élèves qui ont croisé ma route : votre regard

neuf sur ces sujets que l'on croit pourtant bien connaître force l'humilité.

Administrativement, j'ai pu compter sur la compétence sans faille de Sylvie Dervin puis Fabienne Brosse pour le laboratoire, et de Suzanne Thuron – dite : la formidable – pour l'Ecole Doctorale. Un grand merci, continuez de simplifier la vie de ceux qui arriveront ensuite ! Merci également à Quentin Touzé, de la bibliothèque, pour m'avoir indiqué que la section remerciements n'avait pas de limite de pages.

Une thèse, c'est aussi des camarades sur le même bateau. Merci aux collègues Adrien Dekkers, Romain Pascual, Antonin Della Noce, et à tous les autres, pour ces parties de babyfoot et tant de délicieux repas du CROUS. Merci à Salim Nibouche pour vouer avec moi un culte à La Monade. Mention spéciale à Brice Hannebicque, qui se situe une année en avance depuis la nuit des temps, et me communique l'emplacement des chausse-trappes au fur et à mesure qu'il les découvre. Bravo pour là où tu es maintenant, tu le mérites[1]. A mon acolyte Erwan Mahé, je suis heureux que les coïncidences de la vie m'aient donné l'opportunité de mieux te connaître. Nos anecdotes du midi, nos sessions de natation à la Vague ou à Orsay, nos randonnées le long de l'Yvette et nos fous rires dans l'open space me manqueront.

Produire une thèse est une activité qui paraît si solitaire, et pourtant la liste des amis qui m'ont entouré sans faillir durant cette épreuve est encore longue. Par qui commencer ? Probablement par Jean, qui il y a déjà douze ans (oui, ça fait mal) m'informait qu'on pouvait faire des maths par plaisir. De rival à meilleur ami, il n'y avait qu'un pas et un passage chez le coiffeur. Bien des années d'études plus tard, les quatre thèses parallèles dans quatre disciplines différentes de ces deux couples curieusement symétriques, Alice et Jean versus Edwige et Alexandre, sont rythmées par de bienvenues rencontres où les maîtres mots sont jeux de société et brownie. Merci à vous deux pour être là. Merci à tous les membres de la Compagnie du Navet, fidèles au poste et aux réveillons du Nouvel An : Nicolas V, pour m'avoir informé à temps du Problème Récurrent, Pierre, pour être à la fois témoin, streamer et fan numéro 1, Alexandre B, pour partager en plus d'un prénom une passion malsaine pour le split push, Nicolas D, pour sa tempérance et sa passion pour le rôle d'avocat du diable, Joseph, pour me dire que c'était bien tenté alors que je viens de donner ma dame en un coup, Corentin, pour être Corentin, et Edwige, pour les raisons que l'on

[1] J'espère juste que tu n'es pas tombé dans un trou avec des pics entre le moment où j'écris ces lignes et celui où tu les lis.

verra plus tard. Dans la famille rythmer une thèse, j'appelle maintenant les rôlistes de la bande, dont la longue campagne de D&D a vu naître et se clore ce doctorat : merci à Braise Pascal, Ryze, Uleek et au MJ pour avoir emmené ce pataud de Bœuf dans tant d'aventures. C'est l'occasion de remercier Héloïse, une thésarde de plus, pour son enthousiasme débordant et la qualité de sa peinture. Dans un univers parallèle, merci à Nyanyel, Gemma, Howerd et feu Irwen pour vos vagabondages avec Larsen Lutin dans la vallée Picorine. Merci à Camille et Jonathan, des amis comme on n'en fait plus, toujours prêts à résoudre une énigme ou à lancer une raclette. Merci aux quatre autres Assa-Cinq des Templiers, mes amis de toujours, pour m'avoir permis si souvent de me changer les idées autour d'une bière, d'un Mario Kart ou d'un film estampillé The Asylum : Nicolas T dit Toto, Samuel dit Umut, Romain dit Rufus, Mathia dit Matmut. Votre présence est inestimable.

Merci à Ludwig van Beethoven, qui a rempli trois semaines d'oreilles pendant la phase de rédaction, suite à un conseil trouvé un recoin de la thèse d'Alexis Saurin. Merci à la synthwave pour avoir rempli les 153 autres semaines, en me renvoyant dans le passé, loin vers des temps nostalgiques que je n'ai pas vécus. Merci à ma guitare pour m'avoir permis d'exprimer ce qu'il n'était pas possible d'exprimer autrement.

Ce manuscrit a été rédigé en grande partie dans une certaine maison normande, où je me suis reclus deux longs mois qui furent les plus productifs que j'aie jamais connus. Les mots sont insuffisants pour exprimer ma gratitude envers mes grands-parents, Papy Cool et Mamie B, qui m'ont hébergé et se sont assurés que je ne manque de rien pendant ce sprint final. Sans vous, j'y serais encore. Merci également à Mamie Cotte, aux Goy de Gravigny et de Marcilly-la-Campagne, aux Lesage d'Aix, pour avoir envers et contre tout cherché à comprendre en quoi consistait ma thèse, et m'avoir demandé régulièrement des nouvelles alors que mes réponses incompréhensibles en auraient découragé plus d'un.

Merci à Baptiste et Claire, un frère, une sœur, pour avoir grandi et continuer de grandir ensemble. Je vous aime. Si vous vous demandez jusqu'à quand j'ai tenu avant de pleurer en rédigeant ce texte, c'était maintenant. Merci à Papa et Maman, vous qui m'avez toujours plus que soutenu, cru, encouragé, accompagné, aidé, sauvé parfois, aimé. Je vous aime aussi.

Au cours de cette thèse, j'ai rejoint, pour de bon, une autre famille. Merci à Benoît et Anne pour m'avoir accepté dès le tout début. Avec vous, les légendes sur les beaux-parents ont du souci à se faire. Merci à Marie-Christine, à Marie, à Clara, à Renaud, pour leur accueil dans la famille et leur présence à ce féérique mariage en tout petit comité. D'ailleurs, merci à Manu pour son accompagnement de Père hors pair. Merci aussi à Nicolas B et Tom, qui peut-être finiront par me rejoindre : j'aurais aimé me réunir avec vous et les autres cet été - ce n'est que partie remise. Avant de passer à la personne la plus importante de ces remerciements, il faut en mentionner un autre qui a rejoint une nouvelle famille. Bienvenue à Edouard, j'espère qu'un jour tu auras moins peur.

J'ai attendu jusqu'au dernier moment avant de parler d'elle, car je sais qu'elle sera gênée. Mais tu ne t'en sortiras pas comme ça. Edwige Nortier, je suis si heureux de t'avoir épousée. Merci pour ton soutien, ton rire, ton incorrigibilité, ta fragilité et ta force, ton amour. Toi qui as commencé cette thèse ma petite amie et l'a finie ma femme, j'ai hâte de ce que nous allons construire ensemble.

Abstract

Monads are widely used in computer science to model computational effects. To represent complex systems, compositionality of monads is therefore crucial. One can usually compose two monads using distributive laws. When no distributive law exists, it is sometimes still possible to recover what looks like a composite effect by using a *weak* distributive law. The phenomenon occurs when combining probabilistic choice with non-deterministic choice, or when combining non-deterministic choice with itself.

This thesis leverages and enhances the framework of weak distributive laws towards applications in computer science. Firstly, we focus on the two most-known examples where distributive laws fail in the category of sets. The origin of scattered results of the literature is explained through the lens of weak distributive laws. This includes composition of equational theories for non-determinism and probability as well as coalgebraic constructions for probabilistic automata and alternating automata. Secondly, aiming at applications in the semantics of programming languages, we study how to obtain laws in other categories. Notably, we generalise weak self-distribution of non-deterministic choice to arbitrary toposes and compact Hausdorff spaces.

Contents

Introduction

There is a recent interest in probabilistic programming amongst the theoretical computer science community [144, 74, 73, 133, 48, 40, 49], fitting within a long-established tradition of studying foundations of computational effects. These works, as well as the two standard abstractions of computational effects that are Lawvere theories [120, 75] and monads [109, 119], are all formulated in the language of category theory.

Mathematics is the art of giving the same name to different things, said Henri Poincaré [121]. Category theory exports this philosophy to a whole new level. Designed by Eilenberg and MacLane in the 1940s, category theory provides a unifying structuralist viewpoint on mathematics and has been importantly used in theoretical computer science. As an abstract standpoint, it encompasses a wide range of constructions, leading to the identification of common structures or links that a more down-to-earth approach would be unlikely to underline. As a research tool, it is a body of heuristics allowing to find meaningful mathematical constructions. As a common language, it makes science more communicable between communities of computer scientists, logicians, mathematicians, physicists – e.g., the topic of this computer science thesis, weak distributive laws, has been introduced independently by a mathematical physicist (Gabriella Böhm) and a mathematician (Ross Street).

Context

Monads. *Monads* are a central concept of category theory. We shall follow the approach initiated by Moggi [109] by thinking of monads as *computational effects*. In a nutshell, a monad consists of three pieces of data that can be interpreted as the effect itself, a procedure to create an effect, and a procedure to collapse two of these effects into one. Coming back to probability, the essence of randomness is first depicted by the renowned Giry monad [56], which gives rise to many variations including the probabilistic powerdomain of Jones and Plotkin [84]. There is recent work exhibiting a monad with a probabilistic flavour in the category of so-called quasi-Borel spaces [74, 73], which is well-behaved for programming purposes – but also there are other approaches somewhat orthogonal to monads, e.g. [48, 40, 49]. Before the non-categorically-aware reader is lost in the technical aspects of this work, we shall provide a concrete example of a monad and what it intuitively means.

Non-deterministic choice. The *powerset monad*, importantly used in the sequel, abstracts away the concept of a *non-deterministic choice*. Consider a set of all possible outcomes denoted by X with no further structure. Imagine a black-box system that is able, given a certain number of outcomes $x_1, \dots, x_n$, to choose one outcome between them. As a user of the system, there is no way to guess in advance what will be the chosen outcome. A convenient mathematical way to store choice options is to consider them as a subset $U \subseteq X$. This defines how the effect of *non-deterministic choice* structures sets: it transforms any set X into its *powerset* $\{U \subseteq X\}$. When the user feeds a subset U to the system, they do not know the outcome, but they *do know* that the outcome belongs to U. Now, apply a function $f : X \to Y$ to the outcome. The user still does not know what element $y \in Y$ is obtained, but they *do know* that it belongs to the direct image $f(U) = \{f(x) \mid x \in U\}$. Therefore, non-deterministic choice acts on functions $f : X \to Y$ by transforming them into their direct image $f : \{U \subseteq X\} \to \{V \subseteq Y\}$. These two transformations define the effect itself. Now, imagine no choice is given i.e. we consider a single element $x \in X$. A simple procedure to create a choice is to *structure* x into the singleton subset $\{x\}$: the system is given a choice between all elements of the singleton, consequently it is forced to choose x. Last, imagine there is a *choice between choices*: the system must choose an outcome amongst a certain number of subsets $U_1, \dots, U_n$ modelling choices, to then choose an outcome inside the chosen subset. Equivalently, it can perform a single choice in the subset $\bigcup_{1 \leq i \leq n} U_i$. Therefore, collapsing two non-deterministic choices into one is embodied by the *union* operation. Categorically, the operations described in this paragraph correspond to a functor P on Set mapping a set to its powerset and a function to its direct image, along with two natural transformations $\eta^{\mathbf{P}} : 1 \to P$, $\mu^{\mathbf{P}} : PP \to P$ satisfying $\eta^{\mathbf{P}}_X(x) = \{x\}$ and $\mu^{\mathbf{P}}_X(\mathcal{U}) = \bigcup \mathcal{U}$.

Compositionality. The above sketch of the powerset monad $\mathbf{P}$ is just one out of the many monads of interest in computer science. The prominent other example employed in this work is the distribution monad $\mathbf{D}$ modelling probabilistic choice. There are also monads for more common computational effects such as exceptions, lists, trees, inputs, or outputs. Complex systems may exhibit a behaviour that is related to two or more monads. In this context, the compositionality of monads is crucial. There is also an algebraic interest in understanding monad compositionality, related to the algebraic approach of computational effects [75]. Although category theory is especially well-suited to provide compositionality results, there is no general way to compose two monads and get a third monad. Even in some apparently simple cases, there may be no composite monad. For example, there is no monad on the double powerset functor [91] i.e. performing two non-deterministic choices in a row does not yield a proper notion of effect. Consequently, the framework of monads slightly clashes with what one would expect from a theoretical tool for computer science, where compositionality is a key feature.

Distributive laws. A standard technique to generate a new monad out of two monads is to use Beck's theory of distributive laws [7]. Given two monads S and T, a distributive law is a natural transformation $\mathsf{TS} \to \mathsf{ST}$ formally allowing to swap the order in which these monads are applied. Provided four compatibility axioms are satisfied – one for each creation and collapse procedures for T and S – a composite monad ST can then be defined. A substantial advantage of distributive laws is their ubiquity: they are equivalently liftings of S to T-algebras and extensions of T to free S-algebras. Concretely, this correspondence makes distributive laws the edge of a three-sided coin. Depending on the question, it may be more relevant to rather look at heads or tails.

Coalgebras. One of the numerous applications of monads in computer science occurs in the theory of coalgebras. Since the seminal work of Rutten [131], the coalgebraic approach has become a well-established body of techniques for modelling various coinductive behaviours. Indeed, dually to initial algebras producing a notion of induction, final coalgebras come with a notion of coinduction. Consequently, coalgebras are particularly well-tailored to study state-based systems, bisimulations, or modal logic. Broadly speaking, a coalgebra can be thought of as the categorical version of an automaton. In this vision, a monad embodies the *branching type* of a system. For example, non-deterministic automata are the combination of two simple concepts: the one of an automaton (comprising e.g. labels in an alphabet, initial and final states) and non-deterministic choice. Formally, they are coalgebras for a composite functor involving the powerset monad P. Similarly, Markov chains, which are the simplest probabilistic state-based systems one can probably imagine, are just coalgebras for the distribution monad D. Amongst the celebrated successes of coalgebra theory lies the generalised powerset construction, which universalises the transformation of a non-deterministic automaton into a deterministic automaton – this will be given a detailed presentation in Chapter 4. For the construction to be sound, it needs to rely on a distributive law. The generalised powerset construction has applications in verifying program equivalence: in a system that has been determinised using a distributive law, bisimilarity can be checked efficiently by using so-called *up-to techniques* [21].

State of the Art

No-go theorems. The PhD thesis of Varacca [152] is the first to exhibit a concrete case where a distributive law cannot exist. Following an idea of Plotkin, Varacca proves that there is no possible distributive law of type $\mathsf{DP} \to \mathsf{PD}$, i.e., probability does not distribute over non-determinism. Using similar methods, Klin and Salamanca prove that there is no distributive law of type $\mathsf{PP} \to \mathsf{PP}$ [91], as a follow-up to the realisation that the literature was erroneously claiming the contrary. Identifying that the arguments underlying both proofs can be viewed as being algebraic rather than categorical, Zwart and Marsden [162] carry out a systematic study of what they called *no-go theorems for distributive laws*. Their generic results, focusing on the equational theories of finitary monads on Set, exhibit $\mathsf{DP} \to \mathsf{PD}$ and $\mathsf{PP} \to \mathsf{PP}$ as two of the many examples for which there cannot be any distributive law. The PhD thesis of Maaike Zwart [161] describes a number of no-go theorems thoroughly and gives many other examples where the existence of a distributive law fails. No-go theorems are essential in the sense that they put unbridgeable boundaries to where there can or cannot be a distributive law.

Composing monads all the same. The annoying but unavoidable consequence of no-go theorems is that systems combining two uncomposable monads are difficult to model. This is an issue e.g. in domain theory. There, the non-existence of a distributive law between the powerset monad and the distribution monad raises challenges for combining powerdomains [118] and probabilistic powerdomains [84, 86, 134] addressed by many authors [108, 58, 148, 88]. Coalgebraically speaking, the absence of laws $\mathsf{PP} \to \mathsf{PP}$ and $\mathsf{DP} \to \mathsf{PD}$ relates to a strewn-with-obstacles history in the modelling of alternating automata and probabilistic automata, respectively. Lacking a distributive law $\mathsf{TS} \to \mathsf{ST}$, one can still try and compose monads using dedicated methods, not relying on categorical distributive laws. The basic idea consists in defining directly a monad involving both features of T and S. The two main pitfalls are then finding the right definition for such a monad and arguing for its canonicity. For instance, Bonchi, Silva, and Sokolova [23] understand the determinisation of probabilistic automata from a coalgebraic perspective by devising a framework of so-called quasi-lax liftings. This way, they circumvent the non-existence of a proper distributive law. Similarly, Klin and Rot [90] successfully retrieve the semantics of alternating automata using the framework of so-called forgetful logics. In both cases, the link with distributive laws is lost in the process, whereas the whole constructions are desperately close to what one would get using such a law.

Algebraic approaches. A standard technique consists in using the fact that monads in Set correspond to equational theories. It is often the case that given equational theories for S and T, there is an obvious *algebraic* distributive law to impose between their operations. Adding this equation to the theory generates a new monad that may or may not intertwine smoothly the features of S and T. Afterwards, there still is the option to slightly modify one of the two monads to retrieve a construction that relies on a categorical distributive law. This general algebraic perspective exemplifies in Varacca's idea of replacing the distribution monad with a monad of so-called *indexed valuations* [151], and has also been used in [88, 25, 39]. A downside of all these algebraic methods, manipulating terms and variables explicitly, is that they are restricted to monads on the category Set, whereas distributive laws can operate in any category.

Weak distributive laws. Another path towards the composition of seemingly uncomposable monads emerges from work on so-called *weak distributive laws*, that is, distributive laws whose required axioms have been weakened. These have first been introduced independently by Street [146] and Böhm [16]. A decade later, Richard Garner publishes the paper *The Vietoris monad and weak distributive laws* [54], where he leverages a particular notion of weak distributive law to take a fresh look at the composition of some Set monads. He peripherally shows that there is a canonical weak distributive law $\mathsf{PP} \to \mathsf{PP}$.

Approach and Contributions

In this thesis, we use weak distributive laws to provide monad compositions that are meaningful to theoretical computer science. We focus on cases where distributive laws are impeached by no-go theorems, our two leading examples being $\mathbf{PP} \to \mathbf{PP}$ and $\mathbf{DP} \to \mathbf{PD}$. We show how known constructions, such as the convex powerset monad and generalised powerset constructions for probabilistic or alternating automata, arise as mere instances of the theory of weak distributive laws. Our approach being categorical rather than algebraic, we are able to generalise some examples to other categories than Set. A more detailed account of our contributions follows.

Further results on the general theory of weak distributive laws are obtained, including a new theorem to produce such laws.

- Any monad morphism yields a weak distributive law (Theorem 2.11) with a trivial composite monad (Proposition 2.12).

The way Garner weakens distributive laws $\mathbf{TS} \to \mathbf{ST}$ is by suppressing the axiom stating compatibility with the unit of $\mathbf{T}$. We describe what the theory becomes if one chooses instead to drop the axiom stating compatibility with the unit of $\mathbf{S}$.

- Coweak distributive laws constitute a dual framework (Section 2.3).

Once we are able to compose some otherwise uncomposable monads using (co)weak distributive laws, the question of composing *more* monads remains. In the case of distributive laws, this has been studied by Cheng under the name *iterated distributive laws*.

- Iterated distributive laws partially generalise to the (co)weak framework (Section 2.4).

Turning our attention to concrete examples of non-trivial weak distributive laws, we study the law $\mathbf{PP} \to \mathbf{PP}$ proposed by Garner in [54] and show that weak distributive laws are an appropriate framework to model the combination of non-deterministic choice and probabilistic choice.

- There is a unique monotone weak distributive law $\mathbf{DP} \to \mathbf{PD}$ (Theorem 3.12).

- The corresponding weak lifting is the convex powerset monad on convex algebras (equations (3.41) to (3.48)).

- The weak composite monad is the monad of convex subsets of distributions (Definition 3.15).

- The weak distributive law recovers an algebraic presentation for the above (Theorem 3.16).

Some variations are also introduced. Notably, it is possible to drop the finite support condition on the distribution monad (Theorem 3.23). The case $\mathsf{PD} \to \mathsf{DP}$ is also lightly discussed, with the following negative result.

- There is no (co)weak distributive law $\mathsf{PD} \to \mathsf{DP}$ related to the law $\mathsf{PP} \to \mathsf{PP}$ via the support monad morphism (Proposition 3.28).

Continuing with our general agenda of bringing weak distributive laws into computer science, we explain how classical results of coalgebra theory can be adapted to weak distributive laws.

- Generalised powerset construction is still available for weak distributive laws (Propositions 4.8, 4.9, 4.10).
- Up-to techniques stemming from generalised powerset construction remain compatible (Theorem 4.27).

These abstract results are instantiated on alternating automata and probabilistic automata, explaining the deep origin of determinisation procedures described in [81, 90, 23].

The semantics of programming languages is most of the time derived in categories with a richer structure than Set. For instance, in domain theory, these categories often have a topological flavour, while in probabilistic programming, the recently-built category of quasi-Borel spaces [74] is a fine choice for interpreting higher-order programs. Therefore it is beneficial that the categorical approach to weak distributive laws does not restrict the scope to the category of sets, contrary to most algebraic methods. With this in mind, we suggest as a motivating objective to generalise the monotone weak distributive law $\mathsf{DP} \to \mathsf{PD}$ to a category that can model *continuous* probability and *continuous* non-determinism. The category of compact Hausdorff spaces is selected for this purpose, as it is somewhat similar to Set and possesses monads interpreting both desired effects.

Our investigations reveal that regular categories are a convenient setting to abstractly state the key result generating monotone laws. In the particular case of toposes, and using results of de Moor [43], we are able to generalise the following Set facts.

- There is no distributive law $\mathsf{PP} \to \mathsf{PP}$ (Proposition 6.19).
- There is a unique monotone weak distributive law $\mathsf{PP} \to \mathsf{PP}$ (Theorem 6.21 and Proposition 6.22).

We provide formal Coq proofs of the new results, standing for constructive proofs performed in the internal logic of the topos. Finally, we study the non-topos of compact Hausdorff spaces, in which non-deterministic choice and probabilistic choice are interpreted by the Vietoris monad V and the Radon monad R, respectively. Drawing inspiration from work on closed relations and continuous relations [9, 10], we provide a result to generate monotone laws in this setting.

- For any monad T on compact Hausdorff spaces, assuming three conditions on its functor and one condition on its multiplication, there is a monotone weak distributive law $\mathsf{TV} \to \mathsf{VT}$ (Corollary 7.11).

After verifying that there cannot be any distributive law of type $\mathsf{VV} \to \mathsf{VV}$ and $\mathsf{RV} \to \mathsf{VR}$, we prove the following results.

- There is a monotone weak distributive law $\mathsf{VV} \to \mathsf{VV}$ (Theorem 7.19).
- At least two out of four conditions required for a monotone weak distributive law $\mathsf{RV} \to \mathsf{VR}$ hold (Propositions 7.26 and 7.28).

Our quest for composing continuous non-deterministic choice with continuous probabilistic choice halts with a conjecture that the two last conditions hold, under which we provide an expression of the resulting weak distributive law $\mathsf{RV} \to \mathsf{RV}$.

Related Work

Our work is overall closely related to the recent paper of Garner [54]. Following our LICS paper [62], interest was raised in the computer science community for weak distributive laws. Bonchi and Santamaria [22] devised a weak distributive law for the monad of semimodules over a semiring, which can be seen as a generalisation of our weak distributive law $\mathsf{DP} \to \mathsf{PD}$. Links between semialgebras and weak distributive laws have also been investigated by Petrişan and Sarkis in an upcoming paper [116].

The PhD thesis of Parlant [112] proposes an algebraic take to explain how the generalised powerset construction of alternating automata and probabilistic automata arise. This viewpoint provides interesting algebraic insights and is partly based on the quasi-lax liftings approach of [23], which turns out to be weak liftings in disguise.

One of our technical results, stating that the multiplication of the distribution monad is weakly cartesian (Proposition 3.10) was proved independently by Fritz and Perrone [52]. Their proof turns out to be simpler, and we reuse it to generalise the result to the countable distribution monad (Theorem 3.23).

Collaborations

In this short section, we would like to credit:

- Daniela Petrişan, for first identifying that something had to be done with weak distributive laws concerning the everlasting problem of combining probability and non-determinism;
- Daniela Petrişan again, for continuously delivering fruitful ideas and technical support that had an impact on most parts of the thesis;
- Marc Aiguier, for pointing out to us that there was a powerset monad in toposes, making it possible to discover the work of de Moor [43];
- Richard Garner, for hinting that the Kleisli category of the Vietoris monad could itself be of interest for weak distributive laws.

In the body of the thesis, all uncredited propositions and theorems are original works.

Outline

In Chapter 1, we recall basic notions of category theory that will be used in the whole presentation. Chapter 2 presents the framework of weak distributive laws in Garner's style. We provide a detailed presentation, with intuitive explanations of what is happening, sometimes using string diagrams. Trivial weak distributive laws, coweak distributive laws, and iterated weak distributive laws are introduced. Chapter 3 contains the two pivotal examples of the thesis: the monotone weak distributive laws $\mathsf{PP} \to \mathsf{PP}$ and $\mathsf{DP} \to \mathsf{PD}$ are derived and studied, along with some variations and a discussion about the case $\mathsf{PD} \to \mathsf{DP}$. Chapter 4 gives coalgebraic applications, extending both the generalised powerset construction and the corresponding up-to techniques to weak distributive laws. We provide a range of examples for probabilistic automata and alternating automata, explaining the origin of some results of Bonchi, Silva, Sokolova, Jacobs, Klin, and Rot. Chapters 3 and 4 together contain and extend some results that have been published in

- [62] Alexandre Goy, Daniela Petrişan. Combining probabilistic and non-deterministic choice via weak distributive laws. LICS 2020.

Chapter 5 is an interlude towards the second part of the thesis, consisting in searching weak distributive laws outside of Set. It provides standard material about regular categories. In Chapter 6, we study the case of toposes, in which results about the distributive law $\mathbf{PP} \to \mathbf{PP}$ are generalised. Finally, Chapter 7 introduces compact Hausdorff spaces, their closed and continuous relations, and the two monads of interest $\mathbf{V}$ (Vietoris) and $\mathbf{R}$ (Radon). We build the monotone weak distributive law $\mathbf{VV} \to \mathbf{VV}$ and provide some first steps toward the existence of a monotone weak distributive law $\mathbf{RV} \to \mathbf{VR}$. Chapters 5, 6, and 7 together contain and extend some results that have been published in

- [63] Alexandre Goy, Daniela Petrişan, Marc Aiguier. Powerset-like monads weakly distribute over themselves in toposes and compact Hausdorff spaces. ICALP 2021.

Chapter 1

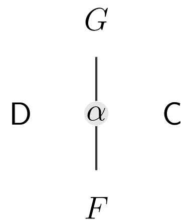
Preliminaries

The reader is supposed to be familiar with the following basic notions of category theory: category, monomorphism and epimorphism, functor, natural transformation, limit and colimit. In the first chapters we will mainly work in a generic category denoted by C and give many examples in the category Set of sets and functions. A table fixing notation is supplied at the end of the thesis.

1.1 String Diagrams

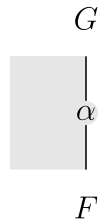
Lots of category theoretic proofs are usually performed by pasting commutative diagrams. This practice, known as *diagram chasing*, has some undeniable strengths. For instance, it is inherently compositional, it perfectly retains the type information of morphisms, and it provides a direct visual overview of the global proof. But diagram chasing also suffers from serious drawbacks. Large commutative diagrams can quickly become unreadable for two main reasons. The first one is that commutative diagrams involve a loot of *bookkeeping*, in the sense that functoriality and naturality are explicitly depicted. The second is that from a certain proof size, the human eye accommodates better a linear, equational-style reasoning, than a global and somewhat disorganised proof. This section provides a brief overview of the graphical formalism of string diagrams in category theory, which can be a relevant alternative to commutative diagrams for certain proofs. We only present a tiny fragment of this formalism, because we exclusively aim at using it for statements of the form $\alpha = \beta$, where α and β are natural transformations defined in terms of other natural transformations. We refer the reader to [103] for a much more complete introduction.

Orientation conventions vary in the literature. In this document, string diagrams are read from bottom to top. A natural transformation $\alpha : F \to G$ between two functors $F, G : \mathsf{C} \to \mathsf{D}$ is represented by a string



(1.1)

Instead of explicitly indicating the categories, we may use colouring. By convention, the portions of the space corresponding to the base category C are always uncoloured. The previous diagram rewrites as



(1.2)

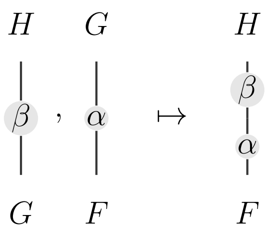
In the rest of this overview, all functors are endofunctors on C. The identity natural transformation $\mathrm{id}_F : F \to F$ is simply denoted by an unmarked string.

(1.3)

The identity functor 1 is not depicted. For example a natural transformation $\kappa : 1 \to K$ looks like

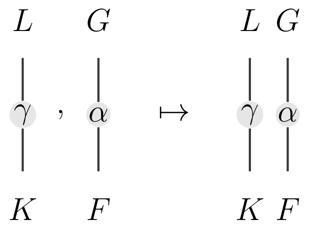
(1.4)

Vertical composition of $\alpha : F \to G$ and $\beta : G \to H$ is denoted by $\beta \circ \alpha : F \to H$. Recall that it is defined by $(\beta \circ \alpha)_X = \beta_X \circ \alpha_X$. Vertical composition is represented in string diagrams by vertical glueing



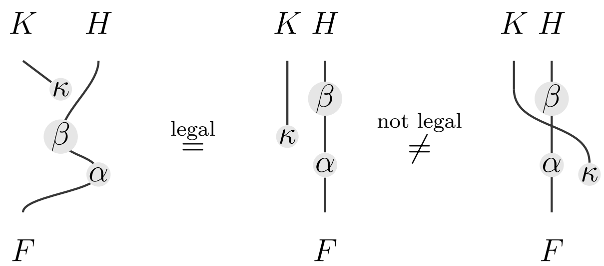
(1.5)

Horizontal composition of $\alpha : F \to G$ and $\gamma : K \to L$ is denoted by $\gamma\alpha : KF \to LG$. Recall that it is defined by $\gamma\alpha = \gamma G \circ K\alpha = L\alpha \circ \gamma F$ (both expressions are equal by naturality). Horizontal composition is represented in string diagrams by horizontal juxtaposition



(1.6)

Deformations of string diagrams preserve the meaning provided they are topologically mild. For example, strings can be expanded or shrinked, but not cross each other. Nodes representing natural transformations can be dragged along strings, but not slide past each other. Legal deformations encode what we called *bookkeeping*, e.g. naturality.



(1.7)

In the sequel, we will use some graphical notation to denote natural transformations without naming them explicitly on the string diagram. For readability, we may use different colours to mark different functors, though in a redundant way: in each case the reader should be able to infer what is represented using only the type information and the shape of the string diagram. In particular, (non-identity) functors will always be displayed on top and bottom of every diagram to help the reader typing expressions.

1.2 Monads

The concept of a monad is a typical example of a mathematical structure devised in category theory, brought into the scope of theoretical computer science, and now enjoying far-reaching ramifications throughout the field. Monads have various interpretations, but we shall follow those of Moggi, Plotkin and Power stating that a monad is a *notion of computation* [109] or a *computational effect* [119]. With respect to this interpretation, monads are concretely used in purely functional programming languages [157] to implement imperative effects such as exceptions, input, or output. The language Haskell, for instance, has a class `Monad` that can be instantiated to recover some monads presented in the sequel, such as the maybe monad, the list monad, and the reader monad. In this presentation, we shall stick to a purely categorical vision of monads. Informally, a monad consists of

- a functor T, modelling the structure of the effect
- a unit natural transformation $\eta^{\mathsf{T}} : 1 \to T$, implementing effect creation by adding one structured layer
- a multiplication natural transformation $\mu^{\mathsf{T}} : TT \to T$, implementing effect destruction by collapsing two structured layers into one single structure layer

satisfying some compatibility conditions, namely

- collapsing after creating is the identity transformation (both for outer layer creation and for inner layer creation)
- collapsing three layers into one layers can be done in any order (outer layers first or inner layers first)

1.2.1 Monads and Examples

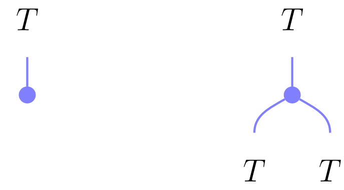
Definition 1.1 (Monad)**.** A *monad* T on a category C is a triple $(T, \eta^{\mathsf{T}}, \mu^{\mathsf{T}})$ such that $T : \mathsf{C} \to \mathsf{C}$ is a functor, $\eta^{\mathsf{T}} : 1 \to T$ is a natural transformation called the unit and $\mu^{\mathsf{T}} : TT \to T$ is a natural transformation called the multiplication. These data should satisfy the following equations:

$$\mu^{\mathsf{T}} \circ T\eta^{\mathsf{T}} = \mathrm{id}_T = \mu^{\mathsf{T}} \circ \eta^{\mathsf{T}} T \qquad \text{(unit axioms)}$$
$$\mu^{\mathsf{T}} \circ T\mu^{\mathsf{T}} = \mu^{\mathsf{T}} \circ \mu^{\mathsf{T}} T \qquad \text{(associativity axiom)}$$

Equivalently, using commutative diagrams:

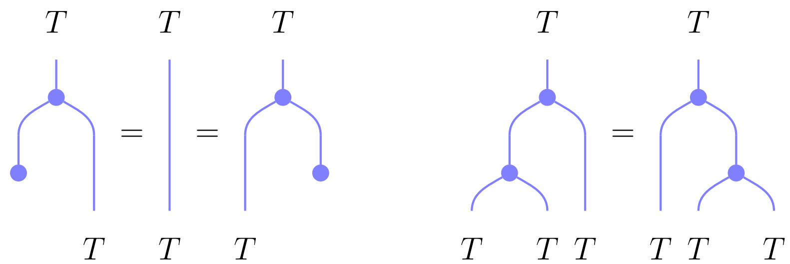
$$\begin{tikzcd} T \arrow[r, "T\eta^{\mathsf{T}}"] \arrow[d, "\eta^{\mathsf{T}} T"'] \arrow[dr, equal] & TT \arrow[d, "\mu^{\mathsf{T}}"] \\ TT \arrow[r, "\mu^{\mathsf{T}}"'] & T \end{tikzcd} \qquad \begin{tikzcd} TTT \arrow[r, "T\mu^{\mathsf{T}}"] \arrow[d, "\mu^{\mathsf{T}} T"'] & TT \arrow[d, "\mu^{\mathsf{T}}"] \\ TT \arrow[r, "\mu^{\mathsf{T}}"'] & T \end{tikzcd} \tag{1.8}$$

Using string diagrams, the monad unit and multiplication are respectively represented by



(1.9)

and the equations are given by the following transformations



(1.10)

Example 1.2 (Identity monad)**.** Let us begin with a trivial example. Any category C has an *identity monad* $\mathbf{\mathsf{Id}}$ whose functor is the identity functor, with unit and multiplication both being the identity natural transformation id.

Example 1.3 (Maybe monad)**.** The *maybe monad* $(- + \mathbf{1})$ on Set is defined by the following data

- $(X + 1)$ is the disjoint union of X and the singleton $1 = \{*\}$. Let the canonical injections be

$$\mathsf{inl} : X \to (X + 1)$$
$$\mathsf{inr} : 1 \to (X + 1)$$

- $(f + 1) : (X + 1) \to (Y + 1)$ is given by

$$(f + 1)(\mathsf{inl}(x)) = \mathsf{inl}(f(x))$$
$$(f + 1)(\mathsf{inr}(*)) = \mathsf{inr}(*)$$

- $\eta_X^{(+\mathbf{1})} : X \to (X + 1)$ is left injection $\eta_X^{(+\mathbf{1})}(x) = \mathsf{inl}(x)$

- $\mu_X^{(+\mathbf{1})} : ((X + 1) + 1) \to (X + 1)$ merges the two copies of $\{*\}$

$$\mu_X^{(+\mathbf{1})}(\mathsf{inl}(\mathsf{inl}(x))) = \mathsf{inl}(x)$$
$$\mu_X^{(+\mathbf{1})}(\mathsf{inl}(\mathsf{inr}(*))) = \mathsf{inr}(*)$$
$$\mu_X^{(+\mathbf{1})}(\mathsf{inr}(*)) = \mathsf{inr}(*)$$

Example 1.4 (Powerset monad)**.** The *powerset monad* $\mathbf{P}$ on Set is defined by

- PX is the set of all subsets of X
- $Pf : PX \to PY$ is the direct image function $Pf(U) = \{f(x) \mid x \in U\}$
- $\eta^{\mathbf{P}}_X : X \to PX$ is the singleton function $\eta^{\mathbf{P}}_X(x) = \{x\}$
- $\mu^{\mathbf{P}}_X : PPX \to PX$ is the union function $\mu^{\mathbf{P}}_X(\mathcal{U}) = \bigcup \mathcal{U}$

Example 1.5 (Powerset monad variations)**.** By restricting the powerset functor but still using direct images, singleton, and union, we can obtain many variations of the powerset monad:

- $P_f X = \{U \in PX \mid U \text{ finite}\}$ yields the *finite powerset monad* $\mathbf{P}_f$
- $P^* X = \{U \in PX \mid U \neq \emptyset\}$ yields the *non-empty powerset monad* $\mathbf{P}^*$
- $P^*_f X = P_f X \cap P^* X$ yields the *finite non-empty powerset monad* $\mathbf{P}^*_f$

Example 1.6 (Distribution monad)**.** The *distribution monad* $\mathbf{D}$ on Set is defined by

- $DX = \{\varphi : X \to [0,1] \mid \sum_{x \in X} \varphi(x) = 1 \text{ and } \mathsf{supp}(\varphi) \text{ is finite}\}$ where the notation supp stands for the *support* defined by $\mathsf{supp}(\varphi) = \{x \in X \mid \varphi(x) \neq 0\}$. Elements of DX are called (finitely supported probability) distributions and can be formally denoted by $\varphi = \sum_{x \in X} \varphi_x \cdot x$, where $\varphi_x = \varphi(x)$
- $Df : DX \to DY$ computes pushforward distributions $Df(\varphi) = \sum_{x \in X} \varphi_x \cdot f(x)$ i.e. for any $y \in Y$, $Df(\varphi)(y) = \sum_{x \in f^{-1}(\{y\})} \varphi(x)$
- $\eta^{\mathbf{D}}_X : X \to DX$ computes the Dirac distribution $\eta^{\mathbf{D}}_X(x) = 1 \cdot x$
- $\mu^{\mathbf{D}}_X : DDX \to DX$ takes a weighted average $\mu^{\mathbf{D}}_X(\Phi) = \sum_{x \in X} \left(\sum_{\varphi \in DX} \Phi_\varphi \varphi_x\right) \cdot x$

Example 1.7 (Countable distribution monad)**.** By extending $\mathbf{D}$ to probability distributions with arbitrary support, one gets the *countable distribution monad* $\mathbf{D}_\omega$. This means that $D_\omega X = \{\varphi : X \to [0,1] \mid \sum_{x \in X} \varphi(x) = 1\}$ and the other data are defined as for $\mathbf{D}$. The terminology *countable* comes from the fact that for any such distribution φ, $\mathsf{supp}(\varphi)$ is at most countable [140, Proposition 2.1.2].

Example 1.8 (Abelian group monad)**.** The *Abelian group monad* $\mathbf{A}$ on Set is defined similarly to the distribution monad, as follows:

- AX is the free Abelian group on X i.e. the set of all functions $\pi : X \to \mathbb{Z}$ with finite support, written formally as $\pi = \sum_{x \in X} \pi_x \cdot x$
- $Af : AX \to AY$ is given by $Af(\pi) = \sum_{x \in X} \pi_x \cdot f(x)$
- $\eta^{\mathsf{A}}_X : X \to AX$ is given by $\eta^{\mathsf{A}}_X(x) = 1 \cdot x$
- $\mu^{\mathsf{A}}_X : AAX \to AX$ is given by $\mu^{\mathsf{A}}_X(\Pi) = \sum_{x \in X} \left(\sum_{\pi \in AX} \Pi_\pi \times \pi_x\right) \cdot x$

Example 1.9 (List monad)**.** The *list monad* L on Set is defined by

- LX is the set of lists of elements of X
- $Lf : LX \to LY$ is the map operation $Lf([x_1, \dots, x_n]) = [f(x_1), \dots, f(x_n)]$
- $\eta^{\mathsf{L}}_X : X \to LX$ creates a one-element list $\eta^{\mathsf{L}}_X(x) = [x]$
- $\mu^{\mathsf{L}}_X : LLX \to LX$ concatenates $\mu^{\mathsf{L}}_X([L_1, \dots, L_n]) = L_1 + \cdots + L_n$

Example 1.10 (Multiset monad)**.** The *multiset monad* M on Set is defined by

- MX is the set of lists of elements of X up to ordering (equivalently, functions $X \to \mathbb{N}$ with finite support)
- $Mf : MX \to MY$ is the map operation $Mf(\llbracket x_1, \dots, x_n \rrbracket) = \llbracket f(x_1), \dots, f(x_n) \rrbracket$
- $\eta^{\mathsf{M}}_X : X \to MX$ creates a one-element multiset $\eta^{\mathsf{M}}_X(x) = \llbracket x \rrbracket$
- $\mu^{\mathsf{M}}_X : MMX \to MX$ concatenates $\mu^{\mathsf{M}}_X(\llbracket M_1, \dots, M_n \rrbracket) = M_1 + \cdots + M_n$

Example 1.11 (Reader monad)**.** Let A be a fixed set. The *reader monad* R on Set is defined by

- $RX = X^A$, the set of functions of type $A \to X$
- $Rf : X^A \to Y^A$ is just composition $Rf(h) = f \circ h$
- $\eta^{\mathsf{R}}_X : X \to X^A$ outputs the constant function $\eta^{\mathsf{R}}_X(x) = \lambda a.x$
- $\mu^{\mathsf{R}}_X : (X^A)^A \to X^A$ computes the diagonal function $\mu^{\mathsf{R}}_X(H) = \lambda a.H(a)(a)$

Example 1.12 (Filter monad [42], ultrafilter monad [98])**.** For any $\mathcal{U} \in PPX$ consider the properties

(i) $\mathcal{U} \neq \emptyset$

(ii) for all $U, V \in PX, (U \in \mathcal{U} \wedge V \in \mathcal{U}) \iff U \cap V \in \mathcal{U}$

(iii) for all $U \in PX$, either $U \in \mathcal{U}$ or $U^c \in \mathcal{U}$ (exclusive)

A filter is a $\mathcal{U}$ satisfying (i), (ii). An ultrafilter additionally satisfies (iii). For any $x \in X$, let $\mathcal{U}_x = \{U \in PX \mid x \in U\}$ be the principal ultrafilter generated by x. The *filter monad* $\mathbf{F}$ is defined by

- $FX = \{\mathcal{U} \in PPX \mid \mathcal{U} \text{ filter}\}$
- $Ff : FX \to FY$ is given by $Ff(\mathcal{U}) = \{V \in PY \mid f^{-1}(V) \in \mathcal{U}\}$
- $\eta^{\mathbf{F}}_X : X \to FX$ computes the principal ultrafilter $\eta^{\mathbf{F}}_X(x) = \mathcal{U}_x$
- $\mu^{\mathbf{F}}_X : FFX \to FX$ is given by $\mu^{\mathbf{F}}_X(\mathfrak{U}) = \bigcup\{\bigcap \mathfrak{u} \mid \mathfrak{u} \in \mathfrak{U}\}$

The *ultrafilter monad* $\boldsymbol{\beta}$ is defined by $\beta X = \{\mathcal{U} \in PPX \mid \mathcal{U} \text{ ultrafilter}\}$, the rest being as for $\mathbf{F}$.

Let us end this list of examples with a monad similar to the powerset monad $\mathbf{P}$ but not living in Set.

Example 1.13 (Vietoris monad [155])**.** Let KHaus be the category of compact Hausdorff spaces and continuous functions. For any compact Hausdorff space X with topology τ_X, let VX be the set of all closed subsets of X with the topology generated by all sets of the form $\{C \in VX \mid C \subseteq U\}$ and $\{C \in VX \mid C \cap U \neq \emptyset\}$, where U ranges over τ_X. This extends to a monad called the *Vietoris monad*, defined by

- $Vf : VX \to VY$ is the direct image function $Vf(C) = \{f(x) \mid x \in C\}$
- $\eta^{\mathbf{V}}_X : X \to VX$ is the singleton function $\eta^{\mathbf{V}}_X(x) = \{x\}$
- $\mu^{\mathbf{V}}_X : VVX \to VX$ is the union function $\mu^{\mathbf{V}}_X(\mathcal{C}) = \bigcup \mathcal{C}$

The reader is invited to refer to Chapter 7 for more details about the Vietoris monad.

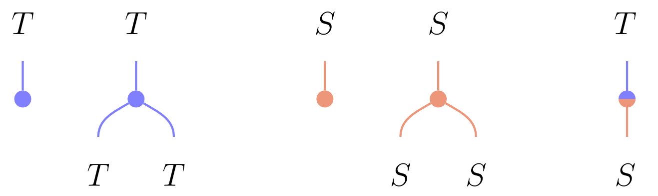
The similarity between two monads can be grasped by monad morphisms.

Definition 1.14 (Monad morphism)**.** Let $\mathbf{T}$, $\mathbf{S}$ be monads on a category C. A *monad morphism* from $\mathbf{S}$ to $\mathbf{T}$ is a natural transformation $\gamma : S \to T$ such that

$$\gamma \circ \eta^{\mathbf{S}} = \eta^{\mathbf{T}} \tag{1.11}$$

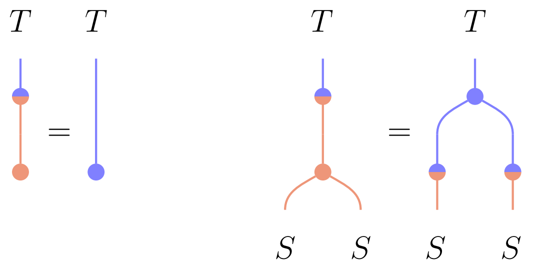
$$\gamma \circ \mu^{\mathbf{S}} = \mu^{\mathbf{T}} \circ \gamma\gamma \tag{1.12}$$

String diagrammatically, these data are



(1.13)

such that



(1.14)

Example 1.15. For any monad $\mathbf{T}$ on a category C, the identity natural transformation $\mathrm{id}_T : T \to T$ induces a monad morphism $\mathbf{T} \to \mathbf{T}$.

Example 1.16. The support operation induces a monad morphism $\mathsf{supp} : \mathbf{D} \to \mathbf{P}$. The axioms respectively state that the support of a Dirac distribution is the corresponding singleton, and that the support of the weighted average of a distribution $\Phi \in DDX$ is the union of the supports of every $\varphi \in \mathsf{supp}(\Phi)$. The support also defines a monad morphism from $\mathbf{D}$ into $\mathbf{P}_f$, $\mathbf{P}^*$, and $\mathbf{P}_f^*$. The support as a monad morphism has been used for example in [81].

1.2.2 Monads Arising from Adjunctions

There are several equivalent definitions of adjunctions [96], but we shall focus on the one that relates easily to monads.

Definition 1.17 (Adjunction)**.** Let C and D be categories. An *adjunction* $(L, R, \eta, \epsilon) : \mathsf{C} \to \mathsf{D}$ consists of

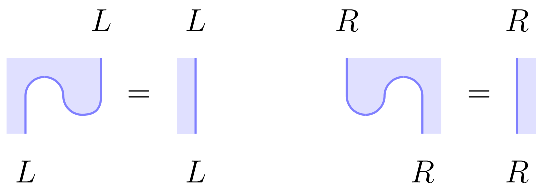
- A functor $L : \mathsf{C} \to \mathsf{D}$, the *left adjoint*
- A functor $R : \mathsf{D} \to \mathsf{C}$, the *right adjoint*
- A natural transformation $\eta : 1 \to RL$, the *unit*
- A natural transformation $\epsilon : LR \to 1$, the *counit*

such that $\epsilon L \circ L\eta = 1$ and $R\epsilon \circ \eta R = 1$.

In string diagrams, the unit and the counit are respectively

$$(1.15)$$

such that



$$(1.16)$$

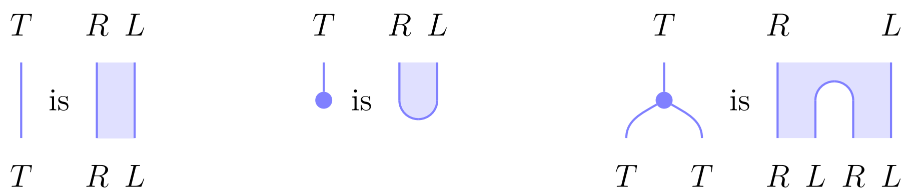
The notation $L \dashv R$ denotes that L is left adjoint to R, with unit and counit left implicit. We will also use the notation

$$\begin{tikzcd} \mathsf{D} \arrow[r, shift right=1.5ex, "R"'] \arrow[r, phantom, "\perp"] & \mathsf{C} \arrow[l, shift right=1.5ex, "L"'] \end{tikzcd} \tag{1.17}$$

We are mainly interested in the link between adjunctions and monads. A first important fact is that contrary to monads, adjunctions compose nicely. Indeed, if both $(L_1, R_1, \eta_1, \epsilon_1) : \mathsf{C} \to \mathsf{D}$ and $(L_2, R_2, \eta_2, \epsilon_2) : \mathsf{D} \to \mathsf{E}$ are adjunctions, then $(L_2L_1, R_1R_2, R_1\eta_2L_1 \circ \eta_1, \epsilon_2 \circ L_2\epsilon_1R_2) : \mathsf{C} \to \mathsf{E}$ is an adjunction. Diagrammatically:

$$\begin{tikzcd} \mathsf{E} \arrow[r, shift right=1.5ex, "R_2"'] \arrow[r, phantom, "\perp"] & \mathsf{D} \arrow[l, shift right=1.5ex, "L_2"'] \arrow[r, shift right=1.5ex, "R_1"'] \arrow[r, phantom, "\perp"] & \mathsf{C} \arrow[l, shift right=1.5ex, "L_1"'] \end{tikzcd} \quad = \quad \begin{tikzcd} \mathsf{E} \arrow[r, shift right=1.5ex, "R_1R_2"'] \arrow[r, phantom, "\perp"] & \mathsf{C} \arrow[l, shift right=1.5ex, "L_2L_1"'] \end{tikzcd} \tag{1.18}$$

Any adjunction (L, R, η, ϵ) yields a monad $(RL, \eta, R\epsilon L)$, with the unit of the monad being the unit of the adjunction:



$$(1.19)$$

Conversely, any monad $\mathbf{T}$ can be obtained in this way from an adjunction. There may actually be many adjunctions generating $\mathbf{T}$. The two extremal examples are the Kleisli adjunction and the Eilenberg-Moore adjunction.

Definition 1.18 (Kleisli category)**.** The *Kleisli category* of $\mathbf{T}$ is the category $\mathsf{Kl}(\mathbf{T})$ with the same objects as C, and where a morphism $f : X \nrightarrow Y$ is a morphism $f : X \to TY$ in C. The identity morphism $X \nrightarrow X$ is $\eta^{\mathbf{T}}_X$. The composition of $f : X \nrightarrow Y$ and $g : Y \nrightarrow Z$ is $g \bullet f = \mu^{\mathbf{T}}_Z \circ Tg \circ f$.

Note that the notation $\nrightarrow$ distinguishes morphisms in $\mathsf{Kl}(\mathbf{T})$, and the notation $\bullet$ distinguishes composition in $\mathsf{Kl}(\mathbf{T})$.

By defining

- $F_{\mathbf{T}} : \mathsf{C} \to \mathsf{Kl}(\mathbf{T})$ by $F_{\mathbf{T}}X = X$ and $F_{\mathbf{T}}f = \eta^{\mathbf{T}}_Y \circ f$ for any $f : X \to Y$
- $U_{\mathbf{T}} : \mathsf{Kl}(\mathbf{T}) \to \mathsf{C}$ by $U_{\mathbf{T}}X = TX$ and $U_{\mathbf{T}}f = \mu^{\mathbf{T}}_Y \circ Tf$ for any $f : X \nrightarrow Y$
- $\eta : 1 \to U_{\mathbf{T}}F_{\mathbf{T}}$ by $\eta_X = \eta^{\mathbf{T}}_X$
- $\epsilon : F_{\mathbf{T}}U_{\mathbf{T}} \nrightarrow 1$ by $\epsilon_X = \mathrm{id}_{TX}$

we get an adjunction $F_{\mathbf{T}} \dashv U_{\mathbf{T}}$ generating $\mathbf{T}$, which is initial amongst these adjunctions and is called the Kleisli adjunction.

A $\mathbf{T}$-algebra is a pair (X, x) where X is an object of C and $x : TX \to X$ is a morphism such that

$$x \circ \eta^{\mathbf{T}}_X = \mathrm{id}_X \qquad \text{(unit axiom)}$$
$$x \circ \mu^{\mathbf{T}}_X = x \circ Tx \qquad \text{(associativity axiom)}$$

In commutative diagrams, this is

$$\begin{tikzcd} X \arrow[r, "\eta^{\mathbf{T}}_X"] \arrow[dr, equal] & TX \arrow[d, "x"] \\ & X \end{tikzcd} \qquad \begin{tikzcd} TTX \arrow[r, "\mu^{\mathbf{T}}_X"] \arrow[d, "Tx"'] & TX \arrow[d, "x"] \\ TX \arrow[r, "x"] & X \end{tikzcd} \tag{1.20}$$

The pair $(TX, \mu^{\mathbf{T}}_X)$ is a $\mathbf{T}$-algebra called the *free* $\mathbf{T}$-algebra on X. Its unit axiom is just one of the unit axioms of $\mathbf{T}$ and its associativity axiom is just the associativity axiom of $\mathbf{T}$.

A morphism of $\mathbf{T}$-algebras $f : (X, x) \to (Y, y)$ is a morphism $f : X \to Y$ in C such that $y \circ Tf = f \circ x$.

$$\begin{tikzcd} TX \arrow[r, "Tf"] \arrow[d, "x"'] & TY \arrow[d, "y"] \\ X \arrow[r, "f"] & Y \end{tikzcd} \tag{1.21}$$

Definition 1.19 (Eilenberg-Moore category)**.** The *Eilenberg-Moore category* of $\mathbf{T}$, denoted by $\mathsf{EM}(\mathbf{T})$, has $\mathbf{T}$-algebras as objects and morphisms of $\mathbf{T}$-algebras as morphisms. Identities and composition are as in C.

By defining

- $F^{\mathbf{T}} : \mathsf{C} \to \mathsf{EM}(\mathbf{T})$ by $F^{\mathbf{T}}X = (TX, \mu_X^{\mathbf{T}})$ and $F^{\mathbf{T}}f = Tf$
- $U^{\mathbf{T}} : \mathsf{EM}(\mathbf{T}) \to \mathsf{C}$ by $U^{\mathbf{T}}(X, x) = X$ and $U^{\mathbf{T}}f = f$
- $\eta : 1 \to U^{\mathbf{T}}F^{\mathbf{T}}$ by $\eta_X = \eta_X^{\mathbf{T}}$
- $\epsilon : F^{\mathbf{T}}U^{\mathbf{T}} \to 1$ by $\epsilon_{(X,x)} = x$, extracting the algebraic structure from the algebra

we get an adjunction $F^{\mathbf{T}} \dashv U^{\mathbf{T}}$ generating $\mathbf{T}$, which is terminal amongst these adjunctions and is called the Eilenberg-Moore adjunction.

Finally, note that $\mathsf{Kl}(\mathbf{T})$ is equivalent to the full subcategory of $\mathsf{EM}(\mathbf{T})$ whose objects are the free algebras $(TX, \mu_X^{\mathbf{T}})$. We now give examples of Kleisli categories and Eilenberg-Moore categories for some specific monads.

Example 1.20 ($\mathsf{Kl}(- + \mathbf{1})$)**.** The Kleisli category of the maybe monad $(- + \mathbf{1})$ is the category of sets and partial functions. Indeed, a morphism $f : X \to (Y + 1)$ can be seen as a partial function $X \rightharpoonup Y$ undefined on the set $f^{-1}(\{\mathsf{inr}(*)\})$.

Example 1.21 ($\mathsf{EM}(- + \mathbf{1})$)**.** The Eilenberg-Moore category of the maybe monad $(- + \mathbf{1})$ is the category of pointed sets and functions preserving the base point. This category is actually equivalent to the category of sets and partial functions [136]. The maybe monad therefore satisfies $\mathsf{Kl}(- + \mathbf{1}) \equiv \mathsf{EM}(- + \mathbf{1})$.

Example 1.22 ($\mathsf{Kl}(\mathbf{P})$)**.** The Kleisli category of the powerset monad $\mathbf{P}$ is isomorphic to the category of sets and relations Rel. Objects of Rel are sets, and morphisms $X \nrightarrow Y$ are relations $R \subseteq X \times Y$. Identities are diagonal relations $\{(x, x) \mid x \in X\}$ and composition of two relations $R \subseteq X \times Y$ and $S \subseteq Y \times Z$ is defined by $S \circ R = \{(x, z) \in X \times Z \mid \exists y \in Y.(x, y) \in R \wedge (y, z) \in S\}$. The isomorphism $\mathsf{Kl}(\mathbf{P}) \cong \mathsf{Rel}$ acts identically on objects, sends a function $f : X \to PY$ to $R_f = \{(x, y) \mid x \in X, y \in f(x)\}$ and sends a relation $R \subseteq X \times Y$ to $f_R = \lambda x.\{y \in Y \mid (x, y) \in R\}$. Note that the submonads of the powerset monad described in Example 1.5 yield Kleisli categories that are wide subcategories of Rel. More precisely, a relation $R \subseteq X \times Y$ is a morphism

- in $\mathsf{Kl}(\mathbf{P}_f)$ when it is image-finite: $\forall x \in X.\{y \in Y \mid (x, y) \in R\}$ is finite;
- in $\mathsf{Kl}(\mathbf{P}^*)$ when it is total: $\forall x \in X.\exists y \in Y.(x, y) \in R$;
- in $\mathsf{Kl}(\mathbf{P}_f^*)$ when it is image-finite and total.

Example 1.23 ($\mathsf{EM}(\mathbf{P})$)**.** The Eilenberg-Moore category of the powerset monad is the category cJSL of complete join-semilattices and (arbitrary-)join-preserving homomorphisms. Variations obtained by restricting $\mathbf{P}$ yield variations obtained by relaxing the conditions on joins:

- $\mathsf{EM}(\mathbf{P}_f)$ objects have all finite joins and morphisms preserve them
- $\mathsf{EM}(\mathbf{P}^*)$ objects have all non-empty joins and morphisms preserve them
- $\mathsf{EM}(\mathbf{P}_f^*)$ objects have all non-empty finite joins and morphisms preserve them

Example 1.24 ($\mathsf{Kl}(\mathbf{D})$)**.** The Kleisli category of the distribution monad is the category of sets and stochastic relations, where a stochastic relation is a function $R : X \times Y \to [0, 1]$ with $\{y \in Y \mid R(x, y) \neq 0\}$ finite and $\sum_{y \in Y} R(x, y) = 1$ for all $x \in X$. The finite-support condition is dropped for the Kleisli category of the countable distribution monad.

Example 1.25 ($\mathsf{EM}(\mathbf{D})$)**.** The Eilenberg-Moore category of the distribution monad $\mathbf{D}$ can be interpreted as a category of *convex algebras* and convex (or affine) maps [23, 88, 142, 147]. The unit axiom of a $\mathbf{D}$-algebra says that the barycenter of a single element is itself, and the associativity axiom stands for barycenter associativity. For the countable distribution monad, convexity has to hold even with respect to countable convex combinations.

Example 1.26 ($\mathsf{EM}(\mathbf{L})$)**.** The Eilenberg-Moore category of the list monad $\mathbf{L}$ is the category Mon of monoids and monoid homomorphisms.

Example 1.27 ($\mathsf{EM}(\mathbf{M})$)**.** Similarly, the Eilenberg-Moore category of the multiset monad $\mathbf{M}$ is the category CMon of *commutative* monoids and monoid homomorphisms.

Example 1.28 ($\mathsf{EM}(\boldsymbol{\beta})$)**.** The Eilenberg-Moore category of the ultrafilter monad $\boldsymbol{\beta}$ is the category KHaus of compact Hausdorff spaces and continuous functions [98, 159]. The Eilenberg-Moore category of the filter monad $\mathbf{F}$ is the category of continuous lattices and functions preserving infima and directed suprema, see [42].

Example 1.29 ($\mathsf{EM}(\mathbf{V})$)**.** One step further, the Eilenberg-Moore category of the Vietoris monad $\mathbf{V}$ on KHaus has been identified by Wyler [159] as the same category of continuous lattices as in the previous example, i.e. $\mathsf{EM}(\mathbf{V}) \cong \mathsf{EM}(\mathbf{F})$.

1.3 Distributive Laws

Let $\mathbf{T} = (T, \eta^{\mathbf{T}}, \mu^{\mathbf{T}})$ and $\mathbf{S} = (S, \eta^{\mathbf{S}}, \mu^{\mathbf{S}})$ be two monads on C. Without additional information, there is no generic way of defining a monad on the composite functor ST. In fact, there are some cases when there is no possible monad structure on the composite functor. Two famous examples are PP [91] and PD [38]. Let us nevertheless try to define a monad on ST using the data of $\mathbf{S}$ and $\mathbf{T}$. There is an obvious candidate for the unit, namely the horizontal composition $\eta^{\mathbf{S}}\eta^{\mathbf{T}} : 1 \to ST$. The choice is less clear for what concerns the multiplication. What is needed is a natural transformation of type $STST \to ST$. Actually, things would be much simpler if we were looking for a natural transformation of type $SSTT \to ST$, because in this situation we would just as well use horizontal composition $\mu^{\mathbf{S}}\mu^{\mathbf{T}} : SSTT \to ST$.

In his seminal paper *Distributive laws* [7], Jon Beck introduced the missing ingredient: a natural transformation $\lambda : TS \to ST$. In the presence of such a natural transformation, a candidate multiplication for ST can be defined as

$$STST \xrightarrow{S\lambda T} SSTT \xrightarrow{\mu^{\mathbf{S}}\mu^{\mathbf{T}}} ST \tag{1.22}$$

Of course, in order for $(ST, \eta^{\mathbf{S}}\eta^{\mathbf{T}}, \mu^{\mathbf{S}}\mu^{\mathbf{T}} \circ S\lambda T)$ to satisfy the monad axioms, the natural transformation λ has to satisfy itself some axioms. When it does, it is called a *distributive law*. In this section, we recall the basics of Beck's theory of distributive laws, and give some examples.

Definition 1.30 (Distributive law). A *distributive law* of type $\mathbf{TS} \to \mathbf{ST}$ is a natural transformation λ of the given type such that the following equations hold.

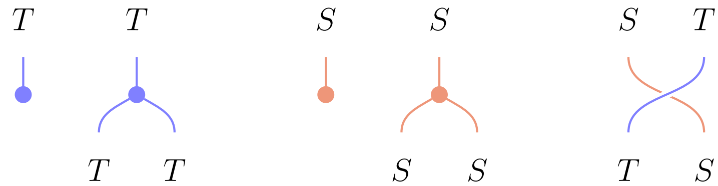
$$\begin{aligned} \lambda \circ T\eta^{\mathbf{S}} &= \eta^{\mathbf{S}}T && (\eta^{\mathbf{S}} \text{ axiom}) \\ \lambda \circ \eta^{\mathbf{T}}S &= S\eta^{\mathbf{T}} && (\eta^{\mathbf{T}} \text{ axiom}) \\ \lambda \circ T\mu^{\mathbf{S}} &= \mu^{\mathbf{S}}T \circ S\lambda \circ \lambda S && (\mu^{\mathbf{S}} \text{ axiom}) \\ \lambda \circ \mu^{\mathbf{T}}S &= S\mu^{\mathbf{T}} \circ \lambda T \circ T\lambda && (\mu^{\mathbf{T}} \text{ axiom}) \end{aligned}$$

In terms of commutative diagrams

$$\begin{tikzcd} & T \arrow[dl, "T\eta^{\mathbf{S}}"'] \arrow[dr, "\eta^{\mathbf{S}}T"] \arrow[d, phantom, "(\eta^{\mathbf{S}})"] & \\ TS \arrow[rr, "\lambda"'] & {} & ST \end{tikzcd} \qquad \begin{tikzcd} & S \arrow[dl, "\eta^{\mathbf{T}}S"'] \arrow[dr, "S\eta^{\mathbf{T}}"] \arrow[d, phantom, "(\eta^{\mathbf{T}})"] & \\ TS \arrow[rr, "\lambda"'] & {} & ST \end{tikzcd}$$

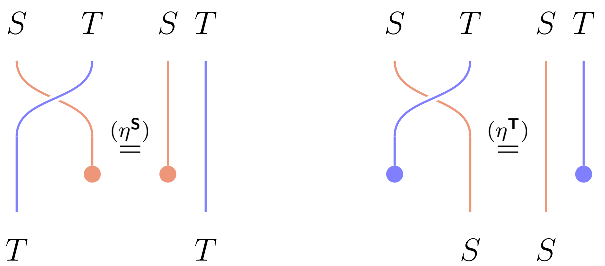
$$\begin{tikzcd} TSS \arrow[r, "\lambda S"] \arrow[d, "T\mu^{\mathbf{S}}"'] & STS \arrow[r, "S\lambda"] \arrow[d, phantom, "(\mu^{\mathbf{S}})"] & SST \arrow[d, "\mu^{\mathbf{S}}T"] \\ TS \arrow[rr, "\lambda"'] & {} & ST \end{tikzcd} \qquad \begin{tikzcd} TTS \arrow[r, "T\lambda"] \arrow[d, "\mu^{\mathbf{T}}S"'] & TST \arrow[r, "\lambda T"] \arrow[d, phantom, "(\mu^{\mathbf{T}})"] & STT \arrow[d, "S\mu^{\mathbf{T}}"] \\ TS \arrow[rr, "\lambda"'] & {} & ST \end{tikzcd}$$

In terms of string diagrams, the data $\eta^{\mathbf{T}}$, $\mu^{\mathbf{T}}$, $\eta^{\mathbf{S}}$, $\mu^{\mathbf{S}}$ and λ are respectively



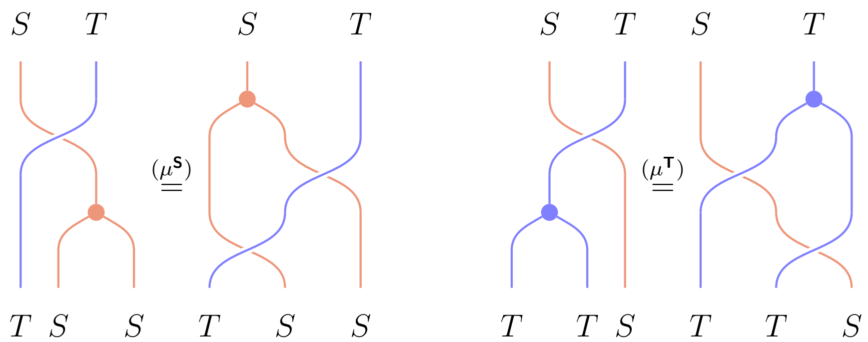
(1.23)

such that



(1.24)

and



(1.25)

Let us immediately give the prototypical example of a distributive law, introduced by Beck in his seminal paper.

Example 1.31 ([7])**.** The following expression defines a distributive law of type **LA** $\to$ **AL** between the list monad and the Abelian group monad.

$$\lambda_X : LAX \to ALX$$
$$[\pi^{(1)}, \ldots, \pi^{(n)}] \mapsto \sum_{(x_1,\ldots,x_n)\in X^n} \left(\prod_{i=1}^{n} \pi^{(i)}_{x_i}\right) [x_1, \ldots, x_n]$$

Interpreting elements in a list as factors in a product, take $n = 2$ and

$$\pi^{(1)} = 1 \cdot a + 1 \cdot b$$
$$\pi^{(2)} = 1 \cdot c + 1 \cdot d$$

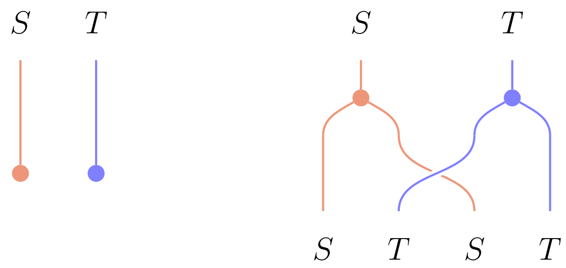
for some $(a, b, c, d) \in X$. We get

$$\lambda_X((a + b)(c + d)) = ac + bc + ad + bd \tag{1.26}$$

so λ embodies the property that *multiplication distributes over addition.*

Remark 1.32. Following the above example, the assertion *there is a distributive law of type* $\mathbf{TS} \to \mathbf{ST}$ will sometimes be shortened as $\mathbf{T}$ *distributes over* $\mathbf{S}$. Conventions around this formulation vary in the literature: some authors define $\mathbf{T}$ *distributes over* $\mathbf{S}$ as witnessing a law of type $\mathbf{ST} \to \mathbf{TS}$. Following discussions with Bartek Klin, a strong advocate of the first formulation, we argue that additionally to being sound with respect to the multiplication / addition example, our convention is easy to remember. We suggest the reader to use the following mnemonic: in the syntax $\mathbf{TS} \to \mathbf{ST}$, the monad $\mathbf{T}$ *distributes over* the monad $\mathbf{S}$ because the symbol $\mathbf{T}$ *jumps over* the symbol $\mathbf{S}$.

From a distributive law $\lambda : \mathbf{TS} \to \mathbf{ST}$, the functor ST inherits a monad structure with unit $1 \xrightarrow{\eta^{\mathbf{S}}\eta^{\mathbf{T}}} ST$ and multiplication $STST \xrightarrow{S\lambda T} SSTT \xrightarrow{\mu^{\mathbf{S}}\mu^{\mathbf{T}}} ST$. String diagrammatically:



(1.27)

This *composite monad* will be denoted by $\mathbf{S} \circ_\lambda \mathbf{T}$, or $\mathbf{S} \circ \mathbf{T}$ when the context is clear. Moreover, the natural transformations $\eta^{\mathbf{S}}T$ and $S\eta^{\mathbf{T}}$ define monad morphisms

$$\eta^{\mathbf{S}}T : \mathbf{T} \to \mathbf{S} \circ \mathbf{T}$$
$$S\eta^{\mathbf{T}} : \mathbf{S} \to \mathbf{S} \circ \mathbf{T}$$

We now give some examples of distributive laws. The first two examples will be used in the sequel to generate examples of iterated weak distributive laws (see Section 2.4).

Example 1.33. The powerset monad distributes over the reader monad via $\lambda : \mathbf{PR} \to \mathbf{RP}$, where $\lambda_X : P(X^A) \to (PX)^A$ is defined by $\lambda_X(U) = \lambda a.\{h(a) \mid h \in U\}$. This distributive law yields a monad structure on the composite functor RP and we denote the composite monad by $\mathbf{R} \circ \mathbf{P}$.

Proof. Given a function $f : X \to Y$, both paths of the naturality diagram are

$$U \in P(X^A) \mapsto \lambda a.\{f(h(a)) \mid h \in U\} \in (PY)^A$$

Given a set X, the $(\eta^{\mathbf{R}})$, $(\mu^{\mathbf{R}})$, $(\eta^{\mathbf{P}})$ and $(\mu^{\mathbf{P}})$ diagrams commute because in each case both paths evaluate the same, respectively, to

$$U \in PX \mapsto \lambda a.U \in (PX)^A$$
$$U \in P((X^A)^A) \mapsto \lambda a.\{H(a)(a) \mid H \in U\} \in (PX)^A$$
$$h \in X^A \mapsto \lambda a.\{h(a)\} \in (PX)^A$$
$$\mathcal{U} \in PP(X^A) \mapsto \lambda a.\{h(a) \mid h \in U \text{ for some } U \in \mathcal{U}\} \in (PX)^A \qquad \square$$

Example 1.34. In a similar fashion, the distribution monad distributes over the reader monad via $\lambda : \mathbf{DR} \to \mathbf{RD}$, where $\lambda_X : D(X^A) \to (DX)^A$ is defined for all $\varphi = \sum_{h \in RX} \varphi_h \cdot h \in D(X^A)$ by $\lambda_X(\varphi) = \lambda a. \sum_{h \in RX} \varphi_h \cdot h(a)$. The corresponding composite monad with functor RD is denoted by $\mathbf{R} \circ \mathbf{D}$.

Proof. Given a function $f : X \to Y$, both paths of the naturality diagram are

$$\varphi \in D(X^A) \mapsto \lambda a. \sum_{h \in RX} \varphi_h \cdot f(h(a)) \in (DY)^A$$

Given a set X, the $(\eta^{\mathbf{R}})$, $(\mu^{\mathbf{R}})$, $(\eta^{\mathbf{D}})$ and $(\mu^{\mathbf{D}})$ diagrams commute because in each case both paths evaluate the same, respectively, to

$$\varphi \in DX \mapsto \lambda a.\varphi \in (DX)^A$$
$$\varphi \in D((X^A)^A)) \mapsto \lambda a. \sum_{H \in (X^A)^A} \varphi_H \cdot H(a)(a) \in (DX)^A$$
$$h \in X^A \mapsto \lambda a.(1 \cdot h(a)) \in (DX)^A$$
$$\Phi \in DD(X^A) \mapsto \lambda a. \sum_{\varphi \in D(X^A)} \sum_{h \in X^A} \Phi_\varphi \varphi_h \cdot h(a) \in (DX)^A \qquad \square$$

Example 1.35 ([99, 2.4.8])**.** The list monad distributes over the powerset monad via $\lambda : \mathbf{LP} \to \mathbf{PL}$ defined by $\lambda_X([U_1, \ldots, U_n]) = \{[x_1, \ldots, x_n] \mid x_i \in U_i\}$, yielding a new monad denoted by $\mathbf{P} \circ \mathbf{L}$.

Example 1.36 ([99, 4.3.4])**.** Similarly, the multiset monad distributes over the powerset monad via $\lambda : \mathbf{MP} \to \mathbf{PM}$ defined by $\lambda_X(\llbracket U_1, \ldots, U_n \rrbracket) = \{\llbracket x_1, \ldots, x_n \rrbracket \mid x_i \in U_i\}$, yielding a new monad denoted by $\mathbf{P} \circ \mathbf{M}$.

Example 1.37 ([80])**.** As recently explicited by Jacobs, the multiset monad distributes over the distribution monad via the so-called *parallel multinomial law* $\lambda : \mathbf{MD} \to \mathbf{DM}$ defined by

$$\lambda_X(\llbracket \varphi_1, \ldots, \varphi_n \rrbracket) = \sum_{(x_1, \ldots, x_n) \in X^n} \prod_{i=1}^{n} \varphi_i(x_i) \cdot \llbracket x_1, \ldots, x_n \rrbracket \tag{1.28}$$

yielding a new monad denoted by $\mathbf{D} \circ \mathbf{M}$.

Example 1.38 ([99, 5.1.6])**.** The list monad distributes over the reader monad via $\lambda : \mathsf{LR} \to \mathsf{RL}$ defined by $\lambda_X([h_1, \ldots, h_n]) = \lambda a.[h_1(a), \ldots, h_n(a)]$, yielding a new monad denoted by $\mathsf{R} \circ \mathsf{L}$.

Example 1.39. For any Set monad T, there is a distributive law $\lambda : (\mathsf{T} + \mathbf{1}) \to \mathsf{T}(- + \mathbf{1})$ defined by $\lambda(\mathsf{inl}(t)) = T\eta_X^{(+\mathbf{1})}(t)$ and $\lambda(\mathsf{inr}(*)) = \eta^{\mathsf{T}}_{(X+1)}(*)$. This yields a monad structure on the functor $T(- + 1)$ denoted by $\mathsf{T} \circ (- + \mathbf{1})$.

Having a distributive law of type $\mathsf{TS} \to \mathsf{ST}$ is actually stronger than having a monad structure on ST. For example, there is no distributive law of type $\mathsf{LL} \to \mathsf{LL}$, but it is possible to define a monad structure with functor LL [161, Remark 5.44]. Distributive laws $\mathsf{TS} \to \mathsf{ST}$ are really about combining *all* the structure of both monads in a smooth way. Equivalently, they are extending T to $\mathsf{Kl}(\mathsf{S})$ or lifting S to $\mathsf{EM}(\mathsf{T})$ [7, 35].

Definition 1.40 (Extension)**.** An *extension* of T to $\mathsf{Kl}(\mathsf{S})$ is a monad $\underline{\mathsf{T}}$ on $\mathsf{Kl}(\mathsf{S})$ such that $\underline{T}F_{\mathsf{S}} = F_{\mathsf{S}}T$ i.e. the diagram below commutes, and such that the natural transformations behave accordingly: $\eta^{\underline{\mathsf{T}}}F_{\mathsf{S}} = F_{\mathsf{S}}\eta^{\mathsf{T}}$ and $\mu^{\underline{\mathsf{T}}}F_{\mathsf{S}} = F_{\mathsf{S}}\mu^{\mathsf{T}}$.

$$\begin{tikzcd} \mathsf{Kl}(\mathsf{S}) \arrow[r, "\underline{T}"] & \mathsf{Kl}(\mathsf{S}) \\ \mathsf{C} \arrow[u, "F_{\mathsf{S}}"] \arrow[r, "T"] & \mathsf{C} \arrow[u, "F_{\mathsf{S}}"'] \end{tikzcd} \tag{1.29}$$

Definition 1.41 (Lifting)**.** A *lifting* of S to $\mathsf{EM}(\mathsf{T})$ is a monad $\overline{\mathsf{S}}$ on $\mathsf{EM}(\mathsf{T})$ such that $U^{\mathsf{T}}\overline{S} = SU^{\mathsf{T}}$ i.e. the diagram below commutes, and such that the natural transformations behave accordingly: $U^{\mathsf{T}}\eta^{\overline{\mathsf{S}}} = \eta^{\mathsf{S}}U^{\mathsf{T}}$ and $U^{\mathsf{T}}\mu^{\overline{\mathsf{S}}} = \mu^{\mathsf{S}}U^{\mathsf{T}}$.

$$\begin{tikzcd} \mathsf{EM}(\mathsf{T}) \arrow[r, "\overline{S}"] \arrow[d, "U^{\mathsf{T}}"'] & \mathsf{EM}(\mathsf{T}) \arrow[d, "U^{\mathsf{T}}"] \\ \mathsf{C} \arrow[r, "S"] & \mathsf{C} \end{tikzcd} \tag{1.30}$$

Now we state the fundamental correspondence theorem, which says that distributive laws, extensions and liftings are three equivalent perspectives.

Theorem 1.42 ([7, 35])**.** *There is a bijective correspondence between*

- *distributive laws of type* $\mathsf{TS} \to \mathsf{ST}$
- *extensions of* T *to* $\mathsf{Kl}(\mathsf{S})$
- *liftings of* S *to* $\mathsf{EM}(\mathsf{T})$

Proof. The proof that distributive laws and liftings are in correspondence is in Beck's paper [7], whereas the proof that they also correspond to extensions is folklore. However, in the whole thesis, it will be important to be able to switch perspectives, so we give the constructions for the following bijections:

$$\text{extensions} \iff \text{distributive laws} \iff \text{liftings}$$

We do not give full details of the verifications, which are long but straightforward.

- extensions $\Rightarrow$ distributive laws

  Let $\underline{\mathbf{T}}$ be an extension of $\mathbf{T}$ to $\mathsf{Kl}(\mathbf{S})$. Let X be an object of C and consider the identity morphism $\mathrm{id}_{SX} : SX \to SX$ as a Kleisli morphism $\mathrm{id}_{SX} : SX \nrightarrow X$. Then, viewing $\underline{T}(\mathrm{id}_{SX}) : TSX \nrightarrow TX$ as a morphism in C again yields the type $TSX \to STX$, which is exactly what is needed for the distributive law. Defining $\lambda_X = \underline{T}(\mathrm{id}_{SX})$, we can check that indeed it verifies the required axioms.

- distributive laws $\Rightarrow$ extensions

  Let $\lambda : \mathbf{TS} \to \mathbf{ST}$ be a distributive law. In order to get an extension of $\mathbf{T}$ to $\mathsf{Kl}(\mathbf{S})$, we are forced to define $\underline{T}X = TX$ for every object X of C. Let $f : X \nrightarrow Y$ be a morphism in $\mathsf{Kl}(\mathbf{S})$. Equivalently, this is a morphism $f : X \to SY$ in C. Then $\lambda_Y \circ Tf : TX \to STY$ can be viewed as a Kleisli morphism of type $TX \nrightarrow TY$. Therefore we define $\underline{T}(f) = \lambda_Y \circ Tf$. Finally, unit and multiplication are forced to be $\eta^{\underline{\mathbf{T}}}_X = \eta^{\mathbf{S}}_{TX} \circ \eta^{\mathbf{T}}_X$ and $\mu^{\underline{\mathbf{T}}}_X = \eta^{\mathbf{S}}_{TX} \circ \mu^{\mathbf{T}}_X$. One can then check that all required properties hold.

- liftings $\Rightarrow$ distributive laws

  Let $\overline{\mathbf{S}}$ be a lifting of $\mathbf{S}$ to $\mathsf{EM}(\mathbf{T})$. Let X be an object of C. Apply $\overline{\mathbf{S}}$ to the $\mathbf{T}$-algebra $(TX, \mu^{\mathbf{T}}_X)$ to get a $\mathbf{T}$-algebra $(STX, \overline{S}\mu^{\mathbf{T}}_X)$ whose type as a morphism is $TSTX \to STX$. Precomposing with $TS\eta^{\mathbf{T}}_X$ yields the type $TSX \to STX$, therefore we can define $\lambda_X = \overline{S}\mu^{\mathbf{T}}_X \circ TS\eta^{\mathbf{T}}_X$ and check all axioms. Note that here, $\overline{S}\mu^{\mathbf{T}}_X$ is a slight notation abuse that denotes the C-morphism extracted from the $\mathbf{T}$-algebra $\overline{S}(TX, \mu^{\mathbf{T}}_X)$. Formally, $\overline{S}\mu^{\mathbf{T}}_X$ is $U^{\mathbf{T}}\epsilon^{\mathbf{T}}_{\overline{S}F^{\mathbf{T}}X}$, where $\epsilon^{\mathbf{T}} : F^{\mathbf{T}}U^{\mathbf{T}} \to 1$ is the counit of the Eilenberg-Moore adjunction.

- distributive laws $\Rightarrow$ liftings

Let $\lambda : \mathsf{TS} \to \mathsf{ST}$ be a distributive law. For any T-algebra (X, x), one can check that the morphism $Sx \circ \lambda_X : TSX \to SX$ defines a T-algebra. Hence we define $\overline{S}(X, x) = (SX, Sx \circ \lambda_X)$. The other data of the lifting are forced to be $\overline{S}f = Sf$, $\eta^{\overline{\mathsf{S}}}_X = \eta^{\mathsf{S}}_X$ and $\mu^{\overline{\mathsf{S}}}_X = \mu^{\mathsf{S}}_X$. These assignments can be shown to satisfy all required properties.

Finally, note that these constructions are inverse to each other. $\square$

In the proof we can remark that once the functor of a monad is extended / lifted, there is at most one extension / lifting of its unit and multiplication. The important parts of the constructions can be summed up as

$$\lambda_X = \underline{T}(\mathrm{id}_{SX}) \qquad \lambda_X = \overline{S}\mu^{\mathsf{T}}_X \circ TS\eta^{\mathsf{T}}_X$$
$$\underline{T}(f) = \lambda_Y \circ Tf \qquad \overline{S}(X, x) = (SX, Sx \circ \lambda_X)$$

Table 1.1 synthetises the interplay between axioms of distributive laws and properties of their corresponding extensions and liftings.

Table 1.1: Property correspondence in bijections of Theorem 1.42

extension	distributive law	lifting
$\underline{T}$ preserves 1	(η^{S}) axiom	$\eta^{\overline{\mathsf{S}}}$ components are algebra morphisms
$\underline{T}$ preserves $\bullet$	(μ^{S}) axiom	$\mu^{\overline{\mathsf{S}}}$ components are algebra morphisms
$\eta^{\underline{\mathsf{T}}}$ is natural	(η^{T}) axiom	any $\overline{S}(X, x)$ satisfies the T-unit axiom
$\mu^{\underline{\mathsf{T}}}$ is natural	(μ^{T}) axiom	any $\overline{S}(X, x)$ satisfies the T-associativity axiom

We now give examples of the extensions and liftings for the distributive laws presented before.

Example 1.43. Consider the distributive law $\lambda : \mathsf{LP} \to \mathsf{PL}$ of Example 1.35. According to Theorem 1.42, this is equivalently an extension of L to the Kleisli category of P, that is Rel, or a lifting of P to the Eilenberg-Moore category of L, that is Mon. The extension is defined on a relation $R \subseteq X \times Y$ by:

$$\underline{L}(R) = \{([x_1, ..., x_n], [y_1, ..., y_n]) \mid n \in \mathbb{N}, (x_i, y_i) \in R\} \tag{1.31}$$

The lifting is defined on a monoid $(M, *, e)$ by

$$\overline{P}(M, *, e) = (PX, (U, V) \mapsto \{x * y \mid x \in U, y \in V\}, \{e\}) \tag{1.32}$$

The monoid $\overline{P}(M, *, e)$ is an example of a *complex algebra*, see [55].

Example 1.44. The distributive law $\lambda : \mathbf{LR} \to \mathbf{RL}$ of Example 1.38 gives rise to an extension of $\mathbf{L}$ to $\mathsf{Kl}(\mathbf{R})$ defined for every $f : X \to Y^A$ (seen as a function $X \times A \to Y$) by

$$\underline{L}f([x_1, ..., x_n]) = \lambda a.[f(x_1, a), ..., f(x_n, a)] \tag{1.33}$$

and to a lifting of $\mathbf{R}$ to Mon defined for every monoid $(M, *, e)$ by

$$\overline{R}(M, *, e) = (X^A, (g, h) \mapsto \lambda a.g(a) * h(a), \lambda a.e) \tag{1.34}$$

Example 1.45. The distributive law $\lambda : (\mathbf{T} + \mathbf{1}) \to \mathbf{T}(- + \mathbf{1})$ of Example 1.39 gives rise to an extension of $(- + \mathbf{1})$ to $\mathsf{Kl}(\mathbf{T})$ defined for every $f : X \to TY$ by

$$\underline{(f + 1)}(\mathsf{inl}(x)) = T\eta_X^{(+\mathbf{1})}(f(x)) \tag{1.35}$$

$$\underline{(f + 1)}(\mathsf{inr}(*)) = \eta^{\mathbf{T}}_{(X+1)}(*) \tag{1.36}$$

and to a lifting of $\mathbf{T}$ to $\mathsf{EM}(- + \mathbf{1})$ defined for every pointed set (X, p) by

$$\overline{T}(X, p) = (TX, \eta^{\mathbf{T}}_X(p)) \tag{1.37}$$

Let $\lambda : \mathbf{TS} \to \mathbf{ST}$ be a distributive law. The category $\mathsf{Alg}(\lambda)$ of λ-algebras is defined as follows. Objects are triples (X, τ, σ) where (X, τ) is a $\mathbf{T}$-algebra and (X, σ) is a $\mathbf{S}$-algebra such that the following diagram commutes. Intuitively, this means that we have an object X with two algebraic structures related by the distributivity axioms encoded by λ.

$$\begin{tikzcd} TSX \arrow[rr, "\lambda_X"] \arrow[d, "T\sigma"'] & & STX \arrow[d, "S\tau"] \\ TX \arrow[dr, "\tau"'] & & SX \arrow[dl, "\sigma"] \\ & X & \end{tikzcd} \tag{1.38}$$

Morphisms $(X, \tau, \sigma) \to (Y, \tau', \sigma')$ are morphisms $f : X \to Y$ in C such that $f : (X, \tau) \to (Y, \tau')$ is a morphism of $\mathbf{T}$-algebras and $f : (X, \sigma) \to (Y, \sigma')$ is a morphism of $\mathbf{S}$-algebras. The free λ-algebra on X is $(STX, S\mu^{\mathbf{T}}_X \circ \lambda_{TX}, \mu^{\mathbf{S}}_{TX})$. The full subcategory of free λ-algebras is denoted by $\mathsf{FAlg}(\lambda)$.

Proposition 1.46 ([7])**.** *Let $\lambda : \mathbf{TS} \to \mathbf{ST}$ be a distributive law. Then the composite monad $\mathbf{S} \circ_\lambda \mathbf{T}$ coincides with the monads described by the composite adjunctions*

$$\begin{tikzcd} \mathsf{EM}(\overline{\mathbf{S}}) \arrow[r, bend right, "U^{\overline{\mathbf{S}}}"'] \arrow[r, phantom, "\perp"] & \mathsf{EM}(\mathbf{T}) \arrow[l, bend right, "F^{\overline{\mathbf{S}}}"'] \arrow[r, bend right, "U^{\mathbf{T}}"'] \arrow[r, phantom, "\perp"] & \mathsf{C} \arrow[l, bend right, "F^{\mathbf{T}}"'] \end{tikzcd} \tag{1.39}$$

$$\begin{tikzcd} \mathsf{Kl}(\underline{\mathbf{T}}) \arrow[r, bend right, "U_{\underline{\mathbf{T}}}"'] \arrow[r, phantom, "\perp"] & \mathsf{Kl}(\mathbf{S}) \arrow[l, bend right, "F_{\underline{\mathbf{T}}}"'] \arrow[r, bend right, "U_{\mathbf{S}}"'] \arrow[r, phantom, "\perp"] & \mathsf{C} \arrow[l, bend right, "F_{\mathbf{S}}"'] \end{tikzcd} \tag{1.40}$$

Moreover, $\mathsf{EM}(\mathbf{S} \circ \mathbf{T}) \cong \mathsf{EM}(\overline{\mathbf{S}}) \cong \mathsf{Alg}(\lambda)$ *and* $\mathsf{Kl}(\mathbf{S} \circ \mathbf{T}) \cong \mathsf{Kl}(\underline{\mathbf{T}}) \equiv \mathsf{FAlg}(\lambda)$.

Proof. The assertions concerning Eilenberg-Moore categories are present in [7]. The fact that the composite Kleisli adjunction yields $\mathbf{S} \circ_\lambda \mathbf{T}$ is an elementary but tedious computation. Isomorphisms are given by

$$\mathsf{EM}(\overline{\mathbf{S}}) \cong \mathsf{Alg}(\lambda)$$
$$((X, \tau), \sigma) \leftrightarrow (X, \tau, \sigma)$$

$$\mathsf{EM}(\mathbf{S} \circ \mathbf{T}) \cong \mathsf{Alg}(\lambda)$$
$$(X, x) \mapsto (X, x \circ \eta^{\mathbf{S}}_{TX}, x \circ S\eta^{\mathbf{T}}_X)$$
$$(X, \sigma \circ S\tau) \leftarrow\!\shortmid (X, \tau, \sigma)$$

The isomorphism $\mathsf{Kl}(\mathbf{S} \circ \mathbf{T}) \cong \mathsf{Kl}(\underline{\mathbf{T}})$ is the identity: these categories have the same objects, morphisms, and a short calculation shows that identities and composition coincide. Finally, the isomorphism $\mathsf{EM}(\mathbf{S} \circ \mathbf{T}) \cong \mathsf{Alg}(\lambda)$ restricts to an isomorphism between the category of free $(\mathbf{S} \circ \mathbf{T})$-algebras (which is equivalent to $\mathsf{Kl}(\mathbf{S} \circ \mathbf{T})$) and $\mathsf{FAlg}(\lambda)$. $\square$

1.4 Algebraic Presentations

Let $\mathcal{V}$ be a countable set whose elements are called variables. A *signature* Σ is a set of symbols $\sigma \in \Sigma$, each symbol having an arity $\mathsf{ar}(\sigma) \in \mathbb{N}$. The set of Σ-*terms* with variables in $\mathcal{V}$ is denoted by $\mathsf{Term}_\Sigma(\mathcal{V})$ and generated by the following grammar

$$t ::= x \mid \sigma(t_1, \ldots, t_n) \tag{1.41}$$

where $x \in \mathcal{V}$, $n \in \mathbb{N}$ and $\sigma \in \Sigma$ such that $\mathsf{ar}(\sigma) = n$.

Let Σ be a signature. A Σ-*algebra* A consists of

- a set, also denoted by A
- for every symbol $\sigma \in \Sigma$ with $\mathsf{ar}(\sigma) = n$, a function $[\sigma]_A : A^n \to A$

A *valuation* is a function $v : \mathcal{V} \to A$. The v-evaluation of a term $t \in \mathsf{Term}_\Sigma(\mathcal{V})$ is defined inductively by

- $[x]_{A,v} = v(x)$ for all $x \in \mathcal{V}$
- $[\sigma(t_1, \ldots, t_n)]_{A,v} = [\sigma]_A([t_1]_{A,v}, \ldots, [t_n]_{A,v})$ for every $n \in \mathbb{N}$, $\sigma \in \Sigma$ such that $\mathsf{ar}(\sigma) = n$

Given a signature Σ, a Σ-*equation* is a pair $(t_1, t_2) \in \mathsf{Term}_\Sigma(\mathcal{V})^2$, more conveniently denoted by $t_1 \approx t_2$. A set of equations E generates a congruence denoted by $\approx_E$ on the set $\mathsf{Term}_\Sigma(\mathcal{V})$ with respect to the operation of term substitution. The equivalence class of a term t is denoted by $[t]_{\approx_E}$. A pair (Σ, E) is called an *equational theory.* A (Σ, E)-*algebra* A is a Σ-algebra such that for every $t_1 \approx t_2$ in E and every valuation $v : \mathcal{V} \to A$, $[t_1]_{A,v} = [t_2]_{A,v}$. The quotient $\mathsf{Term}_\Sigma(\mathcal{V})/\approx_E$ is denoted by $F\mathcal{V}$ and called the free (Σ, E)-algebra on $\mathcal{V}$, with operations defined by

$$[\sigma]_{F\mathcal{V}}([t_1]_{\approx_E}, \ldots, [t_n]_{\approx_E}) = [\sigma(t_1, \ldots, t_n)]_{\approx_E} \tag{1.42}$$

Concretely, in $\mathsf{Term}_\Sigma(\mathcal{V})/\approx_E$, any two terms that are provably equal with respect to equational logic and equations in E are identified. Given any set X, one can construct as well a free (Σ, E)-algebra with variables in X, denoted by $FX = \mathsf{Term}_\Sigma(X)/\approx_E$.

A morphism of (Σ, E)-algebras A and B is a function $f : A \to B$ such that for all $\sigma \in \Sigma$ with arity n and $a_1, \ldots, a_n \in A$, $f([\sigma]_A(a_1, \ldots, a_n)) = [\sigma]_B(f(a_1), \ldots, f(a_n))$. The category of (Σ, E)-algebras and morphisms is denoted by $\mathsf{Alg}(\Sigma, E)$. The free (Σ, E)-algebra construction on a set X induces a functor $F : \mathsf{Set} \to \mathsf{Alg}(\Sigma, E)$, which has a right adjoint $U : \mathsf{Alg}(\Sigma, E) \to \mathsf{Set}$ forgetting the algebraic structure. This adjunction induces a monad $\mathbf{T}$ with functor UF on Set, and $\mathsf{Alg}(\Sigma, E) \cong \mathsf{EM}(\mathbf{T})$ [96, VI.8 Theorem 1].

$$\begin{tikzcd} \mathsf{EM}(\mathbf{T}) \cong \mathsf{Alg}(\Sigma, E) \arrow[r, bend right, "U"'] \arrow[r, phantom, "\perp"] & \mathsf{Set} \arrow[l, bend right, "F"'] \arrow[loop right, "\mathbf{T}"] \end{tikzcd} \tag{1.43}$$

Definition 1.47. Let $\mathbf{T}$ be a monad on Set. If $\mathsf{EM}(\mathbf{T}) \cong \mathsf{Alg}(\Sigma, E)$ for some signature Σ and some set of Σ-equations E, the monad $\mathbf{T}$ is said to be *presented by* (Σ, E).

Note that if they exist, presentations are not unique: a monad can be presented by many different equational theories.

The algebraic way of looking at monads provides intuitive insights about what monads constructs and monad properties mean. To illustrate, we fix s a symbol of arity 1, $*$ an (infix) symbol of arity 2, and $(\Sigma, E) = (\{s, *\}, \emptyset)$ the equational theory over these symbols with no equations.

- An element of TX can be seen as a term with variables in X. We use brackets to denote that the expression is a term, as in e.g. $\langle a * s(b)\rangle$.

- Given a function $f : X \to Y$ between sets of variables, the function Tf substitutes variables inside terms, e.g. if $f(a) = f(b) = c$, $Tf(\langle a * s(b)\rangle) = \langle c * s(c)\rangle$.

- The unit $\eta^{\mathsf{T}}_X : X \to TX$ says that any variable can be considered as a term, e.g. $\langle a\rangle$ is a term.

- The multiplication $\mu^{\mathsf{T}}_X : TTX \to TX$ says that a term of terms can be flattened, e.g. $\langle\langle a * s(b)\rangle * \langle c * s(s(c))\rangle\rangle$ can be flattened to $\langle(a * s(b)) * (c * s(s(c)))\rangle$.

- The unit axiom $\mu^{\mathsf{T}} \circ T\eta^{\mathsf{T}} = \mathrm{id}_T$ means that given a term, if all variables are considered themselves as terms, then flattening outputs the original term. For example, flattening $\langle\langle a\rangle * s(\langle b\rangle)\rangle$ yields $\langle a * s(b)\rangle$.

- A T-algebra (X, x) can be seen as evaluating terms to their value via the function $x : TX \to X$. For example, if $X = \mathbb{N}$ and x interprets s as the successor operation and $*$ as addition, then $x(\langle 0 * s(1)\rangle) = 0 + (1 + 1) = 2$.

- The unit axiom for T-algebras says that for any variable, considering this variable as a term then evaluating it just outputs the original variable i.e. $x(\langle a\rangle) = a$.

- Evaluating a term of terms can be done in two different ways: one can either flatten it and then evaluate the result, or evaluate the inner terms then evaluate the result. The associativity axiom for T-algebras says that these processes are equivalent. For instance, $\langle\langle 1*2\rangle*\langle s(3)*0\rangle\rangle$ can be flattened to $\langle(1*2)*(s(3)*0)\rangle$ then evaluated to $(1+2)+((3+1)+0) = 7$, or the inner terms can be evaluated as $\langle 3 * 4\rangle$, then the resulting term evaluated to $3 + 4 = 7$.

- The unit axiom $\mu^{\mathsf{T}} \circ \eta^{\mathsf{T}} T = \mathrm{id}_T$ and the associativity axiom $\mu^{\mathsf{T}} \circ \mu^{\mathsf{T}} T = \mu^{\mathsf{T}} \circ T\mu^{\mathsf{T}}$ together mean that (TX, μ^{T}_X) is a T-algebra, that is, flattening is itself a correct evaluation procedure.

A monad is *finitary* if its functor commutes with filtered colimits. An important result is that any finitary monad is presented by some equational theory [29, §4.6].

Example 1.48. The finite powerset monad P_f is presented by the equational theory of join-semilattices with bottom. The signature of this theory is $\Sigma = \{\bot, \vee\}$ with $\mathsf{ar}(\bot) = 0$ and $\mathsf{ar}(\vee) = 2$, and equations are

- $x \vee \bot = x$

- $x \vee x = x$
- $x \vee y = y \vee x$
- $(x \vee y) \vee z = x \vee (y \vee z)$

Example 1.49 ([88])**.** The distribution monad D is presented by the equational theory of convex algebras. The signature of this theory is $\Sigma = \{\oplus_r \mid r \in [0,1]\}$, and equations are

- $x \oplus_1 y = x$
- $x \oplus_r x = x$
- $x \oplus_r y = y \oplus_{1-r} x$
- $(x \oplus_p y) \oplus_r z = x \oplus_{pr} \left(y \oplus_{\frac{r-pr}{1-pr}} z\right)$ if $r, p \neq 1$

Equationally, a morphism of convex algebras is just a function f satisfying $f(x \oplus_r y) = f(x) \oplus_r f(y)$, that is, an affine function.

It may happen that some non-finitary monads also have an equational theory, provided arities of symbols in the signature are allowed to have infinite cardinalities, and for each cardinality there is a corresponding set of distinct variables. This is the case of the powerset monad.

Example 1.50 ([29, 22])**.** The powerset monad P is presented by the infinitary theory of complete join-semilattices. The signature of this theory is $\Sigma = \{\bigvee_{i \in I} \mid I \text{ is a set}\}$ (which is actually a proper class) and the equations are

- $\bigvee_{i \in \{*\}} x_i = x_*$
- $\bigvee_{j \in J} x_j = \bigvee_{i \in I} x_{f(i)}$ for every surjective $f : I \to J$
- $\bigvee_{i \in I} x_i = \bigvee_{j \in J} \bigvee_{i \in f^{-1}(\{j\})} x_i$ for every $f : I \to J$

In our approach to monads and distributive laws, algebraic presentations are not central – see [161] and [112] for two recent dissertations about monads and distributive laws relying strongly on algebraic presentations. Still, there is one application of distributive laws to algebraic presentations that we would like to mention. Given a distributive law $\lambda : \mathsf{TS} \to \mathsf{ST}$, Proposition 1.46 provides an isomorphism $\mathsf{EM}(\mathsf{S} \circ_\lambda \mathsf{T}) \cong \mathsf{Alg}(\lambda)$. Therefore, finding an algebraic presentation of the composite monad $\mathsf{S} \circ_\lambda \mathsf{T}$ amounts to finding an equational theory (Σ, E) such that $\mathsf{Alg}(\lambda) \cong \mathsf{Alg}(\Sigma, E)$. Recall that λ-algebras are just triples (X, τ, σ) where (X, τ) is a $\mathbf{T}$-algebra, (X, σ) is a $\mathbf{S}$-algebra, and the following pentagon commutes

$$\begin{tikzcd}
TSX \arrow[rr, "\lambda_X"] \arrow[d, "T\sigma"'] & & STX \arrow[d, "S\tau"] \\
TX \arrow[dr, "\tau"'] & & SX \arrow[dl, "\sigma"] \\
& X &
\end{tikzcd} \tag{1.44}$$

Therefore, if (Σ_T, E_T) is an algebraic presentation of $\mathbf{T}$ and (Σ_S, E_S) is an algebraic presentation of $\mathbf{S}$, the above pentagon intuitively encodes the set of missing equations E such that $(\Sigma_T \cup \Sigma_S, E_T \cup E_S \cup E)$ is an algebraic presentation of the composite monad $\mathbf{S} \circ_\lambda \mathbf{T}$. Using commutativity of the pentagon, one may guess new equations that can eventually lead to an algebraic presentation of the composite monad. Theorem 3.16 will put this method into practice to yield an algebraic presentation for the *monad of convex sets of distributions* out of algebraic presentations for the powerset monad and the distribution monad.

1.5 Iterated Distributive Laws

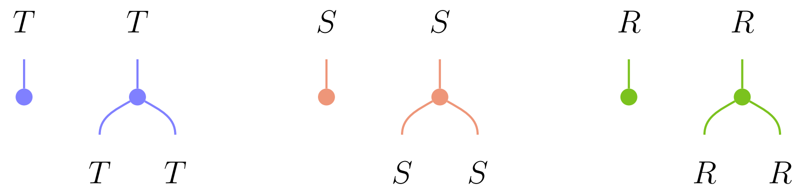
As we have seen before, distributive laws are a convenient tool allowing to compose two monads. This framework being compositional, it is supposed to allow us to compose even more monads in a smooth way. Let $(\mathbf{T}_i)_{i\in\mathbb{N}^*}$ be a sequence of monads. If we know that $\mathbf{T}_1 \dots \mathbf{T}_n$ is a monad, we only need a distributive law of type $\mathbf{T}_{n+1}(\mathbf{T}_1 \dots \mathbf{T}_n) \to (\mathbf{T}_1 \dots \mathbf{T}_n)\mathbf{T}_{n+1}$ to define a monad structure on $T_1 \dots T_n T_{n+1}$. As we stack more and more monads, such distributive laws can become tedious to define, and their diagrams laborious to check. A possible solution is to use results of Cheng [34] about *iterated distributive laws.* The principle of this approach is to generate a monad structure on $T_1 \dots T_n$ by means of $O(n^2)$ distributive laws involving only 2 monads T_i each, and $O(n^3)$ new equations to check involving only 3 monads T_i each. In some sense, this is a trade-off between the number of objects manipulated and their complexity. We present here the main theorem of [34] and give a proof of the case $n = 3$, because the details will be needed in the next chapters.

Theorem 1.51 ([34, Theorem 1.6])**.** *Let $n \in \mathbb{N}$, $n \geq 3$, and let $(T_i)_{1 \leq i \leq n}$ be monads on a category C. For all $1 \leq j < i \leq n$, let $\lambda_{ij} : \mathbf{T}_i\mathbf{T}_j \to \mathbf{T}_j\mathbf{T}_i$ be a distributive law, and assume that for all $1 \leq k < j < i \leq n$ the so-called Yang-Baxter diagram commutes:*

$$\begin{tikzcd}[column sep=small]
& T_jT_iT_k \arrow[r, "T_j\lambda_{ik}"] & T_jT_kT_i \arrow[dr, "\lambda_{jk}T_i"] & \\
T_iT_jT_k \arrow[ur, "\lambda_{ij}T_k"] \arrow[dr, "T_i\lambda_{jk}"'] & & & T_kT_jT_i \\
& T_iT_kT_j \arrow[r, "\lambda_{ik}T_j"'] & T_kT_iT_j \arrow[ur, "T_k\lambda_{ij}"'] &
\end{tikzcd} \tag{1.45}$$

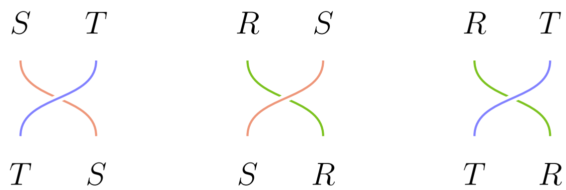
Then for all $1 \leq i < n$, $(T_1 \dots T_i)$ and $(T_{i+1} \dots T_n)$ have a monad structure and there is a distributive law of type $(\mathbf{T}_{i+1} \dots \mathbf{T}_n)(\mathbf{T}_1 \dots \mathbf{T}_i) \to (\mathbf{T}_1 \dots \mathbf{T}_i)(\mathbf{T}_{i+1} \dots \mathbf{T}_n)$. Moreover, the induced monad structures on $T_1 \dots T_n$ are the same, allowing to write unambiguously $\mathbf{T}_1 \dots \mathbf{T}_n$.

Proof. The proof is by induction on n. We will treat only the case $n = 3$ and use string diagrams. For a complete proof using commutative diagrams instead, see [34]. Let $\mathbf{R}$, $\mathbf{S}$, $\mathbf{T}$ be monads on C. In string diagrams:



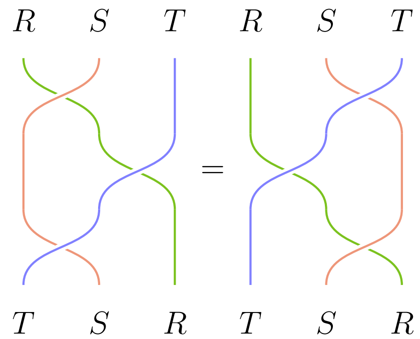
(1.46)

Let $\lambda : \mathbf{TS} \to \mathbf{ST}$, $\sigma : \mathbf{SR} \to \mathbf{RS}$ and $\tau : \mathbf{TR} \to \mathbf{RT}$ be distributive laws.



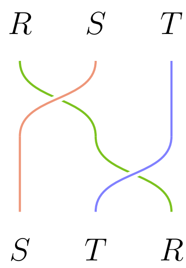
(1.47)

Each of them satisfies equations (1.24) and (1.25). The Yang-Baxter diagram gives the equation



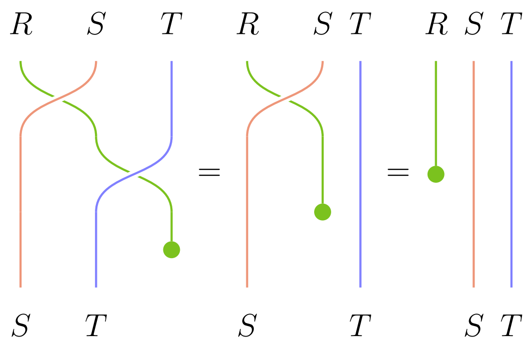
(1.48)

Define the natural transformation $\delta = STR \xrightarrow{S\tau} SRT \xrightarrow{\sigma T} RST$, this is:



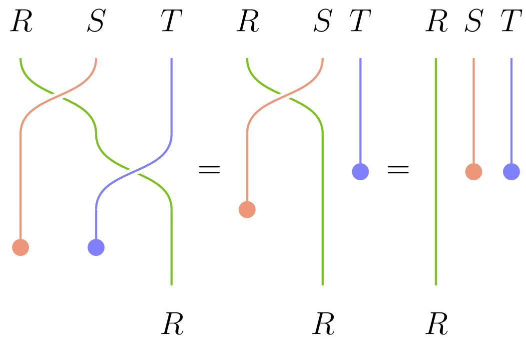
(1.49)

Let us prove that δ is a distributive law of type $(\mathbf{S} \circ_\lambda \mathbf{T})\mathbf{R} \to \mathbf{R}(\mathbf{S} \circ_\lambda \mathbf{T})$. Remind that the monad $\mathbf{S} \circ_\lambda \mathbf{T}$ has unit and multiplication given by equation (1.27). For the $\eta^{\mathbf{R}}$ axiom of δ, we need the $\eta^{\mathbf{R}}$ axioms of both τ and σ:



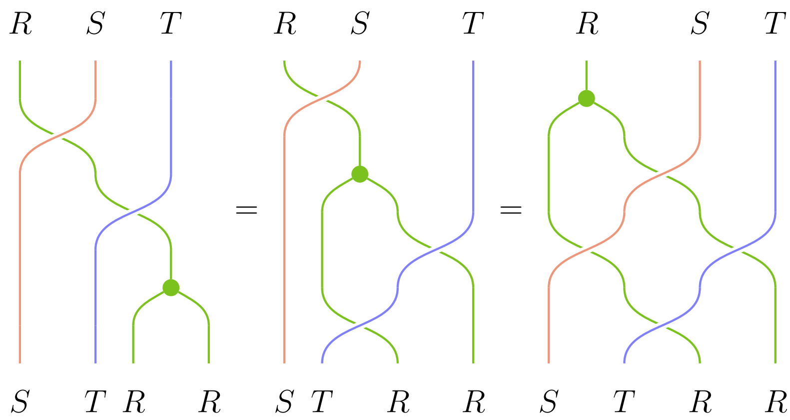
(1.50)

For the $\eta^{\mathbf{S} \circ_\lambda \mathbf{T}}$ axiom of δ, we need the $\eta^{\mathbf{T}}$ axiom of τ and the $\eta^{\mathbf{S}}$ axiom of σ:



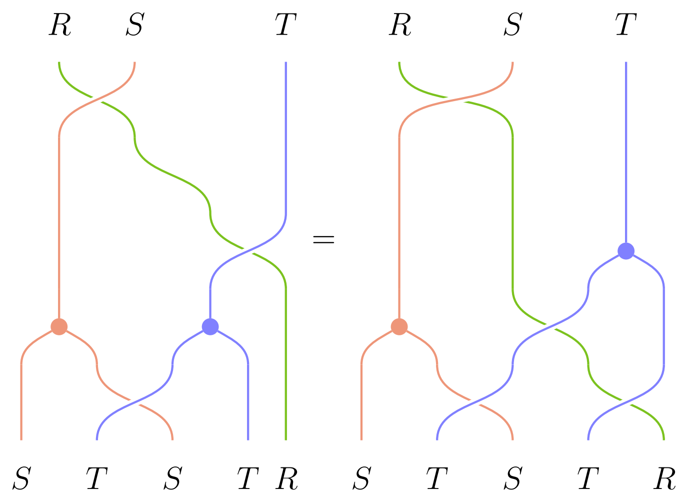
(1.51)

For the $\mu^{\mathbf{R}}$ axiom of δ, we need the $\mu^{\mathbf{R}}$ axioms of both τ and σ:

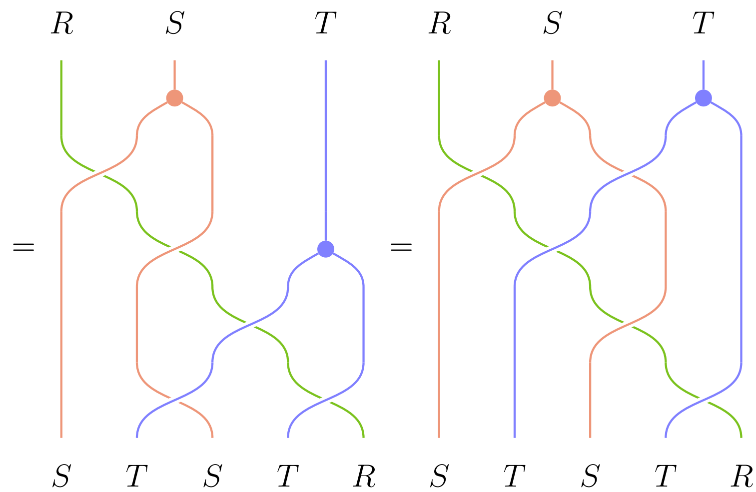


(1.52)

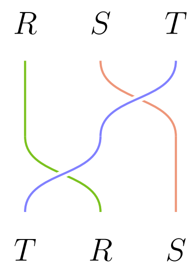
For the $\mu^{\mathsf{S}\circ_\lambda\mathsf{T}}$ axiom of δ, we need the μ^{T} axiom of τ, the μ^{S} axiom of σ, and the Yang-Baxter equation (applied in this order):



(1.53)

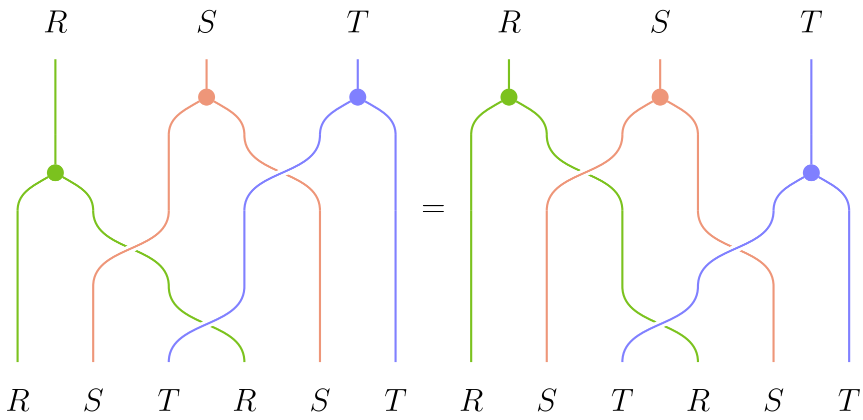


Hence δ is a distributive law. Alternatively, we could have defined a natural transformation $\delta' = TRS \xrightarrow{\tau S} RTS \xrightarrow{R\lambda} RST$ and proved that δ' is a distributive law of type $\mathsf{T}(\mathsf{R}\circ_\sigma\mathsf{S}) \to (\mathsf{R}\circ_\sigma\mathsf{S})\mathsf{T}$:



(1.54)

It remains to check that δ and δ' generate the same multiplication on RST:



(1.55)

□

Chapter 2

The Weak Framework

2.1 Weak Distributive Laws

Distributive laws are sometimes too restrictive, in the sense that simultaneously imposing the four compatibility axioms may make the existence of a distributive law of a certain type impossible. See the thesis of Maaike Zwart [161] for many concrete examples of this situation. In the recent years, authors have tried to mitigate this issue by relaxing some of the compatibility conditions in a number of different ways. Street [146] and Böhm [15] studied different such weakenings, first independently, then together with Lack [16]. A decade later, Richard Garner [54] highlighted a particular case of their framework and used it to prove that the Vietoris monad on the category of compact Hausdorff spaces (i.e. Eilenberg-Moore algebras for the ultrafilter monad) was a *sort of* lifting of the powerset monad. In this section, we present Garner's way of weakening distributive laws.

The different weakenings are slightly intertwined in the literature. We do our best to give appropriate credit to every author. To sum up:

- The general philosophy of weakening axioms and the leitmotiv of splitting idempotents are first due to Street and Böhm, independently.

- The most abstract results, operating in a 2-categorical setting, are due to Böhm, joined later by Street and Lack.

- The specific weakening of axioms that we use in the sequel is due to Garner as a particuliar instance of Böhm's results. The spirit is very close to Street, though no general implication holds between being a Street weak distributive law (in the sense of [146]) and being a Garner weak distributive law (in the sense of [54]). More precisely, if a law is both Garner-weak and Street-weak, then it is a distributive law. An important contribution of Garner is to bring weak distributive laws into concrete situations involving well-known Set functors.

Definition 2.1 (Weak distributive law)**.** A *weak distributive law* of type $\mathbf{TS} \to \mathbf{ST}$ is a natural transformation λ of this type such that the following equations hold.

$$\lambda \circ T\eta^{\mathbf{S}} = \eta^{\mathbf{S}}T \qquad (\eta^{\mathbf{S}} \text{ axiom})$$
$$\lambda \circ T\mu^{\mathbf{S}} = \mu^{\mathbf{S}}T \circ S\lambda \circ \lambda S \qquad (\mu^{\mathbf{S}} \text{ axiom})$$
$$\lambda \circ \mu^{\mathbf{T}}S = S\mu^{\mathbf{T}} \circ \lambda T \circ T\lambda \qquad (\mu^{\mathbf{T}} \text{ axiom})$$

From now on, a distributive law in the sense of Definition 1.30 may sometimes be called a *plain* distributive law, to insist on the fact that it satisfies all axioms.

How does this weakening of distributive laws impact the correspondence of Theorem 1.42? Deleting the ($\eta^{\mathbf{T}}$) axiom acts on Table 1.1 in two ways:

- One cannot prove that $\eta^{\mathbf{T}}_X$ is natural anymore. This can be easily patched by relaxing the statement that extensions comprise a unit, leading directly to the definition of weak extensions.

- One cannot prove that $\overline{S}(X, x)$ satisfies the $\mathbf{T}$-unit axiom anymore. This is more challenging, because this means the traditional $\overline{S}$ is not even a functor: it maps algebras to *semialgebras*, i.e., pairs (Y, y) with $y : TY \to Y$ merely satisfying the $\mathbf{T}$-associativity axiom. Equivalently, $Sx \circ \lambda_X \circ \eta^{\mathbf{T}}_{SX}$ is not the identity, but merely an idempotent. The clever patch to this situation, originating in both Street and Böhm papers, consists in splitting this idempotent. More specifically, in our case Garner puts this idea into action by (1) making a detour into the category of semialgebras and (2) recovering an algebra by splitting the faulty idempotent. The splitting generates natural transformations π and ι that act like algebra-semialgebra translators. This leads to the definition of weak liftings.

Definition 2.2 (Weak extension)**.** A *weak extension* of $\mathbf{T}$ to $\mathsf{Kl}(\mathbf{S})$ is a pair $(\underline{T}, \mu^{\underline{\mathbf{T}}})$ comprising a functor $\underline{T} : \mathsf{Kl}(\mathbf{S}) \to \mathsf{Kl}(\mathbf{S})$ and a natural transformation $\mu^{\underline{\mathbf{T}}} : \underline{T}\underline{T} \to \underline{T}$ such that $\underline{T}F_{\mathbf{S}} = F_{\mathbf{S}}T$ and $\mu^{\underline{\mathbf{T}}}F_{\mathbf{S}} = F_{\mathbf{S}}\mu^{\mathbf{T}}$.

Definition 2.3 (Weak lifting)**.** A *weak lifting* of $\mathbf{S}$ to $\mathsf{EM}(\mathbf{T})$ is a monad $\overline{\mathbf{S}}$ on $\mathsf{EM}(\mathbf{T})$ along with two natural transformations

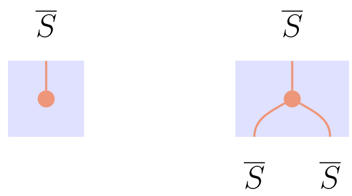
$$\pi : SU^{\mathbf{T}} \to U^{\mathbf{T}}\overline{S} \qquad\qquad \iota : U^{\mathbf{T}}\overline{S} \to SU^{\mathbf{T}}$$

such that $\pi \circ \iota = 1$ and the four following diagrams commute

$$
\begin{tikzcd}
 & U^{\mathsf{T}} \arrow[dl, "\eta^{\mathsf{S}} U^{\mathsf{T}}"'] \arrow[dr, "U^{\mathsf{T}} \eta^{\overline{\mathsf{S}}}"] \arrow[d, phantom, "(\pi.\eta^{\mathsf{S}})"] & \\
SU^{\mathsf{T}} \arrow[rr, "\pi"'] & & U^{\mathsf{T}}\overline{S}
\end{tikzcd}
\qquad
\begin{tikzcd}
 & U^{\mathsf{T}} \arrow[dl, "U^{\mathsf{T}} \eta^{\overline{\mathsf{S}}}"'] \arrow[dr, "\eta^{\mathsf{S}} U^{\mathsf{T}}"] \arrow[d, phantom, "(\iota.\eta^{\mathsf{S}})"] & \\
U^{\mathsf{T}}\overline{S} \arrow[rr, "\iota"'] & & SU^{\mathsf{T}}
\end{tikzcd}
$$

$$
\begin{tikzcd}
SSU^{\mathsf{T}} \arrow[r, "S\pi"] \arrow[d, "\mu^{\mathsf{S}} U^{\mathsf{T}}"'] & SU^{\mathsf{T}}\overline{S} \arrow[r, "\pi\overline{S}"] \arrow[d, phantom, "(\pi.\mu^{\mathsf{S}})"] & U^{\mathsf{T}}\overline{S}\overline{S} \arrow[d, "U^{\mathsf{T}}\mu^{\overline{\mathsf{S}}}"] \\
SU^{\mathsf{T}} \arrow[rr, "\pi"'] & & U^{\mathsf{T}}\overline{S}
\end{tikzcd}
\qquad
\begin{tikzcd}
U^{\mathsf{T}}\overline{S}\overline{S} \arrow[r, "\iota\overline{S}"] \arrow[d, "U^{\mathsf{T}}\mu^{\overline{\mathsf{S}}}"'] & SU^{\mathsf{T}}\overline{S} \arrow[r, "S\iota"] \arrow[d, phantom, "(\iota.\mu^{\mathsf{S}})"] & SSU^{\mathsf{T}} \arrow[d, "\mu^{\mathsf{S}} U^{\mathsf{T}}"] \\
U^{\mathsf{T}}\overline{S} \arrow[rr, "\iota"'] & & SU^{\mathsf{T}}
\end{tikzcd}
$$

A string diagram representation possibly exposes these conditions with greater clarity. The monad $\overline{\mathbf{S}}$ on $\mathsf{EM}(\mathbf{T})$ is

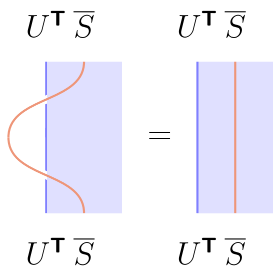


(2.1)

and the natural transformations π, ι respectively are

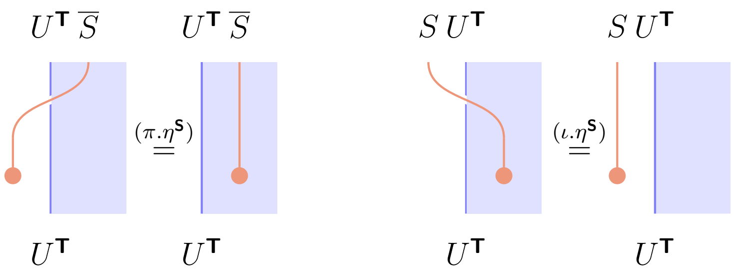
(2.2)

The condition $\pi \circ \iota = 1$ is pictured as



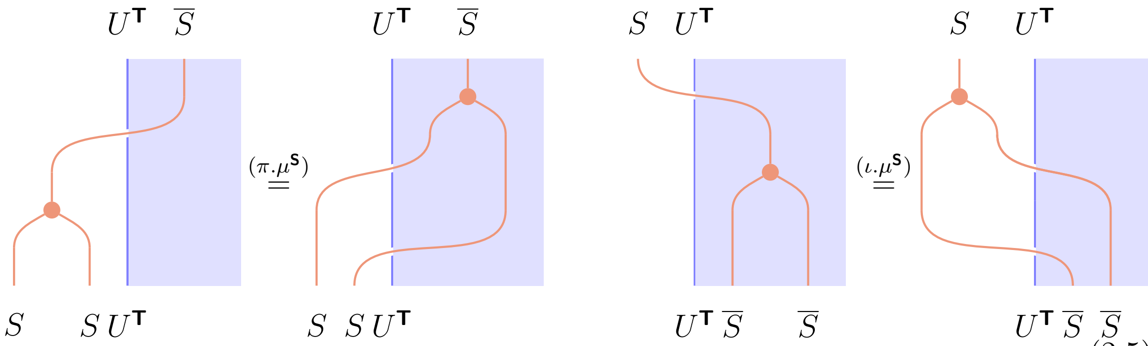
(2.3)

while compatibility conditions with the monad structure are, for the unit



(2.4)

and for the multiplication



(2.5)

To state the correspondence theorem, we will need the notion of split idempotent. An idempotent in C is an endomorphism $e : X \to X$ such that $e \circ e = e$. One says that the idempotent e *splits* (or that e is a split idempotent) when $e = i \circ p$ for some $p : X \to Y$, $i : Y \to X$ such that $p \circ i = 1$. In this situation, p is always an epimorphism and i is always a monomorphism.

Example 2.4. The category Set is *idempotent complete*, i.e., every idempotent splits. Let $e : X \to X$ be an idempotent function and $Y = \{x \in X \mid e(x) = x\}$ be the set of its fixpoints – or equivalently, the image of X under e. Then $p : X \twoheadrightarrow Y$ and $i : Y \hookrightarrow X$ defined by

$$p(x) = e(x) \qquad\qquad i(x) = x \tag{2.6}$$

exhibit e as a split idempotent.

The following theorem is due to Garner [54, Propositions 11 and 13].

Theorem 2.5 (Garner [54])**.** *There is a bijective correspondence between*

- *weak distributive laws of type* $\mathbf{TS} \to \mathbf{ST}$
- *weak extensions of* $\mathbf{T}$ *to* $\mathsf{Kl}(\mathbf{S})$
- *if idempotents split in* C*, weak liftings of* $\mathbf{S}$ *to* $\mathsf{EM}(\mathbf{T})$*.*

Proof. We give only the constructions. The bijection weak distributive laws $\iff$ weak extensions has the very same expression as in Theorem 1.42. We focus on the bijection weak distributive laws $\iff$ weak liftings. For a complete proof, see [54].

- weak liftings $\Rightarrow$ weak distributive laws

Let $\overline{\mathbf{S}}$ be a weak lifting of $\mathbf{S}$ to $\mathsf{EM}(\mathbf{T})$. Let $\pi : SU^{\mathbf{T}} \to U^{\mathbf{T}}\overline{S}$ and $\iota : U^{\mathbf{T}}\overline{S} \to SU^{\mathbf{T}}$ be the weak lifting natural transformations. Recall that for any $\mathbf{T}$-algebra (X, x), the C-morphism of the algebra $\overline{S}(X, x)$ is formally given by $U^{\mathbf{T}}\epsilon^{\mathbf{T}}_{\overline{S}F^{\mathbf{T}}X}$, where $\epsilon^{\mathbf{T}} : F^{\mathbf{T}}U^{\mathbf{T}} \to 1$ is the counit of the Eilenberg-Moore adjunction, defined by $\epsilon^{\mathbf{T}}_{(X,x)} = x : (TX, \mu^{\mathbf{T}}_X) \to (X, x)$. We define the weak distributive law to be

$$\lambda = TS \xrightarrow{TS\eta^{\mathbf{T}}} TSTX \xrightarrow{T\pi F^{\mathbf{T}}} TU^{\mathbf{T}}\overline{S}F^{\mathbf{T}} \xrightarrow{U^{\mathbf{T}}\epsilon^{\mathbf{T}}\overline{S}F^{\mathbf{T}}} U^{\mathbf{T}}\overline{S}F^{\mathbf{T}} \xrightarrow{\iota F^{\mathbf{T}}} ST \tag{2.7}$$

- weak distributive laws $\Rightarrow$ weak liftings

Let $\lambda : \mathbf{TS} \to \mathbf{ST}$ be a weak distributive law. For any $\mathbf{T}$-algebra (X, x), the pair $(SX, Sx \circ \lambda_X)$ is a $\mathbf{T}$-semialgebra, in the sense that it satisfies the associativity axiom but not necessarily the unit axiom. As idempotents split in C, they also split in the category of $\mathbf{T}$-semialgebras[1]. By splitting the idempotent

$$Sx \circ \lambda_X \circ \eta^{\mathbf{T}}_{SX} : (SX, Sx \circ \lambda_X) \to (SX, Sx \circ \lambda_X) \tag{2.8}$$

in the category of $\mathbf{T}$-semialgebras, we actually get a $\mathbf{T}$-algebra $\overline{S}(X, x)$:

$$(SX, Sx \circ \lambda_X) \xrightarrow{\pi_{(X,x)}} \overline{S}(X, x) \xrightarrow{\iota_{(X,x)}} (SX, Sx \circ \lambda_X) \tag{2.9}$$

and this defines simultaneously $\overline{S}$ on objects, ι and π. Note that the action of $\overline{S}(X, x)$ is given by $\pi_{(X,x)} \circ Sx \circ \lambda_X \circ T\iota_{(X,x)}$. On a morphism $f : (X, x) \to (Y, y)$, the functor $\overline{S}$ is defined by

$$\overline{S}f = \overline{S}(X, x) \xrightarrow{\iota_{(X,x)}} SX \xrightarrow{Sf} SY \xrightarrow{\pi_{(Y,y)}} \overline{S}(Y, y) \tag{2.10}$$

The other data of the weak lifting are forced to be $\eta^{\overline{\mathbf{S}}}_{(X,x)} = \pi_{(X,x)} \circ \eta^{\mathbf{S}}_X$ and $\mu^{\overline{\mathbf{S}}}_{(X,x)} = \pi_{(X,x)} \circ \mu^{\mathbf{S}}_X \circ S\iota_{(X,x)} \circ \iota_{\overline{S}(X,x)}$.

Moreover, any two splittings of the same idempotent are isomorphic. Under their identification, the constructions described above are inverse to each other. □

Example 2.6 ([54])**.** The following example of weak distributive law was the centerpiece of Garner's paper. Consider the ultrafilter monad $\boldsymbol{\beta}$ (Example 1.12) and the powerset monad $\mathbf{P}$ (Example 1.4) on Set. There is a weak distributive law $\boldsymbol{\beta}\mathbf{P} \to \mathbf{P}\boldsymbol{\beta}$, defined for all $\mathfrak{v} \in \beta PX$ by

$$\lambda_X(\mathfrak{v}) = \left\{\mathcal{U} \in \beta X \mid \bigcup \mathcal{V} \in \mathcal{U} \text{ for all } \mathcal{V} \in \mathfrak{v}\right\} \tag{2.11}$$

Interestingly, the corresponding weak lifting is the Vietoris monad on the Eilenberg-Moore category of $\boldsymbol{\beta}$, that is, on KHaus.

[1]Implicitly used by Garner [54, Lemma 12], this fact is proved by an easy diagram chase.

In the rest of this section, we assume that idempotents split in C, so that weak distributive laws, weak extensions and weak liftings are equivalent data. From a weak distributive law $\lambda : \mathbf{TS} \to \mathbf{ST}$, there is still a possibility to recover a composite monad on C. Indeed, we can compose the Eilenberg-Moore adjunctions of $\mathbf{T}$ and $\overline{\mathbf{S}}$:

$$\begin{tikzcd} \mathsf{EM}(\overline{\mathbf{S}}) \arrow[r, shift right=1.5ex, "U^{\overline{\mathbf{S}}}"'] \arrow[r, phantom, "\perp"] & \mathsf{EM}(\mathbf{T}) \arrow[l, shift right=1.5ex, "F^{\overline{\mathbf{S}}}"'] \arrow[r, shift right=1.5ex, "U^{\mathbf{T}}"'] \arrow[r, phantom, "\perp"] & \mathsf{C} \arrow[l, shift right=1.5ex, "F^{\mathbf{T}}"'] \end{tikzcd} \tag{2.12}$$

The functor of this new monad is $U^{\mathbf{T}}\overline{S}F^{\mathbf{T}}$ and can be obtained by splitting the idempotent $S\mu^{\mathbf{T}} \circ \lambda T \circ \eta^{\mathbf{T}}ST : ST \to ST$. Let us briefly explain the origin of this idempotent. Recall that a $\mathbf{T}$-semialgebra is a pair (X, x) such that $x : TX \to X$ satisfies $x \circ Tx = x \circ \mu^{\mathbf{T}}_X$. Although nothing is said about $\eta^{\mathbf{T}}_X$, the composition $x \circ \eta^{\mathbf{T}}_X$ in fact always is an idempotent because

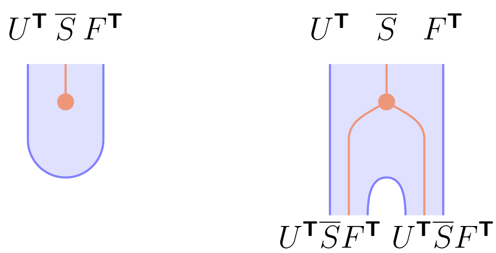
$$x \circ \eta^{\mathbf{T}}_X \circ x \circ \eta^{\mathbf{T}}_X = x \circ Tx \circ \eta^{\mathbf{T}}_{TX} \circ \eta^{\mathbf{T}}_X = x \circ \mu^{\mathbf{T}}_X \circ \eta^{\mathbf{T}}_{TX} \circ \eta^{\mathbf{T}}_X = x \circ \eta^{\mathbf{T}}_X \tag{2.13}$$

Axioms of weak distributive laws are sufficient to entail that $(STX, S\mu^{\mathbf{T}}_X \circ \lambda_{TX})$ is a $\mathbf{T}$-semialgebra. Therefore, $S\mu^{\mathbf{T}}_X \circ \lambda_{TX} \circ \eta^{\mathbf{T}}_{STX}$ is an idempotent so one can split pointwise the natural transformation $S\mu^{\mathbf{T}} \circ \lambda T \circ \eta^{\mathbf{T}}ST : ST \to ST$ to obtain a functor *close to* ST which is $U^{\mathbf{T}}\overline{S}F^{\mathbf{T}}$. If λ is not a distributive law, this means that the functor of the composite monad is *not exactly* the composition of the functors of the monads, but rather a weak version of it (see Chapter 3 for non-trivial examples). The composition of $\mathbf{S}$ and $\mathbf{T}$ via a weak distributive law λ will be denoted by $\mathbf{S} \bullet_\lambda \mathbf{T}$, or just $\mathbf{S} \bullet \mathbf{T}$ if the context is clear. Note that if λ is a distributive law, then $\mathbf{S} \bullet_\lambda \mathbf{T} = \mathbf{S} \circ_\lambda \mathbf{T}$.

Explicitly, the full *weak composite monad* is given by

$$\mathbf{S} \bullet \mathbf{T} = (U^{\mathbf{T}}\overline{S}F^{\mathbf{T}}, U^{\mathbf{T}}\eta^{\overline{\mathbf{S}}}F^{\mathbf{T}} \circ \eta^{\mathbf{T}}, U^{\mathbf{T}}\mu^{\overline{\mathbf{S}}}F^{\mathbf{T}} \circ U^{\mathbf{T}}\overline{S}\epsilon^{\mathbf{T}}\overline{S}F^{\mathbf{T}}) \tag{2.14}$$

where we recall that $\epsilon^{\mathbf{T}} : F^{\mathbf{T}}U^{\mathbf{T}} \to 1$ is the counit of the Eilenberg-Moore adjunction. In string diagrams, the unit and the multiplication are respectively:



(2.15)

The following proposition means that any $(\mathbf{S} \bullet \mathbf{T})$-algebra also carries both a $\mathbf{T}$-algebra and a $\mathbf{S}$-algebra structure.

Proposition 2.7. *The natural transformations*

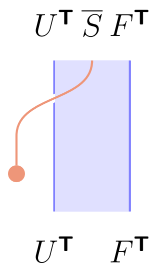
$$T \xrightarrow{\eta^{\mathsf{S}} T} ST \xrightarrow{\pi F^{\mathsf{T}}} S \bullet T$$

$$S \xrightarrow{S\eta^{\mathsf{T}}} ST \xrightarrow{\pi F^{\mathsf{T}}} S \bullet T$$

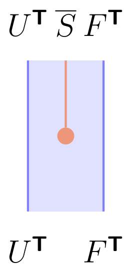
define monad morphisms

$$\pi F^{\mathsf{T}} \circ \eta^{\mathsf{S}} T : \mathbf{T} \to \mathbf{S} \bullet \mathbf{T}$$
$$\pi F^{\mathsf{T}} \circ S\eta^{\mathsf{T}} : \mathbf{S} \to \mathbf{S} \bullet \mathbf{T}$$

Proof. The first one, $\pi F^{\mathsf{T}} \circ \eta^{\mathsf{S}} T$, is depicted as

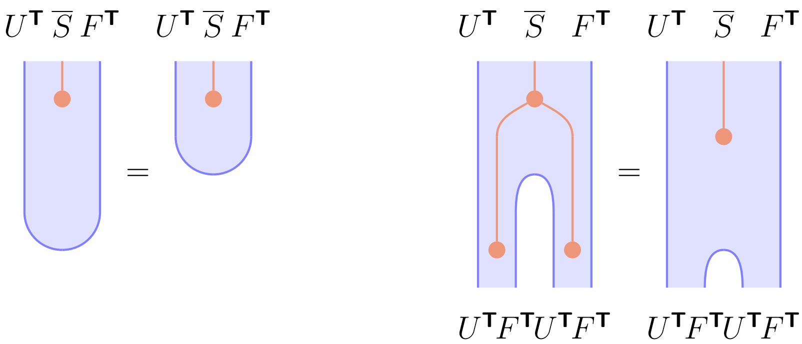


which is equal by $(\pi.\eta^{\mathsf{S}})$ to



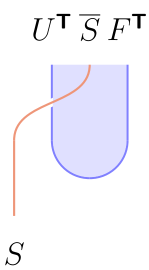
(2.16)

The unit diagram of monad morphisms is trivial and the multiplication diagram commutes because of $\mu^{\overline{\mathsf{S}}}$ associativity.



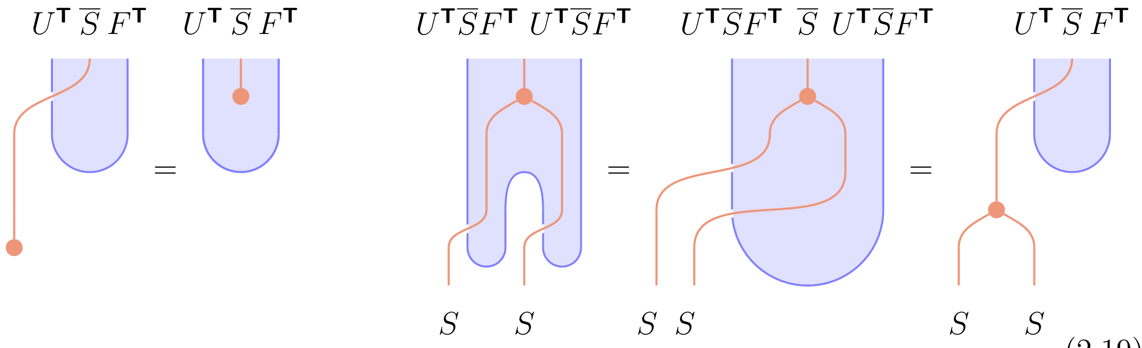
(2.17)

The second one, $\pi F^{\mathsf{T}} \circ S\eta^{\mathsf{T}}$, is depicted as



(2.18)

The unit diagram commutes by the $(\pi.\eta^{\mathsf{S}})$ axiom and the multiplication diagram commutes by the adjunction property $U^{\mathbf{T}}\epsilon^{\mathbf{T}} \circ \eta^{\mathbf{T}}U^{\mathbf{T}} = 1$ and the $(\pi.\mu^{\mathbf{S}})$ axiom.



(2.19)

$\square$

Additionally, the category $\mathsf{Alg}(\lambda)$ can be defined exactly like it was for plain distributive laws. Objects are triples (X, τ, σ) with (X, τ) being a $\mathbf{T}$-algebra and σ being a $\mathbf{S}$-algebra such that the $\sigma \circ S\tau \circ \lambda = \tau \circ T\sigma$, as in the pentagon (1.38). Morphisms are the simultaneous $\mathbf{T}$-algebra and $\mathbf{S}$-algebra morphisms. Proposition 1.46 adapts to the weak framework as follows:

Proposition 2.8 ([54])**.** *There are isomorphisms* $\mathsf{EM}(\mathbf{S} \bullet \mathbf{T}) \cong \mathsf{EM}(\overline{\mathbf{S}}) \cong \mathsf{Alg}(\lambda)$.

Example 2.9 ([54])**.** The weak composite monad on Set corresponding to the weak distributive law $\lambda : \boldsymbol{\beta}\mathbf{P} \to \mathbf{P}\boldsymbol{\beta}$ of Example 2.6 is the filter monad: $\mathbf{P} \bullet_{\lambda} \boldsymbol{\beta} = \mathbf{F}$. The corresponding weak lifting being the Vietoris monad, the identification of $\mathbf{V}$-algebras as continuous lattices $\mathsf{EM}(\mathbf{F}) \cong \mathsf{EM}(\mathbf{V})$ witnessed in Example 1.28 therefore is an instance of Proposition 2.8.

To end this section, we provide a couple of formulas that will be used later to compute the composite monad $\mathbf{S} \bullet \mathbf{T}$ in concrete cases.

Lemma 2.10. *Let* $\mathbf{S} \bullet_{\lambda} \mathbf{T} = (S \bullet T, \eta^{\mathbf{S}\bullet\mathbf{T}}, \mu^{\mathbf{S}\bullet\mathbf{T}})$ *be the composite monad on* C *with respect to a weak distributive law* $\lambda : \mathbf{TS} \to \mathbf{ST}$*. Let* $f : X \to Y$ *be a morphism in* C*. The following equations hold:*

1. $\iota_{F^{\mathbf{T}}Y} \circ (S \bullet T)f = STf \circ \iota_{F^{\mathbf{T}}X}$
2. $\iota F^{\mathbf{T}} \circ \eta^{\mathbf{S}\bullet\mathbf{T}} = \eta^{\mathbf{S}}\eta^{\mathbf{T}}$
3. $\iota F^{\mathbf{T}} \circ \mu^{\mathbf{S}\bullet\mathbf{T}} = \mu^{\mathbf{S}}\mu^{\mathbf{T}} \circ S\lambda T \circ \iota F^{\mathbf{T}}\iota F^{\mathbf{T}}$

Proof. Recall that by the adjunction composition described in equation (2.12) we have the expressions

$$\mathbf{S} \bullet \mathbf{T} = (U^{\mathbf{T}}\overline{S}F^{\mathbf{T}}, U^{\mathbf{T}}\eta^{\overline{\mathbf{S}}}F^{\mathbf{T}} \circ \eta^{\mathbf{T}}, U^{\mathbf{T}}\mu^{\overline{\mathbf{S}}}F^{\mathbf{T}} \circ U^{\mathbf{T}}\overline{S}\epsilon^{\mathbf{T}}\overline{S}F^{\mathbf{T}}) \tag{2.20}$$

1. Naturality of ι for the algebra morphism $F^{\mathsf{T}}f$ directly yields

$$\iota_{F^{\mathsf{T}}Y} \circ U^{\mathsf{T}}\overline{S}F^{\mathsf{T}}f = SU^{\mathsf{T}}F^{\mathsf{T}}f \circ \iota_{F^{\mathsf{T}}X}$$

2. The $(\iota.\eta^{\mathsf{S}})$ diagram precomposed with F^{T} gives

$$\iota F^{\mathsf{T}} \circ U^{\mathsf{T}}\eta^{\overline{\mathsf{S}}}F^{\mathsf{T}} = \eta^{\mathsf{S}}U^{\mathsf{T}}F^{\mathsf{T}}$$

   We get the wanted expression by composing with η^{T} on the right.

3. Let us compute

$$\begin{aligned}
&\iota F^{\mathsf{T}} \circ U^{\mathsf{T}}\mu^{\overline{\mathsf{S}}}F^{\mathsf{T}} \circ U^{\mathsf{T}}\overline{S}\epsilon^{\mathsf{T}}\overline{S}F^{\mathsf{T}} && \\
&= \mu^{\mathsf{S}}U^{\mathsf{T}}F^{\mathsf{T}} \circ S\iota F^{\mathsf{T}} \circ \iota\overline{S}F^{\mathsf{T}} \circ U^{\mathsf{T}}\overline{S}\epsilon^{\mathsf{T}}\overline{S}F^{\mathsf{T}} && (\iota.\mu^{\mathsf{S}}) \text{ diagram} \\
&= \mu^{\mathsf{S}}T \circ S\iota F^{\mathsf{T}} \circ SU^{\mathsf{T}}\epsilon^{\mathsf{T}}\overline{S}F^{\mathsf{T}} \circ \iota F^{\mathsf{T}}U^{\mathsf{T}}\overline{S}F^{\mathsf{T}} && \iota \text{ naturality} \\
&= \mu^{\mathsf{S}}T \circ SS\mu^{\mathsf{T}} \circ S\lambda T \circ ST\iota F^{\mathsf{T}} \circ \iota F^{\mathsf{T}}U^{\mathsf{T}}\overline{S}F^{\mathsf{T}} && (*) \\
&= \mu^{\mathsf{S}}\mu^{\mathsf{T}} \circ S\lambda T \circ \iota F^{\mathsf{T}}\iota F^{\mathsf{T}} &&
\end{aligned}$$

   The transformation performed in $(*)$ is correct because for every object X, $\iota_{F^{\mathsf{T}}X} : (U^{\mathsf{T}}\overline{S}F^{\mathsf{T}}X, U^{\mathsf{T}}\epsilon^{\mathsf{T}}\overline{S}F^{\mathsf{T}}X) \to (STX, S\mu^{\mathsf{T}}_X \circ \lambda_{TX})$ is a semialgebra morphism. $\square$

2.2 Finding Weak Distributive Laws

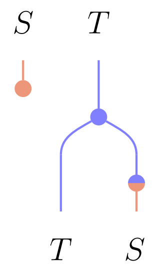
Finding distributive laws is notoriously known as being difficult, with a series of papers from Manes and Mulry devoted to the question [99, 100, 101] including two erroneous claims, corrected in [91, 162]. There is an additional difficulty when it comes to finding weak distributive laws: there are many trivial ones. We will prove here that any monad morphism yields a weak distributive law.

Theorem 2.11. *Let $\gamma : \mathsf{S} \to \mathsf{T}$ be a monad morphism. Then*

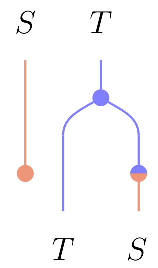
$$TS \xrightarrow{T\gamma} TT \xrightarrow{\mu^{\mathsf{T}}} T \xrightarrow{\eta^{\mathsf{S}}T} ST \tag{2.21}$$

is a weak distributive law. It is a distributive law if and only if $\eta^{\mathsf{S}}\gamma = S\eta^{\mathsf{T}}$.

Proof. We will prove this with string diagrams. The natural transformation defined in equation (2.21) is represented by

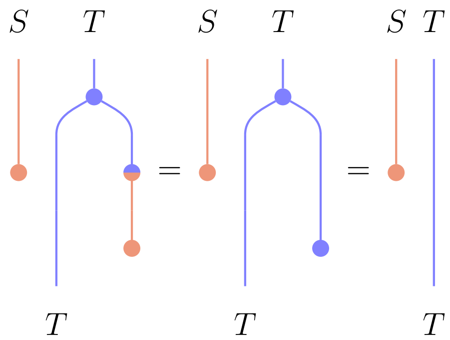


or, equivalently, by



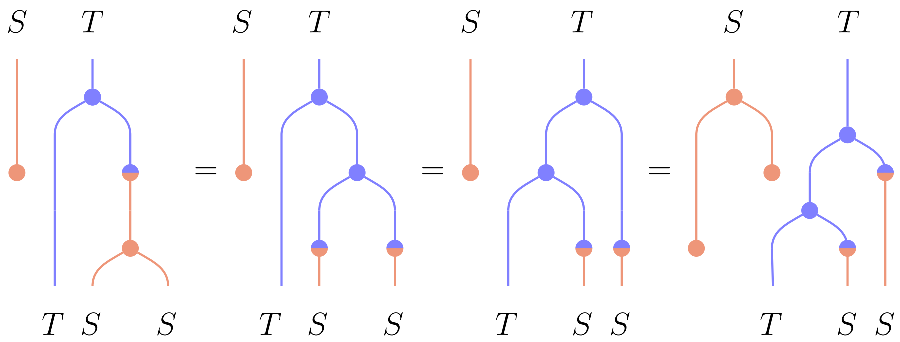
(2.22)

The η^{S} equation holds, using equation (1.11) of monad morphisms and the unit axiom of T:



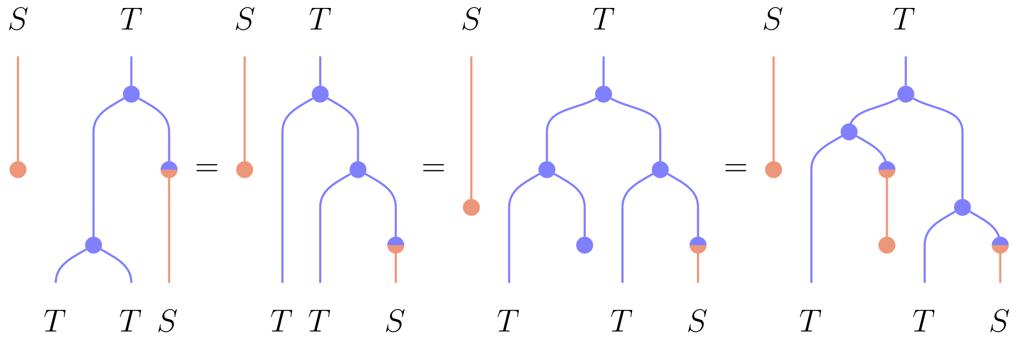
(2.23)

The μ^{S} equation holds, using equation (1.12) of monad morphisms and the associativity axiom of T and the unit axiom of S:



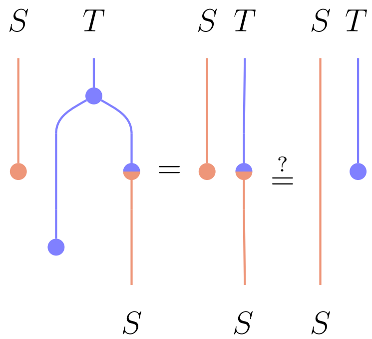
(2.24)

The μ^{T} equation holds, using the associativity axiom of T, the unit axiom of T and equation (1.11).



(2.25)

So all axioms of weak distributive laws are satisfied. The last axiom needed to get a distributive law amounts to checking:



(2.26)

and the $\stackrel{?}{=}$ equality holds if and only if $\eta^{\mathbf{S}}\gamma = S\eta^{\mathbf{T}}$. This completes the proof. $\square$

In particular, for any monad $\mathbf{T}$, the identity monad morphism $\mathrm{id}_T : \mathbf{T} \to \mathbf{T}$ yields a weak distributive law $\eta^{\mathbf{T}}T \circ \mu^{\mathbf{T}} : \mathbf{TT} \to \mathbf{TT}$. We retrieve Zwart's result in [161, Theorem 4.3] stating that $\eta^{\mathbf{T}}T \circ \mu^{\mathbf{T}}$ is a distributive law exactly when $\eta^{\mathbf{T}}T = T\eta^{\mathbf{T}}$, that is, when the monad $\mathbf{T}$ is idempotent.

It is interesting to ask what is the corresponding weak lifting of such a weak distributive law $\lambda = \eta^{\mathbf{S}}T \circ \mu^{\mathbf{T}} \circ T\gamma$.

Proposition 2.12. *Let* $\mathbf{S}$, $\mathbf{T}$ *be monads on a category* C *and* $\lambda = \eta^{\mathbf{S}}\mu^{\mathbf{T}} \circ T\gamma$ *for some monad morphism* $\gamma : \mathbf{S} \to \mathbf{T}$. *There is a weak lifting* $\overline{\mathbf{S}}$ *corresponding to* λ, *moreover* $\overline{\mathbf{S}} = \mathbf{Id}$ *and* $\mathbf{S} \bullet \mathbf{T} = \mathbf{T}$.

Proof. Recall that the possible weak lifting $\overline{S} : \mathsf{EM}(\mathbf{T}) \to \mathsf{EM}(\mathbf{T})$ is obtained on a $\mathbf{T}$-algebra (X, x) by splitting the idempotent $Sx \circ \lambda_X \circ \eta^{\mathbf{T}}_{SX}$, which is in this case

$$\begin{aligned}
Sx \circ \lambda_X \circ \eta^{\mathbf{T}}_{SX} &= Sx \circ \eta^{\mathbf{S}}_{TX} \circ \mu^{\mathbf{T}}_X \circ T\gamma_X \circ \eta^{\mathbf{T}}_{SX} && \lambda \text{ definition} \\
&= \eta^{\mathbf{S}}_X \circ x \circ \mu^{\mathbf{T}}_X \circ \eta^{\mathbf{T}}_{TX} \circ \gamma_X && \eta \text{ naturality} \\
&= \eta^{\mathbf{S}}_X \circ x \circ \gamma_X && \mathbf{T} \text{ monad}
\end{aligned}$$

But note that $x \circ \gamma_X \circ \eta^{\mathbf{S}}_X = x \circ \eta^{\mathbf{T}}_X = \mathrm{id}_X$ because of monad morphism properties and $\mathbf{T}$-algebras properties. The idempotent $Sx \circ \lambda_X \circ \eta^{\mathbf{T}}_{SX}$ is therefore split. Notably, even if not all idempotents split in C, λ has a corresponding weak lifting $\overline{\mathbf{S}}$ with $\iota_{(X,x)} = \eta^{\mathbf{S}}_X$ and $\pi_{(X,x)} = x \circ \gamma_X$. The carrier of $\overline{S}(X, x)$ is X. Using constructions weak distributive laws $\Rightarrow$ weak liftings from Theorem 2.5, we then have

data	from generic formula	after simplifications
action of $\overline{S}(X, x)$	$(x \circ \gamma_X) \circ Sx \circ \lambda_X \circ T\eta^{\mathbf{S}}_X$	x
$\overline{S}f$ where $f : (X, x) \to (Y, y)$	$(y \circ \gamma_Y) \circ Sf \circ \eta^{\mathbf{S}}_X$	f
$\eta^{\overline{\mathbf{S}}}_{(X,x)}$	$(x \circ \gamma_X) \circ \eta^{\mathbf{S}}_X$	id_X
$\mu^{\overline{\mathbf{S}}}_{(X,x)}$	$(x \circ \gamma_X) \circ \mu^{\mathbf{S}}_X \circ S\eta^{\mathbf{S}}_X \circ \eta^{\mathbf{S}}_X$	id_X

Consequently $\overline{\mathbf{S}}$ is the identity monad $\mathbf{Id}$ on $\mathsf{EM}(\mathbf{T})$. It follows that $\mathbf{S} \bullet \mathbf{T} = U^{\mathbf{T}}\overline{\mathbf{S}}F^{\mathbf{T}} \cong U^{\mathbf{T}}\mathbf{Id}F^{\mathbf{T}} = \mathbf{T}$. $\square$

Proposition 2.12 justifies our terminology of *trivial* weak distributive laws. Here are a few examples of them.

Example 2.13. Using the support monad morphism $\mathsf{supp} : \mathbf{D} \to \mathbf{P}$ from Example 1.16, the natural transformation $\lambda : PD \to DP$ defined by $\lambda_X(U) = 1 \cdot \bigcup_{\varphi \in U} \mathsf{supp}(\varphi)$ is a weak distributive law.

Example 2.14. Using $\mathrm{id}_P : \mathbf{P} \to \mathbf{P}$, the natural transformation $\lambda : PP \to PP$ defined by $\lambda_X(\mathcal{U}) = \{\bigcup \mathcal{U}\}$ is a weak distributive law. The functor of the corresponding weak extension of $\mathbf{P}$ to Rel maps a relation R to its graph $\underline{P}R = \{(U, R[U]) \mid U \in PX\}$. Note that $\underline{P}$ is not monotone with respect to relation inclusion.

Example 2.15. Using $\mathrm{id}_D : \mathbf{D} \to \mathbf{D}$, the natural transformation $\lambda : DD \to DD$ defined by $\lambda_X(\Phi) = 1 \cdot \left(\sum_{x \in X} \left(\sum_{\varphi \in DX} \Phi_\varphi \varphi_x\right) \cdot x\right)$ is a weak distributive law.

Example 2.16. One cannot obtain a weak distributive law of type $\mathbf{DP} \to \mathbf{PD}$ using Theorem 2.11, because there is no monad morphism $\mathbf{P} \to \mathbf{D}$. Indeed, $P0 = 1$, $D0 = 0$ and there is no map $1 \to 0$ in Set.

As there are no distributive laws of type $\mathbf{PP} \to \mathbf{PP}$ [91], $\mathbf{DD} \to \mathbf{DD}$ and $\mathbf{PD} \to \mathbf{DP}$ [161], neither of Examples 2.14, 2.15 and 2.13 satisfies the last axiom. More interestingly, these examples show how weak distributive laws obtained via Theorem 2.11 act: they basically collapse all algebraic structure related to $\mathbf{S}$ by using the transformation γ in concert with the multiplication, then use the unit to output data with a trivial $\mathbf{S}$-structure. We would like to find weak distributive laws having a less simplistic behaviour and better benefitting from the full structure of the monads $\mathbf{S}$ and $\mathbf{T}$ in the sense that their weak lifting is *not* the identity monad. To this end, we use a mechanism for obtaining well-behaved laws, consisting in looking only for those laws of type $\mathbf{TS} \to \mathbf{ST}$ that preserve the additional structure that the Kleisli category of $\mathbf{S}$ may possess. The most usual case is to take $\mathsf{C} = \mathsf{Set}$, $\mathbf{S} = \mathbf{P}$ and identify the Kleisli category of $\mathbf{P}$ with the category of sets and relations Rel. In Rel, every hom-set is a poset with respect to relation inclusion. In this setting, laws may or may not interact nicely with the order structure. For instance, the trivial $\mathbf{PP} \to \mathbf{PP}$ law $\mathcal{U} \mapsto \{\bigcup \mathcal{U}\}$ does not preserve the order structure, because its weak extension is not monotone with respect to relation inclusion. Restricting the search scope to *monotone* laws is a classical idea relying on results from Barr [3], used by many authors (see [93] for a survey) and reintroduced by Garner [54] in the theory of weak distributive laws.

A Pos-enriched category is a category C such that every homset is a poset and composition respects the order: for every f, g, h, h' with appropriate domains and codomains

$$\text{if } f \leq g \text{ then } h \circ f \leq h \circ g \text{ and } f \circ h' \leq g \circ h' \tag{2.27}$$

In this context, a functor $F : \mathsf{C} \to \mathsf{C}$ is *locally monotone* if $f \leq g \Rightarrow Ff \leq Fg$.

Definition 2.17 (Monotone law)**.** Assume that $\mathsf{Kl}(\mathbf{S})$ is a Pos-enriched category. A (weak) distributive law $\lambda : \mathbf{TS} \to \mathbf{ST}$ is *monotone* if the functor of its corresponding (weak) extension is locally monotone.

Using the definition of the extension functor corresponding to a weak distributive law (see the proof of Theorem 1.42), we emphasise that λ is monotone when for any $f, g : X \to SY$, if $f \le g$ then $\lambda_Y \circ Tf \le \lambda_Y \circ Tg$. If we take C to be Set and $\mathbf{S}$ to be the powerset monad $\mathbf{P}$, $\mathsf{Kl}(\mathbf{P})$ is the category Rel of sets and relations. Relation inclusion makes Rel a Pos-enriched category.

Remark 2.18. In this chapter, we only consider the Pos-enriched category Rel. Other instances will arise in Chapters 5, 6 and 7.

There is a valuable existence characterisation for monotone laws $\mathbf{TP} \to \mathbf{PT}$, due to Barr for distributive laws [3] and adapted by Garner [54] for weak distributive laws. The key categorical notion to state this result is the one of weak pullbacks, which arises naturally when considering composition of relations seen as spans. In Chapter 5, a more general account of relational composition will be given in these terms.

In a category C, the pullback of two morphisms $f : X \to Z$ and $g : Y \to Z$ is, whenever it exists, the limit of the following cospan.

$$\begin{tikzcd} & Y \arrow[d, "g"] \\ X \arrow[r, "f"'] & Z \end{tikzcd} \tag{2.28}$$

More concretely, the pullback of f and g is an object P along with two morphisms $p_1 : P \to X$ and $p_2 : P \to Y$ such that

- $f \circ p_1 = g \circ p_2$
- for every Q, $q_1 : Q \to X$, $q_2 : Q \to Y$ such that $f \circ q_1 = g \circ q_2$, there is a unique $h : Q \to P$ such that $q_1 = p_1 \circ h$ and $q_2 = p_2 \circ h$

The pullback situation is summed up in the following diagram:

$$\begin{tikzcd} Q \arrow[dr, dashed, "\exists! h"] \arrow[drr, "q_2"] \arrow[ddr, "q_1"'] & & \\ & P \arrow[r, "p_2"'] \arrow[d, "p_1"] & Y \arrow[d, "g"] \\ & X \arrow[r, "f"'] & Z \end{tikzcd} \tag{2.29}$$

A *weak pullback* of f and g is a weak limit of the their cospan: the concrete definition is the same as for the pullback, but the uniqueness requirement for h is dropped. Whereas pullbacks are unique up to isomorphism, there may be many different non-isomorphic weak pullbacks of a same cospan. The following notation denote that P is the pullback, respectively W is a weak pullback, of f and g.

$$\begin{tikzcd} P \arrow[r, "p_2"] \arrow[d, "p_1"'] \arrow[dr, phantom, very near start, "\lrcorner"] & Y \arrow[d, "g"] \\ X \arrow[r, "f"'] & Z \end{tikzcd} \qquad \begin{tikzcd} W \arrow[r, "w_2"] \arrow[d, "w_1"'] \arrow[dr, phantom, very near start, "\cdot\lrcorner"] & Y \arrow[d, "g"] \\ X \arrow[r, "f"'] & Z \end{tikzcd} \tag{2.30}$$

Definition 2.19 (Weakly cartesian functor)**.** A functor $F : \mathsf{C} \to \mathsf{C}$ is *weakly cartesian* if it preserves weak pullbacks i.e. for any weak pullback

$$\begin{tikzcd} W \arrow[r, "w_2"] \arrow[d, "w_1"'] \arrow[dr, phantom, very near start, "\cdot\lrcorner"] & Y \arrow[d, "g"] \\ X \arrow[r, "f"'] & Z \end{tikzcd} \tag{2.31}$$

the following square is a weak pullback

$$\begin{tikzcd} FW \arrow[r, "Fw_2"] \arrow[d, "Fw_1"'] \arrow[dr, phantom, very near start, "\cdot\lrcorner"] & FY \arrow[d, "Fg"] \\ FX \arrow[r, "Ff"'] & FZ \end{tikzcd} \tag{2.32}$$

Definition 2.20 (Weakly cartesian natural transformation)**.** A natural transformation $\alpha : F \to G$ is weakly cartesian if its naturality squares are weak pullbacks i.e. for any $f : X \to Y$ the following square is a weak pullback

$$\begin{tikzcd} FX \arrow[r, "\alpha_X"] \arrow[d, "Ff"'] \arrow[dr, phantom, very near start, "\cdot\lrcorner"] & GX \arrow[d, "Gf"] \\ FY \arrow[r, "\alpha_Y"'] & GY \end{tikzcd} \tag{2.33}$$

In the rest of this section we take C to be Set. The pullback of two functions $f : X \to Z$ and $g : Y \to Z$ always exists and is given by

$$\begin{aligned} &P = \{(x, y) \in X \times Y \mid f(x) = g(y)\} \\ &p_1(x, y) = x \\ &p_2(x, y) = y \end{aligned}$$

Using that pullbacks in Set have the above form, one can easily figure out that a weak pullback of $f : X \to Z$, $g : Y \to Z$ is just a set W and two functions $w_1 : W \to X$, $w_2 : W \to Y$ such that for every $(x, y) \in X \times Y$ satisfying $f(x) = g(y)$, there is a $t \in W$ such that $w_1(t) = x$ and $w_2(t) = y$.

Theorem 2.21 ([3, 93])**.** *Let F, $G : \mathsf{Set} \to \mathsf{Set}$ be functors and $\alpha : F \to G$ be a natural transformation.*

- *There is a locally monotone functor $\underline{F} : \mathsf{Rel} \to \mathsf{Rel}$ such that $F_{\mathbf{P}}F = \underline{F}F_{\mathbf{P}}$ if and only if F is weakly cartesian. In this case, such a locally monotone $\underline{F}$ is unique.*

- *Assume F and G satisfy the previous point. There is a natural transformation $\underline{\alpha} : \underline{F} \to \underline{G}$ such that $F_{\mathbf{P}}\alpha = \underline{\alpha}F_{\mathbf{P}}$ if and only if α is weakly cartesian. In this case, such an $\underline{\alpha}$ is unique.*

One can express $\underline{F}$ as follows. On any set X, $\underline{F}X = FX$. Given a relation $R \subseteq X \times Y$, let $\pi_1 : R \to X$ and $\pi_2 : R \to Y$ be the canonical projections, then

$$\underline{F}R = \{(u, v) \in FX \times FY \mid \exists t \in FR \,.\, F\pi_1(t) = u \text{ and } F\pi_2(t) = v\} \tag{2.34}$$

Note that $\underline{\alpha}$ is forced to be $\underline{\alpha}_X = \alpha_X$ – being weakly cartesian simply makes this a natural transformation in Rel.

As a consequence we have

Theorem 2.22 ([54])**.** *For every monad $\mathbf{T}$ on Set,*

- *there exists at most one monotone distributive law of type $\mathbf{TP} \to \mathbf{PT}$, and there is one if and only if T, $\eta^{\mathbf{T}}$ and $\mu^{\mathbf{T}}$ are weakly cartesian;*

- *there exists at most one monotone weak distributive law of type $\mathbf{TP} \to \mathbf{PT}$, and there is one if and only if T and $\mu^{\mathbf{T}}$ are weakly cartesian.*

Algebraically, Theorem 2.22 can be understood in light of a paper from Gautam [55] explaining how equations defined on a set can be lifted pointwise to its powerset.

Example 2.23. Distributive laws arising from Theorem 2.22 include the laws $\mathbf{LP} \to \mathbf{PL}$ and $\mathbf{MP} \to \mathbf{PM}$ from Examples 1.35 and 1.36, respectively.

As clearly presented in [36], frequently, the unit is not weakly cartesian, whereas the rest of the monad is. This is what happens for the ultrafilter monad.

Example 2.24 ([54])**.** The weak distributive law $\boldsymbol{\beta}\mathbf{P} \to \mathbf{P}\boldsymbol{\beta}$ of Example 2.6 is an instance of Theorem 2.22. Indeed, one can prove that β and $\mu^{\boldsymbol{\beta}}$ are weakly cartesian, but $\eta^{\boldsymbol{\beta}}$ is not. Consequently there is no monotone distributive law of type $\boldsymbol{\beta}\mathbf{P} \to \mathbf{P}\boldsymbol{\beta}$, but there is a unique monotone weak distributive law of this type. Computations lead to the expression of Example 2.6.

In the next chapter, we will use Theorem 2.22 to describe the unique monotone weak distributive laws of type $\mathbf{PP} \to \mathbf{PP}$ and $\mathbf{DP} \to \mathbf{PD}$.

2.3 Coweak Distributive Laws

Deleting the (η^{T}) axiom in the definition of weak distributive laws may seem arbitrary. Although the framework of distributive laws is not strictly symmetric in both monads, intuitively each monad is equally important. So a logical question that arises after discovering weak distributive laws is: what happens if we delete the (η^{S}) axiom instead? As this resembles a dual of a (Garner) weak distributive law, this notion will be called a *coweak distributive law*. This kind of weakening does not seem to be studied in the literature. To the understanding of the author, it cannot be derived as an instance of [54, 146, 15, 16]. Indeed, in these papers authors focus on weak distributive laws with the intention to weakly lift monads to Eilenberg-Moore categories, and we will see that this cannot be achieved if we decide to drop the (η^{S}) axiom instead of the (η^{T}) one. Instead, we will be able to coweakly extend monads to Kleisli categories.

Remark 2.25. There is a long tradition of looking at monads from a 2-categorical perspective [145, 94]. In this light, the fact that liftings to the Eilenberg-Moore category are equivalently extensions to the Kleisli category is a simple instance of categorical duality. See, for example, [125] for a 2-categorical study of distributive laws between a monad and a comonad exploiting duality. This standpoint has been also used to describe weak distributive laws, generalising the interplay between monads and comonads to the weak framework [15, 16]. From an abstract 2-categorical viewpoint, the results about coweak distributive laws we are about to describe are just *formally dual* to the results about weak distributive laws from the previous section.

We hereby define coweak distributive laws as expected:

Definition 2.26 (Coweak distributive law)**.** Let T, S be two monads on a category C. A *coweak distributive law* of type $\mathsf{TS} \to \mathsf{ST}$ is a natural transformation λ of this type such that the following equations hold.

$$\lambda \circ \eta^{\mathsf{T}} S = S\eta^{\mathsf{T}} \qquad (\eta^{\mathsf{T}} \text{ axiom})$$

$$\lambda \circ T\mu^{\mathsf{S}} = \mu^{\mathsf{S}} T \circ S\lambda \circ \lambda S \qquad (\mu^{\mathsf{S}} \text{ axiom})$$

$$\lambda \circ \mu^{\mathsf{T}} S = S\mu^{\mathsf{T}} \circ \lambda T \circ T\lambda \qquad (\mu^{\mathsf{T}} \text{ axiom})$$

By an informal analysis similar to the one of Section 2.1, we are led to the notions of coweak liftings and coweak extensions. Deleting the (η^{S}) axiom impacts Table 1.1 in two ways:

- One cannot prove that $\eta^{\overline{\mathsf{S}}}$ components are algebra morphisms. This can be easily patched by relaxing the statement that liftings comprise a unit. This leads directly to the definition of coweak liftings.

- One cannot prove that $\underline{T}$ preserves Kleisli identities – hence, it is not a functor. This is more of a concern, because this means that the construction of an extension collapses right from the beginning. Equivalently, $\lambda_X \circ T\eta^{\mathsf{S}}_X$ is not a Kleisli identity, but merely a Kleisli idempotent. Our strategy consists in splitting this Kleisli idempotent to recover a functor that will preserve Kleisli identities. This leads to the definition of coweak extensions.

Unsurprisingly, the definition of coweak extensions and coweak liftings is a dual version of weak extensions and weak liftings.

Definition 2.27 (Coweak extension)**.** A *coweak extension* of T to S is a monad $\underline{\mathsf{T}}$ on $\mathsf{Kl}(\mathsf{S})$ along with two natural transformations

$$\pi : F_{\mathsf{S}}T \nrightarrow \underline{T}F_{\mathsf{S}} \qquad\qquad \iota : \underline{T}F_{\mathsf{S}} \nrightarrow F_{\mathsf{S}}T$$

such that $\pi \bullet \iota = 1$ and the four following diagrams commute

$$\begin{tikzcd}
& F_{\mathsf{S}} \arrow[dl, "F_{\mathsf{S}}\eta^{\mathsf{T}}"', "|" marking] \arrow[dr, "\eta^{\underline{\mathsf{T}}}F_{\mathsf{S}}", "|" marking] \arrow[d, phantom, "(\pi.\eta^{\mathsf{T}})"] & \\
F_{\mathsf{S}}T \arrow[rr, "\pi"', "|" marking] & {} & \underline{T}F_{\mathsf{S}}
\end{tikzcd}
\qquad
\begin{tikzcd}
& F_{\mathsf{S}} \arrow[dl, "\eta^{\underline{\mathsf{T}}}F_{\mathsf{S}}"', "|" marking] \arrow[dr, "F_{\mathsf{S}}\eta^{\mathsf{T}}", "|" marking] \arrow[d, phantom, "(\iota.\eta^{\mathsf{T}})"] & \\
\underline{T}F_{\mathsf{S}} \arrow[rr, "\iota"', "|" marking] & {} & F_{\mathsf{S}}T
\end{tikzcd}$$

$$\begin{tikzcd}
F_{\mathsf{S}}TT \arrow[r, "\pi T", "|" marking] \arrow[d, "F_{\mathsf{S}}\mu^{\mathsf{T}}"', "|" marking] & \underline{T}F_{\mathsf{S}}T \arrow[r, "\underline{T}\pi", "|" marking] \arrow[d, phantom, "(\pi.\mu^{\mathsf{T}})"] & \underline{T}\underline{T}F_{\mathsf{S}} \arrow[d, "\mu^{\underline{\mathsf{T}}}F_{\mathsf{S}}", "|" marking] \\
F_{\mathsf{S}}T \arrow[rr, "\pi"', "|" marking] & & \underline{T}F_{\mathsf{S}}
\end{tikzcd}
\qquad
\begin{tikzcd}
\underline{T}\underline{T}F_{\mathsf{S}} \arrow[r, "\underline{T}\iota", "|" marking] \arrow[d, "\mu^{\underline{\mathsf{T}}}F_{\mathsf{S}}"', "|" marking] & \underline{T}F_{\mathsf{S}}T \arrow[r, "\iota T", "|" marking] \arrow[d, phantom, "(\iota.\mu^{\mathsf{T}})"] & F_{\mathsf{S}}TT \arrow[d, "F_{\mathsf{S}}\mu^{\mathsf{T}}", "|" marking] \\
\underline{T}F_{\mathsf{S}} \arrow[rr, "\iota"', "|" marking] & & F_{\mathsf{S}}T
\end{tikzcd}$$

Definition 2.28 (Coweak lifting)**.** A *coweak lifting* of S to $\mathsf{EM}(\mathsf{T})$ is a pair $(\overline{S}, \mu^{\overline{\mathsf{S}}})$ comprising a functor $\overline{S} : \mathsf{EM}(\mathsf{T}) \to \mathsf{EM}(\mathsf{T})$ and a natural transformation $\mu^{\overline{\mathsf{S}}} : \overline{S}\,\overline{S} \to \overline{S}$ such that $U^{\mathsf{T}}\overline{S} = SU^{\mathsf{T}}$ and $U^{\mathsf{T}}\mu^{\overline{\mathsf{S}}} = \mu^{\mathsf{S}}U^{\mathsf{T}}$.

Theorem 2.29. *There is a bijective correspondence between*

- *coweak distributive laws of type* $\mathsf{TS} \to \mathsf{ST}$

- *coweak liftings of* S *to* $\mathsf{EM}(\mathsf{T})$

- *if idempotents split in* $\mathsf{Kl}(\mathsf{S})$*, coweak extensions of* T *to* $\mathsf{Kl}(\mathsf{S})$*.*

Proof. The bijection coweak distributive laws $\iff$ coweak liftings is the same as in Theorem 1.42, dropping the unit $\eta^{\bar{\mathbf{S}}}$ everywhere, and with the same proofs – details are omitted. The bijection coweak distributive laws $\iff$ coweak extensions is more complicated. We give the explicit constructions herein; verifications are long and postponed to Appendix A. According to Remark 2.25, one could simply claim that the following constructions work by abstract duality with weak distributive laws. We find it nevertheless interesting to prove the result with full details: such proofs are rarely found in the literature, even since the original paper from Beck [7]. The duality-aware reader can consider the proof in Appendix A as a full version of Garner's [54, Proposition 13].

- coweak extensions $\Rightarrow$ coweak distributive laws

  Let $\underline{\mathbf{T}}$ be a coweak extension of $\mathbf{T}$ to $\mathsf{Kl}(\mathbf{S})$. Viewing $\mathrm{id}_{SX} : SX \to SX$ as a Kleisli morphism $\mathrm{id}_{SX} : SX \nrightarrow X$, we can define the composition

  $$F_{\mathbf{S}}TSX \xrightarrow{\pi_{SX}} \underline{T}F_{\mathbf{S}}SX = \underline{T}SX \xrightarrow{\underline{T}(\mathrm{id}_{SX})} \underline{T}X = \underline{T}F_{\mathbf{S}}X \xrightarrow{\iota_X} F_{\mathbf{S}}TX \tag{2.35}$$

  Recall that on objects, the free functor $F_{\mathbf{S}}$ acts identically. Therefore, the above morphism, seen in C again, has type $TSX \to STX$ and is defined to be the value of the coweak distributive law λ_X.

- coweak distributive laws $\Rightarrow$ coweak extensions

  Let $\lambda : \mathbf{TS} \to \mathbf{ST}$ be a coweak distributive law. From a Kleisli viewpoint, λ_X has type $TSX \nrightarrow TX$. For any $h : X \nrightarrow Y$ we define $h^+ = \lambda_Y \circ Th : TX \nrightarrow TY$. In particular, let $e_X = (\eta^{\mathbf{S}}_X)^+$. Then $e_X : TX \nrightarrow TX$ is a Kleisli idempotent. Splitting

  $$e_X = TX \xrightarrow{\pi_X} \underline{T}X \xrightarrow{\iota_X} TX \tag{2.36}$$

  yields the wanted natural transformations $\pi : F_{\mathbf{S}}T \nrightarrow \underline{T}F_{\mathbf{S}}$ and $\iota : \underline{T}F_{\mathbf{S}} \nrightarrow F_{\mathbf{S}}T$ and defines $\underline{T}$ on objects. On a morphism $h : X \nrightarrow Y$, define

  $$\underline{T}h = \underline{T}X \xrightarrow{\iota_X} TX \xrightarrow{h^+} TY \xrightarrow{\pi_Y} \underline{T}Y \tag{2.37}$$

  The unit and multiplication are forced to be

  $$\eta^{\underline{\mathbf{T}}}_X = \pi_X \bullet F_{\mathbf{S}}\eta^{\mathbf{T}}_X \tag{2.38}$$
  $$\mu^{\underline{\mathbf{T}}}_X = \pi_X \bullet (F_{\mathbf{S}}\mu^{\mathbf{T}}_X) \bullet \iota_{TX} \bullet \underline{T}(\iota_X) \tag{2.39}$$

Under the identification of isomorphic idempotent splittings, the constructions described above are inverse to each other. $\square$

As for the weak framework, the main purpose of the coweak framework is to get a composite monad built from $\mathbf{S}$ and $\mathbf{T}$. This can be achieved by composing Kleisli adjunctions of the coweak extension:

$$\mathsf{Kl}(\underline{\mathbf{T}}) \underset{U_{\underline{\mathbf{T}}}}{\overset{F_{\underline{\mathbf{T}}}}{\underset{\longrightarrow}{\overset{\longleftarrow}{\perp}}}} \mathsf{Kl}(\mathbf{S}) \underset{U_{\mathbf{S}}}{\overset{F^{\mathbf{S}}}{\underset{\longrightarrow}{\overset{\longleftarrow}{\perp}}}} \mathsf{C} \tag{2.40}$$

The resulting coweak composite monad is denoted by $\mathbf{S} \bullet_\lambda \mathbf{T}$, or just $\mathbf{S} \bullet \mathbf{T}$, and satisfies $\mathsf{Kl}(\mathbf{S} \bullet \mathbf{T}) \cong \mathsf{Kl}(\underline{\mathbf{T}})$.

We now aim at finding examples of coweak distributive laws. A first observation is that there is an obvious dual of Theorem 2.11, that one can prove explicitly by using string diagrams.

Theorem 2.30. *Let $\gamma : \mathbf{T} \to \mathbf{S}$ be a monad morphism. Then*

$$TS \xrightarrow{\gamma S} SS \xrightarrow{\mu^{\mathbf{S}}} S \xrightarrow{S\eta^{\mathbf{T}}} ST \tag{2.41}$$

is a coweak distributive law. It is a distributive law if and only if $\gamma\eta^{\mathbf{T}} = \eta^{\mathbf{S}}T$.

A similar result is present in the literature: [81, Lemma 8] taken with a trivial alphabet means that the natural transformation of Theorem 2.30 satisfies axioms $(\eta^{\mathbf{T}})$ and $(\mu^{\mathbf{T}})$. As we will see later, it is rather rare that idempotents split in a Kleisli category, but in the case of trivial laws the coweak extension always exists: the dual of Proposition 2.12 is

Proposition 2.31. *Let $\mathbf{S}$, $\mathbf{T}$ be monads on a category C and $\lambda = \mu^{\mathbf{S}}\eta^{\mathbf{T}} \circ \gamma S$ for some monad morphism $\gamma : \mathbf{T} \to \mathbf{S}$. There is a coweak extension $\underline{\mathbf{T}}$ corresponding to λ, moreover $\underline{\mathbf{T}} = \mathbf{Id}$ and $\mathbf{S} \bullet \mathbf{T} = \mathbf{S}$.*

Proof. One can check that the Kleisli idempotent $e_X = \lambda_X \circ T\eta^{\mathbf{S}}_X$ is equal to $F_{\mathbf{S}}\eta^{\mathbf{T}}_X \bullet \gamma_X$, where γ_X is seen as a Kleisli morphism of type $TX \nrightarrow X$, and that $\gamma_X \bullet F_{\mathbf{S}}\eta^{\mathbf{T}}_X$ is a Kleisli identity. Then there is a coweak extension with $\underline{T}X = X$, $\pi_X = \gamma_X$, $\iota_X = F_{\mathbf{S}}\eta^{\mathbf{T}}_X$ and one can check using the explicit formulas that $\underline{T}f = f$ and $\eta^{\underline{\mathbf{T}}}$, $\mu^{\underline{\mathbf{T}}}$ both are identity natural transformation in $\mathsf{Kl}(\mathbf{S})$. Then $\underline{\mathbf{T}} = \mathbf{Id}$ and consequently $\mathbf{S} \bullet \mathbf{T} = U_{\mathbf{S}}\mathbf{Id}F_{\mathbf{S}} = \mathbf{S}$. $\square$

For a coweak distributive law $\lambda = S\eta^{\mathsf{T}} \circ \mu^{\mathsf{S}} \circ \gamma S$, the functor of the corresponding coweak lifting maps a T-algebra (X, x) to a T-algebra on SX whose action is given by:

$$Sx \circ \lambda_X = Sx \circ S\eta_X^{\mathsf{T}} \circ \mu_X^{\mathsf{S}} \circ \gamma_{SX} = \mu_X^{\mathsf{S}} \circ \gamma_{SX} \tag{2.42}$$

This time, the coweak lifting is trivial in the sense that it does not even involve the action x. In the case where $\gamma = \mathrm{id}_{\mathsf{T}}$, note that we have $\overline{T} = F^{\mathsf{T}} U^{\mathsf{T}}$.

Example 2.32. Using $\mathsf{supp} : \mathsf{D} \to \mathsf{P}$ as a monad morphism, authors of [81] derive what turns out to be a coweak distributive law of type $\mathsf{DP} \to \mathsf{PD}$ with expression

$$\lambda_X(\Phi) = \Big\{1 \cdot x \mid x \in \bigcup \mathsf{supp}(\Phi)\Big\} \tag{2.43}$$

The functor of the corresponding coweak lifting maps a D-algebra (X, x) to

$$\Big(PX, \varphi \mapsto \bigcup \mathsf{supp}(\varphi)\Big) \tag{2.44}$$

Example 2.33. Another instance also present in [81] is the coweak distributive law of type $\mathsf{PP} \to \mathsf{PP}$ obtained via the identity monad morphism $\mathsf{P} \to \mathsf{P}$, which is given by

$$\lambda_X(\mathcal{U}) = \Big\{\{x\} \mid x \in \bigcup \mathcal{U}\Big\} \tag{2.45}$$

The functor of the corresponding coweak lifting maps a P-algebra (X, x) to (PX, μ_X^{P}).

Again in the coweak setting, trivial laws destroy too much structure to be interesting. Recall that in the case of weak distributive laws, the common strategy to get meaningful laws was the combination of two ideas:

- Use $\mathsf{S} = \mathsf{P}$, the powerset monad, so that $\mathsf{Kl}(\mathsf{S}) = \mathsf{Rel}$ has enjoyable properties: it is Pos-enriched and monotone extensions of functors are easy to characterise.
- First look for a weak extension of T, then use the correspondence to derive the weak distributive law and the weak lifting.

In the setting of coweak distributive laws, these ideas are not relevant anymore:

- Idempotents do not always split in Rel. For example, $\{(0,0), (0,1), (1,1)\}$ is an idempotent that is not split. Therefore the existence of natural transformations ι and π is in general unlikely.
- Coweak extensions are the most intricated notion of the coweak framework, dually to weak liftings being the most intricated notion of the weak framework. Consequently it seems rather counterproductive to look for the coweak extension first.

Remark 2.34. The above reasoning does not strictly forbid the strategy to be carried out. We have seen in the proof of Theorem 2.29 that split idempotents are only needed for the construction coweak distributive law $\Rightarrow$ coweak extension. So if one manages to find a coweak extension in Rel, there are corresponding coweak distributive law and coweak lifting.

The question of how to find useful coweak distributive laws is still open. Dualising the common strategy, it is tempting to try and look for the coweak lifting first. The agenda is then to identify desirable properties of functors in $\mathsf{EM}(\mathbf{T})$, then restrict the search space to coweak liftings that interact nicely with these properties, and then provide abstract results to efficiently check these properties (for monotone extension to Rel, this was weak cartesianness). Of course, there may be other relevant strategies.

To fully exploit Theorem 2.29, we need a monad $\mathbf{S}$ whose Kleisli category is idempotent complete. Even if C is idempotent complete, there is no reason for $\mathsf{Kl}(\mathbf{S})$ to be, as shows the Rel example. Still assuming C is idempotent complete, then $\mathsf{EM}(\mathbf{S})$ is too, so a sufficient condition is $\mathsf{Kl}(\mathbf{S}) \equiv \mathsf{EM}(\mathbf{S})$ i.e. every $\mathbf{S}$-algebra is free. The following handy reformulation was pointed to me by Zhen Lin:

Fact 2.35. *For any monad $\mathbf{S}$ on an idempotent complete category, the Kleisli category of $\mathbf{S}$ is idempotent complete if and only if every retract of a free $\mathbf{S}$-algebra is free.*

All in all, here are some examples of Set monads whose Kleisli category is idempotent complete:

- The maybe monad $(- + \mathbf{1})$, because its Kleisli category and its Eilenberg-Moore category are equivalent.
- The k-vector space monad for any field k, because every vector space has a base i.e. every algebra is free.
- The Abelian group monad $\mathbf{A}$, because a retract of a free algebra is then called a regular-projective algebra, and it is known that every regular-projective Abelian group is free [2, Example 5.2.3].

2.4 Iterated Weak Distributive Laws

One can work at adapting Theorem 1.51 to iterated (co)weak distributive laws in various ways. For the sake of simplicity, we limit ourselves to $n = 3$ monads, so let R, S, T be monads on a category C. Trying to define a law

$$(\mathsf{S} \bullet \mathsf{T})\mathsf{R} \to \mathsf{R}(\mathsf{S} \bullet \mathsf{T})$$

using a weak distributive law $\mathsf{TS} \to \mathsf{ST}$ is tedious because the weak composite monad $\mathsf{S} \bullet \mathsf{T}$ forces the weak lifting natural transformations ι and π to appear everywhere. As we are heading towards concrete examples in Chapter 4, we therefore impose in the above case that the law $\mathsf{TS} \to \mathsf{ST}$ is always a plain distributive law. Under this restriction, there is a proper composite monad $\mathsf{S} \circ \mathsf{T}$ and we can easily define laws of type

$$(\mathsf{S} \circ \mathsf{T})\mathsf{R} \to \mathsf{R}(\mathsf{S} \circ \mathsf{T})$$

using the same formula as in the proof of Theorem 1.51. Now, let us inventory which axioms are used in every part of the proof.

Table 2.1: Axioms required for an iterated distributive law. To prove the axiom of the first column, one needs the hypotheses of the four last columns. Y-B denotes the need for the Yang-Baxter equation.

$\delta : (\mathsf{S} \circ_\lambda \mathsf{T})\mathsf{R} \to \mathsf{R}(\mathsf{S} \circ_\lambda \mathsf{T})$	$\lambda : \mathsf{TS} \to \mathsf{ST}$	$\sigma : \mathsf{SR} \to \mathsf{RS}$	$\tau : \mathsf{TR} \to \mathsf{RT}$	
(η^{R})	plain	(η^{R})	(η^{R})	
(μ^{R})	plain	(μ^{R})	(μ^{R})	
$(\eta^{\mathsf{S}\circ\mathsf{T}})$	plain	(η^{S})	(η^{T})	
$(\mu^{\mathsf{S}\circ\mathsf{T}})$	plain	(μ^{S})	(μ^{T})	Y-B
$\delta' : \mathsf{T}(\mathsf{R} \circ_\sigma \mathsf{S}) \to (\mathsf{R} \circ_\sigma \mathsf{S})\mathsf{T}$	$\lambda : \mathsf{TS} \to \mathsf{ST}$	$\sigma : \mathsf{SR} \to \mathsf{RS}$	$\tau : \mathsf{TR} \to \mathsf{RT}$	
(η^{T})	(η^{T})	plain	(η^{T})	
(μ^{T})	(μ^{T})	plain	(μ^{T})	
$(\eta^{\mathsf{R}\circ\mathsf{S}})$	(η^{S})	plain	(η^{R})	
$(\mu^{\mathsf{R}\circ\mathsf{S}})$	(μ^{S})	plain	(μ^{R})	Y-B

From Table 2.1 we can easily deduce when (co)weakening some of the distributive laws λ, σ and τ produces (co)weak distributive laws δ and δ'. Results are displayed in Table 2.2. Note that the Yang-Baxter condition is required in any case, because it is involved in multiplication axioms, which are required in each sort of law.

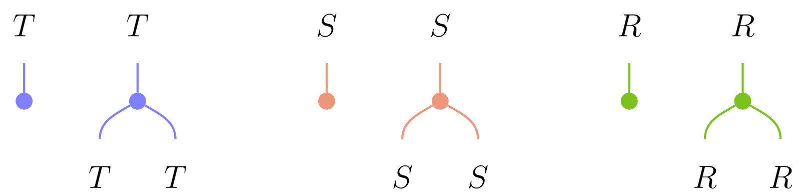
To build easily some concrete examples of such iterated laws, we provide two technical lemmas. Note that there possibly are a plethora of similar constructions – we provide two that are particularly well-suited for the examples we have in mind.

Table 2.2: Iterated (co)weak distributive laws. Given data in the first four columns, one can obtain a law as in the last column

$\lambda : \mathbf{TS} \to \mathbf{ST}$	$\sigma : \mathbf{SR} \to \mathbf{RS}$	$\tau : \mathbf{TR} \to \mathbf{RT}$		$\delta : (\mathbf{S} \circ_\lambda \mathbf{T})\mathbf{R} \to \mathbf{R}(\mathbf{S} \circ_\lambda \mathbf{T})$
plain	plain	plain	Y-B	plain
plain	weak	weak	Y-B	weak
plain	coweak	coweak	Y-B	coweak
$\lambda : \mathbf{TS} \to \mathbf{ST}$	$\sigma : \mathbf{SR} \to \mathbf{RS}$	$\tau : \mathbf{TR} \to \mathbf{RT}$		$\delta' : \mathbf{T}(\mathbf{R} \circ_\sigma \mathbf{S}) \to (\mathbf{R} \circ_\sigma \mathbf{S})\mathbf{T}$
plain	plain	plain	Y-B	plain
weak	plain	weak	Y-B	weak
coweak	plain	coweak	Y-B	coweak

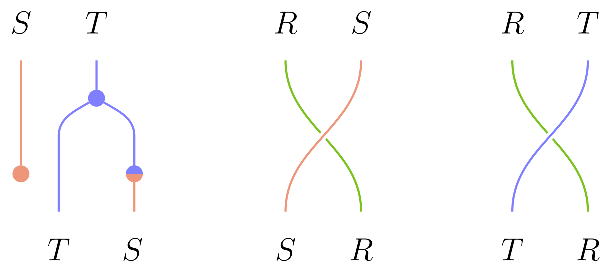
Lemma 2.36. *Let $\sigma : \mathbf{SR} \to \mathbf{RS}$ be a distributive law, $\tau : \mathbf{TR} \to \mathbf{RT}$ be a weak distributive law, and $\gamma : \mathbf{S} \to \mathbf{T}$ be a monad morphism such that $R\gamma \circ \sigma = \tau \circ \gamma R$. Let $\lambda : \mathbf{TS} \to \mathbf{ST}$ be the weak distributive law defined by $\lambda = \eta^{\mathbf{S}} \mu^{\mathbf{T}} \circ T\gamma$. Then λ, σ and τ satisfy the Yang-Baxter condition, so that they yield a weak distributive law $\mathbf{T}(\mathbf{R} \circ_\sigma \mathbf{S}) \to (\mathbf{R} \circ_\sigma \mathbf{S})\mathbf{T}$.*

Proof. We use string diagrams. We have monads



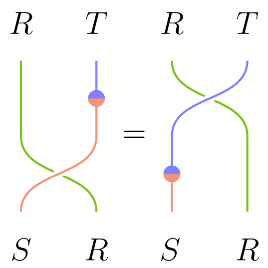
(2.46)

and the respectively weak, plain and weak distributive laws λ, σ and τ



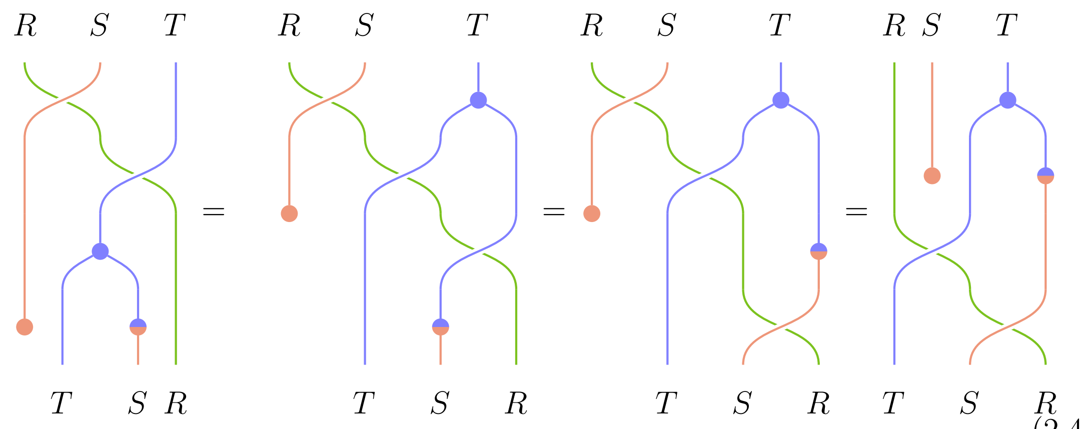
(2.47)

The assumption $R\gamma \circ \sigma = \tau \circ \gamma R$ amounts to



(2.48)

The Yang-Baxter equation is then derivable using the (μ^{T}) axiom of τ, the assumption $R\gamma \circ \sigma = \tau \circ \gamma R$ and the (η^{S}) axiom of σ.



(2.49)

$\square$

Lemma 2.37. *Let $\sigma : \mathsf{SR} \to \mathsf{RS}$ be a distributive law, $\tau : \mathsf{TR} \to \mathsf{RT}$ be a coweak distributive law and $\gamma : \mathsf{T} \to \mathsf{S}$ be a monad morphism such that $R\gamma \circ \tau = \sigma \circ \gamma R$. Let $\lambda : \mathsf{TS} \to \mathsf{ST}$ be the coweak distributive law defined by $\lambda = \mu^{\mathsf{S}}\eta^{\mathsf{T}} \circ \gamma S$. Then λ, σ and τ satisfy the Yang-Baxter condition, so that they yield a coweak distributive law $\mathsf{T}(\mathsf{R} \circ_\sigma \mathsf{S}) \to (\mathsf{R} \circ_\sigma \mathsf{S})\mathsf{T}$.*

Proof. This can be proved as for Lemma 2.36. Properties used are the (μ^{S}) axiom of σ, the assumption $R\gamma \circ \tau = \sigma \circ \gamma R$ and the (η^{T}) axiom of τ. $\square$

We give a few examples involving the following laws in Set between the powerset monad P, the distribution monad D and the reader monad R:

$\lambda^1 : \mathsf{PR} \to \mathsf{RP}$	$\lambda^1_X(U) = \lambda a.\{h(a) \mid h \in U\}$	plain	Example 1.33
$\lambda^2 : \mathsf{DR} \to \mathsf{RD}$	$\lambda^2_X(\varphi) = \lambda a. \sum_{h \in RX} \varphi_h \cdot h(a)$	plain	Example 1.34
$\lambda^3 : \mathsf{PD} \to \mathsf{DP}$	$\lambda^3_X(U) = 1 \cdot \bigcup_{\varphi \in U} \mathsf{supp}(\varphi)$	weak	Example 2.13
$\lambda^4 : \mathsf{PP} \to \mathsf{PP}$	$\lambda^4_X(\mathcal{U}) = \left\{\bigcup \mathcal{U}\right\}$	weak	Example 2.14
$\lambda^5 : \mathsf{DP} \to \mathsf{PD}$	$\lambda^5_X(\Phi) = \left\{1 \cdot x \mid x \in \bigcup \mathsf{supp}(\Phi)\right\}$	coweak	Example 2.32
$\lambda^6 : \mathsf{PP} \to \mathsf{PP}$	$\lambda^6_X(\mathcal{U}) = \left\{\{x\} \mid x \in \bigcup \mathcal{U}\right\}$	coweak	Example 2.33

Example 2.38. Applying Lemma 2.36 with $\sigma = \tau = \lambda^1$, and $\gamma = \mathrm{id}_P$ hence $\lambda = \lambda^4$ requires the identity $R\,\mathrm{id}\,P \circ \lambda^1 = \lambda^1 \circ \mathrm{id}\,RP$, which is trivial. Then, there is a weak distributive law $\mathsf{P}(\mathsf{R} \circ_{\lambda^1} \mathsf{P}) \to (\mathsf{R} \circ_{\lambda^1} \mathsf{P})\mathsf{P}$ defined by

$$\mathcal{U} \in P((PX)^A) \mapsto \lambda a. \left\{ \bigcup_{h \in \mathcal{U}} h(a) \right\} \in (PPX)^A \tag{2.50}$$

Example 2.39. Using the same data as in Example 2.38, but this time applying Lemma 2.37, we obtain a coweak distributive law $\mathbf{P}(\mathbf{R} \circ_{\lambda^1} \mathbf{P}) \to (\mathbf{R} \circ_{\lambda^1} \mathbf{P})\mathbf{P}$ defined by

$$\mathcal{U} \in P((PX)^A) \mapsto \lambda a. \left\{ \{x\} \mid x \in \bigcup_{h \in \mathcal{U}} h(a) \right\} \in (PPX)^A \tag{2.51}$$

Example 2.40. Applying Lemma 2.36 with $\sigma = \lambda^2$, $\tau = \lambda^1$, and $\gamma = \mathsf{supp}$ hence $\lambda = \lambda^3$ requires the identity $(R\mathsf{supp}) \circ \lambda^2 = \lambda^1 \circ (\mathsf{supp}R)$ which can be easily verified. Then, there is a weak distributive law $\mathbf{P}(\mathbf{R} \circ_{\lambda^2} \mathbf{D}) \to (\mathbf{R} \circ_{\lambda^2} \mathbf{D})\mathbf{P}$ defined by

$$\mathcal{U} \in P((DX)^A) \mapsto \lambda a. \left(1 \cdot \bigcup_{h \in \mathcal{U}} \mathsf{supp}(h(a)) \right) \in (DPX)^A \tag{2.52}$$

Example 2.41. Applying Lemma 2.37 with $\sigma = \lambda^1$, $\tau = \lambda^2$ and $\gamma = \mathsf{supp}$ hence $\lambda = \lambda^5$ requires the identity $(R\mathsf{supp}) \circ \lambda^2 = \lambda^1 \circ (\mathsf{supp}R)$, as in Example 2.39. Then, there is a coweak distributive law $\mathbf{D}(\mathbf{R} \circ_{\lambda^1} \mathbf{P}) \to (\mathbf{R} \circ_{\lambda^1} \mathbf{P})\mathbf{D}$ defined by

$$\Phi \in D((PX)^A) \mapsto \lambda a. \left\{ 1 \cdot x \mid x \in \bigcup_{h \in \mathsf{supp}\Phi} h(a) \right\} \in (PDX)^A \tag{2.53}$$

Chapter 3

Combining Probability and Non-Determinism

There are no distributive laws of type $\mathsf{PP} \to \mathsf{PP}$ and $\mathsf{DP} \to \mathsf{PD}$, where we recall that P denotes the powerset monad and D denotes the distribution monad. These two non-examples, both occuring in the category of sets and functions, do have a specific status in the history of the theory of distributive laws.

The non-existence of a distributive law of type $\mathsf{DP} \to \mathsf{PD}$ was first proved by Varacca and Winskel [152, 153] using an idea of Gordon Plotkin. Given a candidate distributive law, the proof consists in keeping track of computations in both paths of some conveniently chosen naturality diagrams. Using three such naturality diagrams and only the unit axioms, a contradiction is obtained. This result is the first concrete no-go theorem concerning distributive laws, and relates to a long history of difficulties when trying to combine non-determinism (P) and probability (D) in domain theory [88, 148, 108, 152, 153] or, more recently, in coalgebraic semantics of Segala systems [23, 25]. The solution proposed by Varacca is to replace D with a monad of *indexed valuations* [151].

The non-existence of a distributive law of type $\mathsf{PP} \to \mathsf{PP}$ is a more recent result. At first, a few transformations were wrongly identified as distributive laws in the literature, first by Manes and Mulry [99], then by Klin and Rot [89, 90]. Klin and Salamanca [91] settled the question by proving that there is no such law, with a method close to the Plotkin argument of [153]. This second negative result relates to the fact that working with two layers of non-determinism often resulted in the use of workarounds – e.g. replacing one powerset monad by a list monad to study coalgebraic semantics of logic programming [27], or replacing powersets by downsets and upsets to study coalgebraic semantics of alternating automata [8].

As it turns out, these two non-examples were the first of a long list. Zwart and Marsden [162] recently produced abstract theorems that automatically forbid distributive laws to exist in Set. Their approach is algebraic: every composite monad corresponds to a composite algebraic theory, and they show that under certain conditions, there is no composite algebraic theory. The results of Zwart and Marsden recover the previous ones about $\mathbf{PP} \to \mathbf{PP}$ and $\mathbf{DP} \to \mathbf{PD}$ and generate new concrete non-examples, e.g. there are no distributive laws of type $\mathbf{PD} \to \mathbf{DP}$ and $\mathbf{DD} \to \mathbf{DD}$. Zwart and Marsden notably imported from physics the terminology 'no-go theorem' to designate this kind of impossibility result. The PhD thesis of Maaike Zwart [161] provides a precise and systematic study of no-go theorems, along with many examples and non-examples using familiar Set monads. Another no-go theorem not falling into the scope of Zwart and Marsden results is the one of Salamanca [135] showing that there is no distributive law $\mathbf{LaP} \to \mathbf{PLa}$, where $\mathbf{La}$ is the free lattice monad.

In this chapter, we show how the concept of weak distributive law helps to mitigate the absence of distributive laws for the two now-classical instances $\mathbf{PP} \to \mathbf{PP}$ and $\mathbf{DP} \to \mathbf{PD}$. Although there are no such distributive laws, there are weak ones, and even *monotone* weak ones. In a first section, we begin by recalling the results of Garner [54] about the unique monotone weak distributive law $\mathbf{PP} \to \mathbf{PP}$. This law will be the basis of several subsequent developments – an application to alternating automata in Chapter 4 and generalisations to toposes and compact Hausdorff spaces in Chapters 6 and 7. In a second section, we shall provide original results about the unique monotone weak distributive law $\mathbf{DP} \to \mathbf{PD}$. These results, finally explaining how to properly combine non-determinism and probability in a non *ad hoc* way, have been published in [62]. In a third section, we will briefly discuss another interesting case: $\mathbf{PD} \to \mathbf{DP}$. Note that one cannot rule out weak distributive laws of this type using monotonicity criteria, as in Theorem 2.22, because $\mathsf{Kl}(\mathbf{D})$ does not have a relevant Pos-enriched structure. However, we will find on the way another property that is arguably desirable for such a law. Unfortunately, we will prove that there is no law satisfying this new property. The question of whether there is a meaningful weak distributive law of type $\mathbf{PD} \to \mathbf{DP}$ therefore remains open.

A further point is that in both $\mathbf{PP} \to \mathbf{PP}$ and $\mathbf{DP} \to \mathbf{PD}$ cases, the impossibility result is actually stronger: there is no *monad structure* on the functors PP [91] and PD [38]. Obtaining composite monads via weak distributive laws does not clash with these results, since in the weak framework the functor of the composite monad is *not* the composition of the monad functors.

3.1 Powerset over Powerset

3.1.1 The Monotone Law

There is a trivial weak distributive law of type $\mathsf{PP} \to \mathsf{PP}$ obtained in Example 2.14 and defined by

$$\lambda_X(\mathcal{U}) = \left\{ \bigcup \mathcal{U} \right\}$$

In this section we will use Theorem 2.22 to prove that there is also a unique monotone weak distributive law of this type. In particular, note that the $\mathsf{PP} \to \mathsf{PP}$ case witnesses that there can be multiple weak distributive laws, even when there is no distributive law.

Frobenius reciprocity law. The following elementary property will be implicitly used on several occasions. For any function $f : X \to Y$ and subsets $U \subseteq X$, $V \subseteq Y$,

$$f(U \cap f^{-1}(V)) = f(U) \cap V \tag{3.1}$$

We now prove basic properties of the powerset monad, which are all well-known results, see e.g. [149] for the following proposition.

Proposition 3.1 ([149])**.** *The powerset functor P is weakly cartesian.*

Proof. Let W be a weak pullback of $f : X \to Z$ and $g : Y \to Z$.

$$\begin{tikzcd} W \arrow[r, "w_2"] \arrow[d, "w_1"'] \arrow[dr, phantom, "\lrcorner" very near start] & Y \arrow[d, "g"] \\ X \arrow[r, "f"'] & Z \end{tikzcd} \tag{3.2}$$

Equivalently, this means that $w_1 \circ w_2^{-1} = f^{-1} \circ g : PY \to PX$. Let us show that the image of this square under P still is a weak pullback.

$$\begin{tikzcd} PW \arrow[r, "Pw_2"] \arrow[d, "Pw_1"'] & PY \arrow[d, "Pg"] \\ PX \arrow[r, "Pf"'] & PZ \end{tikzcd} \tag{3.3}$$

Take $U \subseteq X$, $V \subseteq Y$ and assume $f(U) = g(V)$. We must find some $C \subseteq W$ such that $w_1(C) = U$ and $w_2(C) = V$. Define

$$C = w_1^{-1}(U) \cap w_2^{-1}(V)$$

Obviously C is included in W. Moreover, $f(U) = g(V)$ entails $U \subseteq (f^{-1} \circ g)(V) = (w_1 \circ w_2^{-1})(V)$. Hence

$$w_1(C) = w_1(w_1^{-1}(U) \cap w_2^{-1}(V)) = U \cap w_1(w_2^{-1}(V)) = U \tag{3.4}$$

Similarly, $w_2(C) = V$, so that W is a weak pullback and P is weakly cartesian. $\square$

Proposition 3.2 ([54]). *The unit η^{P} is not weakly cartesian.*

Proof. Consider the naturality square of the unique map $! : \{0,1\} \to \{0\}$.

$$\begin{array}{ccc} \{0,1\} & \xrightarrow{\eta^{\mathsf{P}}_{\{0,1\}}} & P\{0,1\} \\ {\scriptstyle !}\big\downarrow & & \big\downarrow{\scriptstyle P!} \\ \{0\} & \xrightarrow[\eta^{\mathsf{P}}_{\{0\}}]{} & P\{0\} \end{array} \tag{3.5}$$

The pullback of $\eta^{\mathsf{P}}_{\{0\}}$ and $P!$ is $\{(0,\{0\}),(0,\{1\}),(0,\{0,1\})\}$. If the above square were a weak pullback, there would be an element of $\{0,1\}$ mapped to the pullback element $(0,\{0,1\})$ by $\langle !, \eta^{\mathsf{P}}_{\{0,1\}} \rangle$. This is impossible because $\{0,1\}$, as a non-singleton, is not in the image of $\eta^{\mathsf{P}}_{\{0,1\}}$. $\square$

Proposition 3.3 ([54]). *The multiplication μ^{P} is weakly cartesian.*

Proof. Let $f : X \to Y$ be a function and let us show that its naturality square is a weak pullback.

$$\begin{array}{ccc} PPX & \xrightarrow{\mu^{\mathsf{P}}_X} & PX \\ {\scriptstyle PPf}\big\downarrow & & \big\downarrow{\scriptstyle Pf} \\ PPY & \xrightarrow[\mu^{\mathsf{P}}_Y]{} & PY \end{array} \tag{3.6}$$

Let $U \subseteq X$, $\mathcal{V} \subseteq PY$ such that $f(U) = \bigcup \mathcal{V}$. We must find $\mathcal{U} \subseteq PX$ such that $\bigcup \mathcal{U} = U$ and $Pf(\mathcal{U}) = \mathcal{V}$. Define

$$\mathcal{U} = \{U \cap f^{-1}(V) \mid V \in \mathcal{V}\}$$

Then, using standard properties of set-theoretic operations and the hypothesis $f(U) = \bigcup \mathcal{V}$ we obtain:

$$PPf(\mathcal{U}) = \{f(U \cap f^{-1}(V)) \mid V \in \mathcal{V}\} = \{f(U) \cap V \mid V \in \mathcal{V}\} = \{V \mid V \in \mathcal{V}\} = \mathcal{V}$$
$$\bigcup \mathcal{U} = U \cap f^{-1}\left(\bigcup \mathcal{V}\right) = U \cap f^{-1}(f(U)) = U$$
$\square$

By applying Theorem 2.22 we obtain that there is a unique monotone weak distributive law of type $\mathsf{PP} \to \mathsf{PP}$. Let us first get an expression of the weak extension using equation (2.34), then use the correspondence theorem to compute the weak distributive law and the weak lifting. On a relation $R \subseteq X \times Y$ with projections $\pi_1 : R \to X$, $\pi_2 : R \to Y$, the formula

$$\underline{P}R = \{(U,V) \in PX \times PY \mid \exists C \subseteq R.\pi_1(C) = U \text{ and } \pi_2(C) = V\} \tag{3.7}$$

can be more conveniently expressed as

$$\begin{aligned}\underline{P}R = \{(U,V) \in PX \times PY \mid & \forall x \in U.\exists y \in V.(x,y) \in R \text{ and} \\ & \forall y \in V.\exists x \in U.(x,y) \in R\}\end{aligned}$$

which is known, especially when R is a preorder, under the name *Egli-Milner extension* of R. According to the correspondence of Theorem 2.5, the monotone weak distributive law at X is the function λ_X whose graph is $\underline{P}(\ni_X)$, where we recall that $\ni_X$ is the set $\{(U,x) \mid x \in U\} \subseteq PX \times X$. Therefore

Theorem 3.4 ([54])**.** *The unique monotone weak distributive law* $\lambda : \mathsf{PP} \to \mathsf{PP}$ *is given on every* $\mathcal{U} \in PPX$ *by the expression*

$$\lambda_X(\mathcal{U}) = \left\{V \in PX \mid V \subseteq \bigcup \mathcal{U} \textit{ and } \forall U \in \mathcal{U}.U \cap V \neq \emptyset\right\} \tag{3.8}$$

We now compute the corresponding weak lifting accordingly to the procedure described in Theorem 1.42. Let (X, x) be an Eilenberg-Moore algebra for the monad P, that is, a complete join semi-lattice. As λ is a weak distributive law, $Px \circ \lambda_X$ is a P-semialgebra. Hence, precomposing with η^{P} yields an idempotent

$$Px \circ \lambda_X \circ \eta^{\mathsf{P}}_{PX} : (PX, Px \circ \lambda_X) \to (PX, Px \circ \lambda_X) \tag{3.9}$$

mapping a subset $U \subseteq X$ to

$$(Px \circ \lambda_X \circ \eta^{\mathsf{P}}_{PX})(U) = \{x(V) \mid V \in PU \setminus \{\emptyset\}\} \tag{3.10}$$

We may now split the idempotent $Px \circ \lambda_X \circ \eta^{\mathsf{P}}_{PX}$ in the category of P-semialgebras as follows:

$$(PX, Px \circ \lambda_X) \xrightarrow{\pi_{(X,x)}} (S, s) \xrightarrow{\iota_{(X,x)}} (PX, Px \circ \lambda_X) \tag{3.11}$$

By the standard epi-mono factorisation in Set, the carrier S can be computed as the image of the idempotent, that is, the set of its fixpoints.

Definition 3.5 (Up closure, upclosed subset)**.** Let (X, x) be a complete semi-lattice. For any $U \subseteq X$, the set

$$\mathsf{up}_x(U) = \{x(V) \mid V \in PU \setminus \{\emptyset\}\}$$

is called the *x-up closure* of U. It contains all x-joins of *non-empty* subsets of U. We say that U is an *x-upclosed subset* whenever $U = \mathsf{up}_x(U)$ i.e. U is closed under non-empty x-joins.

Therefore, P maps (X, x) to the subalgebra obtained by restriction of $Px \circ \lambda_X$ to x-upclosed subsets.

$$\overline{P}(X, x) = (S, s) \tag{3.12}$$

$$S = \{U \in PX \mid U = \mathsf{up}_x(U)\} \tag{3.13}$$

$$\pi_{(X,x)}(U) = \mathsf{up}_x(U) \tag{3.14}$$

$$\iota_{(X,x)}(U) = U \tag{3.15}$$

The following lemma provides an expression of the action of the $\mathbf{P}$-algebra (S, s).

Lemma 3.6. *For every $\mathcal{U} \in PS$,*

$$s(\mathcal{U}) = \{x(\mathsf{Im}c) \mid c : \mathcal{U} \to X \text{ choice function}\} \tag{3.16}$$

By choice function we mean that for every $U \in \mathcal{U}$, $c(U) \in U$, and Im denotes the image i.e. $\mathsf{Im}c = \{c(U) \mid U \in \mathcal{U}\}$.

Proof. By the splitting in the category of $\mathbf{P}$-semialgebras, we know that s is the restriction of $Px \circ \lambda_X$ to the set of x-upclosed subsets of X:

$$\begin{tikzcd}
PPX \arrow[r, "P\pi_{(X,x)}"] \arrow[d, "\lambda_X"'] & PS \arrow[r, "P\iota_{(X,x)}"] \arrow[dd, "s"'] & PPX \arrow[d, "\lambda_X"] \\
PPX \arrow[d, "Px"'] & & PPX \arrow[d, "Px"] \\
PX \arrow[r, "\pi_{(X,x)}"'] & S \arrow[r, "\iota_{(X,x)}"'] & PX
\end{tikzcd} \tag{3.17}$$

Hence, on every $\mathcal{U} \in PS$ we have

$$s(\mathcal{U}) = \left\{x(V) \mid V \subseteq \bigcup \mathcal{U} \text{ and } \forall U \in \mathcal{U}. U \cap V \neq \emptyset\right\}$$

Now, for any choice function $c : \mathcal{U} \to X$, $\mathsf{Im}c$ is included in $\bigcup \mathcal{U}$ and intersects every $U \in \mathcal{U}$. For the converse inclusion, let $V \subseteq \bigcup \mathcal{U}$ such that V intersects every $U \in \mathcal{U}$.

Then one can define a choice function $c : \mathcal{U} \to X$ by $U \mapsto x(U \cap V)$. Indeed, for any $U \in \mathcal{U}$, $U \cap V$ is a non-empty subset of the x-upclosed subset U, therefore $x(U \cap V) \in U$. To conclude we remark that $x(\mathsf{Im} c) = x(V)$ as follows:

$$\begin{aligned} x(\mathsf{Im} c) &= x(\{x(U \cap V) \mid U \in \mathcal{U}\}) \\ &= (x \circ Px)(\{U \cap V \mid U \in \mathcal{U}\}) \\ &= (x \circ \mu^{\mathbf{P}}_X)(\{U \cap V \mid U \in \mathcal{U}\}) \\ &= x\left(\bigcup_{U \in \mathcal{U}} U \cap V\right) = x(V) \text{ because } V \subseteq \bigcup \mathcal{U} \end{aligned}$$ $\square$

Let $f : (X, x) \to (Y, y)$ be a morphism of complete join semi-lattices. Use the naturality diagram of ι to immediately get, for all x-upclosed subset U:

$$\overline{P} f(U) = Pf(U) \tag{3.18}$$

To compute $\eta^{\overline{\mathbf{P}}}$ and $\mu^{\overline{\mathbf{P}}}$, we shall again take advantage of the fact that ι is a mere inclusion. Let (X, x) be a complete join semi-lattice. The $(\iota.\eta^{\mathbf{P}})$ and $(\iota.\mu^{\mathbf{P}})$ diagrams of weak liftings immediately yield, for every $u \in X$ and every $\mathcal{U}$ in the carrier of $\overline{PP}(X, x)$,

$$\eta^{\overline{\mathbf{P}}}_{(X,x)}(u) = \{u\} \tag{3.19}$$

$$\mu^{\overline{\mathbf{P}}}_{(X,x)}(\mathcal{U}) = \bigcup \mathcal{U} \tag{3.20}$$

To sum up, the *upclosed powerset monad* $\overline{\mathbf{P}}$ maps any complete join-semilattice (X, x) to the complete join-semilattice of its x-upclosed subsets, and acts as the usual powerset monad when it comes to morphisms, unit and multiplication. Using Lemma 2.10, it is easy to obtain an expression of the composite monad $\mathbf{P} \bullet \mathbf{P}$. In the following definition, we keep track of the subsets level using an index notation: U_1 denotes a subset of X, U_2 denotes a subset of the subsets of X, etc.

Definition 3.7. The *monad of upclosed sets of subsets* $\mathbf{P} \bullet \mathbf{P}$ is defined as follows.

$$\begin{aligned} &(P \bullet P)X = \{U_2 \in PPX \mid U_2 \text{ is closed under non-empty unions}\} \\ &(P \bullet P)f(U_2) = \{f(U_1) \mid U_1 \in U_2\} \\ &\eta^{\mathbf{P}\bullet\mathbf{P}}_X(x) = \{\{x\}\} \\ &\mu^{\mathbf{P}\bullet\mathbf{P}}_X(U_4) = \bigcup_{U_3 \in U_4} \left\{\bigcup V_2 \mid V_2 \subseteq \bigcup U_3 \text{ such that } \forall U_2 \in U_3 . U_2 \cap V_2 \neq \emptyset\right\} \end{aligned}$$

3.1.2 Variations

Let us consider a few variations of the monotone weak distributive law $\lambda : \mathsf{PP} \to \mathsf{PP}$ whose expression is recalled below:

$$\lambda_X(\mathcal{U}) = \left\{ V \in PX \mid V \subseteq \bigcup \mathcal{U} \text{ and } \forall U \in \mathcal{U}. U \cap V \neq \emptyset \right\}$$

First, note that in [54], Garner provides a monotone weak distributive law of type $\mathsf{P}_f\mathsf{P} \to \mathsf{PP}_f$, mapping any *finite* $\mathcal{U} \in P_f PX$ to

$$\left\{ V \in PX \textit{ finite} \mid V \subseteq \bigcup \mathcal{U} \text{ and } \forall U \in \mathcal{U}. U \cap V \neq \emptyset \right\} \tag{3.21}$$

that is

$$\begin{aligned} \mathsf{P}_f\mathsf{P} &\to \mathsf{PP}_f \\ \mathcal{U} &\to \lambda_X(\mathcal{U}) \cap P_f X \end{aligned}$$

We note that it can be further restricted to

$$\begin{aligned} \mathsf{P}_f\mathsf{P}_f &\to \mathsf{P}_f\mathsf{P}_f \\ \mathcal{U} &\mapsto \lambda_X(\mathcal{U}) \end{aligned}$$

because if $\mathcal{U} \in P_f P_f X$, then $\lambda_X(\mathcal{U}) \subseteq P(\bigcup \mathcal{U})$ is finite.

By contrast, there seems to be no obvious way to get a weak distributive law of type $\mathsf{PP}_f \to \mathsf{P}_f\mathsf{P}$. For example, with $X = \mathbb{N}$ and $\mathcal{U} = \{\{0, n\} \mid n \in \mathbb{N}^*\} \in PP_f\mathbb{N}$, $\lambda_X(\mathcal{U}) = \{V \subseteq \mathbb{N} \mid 0 \in V \text{ or } \mathbb{N}^* \subseteq V\}$ is infinite, even if we restrict it to finite Vs.

The situation is slightly easier with the non-empty powerset monad P^*. As $\emptyset \in \lambda_X(\mathcal{U}) \iff \mathcal{U} = \emptyset$, the law $\lambda : \mathsf{PP} \to \mathsf{PP}$ restricts to a weak distributive law $\mathsf{P}^*\mathsf{P} \to \mathsf{PP}^*$. Similarly, $\bigcup \mathcal{U} \in \lambda_X(\mathcal{U}) \iff \forall U \in \mathcal{U}. U \neq \emptyset$, so λ restricts to a weak distributive law $\mathsf{PP}^* \to \mathsf{P}^*\mathsf{P}$. Applying both restrictions gives a weak distributive law $\mathsf{P}^*\mathsf{P}^* \to \mathsf{P}^*\mathsf{P}^*$. One can then safely impose non-emptyness of any of the two monads. The whole situation is summed up in Table 3.1.

Table 3.1: Existence of a monotone weak distributive law of type (row $\circ$ column) $\to$ (column $\circ$ row)

	P	P_f	P^*	P_f^*
P	✓	?	✓	?
P_f	✓	✓	✓	✓
P^*	✓	?	✓	?
P_f^*	✓	✓	✓	✓

Variations of the monotone law $\mathbf{PP} \to \mathbf{PP}$ give rise to variations of the monad of upclosed sets of subsets with similar algebras. In [54], Garner identifies $(\mathbf{P} \bullet \mathbf{P}_f)$-algebras with the commutative unital quantales whose multiplication is idempotent. The simplest case is maybe the one of $(\mathbf{P}_f^* \bullet \mathbf{P}_f^*)$-algebras, which are the join-distributive bisemilattices [126] (see also [88]). A join-distributive bisemilattice is a triple $(X, \vee, \wedge)$ such that $(X, \vee)$ and $(X, \wedge)$ are semilattices and the following distributivity equation holds

$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z) \tag{3.22}$$

This claim can be proved with the same method as in the upcoming Theorem 3.16. It consists in using the isomorphism $\mathsf{EM}(\mathbf{P}_f^* \bullet \mathbf{P}_f^*) \cong \mathsf{Alg}(\lambda)$ of Proposition 2.8, and following sets such as $\{\{x\}, \{y, z\}\}$ in the pentagon (1.38) defining λ-algebras.

3.2 Distribution over Powerset

3.2.1 The Monotone Law

We will again use Theorem 2.22 to find the unique monotone weak distributive law of type $\mathbf{DP} \to \mathbf{PD}$. First, we have:

Proposition 3.8 ([110, 45])**.** *The distribution functor D is weakly cartesian.*

We do not provide a proof, since this result is known and its details are not needed in the subsequent developments. The interested reader can refer to [110] for an elementary proof or to [45] for a proof using the max-flow min-cut theorem. Second, we verify that there is no hope for a monotone distributive law:

Proposition 3.9. *The unit $\eta^{\mathbf{D}}$ is not weakly cartesian.*

Proof. The argument is very similar to Proposition 3.2. Consider the naturality square of the unique map $! : \{0, 1\} \to \{0\}$.

$$\begin{tikzcd} \{0,1\} \arrow[r, "\eta^{\mathbf{D}}_{\{0,1\}}"] \arrow[d, "!"'] & D\{0,1\} \arrow[d, "D!"] \\ \{0\} \arrow[r, "\eta^{\mathbf{D}}_{\{0\}}"'] & D\{0\} \end{tikzcd} \tag{3.23}$$

As $D\{0\}$ is a singleton, the pullback of $\eta^{\mathbf{D}}_{\{0\}}$ and $D!$ is the infinite set $\{0\} \times D\{0, 1\}$, so there is no surjective function from $\{0, 1\}$ into it. Therefore, $\eta^{\mathbf{D}}$ is not weakly cartesian. $\square$

However, things work fine for the multiplication:

Proposition 3.10. *The multiplication μ^{D} is weakly cartesian.*

Proof. We provide the proof using Farkas' lemma presented in [62]. In this proof, a distribution $\varphi \in DX$ will be formally defined by the expression

$$\varphi = \sum_{i=1}^{n} p_i \cdot x_i$$

where $(x_i)_{1 \leq i \leq n}$ is an injective enumeration of $\mathsf{supp}(\varphi)$ and $p_i = \varphi(x_i) \in (0, 1]$. Consider the μ^{D} naturality square of a function $f : X \to Y$

$$\begin{tikzcd} DDX \arrow[r, "\mu^{\mathsf{D}}_X"] \arrow[d, "DDf"'] & DX \arrow[d, "Df"] \\ DDY \arrow[r, "\mu^{\mathsf{D}}_Y"'] & DY \end{tikzcd} \tag{3.24}$$

and let us prove that this is a weak pullback. Let $\varphi \in DX$ and $\Psi \in DDY$ be such that $Df(\varphi) = \mu^{\mathsf{D}}_Y(\Psi)$. We must find $\Phi \in DDX$ such that $\mu^{\mathsf{D}}_X(\Phi) = \varphi$ and $DDf(\Phi) = \Psi$. Write formally

$$\Psi = \sum_{j=1}^{m} q_j \cdot \psi_j$$

Assume there are distributions $(\varphi_j)_{1 \leq j \leq m}$ in DX such that the two following conditions are satisfied

$$Df(\varphi_j) = \psi_j \text{ for all } j \in \{1, \dots, m\} \tag{3.25}$$

$$\mu^{\mathsf{D}}_X \left(\sum_{j=1}^{m} q_j \cdot \varphi_j \right) = \varphi \tag{3.26}$$

Then we can formally define $\Phi = \sum_{j=1}^{m} q_j \cdot \varphi_j$. This enumeration of $\mathsf{supp}(\Phi)$ is injective because of equation (3.25). This Φ completes the proof, because equation (3.26) is exactly $\mu^{\mathsf{D}}_X(\Phi) = \varphi$ and equation (3.26) yields

$$DDf(\Phi) = \sum_{j=1}^{m} q_j \cdot Df(\varphi_j) = \sum_{j=1}^{m} q_j \cdot \psi_j = \Psi$$

In the remainder of the proof, we build the distributions $(\varphi_j)_{1 \leq j \leq m}$ by providing a non-negative solution to the following linear system of equations, with variables $\varphi_j(x)$ for $1 \leq j \leq m$ and $x \in \mathsf{supp}(\varphi)$:

$$\begin{cases} \sum_{x \in f^{-1}(\{y\})} \varphi_j(x) = \psi_j(y) \text{ for all } y \in Y, j \in \{1, \dots, m\} \\ \\ \sum_{j=1}^{m} q_j \varphi_j(x) = \varphi(x) \text{ for all } x \in \mathsf{supp}(\varphi) \end{cases} \tag{3.27}$$

First note that it suffices to find, locally for each $y \in Y$, a non-negative solution to the following linear systems with variables $\varphi_j^y(x)$:

$$\begin{cases} \sum_{x \in f^{-1}(\{y\})} \varphi_j^y(x) = \psi_j(y) \text{ for all } j \in \{1, \dots, m\} \\ \\ \sum_{j=1}^m q_j \varphi_j^y(x) = \varphi(x) \text{ for all } x \in f^{-1}(\{y\}) \cap \mathsf{supp}(\varphi) \end{cases} \tag{3.28}$$

Indeed, in the presence of local solutions, one can define a global solution by $\varphi_j(x) = \varphi_j^{f(x)}(x)$. Note that $\varphi_j \in DX$ because $\mathsf{supp}(\varphi_j) \subseteq \mathsf{supp}(\varphi)$ (by the second equation in (3.28)), so it is finitely supported, and

$$\sum_{x \in X} \varphi_j(x) = \sum_{x \in X} \varphi_j^{f(x)}(x) = \sum_{y \in Y} \sum_{x \in f^{-1}(\{y\})} \varphi_j^y(x) = \sum_{y \in Y} \psi_j(y) = 1$$

Therefore we can fix $y \in Y$ and solve the system given by equation (3.28). To see things more clearly, we will express this system using matrices. Writing $f^{-1}(\{y\}) \cap \mathsf{supp}(\varphi) = \{x_1, \dots, x_k\}$ the system of $m \times k$ variables and $m + k$ equations is

$$\begin{cases} \sum_{i=1}^k \varphi_j^y(x_i) = \psi_j(y) \text{ for all } j \in \{1, \dots, m\} \\ \\ \sum_{j=1}^m q_j \varphi_j^y(x_i) = \varphi(x_i) \text{ for all } i \in \{1, \dots, k\} \end{cases} \tag{3.29}$$

In order to abstract away notation related to distributions, variables, coefficients, we define the vectors

$$\begin{aligned} u &= (u_{1,1}, \dots, u_{m,1}, \dots, u_{1,k}, \dots, u_{m,k}) && (3.30) \\ &= (\varphi_1^y(x_1), \dots, \varphi_m^y(x_1), \dots, \varphi_1^y(x_k), \dots, \varphi_m^y(x_k)) && (3.31) \\ v &= (v_1, \dots, v_{m+k}) && (3.32) \\ &= (\psi_1(y), \dots, \psi_m(y), \varphi(x_1), \dots, \varphi(x_k)) && (3.33) \\ q &= (q_1, \dots, q_m) && (3.34) \end{aligned}$$

and the matrix

$$M = \left(\begin{array}{cccc} I_m & I_m & \dots & I_m \\ \hline q & 0 & \dots & 0 \\ 0 & q & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & q \end{array} \right) \in \mathcal{M}_{m+k,mk}(\mathbb{R}) \tag{3.35}$$

where I_m is the identity matrix of size m. To find a non-negative solution to the system $Mu = v$, we can now apply the widely known Farkas lemma.

Lemma 3.11 (Farkas, [50])**.** *Let $M \in \mathcal{M}_{p,q}(\mathbb{R})$ and $b \in \mathbb{R}^p$. Then exactly one of the following statements is true:*

- *There exists some $x \in \mathbb{R}^q$ such that $Mx = b$ and $x \geq 0$.*
- *There exists some $z \in \mathbb{R}^p$ such that $M^T z \geq 0$ and $b^T z < 0$.*

Take $p = m + k$, $q = mk$, $M = M$ and $b = v$. Assume towards a contradiction that there is some $z \in \mathbb{R}^{m+k}$ such that $M^T z \geq 0$ and $v^T z < 0$, that is:

$$z_i + q_i z_{m+j} \geq 0 \text{ for all } i \in \{1, \ldots, m\} \text{ and } j \in \{1, \ldots, k\} \qquad (\alpha_{i,j})$$

$$\sum_{i=1}^{m+k} v_i z_i < 0 \qquad (\beta)$$

Our hypothesis $\mu_Y^{\mathsf{D}}(\Psi) = Df(\varphi)$ applied on y also yields

$$\sum_{i=1}^{m} q_i v_i = \sum_{j=1}^{k} v_{m+j} \qquad (\gamma)$$

For a fixed $j \in \{1, \ldots, k\}$, summing the equations $v_i \times (\alpha_{i,j})$ over $i \in \{1, \ldots, m\}$ gives

$$\sum_{i=1}^{m} v_i z_i + z_{m+j} \sum_{i=1}^{m} q_i v_i \geq 0$$

Using equations (β) and (γ) we obtain

$$z_{m+j} \sum_{l=1}^{k} v_{m+l} \geq -\sum_{i=1}^{m} v_i z_i > \sum_{l=1}^{k} v_{m+l} z_{m+l}$$

Let $S = \left(\sum_{l=1}^{k} v_{m+l} z_{m+l}\right) / \left(\sum_{l=1}^{k} v_{m+l}\right)$ be the weighted average of all z_{m+l}. This quantity is well-defined because v_{m+l} are positive and $k \neq 0$. We just proved that $z_{m+j} > S$ for all j. By taking the weighted average of both sides with respect to the v_{m+j} again, we obtain $S > S$, a contradiction. Therefore, the first statement of Farkas' lemma is true. This gives a non-negative solution u to the equation $Mu = v$. Consequently, μ^{D} is weakly cartesian. □

By applying Theorem 2.22 we obtain that there is a unique monotone weak distributive law of type $\mathsf{DP} \to \mathsf{PD}$. In the same manner as in the previous section, we will now compute the triptych weak extension / weak distributive law / weak lifting and the composite monad. We first compute the weak extension on a relation $R \subseteq X \times Y$ with projections $\pi_1 : R \to X$ and $\pi_2 : R \to Y$. Here equation (2.34) amounts to

$$\underline{D}R = \{(\varphi, \psi) \in DX \times DY \mid \exists \theta \in DR.D\pi_1(\theta) = \varphi \text{ and } D\pi_2(\theta) = \psi\} \qquad (3.36)$$

In probabilistic parlance, $(\varphi, \psi) \in \underline{D}R$ whenever φ and ψ are the marginals of some distribution on R. There is a nice expression of the weak distributive law:

Theorem 3.12. *The unique monotone weak distributive law* $\mathbf{DP} \to \mathbf{PD}$ *is defined on any* $\Phi \in DPX$*, formally expressed as* $\Phi = \sum_{i=1}^{n} p_i \cdot U_i$ *with* $p_i > 0$ *and distinct* U_i*, by*

$$\lambda_X \left(\sum_{i=1}^{n} p_i \cdot U_i \right) = \left\{ \mu_X^{\mathbf{D}} \left(\sum_{i=1}^{n} p_i \cdot \varphi_i \right) \mid \forall i. \varphi_i \in DX \text{ and } \mathsf{supp}\varphi_i \subseteq U_i \right\} \tag{3.37}$$

Proof. According to the correspondence of Theorem 2.5, the monotone weak distributive law at X is the function λ_X whose graph is $\underline{D}(\ni_X)$, where $\ni_X$ is the set $\{(U, x) \mid x \in U\} \subseteq PX \times X$. Therefore

$$\begin{aligned} \lambda_X(\Phi) = \{\varphi \in DX \mid \exists \Theta \in D(\ni_X). \forall U \in PX. \Phi(U) = \sum_{x \in U} \Theta(U, x) \text{ and} \\ \forall x \in X. \varphi(x) = \sum_{U \ni x} \Theta(U, x)\} \end{aligned}$$

Let $\varphi \in \lambda_X(\Phi)$ and $\Theta \in D(\ni_X)$ as above. For any $x \in X$, let $\varphi_i(x) = \Theta(U_i, x)/p_i$, where Θ is extended to $D(PX \times X)$ with value 0 outside of $\ni_X$. Then $\varphi_i \in DX$ because

$$\sum_{x \in X} \varphi_i(x) = (1/p_i) \sum_{x \in U_i} \Theta(U_i, x) = \Phi(U_i)/p_i = 1$$

Furthermore, it is clear that $\mathsf{supp}(\varphi_i) \subseteq U_i$, and for any $x \in X$,

$$\varphi(x) = \sum_{U \ni x} \Theta(U, x) = \sum_{i=1}^{n} \Theta(U_i, x) = \sum_{i=1}^{n} p_i \varphi_i(x) = \mu_X^{\mathbf{D}} \left(\sum_{i=1}^{n} p_i \cdot \varphi_i \right)(x)$$

Conversely, let $\varphi_i \in DX$ with $\mathsf{supp}(\varphi_i) \subseteq U_i$ for every $i \in \{1, \ldots, n\}$ and let us show that $\mu_X^{\mathbf{D}} \left(\sum_{i=1}^{n} p_i \cdot \varphi_i \right) \in \lambda_X(\Phi)$. Define $\Theta(U_i, x) = p_i \varphi_i(x)$ for every i and $\Theta(U, x) = 0$ for other $U \ni x$. Then

$$\sum_{x \in U} \Theta(U, x) = \Phi(U)$$

because both have value p_i on U_is and 0 elsewhere. Moreover

$$\sum_{U \ni x} \Theta(U, x) = \sum_{i=1}^{n} \Theta(U_i, x) = \sum_{i=1}^{n} p_i \varphi_i(x) = \mu_X^{\mathbf{D}} \left(\sum_{i=1}^{n} p_i \cdot \varphi_i \right)(x) \qquad \square$$

Let us compute the weak lifting of $\mathbf{P}$ to the category $\mathsf{EM}(\mathbf{D})$ of convex algebras. For (X, x) a convex algebra, the idempotent

$$Px \circ \lambda_X \circ \eta_{PX}^{\mathbf{D}} : (PX, Px \circ \lambda_X) \to (PX, Px \circ \lambda_X) \tag{3.38}$$

maps a subset $U \subseteq X$ to

$$(Px \circ \lambda_X \circ \eta^{\mathbf{D}}_{PX})(U) = \{x(\varphi) \mid \varphi \in DU\} \tag{3.39}$$

Here, $\varphi \in DU$ is a short way of saying that $\varphi \in DX$ has support included in U. To better formulate what follows, it is useful to introduce the notion of convex subset for a convex algebra.

Definition 3.13 (Convex closure, convex subset)**.** Let (X, x) be a convex algebra. For any $U \subseteq X$, the set

$$\mathsf{conv}_x(U) = \{x(\varphi) \mid \varphi \in DU\} \tag{3.40}$$

is called the *convex closure* of U. We say that U is an *x-convex subset* whenever $U = \mathsf{conv}_x(U)$.

Note that the inclusion $U \subseteq \mathsf{conv}_x(U)$ always holds, because of the unit axiom of $\mathbf{D}$-algebras. Fixpoints of the function $Px \circ \lambda_X \circ \eta^{\mathbf{D}}_{PX}$ are exactly the x-convex subsets of X. From this observation, we deduce that the weak lifting is given by the subalgebra obtained by restriction of $Px \circ \lambda_X$ to x-upclosed subsets.

$$\overline{P}(X, x) = (S, s) \tag{3.41}$$

$$S = \{U \in PX \mid U = \mathsf{conv}_x(U)\} \tag{3.42}$$

$$\pi_{(X,x)}(U) = \mathsf{conv}_x(U) \tag{3.43}$$

$$\iota_{(X,x)}(U) = U \tag{3.44}$$

As for the previous section, we provide an expression of the $\mathbf{D}$-algebra $s : DS \to S$.

Lemma 3.14. *For every $\Phi = \sum_{i=1}^n p_i \cdot U_i \in DS$, with distinct U_i and non-zero p_i,*

$$s(\Phi) = \left\{ x\left(\sum_{i=1}^n p_i \cdot u_i\right) \mid \forall i \in \{1, \ldots, n\}.u_i \in U_i \right\} \tag{3.45}$$

Proof. The proof is very similar to Lemma 3.6, because both weak distributive laws rely on a notion of convexity (see [22, Section 5]). Yet we provide details for the sake of completeness. As s is the restriction of $Px \circ \lambda_X$,

$$\begin{aligned} s(\Phi) &= \left\{ (x \circ \mu^{\mathbf{D}}_X)\left(\sum_{i=1}^n p_i \cdot \varphi_i\right) \mid \forall i.\varphi_i \in DU_i \right\} \\ &= \left\{ (x \circ Dx)\left(\sum_{i=1}^n p_i \cdot \varphi_i\right) \mid \forall i.\varphi_i \in DU_i \right\} \\ &= \left\{ x\left(\sum_{i=1}^n p_i \cdot x(\varphi_i)\right) \mid \forall i.\varphi_i \in DU_i \right\} \end{aligned}$$

To prove equation (3.45), we reason by double inclusion. From $u_i \in U_i$ we can define $\varphi_i = 1 \cdot u_i$. For the other inclusion, from $\varphi_i \in DU_i$ we can define $u_i = x(\varphi_i) \in U_i$ because U_i is x-convex. $\square$

By exploiting ι as in Section 3.1, we also get

$$\overline{P}f(U) = Pf(U) \tag{3.46}$$

$$\eta^{\overline{\mathsf{P}}}_{(X,x)}(u) = \{u\} \tag{3.47}$$

$$\mu^{\overline{\mathsf{P}}}_{(X,x)}(\mathcal{U}) = \bigcup \mathcal{U} \tag{3.48}$$

To sum up, the monad $\overline{\mathsf{P}}$ maps a convex algebra (X, x) to the convex algebra whose carrier is the set of all x-convex subsets of X, with the action x being extended pointwise (as shown in Lemma (3.14)), and $\overline{\mathsf{P}}$ acts as the powerset for what concerns morphisms, unit, and multiplication. In [23], Bonchi, Silva and Sokolova build from scratch a slight variation of this monad comprising only non-empty convex subsets. With respect to their terminology, we call $\overline{\mathsf{P}}$ the *convex powerset monad.* It has to be noted that considerations about the empty set originate in [78], where Jacobs exhibits a similar construction using the free semimodule monad instead of our D. To give a homogeneous account of his construction, Jacobs had to drop the empty set – as stated later in [62, 22], it turns out that this restriction was unnecessary.

Using Lemma 2.10 it is easy to find an expression of the composite monad $\mathsf{P} \bullet \mathsf{D}$ on Set.

Definition 3.15. The *monad of convex sets of distributions* $\mathsf{P} \bullet \mathsf{D}$ is defined by

$$\begin{aligned}
&(P \bullet D)(X) = \{U \in PDX \mid U \text{ is convex with respect to } \mu^{\mathsf{D}}_X\} \\
&(P \bullet D)f(U) = \{Df(\varphi) \mid \varphi \in U\} \\
&\eta^{\mathsf{P}\bullet\mathsf{D}}(x) = \{1 \cdot x\} \\
&\mu^{\mathsf{P}\bullet\mathsf{D}}(\mathcal{U}) = \bigcup_{\Phi \in \mathcal{U}} \left\{ \mu^{\mathsf{D}}_X \left(\sum_{U \in \mathsf{supp}(\Phi)} \Phi_U \cdot \varphi^U \right) \mid \forall U \in \mathsf{supp}(\Phi). \varphi^U \in U \right\}
\end{aligned}$$

The terminology *monad of convex sets of distributions* is borrowed to Mio [105], but it has to be noted that mild variations of this monad appear at various places in the literature [78, 148, 23, 25], where it comes often alongside the convex powerset monad. The ideas underlying both the monad of convex sets of distributions and the convex powerset monad originate in domain theory, where the combination of probabilistic choice and non-deterministic choice was studied extensively by many authors – see the introduction of [88] for a comprehensive account. Notably, these monads are close to the notion of power Kegelspitzen of Keimel and Plotkin in *loc. cit.*

3.2.2 Algebraic Presentation

We prove, using the monotone weak distributive law $\mathsf{DP} \to \mathsf{PD}$, that the monad of convex sets of distributions can be presented by the (infinitary) equational theory of complete convex join-semilattices.

Related work. Using different methods, a similar result was derived by Bonchi, Sokolova and Vignudelli [25, 26] for *finitely generated* convex sets of distributions. Intuitively, this amounts to replacing P with its finite version P_f. Technically, things become slightly more involved. The link between these two algebraic presentations was later established by Bonchi and Santamaria in [22], building on our LICS paper [62]. Quantitative variations of the algebraic presentation have been studied in categories of metric spaces [106].

Theorem 3.16. *The monad of convex sets of distributions* $\mathsf{P} \bullet \mathsf{D}$ *is presented by the equational theory of complete convex join semi-lattices* $(\Sigma, E) = (\Sigma_P \cup \Sigma_D, E_P \cup E_D \cup E_\lambda)$*, where*

- (Σ_P, E_P) *is the equational theory of complete join semi-lattices described in Example 1.50 i.e.* $\Sigma_P = \{\bigvee_{i\in I} \mid I \text{ is a set}\}$ *and* E_P *are equations*
  - $\bigvee_{i\in\{*\}} x_i = x_*$
  - $\bigvee_{j\in J} x_j = \bigvee_{i\in I} x_{f(i)}$ *for every surjective* $f : I \to J$
  - $\bigvee_{i\in I} x_i = \bigvee_{j\in J} \bigvee_{i\in f^{-1}(\{j\})} x_i$ *for every* $f : I \to J$
- (Σ_D, E_D) *is the equational theory of convex algebras described in Example 1.49 i.e.* $\Sigma_D = \{\oplus_r \mid r \in [0, 1]\}$ *and* E_D *are equations*
  - $x \oplus_1 y = x$
  - $x \oplus_r x = x$
  - $x \oplus_r y = y \oplus_{1-r} x$
  - $(x \oplus_p y) \oplus_r z = x \oplus_{pr} \left(y \oplus_{\frac{r-pr}{1-pr}} z\right)$ *if* $r, p \neq 1$

- *E_λ consists of the following equations, for every $r \in [0,1]$ and every set I*

$$x \oplus_r \bigvee_{i \in I} y_i = \bigvee_{i \in I} (x \oplus_r y_i) \tag{3.49}$$

We start with some basic reminders. The isomorphism $\mathsf{EM}(\mathbf{P}) \cong \mathsf{Alg}(\Sigma_P, E_P)$ sends a $\mathbf{P}$-algebra (X, j) – with $j : PX \to X$ seen as an abstract join operator – to the (Σ_P, E_P)-algebra on X defined by

$$\bigvee_{i \in I} x_i = j(\{x_i \mid i \in I\}) \tag{3.50}$$

Conversely, a (Σ_P, E_P)-algebra is sent to the $\mathbf{P}$-algebra (X, j) defined by reading the above equation from right to left.

The isomorphism $\mathsf{EM}(\mathbf{D}) \cong \mathsf{Alg}(\Sigma_D, E_D)$ sends a $\mathbf{D}$-algebra (X, b) – with $b : DX \to X$ seen as an abstract barycenter operator – to the (Σ_D, E_D)-algebra on X defined by

$$x \oplus_r y = b(r \cdot x + (1-r) \cdot y) \tag{3.51}$$

Conversely, a (Σ_D, E_D)-algebra is sent to the $\mathbf{D}$-algebra (X, b) defined by induction on support cardinality as follows

$$b(1 \cdot x) = x \tag{3.52}$$

$$b(r \cdot x + (1-r)\varphi) = x \oplus_r b(\varphi) \text{ where } r \neq 0 \text{ and } x \notin \mathsf{supp}(\varphi) \tag{3.53}$$

In both cases, isomorphisms act on morphisms trivially: a $\mathbf{P}$-algebra morphism f corresponds to the (Σ_P, E_P)-algebra morphism f, and similarly for $\mathbf{D}$.

According to Proposition 2.8, $\mathsf{EM}(\mathbf{P} \bullet \mathbf{D}) \cong \mathsf{Alg}(\lambda)$, where $\lambda : \mathbf{DP} \to \mathbf{PD}$ is the monotone weak distributive law. Therefore, to prove that (Σ, E) presents $\mathbf{P} \bullet \mathbf{D}$, it suffices to prove that $\mathsf{Alg}(\lambda) \cong \mathsf{Alg}(\Sigma, E)$. Recall that $\mathsf{Alg}(\lambda)$ objects are triples (X, b, j) where (X, b) is a $\mathbf{D}$-algebra and (X, j) is a $\mathbf{P}$-algebra, such that the following pentagon commutes:

$$\begin{tikzcd} DPX \arrow[rr, "\lambda_X"] \arrow[d, "Dj"'] & & PDX \arrow[d, "Pb"] \\ DX \arrow[dr, "b"'] & & PX \arrow[dl, "j"] \\ & X & \end{tikzcd} \tag{3.54}$$

and that morphisms between λ-algebras are the simultaneous $\mathbf{D}$- and $\mathbf{P}$-algebra morphisms. The isomorphism $\mathsf{Alg}(\lambda) \cong \mathsf{Alg}(\Sigma, E)$ is simply

$$(X, b, j) \leftrightarrow \left(X, \bigvee_{i \in I}, \oplus_r \right) \tag{3.55}$$

where operators are defined with respect to each other as in equations (3.50), (3.51), (3.52) and (3.53) – and the action on morphisms is the identity. Because (Σ_P, E_P) presents P and (Σ_D, E_D) presents D, the only thing that has to be checked is that (X, b, j) makes the pentagon (3.54) commute if and only if equations of E_λ displayed in (3.49) are satisfied.

Pentagon implies E_λ. Let (X, b, j) be a λ-algebra. Recall that λ is given by the formula of Theorem 3.12:

$$\lambda_X\left(\sum_{i=1}^n p_i \cdot U_i\right) = \left\{\mu_X^{\mathsf{D}}\left(\sum_{i=1}^n p_i \cdot \varphi_i\right) \mid \forall i.\varphi_i \in DX \text{ and } \mathsf{supp}\varphi_i \subseteq U_i\right\}$$

Taking the convex closure of a subset does not increase its supremum:

Lemma 3.17. *For any $A \in PX$, $\bigvee_{x \in A} x = \bigvee_{x \in \mathsf{conv}_b(A)} x$.*

Proof. Follow the distribution $1 \cdot A$ along both paths of the pentagon. The left leg gives

$$(b \circ Dj)(1 \cdot A) = b(1 \cdot j(A)) = j(A) = \bigvee_{x \in A} x$$

by the unit axiom of the D-algebra b, and $\bigvee$ definition. The right leg gives

$$\begin{aligned}(j \circ Pb \circ \lambda_X)(1 \cdot A) &= (j \circ Pb)(\{\varphi \in DX \mid \mathsf{supp}\varphi \subseteq A\}) \\ &= j\{b(\varphi) \mid \varphi \in DX.\mathsf{supp}\varphi \subseteq A\} \\ &= j(\mathsf{conv}_b(A)) \\ &= \bigvee_{x \in \mathsf{conv}_b(A)} x\end{aligned}$$

by definition of conv and $\bigvee$, whence the required equality. ☐

Proposition 3.18. *For any $x \in X$ and $(y_i)_{i \in I} \in X^I$, $x \oplus_r \bigvee_{i \in I} y_i = \bigvee_{i \in I}(x \oplus_r y_i)$.*

Proof. Follow the distribution $r \cdot \{x\} + (1 - r) \cdot \{y_i \mid i \in I\}$ along both paths of the pentagon. The left leg gives

$$\begin{aligned}(b \circ Dj)(r \cdot \{x\} + (1 - r) \cdot \{y_i \mid i \in I\}) &= b(r \cdot j(\{x\}) + (1 - r) \cdot j\{y_i \mid i \in I\}) \\ &= x \oplus_r \bigvee_{i \in I} y_i\end{aligned}$$

by definition of $\bigvee$, $\oplus_r$ and the unit axiom of the $\mathbf{P}$-algebra j. The right leg gives

$$\begin{aligned}
&(j \circ Pb \circ \lambda_X)(r \cdot \{x\} + (1-r) \cdot \{y_i \mid i \in I\}) \\
&= j(\{b(r \cdot x + (1-r)\varphi) \mid \mathsf{supp}\varphi \subseteq \{y_i \mid i \in I\}) && \lambda \text{ definition} \\
&= j(\mathsf{conv}_b(\{b(r \cdot x + (1-r) \cdot y_i) \mid i \in I\})) && \mathsf{conv} \text{ definition} \\
&= j(\{b(r \cdot x + (1-r) \cdot y_i) \mid i \in I\}) && \text{Lemma 3.17} \\
&= \bigvee_{i \in I} (x \oplus_r y_i) && \bigvee \text{ and } \oplus_r \text{ definition}
\end{aligned}$$

whence the required equality. □

E_λ implies pentagon. Let $(X, \bigvee_{i\in I}, \oplus_r)$ be a (Σ, E)-algebra. Equations of E_λ

$$x \oplus_r \bigvee_{i \in I} y_i = \bigvee_{i \in I} (x \oplus_r y_i) \tag{3.56}$$

will be simply referred to as *distributivity axioms*. A first easy result is that these equations hold as well in the other direction:

$$\begin{aligned}
\left(\bigvee_{i\in I} x_i\right) \oplus_r y &= y \oplus_{1-r} \left(\bigvee_{i \in I} x_i\right) \\
&= \bigvee_{i \in I} (y \oplus_{1-r} x_i) \\
&= \bigvee_{i \in I} (x_i \oplus_r y)
\end{aligned}$$

A consequence is double distributivity:

$$\left(\bigvee_{i\in I} x_i\right) \oplus_r \left(\bigvee_{j \in J} y_j\right) = \bigvee_{(i,j) \in I \times J} (x_i \oplus_r y_j) \tag{3.57}$$

The join of two elements will be denoted by $a \vee b = \bigvee_{x \in \{a,b\}} x$. It is noteworthy that as $(X, \bigvee_{i \in I})$ is a complete join-semilattice, we can use any property of complete join-semilattices. For instance, the relation $a \leq b \iff a \vee b = b$ defines an order on X. We begin by proving some lemmas relating this order and the operations $\oplus_r$.

Lemma 3.19. *For all $a, b, c, d \in X$ and $r \in [0,1]$, if $a \leq b$ and $c \leq d$, then $a \oplus_r c \leq b \oplus_r d$.*

Proof. First note that by distributivity $(a \oplus_r c) \vee (b \oplus_r c) = (a \vee b) \oplus_r c = b \oplus_r c$, so $a \oplus_r c \leq b \oplus_r c$. Similarly, $b \oplus_r c \leq b \oplus_r d$, so by transitivity $a \oplus_r c \leq b \oplus_r d$. □

The following lemma shows that a convex combination always lies below the join of its support.

Lemma 3.20. *For every $\varphi \in DX$, $b(\varphi) \leq \bigvee_{x \in \mathsf{supp}(\varphi)} x$.*

Proof. By induction on the cardinality of $\mathsf{supp}(\varphi)$.

- If $\varphi = 1 \cdot x$, the property rewrites as $b(1 \cdot x) \leq \bigvee\{x\}$ i.e. $x \leq x$
- If $\varphi = r \cdot x + (1-r) \cdot y$, with $r \neq 0$ and $x \neq y$, the property rewrites as $x \oplus_r y \leq x \vee y$. As $x \leq x \vee y$ and $y \leq x \vee y$, Lemma 3.19 and $\oplus_r$ idempotency yield $x \oplus_r y \leq (x \vee y) \oplus_r (x \vee y) = x \vee y$.
- Let $n \geq 2$ and assume that the result holds for all φ with support of cardinality at most n. Consider a distribution ψ whose support has cardinality $n+1$ and write it $\psi = r \cdot x + (1-r)\varphi$ where $r \neq 0$, $x \notin \mathsf{supp}(\varphi)$ and the support of φ has cardinality n. Then

$$\begin{aligned} b(\psi) = x \oplus_r b(\varphi) &\leq x \vee b(\varphi) && \text{by a base case} \\ &\leq x \vee \bigvee_{y \in \mathsf{supp}(\varphi)} y && \text{by induction hypothesis} \\ &= \bigvee_{y \in \mathsf{supp}(\psi)} y \end{aligned}$$

  which achieves the proof. □

Lemma 3.21. *For every $A \in PX$, $\bigvee_{x \in A} x = \bigvee_{\varphi \in DX, \mathsf{supp}(\varphi) \subseteq A} b(\varphi)$.*

Proof. For $x \in A$, the distribution $\varphi_x = 1 \cdot x$ has support included in A and $x \leq x = b(1 \cdot x)$. Consequently, $\bigvee_{x \in A} x \leq \bigvee_{\varphi \in DX, \mathsf{supp}(\varphi) \subseteq A} b(\varphi)$. For the converse direction, let $\varphi \in DX$ such that $\mathsf{supp}(\varphi) \subseteq A$. According to Lemma 3.20, $b(\varphi) \leq \bigvee_{x \in \mathsf{supp}(\varphi)} x \leq \bigvee_{x \in A} x$ because $\mathsf{supp}(\varphi) \subseteq A$. Hence, $\bigvee_{\varphi \in DX, \mathsf{supp}(\varphi) \subseteq A} b(\varphi) \leq \bigvee_{x \in A} x$. □

We finally prove the main statement.

Proposition 3.22. *The triple (X, b, j) makes the pentagon* (3.54) *commute.*

Proof. By induction on the cardinality of $\mathsf{supp}(\Phi)$, where $\Phi \in DPX$.

- If $\Phi = 1 \cdot A$ for some $A \in PX$, we must prove that $j(A) = j(\mathsf{conv}_b(A))$, which is nothing but Lemma 3.21.
- Let $n \geq 1$ and assume the pentagon commutes holds when starting from any $\Phi \in DPX$ such that $\mathsf{supp}(\Phi)$ has cardinality n. Let $\Psi \in DPX$ with support of cardinality $n+1$. We can write $\Psi = r \cdot A + (1-r)\Phi$ where $r \neq 0$, $\Phi \in DPX$ has

support of cardinality n and $A \notin \mathsf{supp}(\Phi)$. Denote formally $\Phi = \sum_{1\leq i\leq n} p_i \cdot A_i$. The induction hypothesis applied to Φ means that

$$b\left(\sum_{1\leq i\leq n} p_i \cdot j(A_i)\right) = j\left(\left\{b\left(\sum_{1\leq i\leq n} p_i \cdot b(\varphi_i)\right) \mid \varphi_i \in DX, \mathsf{supp}(\varphi_i) \subseteq A_i\right\}\right) \tag{3.58}$$

Now computing the pentagon for Ψ, the left leg gives

$$j(A) \oplus_r b\left(\sum_{1\leq i\leq n} p_i \cdot j(A_i)\right) \tag{3.59}$$

while the right leg gives

$$\begin{aligned}
& j\left(\left\{b(\varphi) \oplus_r b\left(\sum_{1\leq i\leq n} p_i \cdot b(\varphi_i)\right) \mid \varphi, \varphi_i \in DX, \mathsf{supp}(\varphi) \subseteq A, \mathsf{supp}(\varphi_i) \subseteq A_i\right\}\right) \\
&= \left(\bigvee_{\varphi, \mathsf{supp}(\varphi)\subseteq A} b(\varphi)\right) \oplus_r \left(\bigvee_{\varphi_i, \mathsf{supp}(\varphi_i)\subseteq A_i} b\left(\sum_{1\leq i\leq n} p_i \cdot b(\varphi_i)\right)\right) \text{ by equation 3.57} \\
&= j(A) \oplus_r b\left(\sum_{1\leq i\leq n} p_i \cdot j(A_i)\right)
\end{aligned}$$

where the last step uses jointly Lemma 3.21 and the induction hypothesis. □

This achieves the proof that the monad of convex sets of distributions is presented by the theory of complete convex join-semilattices.

3.2.3 Variations

It turns out that the same year, a simpler proof of Proposition 3.10 was found by Fritz and Perrone [52]. Their proof can be extended to prove that the multiplication of the *countable* distribution monad $\mathbf{D}_\omega$ is weakly cartesian, thus providing a variation of our weak distributive law $\mathbf{DP} \to \mathbf{PD}$ to a law $\mathbf{D}_\omega\mathbf{P} \to \mathbf{PD}_\omega$.

Theorem 3.23. *There is a unique monotone weak distributive law $\lambda : \mathbf{D}_\omega\mathbf{P} \to \mathbf{PD}_\omega$ given by the expected formula: for every $\Phi = \sum_{U\in PX} \Phi_U \cdot U \in D_\omega PX$,*

$$\lambda_X(\Phi) = \left\{\mu_X^{\mathbf{D}_\omega}\left(\sum_{U\in\mathsf{supp}(\Phi)} \Phi_U \cdot \varphi^U\right) \mid \forall U \in \mathsf{supp}(\Phi). \varphi^U \in DX \textit{ and } \mathsf{supp}(\varphi^U) \subseteq U\right\}$$

Proof. The countable distribution functor D_ω is known to be weakly cartesian [140, Lemma 3.5.6]. Note also that $\eta^{\mathbf{D}_\omega}$ is not weakly cartesian, using the same argument as for $\eta^{\mathbf{D}}$. To see that $\mu^{\mathbf{D}_\omega}$ is weakly cartesian, we can adapt the proof of Fritz and

Perrone [52, Proposition 6.4] to the countable case. Let $f : X \to Y$ be a function and consider the naturality square

$$
\begin{array}{ccc}
D_\omega D_\omega X & \xrightarrow{\mu_X^{\mathbf{D}_\omega}} & D_\omega X \\
{\scriptstyle D_\omega D_\omega f}\big\downarrow & & \big\downarrow{\scriptstyle D_\omega f} \\
D_\omega D_\omega Y & \xrightarrow[\mu_Y^{\mathbf{D}_\omega}]{} & D_\omega Y
\end{array}
\tag{3.60}
$$

Let $\varphi \in D_\omega X$ be defined formally as $\varphi = \sum_{x\in X} \varphi_x \cdot x$, that means, $\varphi_x \in [0,1]$ is the coefficient of φ at x, with $\sum_{x\in X} \varphi_x = 1$. Similarly, let $\Psi \in D_\omega D_\omega Y$ be defined formally as $\Psi = \sum_{\psi\in D_\omega Y} \Psi_\psi \cdot \psi$, where for every $\psi \in D_\omega Y$, $\psi = \sum_{y\in Y} \psi_y \cdot y$. Assume $\mu_Y^{\mathbf{D}_\omega}(\Psi) = D_\omega f(\varphi)$, that is, for all $y \in Y$,

$$\sum_{\psi\in D_\omega Y} \Psi_\psi \psi_y = \varphi^*_y \tag{3.61}$$

where φ^*_y stands for $\sum_{x\in f^{-1}(\{y\})} \varphi_x$. For every $\psi \in D_\omega Y$, let

$$\psi_* = \sum_{\substack{x\in X \\ \varphi^*_{f(x)}>0}} \psi_{f(x)} \frac{\varphi_x}{\varphi^*_{f(x)}} \cdot x$$

The construction ψ_* can be seen as a pullback distribution of ψ with respect to f in a sense made precise by the following lemma.

Lemma 3.24. *If $\Psi_\psi > 0$, then $\psi_* \in D_\omega X$ and $D_\omega f(\psi_*) = \psi$.*

Proof. Assume $\Psi_\psi > 0$. To check that $\psi_* \in D_\omega X$ it suffices to check that coefficients sum to 1. And indeed

$$\sum_{\substack{x\in X \\ \varphi^*_{f(x)}>0}} \psi_{f(x)} \frac{\varphi_x}{\varphi^*_{f(x)}} = \sum_{\substack{y\in Y \\ \varphi^*_y>0}} \psi_y = 1$$

The last equality holds because as $\Psi_\psi > 0$, if $\psi_y > 0$ then by equation (3.61) also $\varphi^*_y > 0$. Now let us check that $D_\omega f(\psi_*) = \psi$.

$$D_\omega f(\psi_*) = \sum_{\substack{x\in X \\ \varphi^*_{f(x)}>0}} \psi_{f(x)} \frac{\varphi_x}{\varphi^*_{f(x)}} \cdot f(x) = \sum_{\substack{y\in Y \\ \varphi^*_y>0}} \psi_y \cdot y = \psi$$

Again, the last equality holds because of equation (3.61). $\square$

Now we can define

$$\Phi = \sum_{\substack{\psi\in D_\omega Y \\ \Psi_\psi>0}} \Psi_\psi \cdot \psi_* \in D_\omega D_\omega X$$

It only remains to compute

$$D_\omega D_\omega f(\Phi) = \sum_{\substack{\psi \in D_\omega Y \\ \Psi_\psi > 0}} \Psi_\psi \cdot D_\omega f(\psi_*) = \sum_{\substack{\psi \in D_\omega Y \\ \Psi_\psi > 0}} \Psi_\psi \cdot \psi = \Psi$$

and

$$\begin{aligned}
\mu_X^{\mathbf{D}_\omega}(\Phi) &= \sum_{\substack{\psi \in D_\omega Y \\ \Psi_\psi > 0}} \sum_{\substack{x \in X \\ \varphi^*_{f(x)} > 0}} \Psi_\psi \psi_{f(x)} \frac{\varphi_x}{\varphi^*_{f(x)}} \cdot x \\
&= \sum_{\substack{x \in X \\ \varphi^*_{f(x)} > 0}} \left(\sum_{\substack{\psi \in D_\omega Y \\ \Psi_\psi > 0}} \Psi_\psi \psi_{f(x)} \right) \frac{\varphi_x}{\varphi^*_{f(x)}} \cdot x \\
&= \sum_{\substack{x \in X \\ \varphi^*_{f(x)} > 0}} \varphi_x \cdot x \qquad \text{by equation (3.61)} \\
&= \varphi
\end{aligned}$$

where the last equality holds because whenever $\varphi_x > 0$, then $\varphi^*_{f(x)} > 0$. This achieves the proof that $\mu^{\mathbf{D}_\omega}$ is weakly cartesian. Computation of the weak distributive law formula starting from the weak extension is as for the case $\mathbf{DP} \to \mathbf{PD}$ (Theorem 3.12). □

Another possible variation, hinted in [54], is to replace the distribution monad $\mathbf{D}$ with the S-left-semimodule monad $\mathbf{S}$ for some semiring S. Under certain conditions on S, there is a weak distributive law of type $\mathbf{SP} \to \mathbf{PS}$. This has been thoroughly explored by Bonchi and Santamaria [22], drawing inspiration from [36] and our LICS paper [62]. The monotone weak distributive law $\mathbf{P}_f\mathbf{P} \to \mathbf{PP}_f$ arises as an instance of their framework with S being the Booleans $\{0, 1\}$. Strictly speaking, our two main examples $\mathbf{PP} \to \mathbf{PP}$ and $\mathbf{DP} \to \mathbf{PD}$ do not fall into the scope of [22] – the first one because the outer $\mathbf{P}$ allows for infinite sets i.e. Boolean distributions with arbitrary support; the second one because of the condition $\sum_x \varphi_x = 1$. However, the gist is really close to what Bonchi and Santamaria study. In particular, the expressions given in Lemmas 3.6 and 3.14 are related to an observation made in [22, Section 4], namely that there are *minimal* distributions picking exactly one element out of every subset, and that taking the convex closure of the set of minimal distributions generate all the distributions of interest.

3.3 Powerset over Distribution

In the last section, combination of probability and non-determinism has been studied in the direction $\mathbf{DP} \to \mathbf{PD}$. We would like to make the monads commute in the other direction $\mathbf{PD} \to \mathbf{DP}$ as well. Indeed, systems resolving probabilistic choice before non-deterministic choice are related to the notion of random set and arise in domain theory in the work of Goubault-Larrecq [57, 59, 60]. However, it has been showed by Keimel and Plotkin [88] that adding the obvious equation

$$x \vee (y \oplus_r z) = (x \vee y) \oplus_r (x \vee z) \tag{3.62}$$

to the equational theory makes probabilities totally disappear. In the presence of equation (3.62), the operator $\oplus_r$ acts as a deterministic meet $\wedge$ and the resulting monad is the monad of bisemilattices $(X, \vee, \wedge)$ with $\vee$ distributing over $\wedge$ [126]. By letting go of idempotency in the algebraic theory of $\mathbf{P}$, some authors recover a composite algebraic theory [88, 39] which corresponds to the parallel multinomial distributive law $\mathbf{MD} \to \mathbf{DM}$ of Example 1.37 [80]. See also [41] for a continuous version in the category of measurable spaces. In the same manner as our monotone $\mathbf{DP} \to \mathbf{PD}$ can be regarded as a distorted restriction of the monotone $\mathbf{MP} \to \mathbf{PM}$ from Example 1.36, it could be possible to derive an interesting law $\mathbf{PD} \to \mathbf{DP}$ starting from Jacobs's parallel multinomial law. In this section we sketch some arguments hinting that, quite the contrary, there is little hope to get a meaningful law $\mathbf{PD} \to \mathbf{DP}$.

Consider the two monotone weak distributive laws $\lambda^{\mathbf{P}_f} : \mathbf{P}_f\mathbf{P} \to \mathbf{P}\mathbf{P}_f$ and $\lambda^{\mathbf{D}} : \mathbf{DP} \to \mathbf{PD}$, respectively described in equations (3.21) and (3.37). These laws turn out to be related to each other via the monad morphism $\mathsf{supp} : \mathbf{D} \to \mathbf{P}_f$ of Example 1.16.

Proposition 3.25. *The following diagram commutes:*

$$\begin{tikzcd} DP \arrow[r, "\lambda^{\mathbf{D}}"] \arrow[d, "\mathsf{supp}P"'] & PD \arrow[d, "P\mathsf{supp}"] \\ P_f P \arrow[r, "\lambda^{\mathbf{P}_f}"] & PP_f \end{tikzcd} \tag{3.63}$$

Proof. Let $\Phi \in DPX$ be defined formally as $\Phi = \sum_{1\leq i\leq n} p_i \cdot U_i$. The bottom left path evaluates to

$$\begin{aligned} (\lambda^{\mathbf{P}_f} \circ \mathsf{supp}P)_X(\Phi) &= \lambda^{\mathbf{P}_f}_X(\{U_i \mid 1 \leq i \leq n\}) && (3.64) \\ &= \left\{ V \subseteq X \text{ finite } \mid V \subseteq \bigcup_{1\leq i\leq n} U_i \text{ and } \forall i. V \cap U_i \neq \emptyset \right\} && (3.65) \end{aligned}$$

while the top right path evaluates to

$$(P\mathsf{supp} \circ \lambda^{\mathbf{D}})_X(\Phi) = \left\{ \mathsf{supp}\left(\sum_{1 \le i \le n} p_i \varphi_i \right) \mid \varphi_i \in DX, \mathsf{supp}(\varphi_i) \subseteq U_i \right\} \tag{3.66}$$

$$= \left\{ \bigcup_{1 \le i \le n} \mathsf{supp}(\varphi_i) \mid \varphi_i \in DX, \mathsf{supp}(\varphi_i) \subseteq U_i \right\} \tag{3.67}$$

Let us prove that these two sets are equal. Let $V \subseteq X$ be a finite subset satisfying $V \subseteq \bigcup_{1 \le i \le n} U_i$ and $V \cap U_i \neq \emptyset$. Define, for all $i \in \{1, \ldots, n\}$, the distribution φ_i to be uniform on the finite non-empty set $V \cap U_i$. Then $\mathsf{supp}(\varphi_i) \subseteq U_i$ and $V = \bigcup_{1 \le i \le n} \mathsf{supp}(\varphi_i)$. Conversely, let φ_i be distributions such that $\mathsf{supp}(\varphi_i) \subseteq U_i$. The set $\bigcup_{1 \le i \le n} \mathsf{supp}(\varphi_i)$ is finite, contained in $\bigcup_{1 \le i \le n} U_i$ and its intersection with a fixed U_{i_0} is precisely $\mathsf{supp}(\varphi_{i_0})$ which is non-empty. This completes the proof. □

In the sense of Proposition 3.25, the monad morphism $\mathsf{supp} : \mathbf{D} \to \mathbf{P}_f$ induces a morphism of weak distributive laws $\mathsf{supp} : \lambda^{\mathbf{D}} \to \lambda^{\mathbf{P}_f}$. This fact will impact on the generalised determinisation constructions presented in Chapter 4. It also suggests, as a mechanism to identify relevant laws, to connect them to laws we already have using monad morphisms (or just natural transformations) in the style of the square (3.63). In the upcoming examples, we will use the support monad morphism, either of type $\mathsf{supp} : \mathbf{D} \to \mathbf{P}$ or $\mathsf{supp} : \mathbf{D} \to \mathbf{P}_f$, to illustrate this idea. A first negative result is that the monotone laws $\lambda^{\mathbf{P}} : \mathbf{PP} \to \mathbf{PP}$ and $\lambda^{\mathbf{D}} : \mathbf{DP} \to \mathbf{PD}$ described in equations (3.8) and (3.37) are *not* connected by supp, because

Proposition 3.26. *There is no natural transformation $\alpha : DP \to PD$ such that*

$$\begin{tikzcd} DP \arrow[r, "\alpha"] \arrow[d, "\mathsf{supp}P"'] & PD \arrow[d, "P\mathsf{supp}"] \\ PP \arrow[r, "\lambda^{\mathbf{P}}"] & PP \end{tikzcd} \tag{3.68}$$

Proof. Let X be an infinite set. The Dirac distribution $1 \cdot X \in DPX$ is mapped via the bottom left path first to $\{X\}$, then to $PX \backslash \{\emptyset\}$. In particular $X \in (\lambda^{\mathbf{P}} \circ \mathsf{supp}P)_X(1 \cdot X)$, whereas no infinite set can be an element of $(P\mathsf{supp} \circ \alpha)_X(1 \cdot X)$. □

Similarly,

Proposition 3.27. *There is no natural transformation $\alpha : PD \to DP$ such that*

$$\begin{tikzcd} PD \arrow[r, "\alpha"] \arrow[d, "P\mathsf{supp}"'] & DP \arrow[d, "\mathsf{supp}P"] \\ PP \arrow[r, "\lambda^{\mathbf{P}}"] & PP \end{tikzcd} \tag{3.69}$$

Proof. Follow $\left\{(1/n)\sum_{1\le i\le n} 1\cdot i \mid n\in\mathbb{N}^*\right\} \in PD\mathbb{N}^*$ along the bottom left path. It is mapped first to $\{\{1,\ldots,n\} \mid n\in\mathbb{N}^*\}$, then to the set $\{U\subseteq \mathbb{N}^* \mid 1\in U\}$. Being infinite, this set cannot be in the image of $(\mathsf{supp}P\circ\alpha)_{\mathbb{N}^*}$. □

In both examples, the possibility of having infinite sets makes the connection via supp fail. By contrast, Proposition 3.25 works well precisely because the support is seen as a monad morphism of type $\mathbf{D}\to\mathbf{P}_f$. Can we use this mechanism to discover a law of type $\mathbf{PD}\to\mathbf{DP}$ using $\mathsf{supp}:\mathbf{D}\to\mathbf{P}_f$? This would require a square

$$\begin{tikzcd} PD \arrow[r, "\alpha"] \arrow[d, "P\mathsf{supp}"'] & DP \arrow[d, "\mathsf{supp}P"] \\ PP_f \arrow[r, "?"] & P_fP \end{tikzcd} \tag{3.70}$$

and it has already been noticed that there is no obvious (weak) distributive law $\mathbf{PP}_f\to\mathbf{P}_f\mathbf{P}$. There is a quick patch to this issue: simply consider, instead, the monotone weak distributive law $\lambda:\mathbf{P}_f\mathbf{P}_f\to\mathbf{P}_f\mathbf{P}_f$ at the bottom of the square, and look for a law of type $\alpha:\mathbf{P}_f\mathbf{D}\to\mathbf{DP}_f$ at the top. Unfortunately, it turns out that this situation is impossible again, but the argument is much more subtle.

Proposition 3.28. *There is no natural transformation $\alpha: P_fD\to DP_f$ such that*

$$\begin{tikzcd} P_fD \arrow[r, "\alpha"] \arrow[d, "P_f\mathsf{supp}"'] & DP_f \arrow[d, "\mathsf{supp}P_f"] \\ P_fP_f \arrow[r, "\lambda"] & P_fP_f \end{tikzcd} \tag{3.71}$$

Proof. Assume there is such a natural transformation α. We recall that λ is given for every $\mathcal{U}\in P_fP_fX$ by

$$\lambda_X(\mathcal{U}) = \left\{V\subseteq X \mid V\subseteq \bigcup\mathcal{U} \text{ and } \forall U\in\mathcal{U}.U\cap V\neq\emptyset\right\}$$

Let X be an infinite countable set. One can define, in X, two sequences $(x_n)_{n\in\mathbb{N}^*}$ and $(y_n)_{n\in\mathbb{N}^*}$ whose elements are all distinct. Then, define for any $n\in\mathbb{N}^*$ the following sets of distributions:

$$\Phi_n = \left\{\frac{1}{2}\cdot x_k + \frac{1}{2}\cdot y_k \mid k\in\{1,\ldots,n\}\right\}$$

Fix $n\in\mathbb{N}^*$. Using the hypothesis that the square (3.71) commutes, one can see that the set $\mathsf{supp}_{P_fX}(\alpha_X(\Phi_n))$ is precisely

$$\{U\subseteq\{x_k,y_k \mid k\in\{1,\ldots,n\}\} \mid \forall k\in\{1,\ldots,n\}.(x_k\in U \text{ or } y_k\in U)\} \tag{3.72}$$

In particular, every set U that picks exactly one element out of every pair $\{x_k, y_k\}$ for $k \in \{1, \ldots, n\}$ is in the support of $\alpha_X(\Phi_n)$. There are 2^n such sets. Moreover, all of them have the same probability coefficient in the distribution $\alpha_X(\Phi_n)$. Indeed, for such a set U one can define a function g_U such that for all $k \in \{1, \ldots, n\}$:

- $g_U(x_k)$ is the unique element of $\{x_k, y_k\}$ that belongs to U
- $g_U(y_k)$ is the unique element of $\{x_k, y_k\}$ that does not belong to U

The naturality square of α with respect to g_U shows that the coefficient for U is the same as the one for $\{x_k \mid k \in \{1, \ldots, n\}\}$, hence all these coefficients are equal. Let $p_n \in (0, 1/2^n)$ be this coefficient.

From now on, n is not fixed anymore. Let $f : X \to X$ be a function such that $f(x_n) = x_1$ and $f(y_n) = y_1$ for all $n \in \mathbb{N}^*$. We will now study how Φ_n evolves along the following naturality diagram

$$\begin{tikzcd} P_f D X \arrow[r, "\alpha_X"] \arrow[d, "P_f D f"'] & D P_f X \arrow[d, "D P_f f"] \\ P_f D X \arrow[r, "\alpha_X"] & D P_f X \end{tikzcd} \tag{3.73}$$

First compute the bottom left path:

$$(\alpha_X \circ P_f D f)(\Phi_n) = \alpha_X \left(\left\{ \frac{1}{2} \cdot x_1 + \frac{1}{2} \cdot y_1 \right\} \right) \tag{3.74}$$

$$= p_1 \cdot \{x_1\} + p_1 \cdot \{y_1\} + (1 - 2p_1) \cdot \{x_1, y_1\} \tag{3.75}$$

for the above-defined $p_1 \in (0, 1/2)$. Because of the naturality diagram, this means that p_1 is the sum of all coefficients in $\alpha_X(\Phi_n)$ whose associated subset is mapped to $\{x_1\}$ by $P_f f$. There is a unique such subset with non-zero coefficient, namely $\{x_k \mid k \in \{1, \ldots, n\}\}$, and it has coefficient p_n in $\alpha_X(\Phi_n)$. Henceforth $p_1 = p_n$. This yields that $p_1 \in (0, 1/2^n)$ for every $n \in \mathbb{N}^*$, a contradiction. □

In summary,

- there is no distributive law $\mathsf{PD} \to \mathsf{DP}$ nor $\mathsf{P}_f\mathsf{D} \to \mathsf{DP}_f$ [161],
- there is a trivial weak distributive law of each of these types, arising from the support monad morphism (Example 2.13),
- if λ is a *weak* distributive law of one of these types, then supp does not connect it to the monotone weak distributive laws $\mathsf{PP} \to \mathsf{PP}$ or $\mathsf{P}_f\mathsf{P}_f \to \mathsf{P}_f\mathsf{P}_f$, respectively.

Chapter 4

Applications to Coalgebra

Theoretical computer science comes with a plethora of abstract models of computation that share certain common features. In automata theory, many of these models are state-based, in the sense that they are dealing with certain sets of states, with computations being modelled as certain sets of transitions between states. This is the general structure, but depending on the system, one can implement various additional features in the model, such as labels, initial states, terminal states, inputs, outputs, or several kinds of branching behaviours like non-determinism, probabilities, exceptions... Since the seminal work of Rutten [129, 131], it has been acknowledged by the community that *coalgebra* is a good categorical umbrella that covers most of these state-based systems – providing e.g. generic results about automata determinisation [81] or minimisation [11]. The branching behaviour of such systems is conveniently represented by *monads*, while other features of interest can often be represented by plain *functors*. To combine effects, coalgebraists use different notions of distributive laws – not only distributive laws between two monads as we studied until now in this thesis, but also distributive laws between a monad and a functor [77, 71, 28]. In this chapter, we study what weak distributive laws can bring to coalgebra theory. We begin by a short exposition of coalgebraic modelling for state-based systems, then define monad-functor (co)weak distributive laws and discuss how they fit into the coalgebraic techniques of generalised determinisation and bisimulation up-to.

Our theoretical contribution consists in extending known results in coalgebra theory in order to accommodate *weak* distributive laws. There are many practical consequences, so the chapter will be importantly illustrated by examples. Our two main applications, alternating automata and probabilistic automata, have long been known as being tedious to model coalgebraically – this is a direct consequence from the fact that there is no distributive law $\mathsf{PP} \to \mathsf{PP}$ nor $\mathsf{DP} \to \mathsf{PD}$. The coalgebraic study of alternating automata – modelled by double covariant powerset – had either to use workarounds in other categories [8] or to make use of an elaborated framework [90] of so-called *forgetful logics*. Recently, authors of [23, 24] successfully managed to model probabilistic automata. They achieved this by redoing all constructions manually, using the *ad hoc* notion of quasi-lax lifting to mitigate the absence of distributive law.

More precisely, we aim at explaining the origin of the following results.

- Alternating automata can be coalgebraically turned into non-deterministic automata in a way that preserves the semantics [90, Example 6.8] i.e. the language of a state in the input alternating automaton coincides with the language of the corresponding state in the output non-deterministic automaton.

- Probabilistic automata can be coalgebraically turned into belief-state transformers in a way that preserves the semantics [23, Theorem 28] i.e. distribution bisimilarity in the input probabilistic automaton coincides with bisimilarity in the output belief-state transformer.

- Bisimulation up-to convex hull is a sound up-to technique with respect to the latter transformation [23, Propositions 30 and 31].

The main application of this chapter is the following:

> Using the monotone weak distributive laws $\mathsf{PP} \to \mathsf{PP}$ and $\mathsf{PD} \to \mathsf{DP}$, we are able to derive systematically these results for both alternating automata and probabilistic automata.

In other words, our work unveils the deep origin of some known constructions, by pinpointing that those constructions are akin to the usual ones, when one replaces a distributive law with a weak distributive law.

4.1 Coalgebraic Modelling

Let A be a fixed set seen as an alphabet. In this chapter, we will model labels of different kinds of state-based systems using the reader functor R on Set from Example 1.11, defined as $RX = X^A$ and $Rf(h) = \lambda a.f(h(a))$. For the sake of readability, the reader functor will be simply denoted by $R = (-)^A$. Note also that we will use the word *automaton* loosely, allowing for infinite state spaces.

Let F be an endofunctor on a category C. An F-*coalgebra* – or just coalgebra when the context is clear – is a pair (X, c) consisting of an object X of C and a morphism $c : X \to FX$. A morphism of F-coalgebras $f : (X, c) \to (Y, d)$ is a morphism $f : X \to Y$ in C such that the following square commutes

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "c"'] & Y \arrow[d, "d"] \\ FX \arrow[r, "Ff"] & FY \end{tikzcd} \tag{4.1}$$

The category of F-coalgebras and their morphisms is denoted by $\mathsf{Coalg}(F)$. There is a functor $U^F : \mathsf{Coalg}(F) \to \mathsf{C}$ that forgets the coalgebra structure as $U^F(X, c) = X$ and $U^F f = f$. Table 4.1 displays a few examples of coalgebras in Set. For some detailed examples, see [132] and also [141] for a survey of coalgebras with a probabilistic flavour. A more complete account of coalgebra theory can be found in [79].

Table 4.1: Examples of coalgebras

functor F	F-coalgebras
P	directed graphs
D	Markov chains
$2 \times -^A$	deterministic automata
$2 \times (P-)^A$	non-deterministic automata
$2 \times (PP-)^A$	alternating automata
$(PD-)^A$	probabilistic automata

Coalgebras in Set can be interpreted as modelling several kinds of state-based transition systems. For an F-coalgebra (X, c), X is seen as the set of states of the system, and $c : X \to FX$ is seen as a transition function which specifies how the system evolves step by step. Notably, the coalgebraic modelling of state-based systems does not directly deal with initial states – these can be implemented later, e.g. when considering semantics. Let us give a brief overview of one specific case using the standard example of non-deterministic automata. Coalgebraically, they are F-coalgebras for the functor $F = 2 \times (P-)^A$, defined by

$$FX = 2 \times (PX)^A \tag{4.2}$$

$$Ff = \mathrm{id}_2 \times (Pf)^A \tag{4.3}$$

where 2 stands for the set of Booleans $2 = \{0, 1\}$. That means a non-deterministic automaton (X, c) consists of a state space X and a function $c : X \to 2 \times (PX)^A$ which can be conveniently decomposed as:

- an output function $c_* : X \to 2$, interpreted as specifying if a state $x \in X$ is terminal ($c_*(x) = 1$) or not ($c_*(x) = 0$);
- for each letter $a \in A$, a transition function $c_a : X \to PX$, interpreted as specifying, for any pair of states $(x, y) \in X^2$, if x can directly access y ($y \in c_a(x)$, also denoted by $x \xrightarrow{a} y$) or not ($y \notin c_a(x)$).

Example 4.1. The table on the left defines a non-deterministic automaton with state space $X = \{x_0, x_1, x_2\}$ and alphabet $A = \{a, b\}$. This automaton is displayed on the right using the standard graphical representation. Note that terminal states are underlined.

state x	$c_*(x)$	$c_a(x)$	$c_b(x)$
x_0	0	$\{x_1, x_2\}$	$\{x_0\}$
x_1	0	$\emptyset$	$\{x_1, x_2\}$
x_2	1	$\emptyset$	$\emptyset$

$$\begin{tikzcd} & x_0 \arrow[loop above, "b"] \arrow[dl, "a"'] \arrow[dr, "a"] & \\ x_1 \arrow[loop below, "b"] \arrow[rr, "b"'] & & \underline{x_2} \end{tikzcd} \tag{4.4}$$

Remark 4.2. If for every letter a and every state x, there is exactly one transition of the form $x \xrightarrow{a} ?$, the automaton is actually *deterministic*. Formally, deterministic automata are coalgebras for the functor $2 \times (-)^A$.

One can see that the behaviour of a non-deterministic automaton is fully encapsulated into the function $c : X \to 2 \times (PX)^A$. A more detailed analysis of the functor $F = 2 \times (P-)^A$ shows that it is the combination of two functors, namely the *machine functor* $2 \times (-)^A$ and the functor P from the powerset monad P. This is a common pattern in coalgebraic modelling, where systems are often of the shape

$$X \to GTX \tag{4.5}$$

or

$$X \to TGX \tag{4.6}$$

for some functor G modelling the purely machine-like behaviour of the system (e.g. output, termination, labels) and some monad T modelling the internal branching type of the system (e.g. non-determinism, probabilities, exception handling). It is interesting to ask how a functor G and a monad T can interact, calling the need for a notion of monad-functor distributive law.

4.2 Implementing Weak Distributive Laws

Let T, S be monads and F be an endofunctor on C. One can define distributive laws of type $\mathsf{T}F \to F\mathsf{T}$ to be natural transformations satisfying only the two diagrams related to T, and similarly for distributive laws of type $F\mathsf{S} \to \mathsf{S}F$. Of course, by forgetting additionally the unit diagram, one can define (co)weak versions of these monad-functor distributive laws. The situation is summed up in Table 4.2.

Table 4.2: Axioms required from ((co)weak) distributive laws

	$\mathsf{TS} \to \mathsf{ST}$			$\mathsf{T}F \to F\mathsf{T}$		$F\mathsf{S} \to \mathsf{S}T$	
axiom	plain	weak	coweak	plain	weak	plain	coweak
(η^{S})	✓	✓	×			✓	×
(μ^{S})	✓	✓	✓			✓	✓
(η^{T})	✓	×	✓	✓	×		
(μ^{T})	✓	✓	✓	✓	✓		

In the continuation of Theorem 1.42, there is a bijective correspondence between distributive laws of type $\mathsf{T}F \to F\mathsf{T}$ and liftings of F to $\mathsf{EM}(\mathsf{T})$, and between distributive laws of type $F\mathsf{S} \to \mathsf{S}F$ and extensions of F to $\mathsf{Kl}(\mathsf{S})$. In the same manner, Theorems 2.5 and 2.29 can be easily adapted to the monad-functor case.

Proposition 4.3. *If idempotents split in C, there is a bijective correspondence between*

- *weak distributive laws of type $\mathsf{T}F \to F\mathsf{T}$*
- *weak liftings of F to $\mathsf{EM}(\mathsf{T})$ i.e. functors $\overline{F} : \mathsf{EM}(\mathsf{T}) \to \mathsf{EM}(\mathsf{T})$ along with natural $\pi : FU^{\mathsf{T}} \to U^{\mathsf{T}}\overline{F}$ and $\iota : U^{\mathsf{T}}\overline{F} \to FU^{\mathsf{T}}$ such that $\pi \circ \iota = 1$*

If idempotents split in $\mathsf{Kl}(\mathsf{S})$, there is a bijective correspondence between

- *coweak distributive laws of type $F\mathsf{S} \to \mathsf{S}F$*
- *coweak extensions of F to $\mathsf{Kl}(\mathsf{S})$ i.e. functors $\underline{F} : \mathsf{Kl}(\mathsf{S}) \to \mathsf{Kl}(\mathsf{S})$ along with natural $\pi : F_{\mathsf{S}}F \nrightarrow \underline{F}F_{\mathsf{S}}$ and $\iota : \underline{F}F_{\mathsf{S}} \nrightarrow F_{\mathsf{S}}F$ such that $\pi \bullet \iota = 1$.*

Proof. One can use the same constructions as in Theorem 2.5 and 2.29, and prove all needed equations in the same way as in these theorems. □

What about iterated monad-functor distributive laws? Let R, S, T be monads on C and F, G, H be endofunctors on C. Keeping our requirement that whenever a composite monad appears, this monad has to come from a plain distributive law, there are basically four different families of iterated monad-functor distributive laws. Namely, there are

- Laws of shape $(\mathbf{S} \circ \mathbf{T})F \to F(\mathbf{S} \circ \mathbf{T})$
- Laws of shape $\mathbf{T}(FG) \to (FG)\mathbf{T}$
- Laws of shape $H(\mathbf{R} \circ \mathbf{S}) \to (\mathbf{R} \circ \mathbf{S})H$
- Laws of shape $(GH)\mathbf{R} \to \mathbf{R}(GH)$

Furthermore, these iterated laws can be plain, weak or coweak, depending of the input data. In the continuation of Table 2.2 presenting possible combinations for monad-monad laws, we provide in Table 4.3 an overview of possible combinations for monad-functor laws.

Table 4.3: Iterated monad-functor (co)weak distributive laws. Given data in the first four columns, one can obtain a law as in the last column

$\mathbf{TS} \to \mathbf{ST}$	$\mathbf{S}F \to F\mathbf{S}$	$\mathbf{T}F \to F\mathbf{T}$		$(\mathbf{S} \circ \mathbf{T})F \to F(\mathbf{S} \circ \mathbf{T})$
plain	plain	plain	Y-B	plain
plain	weak	weak	Y-B	weak
$\mathbf{T}G \to G\mathbf{T}$		$\mathbf{T}F \to F\mathbf{T}$		$\mathbf{T}(FG) \to (FG)\mathbf{T}$
plain		plain		plain
weak		weak		weak
$H\mathbf{S} \to \mathbf{S}H$	$\mathbf{SR} \to \mathbf{RS}$	$H\mathbf{R} \to \mathbf{R}H$		$H(\mathbf{R} \circ \mathbf{S}) \to (\mathbf{R} \circ \mathbf{S})H$
plain	plain	plain	Y-B	plain
coweak	plain	coweak	Y-B	coweak
	$G\mathbf{R} \to \mathbf{R}G$	$H\mathbf{R} \to \mathbf{R}H$		$(GH)\mathbf{R} \to \mathbf{R}(GH)$
	plain	plain		plain
	coweak	coweak		coweak

We are especially interested in the $\mathbf{T}(FG) \to (FG)\mathbf{T}$ section of Table 4.3, because these laws are relevant with respect to generalised determinisation. In particular, note the following examples.

Example 4.4. Let λ be the unique monotone weak distributive law of type $\mathbf{PP} \to \mathbf{PP}$ described in Chapter 3 and $\tau : \mathbf{PR} \to \mathbf{RP}$ be the distributive law of Example 1.33. Forgetting one of the monad structures, these are a weak $\lambda : \mathbf{P}P \to P\mathbf{P}$ and a plain $\tau : \mathbf{P}(-^A) \to (\mathbf{P}-)^A$. Without needing any Yang-Baxter diagram, we can combine them to form a weak distributive law $\mathbf{P}(P-)^A \to (P\mathbf{P}-)^A$.

Example 4.5. Similarly, from the unique monotone weak distributive law $\lambda : \mathbf{DP} \to \mathbf{PD}$ of Chapter 3 and the distributive law $\tau : \mathbf{DR} \to \mathbf{RD}$ of Example 1.34, we can directly form a weak distributive law $\mathbf{D}(P-)^A \to (P\mathbf{D}-)^A$.

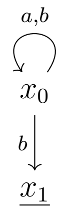
4.3 Generalised Determinisation

Determinisation, also called powerset construction, is a well-known procedure that builds, out of any non-deterministic automaton $c : X \to 2 \times (PX)^A$, a deterministic automaton $c^\# : PX \to 2 \times (PX)^A$ with state space PX. The determinised $c^\#$ is defined, for every state $U \in PX$, by

- ${c^\#}_*(U) = \max\{c_*(x) \mid x \in U\}$
- for every letter $a \in A$, ${c^\#}_a(U) = \bigcup_{x \in U} c_a(x)$

Example 4.6. Consider the non-deterministic automaton c with state space $X = \{x_0, x_1\}$ over the alphabet $A = \{a, b\}$ depicted below.

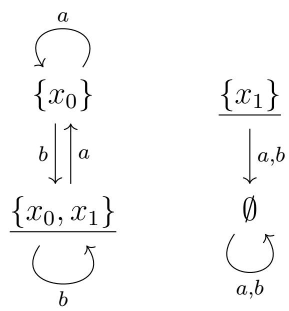
state x	$c_*(x)$	$c_a(x)$	$c_b(x)$
x_0	0	$\{x_0\}$	$\{x_0, x_1\}$
x_1	1	$\emptyset$	$\emptyset$



(4.7)

Its determinisation is the following deterministic automaton with state space PX.

state U	${c^\#}_*(U)$	${c^\#}_a(U)$	${c^\#}_b(U)$
$\emptyset$	0	$\emptyset$	$\emptyset$
$\{x_0\}$	0	$\{x_0\}$	$\{x_0, x_1\}$
$\{x_1\}$	1	$\emptyset$	$\emptyset$
$\{x_0, x_1\}$	1	$\{x_0\}$	$\{x_0, x_1\}$



(4.8)

It turns out that this construction can be captured coalgebraically using a distributive law [138, 139, 81]. Concretely, there is a join-semilattice structure on $2 \times (PX)^A$ arising from a distributive law, and the universal morphism from the free semilattice PX making the following diagram commute is $c^\#$.

$$\begin{tikzcd} X \arrow[r, "c"] \arrow[d, "\eta^{\mathsf{P}}_X"'] & 2 \times (PX)^A \\ PX \arrow[ur, dashed, "c^\#"'] & \end{tikzcd} \tag{4.9}$$

This situation of turning non-deterministic automata into deterministic automata is abstracted into a functor

$$\mathsf{Coalg}(FT) \to \mathsf{Coalg}(F)$$

where F generalises $2 \times (-)^A$ and $\mathbf{T}$ generalises $\mathbf{P}$. We give a presentation of this so-called *generalised determinisation*, or *generalised powerset construction*, close to the general account of [81]. In the context of coalgebras, saying that T lifts to a functor $\overline{T} : \mathsf{Coalg}(FT) \to \mathsf{Coalg}(F)$ means that the following square commutes:

$$\begin{tikzcd}
\mathsf{Coalg}(FT) \arrow[r, "\overline{T}"] \arrow[d, "U^{FT}"'] & \mathsf{Coalg}(F) \arrow[d, "U^F"] \\
\mathsf{C} \arrow[r, "T"] & \mathsf{C}
\end{tikzcd}$$

The following proposition is actually stronger. It explains how the lifted functor also *factors* through $\mathsf{EM}(\mathbf{T})$, as hinted in diagram (4.9).

Proposition 4.7 ([81]). *Let $\lambda : \mathbf{T}F \to F\mathbf{T}$ be a distributive law and $\overline{F}$ be the corresponding lifting of F to $\mathsf{EM}(\mathbf{T})$. Then one can lift the Eilenberg-Moore functors $F^{\mathbf{T}} : \mathsf{C} \to \mathsf{EM}(\mathbf{T})$ and $U^{\mathbf{T}} : \mathsf{EM}(\mathbf{T}) \to \mathsf{C}$ to categories of coalgebras as follows: there are functors $\widehat{F^{\mathbf{T}}}$, $\widehat{U^{\mathbf{T}}}$ such that the following diagram commutes*

$$\begin{tikzcd}
\mathsf{Coalg}(FT) \arrow[r, "\widehat{F^{\mathbf{T}}}"] \arrow[d, "U^{FT}"'] & \mathsf{Coalg}(\overline{F}) \arrow[r, "\widehat{U^{\mathbf{T}}}"] \arrow[d, "U^{\overline{F}}"'] & \mathsf{Coalg}(F) \arrow[d, "U^F"] \\
\mathsf{C} \arrow[r, "F^{\mathbf{T}}"] & \mathsf{EM}(\mathbf{T}) \arrow[r, "U^{\mathbf{T}}"] & \mathsf{C}
\end{tikzcd} \tag{4.10}$$

Proof. Recall that $T = U^{\mathbf{T}}F^{\mathbf{T}}$. For any FT-coalgebra (X, c), define $\widehat{F^{\mathbf{T}}}(X, c)$ to be

$$F^{\mathbf{T}}X \xrightarrow{F^{\mathbf{T}}c} F^{\mathbf{T}}FTX = F^{\mathbf{T}}FU^{\mathbf{T}}F^{\mathbf{T}}X = F^{\mathbf{T}}U^{\mathbf{T}}\overline{F}F^{\mathbf{T}}X \xrightarrow{\epsilon^{\mathbf{T}}_{\overline{F}F^{\mathbf{T}}X}} \overline{F}F^{\mathbf{T}}X \tag{4.11}$$

and $\widehat{F^{\mathbf{T}}}f = F^{\mathbf{T}}f$ for any FT-coalgebra morphism f. This yields a functor $\widehat{F^{\mathbf{T}}}$ such that $U^{\overline{F}}\widehat{F^{\mathbf{T}}} = F^{\mathbf{T}}U^{FT}$. Now, let $((X, x), d)$ be an $\overline{F}$-coalgebra in $\mathsf{EM}(\mathbf{T})$. This means (X, x) is a $\mathbf{T}$-algebra and $d : (X, x) \to \overline{F}(X, x)$ is a $\mathbf{T}$-algebra morphism. Define $\widehat{U^{\mathbf{T}}}((X, x), d)$ to be

$$X = U^{\mathbf{T}}(X, x) \xrightarrow{U^{\mathbf{T}}d} U^{\mathbf{T}}\overline{F}(X, x) = FU^{\mathbf{T}}(X, x) = FX \tag{4.12}$$

and $\widehat{U^{\mathbf{T}}}f = U^{\mathbf{T}}f$ for any $\overline{F}$-coalgebra morphism f. This yields a functor $\widehat{U^{\mathbf{T}}}$ such that $U^F\widehat{U^{\mathbf{T}}} = U^{\mathbf{T}}U^{\overline{F}}$. □

The functor $\widehat{U^{\mathbf{T}}}\widehat{F^{\mathbf{T}}}$ as defined in the proof of Proposition 4.7 performs generalised determinisation by transforming an FT-coalgebra

$$X \xrightarrow{c} FTX \tag{4.13}$$

into the F-coalgebra

$$TX \xrightarrow{Tc} TFTX \xrightarrow{\lambda_{TX}} FTTX \xrightarrow{F\mu^{\mathbf{T}}_X} FTX \tag{4.14}$$

Equation (4.14) is the action of the $\mathbf{T}$-algebra from equation (4.11). One could perform this construction functorially with any natural transformation $\lambda : TF \to FT$, but when $\lambda : \mathbf{T}F \to F\mathbf{T}$ is a distributive law, Proposition 4.7 ensures that this procedure is sound, in the sense that it factors through $\mathsf{EM}(\mathbf{T})$ (see also [81]).

By instantiating the generalised determinisation procedure with $F = 2 \times (-)^A$, $\mathbf{T} = \mathbf{P}$, and the distributive law $\lambda : \mathbf{P}(2 \times -^A) \to 2 \times (\mathbf{P}-)^A$ defined by

$$\lambda_X(U) = (\max\{b \mid (b,h) \in U\}, \lambda a.\{h(a) \mid (b,h) \in U\}) \tag{4.15}$$

one retrieves the usual determinisation $(X, c) \mapsto (PX, c^{\#})$ of non-deterministic automata. We will keep using the notation $(-)^{\#}$ to denote *generalised* determinisation for further examples, i.e. $(-)^{\#} = \widehat{U^{\mathbf{T}}}\widehat{F^{\mathbf{T}}}$.

As may be guessed by looking at the proof of Proposition 4.7, generalised determinisation can be adapted without much effort to the case when $\lambda : \mathbf{T}F \to F\mathbf{T}$ is only a weak distributive law.

Proposition 4.8. *Assume idempotents split in* C. *Let* $\lambda : \mathbf{T}F \to F\mathbf{T}$ *be a weak distributive law and* $\overline{F}$ *be the corresponding weak lifting of* F *to* $\mathsf{EM}(\mathbf{T})$, *with* $\pi : FU^{\mathbf{T}} \to U^{\mathbf{T}}\overline{F}$ *and* $\iota : U^{\mathbf{T}}\overline{F} \to FU^{\mathbf{T}}$. *Then there are functors* $\widehat{F^{\mathbf{T}}}$, $\widehat{U^{\mathbf{T}}}$ *such that the following diagram commutes*

$$\begin{tikzcd}
\mathsf{Coalg}(FT) \arrow[r, "\widehat{F^{\mathbf{T}}}"] \arrow[d, "U^{FT}"'] & \mathsf{Coalg}(\overline{F}) \arrow[r, "\widehat{U^{\mathbf{T}}}"] \arrow[d, "U^{\overline{F}}"'] & \mathsf{Coalg}(F) \arrow[d, "U^F"] \\
\mathsf{C} \arrow[r, "F^{\mathbf{T}}"] & \mathsf{EM}(\mathbf{T}) \arrow[r, "U^{\mathbf{T}}"] & \mathsf{C}
\end{tikzcd} \tag{4.16}$$

Proof. Just define $\widehat{F^{\mathbf{T}}}(X, c)$ to be

$$F^{\mathbf{T}}X \xrightarrow{F^{\mathbf{T}}c} F^{\mathbf{T}}FTX \xrightarrow{F^{\mathbf{T}}\pi_{F^{\mathbf{T}}X}} F^{\mathbf{T}}U^{\mathbf{T}}\overline{F}F^{\mathbf{T}}X \xrightarrow{\epsilon^{\mathbf{T}}_{\overline{F}F^{\mathbf{T}}X}} \overline{F}F^{\mathbf{T}}X \tag{4.17}$$

and $\widehat{U^{\mathbf{T}}}((X, x), d)$ to be

$$X = U^{\mathbf{T}}(X, x) \xrightarrow{U^{\mathbf{T}}d} U^{\mathbf{T}}\overline{F}(X, x) \xrightarrow{\iota_{(X,x)}} FU^{\mathbf{T}}(X, x) = FX \tag{4.18}$$

□

The determinisation functor can be computed with the same expression as for plain distributive laws:

Proposition 4.9. *Under the assumptions of Proposition 4.8, the functor* $\widehat{U^{\mathbf{T}}}\widehat{F^{\mathbf{T}}}$ *maps an* FT*-coalgebra*

$$X \xrightarrow{c} FTX \tag{4.19}$$

to the F*-coalgebra*

$$TX \xrightarrow{Tc} TFTX \xrightarrow{\lambda_{TX}} FTTX \xrightarrow{F\mu^{\mathbf{T}}_X} FTX \tag{4.20}$$

Proof. Expressions of $\widehat{F^{\mathbf{T}}}$ and $\widehat{U^{\mathbf{T}}}$ show that $\widehat{U^{\mathbf{T}}}\widehat{F^{\mathbf{T}}}(X, c)$ is

$$TX \xrightarrow{Tc} TFTX \xrightarrow{T\pi_{F^{\mathbf{T}}X}} TU^{\mathbf{T}}\overline{F}F^{\mathbf{T}}X \xrightarrow{U^{\mathbf{T}}\epsilon^{\mathbf{T}}_{\overline{F}F^{\mathbf{T}}X}} U^{\mathbf{T}}\overline{F}F^{\mathbf{T}}X \xrightarrow{\iota_{F^{\mathbf{T}}}} FTX \tag{4.21}$$

The following commutative diagram suffices to prove that F-coalgebras described in (4.20) and (4.21) coincide.

$$\begin{tikzcd}
TFT \arrow[rr, equal] \arrow[dr, "TF\eta^{\mathbf{T}}T"'] \arrow[dddd, "\lambda T"'] & & TFT \arrow[r, "T\pi F^{\mathbf{T}}"] & TU^{\mathbf{T}}\overline{F}F^{\mathbf{T}} \arrow[ddd, "U^{\mathbf{T}}\epsilon^{\mathbf{T}}\overline{F}F^{\mathbf{T}}"] \\
& TFTT \arrow[ur, "TFU^{\mathbf{T}}\epsilon^{\mathbf{T}}F^{\mathbf{T}}"'] \arrow[d, "T\pi F^{\mathbf{T}}T"'] & & \\
& TU^{\mathbf{T}}\overline{F}F^{\mathbf{T}}T \arrow[l, phantom, "\text{def}"] \arrow[uurr, "TU^{\mathbf{T}}\overline{F}\epsilon^{\mathbf{T}}F^{\mathbf{T}}"'] \arrow[d, "U^{\mathbf{T}}\epsilon^{\mathbf{T}}\overline{F}F^{\mathbf{T}}T"] & \text{nat} & \\
& U^{\mathbf{T}}\overline{F}F^{\mathbf{T}}T \arrow[rr, "U^{\mathbf{T}}\overline{F}\epsilon^{\mathbf{T}}F^{\mathbf{T}}"'] \arrow[dl, "\iota F^{\mathbf{T}}T"] & & U^{\mathbf{T}}\overline{F}F^{\mathbf{T}} \arrow[d, "\iota F^{\mathbf{T}}"] \\
FTT \arrow[rrr, "F\mu^{\mathbf{T}} = FU^{\mathbf{T}}\epsilon^{\mathbf{T}}F^{\mathbf{T}}"'] & & & FT
\end{tikzcd} \tag{4.22}$$

The left pentagon commutes because of the construction of a weak distributive law with respect to a weak lifting (see equation (2.7)). The *nat*-marked square commutes by naturality of $\epsilon^{\mathbf{T}} : F^{\mathbf{T}}U^{\mathbf{T}} \to 1$. The top triangle commutes by the monad property $\mu^{\mathbf{T}} \circ \eta^{\mathbf{T}}T = 1$. Finally, the two unmarked squares commute by naturality of ι and π. ☐

A further remark is that the first step of this stronger generalised determinisation can itself factor through another interesting category of coalgebras. Recall that any weak distributive law of type $\mathbf{TS} \to \mathbf{ST}$ yields a weak composite monad $\mathbf{S} \bullet \mathbf{T}$ with functor $S \bullet T = U^{\mathbf{T}}\overline{S}F^{\mathbf{T}}$. In the same way, any weak distributive law of type $\mathbf{T}F \to F\mathbf{T}$ yields a weak composite functor $F \bullet T = U^{\mathbf{T}}\overline{F}F^{\mathbf{T}}$, and

Proposition 4.10. *Under the assumptions of Proposition 4.9, the functor* $\widehat{F^{\mathbf{T}}}$ *factors through* $\mathsf{Coalg}(F \bullet T)$ *as in*

$$\begin{tikzcd}
\mathsf{Coalg}(FT) \arrow[r] \arrow[d] \arrow[rr, bend left, "\widehat{F^{\mathbf{T}}}"] & \mathsf{Coalg}(F \bullet T) \arrow[r] \arrow[d] & \mathsf{Coalg}(\overline{F}) \arrow[d] \\
\mathsf{C} \arrow[r, "\mathrm{id}"] \arrow[rr, bend right, "F^{\mathbf{T}}"'] & \mathsf{C} \arrow[r, "F^{\mathbf{T}}"] & \mathsf{EM}(\mathbf{T})
\end{tikzcd} \tag{4.23}$$

Proof. The functor $\mathsf{Coalg}(FT) \to \mathsf{Coalg}(F \bullet T)$ simply maps (X, c) to $(X, \pi_{F^{\mathbf{T}}X} \circ c)$. The functor $\mathsf{Coalg}(F \bullet T) \to \mathsf{Coalg}(\overline{F})$ maps (X, d) to $(F^{\mathbf{T}}X, \epsilon^{\mathbf{T}}_{\overline{F}F^{\mathbf{T}}X} \circ F^{\mathbf{T}}d)$. $\square$

To sum up, generalised determinisation with respect to a weak distributive law can be decomposed into three main steps as in

$$\begin{array}{ccccccc}
\mathsf{Coalg}(FT) & \longrightarrow & \mathsf{Coalg}(F \bullet T) & \longrightarrow & \mathsf{Coalg}(\overline{F}) & \longrightarrow & \mathsf{Coalg}(F) \\
(X, c) & \longmapsto & (X, c^{\dagger}) & \longmapsto & ((TX, \mu^{\mathbf{T}}_X), c^{\#}) & \longmapsto & (TX, c^{\#})
\end{array} \tag{4.24}$$

The well-established terminology *generalised determinisation* can be confusing in the case where F features some kind of non-determinism, because then the coalgebra $TX \to FTX$ is not *deterministic* in the classical sense. Similar problems arise with the terminology *generalised powerset construction* when there is no actual powerset. Notably, this terminology clash is annoying for our two main applications, which are alternating automata and probabilistic automata. In [81], authors use the name *non-determinisation* to denote that they are performing a generalised determinisation that does not output a deterministic system. We find this terminology still confusing. In examples, we will use the non-standard terminology *algebraic expansion* to denote generalised determinisation. This terminology has two advantages:

- it does not mention determinism nor powerset at all;
- it insists on the fact that whereas the procedure simplifies the branching type of the system, the counterpart is to ensure an algebraic structure on the new expanded state space.

4.3.1 Alternating Automata

Alternating state-based systems have been introduced in [33]. Their main feature, called *alternation*, is to combine two sorts of non-deterministic choice. Some states are existential, meaning that the system transitions non-deterministically into *at least one* accessible state, as in standard non-deterministic automata. Some states are universal, meaning that the system transitions non-deterministically into *all* accessible states. In the coalgebraic practice [90, 8], these two sorts of choice are hidden into the transitions and structured into two distincts layers. Formally, alternating automata are coalgebras for the functor $2 \times (PP-)^A$. The absence of a distributive law $\mathsf{PP} \to \mathsf{PP}$ makes difficult the coalgebraic study of alternating automata, and several algebraic expansions have been devised. None of them is fully canonical, either because the intended semantics is not respected [81], because there is a detour outside of the category of sets [8], or because the procedure does not rely on a distributive law [90].

Language. Let $c = \langle c_*, (c_a)_{a\in A}\rangle : X \to 2 \times (PPX)^A$ be an alternating automaton with output denoted by $c_* : X \to 2$ and transitions denoted by $c_a : X \to PPX$. The standard semantics of c is a map $[\![-]\!] : X \to 2^{A^*}$ mapping a state x, seen as the initial state, to the language it generates. This map is defined inductively by

$$[\![x]\!](\varepsilon) = c_*(x) \tag{4.25}$$

$$[\![x]\!](aw) = \bigvee_{V\in c_a(x)} \bigwedge_{y\in V} [\![y]\!](w) \qquad \text{where } a \in A, w \in A^* \tag{4.26}$$

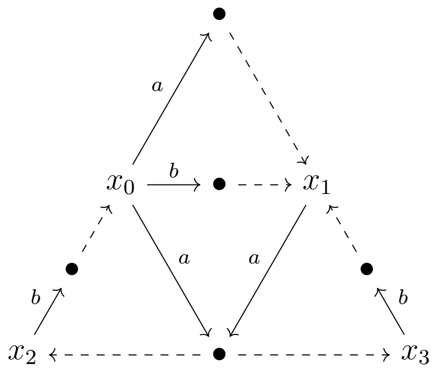
In words, a state x accepts the empty word iff x is a final state, and x accepts a word aw iff x has at least one a-accessible subset U such that every state in U accepts w.

Outputs were temporarily included in the presentation to provide semantic insights. With this picture in mind, and for the sake of simplicity, we forget about outputs to concentrate on the branching structure. From now on, *alternating automata* are $(PP-)^A$ coalgebras. Let $c : X \to (PPX)^A$ be an alternating automaton. The assertion $V \in c_a(x)$ will be also denoted by $x \xrightarrow{a} V$.

Example 4.11. The table on the left defines an alternating automaton with states $X = \{x_0, x_1, x_2, x_3\}$ and alphabet $A = \{a, b\}$. Equivalently, this automaton is represented on the right. Each a-labelled arrow points to a symbol $\bullet$ representing a set

$V \in c_a(x)$, whose elements are pointed by dashed arrows.

state x	$c_a(x)$	$c_b(x)$
x_0	$\{\{x_1\}, \{x_2, x_3\}\}$	$\{\{x_1\}\}$
x_1	$\{\{x_2, x_3\}\}$	$\emptyset$
x_2	$\emptyset$	$\{\{x_0\}\}$
x_3	$\emptyset$	$\{\{x_1\}\}$



(4.27)

It is apparent that PP is a composite branching type using twice the powerset monad $\mathbf{P}$, the outer being interpreted disjunctively, and the inner conjunctively. According to Proposition 4.9, the algebraic expansion of an alternating automaton with respect to some (weak or plain) distributive law $\lambda : \mathbf{P}(P-)^A \to (P\mathbf{P}-)^A$ can be directly computed by the formula

$$c^\# = PX \xrightarrow{Pc} P((PPX)^A) \xrightarrow{\lambda_{PX}} (PPPX)^A \xrightarrow{(P\mu^{\mathbf{P}}_X)^A} (PPX)^A \tag{4.28}$$

We can list three different algebraic expansion procedures for alternating automata using three different laws, respectively using:

(0) the weak distributive law $\lambda^{(0)}$ from Example 2.38, defined by

$$\lambda^{(0)}_X(\mathcal{U}) = \lambda a. \left\{ \bigcup_{h \in \mathcal{U}} h(a) \right\} \tag{4.29}$$

The algebraic expansion of c via $\lambda^{(0)}$ is expressed as

$$c^\#{}_a(U) = \left\{ \bigcup_{x \in U} \bigcup c_a(x) \right\} \tag{4.30}$$

which means that $c^\#$ has exactly one a-transition for every letter $a \in A$, i.e., it actually is a *deterministic* automaton on PX. The unique transition out of a state U is given by

$$U \overset{a}{\mapsto} \{\text{states } y \text{ such that } x \xrightarrow{a} V \text{ for some } x \in U \text{ and some } V \ni y\}$$

From every subset, the deterministic automaton $c^\#$ transitions to the set of all states directly visible by c from elements of this subset. This algebraic expansion amounts to the procedure consisting in (i) collapsing the two layers of non-determinism to get a non-deterministic automaton and (*ii*) performing a standard determinisation. A lot of information is lost – this is a direct consequence of the fact that the weak lifting of $\lambda^{(0)}$ is trivial in the sense of Proposition 2.12 and Example 2.14.

(1) the plain distributive law $\lambda^{(1)}$ from Example 2.39, defined by

$$\lambda^{(1)}_X(\mathcal{U}) = \lambda a. \left\{ \{x\} \mid x \in \bigcup_{h \in \mathcal{U}} h(a) \right\} \tag{4.31}$$

Note that this law is a coweak distributive law between monads $\mathbf{P}$ and $(\mathbf{R} \circ \mathbf{P})$, but it becomes a plain distributive law between the monad $\mathbf{P}$ and the functor $(P-)^A$. The algebraic expansion of c via $\lambda^{(1)}$ can be expressed as

$$c^{\#}{}_a(U) = \bigcup_{x \in U} c_a(x) \tag{4.32}$$

or otherwise said, we have

$$(U \xrightarrow{a} V \text{ in } c^{\#}) \text{ if and only if } (x \xrightarrow{a} V \text{ in } c, \text{ for some } x \in U)$$

From every subset, the non-deterministic automaton $c^{\#}$ collects all transitions that can be performed in c from an element of this subset. Under the natural isomorphism $P(A \times X) \simeq (PX)^A$, the distributive law $\lambda^{(1)}$ can be rewritten as having type

$$\mathbf{P}P(A \times -) \to P(A \times \mathbf{P}-) \tag{4.33}$$

A little computation shows that this corresponds exactly to the law of Jacobs, Silva and Sokolova in [81, Section 5.3] – and indeed, their algebraic expansion of alternating automata coincides with ours.

(2) the weak distributive law $\lambda^{(2)}$ from Example 4.4, defined by

$$\lambda^{(2)}_X(\mathcal{U}) = \lambda a. \left\{ V \subseteq X \mid V \subseteq \bigcup_{h \in \mathcal{U}} h(a) \text{ and } \forall h \in \mathcal{U}. V \cap h(a) \neq \emptyset \right\} \tag{4.34}$$

The determinisation of c via $\lambda^{(2)}$ is expressed as

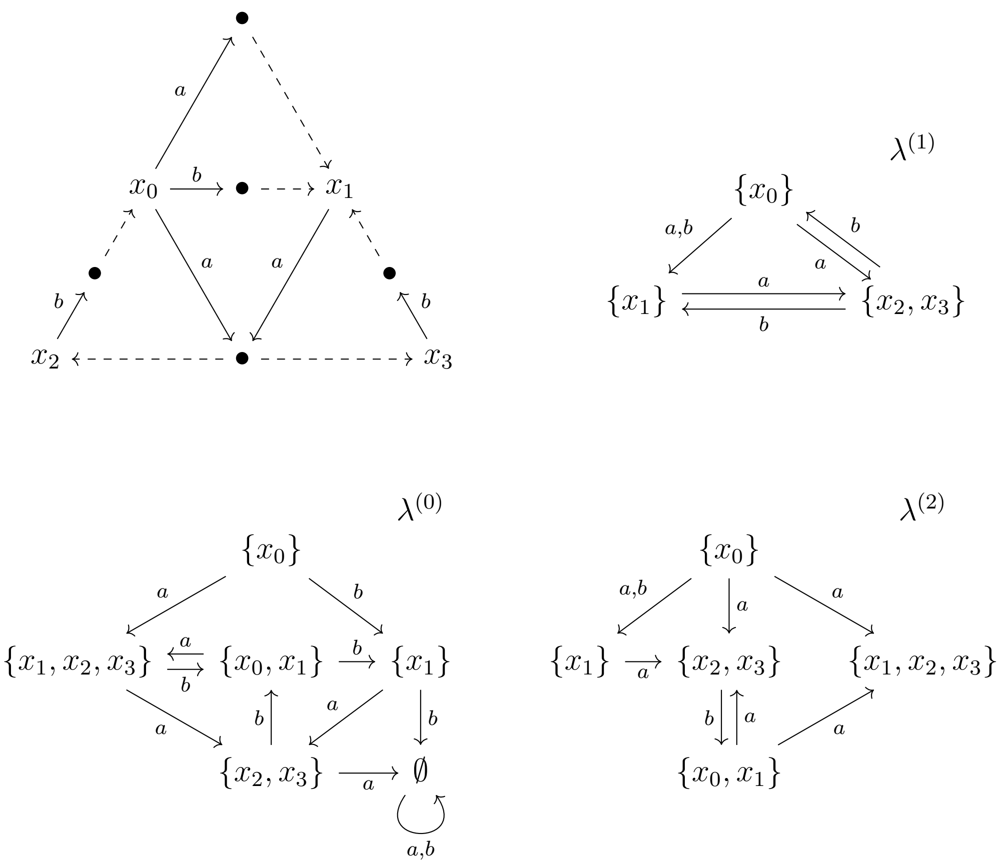
$$c^{\#}{}_a(U) = \left\{ \bigcup \mathcal{V} \mid \mathcal{V} \subseteq \bigcup_{x \in U} c_a(x) \text{ and } \forall x \in U. \mathcal{V} \cap c_a(x) \neq \emptyset \right\} \tag{4.35}$$

More concretely, transitions of $c^{\#}$ of the form $U \xrightarrow{a}$ (something) are generated by the following process. For every $x \in U$, pick any non-zero number of Vs such that $x \xrightarrow{a} V$ in c, and let V_x be the union of these selected Vs. Then $U \xrightarrow{a} \bigcup_{x \in U} V_x$. This transformation of alternating automata into non-deterministic automata was already considered by Klin and Rot in [90], where the crucial natural transformation $PP \to PP$ was merely identified as a non-distributive law. They stated

> *We do not know how to model precisely the standard transformation of alternating automata into nondeterministic ones in our framework.* [90, p.22]

Now we can argue that Klin and Rot algebraic expansion *should actually be* the standard transformation of alternating automata into nondeterministic ones, as it comes directly from the unique monotone weak distributive law $\mathbf{PP} \to \mathbf{PP}$. This gives an intuition about why in [90] the semantics is preserved, and we will see in Section 4.4 that it also unlocks the use of bisimulations up-to.

Example 4.12. Let us display these three algebraic expansions for the alternating automaton from Example 4.11. Only the part accessible from the state $\{x_0\}$ is represented.



The expansion via $\lambda^{(0)}$ is not really interesting in itself. Graphically, it amounts to considering all dashed arrows as non-dashed arrows, then computing the determinised automaton. A more fruitful idea is to compare the expansions using $\lambda^{(1)}$ and $\lambda^{(2)}$. Starting from $\{x_0\}$, the two resulting non-deterministic automata have, at first sight, a similar behaviour. But already, the $\lambda^{(2)}$ one has a transition $\{x_0\} \xrightarrow{a} \{x_1, x_2, x_3\}$ obtained from the union of the transitions $x_0 \xrightarrow{a} \{x_1\}$ and $x_0 \xrightarrow{a} \{x_2, x_3\}$ in the alternating automaton, whereas the $\lambda^{(1)}$ one does not. Another interesting difference arises when looking at what the transitions $x_2 \xrightarrow{b} \{x_0\}$ and $x_3 \xrightarrow{b} \{x_1\}$ become. Using $\lambda^{(1)}$, they give rise to two different transitions $\{x_2, x_3\} \xrightarrow{b} \{x_0\}$ and $\{x_2, x_3\} \xrightarrow{b} \{x_1\}$. Using $\lambda^{(2)}$, they are combined into a single transition $\{x_2, x_3\} \xrightarrow{b} \{x_0, x_1\}$, which is arguably a more desirable behaviour. Indeed, a good notion of composite transition should intuitively mix both the starting states and the ending states.

When using outputs – and appropriate natural transformations making them interact with the rest of the coalgebra – the $\lambda^{(2)}$ algebraic expansion gives rise to the intended semantics $[\![-]\!]$, as described in [90, Examples 4.5 and 6.8]. This is not the case for the $\lambda^{(0)}$ and $\lambda^{(1)}$ algebraic expansions, for reasons that we detail now.

A logical choice. In light of the two-fold purpose of P in the semantics of alternating automata, we provide an interpretation of our three laws $\mathsf{PP} \to \mathsf{PP}$ in terms of propositional logic. When interpreted disjunctively (resp. conjunctively), the powerset will be denoted dy $P_\vee$ (resp. $P_\wedge$). Let X be a set whose elements are considered as propositional variables. As visible in equation (4.26), an element $\mathcal{U} \in P_\vee P_\wedge X$ can be thought of as a formula in disjunctive normal form

$$\bigvee_{U \in \mathcal{U}} \bigwedge_{x \in U} x \tag{4.36}$$

while an element $\mathcal{U} \in P_\wedge P_\vee X$ can be thought of as a formula in conjunctive normal form

$$\bigwedge_{U \in \mathcal{U}} \bigvee_{x \in U} x \tag{4.37}$$

These will be simply denoted by $\bigvee \bigwedge \mathcal{U}$ and $\bigwedge \bigvee \mathcal{U}$. What is expected of a law λ : $\mathsf{PP} \to \mathsf{PP}$ is, in the context of alternating automata, to transform a conjunctive normal form into a disjunctive normal form:

$$\lambda_X : P_\wedge P_\vee X \to P_\vee P_\wedge X \tag{4.38}$$

Let us examine how $\lambda^{(0)}$, $\lambda^{(1)}$ and $\lambda^{(2)}$ carry out this task. Equivalence of propositional formulas is denoted by $\equiv$.

- First note that the natural transformation id_{PP} does not satisfy any of the distributive law axioms. In logical terms, it swaps conjunctions and disjunctions:

$$\bigwedge\bigvee\mathcal{U} \mapsto \bigvee\bigwedge\mathcal{U} \tag{4.39}$$

- The trivial weak distributive law underlying $\lambda^{(0)}$ maps $\mathcal{U}$ to $\{\bigcup\mathcal{U}\}$. In logical terms, and in compliance with Proposition 2.12, the disjunctive interpretation $P_\vee$ disappears:

$$\bigwedge\bigvee\mathcal{U} \mapsto \bigvee\bigwedge\left\{\bigcup\mathcal{U}\right\} \equiv \bigwedge\left(\bigcup\mathcal{U}\right) \equiv \bigwedge\bigwedge\mathcal{U} \tag{4.40}$$

- The trivial coweak distributive law underlying $\lambda^{(1)}$ maps $\mathcal{U}$ to $\{\{x\} \mid x \in \bigcup\mathcal{U}\}$. In logical terms, and in compliance with Proposition 2.31, the conjunctive interpretation $P_\wedge$ disappears:

$$\bigwedge\bigvee\mathcal{U} \mapsto \bigvee\bigwedge\left\{\{x\} \mid x \in \bigcup\mathcal{U}\right\} \equiv \bigvee\left(\bigcup\mathcal{U}\right) \equiv \bigvee\bigvee\mathcal{U} \tag{4.41}$$

- The monotone weak distributive law underlying $\lambda^{(2)}$ maps $\mathcal{U}$ to

$$\left\{V \subseteq \bigcup\mathcal{U} \mid \forall U \in \mathcal{U}, U \cap V \neq \emptyset\right\}$$

  In logical terms, a short calculation shows that

$$\bigwedge\bigvee\mathcal{U} \mapsto \bigvee\bigwedge\left\{V \subseteq \bigcup\mathcal{U} \mid \forall U \in \mathcal{U}, U \cap V \neq \emptyset\right\} \equiv \bigwedge\bigvee\mathcal{U} \tag{4.42}$$

($\Rightarrow$) If the disjunctive normal form is satisfied, then there is a $V \subseteq \bigcup\mathcal{U}$ intersecting every element of $\mathcal{U}$, such that all variables in V are satisfied. For every $U \in \mathcal{U}$, we can pick in V some satisfied $x_U \in U$, so the conjunctive normal form is satisfied.

($\Leftarrow$) If the conjunctive normal form is satisfied, we can pick a satisfied $x_U \in U$ for every $U \in \mathcal{U}$. Then the set $V = \{x_U \mid U \in \mathcal{U}\}$ witnesses that the disjunctive normal form is satisfied.

Therefore, the laws $\lambda^{(0)}$ and $\lambda^{(1)}$ alterate the truth value of the input formula, whereas the law $\lambda^{(2)}$ transforms the input conjunctive normal form into an equivalent disjunctive normal form. This is concrete evidence that the law $\lambda^{(2)}$ is the logical choice for devising an algebraic expansion of alternating automata.

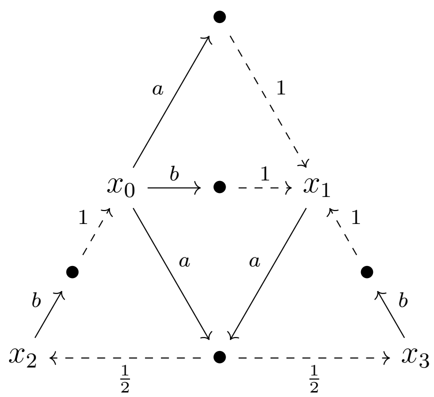
4.3.2 Probabilistic Automata

There are many notions of automata involving probabilities, for which we refer to the survey [141]. Following [23], we define probabilistic automata as $(PD-)^A$ coalgebras. These are also known under the name of *simple Segala systems* [141, 81]. Semantically, both their languages [150] and their bisimulations [23, 24] have been studied – we focus only on the latter, in the next section.

Let $c : X \to (PDX)^A$ be a probabilistic automaton. For any letter $a \in A$, the transition map $c_a : X \to PDX$ assigns to every state x some (finitely supported) distributions on states. These systems combine non-determinism – in the sense that a state can transition to any number of distributions – and probability.

Example 4.13. By assigning a probability to every dashed arrow in Example 4.11, we obtain the following example of a probabilistic automaton.

state x	$c_a(x)$	$c_b(x)$
x_0	$\{1 \cdot x_1, \frac{1}{2} \cdot x_2 + \frac{1}{2} \cdot x_3\}$	$\{1 \cdot x_1\}$
x_1	$\{\frac{1}{2} \cdot x_2 + \frac{1}{2} \cdot x_3\}$	$\emptyset$
x_2	$\emptyset$	$\{1 \cdot x_0\}$
x_3	$\emptyset$	$\{1 \cdot x_1\}$



(4.43)

According to Proposition 4.9, the algebraic expansion of a probabilistic automaton with respect to some (weak or plain) distributive law $\lambda : \mathbf{D}(P-)^A \to (P\mathbf{D}-)^A$ can be directly computed by the formula

$$c^\# = DX \xrightarrow{Dc} D((PDX)^A) \xrightarrow{\lambda_{DX}} (PDDX)^A \xrightarrow{(P\mu_X^{\mathbf{D}})^A} (PDX)^A \tag{4.44}$$

Let us consider the two following algebraic expansions of probabilistic automata, with

(1*) the plain distributive law $\lambda^{(1^*)}$ from Example 2.41, defined by

$$\lambda_X^{(1^*)}(\Phi) = \lambda a. \left\{ 1 \cdot x \mid x \in \bigcup_{h \in \mathsf{supp}\Phi} h(a) \right\} \tag{4.45}$$

This law is coweak between monads $\mathbf{D}$ and $\mathbf{R} \circ \mathbf{P}$, but plain between $\mathbf{D}$ and $(P-)^A$. The algebraic expansion of c via $\lambda^{(1^*)}$ can be expressed as

$$c^{\#}{}_a(\varphi) = \bigcup_{x \in \mathsf{supp}(\varphi)} c_a(x) \tag{4.46}$$

or in other words

$$(\varphi \xrightarrow{a} \psi \text{ in } c^{\#}) \text{ if and only if } (x \xrightarrow{a} \psi \text{ in } c, \text{ for some } x \in \mathsf{supp}(\varphi))$$

Once again, under the natural isomorphism $P(A \times X) \simeq (PX)^A$, one can rewrite $\lambda^{(1^*)}$ as having type

$$\mathbf{D}P(A \times -) \to P(A \times \mathbf{D}-) \tag{4.47}$$

and see that the obtained expression, as well as the resulting algebraic expansion, are the same as for [81, Section 5.2]. This algebraic expansion originates in non-coalgebraic form in [37].

(2*) the weak distributive law $\lambda^{(2^*)}$ from Example 4.5, defined by

$$\lambda_X^{(2^*)}(\Phi) = \lambda a. \left\{ \mu_X^{\mathbf{D}} \left(\sum_{h \in \mathsf{supp}\Phi} \Phi_h \cdot \varphi^h \right) \mid \forall h \in \mathsf{supp}\Phi . \mathsf{supp}(\varphi^h) \subseteq h(a) \right\} \tag{4.48}$$

The algebraic expansion of c via $\lambda^{(2^*)}$ is expressed as

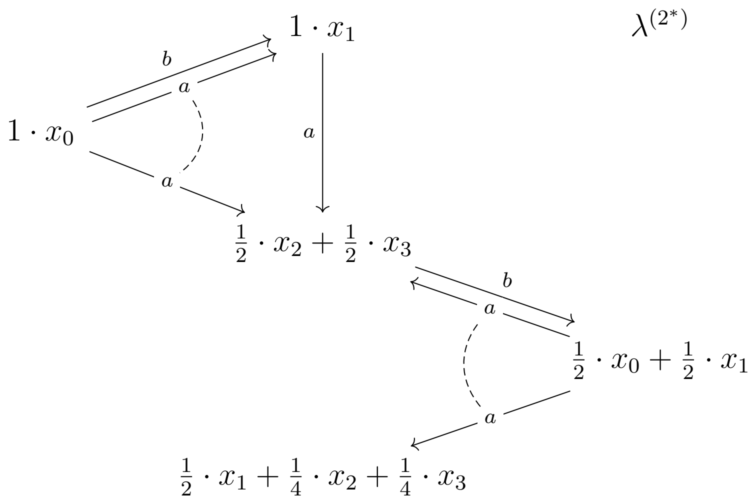
$$c^{\#}{}_a(\varphi) = \left\{ \mu_X^{\mathbf{D}} \left(\sum_{x \in \mathsf{supp}\varphi} \varphi_x \cdot \mu_X^{\mathbf{D}}(\Phi^x) \right) \mid \mathsf{supp}(\Phi^x) \subseteq c_a(x) \right\} \tag{4.49}$$

which means that transitions of shape $\varphi \xrightarrow{a} ?$ in $c^{\#}$ are generated by the following process. For every $x \in \mathsf{supp}(\varphi)$, let θ^x be in the convex closure of the set of distributions θ such that $x \xrightarrow{a} \theta$ in c. Then $\varphi \xrightarrow{a} \sum_{x \in \mathsf{supp}(\varphi)} \varphi_x \theta^x$. The resulting non-deterministic automaton $c^{\#}$ coincides with the belief-state transformer obtained by Bonchi, Silva and Sokolova in [23].

Example 4.14. Let us compute both algebraic expansions on the probabilistic automaton of Example 4.13. The algebraic expansion with respect to $\lambda^{(1^*)}$ consists of a mere copy of the $\lambda^{(1)}$ one for alternating automata.

$$\begin{tikzcd}
 & 1 \cdot x_0 \arrow[dl, "{a,b}"'] \arrow[dr, shift left, "a"] & \lambda^{(1^*)} \\
1 \cdot x_1 \arrow[rr, shift left, "a"] & & \frac{1}{2} \cdot x_2 + \frac{1}{2} \cdot x_3 \arrow[ll, shift left, "b"] \arrow[ul, shift left, "b"]
\end{tikzcd} \tag{4.50}$$

The $\lambda^{(2^*)}$ algebraic expansion makes more sense, because it takes into account the wide range of *convex* transitions allowed by non-determinism. Intuitively, in the state $1 \cdot x_0$ and when reading letter a, the system can non-deterministically choose to view $1 \cdot x_0$ as the combination $(1-t) \cdot x_0 + t \cdot x_0$, with the first x_0 transitioning towards $1 \cdot x_1$ and the second x_0 transitioning towards $\frac{1}{2} \cdot x_2 + \frac{1}{2} \cdot x_3$. The $(1 \cdot x_0)$-accessible portion of the expanded automaton looks like this:



(4.51)

Here, dashed arcs between two transitions $\varphi \xrightarrow{a} \theta_1$ and $\varphi \xrightarrow{a} \theta_2$ denote all transitions into the convex closure of $\{\theta_1, \theta_2\}$. For instance, starting from $1 \cdot x_0$ and reading an a, the system can transition to any distribution $\theta_t = (1-t) \cdot x_1 + \frac{t}{2} \cdot x_2 + \frac{t}{2} \cdot x_3$ for $t \in [0,1]$. Only the two extremal cases $\theta_0 = 1 \cdot x_1$ and $\theta_1 = \frac{1}{2} \cdot x_2 + \frac{1}{2} \cdot x_3$ are displayed. There is no a transition out of θ_t for $t \in (0,1)$, because in the starting probabilistic automaton there is no letter labelling simultaneously some transitions out of x_1, x_2 and x_3. Now, consider the state $\varphi = \frac{1}{2} \cdot x_0 + \frac{1}{2} \cdot x_1$ reading letter a. As $1 \cdot x_0$ can transition to any θ_t and $1 \cdot x_1$ can only transition to θ_1, the composite φ can transition to any distribution of the form $\frac{1}{2}\theta_t + \frac{1}{2}\theta_1$. Again, only extremal transitions are represented.

Remark 4.15. As can be seen in the examples, both following procedures output the same system:

- forgetting probabilities of a probabilistic automaton and algebraically expanding the resulting alternating automaton via $\lambda^{(i)}$
- algebraically expanding a probabilistic automaton via $\lambda^{(i^*)}$ and forgetting probabilities in states of the resulting non-deterministic automaton

This phenomenon happens for $i \in \{1, 2\}$ and is a direct consequence of supp acting like a (weak) distributive law morphism $\lambda^{(i^*)} \to \lambda^{(i)}$ (see Proposition 3.25).

Related Work. In his PhD thesis [112], Louis Parlant manages to derive the $\lambda^{(2)}$ and $\lambda^{(2^*)}$ algebraic expansions using algebraic methods in the context of monoidal monads. He shows that the need for restricting to specific subsets stems from equation preservation, which is required to form a monad on Eilenberg-Moore algebras. His work constitutes a first step in acknowledging that these algebraic expansions are *meaningful on a categorical level* [112, Section 6.3.3]. Our results confirm this vision by putting distributive laws again into the picture. Indeed, these algebraic expansions are canonical in the sense that they arise from the unique monotone weak distributive law of a given type.

4.4 Bisimulations Up-To

Generalised determinisation (= algebraic expansion) of a coalgebra $X \to FTX$ has been described as *sound* because it factors through the category of $\mathbf{T}$-algebras. From a semantic point of view, a powerful consequence is that it allows using up-to techniques for bisimulations. This section will introduce a coalgebraic notion of bisimulation, the framework of bisimulations up-to, and present how distributive laws can play a role in proving compatibility of up-to techniques. Next, we will remark that weak distributive laws can play the same role. As for the previous section, concepts will be illustrated by examples of (non-)deterministic, alternating and probabilistic automata.

There are several coalgebraic notions of bisimulation [143]. To keep things simple, we will stick to an elementary definition in Set. Intuitively, a bisimulation is a relation on a given set such that if a pair (u, v) belongs to the relation, then u and v *behave the same*. The desired notion of *behaviour* has to be specified by a parameter coalgebra describing how u and v can evolve. Technically, for a coalgebra with state space X, a bisimulation is a postfixed point of a well-chosen monotone operator on the complete lattice of relations $P(X \times X)$. We choose to define bisimulations in Hermida-Jacobs style, that is, using a monotone operator derived from relation lifting [72, 130]. See [143] for an overview of how this relates, e.g., to kernel bisimulations used in [23] for probabilistic automata.

Definition 4.16 (Relation lifting). Let $F : \mathsf{Set} \to \mathsf{Set}$ be a functor. For any relation $R \subseteq X \times Y$, the *relation lifting* of R (with respect to F) is defined by

$$\underline{F}R = \{(u, v) \in FX \times FY \mid \exists t \in FR. F\pi_1(t) = u \text{ and } F\pi_2(t) = v\} \tag{4.52}$$

The notation $\underline{F}$ is not a coincidence, because we already have seen this expression before. Equation (4.52) is identical to equation (2.34) of Chapter 2. Using Theorem 2.21 we can see that this $\underline{F}$ defines a Rel endofunctor precisely when F is weakly cartesian – in which case $\underline{F}$ is an extension corresponding to the unique monotone distributive law of type $F\mathbf{P} \to \mathbf{P}F$. For the moment, we do not require F to be weakly cartesian, but this condition will naturally come up in the subsequent developments.

Definition 4.17 (Bisimulation, bisimilarity)**.** Let F be a Set endofunctor and (X, c) be an F-coalgebra. The expression

$$b(R) = (c \times c)^{-1}(\underline{F}R) \tag{4.53}$$

defines a monotone operator $b : P(X \times X) \to P(X \times X)$. A *bisimulation* is a relation $R \in P(X \times X)$ such that $R \subseteq b(R)$. Unravelling definitions, a bisimulation is a relation R such that for all $(u, v) \in R$, there is $t \in FR$ such that $(F\pi_1(t), F\pi_2(t)) = (c(u), c(v))$. By application of the Knaster-Tarski theorem, there is a greatest bisimulation called *bisimilarity*, which is the union of all bisimulations. Two states $u, v \in X$ are *bisimilar*, notation $u \sim_c v$, if there is a bisimulation R such that $(u, v) \in R$.

This definition makes explicit the *coinduction proof principle*. To prove that two states behave the same (in the sense of bisimilarity), it suffices to exhibit a bisimulation that relates them. The coinduction proof principle can be formally stated as

$$(u, v) \in R \subseteq b(R) \Rightarrow u \sim_c v \tag{4.54}$$

Example 4.18. A bisimulation for a deterministic automaton $c : X \to 2 \times X^A$ is a relation $R \subseteq X \times X$ such that for every $(u, v) \in R$,

- u is terminal if and only if v is terminal
- for all $a \in A$, if $u \xrightarrow{a} u'$ and $v \xrightarrow{a} v'$ then $(u', v') \in R$

Consider the (infinite) deterministic automaton with alphabet $A = \{a\}$ and state space $X = \{x_e, x_o\} \cup \{x_n \mid n \in \mathbb{N}^*\}$ depicted below

$$x_o \underset{a}{\overset{a}{\rightleftarrows}} \underline{x_e} \qquad x_1 \xrightarrow{a} \underline{x_2} \xrightarrow{a} x_3 \xrightarrow{a} \underline{x_4} \xrightarrow{a} \cdots \tag{4.55}$$

Then $\{(x_o, x_{2n-1}) \mid n \in \mathbb{N}^*\} \cup \{(x_e, x_{2n}) \mid n \in \mathbb{N}^*\}$ is a bisimulation, so that x_o and x_1 are bisimilar.

Example 4.19. A bisimulation for a non-deterministic automaton $c : X \to 2 \times (PX)^A$ is a relation $R \subseteq X \times X$ such that for every $(u, v) \in R$,

- u is terminal if and only if v is terminal
- for all $a \in A$,
  - if $u \xrightarrow{a} u'$ then there is v' such that $v \xrightarrow{a} v'$ and $(u', v') \in R$
  - if $v \xrightarrow{a} v'$ then there is u' such that $u \xrightarrow{a} u'$ and $(u', v') \in R$

Consider the (infinite) non-deterministic automaton with alphabet $A = \{a\}$ and state space $X = \{x_0, x_\infty\} \cup \{y_n \mid n \in \mathbb{N}\}$ depicted below

$$\begin{tikzcd}
x_\infty \arrow[loop above, "a"] \arrow[d, "a"] & y_1 \arrow[r, "a"] \arrow[d, "a"] & y_2 \arrow[r, "a"] \arrow[dl, "a"] & y_3 \arrow[r, "a"] \arrow[dll, "a"] & \cdots \\
\underline{x_0} & \underline{y_0} & & &
\end{tikzcd} \tag{4.56}$$

Then $\{(x_0, y_0)\} \cup \{(x_\infty, y_n) \mid n \in \mathbb{N}^*\}$ is a bisimulation and e.g. x_∞ and y_1 are bisimilar.

Being a bisimulation is a strong condition: every pair in the bisimulation possibly forces many other pairs to be contained in it as well. As a result, bisimulations often contain a lot of elements, so it can be resource-consuming to find a bisimulation relating two states, as well as to algorithmically check that this is indeed a bisimulation. A possible way to deal with this issue is to use up-to techniques. Such techniques have been introduced specifically to enhance the coinduction proof principle and have attracted a lot of attention in the last decade [122, 124, 21, 127, 19]. Bisimulations detect when two states of an automaton have the same behaviour, whereas bisimulations up-to detect that two states have the same behaviour by proving that they are related *up-to some sound function.*

Definition 4.20 (Bisimulation up-to)**.** Let (X, c) be an F-coalgebra. A monotone map $f : P(X \times X) \to P(X \times X)$ is

- *sound* if for every $R \in P(X \times X)$, $R \subseteq (b \circ f)(R) \Rightarrow R \subseteq \sim_c$
- *compatible* if for every $R \in P(X \times X)$, $f \circ b \subseteq b \circ f$

A *bisimulation up-to* f is a $R \in P(X \times X)$ such that $R \subseteq (b \circ f)(R)$.

Recall that the monotone map b depends on c and has been defined in Equation (4.53).

Every sound f yields a *coinduction up-to f proof principle*

$$(u, v) \in R \subseteq (b \circ f)(R) \Rightarrow u \sim_c v \tag{4.57}$$

but unfortunately, sound maps are not closed under composition. However, compatibility entails soundness, and compatible functions enjoy very good compositional properties.

Proposition 4.21 ([123, 20, 128])**.** *For any F-coalgebra (X, c), the following monotone maps are compatible.*

- $\mathsf{id} : R \mapsto R$
- $\mathsf{refl} : R \mapsto \{(u, u) \mid u \in X\}$
- $\mathsf{sym} : R \mapsto \{(v, u) \mid (u, v) \in R\}$
- $\mathsf{trans} : R \mapsto \{(u, w) \mid \exists v.(u, v) \in R \text{ and } (v, w) \in R\}$ – *if F is weakly cartesian*
- *$f \circ g$, for f and g compatible*
- *$f \cup g$, for f and g compatible*
- *f^ω i.e. $\bigcup_{n \in \mathbb{N}} f^n$, for f compatible*

Intuitively, refl implements reflexive closure, sym implements symmetric closure, and trans implements a step towards transitive closure. Note that compatibility of trans is subject to the requirement that F is weakly cartesian. This makes sense because transitivity is the same as relational composition, and we know that F behaves well with respect to relational composition precisely when it is weakly cartesian (Theorem 2.21, see also [128, Theorem 3] and [20]).

The last base block we need is *contextual closure*, standing for closure under the algebraic operations of some algebraic structure.

Definition 4.22 (Contextual closure)**.** Let (X, x) be a T-algebra, that is, a morphism $x : TX \to X$. Contextual closure with respect to x is a monotone map $\mathsf{cont}_x : P(X \times X) \to P(X \times X)$ defined by

$$\mathsf{cont}_x(R) = \{((x \circ T\pi_1)(t), (x \circ T\pi_2)(t)) \mid t \in TR\} \tag{4.58}$$

Contextual closure is compatible, provided the coalgebraic structure (X, c) and the algebraic structure (X, x) interact well.

Proposition 4.23 ([128, Theorem 4])**.** *Let (X, c) be an F-coalgebra, (X, x) be a T-algebra and $\lambda : TF \to FT$ be a natural transformation. We say (X, x, c) is a* λ-bialgebra *if the following diagram commutes.*

$$\begin{tikzcd}
TFX \arrow[rr, "\lambda_X"] & & FTX \arrow[d, "Fx"] \\
TX \arrow[u, "Tc"] \arrow[dr, "x"'] & & FX \\
& X \arrow[ur, "c"'] &
\end{tikzcd} \tag{4.59}$$

If (X, x, c) is a λ-bialgebra, then cont_x is compatible.

We can now define the up-to technique we are interested in, *congruence closure.*

Definition 4.24. Let (X, x) be a T-algebra. Congruence closure is a monotone map $\mathsf{congr}_x : P(X \times X) \to P(X \times X)$ defined by

$$\mathsf{congr}_x = (\mathsf{refl} \cup \mathsf{sym} \cup \mathsf{trans} \cup \mathsf{cont}_x)^\omega \tag{4.60}$$

According to Propositions 4.21 and 4.23, we have

Proposition 4.25 ([128, Theorem 4])**.** *Assume F is weakly cartesian. Then for any λ-bialgebra (X, x, c), congr_x is compatible.*

Up-to congruence is a particularly well-suited technique for systems on which one can perform generalised determinisation, because of the following result.

Proposition 4.26. *Let $\lambda : \mathbf{T}F \to F\mathbf{T}$ be a weak distributive law and $(-)^\#$ be the corresponding generalised determinisation procedure. Then, for any FT-coalgebra (X, c), the triple $(TX, \mu^{\mathbf{T}}_X, c^\#)$ is a λ-bialgebra.*

Proof. The proof consists of a simple diagram chase:

$$\begin{tikzcd}
TFTX \arrow[rrrr, "\lambda_{TX}"] & & & & FTTX \arrow[dd, "F\mu^{\mathbf{T}}_X"] \\
\arrow[r, phantom, "\text{def}"] & TFTTX \arrow[ul, "TF\mu^{\mathbf{T}}_X"'] \arrow[rr, "\text{nat}", "\lambda_{TTX}"'] \arrow[r, phantom, "(\mu^{\mathbf{T}})\text{ axiom}" near end, shift right=6] & & FTTTX \arrow[ur, "FT\mu^{\mathbf{T}}_X"'] \arrow[d, "F\mu^{\mathbf{T}}_{TX}"] \arrow[r, phantom, "\text{assoc}"] & \\
TTX \arrow[uu, "Tc^\#"] \arrow[r, "TTc"] \arrow[drr, "\mu^{\mathbf{T}}_X"'] & TTFTX \arrow[u, "T\lambda_{TX}"] \arrow[r, "\mu^{\mathbf{T}}_{FTX}"] \arrow[d, phantom, "\text{nat}"] & TFTX \arrow[r, "\lambda_{TX}"] & FTTX \arrow[r, "F\mu^{\mathbf{T}}_X"] \arrow[d, phantom, "\text{def}"] & FTX \\
& & TX \arrow[u, "Tc"] \arrow[urr, "c^\#"'] & &
\end{tikzcd} \tag{4.61}$$

□

Proposition 4.26 is folklore in the case where λ is a plain distributive law, and the proof would look the same. To clarify, our contribution is to combine three observations:

- generalised determinisation can still be performed with a weak distributive law (Proposition 4.8)
- the formula for $(-)^{\#}$ remains the same (Proposition 4.9)
- the standard proof of Proposition 4.26, where λ is supposed to be a distributive law, does not use the $(\eta^{\mathbf{T}})$ axiom

To sum up, the main result of this section can be formulated as follows – examples of applications will be given just beyond.

Theorem 4.27. *Let F be an endofunctor and $\mathbf{T}$ be a monad on Set, along with a weak distributive law $\lambda : \mathbf{T}F \to F\mathbf{T}$. Any FT-coalgebra (X, c) yields an F-coalgebra $(TX, c^{\#})$ for which $\mathsf{cont}_{\mu_X^{\mathbf{T}}}$ is a compatible. If F is weakly cartesian, then $\mathsf{congr}_{\mu_X^{\mathbf{T}}}$ is compatible as well.*

Proof. By Proposition 4.26, $(TX, \mu_X^{\mathbf{T}}, c^{\#})$ is a λ-bialgebra, so $\mathsf{cont}_{\mu_X^{\mathbf{T}}}$ is compatible by Proposition 4.23. When F is weakly cartesian, Proposition 4.25 gives compatibility of $\mathsf{congr}_{\mu_X^{\mathbf{T}}}$. □

Before diving into more complex systems like alternating automata and probabilistic automata, we present a basic example of how up-to techniques can accelerate computation of bisimulations. This example is based on the work of Bonchi and Pous [21]. By exploiting bisimulations up-to congruence, they devised a Hopcroft-Karp-like algorithm – implicitly using a distributive law – capable of checking equivalence of non-deterministic automata. In many cases, their algorithm is faster by an order of magnitude than other cutting-edge algorithms [21, Table 2].

Example 4.28. Let us see what means applying Theorem 4.27 with the weakly cartesian functor $F = 2 \times (-)^A$, $\mathbf{T} = \mathbf{P}$ and $\lambda_X(U) = (\max\{b \mid (b, h) \in U\}, \lambda a.\{h(a) \mid (b, h) \in U\})$ as in equation (4.15). As we have explained before, the generalised determinisation of a non-deterministic automaton with respect to λ is the standard determinisation procedure, also called powerset construction. In the resulting deterministic automaton, $\mathsf{congr}_{\mu_X^{\mathbf{P}}}$ (congruence closure with respect to the algebraic operation of *union*) is a compatible, therefore sound, up-to technique. Thus, any two states whose pair (U, V) is contained in a bisimulation up-to congruence are bisimilar in the sense of deterministic automata (see Example 4.18). Consider the following non-deterministic automaton c:

$$\begin{tikzcd}
& x_2 \arrow[dr, "a"] & & \\
\underline{x_0} \arrow[ur, "a"] \arrow[rr, shift left, "a"] & & \underline{x_1} \arrow[ll, shift left, "a"] & \underline{y} \arrow[loop above, "a"]
\end{tikzcd} \tag{4.62}$$

The determinised of c via λ is $c^{\#}$, partially represented below:

$$\begin{tikzcd}
\underline{\{x_0\}} \arrow[r, "a"] & \underline{\{x_1, x_2\}} \arrow[r, "a"] & \underline{\{x_0, x_1\}} \arrow[r, "a"] & \underline{\{x_0, x_1, x_2\}} \arrow[loop above, "a"] & \underline{\{y\}} \arrow[loop above, "a"]
\end{tikzcd} \tag{4.63}$$

One can easily see that $\{x_0\} \sim_{c^{\#}} \{y\}$. Actually, the smallest bisimulation containing $(\{x_0\}, \{y\})$ has cardinal 4.

$$R = \{(\{x_0\}, \{y\}), (\{x_1, x_2\}, \{y\}), (\{x_0, x_1\}, \{y\}), (\{x_0, x_1, x_2\}, \{y\})\} \tag{4.64}$$

However, to prove $\{x_0\} \sim_{c^{\#}} \{y\}$ it suffices to check that the (smaller) relation

$$R_0 = \{(\{x_0\}, \{y\}), (\{x_1, x_2\}, \{y\})\} \tag{4.65}$$

is a bisimulation up-to $\mathsf{congr}_{\mu^{\mathsf{P}}_X}$, as follows:

- for $(\{x_0\}, \{y\})$: $\{x_0\} \xrightarrow{a} \{x_1, x_2\}$ and $\{y\} \xrightarrow{a} \{y\}$, and we have that the pair $(\{x_1, x_2\}, \{y\}) \in \mathsf{congr}_{\mu^{\mathsf{P}}_X}(R_0)$ because it already is in R_0
- for $(\{x_1, x_2\}, \{y\})$: $\{x_1, x_2\} \xrightarrow{a} \{x_0, x_1\}$ and $\{y\} \xrightarrow{a} \{y\}$, and we have that the pair $(\{x_0, x_1\}, \{y\}) \in \mathsf{congr}_{\mu^{\mathsf{P}}_X}(R_0)$ because, denoting $\mathsf{congr}_{\mu^{\mathsf{P}}_X}(R_0)$ by the infix symbol $\equiv$,

$$\begin{aligned}
\{x_0, x_1\} &= \{x_0\} \cup \{x_1\} & & \\
&\equiv \{y\} \cup \{x_1\} & & \text{because } (\{x_0\}, \{y\}) \in R_0 \\
&\equiv \{x_1, x_2\} \cup \{x_1\} & & \text{because } (\{x_1, x_2\}, \{y\}) \in R_0 \\
&= \{x_1, x_2\} & & \\
&\equiv \{y\} & & \text{because } (\{x_1, x_2\}, \{y\}) \in R_0
\end{aligned}$$

For any algebraic expansion $c \mapsto c^{\#}$ using a weak distributive law, three levels of bisimilarity semantics arise. We designate them by the following terminology:

- Strong bisimilarity. Two states x, y in X are strongly bisimilar if $x \sim_c y$ i.e. strong bisimilarity is the standard notion of bisimilarity for c.
- Combination bisimilarity. Two states x, y in X are combination bisimilar if $x \sim_{c^\dagger} y$, where $c^\dagger$ is the intermediate coalgebra obtained in Proposition 4.10.
- Algebraic bisimilarity. Two states x, y in X are algebraically bisimilar if they give rise to bisimilar states in the algebraic expansion i.e. if $\eta^{\mathbf{T}}_X(x) \sim_{c^\#} \eta^{\mathbf{T}}_X(y)$.

Combination bisimilarity emerges precisely because of the weakness of the distributive law. Intuitively, this semantics copes with the fact that the transition structure of c is not stable enough by adding the possibility to combine transitions in $c^\dagger$. The two main examples are

- for alternating automata and $\lambda^{(2)}$, one can combine any non-zero number of transitions $x \xrightarrow{a} U_i$ in c into a unified transition $x \xrightarrow{a} \bigcup_i U_i$ in $c^\dagger$
- for probabilistic automata and $\lambda^{(2^*)}$, this is convex bisimilarity [105], obtained by combining any finite (non-zero) number of transitions $x \xrightarrow{a} \theta_i$ in c into any convex combination $x \xrightarrow{a} \sum_i p_i\theta_i$ in $c^\dagger$

and a degenerate case is

- for alternating automata and $\lambda^{(0)}$, one *must* combine all possible transitions $x \xrightarrow{a} U_i$ in c into a unique transition $x \xrightarrow{a} \bigcup_i U_i$ in $c^\dagger$.

For probabilistic automata, these three notions of bisimilarity have been gathered in [23] under the respective names *strong probabilistic bisimilarity*, *convex bisimilarity* and *distribution bisimilarity*. Bonchi *et al.* remarked that they indeed arise from some kind of algebraic expansion, and stated about bisimulation up-to convex hull for distribution (= algebraic) bisimilarity:

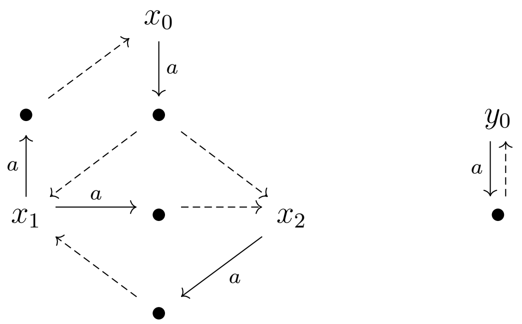
Unfortunately, the lack of a suitable distributive law [153] makes it impossible to reuse the abstract results in [20]. Fortunately, we can redo all the proofs by adapting the theory in [124] to probabilistic automata.

By exhibiting the underlying weak distributive law $\lambda^{(2^*)}$ and using Theorem 4.27, we are actually able to uniformly get compatibility of up-to context/congruence and check more easily algebraic bisimilarity.

Remark 4.29. For the plain distributive laws $\lambda^{(1)}$ and $\lambda^{(1^*)}$, obtained respectively using trivial coweak distributive laws $\mathbf{PP} \to \mathbf{PP}$ and $\mathbf{DP} \to \mathbf{PD}$, strong bisimilarity coincides with algebraic bisimilarity (for probabilistic automata this was already stated in a footnote of [23]).

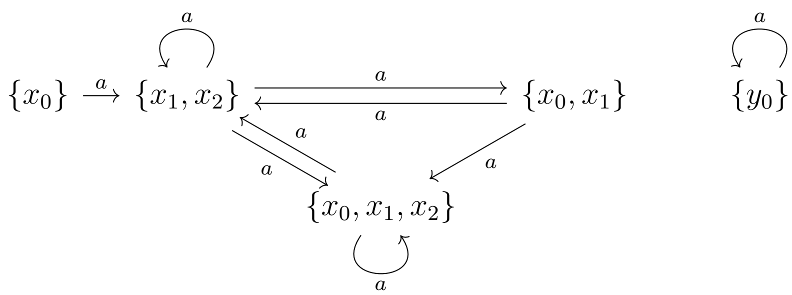
Let us give examples of bisimulations up-to using (non trivial) weak distributive laws. We consider the $\lambda^{(2)}$ algebraic expansion of alternating automata and the $\lambda^{(2^*)}$ algebraic expansion of probabilistic automata. By Theorem 4.27, context closure is compatible. Furthermore, the functor $(P-)^A$ is weakly cartesian by composition of weakly cartesian functors, so in both cases congruence closure is compatible as well.

Example 4.30. Let c be the following alternating automaton with alphabet $A = \{a\}$ and state space $X = \{x_0, x_1, x_2, y_0\}$.



(4.66)

Its $\lambda^{(2)}$ algebraic expansion $c^\#$ is



(4.67)

Let us prove that $\{x_0\} \sim_{c^\#} \{y_0\}$ using coinduction. A first way is to define a bisimulation $R \subseteq PX \times PX$ containing $(\{x_0\}, \{y_0\})$. The smallest such R is

$$\{(\{x_0\}, \{y_0\}), (\{x_1, x_2\}, \{y_0\}), (\{x_0, x_1\}, \{y_0\}), (\{x_0, x_1, x_2\}, \{y_0\})\} \tag{4.68}$$

Another option is to define a bisimulation up-to congruence containing $(\{x_0\}, \{y_0\})$. For alternating automata, as we factor through $\mathsf{EM}(\mathbf{P})$, context closure is given for any $R \subseteq PX \times PX$ by

$$\mathsf{cont}_{\mu^{\mathbf{P}}_X}(R) = \left\{ \left(\bigcup_{(U,V)\in S} U, \bigcup_{(U,V)\in S} V \right) \mid S \subseteq R \right\} \tag{4.69}$$

Define the relation

$$R_0 = \{(\{x_0\}, \{y_0\}), (\{x_1, x_2\}, \{y_0\})\} \tag{4.70}$$

Let us denote $\mathsf{congr}_{\mu^{\mathsf{P}}_X}(R_0)$ by the infix operator $\equiv$. To check that R_0 is a bisimulation up-to congruence, we must prove that for all $(U, V) \in R_0$,

- for every $U \xrightarrow{a} U'$ in $c^{\#}$, there is V' such that $V \xrightarrow{a} V'$ and $U' \equiv V'$
- for every $V \xrightarrow{a} V'$ in $c^{\#}$, there is U' such that $U \xrightarrow{a} U'$ and $U' \equiv V'$

Verifications for the pair $(\{x_0\}, \{y_0\})$ are immediate. For the pair $(\{x_1, x_2\}, \{y_0\})$, we must check three things.

- Check that $\{x_1, x_2\} \equiv \{y_0\}$. This is immediate because $(\{x_1, x_2\}, \{y_0\}) \in R_0$.
- Check that $\{x_0, x_1\} \equiv \{y_0\}$. We have

$$\begin{aligned}
\{x_0, x_1\} &= \{x_0\} \cup \{x_1\} && \\
&\equiv \{y_0\} \cup \{x_1\} && \text{because } (\{x_0\}, \{y_0\}) \in R_0 \\
&\equiv \{x_1, x_2\} \cup \{x_1\} && \text{because } (\{x_1, x_2\}, \{y_0\}) \in R_0 \\
&= \{x_1, x_2\} && \\
&\equiv \{y_0\} && \text{because } (\{x_1, x_2\}, \{y_0\}) \in R_0
\end{aligned}$$

- Check that $\{x_0, x_1, x_2\} \equiv \{y_0\}$. This is obtained from the previous results

$$\begin{aligned}
\{x_0, x_1, x_2\} &= \{x_0, x_1\} \cup \{x_1, x_2\} && \\
&\equiv \{y_0\} \cup \{x_1, x_2\} && \text{because } \{x_0, x_1\} \equiv \{y_0\} \\
&\equiv \{y_0\} \cup \{y_0\} && \text{because } \{x_1, x_2\} \equiv \{y_0\} \\
&= \{y_0\} &&
\end{aligned}$$

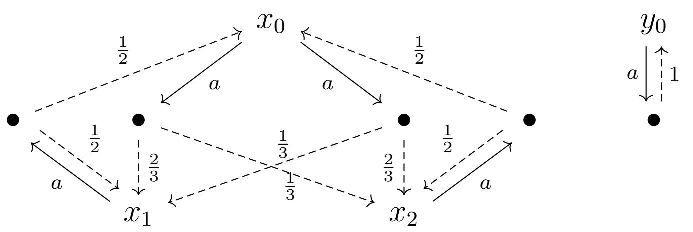
Hence R_0 is a bisimulation up-to congruence and $\{x_0\} \sim_{c^{\#}} \{y_0\}$.

Remark 4.31. In [90], Klin and Rot expressed the following concern about the fact that in the $\lambda^{(2)}$ algebraic expansion of alternating automata, ${c^{\#}}_a(U)$ has to be closed under non-empty unions:

> *The reachable part of the nondeterministic automaton may be larger than the one obtained by the standard procedure. As a result, although our determinization is correct, it may be less efficient than the standard one.* [90, p.22]

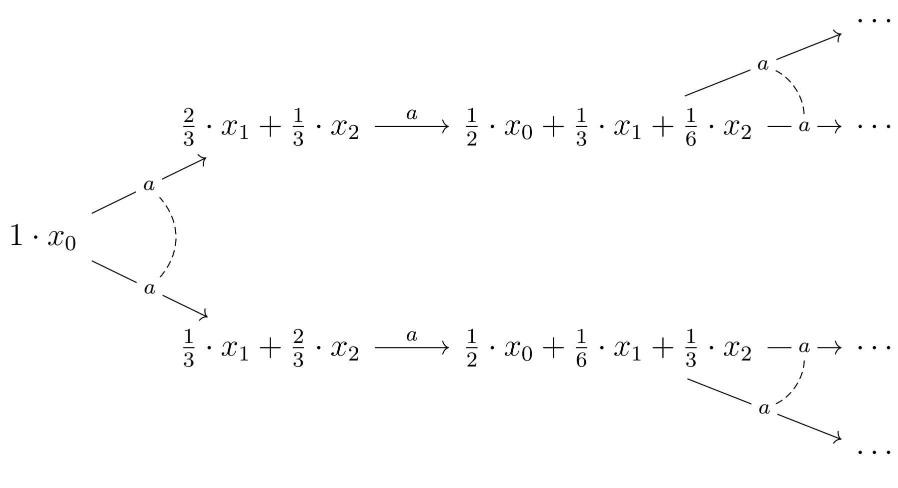
By *the standard one*, Klin and Rot mean *the same procedure, without closing $c^{\#}{}_a(U)$ under non-empty unions.* As we have seen in Example 4.30, bisimulations up-to congruence can be applied to accelerate checks of algebraic bisimilarity. A further remark is that verifications concerning composite states (e.g. $\{x_0, x_1, x_2\}$) are fully determined by verifications concerning states present in the so-called standard determinisation (here, $\{x_0\}$ and $\{x_1, x_2\}$). Thus, additionally from being categorically canonical, the $\lambda^{(2)}$ algebraic expansion is as efficient as the standard one for what concerns algebraic bisimilarity.

Example 4.32. Let c be the following probabilistic automaton with alphabet $A = \{a\}$ and state space $X = \{x_0, x_1, x_2, y_0\}$.



(4.71)

Its $\lambda^{(2^*)}$ algebraic expansion $c^{\#}$ is a non-deterministic automaton with state space DX. In $c^{\#}$, the state $1 \cdot y_0$ has just one a-labelled loop and the state $1 \cdot x_0$ generates the following tree, both vertically and horizontally infinite:



(4.72)

Let us prove that $1 \cdot x_0 \sim_{c^{\#}} 1 \cdot y_0$ using coinduction. A first way is to define a bisimulation $R \subseteq DX \times DX$ containing $(1 \cdot x_0, 1 \cdot y_0)$. The smallest such R consists of all pairs $(\varphi, 1 \cdot y_0)$ where φ ranges over the $(1 \cdot x_0)$ accessible part of $c^{\#}$, partially represented above. Then, any bisimulation is infinite (and even uncountable). Another option is to define a bisimulation up-to congruence R_0 containing $(1 \cdot x_0, 1 \cdot y_0)$. For probabilistic automata, factorisation through $\mathsf{EM}(\mathbf{D})$ yields a context closure given for all $R \subseteq DX \times DX$ by

$$\mathsf{cont}_{\mu^{\mathbf{D}}_X}(R) = \left\{ \left(\sum_{i=1}^{n} p_i \varphi_i, \sum_{i=1}^{n} p_i \psi_i \right) \mid n \in \mathbb{N}, p_i > 0, \sum_{i=1}^{n} p_i = 1, \forall i.(\varphi_i, \psi_i) \in R \right\} \quad (4.73)$$

Define for all $\alpha \in [0,1]$ the distribution $\varphi_\alpha = \frac{2-\alpha}{3} \cdot x_1 + \frac{1+\alpha}{3} \cdot x_2$. These are all the distributions accessible from $1 \cdot x_0$ in one step. Now define

$$R_0 = \{(1 \cdot x_0, 1 \cdot y_0), (\varphi_0, 1 \cdot y_0), (\varphi_1, 1 \cdot y_0)\} \quad (4.74)$$

Let us denote $\mathsf{congr}_{\mu^{\mathbf{D}}_X}(R_0)$ by the infix operator $\equiv$ again and check that R_0 is a bisimulation up-to congruence, i.e. prove that for all $(\varphi, \psi) \in R_0$

- for every $\varphi \xrightarrow{a} \varphi'$ in $c^{\#}$, there is ψ' such that $\psi \xrightarrow{a} \psi'$ and $\varphi' \equiv \psi'$
- for every $\psi \xrightarrow{a} \psi'$ in $c^{\#}$, there is φ' such that $\varphi \xrightarrow{a} \varphi'$ and $\varphi' \equiv \psi'$

Consider, first, the pair $(1 \cdot x_0, 1 \cdot y_0)$. Here we need to check that for every $\alpha \in [0,1]$, $\varphi_\alpha \equiv 1 \cdot y_0$. As $(\varphi_0, 1 \cdot y_0) \in R_0$ and $(\varphi_1, 1 \cdot y_0) \in R_0$, the result is immediate for $\alpha \in \{0,1\}$. Now, for convex combinations φ_α with $\alpha \in (0,1)$, we have

$$\begin{aligned}
\varphi_\alpha &= (1-\alpha)\varphi_0 + \alpha\varphi_1 \\
&\equiv (1-\alpha)(1 \cdot y_0) + \alpha\varphi_1 && \text{because } (\varphi_0, 1 \cdot y_0) \in R_0 \\
&\equiv (1-\alpha)(1 \cdot y_0) + \alpha(1 \cdot y_0) && \text{because } (\varphi_1, 1 \cdot y_0) \in R_0 \\
&= 1 \cdot y_0
\end{aligned}$$

We could have performed both $\equiv$ steps at once using only context closure. Though congruence closure is not formally needed here, it allows to split the reasoning in several simple steps, resulting in more local / compositional proofs.

By symmetry, to end the proof it suffices to consider the pair $(\varphi_0, 1 \cdot y_0)$ of R_0. We must check that

$$\frac{1}{2} \cdot x_0 + \frac{1}{3} \cdot x_1 + \frac{1}{6} \cdot x_2 \equiv 1 \cdot y_0 \quad (4.75)$$

and indeed

$$\begin{aligned}
\frac{1}{2} \cdot x_0 + \frac{1}{3} \cdot x_1 + \frac{1}{6} \cdot x_2 &= \frac{1}{2} \cdot x_0 + \frac{1}{2}\varphi_0 \\
&\equiv \frac{1}{2} \cdot y_0 + \frac{1}{2}\varphi_0 && \text{because } (1 \cdot x_0, 1 \cdot y_0) \in R_0 \\
&\equiv \frac{1}{2} \cdot y_0 + \frac{1}{2} \cdot y_0 && \text{because } (\varphi_0, 1 \cdot y_0) \in R_0 \\
&= 1 \cdot y_0
\end{aligned}$$

so that R_0 is a bisimulation up-to congruence and $1 \cdot x_0 \sim_{c^{\#}} 1 \cdot y_0$.

Remark 4.33. It was shown by Bonchi *et al* [23] that for a finite probabilistic automaton c, it is always possible to find a finite bisimulation up-to congruence witnessing algebraic bisimilarity, despite the cardinality explosion resulting from algebraic expansion.

Chapter 5

Interlude

Although the theory of distributive laws can be used in any category, in the preceding chapters we focused mainly on examples living in the category Set. For a first exposition, the category of sets is indeed convenient: it is intuitive to grasp, contains many interesting monads, and is widely used in computer science, especially in coalgebra theory. A crucial advantage of Set is also the well-established correspondence between monads and algebraic theories. However, the semantics of programming languages commonly relies on categories that are richer than Set. They include categories related to dcpos in domain theory, or the recently introduced category of quasi-Borel spaces in probabilistic programming [74]. Combination of effects in these more general contexts is a problem that regularly attracts the attention of computer scientists, notably combination of non-deterministic choice and probabilistic choice [88, 59].

In the last chapters of this thesis, we find it interesting to start applying the theory of weak distributive laws in categories other than Set. We do not directly aim at the categories mentioned above, but will rather try, as a first approach, to explore categories that are similar to Set. The motivating question is

Are there some weak distributive laws beyond Set*?*

Of course, the answer is positive, because by Theorem 2.11 any monad morphism yields a trivial weak distributive law. So a better motivating question is

Are there some non-trivial weak distributive laws beyond Set*?*

Theorem 2.22 is the only general result at our disposal for generating non-trivial weak distributive laws in Set. Its gist is to extend canonically one of the monads to a category of relations. To generate weak distributive laws similarly in another category C, we can guess that this category should possess a good notion of *relation*. Such categories are known as *regular categories*. In a nutshell, any regular category C

yields a well-behaved category with the same objects as in C but where morphisms can be regarded as relations between objects of C. Functors and natural transformations that are nearly cartesian – a weakening of being weakly cartesian – can be seen to have *relational extensions*.

This short chapter introduces regular categories and explains how Theorem 2.22 generalises to this framework. The chapter in itself contains no original results. It is more of a compilation of well-established facts in the literature, and an introduction to the last two chapters of the thesis. The principal reference here is the seminal paper from Carboni, Kelly and Wood [32] – see also [29]. We will then proceed to our primary goal:

- In Chapter 6, we will examine *toposes*. They are exactly the regular categories in which relational extensions are just extensions with respect to the Kleisli category of some notion of powerset monad. In particular, Set falls into the scope of Chapter 6.
- In Chapter 7, we work in the regular category KHaus of compact Hausdorff spaces and continuous functions, which is not a topos. A bit of further work will be required to establish the link between relational extensions and extensions in a Kleisli category.

Chapters 5, 6 and 7 together are based on the ICALP paper *Powerset-like monads weakly distribute over themselves in toposes and compact Hausdorff spaces* [63].

5.1 Regular Categories

Regular categories have been introduced by Barr [4] and have several equivalent definitions. We shall follow the approach of [32]. To state a definition of regular categories, we need to recall the notion of subobject. Monomorphisms and epimorphisms will be respectively denoted by the standard symbols $\hookrightarrow$ and $\twoheadrightarrow$.

Subobjects. Let X be an object of a category. Two monomorphisms $m : U \hookrightarrow X$, $n : V \hookrightarrow X$ are called isomorphic if there is an isomorphism $k : U \to V$ such that $m = n \circ k$. An isomorphism class of monomorphisms into X is called a *subobject* of X. In the sequel, for simplicity we will not always distinguish between a subobject and its isomorphism class. For two subobjects $m : U \hookrightarrow X$, $n : V \hookrightarrow X$ of a same object X, say $m \leq n$ when there exists a morphism $h : U \to V$ such that $m = n \circ h$. This defines the *subobject order* on the class of subobjects of X.

Definition 5.1 (Regular category)**.** A *regular category* is a category C satisfying the following conditions:

- C is finitely complete i.e. has all finite limits
- for each morphism $f : X \to Y$, there is a smallest subobject $i : U \hookrightarrow Y$ such that $f = i \circ e$ for some $e : X \to U$
- the pullback of a strong epimorphism is a strong epimorphism

For clarity, we recall that the notion of epimorphism has many variations. A morphism $e : X \to Y$ in C is

- an *epimorphism* when for all $f, g : Y \to Z$, if $f \circ e = g \circ e$ then $f = g$
- a *strong epimorphism* when for any monomorphism $i : X' \hookrightarrow Y'$ and any morphisms $f : X \to X'$, $g : Y \to Y'$ such that $g \circ e = i \circ f$, there is a morphism $h : Y \to X'$ such that $h \circ e = f$ and $i \circ h = g$
- a *regular epimorphism* when there are $f, g : Z \to X$ such that e is the coequaliser of f and g
- a *split epimorphism* when there exists $i : Y \to X$ satisfying $e \circ i = \mathrm{id}_Y$.

In Set, epimorphisms are just surjective functions and all the above notions coincide. But in general – see e.g. [1] – one only has the chain of implications

split epimorphism ⇒ regular epimorphism ⇒ strong epimorphism ⇒ epimorphism

In a regular category, regular epimorphisms and strong epimorphisms coincide – therefore we will only use the terminology regular epimorphism.

In the rest of this chapter, C is assumed to be regular. In a factorisation $f = i \circ e$, if i is the smallest subobject as in the second condition of Definition 5.1, then the morphism e is a regular epimorphism [32, §1.3]. Consequently, any morphism f has a factorisation $f = i \circ e$ where i is a monomorphism and e is a regular epimorphism. The monomorphism obtained as the isomorphism class of i is called the *image* of f. Note also that

Lemma 5.2. *In a regular category C, every idempotent splits.*

Proof. Let $f : X \to X$ such that $f \circ f = f$. Write $f = i \circ e$ for some monomorphism $i : U \hookrightarrow X$ and some regular epimorphism $e : X \twoheadrightarrow U$. Then $i \circ e \circ i \circ e = i \circ e$. As i is mono and e is epi, this implies $e \circ i = \mathrm{id}_U$, which exhibits f as a split idempotent. □

Being finitely complete, C has all pullbacks. In this context, weak pullbacks are squares such that the universal morphism into the pullback is a split epimorphism. Requiring a regular epimorphism instead leads to the (weaker) notion of a near pullback [137, Proposition 1.5.2]. Explicitly,

Definition 5.3 (Near pullback)**.** For any two morphisms $f : X \to Z$ and $g : Y \to Z$ in C, a square

$$\begin{tikzcd} N \arrow[r, "n_2"] \arrow[d, "n_1"'] & Y \arrow[d, "g"] \\ X \arrow[r, "f"'] & Z \end{tikzcd} \tag{5.1}$$

is a *near pullback* if the universal morphism into the pullback $h : N \to P$ is a regular epimorphism.

$$\begin{tikzcd} N \arrow[dr, "h"] \arrow[drr, bend left, "n_2"] \arrow[ddr, bend right, "n_1"'] & & \\ & P \arrow[r, "p_2"] \arrow[d, "p_1"] \arrow[dr, phantom, "\lrcorner", very near start] & Y \arrow[d, "g"] \\ & X \arrow[r, "f"'] & Z \end{tikzcd} \tag{5.2}$$

This induces some notions of nearly cartesian functor and natural transformation.

Definition 5.4 (Nearly cartesian functor)**.** A functor $F : \mathsf{C} \to \mathsf{C}$ is *nearly cartesian* if it maps pullbacks to near pullbacks.

Definition 5.5 (Nearly cartesian natural transformation)**.** A natural transformation $\alpha : F \to G$ is *nearly cartesian* if its naturality squares are near pullbacks.

5.2 Relations in a Regular Category

Following [32, § 1.4], every regular category C induces a category of objects and relations denoted by $\mathsf{Rel}(\mathsf{C})$.

- The objects of $\mathsf{Rel}(\mathsf{C})$ are the objects of C.
- A morphism $r : X \rightsquigarrow Y$ in $\mathsf{Rel}(\mathsf{C})$ is a subobject of $X \times Y$ in C and is called a relation. Equivalently, a relation is a jointly monic span i.e. a pair of morphisms $r_1 : R \to X$, $r_2 : R \to Y$ in C such that $\langle r_1, r_2 \rangle : R \hookrightarrow X \times Y$ is a monomorphism. In this view a relation can be represented by

$$\begin{tikzcd} & R \arrow[dl, "r_1"'] \arrow[dr, "r_2"] & \\ X & & Y \end{tikzcd} \tag{5.3}$$

- Composition of relations is performed using pullbacks. More precisely, in the presence of two relations $r : X \rightsquigarrow Y$ and $s : Y \rightsquigarrow Z$ given by monomorphisms $\langle r_1, r_2 \rangle : R \hookrightarrow X \times Y$ and $\langle s_1, s_2 \rangle : S \hookrightarrow Y \times X$, consider the following pullback

$$\begin{tikzcd}
 & & \Theta \arrow[dl, "\theta_1"'] \arrow[dr, "\theta_2"] \arrow[d, phantom, "\vee"] & & \\
 & R \arrow[dl, "r_1"'] \arrow[dr, "r_2"] & {} & S \arrow[dl, "s_1"'] \arrow[dr, "s_2"] & \\
X & & Y & & Z
\end{tikzcd} \tag{5.4}$$

  The relational composition $s \cdot r : X \rightsquigarrow Z$ is defined as being the image of the morphism $\theta = \langle r_1 \circ \theta_1, s_2 \circ \theta_2 \rangle : \Theta \to X \times Z$.

- The identity relation $X \rightsquigarrow X$ is the diagonal monomorphism $\langle \mathrm{id}_X, \mathrm{id}_X \rangle : X \hookrightarrow X \times X$.

The notation $\rightsquigarrow$ distinguishes morphisms in $\mathsf{Rel}(\mathsf{C})$, and the notation $\cdot$ distinguishes composition in $\mathsf{Rel}(\mathsf{C})$. There are two important functors exposing the underlying structure of $\mathsf{Rel}(\mathsf{C})$:

- The *graph functor* $\mathcal{G} : \mathsf{C} \to \mathsf{Rel}(\mathsf{C})$ is the identity on objects and maps a morphism $f : X \to Y$ to the relation $\langle \mathrm{id}_X, f \rangle : X \rightsquigarrow Y$.

- The *transpose functor* $(-)^\circ : \mathsf{Rel}(\mathsf{C})^{\mathsf{op}} \to \mathsf{Rel}(\mathsf{C})$ is the identity on objects and maps a relation $\langle r_1, r_2 \rangle : X \rightsquigarrow Y$ to the relation $\langle r_2, r_1 \rangle : Y \rightsquigarrow X$. Note that it is a contravariant involution: $r^{\circ\circ} = r$.

These functors emphasise properties that are expected of every well-behaved category of relations. The graph functor stresses that every morphism can be canonically identified as a relation between its domain and its codomain. The transpose functor establishes a symmetry between domains and codomains of relations. Graph and transpose are intertwined via the fundamental decomposition equation

$$r = \mathcal{G}r_2 \cdot \mathcal{G}{r_1}^\circ \tag{5.5}$$

which holds for every relation $r = \langle r_1, r_2 \rangle$. Most of the time, the symbol $\mathcal{G}$ will be omitted for more readability. For example, the decomposition equation can be rewritten $r = r_2 \cdot {r_1}^\circ$, and the functoriality of $\mathcal{G}$ is just $g \circ f = g \cdot f$.

A further feature is the order on relations, which is inherited from the subobject order. Let $r : X \rightsquigarrow Y$ and $s : X \rightsquigarrow Y$ be relations, given by monomorphisms

$r : R \hookrightarrow X \times Y$ and $s : S \hookrightarrow X \times Y$. We recall that saying $r \leq s$ when there exists $h : R \to S$ such that $r = s \circ h$ defines an order on the collection of relations of type $X \rightsquigarrow Y$. This order is compatible with composition in the sense of equation (2.27). If C is well-powered (i.e. the collection of subobjects of X is always a set), this makes $\mathsf{Rel}(\mathsf{C})$ into a Pos-enriched category. Even without well-poweredness, we can still define a functor $F : \mathsf{Rel}(\mathsf{C}) \to \mathsf{Rel}(\mathsf{C})$ to be *locally monotone* if $r \leq s \Rightarrow Fr \leq Fs$.

5.3 Relational Extensions

Functors and natural transformations in C can sometimes be extended to relations. The situation is similar to Theorem 2.21, with the graph functor $\mathcal{G} : \mathsf{C} \to \mathsf{Rel}(\mathsf{C})$ playing the role of the traditional graph functor $F_{\mathsf{P}} : \mathsf{Set} \to \mathsf{Rel}$. The following theorem is obtained by reformulating and merging some results from [32, § 4.3] and [137, Corollary 1.5.7]. It will be the basis of subsequent developments in Chapters 6 and 7.

Theorem 5.6 ([32, 137]). *Let F, $G : \mathsf{C} \to \mathsf{C}$ be endofunctors on a regular category C and $\alpha : F \to G$ be a natural transformations.*

- *There is a locally monotone functor $\mathsf{Rel}(F) : \mathsf{Rel}(\mathsf{C}) \to \mathsf{Rel}(\mathsf{C})$ such that $\mathcal{G}F = \mathsf{Rel}(F)\mathcal{G}$ if and only if F preserves regular epimorphisms and is nearly cartesian. In this case, such a locally monotone $\mathsf{Rel}(F)$ is unique.*

- *Assume F and G satisfy the previous point. There is a natural transformation $\mathsf{Rel}(\alpha) : \mathsf{Rel}(F) \to \mathsf{Rel}(G)$ such that $\mathcal{G}\alpha = \mathsf{Rel}(\alpha)\mathcal{G}$ if and only if α is nearly cartesian. In this case, such a $\mathsf{Rel}(\alpha)$ is unique.*

Proof (sketch). We provide a short proof of the uniqueness of $\mathsf{Rel}(F)$, which appeared in [43, §5.3.11]. First observe [32, § 1.2] that for every $f : X \to Y$, f° is the unique relation $Y \rightsquigarrow X$ such that $\mathrm{id}_X \leq f^\circ \cdot f$ and $f \cdot f^\circ \leq \mathrm{id}_Y$. For any locally monotone functor $H : \mathsf{Rel}(\mathsf{C}) \to \mathsf{Rel}(\mathsf{C})$ and any $f : X \to Y$ in C

$$\mathrm{id}_{HX} = H(\mathrm{id}_X) \leq H(f^\circ \cdot f) = H(f^\circ) \cdot Hf$$
$$Hf \cdot H(f^\circ) = H(f \cdot f^\circ) \leq H(\mathrm{id}_Y) = \mathrm{id}_{HY}$$

hence $(Hf)^\circ = H(f^\circ)$. Now assume $\mathsf{Rel}(F)$ is a locally monotone functor extending F in the sense that $\mathcal{G}F = \mathsf{Rel}(F)\mathcal{G}$. On objects this forces $\mathsf{Rel}(F)X = FX$. On any morphism $f : X \to Y$ seen as a relation $f : X \rightsquigarrow Y$ we have $\mathsf{Rel}(F)f = Ff$. Now for an arbitrary relation $r = \langle r_1, r_2 \rangle : X \rightsquigarrow Y$, use the decomposition $r = r_2 \cdot {r_1}^\circ$ to get

$$\mathsf{Rel}(F)(r) = \mathsf{Rel}(F)r_2 \cdot \mathsf{Rel}(F)({r_1}^\circ) = \mathsf{Rel}(F)r_2 \cdot (\mathsf{Rel}(F)r_1)^\circ = Fr_2 \cdot (Fr_1)^\circ \quad (5.6)$$

The fact that the expression (5.6) defines a locally monotone functor exactly when F preserves regular epimorphisms and is nearly cartesian is explicitly proved in [32, §4.3]. For the result about natural transformations, there is at most one such $\mathsf{Rel}(\alpha)$ because $\mathcal{G}\alpha = \mathsf{Rel}(\alpha)\mathcal{G}$ entails $\mathsf{Rel}(\alpha)_X = \alpha_X$. The fact that this expression defines a natural transformation between $\mathsf{Rel}(\mathsf{C})$ endofunctors if and only if α is nearly cartesian is proved in [137, Corollary 1.5.7]. $\square$

From the proof of Theorem 5.6 we can retain that if it exists, the *relational extension* of a functor $F : \mathsf{C} \to \mathsf{C}$ is given on a morphism $r = \langle r_1, r_2 \rangle$ by

$$\mathsf{Rel}(F)(r) = Fr_2 \cdot Fr_1{}^{\circ} \tag{5.7}$$

Obviously uniqueness also entails that if functors $\mathsf{Rel}(F)$ and $\mathsf{Rel}(G)$ exist, then $\mathsf{Rel}(GF)$ exists as well and $\mathsf{Rel}(GF) = \mathsf{Rel}(G)\mathsf{Rel}(F)$ (see also [32, § 4.4])

To conclude this chapter, we show how the notions presented here instantiate with $\mathsf{C} = \mathsf{Set}$.

Example 5.7. The category Set is regular, and

- a subobject of a set X is just a subset $U \subseteq X$
- the image of a function $f : X \to Y$ is just the usual direct image $f(X)$
- as regular epimorphisms coincide with split epimorphisms, near pullbacks are just weak pullbacks, therefore
- a nearly cartesian functor is a functor that maps pullbacks to weak pullbacks, i.e. weakly preserves pullbacks – a property known to be equivalent to being weakly cartesian in Set [70, Lemma 2.6]
- a nearly cartesian natural transformation is just a weakly cartesian natural transformation
- the category $\mathsf{Rel}(\mathsf{Set})$ is just the usual category Rel of sets and relations
- the graph functor is the free functor $F_{\mathbf{P}}$ into $\mathsf{Kl}(\mathbf{P})$, computing the graph $f \mapsto \{(x, f(x)) \mid x \in X\}$ and the transpose functor computes the usual converse of a relation i.e. $R \mapsto \{(y, x) \mid (x, y) \in R\}$
- Theorem 5.6 is just Theorem 2.21 and equation (5.7) is just equation (2.34) recalled below for a relation $R \subseteq X \times Y$:

$$\underline{F}R = \{(u, v) \in FX \times FY \mid \exists t \in FR.F\pi_1(t) = u \text{ and } F\pi_2(t) = v\}$$

Chapter 6

Toposes

As seen in the previous chapter, every regular category yields a category of relations and a way to canonically extend functors and natural transformations, provided a few conditions are met. This is materialised by Theorem 5.6, which is the direct generalisation of Theorem 2.21 from the Set case. The general theorem does not involve Kleisli categories anymore because the free relation functor $F_{\mathbf{P}}$ has been replaced with an abstract graph functor $\mathcal{G}$ unrelated to monads. Consequently, the close connection of this result with the theory of distributive laws is seemingly lost. One may ask

When does $\mathcal{G} = F_{\mathbf{S}}$ for some monad $\mathbf{S}$ on C?

It turns out that the answer is: *exactly when C is a topos.* By a topos, we mean an elementary topos in the sense of Lawvere-Tierney. Toposes may be seen as a categorical generalisation of sets: they comprise notions akin to subsets, characteristic functions, powerset, preimage, direct image, intersections, unions, singletons, and of course, relations. In particular, using the constructions which generalise powerset, direct image, singleton and union, we can recover a formal generalisation of the powerset monad in any topos, denoted by $\mathbf{\exists}$. The graph functor then coincides with the free Kleisli functor of this monad, as in Set. Following the literature, we will simply call $\mathbf{\exists}$ the *powerset monad* of the topos. In this chapter, we introduce toposes and their powerset monad to state a generalised extension theorem in Kleisli-style again – our approach being close to the one of Oege de Moor [43, 44, 17]. As a direct consequence, we formulate a criterion to detect monotone (weak) distributive laws of type $\mathbf{T}\mathbf{\exists} \to \mathbf{\exists}\mathbf{T}$. Our main contribution consists in generalising the Set result that there is a unique monotone weak distributive law of type $\mathbf{PP} \to \mathbf{PP}$ to the case $\mathbf{\exists\exists} \to \mathbf{\exists\exists}$. More precisely, this law is proved to be a distributive law if and only if the topos is degenerate.

6.1 Preliminaries

Standard textbooks about toposes include [5, 95, 104, 83]. The introduction presented here is based on the book of Borceux [30].

Cartesian closed category. An exponential object is an object X^Y with a morphism $\mathsf{ev} : X^Y \times Y \to X$ such that for any object Z and any morphism $e : Z \times Y \to X$, there exists a unique morphism $u : Z \to X^Y$ such that

$$\begin{tikzcd} Z \times Y \arrow[d, "u \times \mathrm{id}_Y"'] \arrow[dr, "e"] & \\ X^Y \times Y \arrow[r, "\mathsf{ev}"'] & X \end{tikzcd} \tag{6.1}$$

In the above picture, the morphisms e and u are called *exponential transposes* of each other. A *cartesian closed category* is a category C with finite products and exponentials.

Subobject classifier. Let C be a finitely complete category. Denote its terminal object by 1. For every object X, the unique morphism into the terminal object is denoted by $!_X : X \to 1$. A *subobject classifier* is a pair comprising an object Ω of C and a monomorphism $\mathsf{true} : 1 \hookrightarrow \Omega$ such that for every subobject $m : A \hookrightarrow X$ there is a unique morphism $\chi_m : X \to \Omega$ such that the following diagram is a pullback

$$\begin{tikzcd} A \arrow[r, "!_A"] \arrow[d, hook, "m"'] \arrow[dr, phantom, "\lrcorner", very near start] & 1 \arrow[d, hook, "\mathsf{true}"] \\ X \arrow[r, "\chi_m"'] & \Omega \end{tikzcd} \tag{6.2}$$

The morphism χ_m is called the *characteristic morphism* of m.

Definition 6.1 (Topos)**.** A topos is a category C such that

- C is finitely complete, i.e. has all finite limits
- C is cartesian closed
- C has a subobject classifier

Example 6.2 (Sets)**.** The category Set is a topos. The subobject classifier is $\Omega = \{0, 1\}$, with $\mathsf{true} : 1 \hookrightarrow \{0, 1\}$ picking 1. The characteristic morphism of a subset $A \subseteq X$ is its characteristic function $\chi_A : X \to \{0, 1\}$ defined by $\chi_A(x) = 1$ if and only if $x \in A$. The category FinSet of *finite* sets and functions is also a topos.

Example 6.3 (Graphs [156])**.** The category Graph is a topos. In this category, an object consists of a set of nodes N, a set of arcs A, and two maps $A \to N$ precising respectively the source and the target of every arc. Morphisms are the usual graph homomorphisms i.e. functions on nodes and edges that preserve sources and targets. The subobject classifier Ω of Graph is the graph

$$\begin{tikzcd} \bot_N \arrow[loop left, "\bot_A"] \arrow[r, bend left, "t"] & \top_N \arrow[l, bend left, "s"] \arrow[loop above, "\top_A"] \arrow[loop below, "(s,t)"] \end{tikzcd} \tag{6.3}$$

with $\mathsf{true} : 1 \hookrightarrow \Omega$ picking the node $\top_N$ and the arrow $\top_A$. As explained in [156], given a subgraph $H \subseteq G$, the characteristic morphism $\chi_H : G \to \Omega$ maps to $\top_N$ exactly the nodes of H and to $\top_A$ exactly the arcs of H. Arcs not in H but for which source/target/both are in H are mapped respectively to $s/t/(s,t)$.

Example 6.4 (Presheaves)**.** For any small category C, the functor category $\mathsf{Set}^{\mathsf{C}^{\mathsf{op}}}$ is a topos called the *presheaf topos* on C. Its subobject classifier Ω is the presheaf that sends an object X of C to its *sieves*, that is, the set of subobjects of the presheaf $\mathsf{Hom}_{\mathsf{C}}(-, X)$. The morphism $\mathsf{true} : 1 \to \Omega$ is the natural transformation that picks the maximal sieve. The categories Set and Graph are examples of presheaf toposes, based respectively on the small categories $*$ and $N \underset{t}{\overset{s}{\rightrightarrows}} A$ (where identities are omitted).

Example 6.5 (Nominal sets [117, 115])**.** Let $\sigma\mathbb{N}$ be the group of all bijective functions $\pi : \mathbb{N} \to \mathbb{N}$ such that $\{n \in \mathbb{N} \mid \pi(n) \neq n\}$ is finite. A *nominal set* is a set X equipped with a $\sigma\mathbb{N}$-action $\cdot : \sigma\mathbb{N} \times X \to X$ such that every element $x \in X$ is *finitely suppported*, that is

$$\exists S \subseteq \mathbb{N} \text{ finite such that } (\forall s \in S. \pi(s) = s) \Rightarrow \pi \cdot x = x \tag{6.4}$$

An equivariant function between nominal sets is a function $f : X \to Y$ such that $f(\pi \cdot x) = \pi \cdot f(x)$ for every $x \in X$ and $\pi \in \sigma\mathbb{N}$. The category Nom of nominal sets and equivariant functions is a topos. The subobject classifier is $\Omega = \{0, 1\}$ with the discrete action i.e. $\pi \cdot x = x$ for $x \in \{0, 1\}$, with true picking 1. Subobjects of X are the equivariant subsets $A \subseteq X$, i.e., such that for all $\pi \in \sigma\mathbb{N}$ and $x \in A$, $\pi \cdot x \in A$ – the characteristic morphism $\chi_A : X \to \Omega$ is the usual characteristic function.

Toposes possess several pleasant categorical properties [30, Chapter 5]. In particular, for any topos C the following holds.

- C is finitely cocomplete, i.e., has all finite colimits. The initial object of a topos is denoted by 0.
- C is regular, hence there is a category $\mathsf{Rel}(\mathsf{C})$ of objets and relations.
- C is balanced, i.e. every monomorphic epimorphism is an isomorphism.
- Epimorphisms and regular epimorphisms coincide.

Example 6.6 (Degenerate topos)**.** The category with one object and one morphism is a topos called the degenerate topos. Trivially, $\Omega = 1$. More broadly, any topos such that $0 \cong 1$ is equivalent to the degenerate topos and therefore called degenerate.

The object of subobjects. Let C be a topos with subobject classifier $\mathsf{true} : 1 \hookrightarrow \Omega$. By definition of the subobject classifier, subobjects of an object X are in bijection with morphisms of type $X \to \Omega$. Indeed, a subobject is mapped to its characteristic morphism, and a morphism $X \to \Omega$ is mapped to its pullback along true. Morphisms of type $X \to \Omega$ are themselves in bijection with morphisms of type $1 \to \Omega^X$, by exponential transposition and the isomorphism $1 \times X \cong X$. In categorical parlance, a morphism $1 \to Z$ is called a *global element* of Z. Global elements of Ω^X being equivalently subobjects of X, it makes sense to think about Ω^X as the *object of subobjects of* X. Usually, a morphism $\varphi : X \to \Omega$ is regarded as a formula with one variable in X, and the subobject it classifies is denoted by $\{x : X \mid \varphi(x)\}$.

Intersection, equality, singleton, membership. We sketch how four selected notions of set theory generalise to toposes.

- Consider the morphism $\langle \mathsf{true}, \mathsf{true} \rangle : 1 \hookrightarrow \Omega \times \Omega$. This is a monomorphism whose characteristic morphism will be called conjunction and denoted by $\wedge : \Omega \times \Omega \to \Omega$. Consider any object X and two subobjects $m : A \hookrightarrow X$, $n : B \hookrightarrow X$. The intersection of m and n, denoted by $m \cap n$, is defined as the subobject whose characteristic morphism is

$$X \xrightarrow{\langle \chi_m, \chi_n \rangle} \Omega \times \Omega \xrightarrow{\wedge} \Omega \tag{6.5}$$

- Consider any object X and two subobjects $m : A \hookrightarrow X$, $n : B \hookrightarrow X$. The morphism $\langle \mathrm{id}_X, \mathrm{id}_X \rangle : X \hookrightarrow X \times X$ is a monomorphism whose characteristic morphism will be denoted by $=_X : X \times X \to \Omega$. Notably, for two global elements $a, b : 1 \to X$, the global element $a =_X b$ defined by

$$1 \xrightarrow{\langle a,b \rangle} X \times X \xrightarrow{=_X} \Omega \tag{6.6}$$

is equal to true if and only if $a = b$.

- Let X be an object. Taking the exponential transpose of the equality morphism $=_X : X \times X \to \Omega$, we obtain a morphism that will be denoted by $\{-\}_X : X \to \Omega^X$ and generalises the notion of singleton. Notably, for any global element $a : 1 \to X$, the global element

$$1 \xrightarrow{a} X \xrightarrow{\{-\}_X} \Omega^X \tag{6.7}$$

corresponds to the subobject $a : 1 \hookrightarrow X$.

- Let X be an object. The exponential transpose of the identity $\mathrm{id}_{\Omega^X} : \Omega^X \to \Omega^X$ is denoted by $\in_X : X \times \Omega^X \to \Omega$ and its corresponding subobject of $X \times \Omega^X$ generalises the membership relation.

Internal logic. Toposes have access to a rich internal logic, stemming from the above constructions and similar constructions for disjunction, implication, negation and quantifiers. Every valid statement of intuitionistic set theory holds in any topos. By intuitionistic set theory, we mean intuitionistic predicate calculus with relations $=$ and $\in$ satisfying the most undisputed axioms of set theory (extensionality, pair, union, powerset, replacement and comprehension, empty set). As toposes are only assumed to have finite (co)limits, the logic is finitist i.e. one can build explicitly finite sets, but not infinite ones – in particular note that there is no axiom of infinity. The logic being constructive, the law of excluded middle cannot be used in general, as well as the axiom of choice. By contrast, set extensionality and functional extensionality hold [30, Theorem 6.9.2, Proposition 6.10.2]. The internal logic is a powerful tool whose rigorous presentation cannot be contained in just a few pages. We will use it both to introduce some known constructions such as the powerset monad of toposes, and to prove generalisations of Set results about them. We encourage the non-familiar reader to either consider the internal logic arguments at the intuitive level of *doing intuitionistic set-theoretic logic in an abstract categorical framework*, or to first read [30, Chapter 6] for a gentle, yet fully rigorous, introduction to the internal logic.

Mitchell-Bénabou language. We very briefly sketch the language of the internal logic and relate some simple formulas to the underlying categorical constructions. For technical details, e.g. variable layout, the reader is again referred to [30]. Terms and formulas of the Mitchell-Bénabou language are defined inductively and given an interpretation. Each term t has a type which is an object X of the topos, with notation $t : X$. The terms of type Ω are called formulas. The interpretation of a term t of type X is a morphism $[t]$ of codomain X, and whose domain stands for variable context. Note that the interpretation of a formula is, as expected, a morphism into Ω.

- For each object X, there are variable terms $x : X$ interpreted as $[x] = \mathrm{id}_X : X \to X$. When seen in a context comprising more variables, the interpretation is a projection, e.g. if the term x is considered to have free variables $x : X$, $y : Y$ then $[x] = \pi_1 : X \times Y \to X$ is the canonical projection. In the sequel, considerations about variable context are left implicit.
- For each global element $c : 1 \to X$, there is a constant term $c : X$ interpreted as $[c] = c : 1 \to X$.
- For each morphism $f : X \to Y$ and each term $t : X$, there is a term $f(t) : Y$ interpreted as $[f(t)] = f \circ [t]$.
- For terms $t_1 : X$, $t_2 : Y$, there is a term $\langle t_1, t_2 \rangle : X \times Y$ interpreted as $[\langle t_1, t_2 \rangle] = \langle [t_1], [t_2] \rangle$.
- For a formula $\varphi : \Omega$ with free variables x, y, there is a term $\{x : X \mid \varphi(x, y)\} : \Omega^X$ with free variable y interpreted as the exponential transpose $[\{x : X \mid \varphi(x, y)\}] : Y \to \Omega^X$ of $[\varphi] : X \times Y \to \Omega$.
- For terms $t_1, t_2 : X$, the formula $t_1 = t_2$ is interpreted as $[t_1 = t_2] = (=_X \circ \langle [t_1], [t_2] \rangle)$.
- For terms $s : X$, $t : \Omega^X$, the formula $s \in t$ is interpreted as $[s \in t] = \in_X \circ \langle [s], [t] \rangle$.
- Formulas can be composed using Boolean operators and quantifiers, interpreted by categorical constructions of the topos, e.g. conjunction is interpreted using the morphism $\wedge : \Omega \times \Omega \to \Omega$.

Example 6.7. Given subobjects $m : A \hookrightarrow X$ and $n : B \hookrightarrow X$ and a variable $x : X$, the subobject $\{x : X \mid \chi_m(x) \wedge \chi_n(x)\}$ is classified by the interpretation of the formula $\chi_m(x) \wedge \chi_n(x)$, i.e.

$$\begin{aligned}
[\chi_m(x) \wedge \chi_n(x)] &= \wedge\langle[\chi_m(x)], [\chi_n(x)]\rangle \\
&= \wedge\langle\chi_m \circ [x], \chi_n \circ [x]\rangle \\
&= \wedge\langle\chi_m, \chi_n\rangle
\end{aligned}$$

which by definition corresponds to the intersection of subobjects $m \cap n$. More generally, for φ, ψ formulas, $\{x : X \mid \varphi(x) \wedge \psi(x)\} = \{x : X \mid \varphi(x)\} \cap \{x : X \mid \psi(x)\}$.

Using the Mitchell-Bénabou language, many categorical constructions can be expressed similarly to how they would be in Set. For example, the pullback of two morphisms $f : X \to Z$ and $g : Y \to Z$ is the subobject $\{(x, y) : X \times Y \mid f(x) = g(y)\}$ [30, Proposition 6.10.3].

Valid formula. A formula φ of the Mitchell-Bénabou language with one free variable $x : X$ is *valid*, written $x : X \vdash \varphi(x)$, if $[\varphi] = X \xrightarrow{!_X} 1 \xrightarrow{\mathsf{true}} \Omega$.

Example 6.8. For a morphism $f : X \to Y$, let us prove that

$$x : X \vdash f(x) = f(x) \tag{6.8}$$

Using the inductive definition of the interpretation, the formula $f(x) = f(x)$ is interpreted as $=_Y \circ\langle f, f\rangle$. We must therefore prove that $=_Y \circ\langle f, f\rangle = \mathsf{true}\circ !_X$. The wanted equation is a consequence of commutation of the following diagram, where the left square commutes trivially and the right square is the definition of $=_Y$.

$$\begin{tikzcd}
X \arrow[r, "f"] \arrow[d, "\mathrm{id}_X"'] & Y \arrow[r, "!_Y"] \arrow[d, "\langle \mathrm{id}_Y, \mathrm{id}_Y\rangle"'] & 1 \arrow[d, "\mathsf{true}"] \\
X \arrow[r, "\langle f, f\rangle"'] & Y \times Y \arrow[r, "=_Y"'] & \Omega
\end{tikzcd} \tag{6.9}$$

Example 6.9. For an object X, let us prove that $x : X \vdash \{x\}_X = \{x' : X \mid x = x'\}$. Using the inductive definition of the interpretation,

$$[\{x\}_X = \{x' : X \mid x = x'\}] \tag{6.10}$$

$$= =_{\Omega^X} \circ\langle[\{x\}_X], [\{x' : X \mid x = x'\}]\rangle \tag{6.11}$$

$$= =_{\Omega^X} \circ\langle\{-\}_X, \{-\}_X\rangle \tag{6.12}$$

We detail the last step. The interpretation $[\{x' : X \mid x = x'\}] : X \to \Omega^X$ is the exponential transpose of the interpretation $[x = x']$ i.e. $=_X\colon X \times X \to \Omega$, which is precisely the definition of $\{-\}_X$. According to the previous example, the expression (6.12) is equal to $\mathsf{true} \circ !_X$. Therefore $x : X \vdash \{x\}_X = \{x' : X \mid x = x'\}$ holds.

Example 6.10. For a morphism $f : X \to Z$, the property f *is an epimorphism* is equivalent to the following statement in the internal logic [30, Proposition 6.10.2]

$$\vdash \forall(z : Z).\exists(x : X).f(x) = z \tag{6.13}$$

6.2 The Powerset Monad

In this section we define the powerset monad in a topos.

Theorem 6.11 (Freyd [51, §1.911])**.** *Let* C *be a regular category. Then the following are equivalent*

- *the graph functor* $\mathcal{G} : \mathsf{C} \to \mathsf{Rel}(\mathsf{C})$ *has a right adjoint*
- C *is a topos*

Oege de Moor [43, §6.1.1] even uses this property as a definition of toposes – we now follow his approach and notation.

Definition 6.12. The *powerset monad of a topos* C is the monad obtained from the adjunction whose left adjoint is $\mathcal{G}$. It is denoted by $\boldsymbol{\exists}$.

As every free Kleisli functor is a left adjoint, Theorem 6.11 entails that if the graph functor of a regular category C is the free Kleisli functor of some monad, then C is a topos. The converse holds [43, §6.1.10]: in any topos C, $\mathsf{Rel}(\mathsf{C})$ is isomorphic to $\mathsf{Kl}(\boldsymbol{\exists})$. Under this isomorphism we can identify $\mathcal{G} = F_{\boldsymbol{\exists}}$.

The powerset monad of toposes is well-known, and its constructions arise at various places in the literature. As a monad, it appears in many papers of René Guitart under the appellation *the involutive monad in a topos* [66, 67, 68, 69]. According to Guitart [66, Corollaire 1.1.1.b], this monad was first exhibited by Lawvere and Tierney when they introduced (elementary) toposes in 1970 [65]. We now give an explicit description of the monad. For proofs of its expression in the internal logic, see [111, Propositions 4.9, 4.17, 4.19].

- On objects, $\boldsymbol{\exists}X = \Omega^X$ is the object of subobjects of X.

- On a morphism $f : X \to Y$, $\mathbf{\exists} f : \Omega^X \to \Omega^Y$ is known as the existential image, and is internally described by

$$(a : \Omega^X) \vdash \mathbf{\exists} f(a) = \{y : Y \mid \exists x : X.x \in a \wedge f(x) = y\} \tag{6.14}$$

- The unit $\eta^{\mathbf{\exists}}_X : X \to \Omega^X$ is the singleton morphism $\{-\}_X$. Recall that it is the exponential transpose of the characteristic morphism of $\langle \mathrm{id}_X, \mathrm{id}_X \rangle : X \hookrightarrow X \times X$, and that we proved in Example 6.9 that an internal description is given by

$$(x : X) \vdash \eta^{\mathbf{\exists}}_X(x) = \{x' : X \mid x = x'\} \tag{6.15}$$

- The multiplication $\mu^{\mathbf{\exists}}_X : \Omega^{\Omega^X} \to \Omega^X$ is most conveniently defined using the internal logic, as

$$(t : \Omega^{\Omega^X}) \vdash \mu^{\mathbf{\exists}}_X(t) = \{x : X \mid \exists s : \Omega^X.x \in s \wedge s \in t\} \tag{6.16}$$

Example 6.13. In Set, the monad $\mathbf{\exists}$ coincides with the usual powerset monad $\mathbf{P}$.

Example 6.14. The powerset monad of a presheaf topos is just the powerset monad of Set computed pointwise. For example, let us consider the topos Graph of graphs and graph homomorphisms. Let $G = (N, A)$ be a graph, with $s : A \to N$ being the source map and $t : A \to N$ being the target map.

- Nodes of Ω^G correspond to subsets $U \subseteq N$. From node U to node V, there is one arc per subset $B \subseteq A$ such that $s(B) \subseteq U$ and $t(B) \subseteq V$.
- Given a graph homomorphism $f : G \to H$, $\mathbf{\exists} f : \Omega^G \to \Omega^H$ maps
  - a node U to the node $f(U)$
  - an arc $U \xrightarrow{A} V$ to the arc $f(U) \xrightarrow{f(A)} f(V)$
- The unit $\eta^{\mathbf{\exists}}_G : G \to \Omega^G$ maps
  - a node x to the node $\{x\}$
  - an arc $x \xrightarrow{a} y$ to the arc $\{x\} \xrightarrow{\{a\}} \{y\}$
- The multiplication $\mu^{\mathbf{\exists}}_G : \Omega^{\Omega^G} \to \Omega^G$ maps
  - a node $\mathcal{U}$ to the node $\bigcup \mathcal{U}$
  - an arc $\mathcal{U} \xrightarrow{\mathcal{A}} \mathcal{V}$ to the arc $\bigcup \mathcal{U} \xrightarrow{\bigcup \mathcal{A}} \bigcup \mathcal{V}$

Example 6.15. In the topos Nom of nominal sets and equivariant functions, $\exists X = \Omega^X$ is the set of all finitely supported subsets of X, with action given by $\pi \cdot U = \{\pi \cdot x \mid x \in U\}$ for every $U \in \Omega^X$ and $\pi \in \sigma\mathbb{N}$. The rest of the monad is defined as for the powerset monad. This monad has been used recently in [160].

Relational extensions are exactly monotone extensions (in the sense of Definition 1.40) in the Kleisli category of the powerset monad. We recall that the qualifier *monotone* is with respect to the order on relations inherited from the subobject order. This identification results in recovering a theorem generating monotone (weak) distributive laws.

Theorem 6.16. *For every monad* $\mathbf{T}$ *on a topos* C*,*

- *there exists at most one monotone distributive law of type* $\mathbf{T}\exists \to \exists\mathbf{T}$*, and there is one if and only if* T *preserves epimorphisms and* T*,* $\eta^{\mathbf{T}}$ *and* $\mu^{\mathbf{T}}$ *are nearly cartesian;*

- *there exists at most one monotone weak distributive law of type* $\mathbf{T}\exists \to \exists\mathbf{T}$*, and there is one if and only if* T *preserves epimorphisms and* T *and* $\mu^{\mathbf{T}}$ *are nearly cartesian.*

Proof. Apply Theorem 5.6 using $\mathcal{G} = F_{\exists}$ and $\mathsf{Rel}(\mathsf{C}) = \mathsf{Kl}(F_{\exists})$ to get, under the conditions above, existence and uniqueness of a locally monotone (weak) extension. Equivalently, this is a monotone (weak) distributive law. $\square$

The restriction of Theorem 6.16 for distributive laws $T\exists \to \exists T$ already appears in [43, Theorem 6.2.4].

6.3 Powerset over Powerset

In this section, we apply Theorem 6.16 to $\mathbf{T} = \exists$. New results are marked by the symbol ✓. They are generalisations of known Set statements that can be extended to arbitrary toposes because they only rely on internal constructive arguments. To provide concrete evidence of this intuitionistic flavour, all results marked by the symbol ✓ have been derived using the proof assistant Coq [76]. The idea underlying this formalisation is that the fragment of Coq comprising first-order logic and the basic set-theoretical constructions required for defining the powerset monad is coherent with the intuitionistic set theory internal to toposes. In particular, the subobject classifier Ω is represented by the type `Prop`. No external libraries are required. The code is provided in Appendix B and on Github [61].

Assumptions about the functor $\exists$ have already been verified in [43, Proposition 6.5.1] – in fact, de Moor proves the stronger result that $\exists$ is weakly cartesian.

Proposition 6.17 (de Moor). *The powerset functor $\exists$ preserves epimorphisms and is nearly cartesian.*

So we already know there is a monotone distributive law $\boldsymbol{\exists}\exists \to \exists\boldsymbol{\exists}$ between the functor $\exists$ and the monad $\boldsymbol{\exists}$. We provide the missing results in order to find a weak distributive law bewteen monads.

Proposition 6.18 ($\checkmark$ = `eta_nearly_cartesian`). *The unit $\eta^{\boldsymbol{\exists}}$ is nearly cartesian if and only if* C *is degenerate.*

Proof (sketch). There is an apprehensible proof using mainly the categorical definition of $\eta^{\boldsymbol{\exists}}$. We provide a sketch – details are postponed to Section 6.3.1. If C is degenerate, then every natural transformation is trivially nearly cartesian. Conversely, assume $\eta^{\boldsymbol{\exists}}$ is nearly cartesian. Components of $\eta^{\boldsymbol{\exists}}$ are monomorphisms, inducing that $\eta^{\boldsymbol{\exists}}$ is cartesian i.e. naturality squares are pullbacks. In particular, we have the following pullback

$$\begin{tikzcd} \Omega \arrow[r, "!_\Omega"] \arrow[d, "\eta^{\boldsymbol{\exists}}_\Omega"'] \arrow[dr, phantom, "\lrcorner", very near start] & 1 \arrow[d, "\eta^{\boldsymbol{\exists}}_1"] \\ \Omega^\Omega \arrow[r, "\exists !_\Omega"'] & \Omega \end{tikzcd} \tag{6.17}$$

Note that when $\mathsf{C} = \mathsf{Set}$, $\Omega \cong \{0,1\}$ and $1 \cong \{0\}$, so this square is the usual counterexample of Proposition 3.2. Let $\mathsf{full} : 1 \to \Omega^\Omega$ be the morphism that picks the maximal subobject of Ω. The pasting law for pullbacks yields $P \cong Q$, where P and Q are defined by the following pullbacks

$$\begin{tikzcd} P \arrow[r] \arrow[d, "!_P"'] \arrow[dr, phantom, "\lrcorner", very near start] & \Omega \arrow[d, "\eta^{\boldsymbol{\exists}}_\Omega"] \\ 1 \arrow[r, "\mathsf{full}"'] & \Omega^\Omega \end{tikzcd} \qquad \begin{tikzcd} Q \arrow[r, "!_Q"] \arrow[d, "!_Q"'] \arrow[dr, phantom, "\lrcorner", very near start] & 1 \arrow[d, "\eta^{\boldsymbol{\exists}}_1"] \\ 1 \arrow[r, "\exists !_\Omega \circ \mathsf{full}"'] & \Omega \end{tikzcd} \tag{6.18}$$

Furthermore, one can prove that $P \cong P \times \Omega$ and $Q \cong 1$. Combining these results yields $\Omega \cong 1$, and this entails that the topos C is degenerate. $\square$

It is noteworthy that the Set counterexample does still generate a counterexample in an arbitrary non-degenerate topos. A consequence is that if C is not degenerate, there is no *monotone* distributive law of type $\boldsymbol{\exists\exists} \to \boldsymbol{\exists\exists}$. More generally

Proposition 6.19 (✓ = `dlaw_degenerate`)**.** *There is a distributive law of type* $\mathbf{\exists\exists} \to \mathbf{\exists\exists}$ *if and only if* C *is degenerate.*

Proof. The $\mathbf{PP} \to \mathbf{PP}$ counterexample of Klin and Salamanca [91] can be defined internally, because it is based on two finite sets, interpreted as finite coproducts in toposes. The rest of their proof only relies on constructive arguments. Their proof has been formalised step by step in our Coq file. □

Proposition 6.20 (✓ = `mu_nearly_cartesian`)**.** *The multiplication* $\mu^{\mathbf{\exists}}$ *is nearly cartesian.*

Proof. Our proof consists in mimicking the intuitionistic proof of the Set case. We provide details in Section 6.3.2. □

In conclusion

Theorem 6.21. *In any topos* C*, there is a unique monotone weak distributive law of type* $\mathbf{\exists\exists} \to \mathbf{\exists\exists}$*. This is a distributive law if and only if* C *is degenerate.*

Proof. Apply Theorem 6.16 using Propositions 6.17, 6.18 and 6.20. □

Proposition 6.22 (✓ = `monotone_weak_dlaw`)**.** *In the internal logic, the unique monotone weak distributive law* λ *of type* $\mathbf{\exists\exists} \to \mathbf{\exists\exists}$ *can be expressed as*

$$\begin{aligned}(t : \Omega^{\Omega^X}) \vdash \lambda_X(t) = \{s : \Omega^X \mid {} & (\forall (x : X), x \in s \to x \in \mu^{\mathbf{\exists}}_X(t)) \\ & \wedge \forall (u : \Omega^X). u \in t \to \exists (x : X). x \in u \wedge x \in s\}\end{aligned} \tag{6.19}$$

Proof. In [43, Proposition 6.5.7], de Moor shows that the extension $\underline{\exists}$ of $\exists$ to $\mathsf{Rel}(\mathsf{C})$ can be computed using the Egli-Milner formula. Internally, this means that for every subobject r of $X \times Y$, written $r = \{(x, y) : X \times Y \mid (x, y) \in r\}$,

$$\underline{\exists} r = \{(u, v) : \Omega^X \times \Omega^Y \mid \forall x : X, x \in u \to \exists y : Y, y \in v \wedge (x, y) \in r \tag{6.20}$$

$$\wedge \forall y : Y, y \in v \to \exists x : X, x \in u \wedge (x, y) \in r\} \tag{6.21}$$

Then, computing the morphism whose graph is $\underline{\exists}(\ni_X)$ leads constructively to the expression of equation (6.19). □

Example 6.23. In the topos Graph, this law consists in computing the Set law $\lambda : \mathbf{PP} \to \mathbf{PP}$ on both nodes and arcs, i.e., it maps

- a node $\mathcal{U}$ to the node $\lambda_N(\mathcal{U})$

- an arc $\mathcal{U} \xrightarrow{\mathcal{A}} \mathcal{V}$ to the arc $\lambda_N(\mathcal{U}) \xrightarrow{\lambda_A(\mathcal{A})} \lambda_N(\mathcal{V})$

Example 6.24. In the topos Nom, according to the *finite support principle* [117, §2.5], the internal logic stands for the usual logic of Set provided any quantification over functions or subsets is restricted to the finitely supported ones. Therefore, the monotone weak distributive law $\lambda : \exists\exists \to \exists\exists$ expresses as follows for all $\mathcal{U}$ finitely supported subset of the finitely supported subsets of the nominal set X,

$$\lambda_X(\mathcal{U}) = \Big\{V \subseteq X \textit{ finitely supported} \mid V \subseteq \bigcup \mathcal{U} \text{ and } \forall U \in \mathcal{U}. U \cap V \neq \emptyset\Big\} \tag{6.22}$$

6.3.1 Proof for the Unit

In this section we provide a complete proof that the unit $\eta^{\exists}$ is nearly cartesian if and only if the topos is degenerate. A natural transformation is *cartesian* if its naturality squares are pullbacks.

Lemma 6.25. *The unit $\eta^{\exists}$ is nearly cartesian if and only if it is cartesian.*

Proof. Clearly, cartesian entails nearly cartesian. For the other direction, consider the naturality square of a morphism $f : X \to Y$ and the pullback of $\eta^{\exists}_Y$ along $\exists f$

$$\begin{tikzcd}
X \arrow[dr, "h"] \arrow[drr, bend left, "\eta^{\exists}_X"] \arrow[ddr, bend right, "f"'] & & \\
& P \arrow[r, "p_1"] \arrow[d, "p_2"'] \arrow[dr, phantom, "\lrcorner", very near start] & \Omega^X \arrow[d, "\exists f"] \\
& Y \arrow[r, "\eta^{\exists}_Y"'] & \Omega^Y
\end{tikzcd} \tag{6.23}$$

As $\eta^{\exists}$ is nearly cartesian, h is an epimorphism. Components of $\eta^{\exists}$ are monomorphisms [97, Chapter IV, Lemma 1], so from $\eta^{\exists}_X = p_1 \circ h$ we deduce that the morphism h is a monomorphism as well. A topos is balanced, hence h is an isomorphism and the naturality square of f is a pullback. $\square$

In the rest of the proof we often use the identification $\Omega \cong 1 \times \Omega$. The monomorphism $\mathrm{id}_\Omega : \Omega \hookrightarrow \Omega$, representing the maximal subobject of Ω, has a characteristic morphism $\chi_{\mathrm{id}_\Omega} : \Omega \to \Omega$ whose exponential transpose is denoted by $\mathsf{full} : 1 \to \Omega^\Omega$. Seen as a global element, full picks the maximal element of the poset of subobjects Ω^Ω i.e.

$$x : 1 \vdash \mathsf{full}(x) = \{\omega : \Omega \mid \mathsf{true}\} \tag{6.24}$$

Consider the following pullbacks

$$\begin{tikzcd} P \arrow[r, "p"] \arrow[d, "!_P"'] \arrow[dr, phantom, "\lrcorner", very near start] & \Omega \arrow[d, "\eta^{\exists}_{\Omega}"] \\ 1 \arrow[r, "\mathsf{full}"'] & \Omega^{\Omega} \end{tikzcd} \qquad \begin{tikzcd} Q \arrow[r, "!_Q"] \arrow[d, "!_Q"'] \arrow[dr, phantom, "\lrcorner", very near start] & 1 \arrow[d, "\eta^{\exists}_{1}"] \\ 1 \arrow[r, "\exists!_{\Omega}\circ\mathsf{full}"'] & \Omega \end{tikzcd} \tag{6.25}$$

Lemma 6.26. *If $\eta^{\exists}$ is cartesian, then $P \cong Q$.*

Proof. In the diagram

$$\begin{tikzcd} P \arrow[r, "p"] \arrow[d, "!_P"'] \arrow[dr, phantom, "\lrcorner", very near start] & \Omega \arrow[r, "!_{\Omega}"] \arrow[d, "\eta^{\exists}_{\Omega}"'] \arrow[dr, phantom, "\lrcorner", very near start] & 1 \arrow[d, "\eta^{\exists}_{1}"] \\ 1 \arrow[r, "\mathsf{full}"'] & \Omega^{\Omega} \arrow[r, "\exists!_{\Omega}"'] & \Omega \end{tikzcd} \tag{6.26}$$

the left square is the pullback that defines P and the right square is the pullback naturality diagram for $!_{\Omega} : \Omega \to 1$. Therefore, by the pasting law for pullbacks, the whole rectangle is itself a pullback of $\eta^{\exists}_{1}$ and $\exists!_{\Omega} \circ \mathsf{full}$, so that $P \cong Q$. ☐

Lemma 6.27. $Q \cong 1$

Proof. Note that 1 is the pullback of a cospan $1 \longrightarrow Z \longleftarrow 1$ if and only if the two morphisms of the cospan are equal. The fact $Q \cong 1$ is therefore equivalent to $\eta^{\exists}_{1} = \exists!_{\Omega} \circ \mathsf{full}$. This can be proved in the internal logic using the fact that 1 plays the role of the unit type, see [111, 5.4].

$$\begin{aligned} & x : 1 \vdash \eta^{\exists}_{1}(x) = \exists!_{\Omega}(\mathsf{full}(x)) \\ \iff & x : 1, y : 1 \vdash y \in \eta^{\exists}_{1}(x) \leftrightarrow y \in \exists!_{\Omega}(\mathsf{full}(x)) && \text{(extensionality)} \\ \iff & x : 1, y : 1 \vdash y = x \leftrightarrow \exists\omega : \Omega.\omega \in \mathsf{full}(x) \wedge y = !_{\Omega}(\omega) && \text{(definition of } \eta^{\exists} \text{ and } \exists) \\ \iff & x : 1 \vdash \exists\omega : \Omega.\omega \in \mathsf{full}(x) && \text{(1 unit type)} \\ \iff & \vdash \exists\omega : \Omega.\mathsf{true} && \text{(definition of } \mathsf{full}) \end{aligned}$$

and the latest assumption holds in any topos. ☐

Lemma 6.28. $P \cong P \times \Omega$

Proof. First, consider

$$\begin{tikzcd} P\times\Omega \arrow[r, "!_P\times \mathrm{id}_{\Omega}"] \arrow[d, hook, "\mathrm{id}_{P\times\Omega}"'] \arrow[dr, phantom, "\lrcorner", very near start] & \Omega \arrow[r, "!_{\Omega}"] \arrow[d, hook, "\mathrm{id}_{\Omega}"'] \arrow[dr, phantom, "\lrcorner", very near start] & 1 \arrow[d, hook, "\mathsf{true}"] \\ P\times\Omega \arrow[r, "!_P\times \mathrm{id}_{\Omega}"'] & \Omega \arrow[r, "\chi_{\mathrm{id}_{\Omega}}"'] & \Omega \end{tikzcd} \tag{6.27}$$

The right square of the diagram above is a pullback by definition of $\chi_{\mathrm{id}_\Omega}$. The left square can be easily checked to be a pullback along the identity morphism. By pasting pullbacks, we get that the outer rectangle is a pullback. Second, consider

$$\begin{tikzcd}
P \arrow[r, "p"] \arrow[d, hook, "\langle \mathrm{id}_P, p\rangle"'] \arrow[dr, phantom, "\lrcorner", very near start] & \Omega \arrow[r, "!_\Omega"] \arrow[d, hook, "\langle \mathrm{id}_\Omega, \mathrm{id}_\Omega\rangle"'] \arrow[dr, phantom, "\lrcorner", very near start] & 1 \arrow[d, hook, "\mathsf{true}"] \\
P \times \Omega \arrow[r, "p \times \mathrm{id}_\Omega"'] & \Omega \times \Omega \arrow[r, "\chi_{\langle \mathrm{id}_\Omega, \mathrm{id}_\Omega\rangle}"'] & \Omega
\end{tikzcd} \tag{6.28}$$

The right square of the diagram above is a pullback by definition of $\chi_{\langle \mathrm{id}_\Omega, \mathrm{id}_\Omega\rangle}$. The left square is also a pullback, because it trivially commutes and for every X as in the below left diagram, $u : X \to P$ is the unique morphism making the below right diagram commute.

$$\begin{tikzcd}
X \arrow[r, "w"] \arrow[d, "\langle u,v\rangle"'] & \Omega \arrow[d, "\langle \mathrm{id}_\Omega, \mathrm{id}_\Omega\rangle"] \\
P \times \Omega \arrow[r, "p \times \mathrm{id}_\Omega"'] & \Omega \times \Omega
\end{tikzcd}
\qquad
\begin{tikzcd}
X \arrow[dr, "u"] \arrow[drr, bend left, "w"] \arrow[ddr, bend right, "\langle u,v\rangle"'] & & \\
& P \arrow[r, "p"] \arrow[d, "\langle \mathrm{id}_P, p\rangle"'] \arrow[dr, phantom, "\lrcorner", very near start] & \Omega \arrow[d, "\langle \mathrm{id}_\Omega, \mathrm{id}_\Omega\rangle"] \\
& P \times \Omega \arrow[r, "p \times \mathrm{id}_\Omega"'] & \Omega \times \Omega
\end{tikzcd} \tag{6.29}$$

Hence, by pasting pullbacks, the outer rectangle in (6.28) is a pullback. Let $\mathsf{ev} : \Omega^\Omega \times \Omega \to \Omega$ be the evaluation morphism from the cartesian closed structure of C. By definition of full as the exponential transpose of $\chi_{\mathrm{id}_\Omega}$ and $\eta^{\exists}_\Omega$ as the exponential transpose of $\chi_{\langle \mathrm{id}_\Omega, \mathrm{id}_\Omega\rangle}$, we have

$$\mathsf{ev} \circ (\mathsf{full} \times \mathrm{id}_\Omega) = \chi_{\mathrm{id}_\Omega} \tag{6.30}$$

$$\mathsf{ev} \circ (\eta^{\exists}_\Omega \times \mathrm{id}_\Omega) = \chi_{\langle \mathrm{id}_\Omega, \mathrm{id}_\Omega\rangle} \tag{6.31}$$

Starting from commutation of the pullback square defining P, we have

$$\mathsf{full} \circ !_P = \eta^{\exists}_\Omega \circ p$$

$$(\mathsf{full} \times \mathrm{id}_\Omega) \circ (!_P \times \mathrm{id}_\Omega) = (\eta^{\exists}_\Omega \times \mathrm{id}_\Omega) \circ (p \times \mathrm{id}_\Omega)$$

$$\chi_{\mathrm{id}_\Omega} \circ (!_P \times \mathrm{id}_\Omega) = \chi_{\langle \mathrm{id}_\Omega, \mathrm{id}_\Omega\rangle} \circ (p \times \mathrm{id}_\Omega) \qquad \text{(apply } \mathsf{ev} \text{ on the left)}$$

The last equality shows that the outer squares of Diagrams (6.27) and (6.28) are isomorphic pullbacks, so that $P \times \Omega \cong P$. □

A last lemma is needed

Lemma 6.29. *In any topos, if* $1 \cong \Omega$ *then* $0 \cong 1$.

Proof. Assuming $1 \cong \Omega$, there is only one morphism of type $1 \to \Omega$. The following squares, classifying respectively the monomorphisms $!_1 : 1 \hookrightarrow 1$ and $!_0 : 0 \hookrightarrow 1$, are then pullbacks of the same cospan, so $0 \cong 1$.

$$\begin{tikzcd} 1 \arrow[r, "!_1"] \arrow[d, hook, "!_1"'] \arrow[dr, phantom, "\lrcorner", very near start] & 1 \arrow[d, hook, "\mathsf{true}"] \\ 1 \arrow[r, "\chi_{!_1}"'] & \Omega \end{tikzcd} \qquad \begin{tikzcd} 0 \arrow[r, "!_0"] \arrow[d, hook, "!_0"'] \arrow[dr, phantom, "\lrcorner", very near start] & 1 \arrow[d, hook, "\mathsf{true}"] \\ 1 \arrow[r, "\chi_{!_0}"'] & \Omega \end{tikzcd} \tag{6.32}$$

□

We finally prove

Theorem 6.30. *The unit $\eta^{\exists}$ is nearly cartesian if and only if the topos is degenerate.*

Proof. In a degenerate topos, any natural transformation is trivially nearly cartesian. Conversely, assume $\eta^{\exists}$ is nearly cartesian. Then by Lemma 6.25, $\eta^{\exists}$ is cartesian. By Lemmas 6.26, 6.27 and 6.28, we have $P \cong Q$, $Q \cong 1$ and $P \cong P \times \Omega$, therefore $1 \cong \Omega$. By Lemma 6.29, $0 \cong 1$ so the topos is degenerate. □

6.3.2 Proof for the Multiplication

In this section we provide a complete proof in the internal logic that the multiplication $\mu^{\exists}$ is nearly cartesian. In fact, we even prove that it is weakly cartesian. As opposed to the unit case where a counterexample had to be found, the internal logic is particularly well-suited for such a positive result. Moreover, not having to manipulate the explicit categorical definition of the multiplication is highly practical.

Let $f : X \to Y$ be a morphism in C. The pullback P of $\exists f : \exists X \to \exists Y$ and $\mu^{\exists}_Y : \exists\exists Y \to \exists Y$ is a subobject of $\Omega^X \times \Omega^{\Omega^Y}$ described internally by

$$P = \{(s,t) : \Omega^X \times \Omega^{\Omega^Y} \mid \exists f(s) = \mu^{\exists}_Y(t)\} \tag{6.33}$$

so that

$$(s,t) : P \vdash \exists f(s) = \mu^{\exists}_Y(t) \tag{6.34}$$

Proving that the universal morphism $h : \Omega^{\Omega^X} \to P$ is a split epimorphism is equivalent to finding a $k : P \to \Omega^{\Omega^X}$ such that $\mu^{\exists}_X \circ k = p_1$ and $\exists\exists f \circ k = p_2$, where $p_1 : P \to \Omega^X$, $p_2 : P \to \Omega^{\Omega^Y}$ are the projections. In the Set case (Proposition 3.3), we used to define

$$(U, \mathcal{V}) \mapsto \{U \cap f^{-1}(V) \mid V \in \mathcal{V}\}$$

Let us implement this idea in the internal logic by defining

$$(s,t) : P \vdash k(s,t) = \{u : \Omega^X \mid \exists v : \Omega^Y . v \in t \wedge u = s \cap f^{-1}(v)\} \tag{6.35}$$

where the intersection and the preimage are respectively defined by

$$u : \Omega^X, u' : \Omega^X \vdash u \cap u' = \{x : X \mid x \in u \wedge x \in u'\} \tag{6.36}$$

$$v : \Omega^Y \vdash f^{-1}(v) = \{x : X \mid f(x) \in v\} \tag{6.37}$$

It only remains to prove the following judgements

$$(s,t) : P \vdash \mu^{\exists}_X(k(s,t)) = s$$

$$(s,t) : P \vdash \exists\exists f(k(s,t)) = t$$

Proposition 6.31. $(s,t) : P \vdash \mu^{\exists}_X(k(s,t)) = s$

Proof.

$$\begin{aligned}
&(s,t) : P \vdash \mu^{\exists}_X(k(s,t)) = s \\
&\iff \quad \text{(extensionality)} \\
&(s,t) : P, x : X \vdash x \in \mu^{\exists}_X(k(s,t)) \leftrightarrow x \in s \\
&\iff \quad (\mu^{\exists} \text{ definition}) \\
&(s,t) : P, x : X \vdash (\exists u : \Omega^X . x \in u \wedge u \in k(s,t)) \leftrightarrow x \in s \\
&\iff \quad (k \text{ definition}) \\
&(s,t) : P, x : X \vdash (\exists u : \Omega^X . x \in u \wedge \exists v : \Omega^Y . v \in t \wedge u = f^{-1}(v) \cap s) \leftrightarrow x \in s \\
&\iff \quad (\text{simplification of } \exists u : \Omega^X . u = f^{-1}(v) \cap s) \\
&(s,t) : P, x : X \vdash (\exists v : \Omega^Y . x \in f^{-1}(v) \cap s \wedge v \in t) \leftrightarrow x \in s \\
&\iff \quad (\text{definition of } f^{-1} \text{ and } \cap) \\
&(s,t) : P, x : X \vdash (\exists v : \Omega^Y . f(x) \in v \wedge x \in s \wedge v \in t) \leftrightarrow x \in s \\
&\iff \quad (\text{definition of } \mu^{\exists}) \\
&(s,t) : P, x : X \vdash (f(x) \in \mu^{\exists}_Y(t) \wedge x \in s) \leftrightarrow x \in s \\
&\iff \quad (P \text{ pullback, equation (6.34)}) \\
&(s,t) : P, x : X \vdash (f(x) \in \exists f(s) \wedge x \in s) \leftrightarrow x \in s \\
&\iff \quad (\text{definition of } \exists) \\
&(s,t) : P, x : X \vdash (\exists x' : X . f(x) = f(x') \wedge x' \in s \wedge x \in s) \leftrightarrow x \in s \\
&\iff \quad (\text{simplification using } x \text{ as an existential witness for } x') \\
&(s,t) : P, x : X \vdash x \in s \leftrightarrow x \in s
\end{aligned}$$

and the last judgement is a tautology. □

Proposition 6.32. $(s,t) : P \vdash \exists\exists f(k(s,t)) = t$

Proof.

$$(s,t) : P \vdash \exists\exists f(k(s,t)) = t$$
$$\iff \quad \text{(extensionality)}$$
$$(s,t) : P, v : \Omega^Y \vdash v \in \exists\exists f(k(s,t)) \leftrightarrow v \in t$$
$$\iff \quad (\exists \text{ definition})$$
$$(s,t) : P, v : \Omega^Y \vdash (\exists u : \Omega^X . u \in k(s,t) \wedge \exists f(u) = v) \leftrightarrow v \in t$$
$$\iff \quad (k \text{ definition})$$
$$(s,t) : P, v : \Omega^Y \vdash (\exists u : \Omega^X . \exists b : \Omega^Y . b \in t \wedge u = f^{-1}(b) \cap s \wedge \exists f(u) = v) \leftrightarrow v \in t$$
$$\iff \quad (\text{simplification of } \exists u : \Omega^X . u = f^{-1}(b) \cap s)$$
$$(s,t) : P, v : \Omega^Y \vdash (\exists b : \Omega^Y . b \in t \wedge \exists f(f^{-1}(b) \cap s) = v) \leftrightarrow v \in t$$
$$\iff \quad \text{(Frobenius reciprocity law [97, page 204])}$$
$$(s,t) : P, v : \Omega^Y \vdash (\exists b : \Omega^Y . b \in t \wedge b \cap \exists f(s) = v) \leftrightarrow v \in t$$
$$\iff \quad \text{(extensionality)}$$
$$(s,t) : P, v : \Omega^Y, y : Y \vdash (\exists b : \Omega^Y . b \in t \wedge (y \in b \cap \exists f(s)) \leftrightarrow y \in v) \leftrightarrow v \in t$$
$$\iff \quad (\text{definition of } \cap)$$
$$(s,t) : P, v : \Omega^Y, y : Y \vdash (\exists b : \Omega^Y . b \in t \wedge (y \in b \wedge y \in \exists f(s)) \leftrightarrow y \in v) \leftrightarrow v \in t$$
$$\iff \quad (P \text{ pullback, equation (6.34)})$$
$$(s,t) : P, v : \Omega^Y, y : Y \vdash (\exists b : \Omega^Y . b \in t \wedge (y \in b \wedge y \in \mu^{\exists}_Y(t)) \leftrightarrow y \in v) \leftrightarrow v \in t$$
$$\iff \quad (\text{definition of } \mu^{\exists})$$
$$(s,t) : P, v : \Omega^Y, y : Y \vdash (\exists b : \Omega^Y . b \in t \wedge (y \in b \wedge \exists c : \Omega^Y . y \in c \wedge c \in t) \leftrightarrow y \in v) \leftrightarrow v \in t$$
$$\iff \quad (\text{simplification using } b \text{ as an existential witness for } c)$$
$$(s,t) : P, v : \Omega^Y, y : Y \vdash (\exists b : \Omega^Y . b \in t \wedge (y \in b \leftrightarrow y \in v)) \leftrightarrow v \in t$$
$$\iff \quad \text{(extensionality)}$$
$$(s,t) : P, v : \Omega^Y \vdash (\exists b : \Omega^Y . b \in t \wedge b = v) \leftrightarrow v \in t$$
$$\iff \quad (\text{simplification of } \exists b : \Omega^Y . b = v)$$
$$(s,t) : P, v : \Omega^Y \vdash v \in t \leftrightarrow v \in t$$

and the last judgement is a tautology. $\square$

Chapter 7

Compact Hausdorff Spaces

In this chapter, we work within the regular category KHaus of compact Hausdorff spaces and continuous functions. First, we define in detail the Vietoris monad $\mathbf{V}$ and explain how it can be used to find monotone laws of type $\mathbf{TV} \to \mathbf{VT}$. Next, we apply this framework to $\mathbf{T} = \mathbf{V}$ and derive a monotone weak distributive law $\mathbf{VV} \to \mathbf{VV}$. In the last section, we define the Radon monad $\mathbf{R}$, a probability monad which can be seen as a continuous version of the distribution monad, and we provide partial results towards the existence of a monotone weak distributive law of type $\mathbf{RV} \to \mathbf{VR}$.

Notation

We fix some notation that will be particularly used in the chapter. For a function $f : X \to Y$ between sets, the direct image of a $U \subseteq X$ is denoted by $f(U)$ and the preimage of a $V \subseteq Y$ is denoted by $f^{-1}(V)$. These notions are extended to relations $R \subseteq X \times Y$: the direct image of a $U \subseteq X$ is defined by

$$R[U] = \{y \in Y \mid \exists x \in U.(x, y) \in R\} \tag{7.1}$$

and the preimage of a $V \subseteq Y$ is defined by

$$R^{-1}[V] = \{x \in X \mid \exists y \in V.(x, y) \in R\} \tag{7.2}$$

For singletons, we allow the slight notation abuses $R[x] = R[\{x\}]$ and $R^{-1}[x] = R^{-1}[\{x\}]$. The complement of $U \subseteq X$ with respect to the ambient set X is denoted by U^c.

7.1 Preliminaries

The category of compact Hausdorff spaces and continuous functions, as well as the Vietoris monad, were succinctly defined in the previous chapters. As they are the focus of the current chapter, we give a more detailed exposition of them.

A topological space consists of a set X and a topology τ_X on X. For readability, a topological space (X, τ_X) will simply be denoted by X. Recall that it is

- *compact* if every open cover of X has a finite subcover
- *Hausdorff* if any two distinct points have disjoint neighbourhoods

A function between topological spaces $f : X \to Y$ is *continuous* if for all $U \in \tau_Y$, $f^{-1}(U) \in \tau_X$. Compact Hausdorff spaces and continuous functions form a category denoted by KHaus.

Example 7.1. The prototypical example of compact Hausdorff space is the *unit interval* $[0, 1]$ with the standard topology from the reals.

Example 7.2. Any finite discrete topological space is compact Hausdorff.

Example 7.3. The Cantor set 2^ω, defined as the product of a countable number of copies of the discrete space $\{0, 1\}$, is compact Hausdorff.

We now explore some categorical properties of KHaus.

- An important fact we already mentioned is that KHaus is isomorphic to the Eilenberg-Moore category of the ultrafilter monad $\boldsymbol{\beta}$ [98]. This is due to the fact that an ultrafilter yields a notion of convergence, and that given a set X, a $\boldsymbol{\beta}$-algebra $\beta X \to X$ can be seen as stating existence and unicity of the limit for every ultrafilter. Existence corresponds to compactness, and unicity corresponds to Hausdorffness.
- A first consequence is that KHaus is complete and limits can be computed as in Set and given the initial topology afterwards [29, Theorem 4.3.5].
- A second consequence is that KHaus is regular [29, Theorem 4.3.5], hence there is a category $\mathsf{Rel}(\mathsf{KHaus})$ of compact Hausdorff spaces and relations.
- The category KHaus is not a topos, so the results of Chapter 6 cannot be applied. However, KHaus is a pretopos [102]. For our purposes, what counts is that this implies that regular epimorphisms coincide with epimorphisms, and also that KHaus is balanced i.e. being an isomorphism is equivalent to being both a monomorphism and an epimorphism.

- Furthermore, the epimorphisms are the continuous surjections, the monomorphisms are the continuous injections, so by the previous point the isomorphisms are the continuous bijections.

Topologically speaking, compact Hausdorff spaces are very well-behaved. For example, the direct image of a closed subset is a closed subset, singletons are always closed, diagonals $\{(x,x) \mid x \in X\} \subseteq X \times X$ (and more generally graphs of continuous functions) are closed in the product topology. These overall good properties notably lead to the Vietoris monad's well-definedness. The ideas underlying this monad have been introduced by Vietoris in 1922 [155]. For a detailed categorical treatment, see [159]; for examples of coalgebraic applications, see [31, 12]. We recall the definition of $\mathbf{V}$. For a compact Hausdorff space X, let VX be the set of all closed subsets of X:

$$VX = \{C \subseteq X \mid C \textit{ is closed with respect to } \tau_X\} \tag{7.3}$$

Following Wyler [159, Remark 5.8], VX contains the empty set. For any set $U \subseteq X$, let

$$\Box U = \{C \in VX \mid C \subseteq U\} \tag{7.4}$$

$$\Diamond U = \{C \in VX \mid C \cap U \neq \emptyset\} \tag{7.5}$$

The set VX is endowed with the Vietoris topology, generated by the subbase

$$\{\Box U \mid U \in \tau_X\} \cup \{\Diamond U \mid U \in \tau_X\} \tag{7.6}$$

Concretely, the open sets of VX are arbitrary unions of finite intersections of subsets $\Box U$, $\Diamond U$ with U open. As $\Box U_1 \cap \Box U_2 = \Box(U_1 \cap U_2)$ and $\Box X = VX$, a base of the Vietoris topology is given by sets of the form

$$\Box U_0 \cap \bigcap_{1 \leq i \leq n} \Diamond U_i \tag{7.7}$$

where $n \in \mathbb{N}$ and $U_0, \ldots, U_n \in \tau_X$. Another remark is that the constructors of the Vietoris topology preserve closed sets: if $C \in VX$, then $\Box C \in VVX$ and $\Diamond C \in VVX$.

The Vietoris topology of a compact Hausdorff space is compact Hausdorff again. Defining $Vf(C)$ to be the direct image $f(C)$ for every continuous $f : X \to Y$ therefore yields a functor $V : \mathsf{KHaus} \to \mathsf{KHaus}$. This functor is then extended to a monad $\mathbf{V}$ by mimicking powerset operations as $\eta^{\mathbf{V}}_X(x) = \{x\}$ and $\mu^{\mathbf{V}}_X(\mathcal{C}) = \bigcup \mathcal{C}$.

Remark 7.4. The notation $\Box$, $\Diamond$ is borrowed from modal logic, because of its close links with Vietoris topologies – see [154] for an overview.

As mentioned in Example 1.29, the category $\mathsf{EM}(\mathbf{V})$ can be identified to a category of so-called continuous lattices and functions preserving infima and directed suprema [159]. We shall shortly give a description of the Kleisli category $\mathsf{Kl}(\mathbf{V})$. To conclude this introductory section, note that the Vietoris monad coincides with the powerset monad for finite discrete spaces.

Proposition 7.5. *The Vietoris monad restricts to the full subcategory of* KHaus *whose objects are finite discrete spaces. More precisely,*

- *for any finite discrete space X, VX is the set PX with the discrete topology*
- *moreover $\eta^{\mathbf{V}}_X = \eta^{\mathbf{P}}_X$, $\mu^{\mathbf{V}}_X = \mu^{\mathbf{P}}_X$ and for any (trivially continuous) function $f : X \to Y$ between finite discrete spaces, $Vf = Pf$*

Proof. Let X be finite and discrete. Every subset of X is closed, so the underlying set of VX is PX. To show that the topology of VX is discrete, consider any singleton $\{U\}$, where $U \subseteq X$, and simply remark that it is open as a finite intersection of open sets:

$$\{U\} = \Box U \cap \bigcap_{x \in U} \Diamond\{x\}$$

This proves the first point. The second point is immediate from the definition of $\eta^{\mathbf{V}}$, $\mu^{\mathbf{V}}$ and V on morphisms. □

7.2 Relational Extensions versus Vietoris Extensions

For a moment, let us forget the general theory of relations in regular categories and try to define directly what could be a relevant notion of relation in KHaus. A reasonable demand is that every continuous function $f : X \to Y$ between compact Hausdorff spaces can be regarded as a relevant relation. For such a function, note that the following three properties hold:

- for every $C \in VX$, $f(C) \in VY$ by compactness
- for every $C \in VY$, $f^{-1}(C) \in VX$ by continuity
- for every $U \in \tau_Y$, $f^{-1}(U) \in \tau_X$ by continuity

This hints how to define properties of interest for a relation $R \subseteq X \times Y$.

Definition 7.6. Let $R \subseteq X \times Y$, where X and Y are compact Hausdorff spaces. Consider the properties

1. for every $C \in VX$, $R[C] \in VY$
2. for every $C \in VY$, $R^{-1}[C] \in VX$
3. for every $U \in \tau_Y$, $R^{-1}[U] \in \tau_X$

The relation R is *closed* if it satisfies properties (1) and (2). The relation R is *continuous* if it satisfies additionally property (3).

Note that a relation $R \subseteq X \times Y$ is closed if and only if R is closed as a subset of the product topology of $X \times Y$ [10, Lemma 3.2]. Continuous relations are a bit harder to grasp. For example, any closed subset of $X \times Y$ which is the union of the graphs of an arbitrary family of continuous functions $X \to Y$ is itself continuous. Here are some examples in the compact Hausdorff space $[0,1]$. Below are displayed respectively a continuous function, a continuous relation, and a closed relation.

(7.8)

For any continuous relation $R \subseteq X \times Y$, as Y is both closed and open in itself, $R^{-1}[Y]$ is both closed and open in X. When X is connected, this implies R is either empty or total. Consequently, the third relation pictured above is not continuous.

One can check that each of properties (1)-(3) is satisfied by the identity relation and preserved by the usual composition of relations (as defined in Example 1.22). Therefore there are categories of compact Hausdorff spaces and closed / continuous relations. These categories have been studied in [10] where they are respectively denoted by $\mathsf{KHaus}^{\mathsf{R}}$ and $\mathsf{KHaus}^{\mathsf{C}}$. The following propositions pinpoint that they both emerge from generic constructions.

Proposition 7.7. *The category of compact Hausdorff spaces and closed relations is isomorphic to* $\mathsf{Rel}(\mathsf{KHaus})$.

Proof. The isomorphism is the identity on objects. A closed relation $R \subseteq X \times Y$ is mapped to the subobject given by the inclusion monomorphism of type $R \hookrightarrow X \times Y$. Conversely, a subobject given by a monomorphism $r : R \hookrightarrow X \times Y$ is mapped to the closed relation $Vr(R) \subseteq X \times Y$, depending only on the equivalence class of r.

These transformations clearly are inverse to each other. Functoriality is immediate: pullbacks in KHaus are computed as in Set, so under the above identification both compositions coincide. $\square$

Proposition 7.8 ([10, Remark 4.18])**.** *The category of compact Hausdorff spaces and continuous relations is isomorphic to* $\mathsf{Kl}(\mathbf{V})$*.*

Proof. From continuous relations to $\mathsf{Kl}(\mathbf{V})$, let F be the identity on objects and $FR = \lambda x.R[x]$. In the other direction, let G be the identity on objects and $Gf = \{(x, y) \mid y \in f(x)\}$ for every morphism $f : X \to VY$. Provided these expressions are well-defined, they induce functors that are inverse to each other, as in the case $\mathsf{Rel} \cong \mathsf{Kl}(\mathbf{P})$ – see Example 1.22. We verify next the well-definedness of F and G.

- For any continuous $R \subseteq X \times Y$, $FR : X \to VY$ is a continuous function. Indeed, it is well-defined because of property (1), and for any $U \in \tau_Y$, $(FR)^{-1}(\Box U) = R^{-1}[U^c]^c$ and $(FR)^{-1}(\Diamond U) = R^{-1}[U]$ are open sets by properties (2) and (3), respectively.

- For any continuous $f : X \to VY$, Gf is a continuous relation. Indeed

  1. for any $C \in VX$, $(Gf)[C] = (\mu^{\mathbf{V}}_X \circ Vf)(C) \in VY$
  2. for any $C \in VY$, $(Gf)^{-1}[C] = f^{-1}(\Box(C^c))^c \in VX$ by continuity of f
  3. for any $U \in \tau_Y$, $(Gf)^{-1}[U] = f^{-1}(\Diamond U) \in \tau_X$ by continuity of f $\square$

Properties (1) and (2) of closed relations yield a domain-codomain symmetry embodied by the transpose involution $(-)^\circ : \mathsf{Rel}(\mathsf{KHaus})^{\mathsf{op}} \to \mathsf{Rel}(\mathsf{KHaus})$. Adding the sole property (3) breaks this symmetry for continuous relations. Imposing a fourth property

4. for every $U \in \tau_X$, $R[U] \in \tau_X$

would bring back a transpose involution, but also restrict the scope to *open* functions of KHaus. Relations satisfying (1)-(4) are called *interior relations* and have been also studied in [10].

To sum up, KHaus, $\mathsf{Kl}(\mathbf{V})$ and $\mathsf{Rel}(\mathsf{KHaus})$ have the same objects, but more and more morphisms, as show the inclusions

$$\begin{tikzcd} \mathsf{KHaus} \arrow[rr, "\mathcal{G}"] \arrow[dr, "F_{\mathbf{V}}"'] & & \mathsf{Rel}(\mathsf{KHaus}) \\ & \mathsf{Kl}(\mathbf{V}) \arrow[ur] & \end{tikzcd} \tag{7.9}$$

where the functor $\mathsf{Kl}(\mathbf{V}) \to \mathsf{Rel}(\mathsf{KHaus})$ simply forgets property (3). Unveiling this structure immediately induces

Proposition 7.9. *Let F, G : $\mathsf{Rel}(\mathsf{KHaus}) \to \mathsf{Rel}(\mathsf{KHaus})$ be endofunctors and α : $F \to G$ be a natural transformation.*

- *The functor F restricts to $\mathsf{Kl}(\mathbf{V})$ if and only if it preserves continuous relations.*
- *Assume F and G satisfy the previous point, then α restricts to $\mathsf{Kl}(\mathbf{V})$ if and only if its components are continuous relations.*

Combining this proposition with Theorem 5.6 yields

Theorem 7.10. *Let F, G : $\mathsf{KHaus} \to \mathsf{KHaus}$ be endofunctors and $\alpha : F \to G$ be a natural transformation.*

- *There is a locally monotone functor $\underline{F} : \mathsf{Kl}(\mathbf{V}) \to \mathsf{Kl}(\mathbf{V})$ such that $F_{\mathbf{V}}F = \underline{F}F_{\mathbf{V}}$, obtained by restriction of a relational extension of F to $\mathsf{Rel}(\mathsf{KHaus})$, if and only if F preserves continuous surjections, is nearly cartesian, and $\mathsf{Rel}(F)$ preserves continuous relations. In this case, $\underline{F}$ is unique.*
- *Assume F and G satisfy the previous point. There is a natural transformation $\underline{\alpha} : \underline{F} \to \underline{G}$ such that $F_{\mathbf{V}}\alpha = \underline{\alpha}F_{\mathbf{V}}$ if and only if α is nearly cartesian. In this case, $\underline{\alpha}$ is unique.*

Proof. The proof is made easy by using preceding results.

- By Theorem 5.6, the relational extension exists if and only if F is nearly cartesian and F preserves regular epimorphisms = epimorphisms = continuous surjections. By Proposition 7.9, this restricts to an $\underline{F}$ in $\mathsf{Kl}(\mathbf{V})$ if and only if $\mathsf{Rel}(F)$ preserves continuous relations. The functor $\underline{F}$ is locally monotone because $\mathsf{Rel}(F)$ is so, and $F_{\mathbf{V}}F = \underline{F}F_{\mathbf{V}}$ because $\mathcal{G}F = \mathsf{Rel}(F)\mathcal{G}$. Finally, such an $\underline{F}$ is unique by uniqueness of the relational extension.
- The equation $F_{\mathbf{V}}\alpha = \underline{\alpha}F_{\mathbf{V}}$ forces components of $\underline{\alpha}$ to be $\underline{\alpha}_X = \alpha_X$, so if $\underline{\alpha}$ exists, it is unique. By Theorem 5.6, this defines a natural transformation $\mathsf{Rel}(\alpha)$: $\mathsf{Rel}(F) \to \mathsf{Rel}(G)$ if and only if α is nearly cartesian. By Proposition 7.9, $\mathsf{Rel}(\alpha)$ restricts to $\mathsf{Kl}(\mathbf{V})$ if and only if its components are continuous relations – which is the case, since components of $\mathsf{Rel}(\alpha)$ are graphs of continuous functions. $\square$

Finally we can state the main result concerning (weak) distributive laws, that will be very useful in the next sections.

Corollary 7.11. *For every monad* $\mathbf{T}$ *on* KHaus*,*

- *there exists at most one (monotone) distributive law of type* $\mathbf{TV} \to \mathbf{VT}$ *such that the functor of the corresponding extension is the restriction of a relational extension of* T*. There is one if and only if* T *preserves continuous surjections, is nearly cartesian,* $\mathsf{Rel}(T)$ *preserves continuous relations and* $\eta^{\mathbf{T}}$*,* $\mu^{\mathbf{T}}$ *are nearly cartesian.*

- *there exists at most one (monotone) weak distributive law of type* $\mathbf{TV} \to \mathbf{VT}$ *such that the functor of the corresponding extension is the restriction of a relational extension of* T*. There is one if and only if* T *preserves continuous surjections, is nearly cartesian,* $\mathsf{Rel}(T)$ *preserves continuous relations and* $\mu^{\mathbf{T}}$ *is nearly cartesian.*

According to the correspondence theorems, in each case the graph of $\lambda_X : TVX \to VTX$ is $\underline{T}(\ni_X)$, where $\ni_X = \{(C, x) \in VX \times X \mid x \in C\}$.

7.3 Vietoris over Vietoris

Using the identity monad morphism $\mathbf{V} \to \mathbf{V}$, there is

- a trivial weak distributive law $\mathbf{VV} \to \mathbf{VV}$ defined by $\mathcal{C} \mapsto \{\bigcup \mathcal{C}\}$ – the non-monotone functor of the corresponding weak extension satisfies

$$\underline{V}R = \{(C, R[C]) \mid C \in VX\} \tag{7.10}$$

- a trivial coweak distributive law $\mathbf{VV} \to \mathbf{VV}$ defined by $\mathcal{C} \mapsto \{\{x\} \mid x \in \bigcup \mathcal{C}\}$

This is basically a restatement of Examples 2.14 and 2.33 in KHaus. We are rather interested in generalising the *monotone* weak distributive law $\mathbf{PP} \to \mathbf{PP}$ to KHaus.

7.3.1 The Monotone Law

In this section, we apply Corollary 7.11 with $\mathbf{T} = \mathbf{V}$. We proceed to check the required assumptions. As expected, most of them follow the same line as for sets and toposes; only preservation of continuous relations requires new arguments.

Proposition 7.12. *The Vietoris functor* V *preserves continuous surjections.*

Proof. Let $f : X \to Y$ be a continuous surjective function between compact Hausdorff spaces. Then, the continuous $Vf : VX \to VY$ is surjective: for any $C \in VY$, take $K = f^{-1}(C) \in VX$. Then $Vf(K) = f(K) = f(f^{-1}(C)) = C$ by surjectivity of f. $\square$

Proposition 7.13. *The Vietoris functor V is nearly cartesian.*

Proof. Consider a pullback in KHaus

$$\begin{tikzcd} P \arrow[r, "p_2"] \arrow[d, "p_1"'] \arrow[dr, phantom, "\lrcorner", very near start] & Y \arrow[d, "g"] \\ X \arrow[r, "f"'] & Z \end{tikzcd} \tag{7.11}$$

One must check that the universal morphism $h : VP \to Q$ depicted below is a continuous surjection.

$$\begin{tikzcd} VP \arrow[dr, "h"] \arrow[drr, bend left, "Vp_2"] \arrow[ddr, bend right, "Vp_1"'] & & \\ & Q \arrow[r, "q_1"] \arrow[d, "q_2"] \arrow[dr, phantom, "\lrcorner", very near start] & VY \arrow[d, "Vg"] \\ & VX \arrow[r, "Vf"'] & VZ \end{tikzcd} \tag{7.12}$$

Using that pullbacks in KHaus are computed as in Set, this amounts to proving that for any $C_X \in VX$, $C_Y \in VY$ such that $Vf(C_X) = Vg(C_Y)$, there exists a $C_P \in VP$ such that $Vp_1(C_P) = C_X$ and $Vp_2(C_P) = C_Y$. We take

$$C_P = p_1^{-1}(C_X) \cap p_2^{-1}(C_Y) \tag{7.13}$$

We have $C_P \in VP$ by continuity of the projections p_1, p_2. The rest of the proof is exactly as for Proposition 3.1. $\square$

Hence, the Vietoris functor has a relational extension $\mathsf{Rel}(V)$ on $\mathsf{Rel}(\mathsf{KHaus})$. For the next proposition, we need a technical lemma expressing how the basic opens of the Vietoris topology are transported by relational preimages.

Lemma 7.14. *Let $R \subseteq X \times Y$ be a continuous relation between compact Hausdorff spaces and $(U_i)_{0 \leq i \leq n}$ be open sets of Y, where $n \geq 0$. Then*

$$(\mathsf{Rel}(V)(R))^{-1}\left[\Box U_0 \cap \bigcap_{1 \leq i \leq n} \Diamond U_i\right] = \Box R^{-1}[U_0] \cap \bigcap_{1 \leq i \leq n} \Diamond R^{-1}[U_0 \cap U_i] \tag{7.14}$$

Proof. According to equation (5.7), we dispose of the direct expression $\mathsf{Rel}(V)(R) = Vr_2 \cdot Vr_1{}^{\circ}$, where $r_1 : R \to X$, $r_2 : R \to Y$ are the projections and $\cdot$ is the usual composition of relations. Explicitly, this gives an Egli-Milner-like formula

$$\begin{aligned} \mathsf{Rel}(V)(R) = \{(C, D) \in VX \times VY \mid & \forall x \in C. \exists y \in D. (x, y) \in R \\ & \text{and } \forall y \in D. \exists x \in C. (x, y) \in R\} \end{aligned}$$

Then, a subset $C \in VX$ belongs to the left-hand side of (7.14) if and only if there exists $D \in VY$ such that

- $\forall x \in C. \exists y \in D. (x, y) \in R$
- $\forall y \in D. \exists x \in C. (x, y) \in R$
- $D \subseteq U_0$
- $\forall i \in \{1, \ldots, n\}. D \cap U_i \neq \emptyset$

A subset $C \in VX$ belongs to the right-hand side if and only if

- ∘ $\forall x \in C. \exists y \in U_0. (x, y) \in R$
- ∘ $\forall i \in \{1, \ldots, n\}. \exists (x_i, y_i) \in R \cap (C \times (U_0 \cap U_i))$

($\subseteq$) Let $C \in VX$ belonging to the left-hand side and $D \in VY$ satisfying the four • properties.

- ∘ Let $x \in C$. Then we can find $y \in D \subseteq U_0$ such that $(x, y) \in R$.
- ∘ Let $i \in \{1, \ldots, n\}$. As $D \cap U_i \neq \emptyset$, we can find $y_i \in D \cap U_i \subseteq U_0 \cap U_i$. By another assumption there is $x_i \in C$ such that $(x_i, y_i) \in R$. Hence $(x_i, y_i) \in R \cap (C \times (U_0 \cap U_i))$.

($\supseteq$) For the other inclusion, first note that every compact Hausdorff space is a regular space in the sense of [158, §14]. According to [158, Theorem 14.3], regular spaces satisfy the following property:

For every $y \in U$ open, there is an open W such that $y \in W$ and $\overline{W} \subseteq U$. (reg)

Let $C \in VX$ such that the two assumptions ∘ hold. For every $x \in C$ we fix $y_x \in U_0$ such that $(x, y_x) \in R$. For every $i \in \{1, \ldots, n\}$ we fix $(x_i, y_i) \in R$ such that $x_i \in C$ and $y_i \in U_0 \cap U_i$. By appling the property (reg) in Y, we get for every $x \in C$ a $W_x \in \tau_Y$ such that $y_x \in W_x$ and $\overline{W_x} \subseteq U_0$, and for every $i \in \{1, \ldots, n\}$ a $W_i \in \tau_Y$ such that $y_i \in W_i$ and $\overline{W_i} \subseteq U_0 \cap U_i$. Note that for every $x \in C$, the pair $(x, y_x) \in R$ witnesses that $x \in R^{-1}[W_x]$. Therefore

$$C \subseteq \bigcup_{x \in C} R^{-1}[W_x] \tag{7.15}$$

By continuity of R, this is an open cover of the compact C, so we can extract a finite subcover

$$C \subseteq \bigcup_{1 \leq k \leq m} R^{-1}[W_{x_k}] \tag{7.16}$$

Now we define

$$K = \bigcup_{1 \le i \le n} \overline{W_i} \cup \bigcup_{1 \le k \le m} \overline{W_{x_k}} \tag{7.17}$$

As a finite union of closed subsets, $K \in VY$. Then $((C \times K) \cap R)$ is a closed subset of R so that we can finally define

$$D = Vr_2((C \times K) \cap R) \in VY \tag{7.18}$$

It remains to check that D satisfies the four $\bullet$ properties.

- Let $x \in C$. By equation (7.16), there is a $k \in \{1, \dots, m\}$ and $y \in W_{x_k} \subseteq K$ such that $(x, y) \in R$, hence also $y \in D$.
- Let $y \in D$. By definition of D, there is $x \in C$ such that $(x, y) \in R$.
- The $\overline{W_x}$ and $\overline{W_i}$ are all included in U_0. Then, $D \subseteq K \subseteq U_0$.
- Let $i \in \{1, \dots, n\}$. The pair $(x_i, y_i) \in (C \times K) \cap R$ witnesses that $y_i \in D$, and by definition $y_i \in U_i$, whence $D \cap U_i \neq \emptyset$. $\square$

Proposition 7.15. *The relational extension* $\mathsf{Rel}(V)$ *of the Vietoris functor* V *preserves continuous relations.*

Proof. Let $R \subseteq X \times Y$ be a continuous relation and let us show that $\mathsf{Rel}(V)(R)$ is continuous, i.e., satisfies property (3):

$$\forall U \in \tau_{VY}.(\mathsf{Rel}(V)(R))^{-1}[U] \in \tau_{VX} \tag{7.19}$$

According to Lemma 7.14, the above property is true if U of the form

$$\Box U_0 \cap \bigcap_{1 \le i \le n} \Diamond U_i \tag{7.20}$$

where $n \geq 0$ and $(U_i)_{0 \le i \le n}$ are open subsets of Y. Subsets of that form constitute a basis of τ_{VY}. Because relational preimages commute with arbitrary unions, property (3) then holds for any $U \in \tau_{VY}$. Note that the proof cannot be simplified further by looking directly at subbasic opens $\Box U$ and $\Diamond U$, because relational preimages do not commute with intersections. $\square$

Hence $\mathsf{Rel}(V)$ restricts to a locally monotone extension $\underline{V} : \mathsf{Kl}(\mathbf{V}) \to \mathsf{Kl}(\mathbf{V})$. Considering the monad structure, we can first observe that, as expected, the unit is not well-behaved.

Proposition 7.16. *The Vietoris unit η^{V} is not nearly cartesian.*

Proof. Endow the finite sets $\{0\}$ and $\{0,1\}$ with the discrete topology. Then the naturality square

$$\begin{tikzcd}
\{0,1\} \arrow[r, "\eta^{\mathsf{V}}_{\{0,1\}}"] \arrow[d, "!"'] & V\{0,1\} \arrow[d, "V!"] \\
\{0\} \arrow[r, "\eta^{\mathsf{V}}_{\{0\}}"'] & V\{0\}
\end{tikzcd} \tag{7.21}$$

is a counterexample to nearly cartesianness, with the same argument as in Proposition 3.2. $\square$

Hence no distributive law of type $\mathsf{VV} \to \mathsf{VV}$ is obtainable from Corollary 7.11. More generally,

Proposition 7.17. *There is no distributive law of type $\mathsf{VV} \to \mathsf{VV}$.*

Proof. As in the last proposition, the trick consists in observing that Klin and Salamanca's proof of non-existence of a distributive law $\mathsf{PP} \to \mathsf{PP}$ [91, Theorem 2.4] only makes use of finite sets. According to Proposition 7.5, V coincides with P on finite discrete spaces, so their proof adapts to KHaus. $\square$

Proposition 7.18. *The Vietoris multiplication μ^{V} is nearly cartesian.*

Proof. Let $f : X \to Y$ be a continuous function. We must show that in the pullback of its μ^{V} naturality square, the continuous function h is surjective:

$$\begin{tikzcd}
VVX \arrow[dr, "h"] \arrow[drr, bend left, "\mu^{\mathsf{V}}_X"] \arrow[ddr, bend right, "VVf"'] & & \\
& P \arrow[r, "p_2"'] \arrow[d, "p_1"] \arrow[dr, phantom, "\lrcorner", very near start] & VX \arrow[d, "Vf"] \\
& VVY \arrow[r, "\mu^{\mathsf{V}}_Y"'] & VY
\end{tikzcd} \tag{7.22}$$

Given $C \in VX$ and $\mathcal{D} \in VVY$ be such that $Vf(C) = \mu^{\mathsf{V}}_Y(\mathcal{D})$, let us find a $\mathcal{C} \in VVX$ such that $\mu^{\mathsf{V}}_X(\mathcal{C}) = C$ and $VVf(\mathcal{C}) = \mathcal{D}$. Our usual candidate

$$\mathcal{U} = \{C \cap f^{-1}(D) \mid D \in \mathcal{D}\} \tag{7.23}$$

may not be a closed subset of VX. Instead, we suggest as a new candidate

$$\mathcal{C} = \{K \in VX \mid Vf(K) \in \mathcal{D} \text{ and } K \subseteq C\} = (Vf)^{-1}(\mathcal{D}) \cap \Box C \in VVX \tag{7.24}$$

Note that $\mathcal{U} \subseteq \mathcal{C}$. Reusing the proof of Proposition 3.3, this yields $C = \mu^{\mathbf{P}}_X(\mathcal{U}) \subseteq \mu^{\mathbf{V}}_X(\mathcal{C})$ and $\mathcal{D} = PPf(\mathcal{U}) \subseteq VVf(\mathcal{C})$. The converse inclusions are immediate by the definition of $\mathcal{C}$. $\square$

Applying Corollary 7.11, we finally obtain

Theorem 7.19. *There is a monotone weak distributive law* $\lambda : \mathbf{VV} \to \mathbf{VV}$ *given by*

$$\lambda_X(\mathcal{C}) = \Big\{K \in VX \mid K \subseteq \bigcup \mathcal{C} \text{ and } \forall C \in \mathcal{C}. K \cap C \neq \emptyset\Big\} \tag{7.25}$$

Exploiting the Vietoris modalities $\Box$ and $\Diamond$, there is a more symmetrical expression

$$\lambda_X(\mathcal{C}) = \left(\Box \bigcup_{C \in \mathcal{C}} C\right) \cap \left(\bigcap_{C \in \mathcal{C}} \Diamond C\right) \tag{7.26}$$

which is reminiscent from modal logic's nabla modality [6, 110], defined on sets of formulas Φ by

$$\nabla \Phi = \Box \bigvee \Phi \wedge \bigwedge \Diamond \Phi \tag{7.27}$$

Note also that, as stated e.g. in [154], the Vietoris topology on VX is equivalently generated by all sets of the form

$$\nabla\{U_1, \ldots, U_n\} = \left\{K \in VX \mid K \subseteq \bigcup_{i=1}^{n} U_i \text{ and } \forall i \in \{1, \ldots, n\}. K \cap U_i \neq \emptyset\right\} \tag{7.28}$$

where the U_i range over open sets of X. As a lovely conclusion, the monotone weak distributive law $\mathbf{VV} \to \mathbf{VV}$ that we just described is literally encoding the Vietoris topology.

7.3.2 Variations

We mention a slight variation of the monotone law $\lambda : \mathbf{VV} \to \mathbf{VV}$. A topological space is *totally disconnected* if its only connected components are the singletons. A compact, Hausdorff, totally disconnected space is called a *Stone space*. For example, the Cantor set 2^ω and finite discrete spaces are Stone, but $[0, 1]$ is not. The category of Stone spaces and continuous functions is denoted by Stone – it is a full subcategory of KHaus. The category Stone is widely used in logic because of the celebrated Stone duality, which identifies Stone spaces with Boolean algebras [82], and its many extensions. Notably, the Vietoris monad is known to restrict to a monad on Stone that we still denote by $\mathbf{V}$. There, the topology of VX is generated by all the $\Box U$, $\Diamond U$ where U only ranges over the *clopen* sets of X i.e. sets that are simultaneously closed and open. It is clear that the λ of Theorem 7.19 still defines a weak distributive law of type $\mathbf{VV} \to \mathbf{VV}$ in Stone.

A first remark is that the Vietoris functor on Stone is not weakly cartesian [14], but is known to be nearly cartesian [143], which is coherent with our results.

In Stone, the category $\mathsf{Coalg}(V)$ has been identified as the category of *descriptive Kripke frames*, dual to the category of *Boolean algebras with operators* (Jónsson-Tarski duality [85]). These are called respectively *descriptive general frames* and *modal algebras* in [92]. On the other hand, $\mathsf{Coalg}(VV)$ is the category of so-called *descriptive INL-frames*, where INL stands for *instantial neighbourhood logic*, dual to the category of *Boolean algebras with instantial operators* [13]. Using our results of Chapter 4, it is possible to perform a generalised determinisation of descriptive INL-frames into descriptive Kripke frames. As for alternating automata in Set, there are at least three ways to do this, using respectively the trivial weak law, the trivial coweak law, and the monotone weak law $\mathbf{VV} \to \mathbf{VV}$. It would be interesting to study this procedure from the dual viewpoint: what means generalised determinisation with respect to Boolean algebra with (instantial) operators?

7.4 Radon over Vietoris

Our original motivation for working in the category KHaus was to generalise the monotone weak distributive law $\mathbf{DP} \to \mathbf{PD}$ to a continuous framework. The regular category of compact Hausdorff spaces seems adequate in this regard, because it possesses both

- a powerset-like monad, the Vietoris monad $\mathbf{V}$
- a distribution-like monad, the Radon monad $\mathbf{R}$

This section introduces the Radon monad and takes some first steps towards the existence of a monotone weak distributive law of type $\mathbf{RV} \to \mathbf{VR}$.

First let us give a few reminders on measure theory. This exposition follows the lines of a preceding paper of ours [64] – we refer the reader to standard textbooks [46, 18] for a more detailed account.

Measurable spaces. A *σ-algebra* on a set X is a subset of subsets $\Sigma_X \in PPX$ that contains X and is closed under complement and countable unions. A pair (X, Σ_X) is called a *measurable space*. For any $\mathcal{U} \in PPX$, there is a smallest σ-algebra containing $\mathcal{U}$, called the σ-algebra *generated by* $\mathcal{U}$, denoted by $\sigma(\mathcal{U})$. This σ-algebra $\sigma(\mathcal{U})$ is simply the intersection of all σ-algebras containing $\mathcal{U}$. Every topological space (X, τ_X) generates a measurable space $(X, \sigma(\tau_X))$. In this case, the σ-algebra generated by the open sets of X is called the *Borel σ-algebra*, and its elements are the *Borel sets*. In what follows, $\mathbb{R}$ is endowed with its usual topology and the corresponding Borel σ-algebra.

Measurable functions. Given measurable spaces (X, Σ_X) and (Y, Σ_Y), a *measurable function* $f : (X, \Sigma_X) \to (Y, \Sigma_Y)$ is a function $f : X \to Y$ such that for every $B \in \Sigma_Y$, $f^{-1}(B) \in \Sigma_X$. Notably, if Σ_X and Σ_Y are Borel σ-algebras, any continuous function $X \to Y$ is measurable.

Measures. Given a measurable space (X, Σ_X), a *measure* on it is a non-negative function $m : \Sigma \to \mathbb{R} \cup \{\infty\}$ such that $m(\emptyset) = 0$ and for every countable sequence $(B_n)_{n\in\mathbb{N}}$ of pairwise disjoint sets in Σ, $m\left(\bigcup_{n\in\mathbb{N}} B_n\right) = \sum_{n\in\mathbb{N}} m(B_n)$. This entails that m is monotone with respect to inclusion. A measure on a Borel σ-algebra is called a *Borel measure*. A measure m such that $m(X) = 1$ is called a *probability measure*.

Integration. Let (X, Σ_X) be a measurable space and m a measure on it. Let us define, whenever it exists, the integral of a measurable function $f : (X, \Sigma_X) \to \mathbb{R}$, denoted by $\int_X f \mathrm{d}m$.

> *Step 1.* If $f(X) = \{\alpha_1, \ldots, \alpha_n\}$ for some non-negative $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$, then f is called a *simple* function and its integral is $\int_X f \mathrm{d}m = \sum_{i=1}^n \alpha_i m(f^{-1}(\{\alpha_i\}))$.
>
> *Step 2.* If f is non-negative, let $\int_X f \mathrm{d}m = \sup\left\{\int_X g \mathrm{d}m \mid g \le f, g \text{ simple}\right\} \in [0, \infty]$.
>
> *Step 3.* For arbitrary measurable f, decompose $f = f^+ - f^-$ where $f^+ = \sup(f, 0)$ and $f^- = \sup(-f, 0)$ are non-negative. If their integral are not both ∞, define $\int_X f \mathrm{d}m = \int_X f^+ \mathrm{d}m - \int_X f^- \mathrm{d}m$ and say that f is *m-integrable* if this quantity is finite.

Given a measurable function $f : (X, \Sigma_X) \to (Y, \Sigma_Y)$ and a measure m on (X, Σ_X), the *pushforward measure* of m by f is $m \circ f^{-1}$, that is, $(m \circ f^{-1})(B) = m(f^{-1}(B))$ for every $B \in \Sigma_Y$. A standard result is that a measurable $u : (Y, \Sigma_Y) \to \mathbb{R}$ is $(m \circ f^{-1})$-integrable if and only if $u \circ f$ is m-integrable and in this case, $\int_Y u \mathrm{d}(m \circ f^{-1}) = \int_X (u \circ f) \mathrm{d}m$. We now focus on the very specific case of Borel measures on compact Hausdorff spaces.

Definition 7.20. A *Radon probability measure* on a compact Hausdorff space X is a Borel probability measure m such that for every Borel set B,

$$m(B) = \sup\{m(C) \mid C \in VX \text{ and } C \subseteq B\} \tag{7.29}$$

See [18, Section 7.1] for a general definition of Radon measures, which can be seen to coincide with the above notion for compact Hausdorff spaces. Radon probability measures are equivalent to a certain class of linear functionals, see [18, Corollary 7.10.5].

Theorem 7.21 (Riesz-Markov)**.** *Let X be a compact Hausdorff space. Let $C(X)$ be the Banach space of all continuous functions $X \to \mathbb{R}$ with uniform norm. For every non-negative continuous linear functional $\varphi : C(X) \to \mathbb{R}$ of norm 1, there is a unique Radon probability measure m on X such that*

$$\forall u \in C(X). \varphi(u) = \int_X u \mathrm{d}m \tag{7.30}$$

In the theorem, φ non-negative means $u \geq 0 \Rightarrow \varphi(u) \geq 0$ for all $u \in C(X)$. The reciprocal bijection is given by $m \mapsto (u \mapsto \int_X u\mathrm{d}m)$. The canonical identification obtained from the Riesz-Markov theorem entails a logical choice for topologising the set of Radon probability measures. The set RX of all Radon probability measures on a compact Hausdorff space X is endowed with the *vague topology*, that is, the initial topology with respect to the evaluation functions

$$\begin{aligned} \mathsf{ev}_u : & RX \to \mathbb{R} \\ & \varphi \mapsto \varphi(u) \end{aligned}$$

where u describes $C(X)$. The resulting topological space RX is compact Hausdorff again. The next definition shows that this construction extends to a monad.

Definition 7.22 (Radon monad [147, 87])**.** The *Radon monad* $\mathbf{R}$ on KHaus is defined as follows for compact Hausdorff spaces X, Y and continuous functions $f : X \to Y$.

- RX is the space of Radon probability measures on X with the vague topology
- $Rf : RX \to RY$ computes the pushforward measure $Rf(m) = m \circ f^{-1}$
- $\eta^{\mathbf{R}}_X : X \to RX$ computes the Dirac measure on a point, defined for any Borel set B of X by $\eta^{\mathbf{R}}_X(x)(B) = 1$ if $x \in B$, 0 if $x \notin B$

- $\mu^{\mathsf{R}}_X : RRX \to RX$ computes the mixture measure $\mu^{\mathsf{R}}_X(M)(B) = \int_{RX} \mathsf{ev}_{\chi_B} \mathrm{d}M$, where $\chi_B : X \to \mathbb{R}$ is the characteristic function of B so that $\mathsf{ev}_{\chi_B} : RX \to \mathbb{R}$ is just $\mathsf{ev}_{\chi_B}(m) = m(B)$ for every Borel set B of X

The Radon monad is one of the many possible generalisations of the Set distribution monad. However, contrary to the Vietoris monad, the Radon monad does not restrict to the full subcategory of discrete finite spaces. For example, $R\{0,1\}$ is infinite, hence also not discrete (otherwise compactness would fail). The following proposition makes precise the sense in which R generalises D.

Proposition 7.23. *Forgetting topologies, the action of the Radon monad on finite sets can be identified with that of the distribution monad. More precisely,*

- *for every finite X with the discrete topology, the function*

$$\begin{aligned} \kappa_X &: RX \to DX \\ m &\mapsto \sum_{x \in X} m(\{x\}) \cdot x \end{aligned}$$

  is a bijection with inverse

$$\begin{aligned} \kappa_X^{-1} &: DX \to RX \\ \varphi &\mapsto \lambda B \in PX. \sum_{x \in B} \varphi(x) \end{aligned}$$

- *moreover $\eta^{\mathsf{R}}_X = \kappa_X^{-1} \circ \eta^{\mathsf{D}}_X$ and for every $f : X \to Y$ between finite discrete spaces, $Rf = \kappa_Y^{-1} \circ Df \circ \kappa_X$ (where both equations hold in Set)*

Proof. The topology of X is PX, so the Borel σ-algebra of X is also PX.

- The proof that κ_X and κ_X^{-1} are well-defined and inverse to each other is straightforward and left to the reader.

- The first equation is immediate. For any trivially continuous $f : X \to Y$ between finite discrete spaces and every $m \in RX$,

$$\begin{aligned} (\kappa_Y^{-1} \circ Df \circ \kappa_X)(m) &= (\kappa_Y^{-1} \circ Df) \left(\sum_{x \in X} m(\{x\}) \cdot x \right) \\ &= \kappa_Y^{-1} \left(\sum_{x \in X} m(\{x\}) \cdot f(x) \right) \\ &= \lambda B. \sum_{y \in B} \sum_{x \in f^{-1}(\{y\})} m(\{x\}) \\ &= \lambda B. \sum_{y \in B} (m \circ f^{-1})(\{y\}) \\ &= \lambda B. (m \circ f^{-1})(B) = Rf \end{aligned}$$

□

There is a well-behaved notion of support for Radon measures.

Definition 7.24. The *support* of a Radon probability measure $m \in RX$ is defined by

$$\mathsf{supp}(m) = \bigcap\{C \in VX \mid m(C) = 1\} \tag{7.31}$$

Notably, from [53, Proposition 5.2], we have $m(\mathsf{supp}(m)) = 1$, which easily entails that for every Borel subset B, $m(B \cap \mathsf{supp}(m)) = m(B)$.

Remark 7.25. Actually, the monad morphism $\mathsf{supp} : \mathbf{D} \to \mathbf{P}$ in Set generalises to the topological setting. More precisely, in [53], Fritz *et al.* consider the category Top of topological spaces and continuous functions and prove that the support defines a monad morphism from the so-called *valuation monad* to the so-called *hyperspace monad*. Restricting their result to the full subcategory KHaus reveals that $\mathsf{supp} : \mathbf{R} \to \mathbf{V}$ is close to be a monad morphism – actually, all required diagrams commute, but the functions $\mathsf{supp}_X : RX \to VX$ are not in general continuous [114]. What prevents continuity is that Fritz *et al.* use the (coarser) lower Vietoris topology instead of our Vietoris topology. Consequently, supp is not a monad morphism $\mathbf{R} \to \mathbf{V}$. In particular, we cannot obtain a trivial weak distributive law $\mathbf{VR} \to \mathbf{RV}$ nor a trivial coweak distributive law $\mathbf{RV} \to \mathbf{VR}$ using supp.

The path towards a monotone law $\mathbf{RV} \to \mathbf{VR}$ now requires to verify the four assumptions of Corollary 7.11. We are able to prove the first two of them.

Proposition 7.26. *The Radon functor R preserves continuous surjections.*

Proof. Let $f : X \to Y$ be a continuous surjection between compact Hausdorff spaces. Let $n \in RY$, we must find an $m \in RX$ such that $Rf(m) = n$, that is, $m \circ f^{-1} = n$. The existence of such an m is immediate by [18, Theorem 9.1.9]. □

To prove that R is nearly cartesian, we will use the following result of Edwards [47, Proposition 3.3].

Proposition 7.27 (Edwards)**.** *Let X, Y be compact Hausdorff spaces, $m \in RX$, $n \in RY$ and $K \subseteq X \times Y$ be a closed non-empty relation with projections $k_1 : K \to X$, $k_2 : K \to Y$. Then the following statements are equivalent:*

1. *there is a $\theta \in R(X \times Y)$ such that $\mathsf{supp}(\theta) \subseteq K$ and $Rk_1(\theta) = m$ and $Rk_2(\theta) = n$*

2. *for all open $U \subseteq X$, $V \subseteq Y$ such that $(U \times V^c) \cap K = \emptyset$, we have $m(U) \leq n(V)$*

Proposition 7.28. *The Radon functor R is nearly cartesian.*

Proof. Let $f : X \to Z$ and $g : Y \to Z$ be continuous functions between compact Hausdorff spaces. Consider their pullback $P = \{(x, y) \in X \times Y \mid f(x) = g(y)\}$ endowed with the initial topology with respect to the projections $p_1 : P \to X$ and $p_2 : P \to Y$. Let $m \in RX$ and $n \in RY$ be such that $Rf(m) = Rg(n)$. To prove nearly cartesianness, we must show that there is a $\lambda \in RP$ such that $Rp_1(\lambda) = m$ and $Rp_2(\lambda) = n$. First, we show that the case $P = \emptyset$ is impossible. If P is empty, then $f(X) \cap g(Y) = \emptyset$, so

$$g^{-1}(f(X)) = g^{-1}(f(X)) \cap g^{-1}(g(Y)) = g^{-1}(f(X) \cap g(Y)) = g^{-1}(\emptyset) = \emptyset \tag{7.32}$$

Henceforth $1 = m(X) = (m \circ f^{-1})(f(X)) = (n \circ g^{-1})(f(X)) = n(\emptyset) = 0$, a contradiction. Remark that P is a closed subset of $X \times Y$. Indeed, the diagonal $\Delta_Z \subseteq Z \times Z$ is closed because Z is Hausdorff, and $P = (f \times g)^{-1}(\Delta_Z)$. We now apply Proposition 7.27. The second condition of this proposition is actually equivalent to

2bis. for all $C \in VX$, $D \in VY$ such that $(C^c \times D) \cap K = \emptyset$, we have $n(D) \leq m(C)$

Let $C \in VX$ and $D \in VY$ such that $(C^c \times D) \cap P = \emptyset$. Let us show that $n(D) \leq m(C)$. Note that $f^{-1}(g(D)) \subseteq C$ because if $f(x) \in g(D)$, there is $y \in D$ such that $f(x) = g(y)$, hence $(x, y) \in P$ and using $(C^c \times D) \cap P = \emptyset$ this forces $x \in C$. Then

$$n(D) \leq n(g^{-1}(g(D))) = (n \circ g^{-1})(g(D)) = (m \circ f^{-1})(g(D)) \leq m(C) \tag{7.33}$$

(NB: it is important to use closed sets instead of open sets, because the direct image of an open set is not necessarily measurable, whereas in KHaus the direct image $g(D)$ of a closed set is closed, hence measurable.) We conclude that *2bis.* holds, so there is a Radon probability measure $\theta \in R(X \times Y)$ with $\mathsf{supp}(\theta) \subseteq P$ such that $Rp_1(\theta) = m$ and $Rp_2(\theta) = n$. The rest of the proof is mere paperwork to obtain a measure on P. Let λ be the restriction of θ to Borel subsets of P. We have $\lambda \in RP$ because $\lambda(P) = \theta(P) \geq \theta(\mathsf{supp}(\theta)) = 1$. Furthermore for every Borel subset B of Z:

$$\begin{aligned}
\lambda(p_1^{-1}(B)) &= \lambda(r_1^{-1}(B) \cap P) \\
&= \theta(r_1^{-1}(B) \cap P) \\
&= \theta(r_1^{-1}(B)) \text{ because } \mathsf{supp}(\theta) \subseteq P \\
&= m(B)
\end{aligned}$$

and similarly $\lambda \circ p_2^{-1} = n$. $\square$

Hence there is a relational extension $\mathsf{Rel}(R) : \mathsf{Rel}(\mathsf{KHaus}) \to \mathsf{Rel}(\mathsf{KHaus})$, expressed as follows – using Proposition 7.27 – for every closed $S \subseteq X \times Y$:

$$\mathsf{Rel}(R)(S) = \{(m, n) \in RX \times RY \mid \forall (C, D) \in VX \times VY. \\ \text{if } (C^c \times D) \cap S = \emptyset \text{ then } n(D) \leq m(C).\}$$

Proposition 7.29. *The Radon unit $\eta^{\mathbf{R}}$ is not nearly cartesian.*

Proof. Endow the finite sets $\{0\}$ and $\{0, 1\}$ with the discrete topology. By Proposition 7.23, identify R with D in the naturality square of the unique continuous function $\{0, 1\} \to \{0\}$. This yields a counterexample to nearly cartesianness, with the same argument as in Proposition 3.9. □

Again, no distributive law $\mathbf{RV} \to \mathbf{VR}$ can be obtained from Corollary 7.11, and more generally

Proposition 7.30. *There is no distributive law of type $\mathbf{RV} \to \mathbf{VR}$.*

Proof. The same argument as in Proposition 7.17 will work, with a bit of extra caution because $\mathbf{R}$ does not strictly restrict to finite discrete sets. The standard proof of nonexistence of a distributive law $\lambda : \mathbf{DP} \to \mathbf{PD}$ [153, Appendix] proceeds as follows. One starts with two finite sets X and Y, one specific element of DPX, and three functions $X \to Y$. Using solely unit axioms and naturality squares of λ on these sets and functions, a contradiction is derived. What counts is that *the functor D is never applied more than once.* Therefore, this proof can still be carried out by endowing every finite set with the discrete topology, replacing $\mathbf{D}$ with $\mathbf{R}$ by Proposition 7.23, and replacing $\mathbf{P}$ with $\mathbf{V}$ by Proposition 7.5. □

We do not know if the remaining assumptions of Corollary 7.11 hold. In their absence, we can only conjecture:

Conjecture 7.31. *The following properties hold:*

- *the relational extension* $\mathsf{Rel}(R)$ *preserves continuous relations*
- *the multiplication $\mu^{\mathbf{R}}$ is nearly cartesian*

The result from Edwards might be crucial to prove or disprove the first assumption. The support not being a monad morphism $\mathbf{R} \to \mathbf{V}$ possibly hints that the Radon and Vietoris topologies in KHaus suffer some intrinsic incompatibility, so that we believe $\mathsf{Rel}(R)$ most likely does not preserve continuous relations. Concerning the second assumption, we note that it seems to be related to the notion of *disintegration of measure*, and that a similar result has been proved in [113, Theorem 2.6.9]. In any case, let us be optimistic and compute what would be the resulting weak distributive law.

Proposition 7.32. *If Conjecture 7.31 holds, then the resulting monotone weak distributive law* $\lambda : \mathsf{RV} \to \mathsf{VR}$ *has the following expressions for* $m \in RVX$*,*

$$\lambda_X(m) = \{n \in RX \mid \forall(\mathcal{C}, D) \in VVX \times VX. \Diamond D \subseteq \mathcal{C} \Rightarrow n(D) \leq m(\mathcal{C})\} \tag{7.34}$$
$$= \left\{n \in RX \mid \forall(\mathcal{C}, D) \in VVX \times VX. \bigcup \mathcal{C} \subseteq D \Rightarrow m(\mathcal{C}) \leq n(D)\right\} \tag{7.35}$$

Proof. The first expression is obtained by computing $\mathsf{Rel}(R)$ on the relation $\ni_X = \{(C, x) \in VX \times X \mid x \in C\}$. The property $(\mathcal{C}^c \times D) \cap \ni_X = \emptyset$ is easily seen to be equivalent to $\Diamond D \subseteq \mathcal{C}$. For the second expression, remark that in Edward's result (Proposition 7.27), the assumptions and statement *1.* are symmetrical in X and Y. Then we can recast statements *2.* and *2bis.* by reversing the roles of X and Y, so that for every closed $S \subseteq X \times Y$,

$$\mathsf{Rel}(R)(S) = \{(m, n) \mid \forall(C, D).\text{ if } (C \times D^c) \cap S = \emptyset \text{ then } m(C) \leq n(D)\} \tag{7.36}$$

Applying this again to $\ni_X$, and remarking that $(\mathcal{C} \times D^c) \cap \ni_X = \emptyset$ is equivalent to $\bigcup \mathcal{C} \subseteq D$, we obtain the second expression. $\square$

Conclusion

In this thesis, we used weak distributive laws to rediscover various constructions of theoretical computer science. By lack of proper distributive laws, these constructions were previously performed outside of any general framework, or within a dedicated *ad hoc* framework. Rediscovered constructions include

- the convex powerset monad, exhibited as a weak lifting
- the algebraic presentation for the monad of convex subsets of distributions
- generalised determinisation procedures for alternating automata and probabilistic automata
- compatibility of bisimulation up-to convex hull for probabilistic automata

We generalised generic results stemming from Beck's theory of distributive laws to obtain generic results for weak distributive laws, from which the above constructions are just instances. An important part of the theory was already set up by Richard Garner – we provided a more detailed account, including

- a formal definition of the dual of a weak distributive law, and a detailed proof of their correspondence theorem
- a way to produce a (co)weak distributive law out of any monad morphism, generalising a result from Maaike Zwart about idempotent monads
- an account of iterated weak distributive laws, adapting results from Eugenia Cheng

Non-trivial weak distributive laws can arguably be seen as a refinement of plain distributive laws in the way they intermingle the features of the two monads. Indeed, the scope of the weak composite monad is restricted to those items for which it really makes sense to talk about both algebraic structures at once. For example, combining powerset and distribution requires a restriction to convex subsets.

Our methods are categorical rather than algebraic. Consequently, there is no restriction to work in the category of sets, contrary to the recent works of Zwart and Parlant. In the second part of the thesis, we took advantage of this opportunity to discover weak distributive laws in other categories, following the trail of laws that are monotone with respect to the Kleisli category of some powerset-like monad. Injecting the spirit of weak distributive laws into various results of the literature, we were able to

- identify toposes as the most general categories where these monotone weak distributive laws coincide with relational extensions
- prove that the powerset monad of toposes weakly self-distributes, extending results from Oege de Moor, and derive these proofs formally in the constructive logic of Coq
- find sufficient conditions for the existence of a monotone weak distributive law in the category of compact Hausdorff spaces
- prove that these conditions are met for the Vietoris monad itself, and provide partial results for the Radon monad. We end up two assumptions away from combination of probabilistic choice and non-deterministic choice in compact Hausdorff spaces.

The last point is an ideal transition to comment on future work and open questions.

- Does Conjecture 7.31 hold, or equivalently, does the monotone weak distributive law of the distribution monad over the powerset monad generalise to the continuous framework of compact Hausdorff spaces?
- For a category with a monad $\mathbf{T}$ such that global elements of TX correspond to subobjects of X, can we obtain a weak distributive law $\mathbf{TT} \to \mathbf{TT}$? Such a result would unify our laws $\mathbf{\exists\exists} \to \mathbf{\exists\exists}$ in toposes and $\mathbf{VV} \to \mathbf{VV}$ in compact Hausdorff spaces. The idea is due to an anonymous reviewer, whom we thank.
- Can we find more weak distributive laws with a relevant semantic content, as well as non-trivial coweak distributive laws? Can we find such laws that do not come from relational extensions? Can other notions of weak distributive laws from Street, Böhm and Lack lead to (re)discover some deep constructions of theoretical computer science? The search space for these interrogations obviously includes the category of sets, toposes, and compact Hausdorff spaces,

but also categories that have been mentioned throughout this thesis but not scanned in detail, such as the quasitopos of quasi-Borel spaces [74], or various categories of metric spaces (see e.g. [107]).

- We are interested in understanding more thoroughly the meaning of weak distributive laws at the algebraic level, because the algebraic approach leads to elegant proofs and a deeper understanding of phenomena in Set.

- According to a recent result of Petrişan and Sarkis [116], under mild conditions, weak distributive laws between a pair of monads can be seen as plain distributive laws between a pair of related monads. This sets the stage for no-go theorems for weak distributive laws using the algebraic methods of [161].

- The above points are especially of interest when applied to combination of monads for which there is no possible distributive law and no known weak distributive law. A curious case related to this situation is $\mathbf{PD} \to \mathbf{DP}$, for which there is a trivial weak distributive law, but possibly no other weak distributive law: would algebraic methods be able to detect that there is no meaningful law?

Appendix A

Proof of the Coweak Correspondence Theorem

In this Appendix we prove in detail the non-trivial correspondence of Theorem 2.29:

$$\text{coweak distributive laws} \iff \text{coweak extensions}$$

Coweak Distributive Laws ⇒ Coweak Extensions

For this implication we assume that idempotents split in $\mathsf{Kl}(\mathbf{S})$. Let λ be a coweak distributive law. For any $h : X \nrightarrow Y$ we define

$$h^+ = \lambda_Y \circ Th : TX \nrightarrow TY \tag{A.1}$$

Lemma A.1. *For any $h : X \nrightarrow Y$, $k : Y \nrightarrow Z$,*

$$(k \bullet h)^+ = k^+ \bullet h^+ \tag{A.2}$$

Proof.

$$\begin{aligned}
k^+ \bullet h^+ &= (\lambda_Z \circ Tk) \bullet (\lambda_Y \circ Th) && \text{definition of } -^+ \\
&= \mu^{\mathbf{S}}_{TZ} \circ S\lambda_Z \circ STk \circ \lambda_Y \circ Th && \text{definition of } \bullet \\
&= \mu^{\mathbf{S}}_{TZ} \circ S\lambda_Z \circ \lambda_{SZ} \circ TSk \circ Th && \text{naturality of } \lambda \\
&= \lambda_Z \circ T\mu^{\mathbf{S}}_Z \circ TSk \circ Th && (\mu^{\mathbf{S}}) \text{ axiom} \\
&= \lambda_Z \circ T(k \bullet h) && \text{definition of } \bullet \\
&= (k \bullet h)^+ && \text{definition of } -^+
\end{aligned}$$

□

Define the natural transformation $e = \lambda \circ T\eta^{\mathbf{S}} : T \to ST$. Note that $e_X = (\eta^{\mathbf{S}}_X)^+$, so that for every $h : X \nrightarrow Y$

$$e_Y \bullet h^+ = (\eta^{\mathbf{S}}_Y)^+ \bullet h^+ = (\eta^{\mathbf{S}}_Y \bullet h)^+ = h^+ \tag{A.3}$$

$$h^+ \bullet e_X = h^+ \bullet (\eta^{\mathbf{S}}_X)^+ = (h \bullet \eta^{\mathbf{S}}_X)^+ = h^+ \tag{A.4}$$

In particular, the morphism $e_X : TX \nrightarrow TX$ is idempotent in $\mathsf{Kl}(\mathbf{S})$. In the context of distributive laws, we would have $e_X = \eta^{\mathbf{S}}_{TX}$, so the operation $(-)^+$ would preserve identities and be the functor extending T to $\mathsf{Kl}(\mathbf{S})$. Moreover, $e : T \to ST$ would be a monad morphism. In our case the $\eta^{\mathbf{S}}$ diagram does not commute, so more work is needed to get a coweak extension.

By splitting the Kleisli idempotent e_X we obtain an object $\underline{T}X$ and morphisms $\pi_X : TX \nrightarrow \underline{T}X$, $\iota_X : \underline{T}X \nrightarrow TX$ such that $\pi_X \bullet \iota_X = \eta^{\mathbf{S}}_{\underline{T}X}$ and $\iota_X \bullet \pi_X = e_X$. This induces equations $\pi_X \bullet e_X = \pi_X$ and $e_X \bullet \iota_X = \iota_X$.

Now, for any $h : X \nrightarrow Y$ we define

$$\underline{T}(h) = \pi_Y \bullet h^+ \bullet \iota_X \tag{A.5}$$

Using equations (A.3) and (A.4), we have

$$\underline{T}h \bullet \pi_X = \pi_Y \bullet h^+ \tag{A.6}$$

$$\iota_Y \bullet \underline{T}h = h^+ \bullet \iota_X \tag{A.7}$$

We now list some other identities that will be often used in the proof.

Lemma A.2. *Let $f : X \to Y$, $h : Y \nrightarrow Z$ and $k : W \nrightarrow X$. Then*

$$h \bullet F_{\mathbf{S}}f = h \circ f \tag{A.8}$$

$$F_{\mathbf{S}}f \bullet k = Sf \circ k \tag{A.9}$$

$$\underline{T}F_{\mathbf{S}}f = \pi_Y \bullet F_{\mathbf{S}}Tf \bullet \iota_X \tag{A.10}$$

Proof. These are simple calculations involving only naturality of $\eta^{\mathbf{S}}$ and the monad axioms for $\mathbf{S}$:

$$h \bullet F_{\mathbf{S}}f = \mu^{\mathbf{S}}_Z \circ Sh \circ \eta^{\mathbf{S}}_Y \circ f = \mu^{\mathbf{S}}_Z \circ \eta^{\mathbf{S}}_{SZ} \circ h \circ f = h \circ f$$

$$F_{\mathbf{S}}f \bullet k = \mu^{\mathbf{S}}_Y \circ S\eta^{\mathbf{S}}_Y \circ Sf \circ k = Sf \circ k$$

For the last identity, remark that by definition of $-^+$, $F_{\mathbf{S}}$ and e:

$$(F_{\mathbf{S}}f)^+ = \lambda_Y \circ T\eta^{\mathbf{S}}_Y \circ Tf = e_Y \circ Tf$$

and then compute

$$\begin{aligned}
\underline{T}F_{\mathsf{S}}f &= \pi_Y \bullet (F_{\mathsf{S}}f)^+ \bullet \iota_X && \text{definition of } \underline{T} \\
&= \pi_Y \bullet (e_Y \circ Tf) \bullet \iota_X && \text{remark} \\
&= \pi_Y \bullet e_Y \bullet F_{\mathsf{S}}Tf \bullet \iota_X && \text{equation (A.8)} \\
&= \pi_Y \bullet F_{\mathsf{S}}Tf \bullet \iota_X && \text{splitting}
\end{aligned}$$

□

Lemma A.3. *The constructions $\underline{T}X$ and $\underline{T}h$ define a functor $\underline{T} : \mathsf{Kl}(\mathbf{S}) \to \mathsf{Kl}(\mathbf{S})$.*

Proof. Identities are preserved:

$$\begin{aligned}
\underline{T}(\eta^{\mathsf{S}}_X) &= \underline{T}F_{\mathsf{S}}(\mathrm{id}_X) && \text{definition of } F_{\mathsf{S}} \\
&= \pi_X \bullet F_{\mathsf{S}}T(\mathrm{id}_X) \bullet \iota_X && \text{equation (A.10)} \\
&= \pi_X \bullet \iota_X && \text{functors preserve id} \\
&= \eta^{\mathsf{S}}_{\underline{T}X} && \text{splitting}
\end{aligned}$$

Composition of $h : X \nrightarrow Y$ and $k : Y \nrightarrow Z$ is preserved:

$$\begin{aligned}
\underline{T}k \bullet \underline{T}h &= \pi_Z \bullet k^+ \bullet \iota_Y \bullet \pi_Y \bullet h^+ \bullet \iota_X && \text{definition of } \underline{T} \\
&= \pi_Z \bullet k^+ \bullet e_Y \bullet h^+ \bullet \iota_X && \text{splitting} \\
&= \pi_Z \bullet k^+ \bullet h^+ \bullet \iota_X && \text{equation (A.3)} \\
&= \pi_Z \bullet (k \bullet h)^+ \bullet \iota_X && \text{equation (A.2)} \\
&= \underline{T}(k \bullet h) && \text{definition of } \underline{T}
\end{aligned}$$

□

Lemma A.4. *The morphisms $(\iota_X)_{X \in \mathsf{C}}$ and $(\pi_X)_{X \in \mathsf{C}}$ define natural transformations $\iota : \underline{T}F_{\mathsf{S}} \to F_{\mathsf{S}}T$ and $\pi : F_{\mathsf{S}}T \to \underline{T}F_{\mathsf{S}}$.*

Proof. Let $f : X \to Y$. Using the previous lemmas,

$$\begin{aligned}
\iota_Y \bullet \underline{T}F_{\mathsf{S}}f &= \iota_Y \bullet \pi_Y \bullet F_{\mathsf{S}}Tf \bullet \iota_X && \text{equation (A.10)} \\
&= e_Y \bullet F_{\mathsf{S}}Tf \bullet \iota_X && \text{splitting} \\
&= (e_Y \circ Tf) \bullet \iota_X && \text{equation (A.8)} \\
&= (STf \circ e_X) \bullet \iota_X && \text{naturality of } e \\
&= F_{\mathsf{S}}Tf \bullet e_X \bullet \iota_X && \text{equation (A.9)} \\
&= F_{\mathsf{S}}Tf \bullet \iota_X && \text{splitting}
\end{aligned}$$

$$
\begin{aligned}
\underline{T}F_{\mathsf{S}}f \bullet \pi_X &= \pi_Y \bullet F_{\mathsf{S}}Tf \bullet \iota_X \bullet \pi_X && \text{equation (A.10)}\\
&= \pi_Y \bullet F_{\mathsf{S}}Tf \bullet e_X && \text{splitting}\\
&= \pi_Y \bullet (STf \circ e_X) && \text{equation (A.9)}\\
&= \pi_Y \bullet (e_Y \circ Tf) && \text{naturality of } e\\
&= \pi_Y \bullet e_Y \bullet F_{\mathsf{S}}Tf && \text{equation (A.8)}\\
&= \pi_Y \bullet F_{\mathsf{S}}Tf && \text{splitting}
\end{aligned}
$$

□

Now we define the unit and multiplication of $\underline{T}$ by

$$\eta^{\underline{\mathsf{T}}}_X = \pi_X \bullet F_{\mathsf{S}}\eta^{\mathsf{T}}_X \tag{A.11}$$

$$\mu^{\underline{\mathsf{T}}}_X = \pi_X \bullet (F_{\mathsf{S}}\mu^{\mathsf{T}}_X) \bullet \iota_{TX} \bullet \underline{T}(\iota_X) \tag{A.12}$$

and proceed to verify the monad requirements for $(\underline{T}, \eta^{\underline{\mathsf{T}}}, \mu^{\underline{\mathsf{T}}})$. First, we need a few new equations.

Lemma A.5. *For all $h : X \nrightarrow Y$,*

$$h^+ \bullet F_{\mathsf{S}}\eta^{\mathsf{T}}_X = F_{\mathsf{S}}\eta^{\mathsf{T}}_Y \bullet h \tag{A.13}$$

$$h^+ \bullet F_{\mathsf{S}}\mu^{\mathsf{T}}_X = F_{\mathsf{S}}\mu^{\mathsf{T}}_Y \bullet h^{++} \tag{A.14}$$

$$e_X \bullet F_{\mathsf{S}}\eta^{\mathsf{T}}_X = F_{\mathsf{S}}\eta^{\mathsf{T}}_X \tag{A.15}$$

$$e_X \bullet F_{\mathsf{S}}\mu^{\mathsf{T}}_X = F_{\mathsf{S}}\mu^{\mathsf{T}}_X \bullet e^+_X \tag{A.16}$$

Proof. Compute

$$
\begin{aligned}
h^+ \bullet F_{\mathsf{S}}\eta^{\mathsf{T}}_X &= h^+ \circ \eta^{\mathsf{T}}_X && \text{equation (A.8)}\\
&= \lambda_Y \circ Th \circ \eta^{\mathsf{T}}_X && \text{definition of } -^+\\
&= \lambda_Y \circ \eta^{\mathsf{T}}_{SY} \circ h && \text{naturality of } \eta^{\mathsf{T}}\\
&= S\eta^{\mathsf{T}}_Y \circ h && (\eta^{\mathsf{T}}) \text{ axiom}\\
&= F_{\mathsf{S}}\eta^{\mathsf{T}}_Y \bullet h && \text{equation (A.9)}
\end{aligned}
$$

$$
\begin{aligned}
h^+ \bullet F_{\mathsf{S}}\mu^{\mathsf{T}}_X &= h^+ \circ \mu^{\mathsf{T}}_X && \text{equation (A.9)}\\
&= \lambda_Y \circ Th \circ \mu^{\mathsf{T}}_X && \text{definition of } -^+\\
&= \lambda_Y \circ \mu^{\mathsf{T}}_{SY} \circ TTh && \text{naturality of } \mu^{\mathsf{T}}\\
&= S\mu^{\mathsf{T}}_Y \circ \lambda_{TY} \circ T\lambda_Y \circ TTh && (\mu^{\mathsf{T}}) \text{ axiom}\\
&= S\mu^{\mathsf{T}}_Y \circ \lambda_{TY} \circ T(h^+) && \text{definition of } -^+\\
&= S\mu^{\mathsf{T}}_Y \circ h^{++} && \text{definition of } -^+\\
&= F_{\mathsf{S}}\mu^{\mathsf{T}}_Y \bullet h^{++} && \text{equation (A.9)}
\end{aligned}
$$

By injecting $h = \eta^{\mathsf{S}}_X$ in equations (A.13) and (A.14) we obtain equations (A.15) and (A.16). □

We now have enough intermediate results to be able to work with commutative diagrams directly in $\mathsf{Kl}(\mathsf{S})$, as the next lemmas show. For the sake of readability, in the subsequent diagrams we use the notation $\rightarrow$ instead of $\nrightarrow$.

Lemma A.6. *Both $\eta^{\underline{\mathbf{T}}} : 1 \nrightarrow \underline{T}$ and $\mu^{\underline{\mathbf{T}}} : \underline{T}\underline{T} \nrightarrow \underline{T}$ are natural transformations.*

Proof. Naturality of $\eta^{\underline{\mathbf{T}}}$ is established by the following commutative diagram. Unmarked regions commute by unfolding the definition of $\eta^{\underline{\mathbf{T}}}$.

$$\begin{tikzcd}
X \arrow[rrr, "h"] \arrow[dr, "F_{\mathsf{S}}\eta^{\mathbf{T}}_X"] \arrow[dd, "\eta^{\underline{\mathbf{T}}}_X"'] & & & Y \arrow[dl, "F_{\mathsf{S}}\eta^{\mathbf{T}}_Y"'] \arrow[dd, "\eta^{\underline{\mathbf{T}}}_Y"] \\
& TX \arrow[r, "h^+"] \arrow[dl, "\pi_X"] \arrow[ur, phantom, "\text{(A.13)}"] & TY \arrow[dr, "\pi_Y"] & \\
\underline{T}X \arrow[rrr, "\underline{T}h"'] & \arrow[ur, phantom, "\text{(A.6)}"] & & \underline{T}Y
\end{tikzcd}$$

Naturality of $\mu^{\underline{\mathbf{T}}}$ is established by the following commutative diagram. Unmarked regions commute by unfolding the definition of $\mu^{\underline{\mathbf{T}}}$.

$$\begin{tikzcd}
\underline{T}\underline{T}X \arrow[rrr, "\underline{T}\underline{T}h"] \arrow[dr, "\underline{T}\iota_X"] \arrow[dddd, "\mu^{\underline{\mathbf{T}}}_X"'] & & & \underline{T}\underline{T}Y \arrow[dl, "\underline{T}\iota_Y"'] \arrow[dddd, "\mu^{\underline{\mathbf{T}}}_Y"] \\
& \underline{T}TX \arrow[r, "\underline{T}(h^+)"] \arrow[d, "\iota_{TX}"'] \arrow[ur, phantom, "\text{(A.7)}"] & \underline{T}TY \arrow[d, "\iota_{TY}"] & \\
& TTX \arrow[r, "h^{++}"] \arrow[d, "F_{\mathsf{S}}\mu^{\mathbf{T}}_X"'] \arrow[ur, phantom, "\text{(A.7)}"] & TTY \arrow[d, "F_{\mathsf{S}}\mu^{\mathbf{T}}_Y"] & \\
& TX \arrow[r, "h^+"] \arrow[dl, "\pi_X"'] \arrow[ur, phantom, "\text{(A.14)}"] & TY \arrow[dr, "\pi_Y"] & \\
\underline{T}X \arrow[rrr, "\underline{T}h"'] & \arrow[ur, phantom, "\text{(A.6)}"] & & \underline{T}Y
\end{tikzcd}$$

□

Lemma A.7. *The equations of monads are satisfied by $\eta^{\underline{\mathbf{T}}}$ and $\mu^{\underline{\mathbf{T}}}$.*

Proof. The following diagram proves that the axiom $\mu^{\underline{\mathbf{T}}} \bullet \underline{T}\eta^{\underline{\mathbf{T}}} = 1$ holds, because the top-right path equals $\pi \bullet \iota = 1$. The top right triangle commutes by the monad axiom

$\mu^{\mathbf{T}} \circ T\eta^{\mathbf{T}} = \mathrm{id}_T$. Other unmarked regions commute by definition of $\eta^{\underline{\mathbf{T}}}$, $\mu^{\underline{\mathbf{T}}}$ and the splitting.

$$
\begin{tikzcd}
\underline{T}X \arrow[rrrr, "\iota_X"] \arrow[dr, dashed, "\underline{T}F_{\mathsf{S}}\eta^{\mathbf{T}}_X"] \arrow[dddd, "\underline{T}\eta^{\underline{\mathbf{T}}}_X"'] & & & & TX \arrow[dl, dashed, "F_{\mathsf{S}}T\eta^{\mathbf{T}}_X"'] \arrow[dr, equal] & \\
& \underline{T}TX \arrow[dr, "\iota_{TX}"] \arrow[dd, "\underline{T}e_X"] \arrow[dddl, "\underline{T}\pi_X"'] & \arrow[u, phantom, "(A.10)"] & TTX \arrow[ll, "\pi_{TX}"'] \arrow[dl, "e_{TX}"'] \arrow[dd, "(e_X)^+"] \arrow[rr, "F_{\mathsf{S}}\mu^{\mathbf{T}}_X"'] & & TX \arrow[ddl, "e_X"'] \arrow[ddd, "\pi_X"] \\
& & TTX \arrow[r, phantom, "(A.4)"] \arrow[dr, "(e_X)^+"'] \arrow[dl, phantom, "(A.7)"] & \arrow[rr, phantom, "(A.16)"] & & \\
& \underline{T}TX \arrow[rr, "\iota_{TX}"'] & & TTX \arrow[r, "F_{\mathsf{S}}\mu^{\mathbf{T}}_X"] & TX \arrow[dr, "\pi_X"] & \\
\underline{T}\underline{T}X \arrow[ur, "\underline{T}\iota_X"'] \arrow[rrrrr, "\mu^{\underline{\mathbf{T}}}_X"'] & & & & & \underline{T}X
\end{tikzcd}
$$

The following diagram proves that the axiom $\mu^{\underline{\mathbf{T}}} \bullet \eta^{\underline{\mathbf{T}}}\underline{T} = 1$ holds, because the top-right path equals $\pi \bullet \iota = 1$. The top right triangle commutes by the monad axiom $\mu^{\mathbf{T}} \circ \eta^{\mathbf{T}}T = \mathrm{id}_T$. Other unmarked regions commute by definition of $\eta^{\underline{\mathbf{T}}}$, $\mu^{\underline{\mathbf{T}}}$ and the splitting.

$$
\begin{tikzcd}
\underline{T}X \arrow[rrr, "\iota_X"] \arrow[dr, dashed, "F_{\mathsf{S}}\eta^{\mathbf{T}}_{\underline{T}X}"] \arrow[dddd, "\eta^{\underline{\mathbf{T}}}_{\underline{T}X}"'] & & & TX \arrow[d, dashed, "F_{\mathsf{S}}\eta^{\mathbf{T}}_{TX}"] \arrow[dr, equal] & \\
& T\underline{T}X \arrow[rr, "\iota^+_X"] \arrow[dr, "\iota^+_X"] \arrow[dddl, "\pi_{\underline{T}X}"'] & \arrow[u, phantom, "(A.13)"] & TTX \arrow[r, "F_{\mathsf{S}}\mu^{\mathbf{T}}_X"'] & TX \arrow[ddd, "\pi_X"] \\
& & TTX \arrow[u, phantom, "(A.3)"] \arrow[ur, "e_{TX}"] \arrow[dr, "\pi_{TX}"] \arrow[dl, phantom, "(A.6)"] & & \\
& & & \underline{T}TX \arrow[uu, "\iota_{TX}"'] & \\
\underline{T}\underline{T}X \arrow[urrr, "\underline{T}\iota_X"'] \arrow[rrrr, "\mu^{\underline{\mathbf{T}}}_X"'] & & & & \underline{T}X
\end{tikzcd}
$$

The following diagram shows that $\mu^{\underline{\mathbf{T}}}$ is associative. Unmarked regions commute by definition of $\mu^{\underline{\mathbf{T}}}$ and splitting properties.

$$
\begin{tikzcd}
\underline{TT}TX \arrow[rrrrrr, "\mu^{\underline{\mathbf{T}}}_{\underline{T}X}"] \arrow[drr, "\underline{T}\iota_{\underline{T}X}"] \arrow[ddr, "\underline{TT}\iota_X"'] \arrow[dddddd, "\underline{T}\mu^{\underline{\mathbf{T}}}_X"'] & & & & & & \underline{TT}X \arrow[ddl, "\underline{T}\iota_X"] \arrow[dddddd, "\mu^{\underline{\mathbf{T}}}_X"] \\
& \arrow[r, phantom, "(A.7)"] & \underline{T}T\underline{T}X \arrow[r, "\iota_{T\underline{T}X}"] \arrow[ddl, "\underline{T}(\iota^+_X)"'] & TT\underline{T}X \arrow[r, "F_{\mathsf{S}}\mu^{\mathbf{T}}_{\underline{T}X}"] \arrow[ddl, "\iota^{++}_X"'] \arrow[d, phantom, "(A.3)"] & T\underline{T}X \arrow[urr, "\pi_{\underline{T}X}"'] \arrow[d, "\iota^+_X"] \arrow[r, phantom, "(A.6)"] & {} & \\
& \underline{TT}TX \arrow[d, "\underline{T}\iota_{TX}"'] \arrow[r, phantom, "(A.7)"] & {} & \text{and} \arrow[d, phantom, "(A.14)"] & TTX \arrow[r, "\pi_{TX}"] \arrow[d, dashed, "e_{TX}"] & \underline{T}TX \arrow[dl, "\iota_{TX}"] & \\
& \underline{T}TTX \arrow[r, "\iota_{TTX}"] \arrow[d, "\underline{T}F_{\mathsf{S}}\mu^{\mathbf{T}}_X"'] \arrow[dr, phantom, "\text{naturality}"] & TTTX \arrow[rr, "F_{\mathsf{S}}\mu^{\mathbf{T}}_{TX}"] \arrow[d, "F_{\mathsf{S}}\mu^{\mathbf{T}}_{TX}"] \arrow[drr, phantom, "\text{associativity}"] & {} & TTX \arrow[dr, "F_{\mathsf{S}}\mu^{\mathbf{T}}_X"] & & \\
& \underline{T}TX \arrow[r, "\iota_{TX}"] \arrow[ddl, "\underline{T}\pi_X"'] \arrow[dr, "\underline{T}e_X"'] & TTX \arrow[rrr, "F_{\mathsf{S}}\mu^{\mathbf{T}}_X"] \arrow[dr, "e^+_X"] \arrow[d, phantom, "(A.7)"] & \arrow[dr, phantom, "(A.16)"] & & TX \arrow[dl, "e_X"] \arrow[ddr, "\pi_X"] & \\
& & \underline{T}TX \arrow[r, "\iota_{TX}"] & TTX \arrow[r, "F_{\mathsf{S}}\mu^{\mathbf{T}}_X"] & TX \arrow[drr, "\pi_X"] & & \\
\underline{TT}X \arrow[urr, "\underline{T}\iota_X"'] \arrow[rrrrrr, "\mu^{\underline{\mathbf{T}}}_X"'] & & & & & & \underline{T}X
\end{tikzcd}
$$

This achieves the proof that $(\underline{T}, \eta^{\underline{\mathbf{T}}}, \mu^{\underline{\mathbf{T}}})$ is a monad. $\square$

To achieve the construction of the coweak extension, we need to establish the four commutative diagrams stating that π and ι are compatible with $\eta^{\mathbf{T}}$ and $\mu^{\mathbf{T}}$.

Lemma A.8. *The coweak lifting diagrams* $(\pi.\eta^{\mathbf{T}})$, $(\iota.\eta^{\mathbf{T}})$, $(\pi.\mu^{\mathbf{T}})$ *and* $(\iota.\mu^{\mathbf{T}})$ *commute.*

Proof. The $(\pi.\eta^{\mathbf{T}})$ diagram comes right from the definition of $\eta^{\underline{\mathbf{T}}}$:

$$\eta^{\underline{\mathbf{T}}}_X = \pi_X \bullet F_{\mathsf{S}}\eta^{\mathbf{T}}_X \tag{A.17}$$

The $(\iota.\eta^{\mathbf{T}})$ diagram commutes because:

$$
\begin{aligned}
\iota_X \bullet \eta^{\underline{\mathbf{T}}}_X &= \iota_X \bullet \pi_X \bullet F_{\mathsf{S}}\eta^{\mathbf{T}}_X && \text{definition of } \eta^{\underline{\mathbf{T}}} \\
&= e_X \bullet F_{\mathsf{S}}\eta^{\mathbf{T}}_X && \text{splitting} \\
&= F_{\mathsf{S}}\eta^{\mathbf{T}}_X && \text{equation (A.15)}
\end{aligned}
$$

The $(\pi.\mu^{\mathbf{T}})$ diagram commutes:

$$
\begin{tikzcd}
TTX \arrow[rr, "\pi_{TX}"] \arrow[rrd, "e_{TX}"] \arrow[rdd, "(e_X)^+"'] \arrow[dddd, "F_{\mathsf{S}}\mu^{\mathbf{T}}_X"'] \arrow[rd, phantom, "(A.4)"] & & \underline{T}TX \arrow[rr, "\underline{T}\pi_X"] \arrow[d, dashed, "\iota_{TX}"] \arrow[rdd, "\underline{T}e_X"] & & \underline{T}\underline{T}X \arrow[ldd, "\underline{T}\iota_X"'] \arrow[dddd, "\mu^{\underline{\mathbf{T}}}_X"] \\
 & & TTX \arrow[ld, "(e_X)^+"] \arrow[d, phantom, "(A.7)"] & & \\
 & TTX \arrow[rd, "F_{\mathsf{S}}\mu^{\mathbf{T}}_X"'] \arrow[ldd, phantom, "(A.16)"] & & \underline{T}TX \arrow[ll, "\iota_{TX}"] & \\
 & & TX \arrow[rrd, "\pi_X"] & & \\
TX \arrow[rru, "e_X"] \arrow[rrrr, "\pi_X"'] & & & & \underline{T}X
\end{tikzcd}
$$

Finally, the $(\iota.\mu^{\mathbf{T}})$ diagram commutes:

$$
\begin{tikzcd}
\underline{T}\underline{T}X \arrow[rr, "\underline{T}\iota_X"] \arrow[rd, "\underline{T}\iota_X"'] \arrow[dddd, "\mu^{\underline{\mathbf{T}}}_X"'] & & \underline{T}TX \arrow[r, "\iota_{TX}"] \arrow[d, phantom, "(A.7)"] & TTX \arrow[dddd, "F_{\mathsf{S}}\mu^{\mathbf{T}}_X"] \\
 & \underline{T}TX \arrow[ru, "\underline{T}e_X"'] \arrow[rd, "\iota_{TX}"] & & \\
 & & TTX \arrow[ruu, "e_X^+"'] \arrow[d, "F_{\mathsf{S}}\mu^{\mathbf{T}}_X"] \arrow[r, phantom, "(A.16)"] & {} \\
 & & TX \arrow[lld, "\pi_X"'] \arrow[rd, "e_X"] & \\
\underline{T}X \arrow[rrr, "\iota_X"'] & & & TX
\end{tikzcd}
$$

□

Coweak Extensions ⇒ Coweak Distributive Laws

Let $\underline{\mathbf{T}}$ be a coweak extension of $\mathbf{T}$ to $\mathsf{Kl}(\mathbf{S})$. Define

$$\lambda_X = TSX \xrightarrow{\pi_{SX}} \underline{T}SX \xrightarrow{\underline{T}(\mathrm{id}_{SX})} \underline{T}X \xrightarrow{\iota_X} TX \tag{A.18}$$

In the proof, we will often use equations (A.8) and (A.9). As the upcoming commutative diagrams all live in $\mathsf{Kl}(\mathbf{S})$, we again use the notation $\to$ instead of $\nrightarrow$ for readability.

Lemma A.9. *The morphisms $(\lambda_X)_{X\in\mathsf{C}}$ define a natural $\lambda : TS \to ST$.*

Proof. Let $f : X \to Y$, we must prove that $STf \circ \lambda_X = \lambda_Y \circ TSf$. This is equivalent to $F_{\mathsf{S}}Tf \bullet \lambda_X = \lambda_Y \bullet F_{\mathsf{S}}TSf$, which is true according to the diagram:

$$
\begin{tikzcd}
TSX \arrow[rrr, "\lambda_X"] \arrow[dr, "\pi_{SX}"] \arrow[ddd, "F_{\mathsf{S}}TSf"'] & & & TX \arrow[ddd, "F_{\mathsf{S}}Tf"] \\
& \underline{T}SX \arrow[r, "\underline{T}(\mathrm{id}_{SX})"'] \arrow[d, "\underline{T}F_{\mathsf{S}}Sf"] \arrow[ur, phantom, "\text{def}"] & \underline{T}X \arrow[ur, "\iota_X"] \arrow[d, "\underline{T}F_{\mathsf{S}}f"'] & \\
\arrow[r, phantom, "\text{nat}"] & \underline{T}SY \arrow[r, "\underline{T}(\mathrm{id}_{SY})"] & \underline{T}Y \arrow[dr, "\iota_Y"] \arrow[r, phantom, "\text{nat}"] & \\
TSY \arrow[rrr, "\lambda_Y"'] \arrow[ur, "\pi_{SY}"'] & & \arrow[u, phantom, "\text{def}"] & TY
\end{tikzcd}
$$

The middle square commutes because

$$\mathrm{id}_{SY} \bullet F_{\mathsf{S}}Sf = \mathrm{id}_{SY} \circ Sf = Sf = Sf \circ \mathrm{id}_{SX} = F_{\mathsf{S}}f \bullet \mathrm{id}_{SX} \tag{A.19}$$

$\square$

Lemma A.10. *The (η^{T}) diagram is satisfied by λ.*

Proof. This amounts to proving that $S\eta^{\mathsf{T}}_X = \lambda_X \circ \eta^{\mathsf{T}}_{SX}$. Note that

$$
\begin{aligned}
&F_{\mathsf{S}}\eta^{\mathsf{T}}_X \bullet \mathrm{id}_{SX} = S\eta^{\mathsf{T}}_X \circ \mathrm{id}_{SX} = S\eta^{\mathsf{T}}_X \\
&\lambda_X \bullet F_{\mathsf{S}}TSf = \lambda_X \circ TSf
\end{aligned}
$$

so it is actually equivalent to proving that $F_{\mathsf{S}}\eta^{\mathsf{T}}_X \bullet \mathrm{id}_{SX} = \lambda_X \bullet F_{\mathsf{S}}TSf$. This is done in the following diagram:

$$
\begin{tikzcd}
SX \arrow[rrr, "\mathrm{id}_{SX}"] \arrow[dr, "\eta^{\mathsf{T}}_{SX}"] \arrow[dd, "F_{\mathsf{S}}\eta^{\mathsf{T}}_{SX}"'] & & & X \arrow[dl, "\eta^{\mathsf{T}}_X"'] \arrow[dd, "F_{\mathsf{S}}\eta^{\mathsf{T}}_X"] \\
(\pi.\eta^{\mathsf{T}}) & \underline{T}SX \arrow[r, "\underline{T}(\mathrm{id}_{SX})"] \arrow[ur, phantom, "\text{nat}"] & \underline{T}X \arrow[dr, "\iota_X"] & (\iota.\eta^{\mathsf{T}}) \\
TSX \arrow[rrr, "\lambda_X"'] \arrow[ur, "\pi_{SX}"] & & \arrow[u, phantom, "\text{def}"] & TX
\end{tikzcd}
$$

$\square$

Lemma A.11. *The* (μ^{S}) *diagram is satisfied by* λ.

Proof. This amounts to showing that $\mu^{\mathsf{S}}_{TX} \circ S\lambda_X \circ \lambda_{SX} = \lambda_X \circ T\mu^{\mathsf{S}}_X$. Note that

$$\begin{aligned}\lambda_X \bullet \lambda_{SX} &= \mu^{\mathsf{S}}_{TX} \circ S\lambda_X \circ \lambda_{SX}\\ \lambda_X \bullet F_{\mathsf{S}}T\mu^{\mathsf{S}}_X &= \lambda_X \circ T\mu^{\mathsf{S}}_X\end{aligned}$$

so it is actually equivalent to proving that $\lambda_X \bullet \lambda_{SX} = \lambda_X \bullet F_{\mathsf{S}}T\mu^{\mathsf{S}}_X$. This is done in the following diagram:

$$\begin{tikzcd}
TSSX \arrow[rrr, "\lambda_{SX}"] \arrow[dr, "\pi_{SSX}"] \arrow[ddd, "F_{\mathsf{S}}T\mu^{\mathsf{S}}_X"'] & & & TSX \arrow[rrr, "\lambda_X"] \arrow[dr, "\pi_{SX}"] \arrow[d, phantom, "\text{split}"] & & & TX \arrow[ddd, equal] \\
\arrow[r, phantom, "\text{nat}"] & \underline{T}SSX \arrow[r, "\underline{T}(\mathrm{id}_{SSX})"'] \arrow[d, dashed, "\underline{T}F_{\mathsf{S}}\mu^{\mathsf{S}}_X"] & \underline{T}SX \arrow[ur, "\iota_{SX}"] \arrow[rr, equal] & & \underline{T}SX \arrow[r, "\underline{T}(\mathrm{id}_{SX})"'] & \underline{T}X \arrow[ur, "\iota_X"'] \arrow[d, equal] & \\
 & \underline{T}SX \arrow[rrrr, "\underline{T}(\mathrm{id}_{SX})"] & & \arrow[d, phantom, "\text{def}"] & & \underline{T}X \arrow[dr, "\iota_X"] & \\
TSX \arrow[ur, "\pi_{SX}"'] \arrow[rrrrrr, "\lambda_X"'] & & & & & & TX
\end{tikzcd}$$

The middle rectangle commutes because

$$\mathrm{id}_{SX} \bullet \mathrm{id}_{SSX} = \mu^{\mathsf{S}}_X \circ S\,\mathrm{id}_{SX} \circ \mathrm{id}_{SSX} = \mu^{\mathsf{S}}_X = \mathrm{id}_{SX} \circ \mu^{\mathsf{S}}_X = \mathrm{id}_{SX} \bullet F_{\mathsf{S}}\mu^{\mathsf{S}}_X$$

□

Lemma A.12. *The* (μ^{T}) *diagram is satisfied by* λ.

Proof. This amounts to showing that $S\mu^{\mathsf{T}}_X \circ \lambda_{TX} \circ T\lambda_X = \lambda_X \circ \mu^{\mathsf{T}}_{SX}$. Note that

$$\begin{aligned}&F_{\mathsf{S}}\mu^{\mathsf{T}}_X \bullet \lambda_{TX} \bullet F_{\mathsf{S}}T\lambda_X = F_{\mathsf{S}}\mu^{\mathsf{T}}_X \bullet (\lambda_{TX} \circ T\lambda_X) = S\mu^{\mathsf{T}}_X \circ \lambda_{TX} \circ T\lambda_X\\ &\lambda_X \bullet F_{\mathsf{S}}\mu^{\mathsf{T}}_{SX} = \lambda_X \circ \mu^{\mathsf{T}}_{SX}\end{aligned}$$

so it is actually equivalent to proving that $F_{\mathsf{S}}\mu^{\mathsf{T}}_X \bullet \lambda_{TX} \bullet F_{\mathsf{S}}T\lambda_X = \lambda_X \bullet F_{\mathsf{S}}\mu^{\mathsf{T}}_{SX}$. This is done in the following diagram:

$$
\begin{tikzcd}
TTSX \arrow[rr, "F_{\mathsf{S}}T\lambda_X"] \arrow[dr, "\pi_{TSX}"] \arrow[dddd, "F_{\mathsf{S}}\mu^{\mathsf{T}}_{SX}"'] \arrow[ddr, phantom, "(\pi.\mu^{\mathsf{T}})"] & & TSTX \arrow[rrr, "\lambda_{TX}"] \arrow[dr, "\pi_{STX}"] & & & TTX \arrow[dddd, "F_{\mathsf{S}}\mu^{\mathsf{T}}_X"] \\
& \underline{T}TSX \arrow[rr, "\underline{T}F_{\mathsf{S}}\lambda_X"] \arrow[rrr, bend right=10, "\underline{T}\lambda_X"'] \arrow[d, "\underline{T}\pi_{SX}"] \arrow[ur, phantom, "\text{nat}"] \arrow[ddrrr, phantom, "\text{def}"] & & \underline{T}STX \arrow[r, "\underline{T}(\mathrm{id}_{STX})"] \arrow[ur, phantom, "\text{def}"] & \underline{T}TX \arrow[ur, "\iota_{TX}"'] \arrow[ddr, phantom, "(\iota.\mu^{\mathsf{T}})"] & \\
& \underline{TT}SX \arrow[rrr, "\underline{TT}(\mathrm{id}_{SX})"'] \arrow[d, "\mu^{\mathsf{T}}_{SX}"] \arrow[drrr, phantom, "\text{nat}"] & & & \underline{TT}X \arrow[u, "\underline{T}\iota_X"'] \arrow[d, "\mu^{\mathsf{T}}_X"] & \\
& \underline{T}SX \arrow[rrr, "\underline{T}(\mathrm{id}_{SX})"'] \arrow[drrr, phantom, "\text{def}"] & & & \underline{T}X \arrow[dr, "\iota_X"'] & \\
TSX \arrow[ur, "\pi_{SX}"'] \arrow[rrrrr, "\lambda_X"'] & & & & & TX
\end{tikzcd}
$$

The (unmarked) flat triangle commutes because

$$\mathrm{id}_{STX} \bullet F_{\mathsf{S}}\lambda_X = \mathrm{id}_{STX} \circ \lambda_X = \lambda_X$$

□

Appendix B

Coq Proofs

This Appendix contains formal Coq proofs for the following results of Chapter 6. These proofs can also be found on Github [61].

- `Theorem eta_nearly_cartesian` is a proof of Proposition 6.18
- `Theorem mu_nearly_cartesian` is a proof of Proposition 6.20
- `Theorem monotone_weak_dlaw` is a proof of Proposition 6.22
- `Theorem dlaw_degenerate` is a proof of Proposition 6.19

```
(** * Powerset weakly distributes over itself in toposes
This file is an appendix of the PhD thesis of Alexandre Goy
untitled "On the Compositionality of Monads via Weak Distributive Laws"
(2021).
It contains formal proofs, valid in [Set]
and more generally in any topos, that:
  1 - [Theorem eta_nearly_cartesian]
      the unit of the powerset monad is nearly cartesian
      iff the topos is degenerate
  2 - [Theorem mu_nearly_cartesian]
      the multiplication of the powerset monad is nearly cartesian
  3 - [Theorem monotone_weak_dlaw]
      the unique monotone weak distributive law from the powerset
      over itself is given by the expected Egli-Milner formula
  4 - [Theorem dlaw_degenerate]
      if there is a distributive law [PP → PP]
      then the topos is degenerate
*)

(** * Preliminaries *)

(** We define the powerset monad in the internal logic of a topos.
```

```
  - [P] is the action of the functor on objects (powerset).
     Note that [Prop] plays the role of the subobject classifier.
  - [im] is the action of the functor on morphisms (direct image).
  - [etaP] is the unit (singleton).
  - [muP] is the multiplication (union).
*)
Definition P X := X → Prop.
Definition im [X] [Y] (f : X → Y) (a : P X) (y : Y) :=
    exists (x : X), a x ∧ f x = y.
Definition etaP X (x : X) (x' : X) := x = x'.
Definition muP X (t : P (P X)) (x : X) := exists (s : P X), s x ∧ t s.

(** In any topos, extensionality holds. *)
Axiom ext : forall X, forall A B : P X,
  (forall (x : X), A x ↔ B x) → A = B.

(** A few more constructs will be needed.
  - [inter] is the intersection of two subobjets.
  - [preim] is the preimage of a subobject under a morphism. *)
Definition inter [X] (A B : P X) (x : X) := A x ∧ B x.
Notation "A ∩ B" := (inter A B) (at level 40).
Definition preim [X] [Y] (f : X → Y) (s' : P Y) (x : X) := s' (f x).

(** * 1 - Unit *)

Inductive terminal := elt.
Definition Prop_into_terminal (P : Prop) := elt.
(** We define [full] to be the maximal subobject of [Prop]. *)
Definition full (P : Prop) := True.

Lemma image_of_full_is_singleton :
    im Prop_into_terminal full = etaP terminal elt.
Proof.
apply ext. unfold im,full,etaP,Prop_into_terminal.
intro. destruct x. split ; intro.
  - reflexivity.
  - exists True. split ; trivial.
Qed.

(** The unit [etaP] is nearly cartesian
if and only if the topos is degenerate. *)
Theorem eta_nearly_cartesian :
  (forall X Y (f : X → Y) (s : P X) (y : Y),
  im f s = etaP Y y →
  exists (x : X), etaP X x = s ∧ f x = y) ↔ (True = False).
Proof.
split ; intros.
```

```
  - specialize (H Prop terminal Prop_into_terminal).
    pose proof H full elt image_of_full_is_singleton.
    destruct H0,H0.
    assert (etaP Prop x True ↔ True) by (rewrite H0 ; intuition).
    assert (etaP Prop x False ↔ True) by (rewrite H0 ; intuition).
    unfold etaP in H2,H3.
    assert (x = True) by (apply H2 ; trivial).
    assert (x = False) by (apply H3 ; trivial). rewrite H4 in H5.
    assumption.
  - exfalso. rewrite ← H. trivial.
Qed.

(** * 2 - Multiplication *)

Lemma frobenius : forall [X] [Y] f (s : P X) (s' : P Y),
    im f (s ∩ (preim f s')) = (im f s) ∩ s'.
Proof.
intros. apply ext. unfold preim,inter,im. intro y.
split ; intros ; destruct H,H,H.
  - rewrite H0 in H1. split ; try assumption.
    exists x. split ; assumption.
  - rewrite ← H1 in H0. exists x. auto.
Qed.

Lemma subset_in_union :
  forall X (t : P (P X)) (s : P X), t s → muP X t ∩ s = s.
Proof.
intros. apply ext. unfold muP,inter. intro. split ; intro.
  - destruct H0. assumption.
  - split ; try assumption. exists s. split ; assumption.
Qed.

(** The multiplication [muP] is nearly cartesian. *)
Theorem mu_nearly_cartesian :
  forall X Y (f : X → Y) (s : P X) (t' : P (P Y)),
  im f s = muP Y t' →
  exists (t : P (P X)), muP X t = s ∧ im (im f) t = t'.
Proof.
intros. exists (im (fun s' ⇒ s ∩ (preim f s')) t'). split ; apply ext.
  - intro. unfold muP,im. split ; intro.
    + destruct H0 as (s0,H0),H0,H1 as (s',H1),H1.
      rewrite ← H2 in H0. destruct H0. assumption.
    + assert ((im f s) (f x))
      by (unfold im ; exists x ; split ; try assumption ; reflexivity).
      rewrite H in H1. destruct H1 as (s', H1), H1.
      exists (s ∩ (preim f s')). split.
      * split ; assumption.
```

```
      * exists s'. auto.
  - intro s'. unfold im. split ; intro.
    + destruct H0 as (s0,H0),H0,H0 as (s'0,H0),H0.
      pose proof frobenius f s s'0. rewrite H2,H in H3.
      rewrite subset_in_union in H3 ; try assumption.
      unfold im in H3. rewrite <- H1, H3. assumption.
    + exists (s ∩ (preim f s')). split.
      * exists s'. auto.
      * pose proof frobenius f s s'. rewrite H in H1.
        rewrite subset_in_union in H1 ; assumption.
Qed.

(** * 3 - Monotone weak distributive law *)

Definition relation [X] [Y] (R : P(X * Y)) := { t : X * Y | R t}.
Definition proj1 [X] [Y] (R : P(X * Y)) (t : relation R) :=
    fst (proj1_sig t).
Definition proj2 [X] [Y] (R : P(X * Y)) (t : relation R) :=
    snd (proj1_sig t).
Definition Pext [X] [Y] (R : P(X * Y)) (u : P X * P Y) :=
    exists (t : P (relation R)),
    im (proj1 R) t = fst u /\ im (proj2 R) t = snd u.

Lemma pair_in_relation [X] [Y] (R : P(X * Y)) :
    forall r : relation R, R (proj1 R r, proj2 R r).
Proof.
intro. destruct r as (t,r0). destruct t.
unfold proj1,proj2. simpl. assumption.
Qed.

(** The relational extension is given by the Egli-Milner formula. *)
Proposition Egli_Milner [X] [Y] (R : P(X * Y)) : forall (u : P X * P Y),
    Pext R u <->
      ((forall x : X, (fst u) x -> exists y : Y, (snd u) y /\ R (x,y))
    /\ (forall y : Y, (snd u) y -> exists x : X, (fst u) x /\ R (x,y))).
Proof.
intro. destruct u as (s,s'). unfold Pext. simpl. split ; intro.
  - destruct H as (t,H),H. rewrite <- H, <- H0. unfold im. split.
    + intros. destruct H1 as (r,H1). destruct H1.
      exists (proj2 R r). split.
      * exists r. auto.
      * rewrite <- H2. apply pair_in_relation.
    + intros. destruct H1 as (r,H1). destruct H1.
      exists (proj1 R r). split.
      * exists r. auto.
      * rewrite <- H2. apply pair_in_relation.
  - destruct H. pose (fun (t0 : relation R) =>
```

```
    s (fst (proj1_sig t0)) ∧ s' (snd (proj1_sig t0))) as t.
    exists t. split.
    + apply ext. intro x. split ; intro.
      * unfold im in H1. destruct H1 as (r,H1),H1. destruct r as (r0,H3).
        unfold t in H1. unfold proj1 in H2. simpl in *.
        destruct H1. rewrite H2 in H1. assumption.
      * specialize (H x H1). destruct H as (y,H),H. unfold im.
        exists (exist R (x,y) H2). unfold t. simpl.
        split ; try split ; assumption.
    + apply ext. intro y. split ; intro.
      * unfold im in H1. destruct H1 as (r,H1),H1.
        destruct r as (r0,H3). unfold t in H1.
        unfold proj2 in H2. simpl in *. destruct H1.
        rewrite H2 in H4. assumption.
      * specialize (H0 y H1). destruct H0 as (x,H0),H0.
        unfold im. exists (exist R (x,y) H2).
        unfold t. simpl. split ; try split ; assumption.
Qed.

Definition ni X (t : P X * X) := (fst t) (snd t).
Definition lambda_m X (t : P (P X)) (s : P X) := Pext (ni X) (t,s).

(** The monotone weak distributive law [lambda] expresses as follows. *)
Theorem monotone_weak_dlaw X : forall (t : P (P X)) (s : P X),
    lambda_m X t s ↔
    (forall x : X, s x → muP X t x)
     ∧ (forall u : (P X), t u → exists (x : X), u x ∧ s x).
Proof.
  intros. unfold lambda_m. unfold muP. split ; intro.
  - apply Egli_Milner in H. unfold fst,snd in H. destruct H. split.
    + intros. specialize (H0 x H1). destruct H0 as (s0,H0),H0.
      exists s0. split ; assumption.
    + intros. specialize (H u H1). destruct H as (x,H),H.
      exists x. split ; assumption.
  - apply Egli_Milner. unfold fst,snd. destruct H. split.
    + intro u. intro. specialize (H0 u H1).
      destruct H0 as (x,H0). exists x. destruct H0. split ; assumption.
    + intro x. intro. specialize (H x H1).
      destruct H as (u,H). exists u. destruct H. split ; assumption.
Qed.

(** * 4 - Distributive law *)

(** We prove that if there is a natural transformation
[lambda : PP → PP] satisyfing both unit axioms, then [True = False].
This is the formalisation of a proof by Klin and Salamanca in
  - (1) "Iterated covariant powerset is not a monad", Theorem 2.4
```

```
itself inspired from a proof of Varacca following an idea of Plotkin
  - (2) "Distributing probability over nondeterminism", Theorem 3.2. *)

Definition subset [X] (s t : P X) := forall x : X, s x → t x.
Notation "s ⊆ t" := (subset s t) (at level 60).
Lemma subset_inter [X] (s t t' : P X) : s ⊆ t → s ⊆ t' → s ⊆ t ∩ t'.
Proof. unfold subset,inter. intuition. Qed.
Lemma subset_refl [X] (s t : P X) : s = t → s ⊆ t.
Proof. unfold subset. intros. rewrite H in H0. assumption. Qed.
Lemma im_nonempty [X] [Y] (f : X → Y) (s : P X) :
    (exists (y : Y), (im f) s y) → (exists (x : X), s x).
Proof. intro. destruct H as (y,H),H,H. exists x. assumption. Qed.

(** The functor [P] preserves preimages. *)
Proposition pres_preim [X] [Y] (f : X → Y) :
    forall (s' : P Y) (s : P X), ((im f) s) ⊆ s' → s ⊆ (preim f s').
Proof.
intros. unfold subset,im,preim. intros. apply H. exists x. auto.
Qed.

(** Preliminaries: we define the operation
  [two] that builds a set out of two elements.
  In the terminology of (1) this is a "non-trivial idempotent term". *)
Definition two [X] (x : X) (y : X) (z : X) := z = x ∨ z = y.

Lemma symmetry [X] : forall (x y : X), two x y = two y x.
Proof. unfold two. intros. apply ext. intuition. Qed.

Lemma idempotence [X] : forall (x : X), two x x = etaP X x.
Proof. intro. apply ext. unfold two,etaP. intuition. Qed.

Lemma im_two [X] [Y] (f : X → Y) : forall (x y : X),
    (im f) (two x y) = two (f x) (f y).
Proof.
intros x x'. apply ext. intro y. unfold im,two. split ; intro.
  − destruct H,H,H ; rewrite ← H0, H ; intuition.
  − destruct H. + exists x. auto. + exists x'. auto.
Qed.

(** From here, the reader is encouraged to read Theorem 2.4 of
Klin & Salamanca paper in parallel. Their words are "quoted". *)

(** "Assume, towards a contradiction,
    that there is such a distributive law."
Using the following axioms,
we will be able to derive that True = False *)
Axiom lambda : forall X, P (P X) → P (P X).
```

```
Axiom naturality : forall X Y (f : X → Y) (t : P (P X)),
      (im (im f)) (lambda X t) = lambda Y ((im (im f)) t).
Axiom unit_1 : forall X (s : P X),
    (lambda X) (etaP (P X) s) = im (etaP X) s.
Axiom unit_2 : forall X (s : P X),
    (lambda X) (im (etaP X) s) = etaP (P X) s.

(** "Consider sets:" (in a topos these are finite coproducts) *)
Inductive A := a | b | c |  d.
Inductive U := u | v.

Lemma inter_two (x y z : A) :
   (y = z) → (two x y) ∩ (two x z) = etaP A x.
Proof.
intros. unfold inter,etaP,two. apply ext. intuition. exfalso. apply H.
rewrite ← H0. assumption.
Qed.

(** "And three functions defined by:" *)
Definition f x := match x with a ⇒ u | b ⇒ u | c ⇒ v | d ⇒ v end.
Definition g x := match x with a ⇒ u | b ⇒ v | c ⇒ u | d ⇒ v end.
Definition h x := match x with a ⇒ u | b ⇒ v | c ⇒ v | d ⇒ u end.

(** "Consider the element" *)
Definition t := two (two a b) (two c d).

(** "Analyse how the three naturality squares for [f], [g] and [h]
act on [t]. Recall that [im] acts on functions by taking direct images,
so in particular:" *)
Lemma Pf_ab : (im f) (two a b) = etaP U u.
Proof. rewrite im_two. simpl. apply idempotence. Qed.

Lemma Pf_cd : (im f) (two c d) = etaP U v.
Proof. rewrite im_two. simpl. apply idempotence. Qed.

Lemma Pg_ab : (im g) (two a b) = two u v.
Proof. rewrite im_two. reflexivity. Qed.

Lemma Ph_ab : (im h) (two a b) = two u v.
Proof. rewrite im_two. reflexivity. Qed.

Lemma Pg_cd : (im g) (two c d) = two u v.
Proof. rewrite im_two. reflexivity. Qed.

Lemma Ph_cd : (im h) (two c d) = two u v.
Proof. rewrite im_two. simpl. apply symmetry. Qed.
```

```
(** Additional lemmas needed in the sequel. *)
Lemma preim_g_u : preim g (etaP U u) = two a c.
Proof.
apply ext. unfold preim,two,etaP. intro.
split ; intros ; destruct x ; simpl in * ; intuition ; discriminate.
Qed.

Lemma preim_g_v : preim g (etaP U v) = two b d.
Proof.
apply ext. unfold preim,two,etaP. intro.
split ; intros ; destruct x ; simpl in * ; intuition ; discriminate.
Qed.

Lemma preim_h_u : preim h (etaP U u) = two a d.
Proof.
apply ext. unfold preim,two,etaP. intro.
split ; intros ; destruct x ; simpl in * ; intuition ; discriminate.
Qed.

Lemma preim_h_v : preim h (etaP U v) = two b c.
Proof.
apply ext. unfold preim,two,etaP. intro.
split ; intros ; destruct x ; simpl in * ; intuition ; discriminate.
Qed.

(** "By naturality and idempotence of [two] we get:" *)
Lemma PPg_t : (im (im g)) t = etaP (P U) (two u v).
Proof.
unfold t. rewrite im_two. rewrite Pg_ab,Pg_cd. apply idempotence.
Qed.

Lemma PPh_t : (im (im h)) t = etaP (P U) (two u v).
Proof.
unfold t. rewrite im_two. rewrite Ph_ab,Ph_cd. apply idempotence.
Qed.

(** "Hence, by a unit law for [lambda]
and naturality squares for [g] and [h] we obtain:"*)
Lemma PPg_lambdaA_t :
  (im (im g)) (lambda A t) = two (etaP U u) (etaP U v).
Proof.
rewrite naturality. rewrite PPg_t. rewrite unit_1. apply im_two.
Qed.

Lemma PPh_lambdaA_t :
  (im (im h)) (lambda A t) = two (etaP U u) (etaP U v).
Proof.
```

```
rewrite naturality. rewrite PPh_t. rewrite unit_1. apply im_two.
Qed.

(** "Which implies that [lambda A t] is non-empty:" *)
Lemma lambdaA_t_nonempty : exists (s : P A), lambda A t s.
Proof.
apply (im_nonempty (im g) (lambda A t)). rewrite PPg_lambdaA_t.
exists (etaP U u). unfold two. intuition.
Qed.

(** "and:" *)
Lemma Pg_s :
  forall (s : P A), lambda A t s → (two (etaP U u) (etaP U v)) (im g s).
Proof.
intros. rewrite ← PPg_lambdaA_t. unfold im. exists s. intuition.
Qed.

Lemma Ph_s :
  forall (s : P A), lambda A t s → (two (etaP U u) (etaP U v)) (im h s).
Proof.
intros. rewrite ← PPh_lambdaA_t. unfold im. exists s. intuition.
Qed.

(** "Now (...) applying the same reasoning to four cases we obtain:" *)

Lemma lambdaA_t : forall (s : P A), lambda A t s →
  ((s ⊆ two a c) ∨ (s ⊆ two b d)) ∧ ((s ⊆ two a d) ∨ (s ⊆ two b c)).
Proof.
intros. split.
  − pose proof Pg_s s H.
    destruct H0 ; apply subset_refl in H0 ; apply pres_preim in H0.
    + rewrite preim_g_u in H0. auto.
    + rewrite preim_g_v in H0. auto.
  − pose proof Ph_s s H.
    destruct H0 ; apply subset_refl in H0 ; apply pres_preim in H0.
    + rewrite preim_h_u in H0. auto.
    + rewrite preim_h_v in H0. auto.
Qed.

(** "Distributing intersections over unions and using
the intersection preservation property, we get: "*)
Lemma lambdaA_t_bis : forall (s : P A),
  lambda A t s → (exists (x : A), s ⊆ etaP A x).
Proof.
intros. apply lambdaA_t in H. destruct H,H,H0.
  − exists a. rewrite ← inter_two with (y := c) (z := d) ;
    try discriminate. apply subset_inter ; assumption.
```

```
  - rewrite symmetry in H,H0. exists c.
    rewrite <- inter_two with (y := a) (z := b) ;
    try discriminate. apply subset_inter ; assumption.
  - rewrite symmetry in H,H0. exists d.
    rewrite <- inter_two with (y := a) (z := b) ;
    try discriminate. apply subset_inter ; assumption.
  - exists b.
    rewrite <- inter_two with (y := c) (z := d) ; try discriminate.
    apply subset_inter ; assumption.
Qed.

(** "Now let us come back to the function [f].
By naturality of [two] we get:" *)
Lemma PPf_t : (im (im f)) t = two (etaP U u) (etaP U v).
Proof. unfold t. rewrite im_two. rewrite Pf_ab,Pf_cd. reflexivity. Qed.

(** "Hence, by the naturality square for [f]
and by a unit law for [lambda]:" *)
Lemma PPf_lambdaA_t : im (im f) (lambda A t) = etaP (P U) (two u v).
Proof.
rewrite naturality. rewrite PPf_t. rewrite <- unit_2.
rewrite im_two. reflexivity.
Qed.

(** "This means that:" *)
Lemma Pf_expression : forall (s : P A),
  (lambda A t s) -> (im f) s = two u v.
Proof.
intros. pose proof PPf_lambdaA_t.
assert (im (im f) (lambda A t) (im f s)
<-> etaP (P U) (two u v) (im f s)) by (rewrite H0 ; intuition).
unfold etaP in H1. symmetry. apply H1. exists s. auto.
Qed.

(** "But this contradicts the assumption that [two] is nontrivial." *)
Theorem dlaw_degenerate : True = False.
Proof.
pose proof lambdaA_t_nonempty. destruct H as (s,H).
pose proof Pf_expression s H.
assert (exists x : A, s ⊆ etaP A x) by (apply lambdaA_t_bis; assumption).
destruct H1 as (x,H1).
assert (im f s u) by (rewrite H0 ; unfold two ; intuition).
assert (im f s v) by (rewrite H0 ; unfold two ; intuition).
destruct H2,H3,H2,H3. pose proof H1 x0 H2.
pose proof H1 x1 H3. rewrite H6 in H7. destruct x0,x1 ; discriminate.
Qed.
```

Bibliography

[1] J. Adámek, H. Herrlich, and G. Strecker. *Abstract and concrete categories : the joy of cats*. Wiley, 1990.

[2] J. Adámek, J. Rosickỳ, and E. Vitale. *Algebraic theories: a categorical introduction to general algebra*, volume 184 of *Cambridge Tracts in Mathematics*. Cambridge University Press, 2010.

[3] M. Barr. Relational algebras. In *Reports of the Midwest Category Seminar IV*, pages 39–55. Springer Berlin Heidelberg, 1970.

[4] M. Barr. Exact categories. In *Exact categories and categories of sheaves*, pages 1–120. Springer Berlin Heidelberg, 1971.

[5] M. Barr. *Toposes, triples and theories*, volume 278 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, 1985.

[6] J. Barwise and L. Moss. *Vicious circles: on the mathematics of non-wellfounded phenomena*, volume 60 of *CSLI Lecture Notes*. CSLI, 1996.

[7] J. Beck. Distributive laws. In *Seminar on Triples and Categorical Homology Theory*, pages 119–140. Springer Berlin Heidelberg, 1969.

[8] M. Bertrand and J. Rot. Coalgebraic determinization of alternating automata. arXiv:1804.02546 [cs.LO], 2018.

[9] G. Bezhanishvili, N. Bezhanishvili, and J. Harding. Modal compact Hausdorff spaces. *Journal of Logic and Computation*, 25(1):1–35, 2015.

[10] G. Bezhanishvili, D. Gabelaia, J. Harding, and M. Jibladze. Compact Hausdorff spaces with relations and Gleason spaces. *Applied Categorical Structures*, 27:663–686, 2019.

[11] N. Bezhanishvili, M. Bonsangue, H. H. Hansen, D. Kozen, C. Kupke, P. Panangaden, and A. Silva. Minimisation in logical form. arXiv:2005.11551 [cs.FL], 2020.

[12] N. Bezhanishvili, J. de Groot, and Y. Venema. Coalgebraic geometric logic. In *Proc. CALCO*, volume 139 of *LIPIcs*, pages 7:1–7:18. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.

[13] N. Bezhanishvili, S. Enqvist, and J. de Groot. Duality for instantial neighbourhood logic via coalgebra. In *Proc. CMCS*, pages 32–54. Springer International Publishing, 2020.

[14] N. Bezhanishvili, G. Fontaine, and Y. Venema. Vietoris bisimulations. *Journal of Logic and Computation*, 20(5):1017–1040, 2008.

[15] G. Böhm. The weak theory of monads. *Advances in Mathematics*, 225(1):1–32, 2010.

[16] G. Böhm, S. Lack, and R. Street. On the 2-categories of weak distributive laws. *Communications in Algebra*, 39(12):4567–4583, 2011.

[17] R. Bird and O. de Moor. *Algebra of programming.* Prentice Hall International series in computer science, 1997.

[18] V. Bogachev. *Measure theory*, volume I, II. Springer-Verlag Berlin Heidelberg, 2007.

[19] F. Bonchi, D. Petrişan, D. Pous, and J. Rot. A general account of coinduction up-to. *Acta Informatica*, 54(2):127–190, 2017.

[20] F. Bonchi, D. Petrişan, D. Pous, and J. Rot. Coinduction up to in a fibrational setting. In *Proc. CSL-LICS*, pages 20:1–20:9. ACM, 2014.

[21] F. Bonchi and D. Pous. Checking NFA equivalence with bisimulations up to congruence. In *Proc. POPL*, page 457–468. ACM, 2013.

[22] F. Bonchi and A. Santamaria. Combining semilattices and semimodules. In *Proc. FoSSaCS*, pages 102–123. Springer, 2021.

[23] F. Bonchi, A. Silva, and A. Sokolova. The power of convex algebras. In *Proc. CONCUR*, volume 85 of *LIPIcs*, pages 23:1–23:18. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2017.

[24] F. Bonchi, A. Silva, and A. Sokolova. Distribution bisimilarity via the power of convex algebras. arXiv:1707.02344v5 [cs.LO], 2021.

[25] F. Bonchi, A. Sokolova, and V. Vignudelli. The theory of traces for systems with nondeterminism and probability. In *Proc. LICS*, pages 1–14, 2019.

[26] F. Bonchi, A. Sokolova, and V. Vignudelli. Presenting convex sets of probability distributions by convex semilattices and unique bases. arXiv:2005.01670 [cs.LO], 2020.

[27] F. Bonchi and F. Zanasi. Bialgebraic semantics for logic programming. *Logical Methods in Computer Science*, 11(1), 2016.

[28] M. Bonsangue, H. H. Hansen, A. Kurz, and J. Rot. Presenting distributive laws. In *Proc. CALCO*, pages 95–109. Springer Berlin Heidelberg, 2013.

[29] F. Borceux. *Handbook of categorical algebra*, volume 2. Cambridge University Press, 1994.

[30] F. Borceux. *Handbook of categorical algebra*, volume 3. Cambridge University Press, 1994.

[31] T. Brengos, M. Miculan, and M. Peressotti. Behavioural equivalences for coalgebras with unobservable moves. *Journal of Logical and Algebraic Methods in Programming*, 84(6):826–852, 2015. Special Issue on Open Problems in Concurrency Theory.

[32] A. Carboni, M. Kelly, and R. Wood. A 2-categorical approach to change of base and geometric morphisms I. *Cahiers de topologie et géométrie différentielle catégoriques*, 32(1):47–95, 1991.

[33] A. Chandra, D. Kozen, and L. Stockmeyer. Alternation. *Journal of the ACM*, 28(1):114–133, 1981.

[34] E. Cheng. Iterated distributive laws. *Mathematical Proceedings of the Cambridge Philosophical Society*, 150(3):459–487, 2011.

[35] E. Cheng. Distributive laws for Lawvere theories. *Compositionality*, 2(1), 2020.

[36] M. Clementino, D. Hofmann, and G. Janelidze. The monads of classical algebra are seldom weakly cartesian. *Journal of Homotopy and Related Structures*, 9:175–197, 2014.

[37] S. Crafa and F. Ranzato. A spectrum of behavioral relations over LTSs on probability distributions. In *Proc. CONCUR*, pages 124–139. Springer Berlin Heidelberg, 2011.

[38] F. Dahlqvist and R. Neves. Compositional semantics for new paradigms: probabilistic, hybrid and beyond. arXiv:1804.04145 [cs.LO], 2018.

[39] F. Dahlqvist, L. Parlant, and A. Silva. Layer by layer – combining monads. In *Proc. ICTAC*, pages 153–172. Springer International Publishing, 2018.

[40] V. Danos and T. Ehrhard. Probabilistic coherence spaces as a model of higher-order probabilistic computation. *Information and Computation*, 209(6):966–991, 2011.

[41] S. Dash and S. Staton. A monad for probabilistic point processes. In *Proc. ACT*, volume 333 of *Electronic Proceedings in Theoretical Computer Science*, pages 19–32. Open Publishing Association, 2020.

[42] A. Day. Filter monads, continuous lattices and closure systems. *Canadian Journal of Mathematics*, 27(1):50–59, 1975.

[43] O. de Moor. *Categories, relations and dynamic programming.* PhD thesis, University of Oxford, 1992.

[44] O. de Moor. Categories, relations and dynamic programming. *Mathematical Structures in Computer Science*, 4(1):33–69, 1994.

[45] E. de Vink and J. Rutten. Bisimulation for probabilistic transition systems: a coalgebraic approach. *Theoretical Computer Science*, 221(1):271–293, 1999.

[46] J. Doob. *Measure theory*, volume 143 of *Graduate Texts in Mathematics*. Springer-Verlag, 1994.

[47] D. Edwards. On the existence of probability measures with given marginals. *Annales de l'Institut Fourier*, 28(4):53–78, 1978.

[48] T. Ehrhard, M. Pagani, and C. Tasson. The computational meaning of probabilistic coherence spaces. In *Proc. LICS*, pages 87–96. IEEE, 2011.

[49] T. Ehrhard, M. Pagani, and C. Tasson. Measurable cones and stable, measurable functions: a model for probabilistic higher-order programming. *Proceedings of the ACM on Programming Languages*, 2(POPL), 2018.

[50] J. Farkas. Theorie der einfachen Ungleichungen. *Journal für die reine und angewandte Mathematik*, 124:1–27, 1902.

[51] P. Freyd and A. Scedrov. *Categories, allegories.* North Holland, 1990.

[52] T. Fritz and P. Perrone. Monads, partial evaluations, and rewriting. In *Proc. MFPS*, volume 352 of *Electronic Notes in Theoretical Computer Science*, pages 129–148, 2020.

[53] T. Fritz, P. Perrone, and S. Rezagholi. Probability, valuations, hyperspace: three monads on Top and the support as a morphism. arXiv:1910.03752v2 [math.GN], 2021.

[54] R. Garner. The Vietoris monad and weak distributive laws. *Applied Categorical Structures*, 28(2):339–354, 2020.

[55] N. D. Gautam. The validity of equations of complex algebras. *Archiv für mathematische Logik und Grundlagenforschung*, 3(3):117–124, 1957.

[56] M. Giry. A categorical approach to probability theory. In *Categorical Aspects of Topology and Analysis*, pages 68–85. Springer Berlin Heidelberg, 1982.

[57] J. Goubault-Larrecq. Continuous capacities on continuous state spaces. In *Proc. ICALP*, volume 4596 of *Lecture Notes in Computer Science*, pages 764–776. Springer, 2007.

[58] J. Goubault-Larrecq. Continuous previsions. In *Proc. CSL*, pages 542–557. Springer Berlin Heidelberg, 2007.

[59] J. Goubault-Larrecq. De Groot duality and models of choice: angels, demons and nature. *Mathematical Structures in Computer Science*, 20(2):169–237, 2010.

[60] J. Goubault-Larrecq and K. Keimel. Choquet-Kendall-Matheron theorems for non-Hausdorff spaces. *Mathematical Structures in Computer Science*, 21(3):511–561, 2011.

[61] A. Goy. Powerset weakly distributes over itself in toposes. `https://github.com/Kilgrobil/weak_distributive_laws`. Accessed: 2021-07-12.

[62] A. Goy and D. Petrişan. Combining probabilistic and non-deterministic choice via weak distributive laws. In *Proc. LICS*, page 454–464. ACM, 2020.

[63] A. Goy, D. Petrişan, and M. Aiguier. Powerset-like monads weakly distribute over themselves in toposes and compact Hausdorff spaces. In *Proc. ICALP*, volume 198 of *LIPIcs*. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021.

[64] A. Goy and J. Rot. (In)finite trace equivalence of probabilistic transition systems. In *Proc. CMCS*, pages 100–121. Springer International Publishing, 2018.

[65] J. Gray. The meeting of the Midwest Category Seminar in Zurich, August 24–30, 1970. In *Reports of the Midwest Category Seminar V*, pages 248–255. Springer Berlin Heidelberg, 1971.

[66] R. Guitart. Monades involutives complémentées. *Cahiers de topologie et géométrie différentielle catégoriques*, 16(1):17–101, 1975.

[67] R. Guitart. Calcul des relations inverses. *Cahiers de topologie et géométrie différentielle catégoriques*, 18(1):67–100, 1977.

[68] R. Guitart. Algebraic universes, 1979. Lecture at the Summer School on Universal Algebra and Ordered Sets.

[69] R. Guitart. Qu'est-ce que la logique dans une catégorie? *Cahiers de topologie et géométrie différentielle catégoriques*, 23(2):115–148, 1982.

[70] P. Gumm. Functors for coalgebras. *Algebra universalis*, 45(2):135–147, 2001.

[71] I. Hasuo, B. Jacobs, and A. Sokolova. Generic trace semantics via coinduction. *Logical Methods in Computer Science*, 3(4), 2007.

[72] C. Hermida and B. Jacobs. Structural induction and coinduction in a fibrational setting. *Information and Computation*, 145(2):107–152, 1998.

[73] C. Heunen, O. Kammar, S. Staton, S. Moss, M. Vákár, A. Ścibior, and H. Yang. The semantic structure of quasi-Borel spaces. *PPS Workshop on Probabilistic Programming Semantics*, 2018.

[74] C. Heunen, O. Kammar, S. Staton, and H. Yang. A convenient category for higher-order probability theory. In *Proc. LICS*. IEEE, 2017.

[75] M. Hyland and J. Power. The category theoretic understanding of universal algebra: Lawvere theories and monads. In *Computation, Meaning, and Logic: articles dedicated to Gordon Plotkin*, volume 172 of *Electronic Notes in Theoretical Computer Science*, pages 437–458. Elsevier, 2007.

[76] INRIA. The coq proof assistant. `https://coq.inria.fr/`. Accessed: 2021-07-12.

[77] B. Jacobs. A bialgebraic review of deterministic automata, regular expressions and languages. In *Algebra, Meaning, and Computation: Essays dedicated to Joseph A. Goguen on the Occasion of His 65th Birthday*, pages 375–404. Springer Berlin Heidelberg, 2006.

[78] B. Jacobs. Coalgebraic trace semantics for combined possibilitistic and probabilistic systems. In *Proc. CMCS*, volume 203 of *Electronic Notes in Theoretical Computer Science*, pages 131–152, 2008.

[79] B. Jacobs. *Introduction to coalgebra: towards mathematics of states and observation.* Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 2016.

[80] B. Jacobs. From multisets over distributions to distributions over multisets. In *Proc. LICS*, 2021.

[81] B. Jacobs, A. Silva, and A. Sokolova. Trace semantics via determinization. *Journal of Computer and System Sciences*, 81(5):859–879, 2015. Selected Papers of CMCS'12.

[82] P. Johnstone. *Stone spaces*, volume 3 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, 1982.

[83] P. Johnstone. *Sketches of an elephant: a topos theory compendium.* Oxford University Press, 2002.

[84] C. Jones and G. Plotkin. A probabilistic powerdomain of evaluations. In *Proc. LICS*, pages 186–195. IEEE, 1989.

[85] B. Jónsson and A. Tarski. Boolean algebras with operators, part I. *American Journal of Mathematics*, 73(4):891–939, 1951.

[86] A. Jung and R. Tix. The troublesome probabilistic powerdomain. In *Proc. Comprox III*, volume 13 of *Electronic Notes in Theoretical Computer Science*, pages 70–91. Elsevier, 1998.

[87] K. Keimel. The monad of probability measures over compact ordered spaces and its Eilenberg–moore algebras. *Topology and its Applications*, 156(2):227–239, 2008.

[88] K. Keimel and G. Plotkin. Mixed powerdomains for probability and nondeterminism. *Logical Methods in Computer Science*, 13(1), 2017.

[89] B. Klin and J. Rot. Coalgebraic trace semantics via forgetful logics. In *Proc. FoSSaCS*, pages 151–166. Springer Berlin Heidelberg, 2015.

[90] B. Klin and J. Rot. Coalgebraic trace semantics via forgetful logics. *Logical Methods in Computer Science*, 12(4), 2016.

[91] B. Klin and J. Salamanca. Iterated covariant powerset is not a monad. In *Proc. MFPS*, volume 341 of *Electronic Notes in Theoretical Computer Science*, pages 261–276, 2018.

[92] C. Kupke, A. Kurz, and Y. Venema. Stone coalgebras. *Theoretical Computer Science*, 327(1):109–134, 2004. Selected Papers of CMCS'03.

[93] A. Kurz and J. Velebil. Relation lifting, a survey. *Journal of Logical and Algebraic Methods in Programming*, 85(4):475–499, 2016.

[94] S. Lack and R. Street. The formal theory of monads II. *Journal of Pure and Applied Algebra*, 175(1):243–265, 2002. Special Volume celebrating the 70th birthday of Professor Max Kelly.

[95] J. Lambek and P. Scott. *Introduction to higher order categorical logic.* Cambridge University Press, 1986.

[96] S. Mac Lane. *Categories for the working mathematician*, volume 5 of *Graduate Texts in Mathematics.* Springer-Verlag, 1971.

[97] S. Mac Lane and I. Moerdijk. *Sheaves in geometry and logic: a first introduction to topos theory.* Universitext. Springer New York, 1992.

[98] E. Manes. A triple theoretic construction of compact algebras. In *Seminar on Triples and Categorical Homology Theory*, pages 91–118. Springer Berlin Heidelberg, 1969.

[99] E. Manes and P. Mulry. Monad compositions I: general constructions and recursive distributive laws. *Theory and Applications of Categories*, 18(7):172–208, 2007.

[100] E. Manes and P. Mulry. Monad compositions II: Kleisli strength. *Mathematical Structures in Computer Science*, 18(3):613–643, 2008.

[101] E. Manes and P. Mulry. Near distributive laws. In *Proc. MFPS*, volume 341 of *Electronic Notes in Theoretical Computer Science*, pages 277–295. Elsevier, 2018.

[102] V. Marra and L. Reggio. A characterisation of the category of compact Hausdorff spaces. *Theory and Applications of Categories*, 35(51):1871–1906, 2020.

[103] D. Marsden. Category theory using string diagrams. arXiv:1401.7220v2 [math.CT], 2014.

[104] C. McLarty. *Elementary categories, elementary toposes*, volume 21 of *Oxford Logic Guides*. Clarendon Press, 1992.

[105] M. Mio. Upper-expectation bisimilarity and łukasiewicz μ-calculus. In *Proc. FoSSaCS*, pages 335–350. Springer Berlin Heidelberg, 2014.

[106] M. Mio, R. Sarkis, and V. Vignudelli. Combining nondeterminism, probability, and termination: equational and metric reasoning. In *Proc. LICS*, 2021.

[107] M. Mio and V. Vignudelli. Monads and quantitative equational theories for nondeterminism and probability. In *Proc. CONCUR*, volume 171 of *LIPIcs*, pages 28:1–28:18. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020.

[108] M. Mislove. Nondeterminism and probabilistic choice: obeying the laws. In *Proc. CONCUR*, pages 350–365. Springer Berlin Heidelberg, 2000.

[109] E. Moggi. Notions of computation and monads. *Information and Computation*, 93(1):55–92, 1991. Selected papers of LICS'89.

[110] L. Moss. Coalgebraic logic. *Annals of Pure and Applied Logic*, 96(1):277–317, 1999.

[111] G. Osius. Logical and set theoretical tools in elementary topoi. In *Model Theory and Topoi*, pages 297–346. Springer Berlin Heidelberg, 1975.

[112] L. Parlant. *Monad composition via preservation of algebras*. PhD thesis, University College London, 2020.

[113] P. Perrone. *Categorical probability and stochastic dominance in metric spaces*. PhD thesis, University of Leipzig, 2018.

[114] P. Perrone. Personal communication, 2021.

[115] D. Petrişan. *Investigations into algebra and topology over nominal sets.* PhD thesis, University of Leicester, 2012.

[116] D. Petrişan and R. Sarkis. Semialgebras and weak distributive laws. arXiv:2106.13489 [cs.LO], 2021.

[117] A. Pitts. *Nominal sets: names and symmetry in computer science*, volume 57 of *Cambridge Tracts in Theoretical Computer Science.* Cambridge University Press, 2013.

[118] G. Plotkin. A powerdomain construction. *SIAM Journal on Computing*, 5(3):452–487, 1976.

[119] G. Plotkin and J. Power. Notions of computation determine monads. In *Proc. FoSSaCS*, pages 342–356. Springer Berlin Heidelberg, 2002.

[120] G. Plotkin and J. Power. Computational effects and operations: an overview. In *Proc. Workshop Domains 2002*, volume 73 of *Electronic Notes Theoretical Computer Science*, page 149–163, 2004.

[121] H. Poincaré. *Science et méthode.* Bibliothèque de philosophie scientifique. Flammarion, 1908.

[122] D. Pous. Complete lattices and up-to techniques. In *Proc. APLAS*, pages 351–366. Springer Berlin Heidelberg, 2007.

[123] D. Pous. *Techniques modulo pour les bisimulations. (Up to techniques for bisimulations).* PhD thesis, École Normale Supérieure de Lyon, France, 2008.

[124] D. Pous and D. Sangiorgi. Enhancements of the bisimulation proof method. In *Advanced Topics in Bisimulation and Coinduction*, volume 52 of *Cambridge Tracts in Theoretical Computer Science*, page 233–289. Cambridge University Press, 2011.

[125] J. Power and H. Watanabe. Combining a monad and a comonad. *Theoretical Computer Science*, 280(1):137–162, 2002.

[126] A. Romanowska. On bisemilattices with one distributive law. *Algebra universalis*, 10(1):36–47, 1980.

[127] J. Rot. *Enhanced coinduction.* PhD thesis, Leiden University, 2015.

[128] J. Rot, F. Bonchi, M. Bonsangue, D. Pous, J. Rutten, and A. Silva. Enhanced coalgebraic bisimulation. *Mathematical Structures in Computer Science*, 27(7):1236–1264, 2017.

[129] J. Rutten. Automata and coinduction (an exercise in coalgebra). In *Proc. CONCUR*, pages 194–218. Springer Berlin Heidelberg, 1998.

[130] J. Rutten. Relators and metric bisimulations. In *Proc. CMCS*, volume 11 of *Electronic Notes in Theoretical Computer Science*, pages 252–258. Elsevier, 1998.

[131] J. Rutten. Universal coalgebra: a theory of systems. *Theoretical Computer Science*, 249(1):3–80, 2000.

[132] J. Rutten. *The method of coalgebra: exercises in coinduction.* CWI Amsterdam, 2019.

[133] M. Sabok, S. Staton, D. Stein, and M. Wolman. Probabilistic programming semantics for name generation. *Proc. of the ACM on Programming Languages*, 5(POPL):1–29, 2021.

[134] N. Saheb-Djahromi. Cpo's of measures for nondeterminism. *Theoretical Computer Science*, 12(1):19–37, 1980.

[135] J. Salamanca Téllez. Lattices do not distribute over powerset. *Algebra Universalis*, 81:49, 2020.

[136] L. Schröder. Categories: a free tour. In *Categorical Perspectives*, pages 1–27. Birkhäuser Boston, 2001.

[137] C. Schubert. *Lax algebras: a scenic approach.* PhD thesis, Universität Bremen, 2006.

[138] A. Silva, F. Bonchi, M. Bonsangue, and J. Rutten. Generalizing the powerset construction, coalgebraically. In *Proc. FSTTCS*, volume 8 of *LIPIcs*, pages 272–283. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2010.

[139] A. Silva, F. Bonchi, M. Bonsangue, and J. Rutten. Generalizing determinization from automata to coalgebras. *Logical Methods in Computer Science*, 9(1), 2013.

[140] A. Sokolova. *Coalgebraic analysis of probabilistic systems.* PhD thesis, Technische Universiteit Eindhoven, 2005.

[141] A. Sokolova. Probabilistic systems coalgebraically: a survey. *Theoretical Computer Science*, 412(38):5095–5110, 2011. CMCS Tenth Anniversary Meeting.

[142] A. Sokolova and H. Woracek. Congruences of convex algebras. *Journal of Pure and Applied Algebra*, 219(8):3110–3148, 2015.

[143] S. Staton. Relating coalgebraic notions of bisimulation. *Logical Methods in Computer Science*, 7(1), 2011.

[144] S. Staton, H. Yang, F. Wood, C. Heunen, and O. Kammar. Semantics for probabilistic programming: higher-order functions, continuous distributions, and soft constraints. In *Proc. LICS*, page 525–534. ACM, 2016.

[145] R. Street. The formal theory of monads. *Journal of Pure and Applied Algebra*, 2(2):149–168, 1972.

[146] R. Street. Weak distributive laws. *Theory and Applications of Categories*, 22(12):313–320, 2009.

[147] T. Świrszcz. Monadic functors and convexity. *Bulletin de l'Académie Polonaise des Sciences. Série des sciences math., astr. et phys*, 22(1):39–42, 1974.

[148] R. Tix, K. Keimel, and G. Plotkin. Semantic domains for combining probability and non-determinism. *Electronic Notes in Theoretical Computer Science*, 222:3–99, 2009.

[149] D. Turi. *Functorial operational semantics and its denotational dual.* PhD thesis, Vrije Universiteit, Amsterdam, 1996.

[150] G. van Heerdt, J. Hsu, J. Ouaknine, and A. Silva. Convex language semantics for nondeterministic probabilistic automata. In *Proc. ICTAC*, pages 472–492. Springer International Publishing, 2018.

[151] D. Varacca. The powerdomain of indexed valuations. In *Proc. LICS*, pages 299–308. IEEE, 2002.

[152] D. Varacca. *Probability, nondeterminism and concurrency: two denotational models for probabilistic computation.* PhD thesis, University of Aarhus, 2003.

[153] D. Varacca and G. Winskel. Distributing probability over non-determinism. *Mathematical Structures in Computer Science*, 16(1):87–113, 2006.

[154] Y. Venema and J. Vosmaer. Modal logic and the Vietoris functor. In *Leo Esakia on Duality in Modal and Intuitionistic Logics*, volume 4 of *Outstanding Contributions to Logic*, pages 119–153. Springer Netherlands, 2014.

[155] L. Vietoris. Bereiche zweiter Ordnung. *Monatshefte für Mathematik und Physik*, 32:258–280, 1922.

[156] S. Vigna. A guided tour in the topos of graphs. arXiv:math/0306394 [math.CT], 2003.

[157] P. Wadler. Monads for functional programming. In *Advanced Functional Programming, Proc. Båstad Spring School*, volume 925 of *Lecture Notes in Computer Science*, pages 24–52. Springer, 1995.

[158] S. Willard. *General topology.* Addison-Wesley, 1970.

[159] O. Wyler. Algebraic theories of continuous lattices. In *Continuous Lattices*, pages 390–413. Springer Berlin Heidelberg, 1981.

[160] S. Zetzsche, G. van Heerdt, A. Silva, and M. Sammartino. Canonical automata via distributive law homomorphisms. arXiv:2104.13421 [cs.FL], 2021.

[161] M. Zwart. *On the non-compositionality of monads via distributive laws.* PhD thesis, University of Oxford, 2020.

[162] M. Zwart and D. Marsden. No-go theorems for distributive laws. In *Proc. LICS*, pages 1–13. IEEE, 2019.

Notation

General Notation

2	Booleans $\{0,1\}$
$\mathbb{N}$	Natural numbers $\{0,1,2,\dots,\}$
$\mathbb{N}^*$	Positive natural numbers $\{1,2,\dots\}$
$\mathbb{R}$	Real numbers
$[0,1]$, $(0,1)$	Unit interval (closed, open)
$\mathcal{M}_{p,q}(\mathbb{R})$	Real matrices of size $p \times q$
I_m	Identity matrix of size m
$\_^T$	Matrix transposition
$\varphi = \sum_{x \in X} \varphi_x \cdot x$	Distribution
$\mathsf{supp}(\varphi)$	Support of a distribution
$[x_1, \dots, x_n]$	List
$[\![x_1, \dots, x_n]\!]$	Multiset
$\lambda x.f(x)$, $x \mapsto f(x)$	Function f
U^c	Complement
$\mathsf{Im}(f)$	Direct image of the domain
$f(U)$	Direct image
$f^{-1}(U)$	Preimage
$R[U]$	Relational image
$R^{-1}[U]$	Relational preimage
up_x	Upclosure
conv_x	Convex closure

Category Theory

C, D	Category
Set	Sets and functions
Rel	Sets and relations
cJSL	Complete join-semilattices
Mon	Monoids
CMon	Commutative monoids
Pos	Posets
X, Y	Object
f, g	Morphism

$\circ$	Composition
id, 1	Identity
$\rightarrow$	Morphism typing
$\hookrightarrow$	Monomorphism, subobject
$\twoheadrightarrow$	Epimorphism
F, G	Functor
α, β	Natural transformation
$\beta \circ \alpha$	Vertical composition
$\beta\alpha$	Horizontal composition
$\cong$	Isomorphism
$\equiv$	Equivalence
0	Initial object
1	Terminal object
$X \times Y$	Product
$\langle f, g \rangle$	Pairing
$f \times g$	Product morphism
$X + Y$	Coproduct
inl	Left injection
inr	Right injection
$\lrcorner$	Pullback
$\dot{\lrcorner}$	Weak pullback
$\mathbf{S}$, $\mathbf{T}$	Monad
$\mathbf{Id}$	Identity monad
γ	Monad morphism
$L \dashv R$	Adjunction
η	Unit
ϵ	Counit
μ	Multiplication
$\mathsf{Kl}(\mathbf{T})$	Kleisli category
$\nrightarrow$	Kleisli morphism typing
$\bullet$	Kleisli composition
$F_{\mathbf{T}}$	Free Kleisli functor
$U_{\mathbf{T}}$	Forgetful Kleisli functor
$\mathsf{EM}(\mathbf{T})$	Eilenberg-Moore category
(X, x)	Algebra
$F^{\mathbf{T}}$	Free Eilenberg-Moore functor
$U^{\mathbf{T}}$	Forgetful Eilenberg-Moore functor

Distributive Laws

$\lambda : \mathbf{TS} \rightarrow \mathbf{ST}$	Distributive law (plain, weak, coweak)
$\overline{\mathbf{S}}$	Lifting (plain, weak, coweak)
$\underline{\mathbf{T}}$	Extension (plain, weak, coweak)
π, ι	Naturals of weak lifting or coweak extension
$\mathbf{S} \circ \mathbf{T}$	Composite monad

$\mathbf{S} ◒ \mathbf{T}$	Weak composite monad
$\mathbf{S} ◓ \mathbf{T}$	Coweak composite monad
$\mathsf{Alg}(\lambda)$	Category of λ-algebras
$\mathsf{FAlg}(\lambda)$	Category of free λ-algebras

Presentations

(Σ, E)	Equational theory
$\mathsf{Alg}(\Sigma, E)$	Category of (Σ, E)-algebras
$\vee$	Binary join
$\bigvee$	Join
$\oplus_r$	Convex sum

Set Monads

$(- + \mathbf{1})$	Maybe monad
$\mathbf{P}$	Powerset monad
$\mathbf{P}_f$	Finite powerset monad
$\mathbf{P}^*$	Non-empty powerset monad
$\mathbf{P}_f^*$	Finite non-empty powerset monad
$\mathbf{D}$	Distribution monad
$\mathbf{D}_\omega$	Countable distribution monad
$\mathbf{A}$	Abelian group monad
$\mathbf{L}$	List monad
$\mathbf{M}$	Multiset monad
$\mathbf{R}$	Reader monad
$\mathbf{F}$	Filter monad
$\boldsymbol{\beta}$	Ultrafilter monad
$\mathbf{P} ◒ \mathbf{P}$	Monad of upclosed sets of subsets
$\mathbf{P} ◒ \mathbf{D}$	Monad of convex sets of distributions

Coalgebras

A	Alphabet
$\mathsf{Coalg}(F)$	Category of coalgebras
U^F	Forgetful functor on coalgebras
(X, c)	Coalgebra
c_*	Output function
c_a	Transition function w.r.t letter a
$c^\#$	Algebraic expansion (= determinisation)
$c^\dagger$	Intermediate expansion
$F ◒ T$	Weak composite functor
$[\![-]\!]$	Alternating automaton semantics
$b(R)$	Relation progression
$\sim$	Bisimilarity

cont_x	Contextual closure
congr_x, $\equiv$	Congruence closure

Regular Categories

$\mathsf{Rel}(\mathsf{C})$	Category of objects and relations
$\mathsf{Rel}(F)$	Relational extension of a functor
$\mathsf{Rel}(\alpha)$	Relational extension of a natural
$r = \langle r_1, r_2 \rangle$	Relation
$\rightsquigarrow$	Relation typing
$\cdot$	Relational composition
$\mathcal{G}$	Graph functor
$(-)^\circ$	Transpose functor

Toposes

ev	Exponential evaluation
X^Y	Exponential object
$!_X$	Terminal morphism
$\mathsf{true} : 1 \hookrightarrow \Omega$	Subobject classifier
χ_m	Characteristic morphism
FinSet	Finite sets and functions
Graph	Graphs and graph homomorphisms
$\mathsf{Set}^{\mathsf{C}^{\mathsf{op}}}$	Presheaf topos
Nom	Nominal sets and equivariant functions
$\wedge$	Conjunction morphism
$=_X$	Equality morphism
$\{-\}_X$	Singleton morphism
$\in_X$	Membership morphism
$[t]$, $[\varphi]$	Mitchell-Bénabou interpretation (term, formula)
$x : X \vdash \varphi(x)$	Valid formula
$\mathbf{\exists}$	Powerset monad in a topos
$\checkmark$	Result proved in Coq

Topology, Measures

KHaus	Category of compact Hausdorff spaces
Stone	Category of Stone spaces
τ_X	Topology
U, W	Open subset
C, D	Closed subset
2^ω	Cantor set
$\Box$, $\Diamond$	Vietoris modalities
$\mathbf{V}$	Vietoris monad
$\overline{W}$	Topological closure

Σ_X	Sigma-algebra
$\sigma(\tau_X)$	Borel σ-algebra
B	Borel subset
m	Measure
$m \circ f^{-1}$	Pushforward measure
$\int_X f \mathrm{d}m$	Integral w.r.t. m
$C(X)$	Continuous functions $X \to \mathbb{R}$
ev_u	Integral evaluation
$\mathbf{R}$	Radon monad
$\mathsf{supp}(m)$	Support of a Radon measure

Index

ÉCOLE DOCTORALE
Interfaces:
matériaux, systèmes, usages

Titre : Compositionnalité des monades par lois de distributivité faibles

Mots clés : monade, coalgèbre, théorie des catégories, bisimulation, topos, espace compact Hausdorff

Résumé : Les monades sont un concept de théorie des catégories qui permet de modéliser de façon abstraite la notion d'effet computationnel. La non-compositionnalité des monades est bien connue, mais la théorie des lois de distributivité est un outil classique qui s'est révélé utile pour combiner les effets de plusieurs monades. Dans de nombreux cas, il est impossible de définir une loi de distributivité entre une paire de monades spécifiques. Quand il semble qu'il en existe presque une, il est possible d'utiliser une forme plus faible de loi de distributivité.

Cette thèse étudie les propriétés théoriques des lois de distributivité faibles, introduit une notion duale appelée lois de distributivité cofaibles, et fournit des applications à la théorie des coalgèbres : déterminisation généralisée et techniques up-to pour les bisimulations, avec des exemples pour les automates alternants et les automates probabilistes.

On étudie également des lois de distributivité faibles précises. On calcule l'unique loi de distributivité faible monotone entre la monade des sous-ensembles et la monade des distributions, ce qui permet de combiner le choix probabiliste et le choix non-déterministe de façon canonique. Bien qu'il soit connu que la monade des sous-ensembles se distribue faiblement sur elle-même, ce résultat est généralisé à des topos arbitraires et aux espaces compacts Hausdorff, où le rôle des sous-ensembles est joué par la monade de Vietoris.

Title: On the compositionality of monads via weak distributive laws

Keywords: monad, coalgebra, category theory, bisimulation, topos, compact Hausdorff space

Abstract: Monads are a concept from category theory allowing to model abstractly the notion of computational effect. The non-compositionality of monads is well-known, but the theory of distributive laws is a classical tool that has proved useful to combine effects of several monads. In frequent cases, there is no way of defining a distributive law between a pair of specific monads. When it feels like there almost exists one, a weaker form of distributive law can be used.

This thesis studies theoretical properties of weak distributive laws, introduces a dual notion called coweak distributive laws, and provides applications to coalgebra theory: generalized determinization and up-to techniques for bisimulations, with examples for alternating automata and probabilistic automata.

Some specific weak distributive laws are also studied. The unique monotone weak distributive law between the powerset monad and the distribution monad is derived, allowing to combine probabilistic choice and non-deterministic choice in a canonical way. Although it is known that the powerset monad weakly distributes over itself, this result is generalised to arbitrary toposes and to compact Hausdorff spaces, where the role of the powerset is played by the Vietoris monad.

Maison du doctorat de l'Université Paris-Saclay
2ème étage aile ouest, Ecole normale supérieure Paris-Saclay
4 avenue des Sciences,
91190 Gif sur Yvette, France