Program semantics as Kleisli representations

Fredrik Dahlqvist
University College London
f.dahlqvist@ucl.ac.uk

Renato Neves
INESC TEC (HASLab) & University of Minho
nevrenato@di.uminho.pt

January 25, 2017

Abstract

We introduce a framework for the development of program semantics based on ideas from *representation theory.* Assuming programs form a monoid under sequential composition, they can be *represented* as endomorphisms in the Kleisli category of a monad T. We call such representations *Kleisli representations.* We consider how other program constructs can be added, and equations between them enforced, with the ultimate goal of studying Kleisli representations of Kleene algebras.

We establish several no-go theorems for the Kleisli representability of (fragments of) Kleene algebras for many monads, including the hybrid monad, used in the development of hybrid systems, and the distribution monad which has an analogous role for probabilistic programs. We then suggest strategies around these no-go theorems to establish the representability of (fragments of) Kleene algebras in systems with hybrid or probabilistic features. In the process we establish several new classification results about natural transformations of the type $1 \to T$ and $T \times T \to T$, present several new results about the hybrid and Giry monads, and solve a long-standing open problem about the combination of probabilities with non-determinism.

1 Introduction

1.1 Semantics as representations

Our starting premise is rather conventional: programs are algebraic objects interpreted as actions on a state space. Moreover, these actions are coalgebraic in nature, in the sense that transitions are 'typed' by an endofunctor on the category which the state space inhabits. A transition given by a program instruction can for example be a time-dependent trajectory (as in hybrid systems) or a probability distribution (as in probabilistic systems). Simply put, the overarching aim of this work is to present a principled, simple and flexible formalism which reconciles the algebraic aspect of program instructions with the coalgebraic aspect of their interpretation.

In order to have something concrete to build on we require the following: (1) programs must form a monoid under sequential composition and (2) programs are to be interpreted as coalgebras $X \to TX$ for a monad T. These are mild requirements, and yet they have powerful conceptual consequences. For every state space object X, the monadic data endows the set $\mathrm{End}_{\mathbf{Kl}(T)}(X)$ of endomorphisms of X in the Kleisli category $\mathbf{Kl}(T)$ with the structure of a monoid. This means that we can interpret not just the programs *but also their algebraic* *structure* by defining an interpretation as a *monoid morphism* to $\text{End}_{\mathbf{Kl}(T)}(X)$. The idea of representing an algebraic structure as a collection of endomorphisms is well-known to mathematicians and physicists alike; in fact, it is the central concept of *representation theory* which has been developed extensively in the case of groups and Lie algebras represented as endomorphisms of vector spaces, with important applications ranging from the classification of finite groups to quantum field theory.

The central idea of this work is thus *to view program semantics as representations in a Kleisli category*, or for short as *Kleisli representations.* Note that monoids are the simplest structure for which the concept of representation makes categorical sense, since endomorphisms always form (at least) a monoid. This justifies our starting point. We believe that the Kleisli representation perspective unifies the algebraic and coalgebraic aspects into a simple conceptual entity.

1.2 Defining representations from natural transformations

Programs can usually be assumed to form a more complex algebraic structure than that of a monoid. In this paper we will work our way towards programs as elements of *Kleene algebras* since these are well-studied and interesting structures in theoretical computer science. But what will follow is also intended as a blueprint to study Kleisli representations of arbitrary algebraic structures.

We will interpret the 'non-deterministic choice' operation $+$ via *natural transformations* of the type $T \times T \to T$. This is detailed in Sections 2 and 4, but let us say a few words about naturality. The representation of sequential composition hinges on the multiplication of the monad, which is of course natural. It follows that Kleisli representations of monoids form a category, and in particular that quotient representations and sub-representations are meaningful objects of discourse. This echoes the notion of homomorphism between linear representations of a group (or a Lie algebra), which is key, for example, to defining the fundamental concept of irreducible representation. Similarly, in order for representations of $+$ to form a category, it is necessary that the transformation $T \times T \to T$ interpreting it be natural. Not only does this provide us with natural objects like sub-representations, but from a theoretical computer science perspective, it also allows us to change the 'resolution' of our view of the state space, ignoring certain details if they are not important by forming a coarse-grained quotient of the state space which is compatible with the semantics. In particular it allows, in principle, for the *approximation of semantics.*

1.3 Running examples

In this study about Kleisli representations we will consider several monads. However, in order to keep the paper concise, we decided to provide a special focus to monads closely related to hybrid programs [Höf09, Pla10], and probabilistic ones. Perhaps the notion of hybrid program deserves some words of explanation, as it is less known that its probabilistic counterpart. Briefly put, hybrid programs correspond to transition systems with two transition types: one ('discrete') is used to capture a classic computational operation, the assignment, while the other ('continuous') encodes the evolution of some physical process, like movement or time. Together these transition types make it possible to specify the interaction between a computational device and a physical process, examples ranging from simple thermostats, to complex cruise control systems, aviation systems, and power grids. Recently, a monad to model this sort of behaviour was proposed by the second author *et al.* in [NBHM16]. Intuitively, its Kleisli composition allows hybrid programs to pass the control of a physical process between themselves. It was also shown in [NB16] that its underlying functor plays an important role in providing a semantics for different variants of hybrid automata – the main formalism for hybrid systems – in a uniform way.

1.4 Main contributions

We have just sketched the conceptual contribution of this work, our main technical contributions are as follows:

1. We present several new results classifying natural transformations of the type $\underline{1} \to T$ and $T \times T \to T$ for monads on **Set**, topological spaces, or Polish spaces.

2. In particular we show that:

   - there exist exactly 12 natural transformations $\mathsf{M} \times \mathsf{M} \to \mathsf{M}$, where M is the maybe monad,
   - there exist exactly 16 natural transformations $\mathsf{P} \times \mathsf{P} \to \mathsf{P}$, where P is the powerset monad,

   and provide:

   - a complete classification of all such transformations for the list monad L, the hybrid monad H in **Set** and the multiset monad $\mathsf{B}_{\mathbb{N}}$
   - a complete classification of all such transformations for the Giry monad $\mathsf{G} : \mathbf{Pol} \to \mathbf{Pol}$, and the related 'subdistribution monad' GM. These results exploit the 'Machine' developed in [DDG16a].

3. Based on these results we give several no-go theorems for the Kleisli representability of (fragments of) Kleene algebras for well-known monads.

4. To circumvent no-go theorems we consider combinations of monads. We solve a long-standing open problem by showing that there is no monad structure on PD, where D is the distribution monad ([KG17]). We also show that there exists a distributive law of the non-empty powerset monad over the hybrid monad in **Set**, and of an arbitrary monad T in **Set** over M.

5. We show that these combinations of monad are better-behaved, but that representing the full structure of Kleene algebras remains problematic and probably not desirable.

We assume that the reader has basic knowledge of category theory. The crucial definition of a monad and related notions are briefly detailed in the appendix. The proofs for several theorems of this paper are also available in the appendix.

1.5 Related work

Combining the algebraic aspect of programs with their coalgebraic interpretation in a single mathematical entity is famously exemplified by the work of Turi and Plotkin in [TP97]. The main difference with the work presented here is that [TP97] focuses on distributive laws of monads – which capture the algebraic aspect of programs – over comonads – which capture the semantics of programs – whilst here we have two monads: one which will remain implicit and defines the algebraic structure of programs, and the other which is explicit and defines their behaviour. The bridge between the two is not a distributive law, but additional structure on the latter to interpret the program constructs of the former.

Representations generalise the concept of monoid *action*, which is also combined with the notion of coalgebras in Jacobs' work [Jac00] on hybrid systems. The central concept of [Jac00] is that of a coalgebra together with a monoid action on its carrier, which differs from our notion of Kleisli representation which can be reformulated as a collection of coalgebras 'algebraically indexed' by a monoid.

Kozen's classic probabilistic semantics [Koz85] (or at least its program reduct) is a sophisticated example of Kleisli representation (it interprets a monoid with a $(-)^*$ operation and a cone structure over $\mathbb{Q}$). The probabilistic semantics suggested in [Dob12] does not form a Kleisli representation, since the interpretation of non-deterministic choice and $(-)^*$ is not of the right type, let alone natural. We hope that the results provided in this work can clarify what can or cannot be expected of program interpretations in the Kleisli category of the Giry monad.

To our knowledge, the recent work by Hansen and Kupke [HKL14, HK15] is perhaps the closest in spirit to the present paper since it also seeks to interpret algebraic program constructs via natural transformations, although its focus is directed at developing a coalgebraic version of PDL. Similarly, the work of Plotkin and Power [PP01] on algebraic operations also consider an interpretation of certain program operations by natural transformations in a way that is not dissimilar to ours. Finally, let us mention recent work on ProbNetKAT (see e.g. [FKM+16, SKF+17]) which can also be understood as developing Kleisli representations of Kleene-like algebras for the Giry monad.

2 Kleisli representations

Recall that the quasi-variety of Kleene Algebras (KAs) is defined by a signature consisting of two binary operations ; and +, a unary operation $(-)^*$ and two constants 0 and 1 satisfying Kozen's axioms

K1. $a;(b;c) = (a;b);c$

K2. $1;a = a;1 = a$

K3. $a+(b+c) = (a+b)+c$

K4. $a + b = b + a$

K5. $a + 0 = a$

K6. $a + a = a$

K7. $0;a = 0$

K8. $a;0 = 0$

K9. $a;(b+c) = (a;b)+(a;c)$

K10. $(a+b);c = a;c+b;c$

K11. $1 + (a;a^*) \leq a^*$

K12. $1 + (a^*;a) \leq a^*$

K13. $a;x \leq x \to a^*;x \leq x$

K14. $x;a \leq x \to x;a^* \leq x$

The axioms K1, K2 simply say that a Kleene algebra has a monoidal reduct.

2.1 Kleisli representations for monoids

Given a monoid $(M, \cdot, 1)$, a monad $T : \mathbf{C} \to \mathbf{C}$ and a $\mathbf{C}$-object X, we define a *Kleisli T-representation of M in X*, or simply a *Kleisli representation of M in X* if there is no ambiguity, as a monoid homomorphism

$$\rho : (M, \cdot, 1) \to (\mathrm{End}_{\mathbf{Kl}(T)}(X), \circ_T, \eta^T_X)$$

More generally, for any variety $\mathcal{V}$ extending monoids, we define a *Kleisli representation* of a $\mathcal{V}$-object A as a $\mathcal{V}$-homomorphism $\rho : A \to \mathrm{End}_{\mathbf{Kl}(T)}(X)$ which is a Kleisli representation in the sense defined above for the monoid reduct of A. Clearly such representations may in general not exist, and the purpose of what follows is to find interesting cases where $\mathrm{End}_{\mathbf{Kl}(T)}(X)$ has a structure lying between monoids and Kleene algebras.

Examples

1. A Kleisli $\mathsf{Id}_{\mathbf{Set}}$-representation of a monoid M is simply an action of M on a set. Similarly, a continuous group action can be seen as Kleisli $\mathsf{Id}_{\mathbf{Top}}$-representation of a group object in **Top**.

2. Let **Pol** be the category of Polish spaces and G be the Giry monad over **Pol** ([Gir82]). A Kleisli representation of the monoid $(\mathbb{R}^+, +, 0)$ is simply a stochastic process.

3. Let $\mathsf{F} : \mathbf{Set} \to \mathbf{Vect}$ be the functor building free vector spaces over some chosen field, and let $\mathsf{U} : \mathbf{Vect} \to \mathbf{Set}$ be the corresponding forgetful functor. The composition $\mathsf{UF} : \mathbf{Set} \to \mathbf{Set}$ is a monad, and for any group G, a Kleisli UF-representation of G on a finite set n is simply the usual notion of linear representation of G on the n-dimensional vector space.

4. Any coalgebra $\gamma : X \to TX$ for a monad T is canonically associated with a Kleisli representation of the monoid $\mathbb{N}$ via the map $0 \mapsto \eta^T_X$ and $n \mapsto \gamma^n$, where the composition is of course the Kleisli composition. This is the dynamics which is tacitly understood in most coalgebraic models of computation.

5. Let $\mathsf{F} : \mathbf{Set} \to \mathbf{Mon}$ be the functor that builds a free monoid, let A be a set of actions and consider a function $f : A \to \mathrm{End}_{\mathbf{Kl}(T)}$. This data defines a semantics for composing actions as a Kleisli representation

$$\overline{f} : \mathsf{F}A \to \left(\mathrm{End}_{\mathbf{Kl}(T)}(X), \circ_T, \eta^T_X\right)$$

   In the area of dynamic logics, this is the standard technique that is (implicitly) used to provide a semantics for the composition of actions.

Remark 1. *1. We can view a Kleisli representation of a monoid or of a group as a functor from the corresponding one object category to* $\mathbf{Kl}(T)$*. It is particularly clear from this perspective that representations form a category since representation morphisms are then simply natural transformations. This category is of fundamental importance in (linear) representation theory, see [JS91].*

2. Endomorphisms $X \to TX$ in $\mathbf{Kl}(T)$ are in one-to-one correspondence with endomorphisms $TX \to TX$ in the (Eilenberg-Moore) category of T-algebras, where TX is the free algebra over X. Thus we can interchangeably view the state space as the carrier of a coalgebraic structure, or as being itself an algebraic structure.

2.2 General Kleisli representations

We will place a restriction on what qualifies as a Kleisli representation which is both very restrictive and very *natural*. To justify this restriction let us first define the *category of Kleisli T-representations of a monoid M* (alluded to in the remark above), denoted $\mathbf{Rep}(M, T)$, as the category whose objects are Kleisli T-representations and where an arrow exists between two representations $\rho : M \to \mathrm{End}_{\mathbf{Kl}(T)}(X)$ and $\rho' : M \to \mathrm{End}_{\mathbf{Kl}(T)}(Y)$ whenever there exists a $\mathbf{C}$-arrow $f : X \to Y$ such that $Tf \circ \rho(m) = \rho'(m) \circ f$ for each $m \in M$. In the case of finite-dimensional linear representations of a group, an arrow between two representations ρ, ρ' is equivalent to the usual notion of a homomorphism between ρ and ρ'. Observe the following: if $m, m' \in M$, and $f : X \to Y$ defines an arrow between representations ρ, ρ', the following diagram commutes (in $\mathbf{C}$):

$$
\begin{array}{ccccccc}
X & \xrightarrow{\rho(m)} & TX & \xrightarrow{T\rho(m')} & T^2X & \xrightarrow{\mu_X} & TX \\
\downarrow{\scriptstyle f} & & \downarrow{\scriptstyle Tf} & & \downarrow{\scriptstyle T^2 f} & & \downarrow{\scriptstyle Tf} \\
Y & \xrightarrow[\rho'(m)]{} & TY & \xrightarrow[T\rho'(m')]{} & T^2Y & \xrightarrow[\mu_Y]{} & TY
\end{array}
$$

since the two left-hand squares commute by definition of an arrow between representations and the right-hand square commutes by naturality of μ. What this means is that the naturality of μ makes the interpretation of the monoidal operation compatible with the notion of morphism between representation. It is this – in our view essential – feature which we want to preserve when the signature of the algebraic object to be represented is extended.

When $\mathbf{C}$ has products and a terminal object – which will always be the case in what follows – there are two monoidal structures on the category $[\mathbf{C}, \mathbf{C}]$ of endofunctors on $\mathbf{C}$: the first is given by functor composition and the identity functor, the second by the pointwise products and the constant functor respective to the terminal object, i.e. both $([\mathbf{C}, \mathbf{C}], \circ, \mathsf{Id}_{\mathbf{C}})$ and $([\mathbf{C}, \mathbf{C}], \times, \underline{1})$ are monoidal. We use this doubly-monoidal structure to represent on the one hand the *dynamical* composition operation with the *dynamical* monoidal structure $([\mathbf{C}, \mathbf{C}], \circ, \mathsf{Id}_{\mathbf{C}})$ (viz. as Kleisli representations), and on the other hand the *combinatorial* choice operation with the *combinatorial* monoidal structure $([\mathbf{C}, \mathbf{C}], \times, \underline{1})$. Specifically, given a monad T together with a *natural* transformation $\alpha : T \times T \to T$[1], we can define a new binary operation $\oplus$ on $\mathrm{End}_{\mathbf{Kl}(T)}(X)$ by

$$a \oplus b = \alpha_X \circ \langle a, b \rangle$$

(note the structural similarity with $a \circ_T b = \mu_X \circ (Tb \circ a)$). For a structure $(A, \cdot, 1, +)$ where $(A, \cdot, 1)$ is a monoid and $+$ is another binary operation, we can define a *Kleisli representation* of A as a homomorphism

$$\rho : (A, \cdot, 1, +) \to (\mathrm{End}_{\mathbf{Kl}(T)}, \circ_T, \eta_X, \oplus)$$

To see why α needs to be natural, consider again a morphism $f : \rho \to \rho'$ between two representations of $(A, \cdot, 1, +)$, two elements $m, m' \in A$ and the diagram

$$
\begin{array}{ccccc}
X & \xrightarrow{\langle \rho(m), \rho(m') \rangle} & TX \times TX & \xrightarrow{\alpha_X} & TX \\
\downarrow{\scriptstyle f} & & \downarrow{\scriptstyle Tf \times Tf} & & \downarrow{\scriptstyle Tf} \\
Y & \xrightarrow[\langle \rho'(m), \rho'(m') \rangle]{} & TY \times TY & \xrightarrow[\alpha_Y]{} & TY
\end{array}
$$

The left-hand-side square commutes by definition of a representation morphism, but for the right-hand-square to commute it is essential that α be natural. Once again, for Kleisli representations to be compatible with representation morphisms – and thus obvious notions like sub-representations or quotient representations – we need naturality. Finally, we also want to be able to represent the constant 0. For this we again proceed by analogy with the representation of the constant 1, which is interpreted through the unit of the monoidal structure $([\mathbf{C}, \mathbf{C}], \circ, \mathsf{Id}_{\mathbf{C}})$ by a natural transformation $\eta^T : \mathsf{Id}_{\mathbf{C}} \to T$. Similarly, we interpret the constant 0 through the unit of the monoidal structure $([\mathbf{C}, \mathbf{C}], \times, \underline{1})$ via a natural transformation $\zeta : \underline{1} \to T$. Such a transformation (if it exists, which we discuss below) allows us to define for any object X a map

$$0_X = \zeta_X \circ !_X : X \to TX$$

[1] The functor $T \times T : \mathbf{C} \to \mathbf{C}$ sends a $\mathbf{C}$-object X to $TX \times TX$ and a $\mathbf{C}$-morphism $f : X \to Y$ to the arrow $\langle Tf \circ \pi_1, Tf \circ \pi_2 \rangle : TX \times TX \to TY \times TY$ given by universality of the product $TY \times TY$. For simplicity's sake we will denote this arrow as $Tf \times Tf$.

where $!_X : X \to 1$ is the unique morphism to the terminal object. For a monad T with natural transformations $T \times T \to T$ and $\underline{1} \to T$, if $(A, \cdot, 1., +, 0)$ is an algebraic structure for which $(A, \cdot, 1)$ is a monoid, then we define a *natural Kleisli T-representation* of A as a homomorphism

$$\rho : (A, \cdot, 1, +, 0) \to (\mathrm{End}_{\mathbf{Kl}(T)}, \circ_T, \eta_X, \oplus, 0_X)$$

Again, the naturality of 0 is essential for representation morphisms to be well-defined. In the next sections we will show when transformations of the type $\underline{1} \to T$ and $T \times T \to T$ exist, and if they define representations in which the axioms K3-K10 hold.

Example

We can view the usual notion of representation of a Lie algebra in a finite-dimensional vector space as a Kleisli representation: let $T = \mathsf{UF}$, the monad associated with the free vector space functor $\mathsf{F} : \mathbf{Set} \to \mathbf{Vect}$, and let $\mathfrak{L}$ be a Lie algebra. The transformation $- : T \times T \to T$ defined at each X by $-_X : TX \times TX \to TX, (v, w) \mapsto v - w$ is natural (by linearity). Given a finite set n, we can then define a representation $\rho : \mathfrak{L} \to \mathrm{End}_{\mathbf{Kl}(T)}(n)$ by interpreting the Lie bracket $[\cdot, \cdot]$ as the expected commutator:

$$[v, w] \mapsto -_n \circ \langle \mu_n \circ T\rho(w) \circ \rho(v), \mu_n \circ T\rho(v) \circ \rho(w) \rangle.$$

3 Representability of 0

The following result will give us a completely general characterisation of natural transformations $\underline{1} \to T$.

Theorem 2. *Let $\mathbf{C}$ be a category with an initial object $\emptyset$ and $F : \mathbf{C} \to \mathbf{C}$ be a functor, then for every $\mathbf{C}$-object A there is a bijection*

$$\mathbf{C}(A, F\emptyset) \cong [\mathbf{C}, \mathbf{C}]\,(\underline{A}, F)$$

where $\underline{A}$ is the constant functor defined by A.

Proof. For any $\mathbf{C}$-object X there exists a unique arrow $!_X : \emptyset \to X$, and thus any arrow $m : A \to F\emptyset$ can be extended to an arrow $F!_X \circ m : A \to FX$. Moreover the collection of all such morphisms forms a natural transformation $\alpha(m)$: for any $f : X \to Y$ we have

$$\begin{array}{ccc} A & \xrightarrow{\underline{A}f} & A \\ {\scriptstyle F!_X \circ m}\downarrow & & \downarrow{\scriptstyle F!_Y \circ m} \\ FX & \xrightarrow[Ff]{} & FY \end{array}$$

which commutes since $!_Y = f \circ !_X$ by initiality. We thus have a map $\alpha : \mathbf{C}(A, F\emptyset) \to [\mathbf{C}, \mathbf{C}]\,(\underline{A}, F), m \mapsto \alpha(m)$. Conversely, we can define $\phi : [\mathbf{C}, \mathbf{C}]\,(\underline{A}, F) \to \mathbf{C}(A, F\emptyset), \beta \mapsto \beta_\emptyset$. It is clear that $\phi(\alpha(m)) = m$, to see that $\alpha(\phi(\beta)) = \alpha(\beta_\emptyset) = \beta$ we use the naturality of β and the map $!_X : \emptyset \to X$, and it follows that

$$\begin{array}{ccc} A & \xrightarrow{\underline{A}!_X} & A \\ {\scriptstyle \beta_\emptyset}\downarrow & & \downarrow{\scriptstyle \beta_X} \\ F\emptyset & \xrightarrow[F!_X]{} & FX \end{array}$$

which says exactly that $\alpha(\phi(\beta))_X = \beta_X$ at every X. ◻

Corollary 3. *If* $\mathbf{C}$ *is either* $\mathbf{Set}, \mathbf{Top}$ *or* $\mathbf{Pol}$*, we have*

$$[\mathbf{C}, \mathbf{C}]\,(\underline{1}, F) \cong \mathbf{C}(1, F\emptyset) \cong F\emptyset$$

Note that in **Set** the result above is a trivial consequence of the Yoneda lemma, since $\underline{1}$ is representable as $\hom(\emptyset, -)$. We conclude this section with some examples which will also introduce all the monads that we will examine.

Examples

In what follows X, Y are objects of the category on which the monad is defined, and $f : X \to Y$ is a morphism in this same category.

1. The *maybe monad* $\mathsf{M} : \mathbf{Set} \to \mathbf{Set}$ is defined by

$$\begin{cases} \mathsf{M}X = X + 1 \\ \mathsf{M}f = f + \mathrm{id}_1 : \mathsf{M}X \to \mathsf{M}Y \end{cases}$$

   The unit η^{M} is defined at each X as $\eta^{\mathsf{M}}_X = i_1$, the injection of X into $X + 1$, and the multiplication μ^{M} is defined at each X by $\mu^{\mathsf{M}}_X = \mathrm{id}_X + !_2 : X + 2 \to X + 1$ where $!_2 : 2 \to 1$ is the unique arrow to the terminal object[2]. It follows from Corollary 3 that there exists exactly one natural transformation $\underline{1} \to \mathsf{M}$ defined at each X as the injection of 1 in the coproduct $X + 1$.

2. The *finite list monad* $\mathsf{L} : \mathbf{Set} \to \mathbf{Set}$ is defined by

$$\begin{cases} \mathsf{L}X = \coprod_{n \in \mathbb{N}} X^n \\ \mathsf{L}f = \coprod_{n \in \mathbb{N}} f^n : \mathsf{L}X \to \mathsf{L}Y \end{cases}$$

   The unit η^{L} is defined at each X as $\eta^{\mathsf{L}}_X(x) = [x]$, and the multiplication μ^{L} is defined at each X by the concatenation operation on finite lists of elements of X. Since $\mathsf{L}\emptyset = \{\varepsilon\}$, the empty list, it follows that there exists a unique natural transformation $\underline{1} \to \mathsf{L}$ which selects the empty list at each X.

3. The *hybrid monad* $\mathsf{H} : \mathbf{Top} \to \mathbf{Top}$ is defined by

$$\begin{cases} \mathsf{H}X = \left\{ X^{\mathbb{R}^+} \times [0, \infty] \mid f \circ \min(\_, d) = f \right\} \\ \mathsf{H}f = f^{\mathbb{R}^+} \times \mathrm{id} : \mathsf{H}X \to \mathsf{H}Y \end{cases}$$

   The unit η^{H} is defined at each X by $\eta^{\mathsf{H}}_X(x) = (\underline{x}, 0)$, where $\mathbb{R}^+$ denotes the space $[0, \infty)$, $\underline{x}$ is the constant function on x, and multiplication is defined by

$$\mu^{\mathsf{H}}_X(f, d) = \begin{cases} (\theta_X \circ f, d) +\!\!+ (f\ d) & \text{if } d \neq \infty \\ (\theta_X \circ f, d) & \text{otherwise} \end{cases}$$

   where $\theta_X : \mathsf{H}X \to X$ is the map defined as $\theta_X(f, d) = f(0)$, and $(+\!\!+)$ denotes function concatenation. More details about this particular monad can be found in [NBHM16]. Note that it can be analogously defined in **Set**, the underlying functor of this variant being here denoted by $\mathsf{E} : \mathbf{Set} \to \mathbf{Set}$. It follows from Corollary 3 that there is no natural transformation $\underline{1} \to \mathsf{H}$ or $\underline{1} \to \mathsf{E}$.

[2] For $f : X_1 \to Y_1$ and $g : X_2 \to Y_2$, we use $f + g$ as a short-hand notation for the map $[i_1 \circ f, i_2 \circ g] : X_1 + Y_1 \to X_2 + Y_2$ given by universality of the coproduct $X_2 + Y_2$.

4. The *powerset monad* $\mathsf{P} : \mathbf{Set} \to \mathbf{Set}$ is defined by

$$\begin{cases} \mathsf{P}X = \{U \mid U \subseteq X\} \\ \mathsf{P}f = f[-] : \mathsf{P}X \to \mathsf{P}Y \end{cases}$$

where $f[-]$ denotes the direct image. The unit η^{P} is defined at each X as $\eta^{\mathsf{P}}_X(x) = \{x\}$, and the multiplication μ^{P} is defined at each X by $\mu^{\mathsf{P}}_X(\mathcal{U}) = \bigcup \mathcal{U}$. Note that the non-empty powerset monad $\mathsf{Q} : \mathbf{Set} \to \mathbf{Set}$ together with these monad operations also forms a monad. It follows from Corollary 3 there exists a unique natural transformation $\zeta : \underline{1} \to \mathsf{P}$, which selects the empty set $\emptyset$ at each X, and that there is no natural transformation $\zeta : \underline{1} \to \mathsf{Q}$.

5. Given a semiring S, the *generalized multiset monad* $\mathsf{B}_S : \mathbf{Set} \to \mathbf{Set}$ is defined by

$$\begin{cases} \mathsf{B}_S X = \{\phi : X \to S \mid |\mathrm{supp}(\phi)| < \omega\} \\ \mathsf{B}_S f : \mathsf{B}_S X \to \mathsf{B}_S Y, \phi \mapsto \lambda y. \sum\limits_{x \in f^{-1}(\{y\})} \phi(x) \end{cases}$$

the unit η^{B} is defined at each X by $\eta^{\mathsf{B}}_X(x)(y) = \delta_x(y)$, i.e. 1 if $x = y$ and 0 otherwise. The multiplication μ^{B} is defined at each X by $\mu^{\mathsf{B}}_X(\Phi)(x) = \sum_{\phi \in \mathrm{supp}(\Phi)} \Phi(\phi) \cdot \phi(x)$, where $\cdot$ is the semiring multiplication. It follows from Corollary 3 that there exists a unique natural transformation $\underline{1} \to \mathsf{B}_S$, which selects at each set X the constant map $\underline{0} : X \to S$.

6. The *distribution monad* $\mathsf{D} : \mathbf{Set} \to \mathbf{Set}$ is defined by

$$\begin{cases} \mathsf{D}X = \{p : X \to [0,1] \mid \sum\limits_{x \in X} p(x) = 1, |\mathrm{supp}(p)| < \omega\} \\ \mathsf{D}f : \mathsf{D}X \to \mathsf{D}Y, p \mapsto \lambda y. \sum\limits_{x \in f^{-1}(\{y\})} p(x) \end{cases}$$

The unit is given by $\eta^{\mathsf{D}}_X(x) = \delta_x$, the Dirac delta at x and the multiplication by $\mu^{\mathsf{D}}_X(P)(x) = \sum_{p \in \mathrm{supp}(P)} P(p) \cdot p(x)$. Since $\mathsf{D}\emptyset = \emptyset$ there is no natural transformation $\underline{1} \to \mathsf{D}$.

7. The *Giry monad*: $\mathsf{G} : \mathbf{Pol} \to \mathbf{Pol}$ is defined by

$$\begin{cases} \mathsf{G}(X, \tau) = (\{\text{probabilities on } \sigma(\tau)\}, \\ \qquad\qquad\quad \text{topology of weak convergence}) \\ \mathsf{G}f : \mathsf{G}X \to \mathsf{G}Y, \quad \mu \mapsto f^*\mu \text{ (pushforward measure)} \end{cases}$$

where $\sigma(\tau)$ is the Borel σ-algebra (for all the details see [Gir82] or Chapter 15 of [AB06]). The unit is given by $\eta^{\mathsf{G}}_X(x) = \delta_x$ and the multiplication by $\mu^{\mathsf{G}}_X(P)(B) = \int_{\mathsf{G}X} ev_B dP$, where $ev_B : \mathsf{G}X \to [0,1], \nu \mapsto \nu(B)$. Since $\mathsf{G}\emptyset = \emptyset$, there are no natural transformations $\underline{1} \to \mathsf{G}X$. Note that G restricted to $\mathbf{Pol}_f$, the full subcategory of *finite* Polish spaces, is equivalent to D restricted to the category of finite sets.

We summarize some of our observations in the following 'no-go theorem'.

Theorem 4 (No-Go Theorem I)**.** *Axioms K5, K7 and K8 – and thus Kleene algebras – have no Kleisli representations for the monads* $\mathsf{D}, \mathsf{G}, \mathsf{H}, \mathsf{E}$*, and* Q*.*

4 Representability of $+$, associativity, commutativity, idempotency and unit

For a monad $T : \mathbf{Set} \to \mathbf{Set}$, it is often possible to characterise completely the set of natural transformations $T \times T \to T$ which will interpret the 'non-deterministic choice' operation $+$. Our results require a bit of background material about the presentations of **Set**-valued functors.

4.1 Introduction to the presentation of Set-valued functors

Let $\mathbf{C}$ be a *small* category and $F : \mathbf{C} \to \mathbf{Set}$ be a functor. We define the *category of elements of* F, denoted $\mathbf{El}(F)$, as the category whose objects are pairs (C, α) where C is an object in $\mathbf{C}$ and $\alpha \in FC$. There exists a morphism $\hat{f} : (C, \alpha) \to (D, \beta)$ in $\mathbf{El}(F)$ whenever there exists a morphism $f : C \to D$ such that $Ff(\alpha) = \beta$. For every object (C, α) in $\mathbf{El}(F)$, we will define the *orbit* of (C, α) as all the objects (D, β) which can be reached from (C, α) by a morphism in $\mathbf{El}(F)$. The decomposition of $\mathbf{El}(F)$ in orbits is key to understanding natural transformations involving F, since naturality is only a constraint on objects in the same orbit. The category $\mathbf{El}(F)$ is extremely useful because it allows us to completely reconstruct F. Moreover, this reconstruction process provides us with a presentation of F as a colimit of covariant hom functors.

Theorem 5. *Let* $\mathsf{Y} : \mathbf{C}^{\mathrm{op}} \to [\mathbf{C}, \mathbf{Set}]$ *denote the Yoneda embedding, and* $\mathsf{U}_F : \mathbf{El}(F) \to \mathbf{C}$ *be the forgetful functor sending each pair* (A, α) *to the object* A*, then*

$$F \cong \operatorname{colim}\left(\mathbf{El}(F)^{\mathrm{op}} \xrightarrow{\mathsf{U}_F^{\mathrm{op}}} \mathbf{C}^{\mathrm{op}} \xrightarrow{\mathsf{Y}} [\mathbf{C}, \mathbf{Set}]\right)$$

Proof. See [MM12] I.5 for a proof in the case of presheaves. □

Theorem 5 gives us a way of presenting functors from a *small* category $\mathbf{C}$ to $\mathbf{Set}$ as a colimit of hom functors, but what we really need are presentations of functors $\mathbf{Set} \to \mathbf{Set}$. To move from $\mathbf{C} \to \mathbf{Set}$ to $\mathbf{Set} \to \mathbf{Set}$ we need a few relatively well-known definitions, for further details we refer the reader to the classic [AR94]. Recall that an object A in a category $\mathbf{C}$ is *finitely presentable* if the functor $\hom(A, -)$ preserves filtered colimits. In $\mathbf{Set}$ the finitely presentable objects are precisely the finite sets. A category $\mathbf{C}$ is called *locally finitely presentable* if it is cocomplete and contains a small subcategory $\mathbf{C}_\omega$ of finitely presentable objects such that every object A in $\mathbf{C}$ is the filtered colimit of the canonical diagram $\mathscr{D}_A : \mathbf{C}_\omega \downarrow A \to \mathbf{C}$ sending each arrow of $\mathbf{C}_\omega \downarrow A$ to its domain. The category $\mathbf{Set}$ is finitely presentable, since every set X can written as $\operatorname{colim} \mathscr{D}_X$ with $\mathscr{D}_X : \omega \downarrow X \to \mathbf{Set}$ and ω the subcategory of $\mathbf{Set}$ consisting of elements $n \in \omega$ (this simply says that every set is the union of its finite subsets). Finally, a functor is called *finitary* if it preserves filtered colimits. Finitary functors are entirely determined by their restriction to $\mathbf{C}_\omega$. In fact if $F : \mathbf{C} \to \mathbf{C}$ is a finitary functor on a locally finitely presentable category and $\mathsf{I} : \mathbf{C}_\omega \hookrightarrow \mathbf{C}$ is the inclusion functor, then F can be written as the left Kan extension $F = \mathrm{Lan}_{\mathsf{I}}(F_f)$, where $F_f = F \circ \mathsf{I}$.

Proposition 6. *Let* $F : \mathbf{Set} \to \mathbf{Set}$ *be finitary and let* $\mathsf{I} : \omega \hookrightarrow \mathbf{Set}$ *be the inclusion functor, then*

$$F \cong \operatorname{colim}\left(\mathbf{El}(F_f)^{\mathrm{op}} \xrightarrow{\mathsf{U}_{F_f}^{\mathrm{op}}} \omega^{\mathrm{op}} \xrightarrow{\mathsf{I}^{\mathrm{op}}} \mathbf{Set}^{\mathrm{op}} \xrightarrow{\mathsf{Y}} [\mathbf{Set}, \mathbf{Set}]\right)$$

Proof. This is a simple consequence of the fact that $F = \mathrm{Lan}_{\mathsf{I}}(F_f)$, Theorem 5 applied to F_f, the fact that $\mathrm{Lan}_{\mathsf{I}}(F_f)$ can be expressed as a colimit and that colimits commute. □

Example

The finitary powerset monad $\mathsf{P}_\omega : \mathbf{Set} \to \mathbf{Set}$ is, unsurprisingly, finitary. To simplify the notation let us overload P_ω to also denote its restriction to the small subcategory ω. The category of elements $\mathbf{El}(\mathsf{P}_\omega)$ is given by the pairs $(n, U), U \in \mathsf{P}_\omega n$, and one can easily show that the inclusion of the subcategory $\mathbf{N} \hookrightarrow \mathbf{El}(\mathsf{P}_\omega)$ of pairs of the shape $(n, \{n\})$ is a final functor. The morphisms $(n, \{n\}) \to (m, \{m\})$ in $\mathbf{N}$ are the surjections $p : n \twoheadrightarrow m$ ($\mathsf{P}_\omega p(\{n\}) = \{m\}$ is then automatic). It follows from Proposition 6 that

$$\mathsf{P}_\omega X \cong \operatorname*{colim}_{n \in \mathbf{N}} \hom(n, X) \cong \coprod_{n \in \omega} \hom(n, X)/\sim$$

where $x : n \to X \sim y : m \to X$ iff there exists a surjection $p : n \twoheadrightarrow m$ (or $m \twoheadrightarrow n$) such that $y = x \circ p$ (or $x = y \circ p$), which is to say that they are in the same equivalence class iff $\{x(1), \ldots, x(n)\} = \{y(1), \ldots, y(m)\}$. This presentation is well-known and can be found in e.g. [AGT10].

4.2 Characterising $[\mathbf{Set}, \mathbf{Set}](T \times T, T)$

The following result is a simple application of Proposition 6 and the Yoneda lemma.

Theorem 7. *Let $F : \mathbf{Set} \to \mathbf{Set}$ be a finitary functor and let F_f denote its restriction to ω, then the set $[\mathbf{Set}, \mathbf{Set}](F \times F, F)$ is in one-to-one correspondence with the limit*

$$\lim \left(\mathbf{El}(F_f \times F_f)^{\mathrm{op}} \overset{F_f^{\mathrm{op}} \circ \mathsf{U}^{\mathrm{op}}}{\longrightarrow} \mathbf{Set} \right) \tag{1}$$

Proof. We simply calculate:

$$\begin{aligned}
& [\mathbf{Set}, \mathbf{Set}]\,(F \times F, F) \\
\overset{(1)}{=}\; & [\mathbf{Set}, \mathbf{Set}] \left(\operatorname{colim} \mathbf{El}(F_f \times F_f)^{\mathrm{op}} \overset{\mathsf{Y} \circ \mathsf{I}^{\mathrm{op}} \circ \mathsf{U}^{\mathrm{op}}}{\longrightarrow} [\mathbf{Set}, \mathbf{Set}], F \right) \\
\overset{(2)}{=}\; & \lim\, [\mathbf{Set}, \mathbf{Set}] \left(\mathbf{El}(F_f \times F_f)^{\mathrm{op}} \overset{\mathsf{Y} \circ \mathsf{I}^{\mathrm{op}} \circ \mathsf{U}^{\mathrm{op}}}{\longrightarrow} [\mathbf{Set}, \mathbf{Set}], F \right) \\
\overset{(3)}{=}\; & \lim \left(\mathbf{El}(F_f \times F_f)^{\mathrm{op}} \overset{F_f^{\mathrm{op}} \circ \mathsf{U}^{\mathrm{op}}}{\longrightarrow} \mathbf{Set} \right)
\end{aligned}$$

where (1) is an application of Proposition 6, (2) follows from the fact that the contravariant hom functor sends colimits to limits, and (3) is an application of the Yoneda lemma. □

The hard work consists in computing the limit (1) above. We start with a class of functors for which the limit can be computed rather painlessly.

4.2.1 Coproducts of hom functors

Theorem 8. *If $T : \mathbf{Set} \to \mathbf{Set}$ is a functor expressible as a coproduct of* hom *functors, i.e. if there exists a non-empty family $(X_i)_{i \in I}$ of sets such that $T = \coprod_{i \in I} \hom(X_i, -)$, then*

$$[\mathbf{Set}, \mathbf{Set}]\,(T \times T, T) \cong \prod_{i,j \in I} T(X_i + X_j)$$

A natural transformations $T \times T \to T$ determined by $s \in \prod_{i,j \in I} T(X_i + X_j)$ with component $s_{ij} \in T(X_i + X_j)$

- is commutative *iff for all $i, j \in I$,* *(i) if $X_i, X_j \neq \emptyset$ then $\exists k \in I$ such that $X_k = \emptyset, s_{ij} \in \hom(X_k, X_i + X_j)$ and $s_{ji} \in \hom(X_k, X_j + X_i)$*

(ii) if $X_i = \emptyset$ then $s_{ij} = s_{ji}$.

In particular if $X_i \neq \emptyset$ for all $i \in I$ then there exist no commutative element of $[\mathbf{Set}, \mathbf{Set}](T \times T, T)$.

- is idempotent *iff for each $i \in I$, $s_{ii} \in T(X_i + X_i)$ is a map $s_{ii} : X_i \to X_i + X_i$ s. th. $s_{ii}(x) = x$ for all $x \in X_i$.*
- has a unit $\underline{1} \to T$ *defined by an element of* $\hom(X_i, -)$ *with $X_i = \emptyset$ iff $T\emptyset = \coprod_{i|X_i=\emptyset} 1 \neq \emptyset$ and $s_{ij} = s_{ji} = id_{X_j}$.*

Note that we have not included a criterion for a collection $s \in \prod_{i,j \in I} T(X_i + X_j)$ to define an associative natural transformation. The reason is that such a criterion would demand a very complex relationship between the components s_{ij} of s, and does not provide a easier test for associativity than simply checking associativity itself. Note moreover, that associativity is always satisfiable in Kleisli representations by choosing $\alpha : T \times T \to T$ to be one of the projections. It is however not usually valid in Kleisli representations, i.e. it is not the case that for an arbitrary α, the axiom K3 holds.

Corollary 9. *There are $\mathsf{M}(2) \times \mathsf{M}(1) \times \mathsf{M}(1)$ natural transformations $\mathsf{M} \times \mathsf{M} \to \mathsf{M}$. Such a transformation is commutative only if its '$\mathsf{M}(2)$ coordinate' lies in the second summand of $\mathsf{M}(2)$, idempotent iff its '$\mathsf{M}(2)$ coordinate' is in the first summand of $\mathsf{M}(2)$, and has the unique natural transformation $\underline{1} \to \mathsf{M}$ as a unit iff its '$\mathsf{M}(1)$ coordinates' are in the first summand of $\mathsf{M}(1)$.*

The maybe monad gives us a simple example of transformation α which defines Kleisli representations in which K3 fails. Take $\alpha^{(\star,\star,1)} : \mathsf{M} \times \mathsf{M} \to \mathsf{M}$ defined by the triple $(\star, \star, 1)$. It maps pairs (x, y) with $x, y \neq \star$ to $\star$, the pairs $(x, \star)$ to $\star$ and the pairs $(\star, x)$ to x. For any $x \in X$

$$\begin{aligned} &\alpha_X^{(\star,\star,1)}(\alpha_X^{(\star,\star,1)} \times \mathrm{id}_X(x, \star, x)) = \alpha_X^{(\star,\star,1)}(\star, x) = x \\ \neq &\alpha_X^{(\star,\star,1)}(\mathrm{id}_X \times \alpha_X^{(\star,\star,1)}(x, \star, x)) = \alpha_X^{(\star,\star,1)}(x, x) = \star \end{aligned}$$

The same counter-example can be replicated for L, P and $\mathsf{B}_{\mathbb{N}}$.

Corollary 10. *The natural transformations $\mathsf{L} \times \mathsf{L} \to \mathsf{L}$ are in one-to-one correspondence with $\prod_{i,j \in \mathbb{N}} \mathsf{L}(i + j)$. Moreover,*

- *the commutative ones are all those mapping pairs of non-empty list to the empty list and pairs $(l, \varepsilon) \in X^n \times X^0$ to either l or ϵ, provided that the same choice is made for $(\varepsilon, l) \in X^0 \times X^n$.*
- *the idempotent ones are defined by maps $s_{ii} : i \to i + i$ which maps $x \in i$ to its copy in either the first or the second summand.*
- *the constant transformation to the empty list is a unit the transformations whose $\mathsf{L}(i + \emptyset)$-components are id_i.*

Note in particular that M (resp. L) does not admit a natural transformation $\mathsf{M} \times \mathsf{M} \to \mathsf{M}$ (resp. $\mathsf{L} \times \mathsf{L} \to \mathsf{L}$) which is simultaneously commutative and idempotent.

Corollary 11. *There is no commutative natural transformation $\mathsf{H} \times \mathsf{H} \to \mathsf{H}$*

Proof. Note that the functor $\mathsf{E} : \mathbf{Set} \to \mathbf{Set}$ is isomorphic to

$$\left(\coprod_{r \in \mathbb{R}^+} \hom([0,r], -) \right) + \hom(\mathbb{R}^+, -)$$

It follows from Theorem 8 that there is no commutative natural transformation $\mathsf{E} \times \mathsf{E} \to \mathsf{E}$. This implies that there is no commutative natural transformation $\mathsf{H} \times \mathsf{H} \to \mathsf{H}$. Otherwise, given one such natural transformation, it would be possible to construct a commutative natural transformation $\mathsf{E} \times \mathsf{E} \to \mathsf{E}$ using the fact that the diagram below commutes

$$\begin{array}{ccc} \mathbf{Top} & \xrightarrow{\mathsf{H}} & \mathbf{Top} \\ {\scriptstyle I}\uparrow & & \downarrow{\scriptstyle \mathsf{U}} \\ \mathbf{Set} & \xrightarrow[\mathsf{E}]{} & \mathbf{Set} \end{array}$$

where $I : \mathbf{Set} \to \mathbf{Top}$ is the indiscrete functor. □

As can be seen from $\mathsf{M}, \mathsf{L}, \mathsf{H}$ and E, commutativity is a problematic axiom for the class of functor treated by Theorem 8 as it necessarily lead to trivial natural transformations in the sense that no pair of 'non-abort' instructions can be non-trivially combined. In particular, it is incompatible with idempotency.

Theorem 12 (No-Go Theorem II)**.** *Axiom K4 – and thus KAs – has no Kleisli representation for the monads* H *and* E*. The combination of axioms K4 and K6 – and thus KAs – has no Kleisli representation for the monads* M *and* L*.*

4.2.2 The case of the powerset monad

Theorem 13. *For a cardinal* λ *the set of natural transformations* $(\mathsf{P})^\lambda \to \mathsf{P}$*, where* $(\mathsf{P})^\lambda$ *denotes the* λ*-fold product of* P *with itself, is in one-to-one correspondence with the set of non-increasing maps* $2^\lambda \to 2^\lambda$ *(for the obvious order on* 2^λ*).*

Corollary 14. *There are exactly* 16 *natural transformations* $\mathsf{P} \times \mathsf{P} \to \mathsf{P}$*. The union operation corresponds to the identity map* $\mathrm{id}_{2^2} : 2^2 \to 2^2$ *and is the unique natural transformation* $\mathsf{P} \times \mathsf{P} \to \mathsf{P}$ *which satisfies the axioms K3-K6.*

4.2.3 Generalised restricted multiset functors

We now examine another class of functors which covers several of the examples listed in Section 3. They are all instances of the generalised multiset functor, with some possible restrictions on the multisets. For example the distribution monad can be seen as the $\mathbb{R}$-valued multisets whose total mass is exactly one. Computing the limit (1) of Theorem 7 depends heavily on the choice of semi-ring and on the possible restrictions and may prove extremely difficult. However we do have explicit characterisation in some useful cases.

Proposition 15. *The set of* $[\mathbf{Set}, \mathbf{Set}]\,((\mathsf{B}_{\mathbb{N}})^n, \mathsf{B}_{\mathbb{N}})$ *is in one-to-one correspondence with the set of functions* $\phi : \mathbb{N}^n \to \mathbb{N}^n$ *such that* $\pi_i(\phi(m_1, \ldots, m_n)) = 0$ *whenever* $m_i = 0$*.*

Proof. For clarity we show the result for $n = 2$, the same argument holds for any finite n. When $\mathsf{B}_{\mathbb{N}}$ is applied to a map it creates a map between multisets which preserves the total mass of the multiset. By considering the singleton set 1, it is then clear that $\mathbf{El}(\mathsf{B}_{\mathbb{N}} \times \mathsf{B}_{\mathbb{N}})$ has $\mathbb{N}^2$ orbits. Let us now choose a pair of multisets $((x_1 : i_1, \dots, x_n : i_n), (y_1 : j_1, \dots, y_p : j_p))$ which belongs to the orbit indexed by $(M = \sum_k i_k, N = \sum_k j_k)$. Assume w.l.o.g. that $M \leq N$ and consider the object of $\mathbf{El}(\mathsf{B}_{\mathbb{N}} \times \mathsf{B}_{\mathbb{N}})$ defined by

$$(N, ((1{:}1, \dots, M{:}1, M+1{:}0, \dots, N{:}0), (1{:}1, \dots, N{:}1)))$$

i.e. we assign weight 1 to the first M elements on the left, and we assign weight 1 to all N elements on the right. Clearly this object also belongs to the orbit (M, N). Moreover, it is invariant under all permutations of the first M terms, and it is also invariant under all permutations of the last $N - M$ terms. It follows that for any thread in the limit (1), the component corresponding to the object above must be of the shape

$$(1 : p, \dots, M : p, M+1 : q, \dots, N : q)$$

for some $(p, q) \in \mathbb{N}^2$. Clearly if $M = 0$ or $N = 0$, then we can take $p = 0$ or $q = 0$ accordingly. It remains to show that the choice of (p, q) must be made consistently across the orbit (M, N), but this is immediate since any other object in the (M, N) orbit of $\mathbf{El}(\mathsf{B}_{\mathbb{N}} \times \mathsf{B}_{\mathbb{N}})$ is the codomain of some morphism from $(N, ((1{:}1, \dots, M{:}1, M+1{:}0, \dots, N{:}0), (1{:}1, \dots, N{:}1)))$. ☐

The intuition behind the correspondence with the maps $\phi : \mathbb{N}^2 \to \mathbb{N}^2$ described above is simple: it says that in the orbit (m, n) the natural transformation takes the weighted sum of the two multisets where the weights are given by $\phi(m, n)$. Note that the proof above does not hold for the ring $\mathbb{R}$: assume a multiset $(x : r_1, y : r_2)$, if $\frac{r_1}{r_2}$ is irrational, then it is not possible to find a uniform multiset $(x_1 : \lambda, \dots, x_n : \lambda)$ which maps to $(x : r_1, y : r_2)$.

Theorem 16 (No-Go Theorem III)**.** *The combination of axioms K4 and K6 – and thus KAs – has no Kleisli representation for the monad* $\mathsf{B}_{\mathbb{N}}$.

Proof. It is not hard to check that a natural transformation $\mathsf{B}_{\mathbb{N}} \times \mathsf{B}_{\mathbb{N}} \to \mathsf{B}_{\mathbb{N}}$ is commutative iff the corresponding map $\mathbb{N}^2 \to \mathbb{N}^2$ is. It follows that there is no idempotent and commutative natural transformation $\mathsf{B}_{\mathbb{N}} \times \mathsf{B}_{\mathbb{N}} \to \mathsf{B}_{\mathbb{N}}$ since the commutative transformations at least double the weight of a pair of identical multisets. ☐

Note that the 0 multiset is a unit for any natural transformation $\mathsf{B}_{\mathbb{N}} \times \mathsf{B}_{\mathbb{N}} \to \mathsf{B}_{\mathbb{N}}$, and that the only idempotent natural transformations are those selecting a projection for each orbit.

4.3 Some results for the Giry monad

Due to its importance, we dedicate a section to some results about the Giry monad G. We believe that these results clarify what can be expected of purely probabilistic program semantics. We start with an application of Corollary 3.

Proposition 17. *There is no natural transformation* $\underline{1} \to \mathsf{G}$.

In [DDG16a] and [DDG16b] the first author *et al.* developed a set of criteria (known as 'the Machine') for functors $F, G : \mathbf{Pol} \to \mathbf{Pol}$ under which it can be shown that

$$[\mathbf{Pol}, \mathbf{Pol}]\,(F, G) \simeq [\mathbf{Pol}_f, \mathbf{Pol}_f]\,(F_f, G_f)$$

where F_f is the restriction of F to the category $\mathbf{Pol}_f$, the category of finite Polish spaces (and similarly for G_f). These criteria restrict both the domain and the codomain functors. We refer the reader to [DDG16a] for more details; for our purpose it will be enough to say that the Giry monad G always satisfies the domain and codomain criteria (see [DDG16a] Prop. 5.1) and that finite products of G satisfy the domain criteria (see [DDG16b] Proposition 17). It follows that

$$[\mathbf{Pol}, \mathbf{Pol}]\,(\mathsf{G} \times \mathsf{G}, \mathsf{G}) \simeq [\mathbf{Pol}_f, \mathbf{Pol}_f]\,(\mathsf{G}_f \times \mathsf{G}_f, \mathsf{G}_f) \tag{2}$$

The isomorphism of Eq. (2) allows the following results to be established at the level of finite Polish sets, and then lifted to the entire category **Pol**.

Theorem 18. *The only natural transformations* $\mathsf{G} \times \mathsf{G} \to \mathsf{G}$ *are the convex combinations* $+^{\lambda}$ *defined by the maps* $+^{\lambda}_{X} : \mathsf{G}X \times \mathsf{G}X \to \mathsf{G}X, \lambda \in [0, 1]$ *defined by*

$$(\mu +^{\lambda}_{X} \nu)(A) = \lambda\mu(A) + (1 - \lambda)\nu(A)$$

for any Borelian set A*.*

Corollary 19. *The only commutative natural transformation* $\mathsf{G} \times \mathsf{G} \to \mathsf{G}$ *is the averaging transformation* $+^{1/2}$*.*

Remark 20. *It follows that idempotency is valid for Kleisli* G*-representation. Since idempotency might not be a desirable feature of probabilistic Kleene algebras (see e.g. [FKM+16] or [SKF+17] where the '&' operation is not idempotent), but commutativity usually is (see op. cit. where the '&' operation is commutative), this suggest that a naive representation of the choice operation by a natural transformation* $\mathsf{G} \times \mathsf{G} \to \mathsf{G}$ *is inadequate. This compounds the problem of Proposition 17, and will require a more sophisticated approach, to which we will return in Sections 7 and 8.*

4.4 Summary:

We gather our results in the following table.

Axiom	M	L	H	P	$\mathsf{B}_{\mathbb{N}}$	D	G
K3: Associativity	✓	✓	✓	✓	✓	✓	✓
K4: Commutativity	✓	✓	✗	✓	✓	✓	✓
K5: Unit	✓	✓	✗	✓	✓	✗	✗
K6: Idempotency	✓	✓	✓	✓	✓	✓	✓
K3+K4+K5	✓	✓	✗	✓	✓	✗	✗
K3+K4+K6	✗	✗	✗	✓	✗	✓	✓
K3+K5+K6	✓	✓	✗	✓	✓	✗	✗
K3+K4+K5+K6	✗	✗	✗	✓	✗	✗	✗
Valid axioms					K5	?	K3,K6

Table 1: Satisfiability of axioms K3-K6 in Kleisli representations.

Projections $T \times T \to T$ are always associative and idempotent, which is why we do not have a K3+K6 row in Table 1. The hybrid monad H is clearly the most troublesome and Kleisli H-representations appear to support very few program constructs. We will try to remedy this in section 7. Note also that we have a column for D and a column for G, because at present we do not have a complete classification of natural transformations $\mathsf{D} \times \mathsf{D} \to \mathsf{D}$, but we do for $\mathsf{G} \times \mathsf{G} \to \mathsf{G}$ (Theorem 18). The reason for this can be found in the proof of Theorem 18 which proves the result for rational distributions and then extends it to all distributions by continuity and completeness in **Pol**. This argument fails in **Set**, and we therefore do not know at present how a natural transformation $\mathsf{D} \times \mathsf{D} \to \mathsf{D}$ would act on irrational distributions. This explains why we cannot yet say whether the axioms K3 and K6 are valid for Kleisli D-representation.

5 Distributivity

5.1 The axioms K7 and K8

The satisfiability of the axioms K7 and K8 in Kleisli representations is a surprisingly complex issue. On the one hand the following result always holds:

Proposition 21. *The axiom K7 holds in any Kleisli representation.*

Proof. It follows by the naturality of $0 : \underline{1} \to T$ applied to the map $\eta_X : X \to TX$ that $T\eta_X \circ 0_X = 0_{TX}$, and thus

$$\mu_X \circ 0_{TX} = \mu_X \circ T\eta_X \circ 0_X = 0_X \tag{3}$$

We then have for any Kleisli representation ρ

$$\begin{aligned} \rho(a;0) &:= \mu_X \circ Ta \circ 0_X \circ !_X & \\ &= \mu_X \circ 0_{TX} \circ !_X & \text{Naturality of } 0 \\ &= 0_X \circ !_X & \text{By Eq. (3)} \\ &:= \rho(0) & \end{aligned}$$

□

On the other hand the axiom $a;0 = 0$ need not hold. The following is probably the smallest counterexample: consider the monad M^2 (it is a monad by Theorem 27). It has two natural transformations $\underline{1} \to \mathsf{M}^2$ (by Corollary 3) corresponding to selecting the first or second $+1$ summand. Let us assume that 0 is interpreted by picking the first '+1' summand and consider the constant arrow $a : X \to X + 1 + 1$ to the second '+1' summand. An easy computation shows that the representation of $a;0$, viz. $\mu_X \circ T0_X \circ T!_X \circ a : X \to TX$, is also the constant map to the second '+1' summand, which is clearly different from $0_X : X \to TX$ which is constant to the first '+1' summand. The wisdom gained from this counter-example may be that things go wrong for axiom K8 when programs can crash in more than one way.

We do not know of a pithier criterion guaranteeing $a;0 = 0$ than simply unravelling what K8 means in a Kleisli representation. The axiom is satisfiable for $\mathsf{M}, \mathsf{P}, \mathsf{L}, \mathsf{B}_S$ and V (whilst it does not even make sense for $\mathsf{D}, \mathsf{G}, \mathsf{E}$ and H).

5.2 The axioms K9 and K10

The left and right distribution axioms K9 and K10 need not hold in every Kleisli representation, in fact we shall soon provide a simple counterexample. The following criteria on a choice of natural transformation $\alpha : T \times T \to T$ enforce the validity of K9 and K10 in the corresponding representations.

Theorem 22. *Let* $\mathbf{C}$ *be a category with products,* $T : \mathbf{C} \to \mathbf{C}$ *be a monad, and* $\alpha : T\times T \to T$ *be a natural transformation. If diagram (4) (resp. (5)) commutes then, then the axiom K9 (resp. K10) holds in any Kleisli representation defined by* α*.*

$$\begin{array}{ccc} TTX \times TTX & \xrightarrow{\alpha_{TX}} & TTX \\ \downarrow{\scriptstyle \mu_X \times \mu_X} & & \downarrow{\scriptstyle \mu_X} \\ TX \times TX & \xrightarrow[\alpha_X]{} & TX \end{array} \tag{4}$$

$$\begin{array}{ccc} T(TX \times TX) & \xrightarrow{\langle T\pi_1, T\pi_2 \rangle} & TTX \times TTX \\ {\scriptstyle T\alpha_X}\downarrow & & \downarrow{\scriptstyle \alpha_X \circ \mu_X \times \mu_X} \\ TTX & \xrightarrow[\mu_X]{} & TX \end{array} \tag{5}$$

Proof. Let ρ be a Kleisli representation, we compute

$$\begin{aligned} & \rho(a;(b+c)) \\ := \; & \mu_X \circ Ta \circ \alpha_X \circ \langle b, c\rangle \\ = \; & \mu_X \circ \alpha_{TX} \circ (Ta \times Ta) \circ \langle b, c\rangle && \alpha \text{ natural} \\ = \; & \alpha_X \circ (\mu_X \times \mu_X) \circ (Ta \times Ta) \circ \langle b, c\rangle && \text{Diagram (4)} \\ = \; & \alpha_X \circ \langle \mu_X \circ Ta \circ b, \mu_X \circ Ta \circ c\rangle \\ := \; & \rho((a;b)+(a;c)) \end{aligned}$$

Similarly,

$$\begin{aligned} & \rho((b+c);a) \\ := \; & \mu_X \circ T(\alpha_X \circ \langle b, c\rangle) \circ a \\ = \; & \mu_X \circ T\alpha_X \circ T\langle b, c\rangle \circ a \\ = \; & \alpha_X \circ \mu_X \times \mu_X \circ \langle T\pi_1, T\pi_2\rangle \circ T\langle b, c\rangle \circ a && \text{Diagram (5)} \\ = \; & \alpha_X \circ \mu_X \times \mu_X \circ \langle Tb, Tc\rangle \circ a && \langle Tb, Tc\rangle \text{ unique} \\ = \; & \alpha_X \circ \langle b \circ_T a, c \circ_T a\rangle \\ := \; & \rho((b;a)+(c;a)) \end{aligned}$$

□

It is straightforward to check that diagram (5) commutes for the maybe monad M and any of the 12 natural transformations $\mathsf{M} \times \mathsf{M} \to \mathsf{M}$ described in Corollary 9, the axiom K10 is thus valid for Kleisli M-representations. On the other hand, axiom K9 need not hold as the following example shows. Let $X = \{\top, \bot\}$, and consider the atomic programs a, b, c represented in $\mathrm{End}_{\mathbf{Kl}(\mathsf{M})}(X)$ by:

$$b\,(x) = \top, \qquad c\,(x) = \bot \qquad\qquad a\,(\top) = \top, \qquad a\,(\bot) = \star$$

Suppose moreover that $\alpha : \mathsf{M} \times \mathsf{M} \to \mathsf{M}$ is one of the natural transformations described in Corollary 9 which takes the left projection on the orbit defined by $(1, \emptyset)$. It then follows that

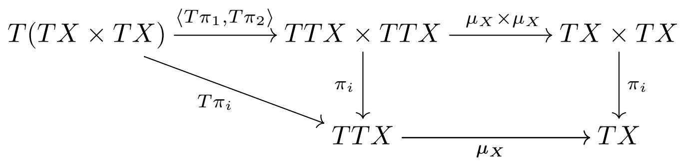
$$a \circ_{\mathsf{M}} (b \oplus c)\,(x) = \mu_X \circ \mathsf{M}a \circ \alpha_X\,(\top, \bot) = \top$$

but

$$(a \circ_{\mathsf{M}} b) \oplus (a \circ_{\mathsf{M}} c)\,(x) = \alpha_X(\top, \star) = \top \neq \star$$

The axiom K9 need not hold in the case of the list monad L either, for exactly the same reason.

It is routine to check that the monad P makes diagrams (4)-(5) commute for any of the 16 choices of α described in Corollary 14, and the axioms K9 and K10 are thus valid in all Kleisli P-representations. Similarly, it is straightforward to check that the diagrams in Theorem 22 also commute for the monads $\mathsf{B}_{\mathbb{N}}, \mathsf{D}$ and G for any choice of α. In each case, the result hinges on the fact that the multiplication of these monads is linear. Moreover, the natural projections $\pi_i : T \times T \to T$ make diagram (4) commute, and, interestingly, the diagram below also commutes.



This shows that diagram (5) commutes, and therefore that all natural projections $\pi_i : T \times T \to T$ respect the distributivity laws. We summarize our findings about distributivity in the following table:

Axiom	M	L	H	P	$B_{\mathbb{N}}$	D	G
K7: 0;a=0	✓	✓	✗	✓	✓	✗	✗
K8: a;0=0	✓	✓	✗	✓	✓	✗	✗
K9: Left dist.	✓	✓	✓	✓	✓	✓	✓
K10: Right dist.	✓	✓	✓	✓	✓	✓	✓
Comment:	K10 only holds for $\alpha_X = \underline{\star}$	K10 only holds for $\alpha_X = \underline{\varepsilon}$		K9-K10 are valid			

Table 2: Satisfiability of axioms K7-K10 in Kleisli representations

6 Representability of $(-)^*$

We now briefly turn our attention to the representation of the troublesome $(-)^*$ operation. Our approach will be to represent the strictly weaker (see [Koz90]) notion of $*$-continuous Kleene algebra, i.e idempotent commutative semirings which satisfy the infinitary equation

$$ab^*c = \sum_n ab^n c$$

where $b^0 = 1$ and $\sum$ is understood as the supremum for the order induced by the $+$ operation. To represent $*$-continuous Kleene algebras we suggest using a natural transformation $\sigma : T^\omega \to T$ to define the Kleisli representation of a term $(a)^*$ as

$$\rho((a)^*) := \sigma_X \circ \langle \eta_X, \rho(a), \rho(a) \circ_T \rho(a), \ldots \rangle.$$

Clearly, some restrictions on σ are desirable. But since we may want to drop some of the axioms of KAs - most notably idempotency in the probabilistic case - these restrictions will have to be tailored to the problem at hand. For now, let us just say that characterising the set

$$[\mathbf{Set}, \mathbf{Set}]\,(T^\omega, T)$$

is very similar to characterising the set $[\mathbf{Set}, \mathbf{Set}]\,(T \times T, T)$ and the techniques described in Section 4 can broadly speaking be re-used. However, one must be careful: the ω-fold product of a finitary functor may not be finitary, so moving to $\aleph_1$-accessible functors may be required to write the limit (1).

As the earlier sections have established, the powerset monad is our only example left standing from the confrontation with the axioms K3-K10. As is well-known it also resists the axioms K11-K14 by interpreting $\sigma : \mathsf{P}^\omega \to \mathsf{P}$ as the union.

Theorem 23 (Kleisli representability theorem)**.** *Kleene algebras (and thus also $*$-continuous Kleene algebras) have Kleisli representations for the powerset monad.*

It is clear that we have barely scratched the surface of a theory of fixed point equations in Kleisli representations, but this work is essentially meant as an initial foray, and we leave this very interesting question to further investigation.

7 Strategies around the no-go theorems

As we have seen in the previous sections, many well-known monads fail to represent at least one of the axioms K3-K10. A way out of these no-go theorems may be to combine two monads in the hope of getting a new monad which can represent a larger subset of the axioms K3-K10.

7.1 Strategy 1: combining with P

The powerset monad is very well-behaved with respect to all the axioms we have encountered so far. So combining a monad of interest with P is a strategy worth exploring. In fact, for every monad T one can define natural transformations

$$\alpha : \mathsf{P}T \times \mathsf{P}T \to \mathsf{P}T, \qquad 0 : \underline{1} \to \mathsf{P}T$$

in the expected manner:

$$\alpha_X(A, B) = A \cup B, \qquad 0_X(\star) = \emptyset$$

Moreover, it straightforward to show that these natural transformations respect the laws K3-K6. So, in principle, less well-behaved monads (like E, D) together with P could give rise to new monads that can be used to represent richer algebraic structures. Unfortunately, in what follows we will see that combining a given monad with P is not a simple task, in some cases even impossible.

7.1.1 Negative result for PD

As already shown in [Var03], there exists no distributive law $\mathsf{DP} \to \mathsf{PD}$. We now present a stronger result which answers a long standing open question (recently raised by Plotkin and Keimel in [KG17]).

Lemma 24. *The only natural transformations $\eta : \mathsf{Id} \to \mathsf{PD}$ are the constant transformation $\eta_X(x) = \emptyset$, and the natural transformation defined by $\eta_X(x) = \{\delta_x\}$.*

Proof. Follows from Corollary 3 and the fact that $\mathsf{PD}1$ has two elements: the empty set and the singleton $\{\delta_1\}$. □

By combining the result above, the monadic laws $\mu \circ \eta_T = \mu \circ T\eta = \mathsf{Id}$, and adapting the proof of [Var03] we can show that:

Theorem 25. *There is no monad structure on PD.*

7.1.2 Positive result for QE

On a more positive note, it is possible to combine the monad Q (non-empty powerset) with E using the following Theorem, and the notion of distributive law (see appendix).

Theorem 26. *There is a distributivity law $\delta : \mathsf{QE} \to \mathsf{EQ}$ defined by*

$$\delta_X(f, d) = \{(g, d) \in \mathsf{E}X \mid g \in f\}$$

where $g \in f$ is shorthand notation for the condition

$$\forall t \in \mathbb{R}^+ . \, g\,(t) \in f\,(t)$$

The proof of the theorem above also shows why it is not possible to combine P with E using this distributive law. This is a significant drawback because we can then use Corollary 3 to show that there is no natural transformation

$$\underline{1} \to \mathsf{QE}$$

which means that we cannot represent the annihilation axioms with QE. Another drawback is that combining monads gives rise to more complex structures which may be significantly more difficult to reason with than the original ones.

7.2 Strategy 2: combining with M.

Our second strategy also consists in combining a monad with another one which supports failure behaviour, this time M. Such a combination is always possible.

Theorem 27. *For every monad T in **Set** (**Top**, or **Pol**) there is a distributivity law δ : $\mathsf{M}T \to T\mathsf{M}$ defined as*

$$\delta_X = [Ti_1, \eta^T_{X+1} \circ i_2]$$

7.2.1 Some results for EM

As was shown by the second author *et al.* in [NBHM16], E : **Set** $\to$ **Set** has a natural transformation

$$\gamma : \mathsf{E} \times \mathsf{E} \to \mathsf{E}(- \times -)$$

which respects certain conditions (technically, those associated with monoidal functors). This additional structure induces a natural transformation

$$\beta : \mathsf{EM} \times \mathsf{EM} \to \mathsf{EM}, \qquad \beta = \mathsf{E}\alpha \circ \gamma_{\mathsf{M},\mathsf{M}}$$

where $\alpha : \mathsf{M} \times \mathsf{M} \to \mathsf{M}$ is an arbitrary natural transformation. Moreover, we have a natural transformation

$$\epsilon : \underline{1} \to \mathsf{EM}X, \qquad \epsilon = T0 \circ \eta^{\mathsf{E}}_{\underline{1}}$$

with $0 : \underline{1} \to \mathsf{M}$.

Remark 28. *This method to construct natural transformations $\beta : \mathsf{EM} \times \mathsf{EM} \to \mathsf{EM}$, ϵ : $\underline{1} \to \mathsf{EM}$ actually applies to any monad in **Set** with a natural transformation*

$$\gamma : T \times T \to T(- \times -)$$

essentially due to Theorem 27.

Theorem 27 and the last construction provides an interesting monad EM whose representation capabilities will be briefly explored next. We start with the following theorem for axioms K3-K6.

Theorem 29. *The axioms K3-K6 are satisfiable in Kleisli representations for the monad* EM.

We now proceed with a brief study of natural transformations $\underline{1} \to \mathsf{EM}$. According to Theorem 2, we have infinitely many natural transformations $\underline{1} \to \mathsf{EM}$ since $\mathsf{EM}\emptyset \cong [0, \infty]$. In this case, each natural transformation $\underline{1} \to \mathsf{EM}$ is a family of maps $\tau_X : \underline{1} \to \mathsf{EM}X, (\star) \mapsto (f, d)$ where $f : \mathbb{R}^+ \to \mathsf{M}X$ is the function that constantly outputs failure, and value d is an arbitrary duration. One interesting choice for the zero construct is the natural transformation $\tau : \underline{1} \to \mathsf{EM}$ that maps to duration zero. And, as expected, law K7 holds. Note, however, that law K8 does not hold if program a outputs any evolution with duration different than zero. Actually, whatever finite duration one chooses for $\tau : \underline{1} \to \mathsf{EM}$ the law will not hold for any program that outputs an evolution with a bigger duration than the chosen one. Interestingly, if the natural transformation $\tau : \underline{1} \to \mathsf{EM}$ chosen is the one that maps to duration infinite, then the same law holds for all programs whose evolutions do not fail in their last point. This means that law K8 is not representable in EM.

7.2.2 Some results for GM

The combination $\mathsf{GM} : \mathbf{Pol} \to \mathbf{Pol}$ (with M topologised in the obvious way) is a monad by Theorem 27. It behaves very similarly to G, but allows failure since $\mathsf{GM}\emptyset = \{\delta_\star\}$, and thus by Corollary 3 there exists a natural transformation $\underline{1} \to \mathsf{GM}$. Moreover, whilst $\mathbf{El}(\mathsf{G})$ has a single orbit, $\mathbf{El}(\mathsf{GM})$ has a collection of orbits labelled by $\lambda \in [0, 1]$. This makes the set of natural transformations $(\mathsf{GM})^2 \to \mathsf{GM}$ rather large, but using the same technique as in the proofs of Theorems 13 and 18 we can describe them concisely as follows.

Theorem 30. *The set* $[\mathbf{Pol}, \mathbf{Pol}]\,((\mathsf{GM})^2, \mathsf{GM})$ *is in one-to-one correspondence with continuous maps*

$$\phi : [0,1] \times [0,1] \to \{(r_1, r_2) \in [0,1]^2 \mid r_1 + r_2 \le 1\}$$

A continuous function ϕ as described above builds a natural transformation $\alpha^\phi : (\mathsf{GM})^2 \to \mathsf{GM}$ in the following way: for a pair of distributions (μ, ν) on $\mathsf{M}X$ (i.e. sub-distributions on X) in the orbit labelled by (λ_1, λ_2) (which can be taken to be the masses allocated to $\star$), $\alpha^\phi_X(\mu, \nu) = \pi_1(\phi(\lambda_1, \lambda_2))\mu + \pi_2(\phi(\lambda_1, \lambda_2))\nu$. Continuity is necessary because the topology on pairs of distribution induces a topology on the orbits of $\mathbf{El}(\mathsf{GM}^2)$, which is the usual topology on $[0, 1]^2$. One can build a continuous function ϕ such that that the associativity axiom K3 and the idempotency axiom K6 both fail, in sharp contrast with the validity of these axioms in Kleisli G-representations. It is also possible to define a function ϕ in such a way that K6 fails but K3+K4+K5 are satisfied simultaneously, or alternatively, in such a way that K3+K4+K5+K6 are all satisfied (by choosing $+^{1/2}$). Finally, note that axioms K8, K9 and K10 are valid for Kleisli GM-representations.

We summarize our findings about the monads QE, EM and GM in Table 3.

8 Conclusions and future work

We hope to have convinced the reader that the Kleisli representation approach to building program semantics is simple, practical and instructive. Simple because the underlying ideas are straightforward, commonly accepted, and their mathematical realisation is a well-known elementary concept in mathematics. Practical, because we have shown many examples of

Axiom	QE	EM	GM
K3: Associativity	✓	✓	✓
K4: Commutativity	✓	✓	✓
K5: Unit	✗	✓	✓
K6: Idempotency	✓	✓	✓
K7: 0;a = 0	✗	✓	✓
K8: a;0 = 0	✗	✗	✓
K9: Left dist.	✓	✓	✓
K10: Right dist.	✓	✓	✓
K3+K4+K5	✗	✓	✓
K3+K4+K6	✓	✗	✓
K3+K5+K6	✗	✓	✓
K3+K4+K5+K6	✗	✗	✓

Table 3: Satisfiability of axioms for hybrid and probabilistic monads combined by M

monads for which Kleisli representations of fragments of Kleene algebras can be classified and characterised in great detail. Finally, instructive because it offers a global and a priori perspective of what can and cannot be expected of program semantics. This perspective works both ways: given a target behaviour, we can make some a priori statements about which program constructs will be supported. For example, we can say that pure hybrid or probabilistic behaviours cannot accommodate `abort` instructions. Conversely, given a programming syntax, we can a priori eliminate or include behaviour types to interpret it adequately. For example, commutative and idempotent binary constructs cannot be interpreted in Kleisli representations for the L, M or H monads.

The work presented here is intended as an initial case study, and we envisage many further applications. Of particular interest is the thorny issue of concurrency. Whilst the exact nature of concurrent Kleene algebras is not entirely settled, some algebraic properties of the $\|$ operator are well-established and the Kleisli representation approach could be used to allow or rule out certain semantics. Dually, algebraic properties of the $\|$ operator could be inferred from well-established semantics by investigating valid axioms in the corresponding Kleisli representations.

The theory of Kleisli representations needs to be developed much further, and two areas of future research stand out in particular. First, we wish to study Kleisli representations over state spaces with some algebraic structure. This is the approach recently taken in the elaboration of a semantics for ProbNetKAT in [FKM+16] and [SKF+17], where the state space is a free semilattice. When a state space X is endowed with a binary operation, say $\odot_X$, and the monad T is monoidal with a 'Fubini transformation' $\otimes : T \times T \to T(- \times -)$, then a new transformation $T \times T \to T$ emerges, given by $T\odot_X \circ \otimes_X$, which is natural w.r.t. to $\odot$-homomorphism (i.e. maps f such that $\odot_X \circ f \times f = f \circ \odot_X$). This is precisely what is done in *op. cit.*, and we believe that understanding this procedure in general is important.

Secondly, representation theory itself suggests many avenues of research. First of all the mathematical notion of a *faithful representation* is essentially the same as the computer science notion of *completeness*. It seems likely that the many proofs of existence of faithful representations could provide inspiration for the construction of completeness proofs. Similarly, the key notion of *irreducible representation* echoes the modularity mantra of computer science: can we isolate 'minimal' Kleisli representations such that every other representation is a combination of minimal ones?

References

[AB06] C. D. Aliprantis and K. Border, *Infinite dimensional analysis: a hitchhiker's guide*, Springer, 2006.

[AGT10] J. Adámek, H P. Gumm, and V. Trnková, *Presentation of set functors: a coalgebraic perspective*, Journal of Logic and Computation **20** (2010), no. 5, 991–1015.

[AR94] J. Adámek and J. Rosicky, *Locally presentable and accessible categories*, vol. 189, Cambridge University Press, 1994.

[DDG16a] F. Dahlqvist, V. Danos, and I. Garnier, *Giry and the machine*, MFPS XXXII, ENTCS, vol. 325, 2016, pp. 85–110.

[DDG16b] ______, *Robustly parameterised higher-order probabilistic models*, CONCUR, LIPIcs, vol. 59, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2016, pp. 23:1–23:15.

[Dob12] E. Doberkat, *A stochastic interpretation of propositional dynamic logic: expressivity*, J. Symb. Log. **77** (2012), no. 2, 687–716.

[FKM+16] N. Foster, D. Kozen, K. Mamouras, M. Reitblatt, and A. Silva, *Probabilistic NetKAT*, ESOP 2016, LNCS, Springer, 2016, pp. 282–309.

[Gir82] M. Giry, *A categorical approach to probability theory*, Categorical aspects of topology and analysis, Springer, 1982, pp. 68–85.

[HK15] H. H. Hansen and C. Kupke, *Weak completeness of coalgebraic dynamic logics*, FICS 2015, EPTCS, vol. 191, 2015, pp. 90–104.

[HKL14] H. H. Hansen, C. Kupke, and R. A. Leal, *Strong completeness for iteration-free coalgebraic dynamic logics*, TCS 2014, Proceedings, LNCS, vol. 8705, Springer, 2014, pp. 281–295.

[Höf09] P. Höfner, *Algebraic calculi for hybrid systems*, Ph.D. thesis, University of Augsburg, 2009.

[Jac00] B. Jacobs, *Object-oriented hybrid systems of coalgebras plus monoid actions*, Th. Comp. Sci. **239** (2000), no. 1, 41–95.

[JS91] A. Joyal and R. Street, *An introduction to Tannaka duality and quantum groups*, Category theory, Springer, 1991, pp. 413–492.

[KG17] Keimel K. and Plotkin G., *Mixed powerdomains for probability and non determinism*, To Appear in LMCS (2017).

[Koz85] D. Kozen, *A probabilistic* PDL, J. Comp. Sys. Sci. **30** (1985), no. 2, 162–178.

[Koz90] D. Kozen, *On Kleene algebras and closed semirings*, Int. Symp. on Math. Found. of Comp. Sc., Springer, 1990, pp. 26–47.

[MM12] S. MacLane and I. Moerdijk, *Sheaves in geometry and logic: A first introduction to topos theory*, Springer, 2012.

[NB16] R. Neves and L. S. Barbosa, *Hybrid automata as coalgebras*, ICTAC 2016, LNCS, vol. 9965, 2016, pp. 385–402.

[NBHM16] R. Neves, L. S. Barbosa, D. Hofmann, and M. A. Martins, *Continuity as a computational effect*, Journal of Logical and Algebraic Methods in Programming (2016).

[Pla10] A. Platzer, *Logical analysis of hybrid systems: Proving theorems for complex dynamics*, Springer, Heidelberg, 2010.

[PP01] G. Plotkin and J. Power, *Semantics for algebraic operations*, ENTCS **45** (2001), 332–345.

[SKF+17] S. Smolka, P. Kumar, N. Foster, D. Kozen, and A. Silva, *Cantor meets Scott*, POPL 2017, ACM, 2017, pp. 557–571.

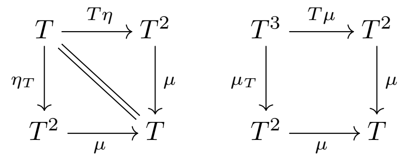
[TP97] D. Turi and G. Plotkin, *Towards a mathematical operational semantics*, LICS 1997, IEEE, 1997, pp. 280–291.

[Var03] Daniele Varacca, *Probability, nondeterminism and concurrency: two denotational models for probabilistic computation*, Ph.D. thesis, BRICS, 2003.

Appendix

8.1 Monads

We briefly recall the definition of a monad and of the Kleisli category associated with a monad. A monad T on a category $\mathbf{C}$ is an endofunctor $T : \mathbf{C} \to \mathbf{C}$ together with a natural transformation $\mu : T^2 \to T$ called the *multiplication of the monad* and a natural transformation $\eta : \mathsf{Id} \to T$ called the *unit of the monad* which satisfy the following commutative diagrams.

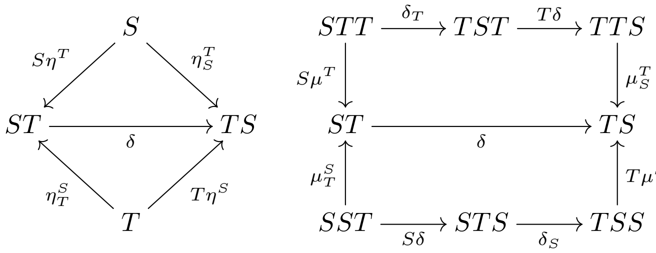


The *Kleisli category* associated with a monad T, and denoted $\mathbf{Kl}(T)$, is the category whose objects are the same as those of $\mathbf{C}$ and whose morphisms $f : X \to Y$ are the morphisms in $\mathbf{C}$ of the form $f : X \to TY$ with the identities $\mathrm{id}_X : X \to X$ being given by the unit η_X of the monad and the composition of Kleisli arrow $f : X \to Y$ and $g : Y \to Z$, which we will denote as $\circ_T$, being defined by $f \circ_T g = \mu_Z \circ Tg \circ f : X \to Z$. The crucial observation for our purpose is that for any object X, the set $\mathrm{End}_{\mathbf{Kl}(T)}(X)$ of endomorphisms of X in $\mathbf{Kl}(T)$ together with the Kleisli composition $\circ_T$ and the unit η_X forms a monoid.

The standard technique to combine two monads (T, η^T, μ^T), (S, η^S, μ^S) resorts to a so called distributive law. More concretely a natural transformation

$$\delta : ST \to TS$$

that makes the following diagrams to commute.



It allows to define the monad $(TS, \eta^{TS}, \mu^{TS})$ where

$$\eta^{TS} = \eta^T_S \circ \eta^S$$
$$\mu^{TS} = \mu^T_S \circ TT\mu^S \circ T\delta_S$$

8.2 Proofs

Proof of Theorem 8. We could compute the limit of Eq. (1), but since T is already 'presented' as a colimit of hom functors, it is easier to take advantage of this presentation and proceed as follows (moreover, it allows us to avoid the question of whether T is finitary or not). In **Set** binary products distribute over arbitrary coproducts. It follows that

$$\begin{aligned} T \times T &= \left(\coprod_{i \in I} \hom(X_i, -)\right) \times \left(\coprod_{j \in I} \hom(X_j, -)\right) \\ &= \coprod_{i,j \in I} (\hom(X_i, -) \times \hom(X_j, -)) \\ &= \coprod_{i,j \in I} \hom(X_i + X_j, -) \end{aligned}$$

where the last step follows from the fact that the contravariant hom functor sends colimits to limits. It follows that

$$\begin{aligned} &[\mathbf{Set}, \mathbf{Set}]\,(T \times T, T) \\ &= [\mathbf{Set}, \mathbf{Set}]\,(\coprod_{i,j \in I} \hom(X_i + X_j, -), T) \\ &= \prod_{i,j \in I} [\mathbf{Set}, \mathbf{Set}]\,(\hom(X_i + X_j, -), T) \\ &\cong \prod_{i,j \in I} T(X_i + X_j) \end{aligned}$$

where the last step is an application of the Yoneda lemma. Given an element $s \in \prod_{i,j \in I} T(X_i + X_j)$ with components $s_{ij} \in T(X_i + X_j)$, the associated natural transformation α^s is defined at each Y by

$$(a, b) \in \hom(X_i, Y) \times \hom(X_j, Y) \mapsto T\,[a, b]\,(s_{ij}) \tag{6}$$

where $[a, b]$ is the coproduct map $X_i + X_j \to Y$.

Commutativity. Let $i, j \in I$ and assume that $X_i, X_j \neq \emptyset$. It follows from condition (i) and Eq. (6) that for any $a : X_i \to Y$ and $b : X_j \to Y$

$$\alpha^s(a, b) = T[a, b](s_{ij}) = \alpha^s(b, a) = T[b, a](s_{ji})$$

since in either case we get the unique map $X_k = \emptyset \to Y$. Suppose now that $X_i = \emptyset$ and that (ii) holds, it then follows once again from Eq. (6) that $\alpha^s(a, b) = \alpha^s(b, a)$ since $s_{ij} = s_{ji} : X_k \to X_i + X_j \cong X_j$. Thus α^s is commutative.

Conversely, assume that α^s is commutative, and let $i, j \in I$. If $X_i = \emptyset$ then by commutativity of α^s and the fact that $X_j \cong X_j + X_i$ we have $s_{ij} := \alpha^s(\mathrm{id}_{X_i}, \mathrm{id}_{X_j}) = \alpha^s(\mathrm{id}_{X_j}, \mathrm{id}_{X_i}) := s_{ji}$. Now assume that $X_i, X_j \neq \emptyset$. We show that (i) must hold by contradiction. Suppose that $s_{ij} \in \hom(X_k, X_i + X_j)$ with X_k of cardinality at least one. By commutativity if we choose the injections $i_1 : X_i \to X_i + X_j$, $i_2 : X_j \to X_i + X_j$ we get

$$s_{ij} := \alpha^s(i_1, i_2) = \alpha^s(i_2, i_1) = [i_2, i_1] \circ s_{ji}$$

i.e. $s_{ji} : X_k \to X_j + X_i$ is the same function as s_{ij} modulo permutation. We assume w.l.o.g. that $|X_i| \leq |X_j|$, i.e. that there exists an injection $m : X_i \rightarrowtail X_j$ and consider the maps $a = i_1 \circ m : X_i \to X_j + X_j$, and $b = i_2 : X_j \to X_j + X_j$. We also define the permutation $\phi : X_j + X_j \to X_j + X_j$ which maps every element of the first component of the coproduct to its counterpart in the second component, and vice versa. Note that no element is fixed under the action of ϕ. We now have

$$\begin{aligned}\alpha^s(a,b) &= \alpha^s(b,a)\\ &= [b,a] \circ [i_2, i_1] \circ s_{ij}\\ &= \phi \circ [a,b] \circ s_{ij}\\ &= \phi \circ \alpha^s(a,b)\end{aligned}$$

which is impossible since for any x, $\phi(x) \neq x$ and s_{ij} having domain of cardinality at least one will will pick at least one such x. Thus s_{ij} must be an empty map $X_k = \emptyset \to X_i + X_j$.

Idempotency. Assume that α^s is idempotent. Each $a \in TY$ lands in a certain component $\hom(X_i, Y), i \in I$, and from the construction above, $\alpha^s_Y(a,a) = T[a,a](s_{ii})$ where $s_{ii} \in T(X_i + X_i)$. By idempotency we clearly need $\alpha^s_Y(a,a) \in \hom(X_i, Y)$ too, and it follows that s_{ii} must be a map $X_i \to X_i + X_i$ such that $[a,a] \circ s_{ii} = a$ for any $a : X \to Y$. In particular for $a = \mathrm{id}_{X_i}$, and the result follows. The converse direction is clear.

Unit. It follows from Corollary 3 that if $T\emptyset = \emptyset$ we cannot interpret 0 at all, let alone have it as a unit. Now assume that $T\emptyset \neq \emptyset$, by our assumption on the shape of T it means that there exists $i \in I$ with $X_i = \emptyset$ and that we interpret 0 via $\zeta : \underline{1} \to \hom(X_i, -)$. For 0 to be a unit for + we need for each $a \in \hom(X_j, Y) \subseteq TY$ that $\alpha_Y(a,0) = \alpha_Y(0,a) = a$. It follows from the construction of α_Y described above that

$$\alpha_Y(a,0) = T[a,0](s_{ji}) = \alpha_Y(0,a) = T[a,0](s_{ij}) = a$$

Since $T[a,0](s_{ji}) = [a,0] \circ s_{ji}$ it must be the case that $s_{ji} : X_j \to X_j + \emptyset = X_j$ and since $[a,0] \circ s_{ji} = a$ for any $a : X_j \to Y$ it is easy to see by considering injective maps a that $s_{ji} = \mathrm{id}_{X_j}$. Similarly $s_{ij} = \mathrm{id}_{X_j}$. ☐

Proof of Corollary 9. The identity functor is representable in **Set** as $\hom(1, -)$, whilst the constant functor $\underline{1}$ is representable as $\hom(\emptyset, -)$ since $\emptyset$ is initial. It follows that

$$\mathsf{M} = \hom(1, -) + \hom(\emptyset, -).$$

And thus by Theorem 8 we have

$$[\mathbf{Set}, \mathbf{Set}]\,(\mathsf{M} \times \mathsf{M}, \mathsf{M}) \cong \mathsf{M}(2) \times \mathsf{M}(1) \times \mathsf{M}(1) \times \mathsf{M}(\emptyset)$$

By Theorem 8 we also have that commutative transformations must pick $s_{11} = \{*\} \in \hom(\emptyset, 1+1)$. By Theorem 8 we know that for a natural transformation $\mathsf{M} \times \mathsf{M} \to \mathsf{M}$ to be idempotent it must be defined by a tuple such that the element of $\mathsf{M}(2) = \hom(1,2) + \hom(\emptyset, 2)$ is a map $1 \to 2$. The condition for the unit follows easily from Theorem 8. ☐

Proof of Theorem 13. Consider any regular cardinal $\kappa > \lambda$, and the κ-accessible version P_κ of P (taking powersets of cardinality less that κ, see [AR94]). Theorem 7 generalises completely straightforwardly to κ-accessible functors, and $[\mathbf{Set}, \mathbf{Set}]\,((\mathsf{P}_\kappa)^\lambda, \mathsf{P}_\kappa)$ is thus given by the limit (1) (with the inclusion functor I suitably modified). To compute this limit, consider a set X and a collection $(U_i)_{i\in\lambda}$ of subsets of X. It is easy to see, by considering what happens at the singleton 1, that $\mathbf{El}((\mathsf{P}_\kappa)^\lambda)$ has 2^λ-orbits; one for each element of $(\mathsf{P}_\kappa 1)^\lambda$. The object $(X, (U_i)_{i\in\lambda})$ in $\mathbf{El}((\mathsf{P}_\kappa)^\lambda)$ belongs to the orbit determined by $(1, (!_X[U_i])_{i\in\lambda})$ where $!_X : X \to 1$, i.e. to the orbit determined by a subset of indices $J \in 2^\lambda$.

We will show that any thread in the limit (1) must pick an element $\bigcup_{k\in K} U_k$ in the copy of $\mathsf{P}_\kappa(X)$ corresponding to $(X,(U_i)_{i\in\lambda})$, for some subset of indices $K \in 2^\lambda$ such that $i \notin K$ whenever $U_i = \emptyset$, and that this choice must be made consistently across the orbit. This will prove that a natural transformation $(\mathsf{P}_\kappa)^\lambda \to \mathsf{P}_\kappa$ is entirely determined by non-increasing maps $2^\lambda \to 2^\lambda$ (mapping J to K).

To prove the claim above consider the object

$$\left(\biguplus_i U_i \times \lambda, \left(\biguplus_i U_i \times \{\epsilon_i\}\right)_{i\in\lambda}\right) \text{ in } \mathbf{El}((\mathsf{P}_\kappa)^\lambda) \tag{7}$$

where $\epsilon_i = i$ if $U_i \neq \emptyset$ and $\emptyset$ else. It belongs to the same orbit as $(X,(U_i)_{i\in\lambda})$ as it is connected to it by the map $f : \biguplus_i U_i \times \lambda \to X$, defined by

$$(u,i) \mapsto \begin{cases} u & \text{if } u \in U_i \\ \text{any } u_i^0 \in U_i & \text{if } u \notin U_i, U_i \neq \emptyset \\ \text{anything} & \text{else} \end{cases}$$

Note now that the object (7) is invariant under all endomorphisms which keep the λ-component constant. This means that for any thread in the limit (1), the component corresponding to the object (7) must contain $\biguplus_i U_i \times \epsilon_i$, i.e. that it will be a union

$$\bigcup_{k\in K}\left(\biguplus_i U_i \times \epsilon_k\right)$$

over some $K \subseteq \lambda$. Moreover, since $\epsilon_k = \emptyset$ when $U_k = \emptyset$ we do indeed have that $J \geq K$ (for the obvious order on 2^λ). By pushing this component of the thread to the $(X,(U_i)_{i\in\lambda})$-component with f, we indeed get $\bigcup_{k\in K} U_k$ as claimed. It remains to check that the choice of $K \subseteq \lambda$ must be made consistently across the orbit. For this consider another object $(X,(V_i)_{i\in\lambda})$ in the same orbit (i.e. $V_i = \emptyset$ iff $U_i = \emptyset$). We can always build the object

$$(X^2, (U_i \times V_i)_{i\in\lambda}) \text{ in } \mathbf{El}((\mathsf{P}_\kappa)^\lambda)$$

which gets mapped to $(X,(U_i)_{i\in\lambda})$ and $(X,(V_i)_{i\in I})$ by the projections maps $\pi_1,\pi_2 : X^2 \to X$. If the $(X,(U_i)_{i\in\lambda})$-component of a thread in the limit (1) is $\bigcup_{k\in K} U_k$, then the $(X^2,(U_i \times V_i)_{i\in\lambda})$-component of the thread must clearly be $\bigcup_{k\in K} U_k \times V_k$, and so the $(X,(V_i)_{i\in I})$-component must be $\bigcup_{k\in K} V_k$.

Finally we need to show that our result for holds for the full powerset monad P. Suppose for the sake of contradiction that $\alpha : (\mathsf{P})^\lambda \to \mathsf{P}$ is not one of the transformations described above, then this must be witnessed at a set X, i.e. there must exist $(U_i)_{i\in\lambda} \in (\mathsf{P}X)^\lambda$ such that $\alpha_X((U_i)_{i\in\lambda})$ is not given by one of the transformations above. But since we can always find a regular cardinal such that $\mathsf{P}_\kappa X = \mathsf{P}X$, this would define a natural transformation $(\mathsf{P}_\kappa)^\lambda \times \mathsf{P}_\kappa \to \mathsf{P}_\kappa$ which is not of the form described above, a contradiction. □

Proof of Theorem 18. The proof that $+^\lambda$ is natural is routine. Recall that the Machine of [DDG16a] tells us that Eq. (2) holds. We show that if $\alpha \in [\mathbf{Pol}_f, \mathbf{Pol}_f](\mathsf{G}_f \times \mathsf{G}_f, \mathsf{G}_f)$ is natural then at every finite space k, the restriction of α_k to rational probability distributions is of the type $+^\lambda$ for some $\lambda \in [0,1]$. The result then extends to any pair of probabilities by continuity and the fact that objects in $\mathbf{Pol}$ are complete.

To start off, we show that

$$\alpha_n\left(\left(\frac{1}{n},\ldots,\frac{1}{n}\right),\left(\frac{1}{n},\ldots,\frac{1}{n}\right)\right) = \left(\frac{1}{n},\ldots,\frac{1}{n}\right) \tag{8}$$

Since $\left(\left(\frac{1}{n},\ldots,\frac{1}{n}\right),\left(\frac{1}{n},\ldots,\frac{1}{n}\right)\right)$ is invariant under any permutation $\phi : n \to n$, it follows by naturality that so must $\alpha_n\left(\left(\frac{1}{n},\ldots,\frac{1}{n}\right),\left(\frac{1}{n},\ldots,\frac{1}{n}\right)\right)$, and equation (8) follows.

Now let $(\mu,\nu) \in \mathsf{G}_f X \times \mathsf{G}_f X$ be a pair of rational probability distributions, we can assume w.l.o.g. that

$$(\mu,\nu) = \left(\left(\frac{p_1}{n},\ldots,\frac{p_k}{n}\right),\left(\frac{q_1}{n},\ldots,\frac{q_k}{n}\right)\right)$$

where $p_i, q_i, 1 \le i \le k$ are integers. This pair of distributions can be expressed as the image of $(\tilde{\mu},\tilde{\nu}) \in \mathsf{G}_f(2n) \times \mathsf{G}_f(2n)$ given by:

$$\left(\left(\frac{1}{n},\ldots,\frac{1}{n},0,\ldots,0\right),\left(0,\ldots,0,\frac{1}{n},\ldots,\frac{1}{n}\right)\right)$$

under the map $f : 2n \to k$,

$$i \mapsto \begin{cases} 1 & \text{if } i \le p_1 \\ j & \text{if } 1+\sum_{l=1}^{j-1} p_l \le i \le \sum_{l=1}^{j} p_l \\ 1 & \text{if } n+1 \le n+q_1 \\ j & \text{if } n+1+\sum_{l=1}^{j-1} q_l \le i \le n+\sum_{l=1}^{j} q_l \end{cases}$$

i.e. $(\mu,\nu) = (\mathsf{G}f \times \mathsf{G}f)(\tilde{\mu},\tilde{\nu})$. By considering all the permutations (ij) and $((n+i)(n+j))$ with $1 \le i,j \le n$, naturality enforces that $\alpha_{2n}(\tilde{\mu},\tilde{\nu})$ is of the shape

$$\alpha_{2n}(\tilde{\mu},\tilde{\nu}) = (\underbrace{\lambda,\ldots,\lambda}_{n},\underbrace{\lambda',\ldots,\lambda'}_{n})$$

for some $\lambda,\lambda' \in [0,1]$.

Consider now the map $g : 2n \to n$ mapping i and $n+i$ to i, $1 \le i \le n$. It follows by naturality and Eq. (8) that

$$\begin{aligned} \left(\frac{1}{n},\ldots,\frac{1}{n}\right) &= \alpha_n\left(\left(\frac{1}{n},\ldots,\frac{1}{n}\right),\left(\frac{1}{n},\ldots,\frac{1}{n}\right)\right) \\ &= \alpha_n(\mathsf{G}g \times \mathsf{G}g(\tilde{\mu},\tilde{\nu}) \\ &= \mathsf{G}g(\alpha_{2n}(\tilde{\mu},\tilde{\nu})) \\ &= \mathsf{G}g(\lambda,\ldots,\lambda,\lambda',\ldots,\lambda') \\ &= (\lambda+\lambda',\ldots,\lambda+\lambda') \end{aligned}$$

It follows that

$$\frac{1}{n} = \lambda + \lambda'$$

and thus $\lambda' = \frac{1}{n} - \lambda$. Another application of naturality gives

$$\begin{aligned} \alpha_k(\mu,\nu) &= \alpha_k(\mathsf{G}f \times \mathsf{G}f(\tilde{\mu},\tilde{\nu}) \\ &= \mathsf{G}f(\alpha_{2n}(\tilde{\mu},\tilde{\nu})) \\ &= \mathsf{G}f(\lambda,\ldots,\lambda,\frac{1}{n}-\lambda,\ldots,\frac{1}{n}-\lambda) \\ &= \left(p_1\lambda + q_1\left(\frac{1}{n}-\lambda\right),\ldots,p_k\lambda + q_k\left(\frac{1}{n}-\lambda\right)\right) \\ &= \left(n\lambda\frac{p_1}{n} + (1-n\lambda)\frac{q_1}{n},\ldots,n\lambda\frac{p_k}{n} + (1-n\lambda)\frac{q_k}{n}\right) \\ &= \mu +_k^{n\lambda} \nu \end{aligned}$$

It remains to verify that $n\lambda$ does not depend on (μ, ν).

Let $(\mu', \nu') \in \mathsf{G}k \times \mathsf{G}k$. We can w.l.o.g. write it as

$$(\mu', \nu') = \left(\left(\frac{p'_1}{m}, \ldots, \frac{p'_k}{m} \right), \left(\frac{q'_1}{m}, \ldots, \frac{q'_k}{m} \right) \right)$$

We can represent both (μ, ν) and (μ', ν') as images of the element $(\tilde{\mu}, \tilde{\nu}) \in \mathsf{G}2mn \times \mathsf{G}2mn$ given by

$$\left(\left(\frac{1}{mn}, \ldots, \frac{1}{mn}, \underbrace{0, \ldots, 0}_{mn} \right), \left(\underbrace{0, \ldots, 0}_{mn}, \frac{1}{mn}, \ldots, \frac{1}{mn} \right) \right)$$

under maps $f : 2mn \to k$

$$i \mapsto \begin{cases} 1 & \text{if } i \leq mp_1 \\ j & \text{if } 1 + \sum_{l=1}^{j-1} mp_l \leq i \leq \sum_{l=1}^{j} mp_l \\ 1 & \text{if } mn + 1 \leq n + mq_1 \\ j & \text{if } mn + 1 + \sum_{l=1}^{j-1} mq_l \leq i \leq n + \sum_{l=1}^{j} mq_l \end{cases}$$

and $f : 2n \to k$ defined similarly (but with the roles of m and n reversed). The commutativity and idempotency argument used above can be re-used to show that

$$\alpha_{2mn}(\tilde{\mu}, \tilde{\nu}) = (\lambda', \ldots, \lambda', \frac{1}{mn} - \lambda', \ldots, \frac{1}{mn} - \lambda')$$

for some $\lambda' \in [0, 1]$. It follows by naturality that

$$\begin{aligned} \alpha_k(\mu, \nu) &= \Big(mp_1 \lambda' + mq_1 \left(\frac{1}{mn} - \lambda' \right), \ldots, \\ &\qquad mp_k \lambda' + mq_k \left(\frac{1}{mn} - \lambda' \right) \Big) \\ &= \Big(mn\lambda' \frac{p_1}{n} + (1 - \lambda' mn) \frac{q_1}{n}, \ldots, \\ &\qquad mn\lambda' \frac{p_k}{n} + (1 - \lambda' mn) \frac{q_k}{n} \Big) \end{aligned}$$

and thus $\lambda' = \frac{\lambda}{m}$. We now have

$$\begin{aligned} \alpha_k(\mu', \nu') =& \mathsf{G}f(\alpha_{2mn}(\tilde{\mu}, \tilde{\nu})) \\ =& \Big(np'_1 \lambda' + nq'_1 \left(\frac{1}{mn} - \lambda' \right), \ldots, \\ &\qquad np'_k \lambda' + nq'_k \left(\frac{1}{mn} - \lambda' \right) \Big) \\ =& \Big(n\lambda \frac{p'_1}{m} + (1 - n\lambda) \frac{q_1}{m}, \ldots, \\ &\qquad n\lambda \frac{p'_k}{m} + (1 - n\lambda) \frac{q_k}{m} \Big) \\ =& \mu' +_k^{n\lambda} \nu' \end{aligned}$$

□

Proof of Theorem 25. We proceed by contradiction. Assume that there exists a unit transformation $\eta : \mathsf{Id} \to \mathsf{PD}$ and a multiplication transformation $\mu : (\mathsf{PD})^2 \to \mathsf{PD}$ for which PD is a monad. By definition of a monad we must have that at any X

$$\mu_X \circ \eta_{\mathsf{PD}X} = \mathrm{id}_{\mathsf{PD}X} \tag{9}$$

and similarly,

$$\mu_X \circ \mathsf{PD}\eta_X = \mathrm{id}_{\mathsf{PD}X}. \tag{10}$$

This set of equations gives us the action of μ_X on very specific inputs, and we will see that it is enough to generate a contradiction.

As shown in Lemma 24, η is either the constant natural transformation to $\emptyset$ or is defined at $x \in X$ by $\eta_X(x) = \{\delta_x\}$. Assume first that η is the constant natural transformation to $\emptyset$, and let X be any non-empty set. Since the cardinality of $\mathsf{PD}X$ is then at least 2, it is clear that there cannot exist a function $\mu_X : \mathsf{PDPD}X \to \mathsf{PD}X$ such that (Eq. 9) holds for any $U \in \mathsf{PD}X$:

$$\mu_X \circ \eta_{\mathsf{PD}X}(U) = \mu_X(\emptyset) = U.$$

We therefore immediately get a contradiction if we assume that η is the constant natural transformation to $\emptyset$.

Next we assume that $\eta_X(x) = \{\delta_x\}$ and follow an argument due to Plotkin. Consider the sets $X = \{a, b, c, d\}$ and $Y = \{a, b\}$, the map $f : X \to Y$ defined by $f(a) = f(c) = a, f(b) = f(d) = b$ and the element $\{\frac{1}{2}\delta_{\{\delta_a,\delta_b\}} + \frac{1}{2}\delta_{\{\delta_c,\delta_d\}}\} \in \mathsf{PDPD}X$. It is straightforward to compute that

$$\begin{aligned}\mathsf{PDPD}f\left(\left\{\frac{1}{2}\delta_{\{\delta_a,\delta_b\}} + \frac{1}{2}\delta_{\{\delta_c,\delta_d\}}\right\}\right) &= \{\delta_{\{\delta_a,\delta_b\}}\}\\ &= \eta_{\mathsf{PD}Y}(\{\delta_a, \delta_b\})\end{aligned}$$

It now follows by naturality and Eq. (9) that

$$\begin{array}{ccc}\{\frac{1}{2}\delta_{\{\delta_a,\delta_b\}} + \frac{1}{2}\delta_{\{\delta_c,\delta_d\}}\} & \longmapsto & \mu_X(\{\frac{1}{2}\delta_{\{\delta_a,\delta_b\}} + \frac{1}{2}\delta_{\{\delta_c,\delta_d\}}\})\\ \downarrow & & \downarrow\\ \{\delta_{\{\delta_a,\delta_b\}}\} & \longmapsto & \mu_Y(\eta_{\mathsf{PD}Y}(\{\delta_a,\delta_b\})) = \{\delta_a,\delta_b\}\end{array}$$

It follows that any distributions in $\mu_X(\{\frac{1}{2}\delta_{\{\delta_a,\delta_b\}} + \frac{1}{2}\delta_{\{\delta_c,\delta_d\}}\})$ must belong to the preimage of $\{\delta_a, \delta_b\}$ under $\mathsf{PD}f$, i.e. to

$$\{p\delta_a + (1-p)\delta_c \mid p \in [0,1]\} \cup \{p\delta_b + (1-p)\delta_d \mid p \in [0,1]\}$$

By considering the map $g : X \to Y$ defined by $g(a) = g(d) = a, g(b) = g(c) = b$, we also get

$$\begin{aligned}\mathsf{PDPD}g\left(\left\{\frac{1}{2}\delta_{\{\delta_a,\delta_b\}} + \frac{1}{2}\delta_{\{\delta_c,\delta_d\}}\right\}\right) &= \{\delta_{\{\delta_a,\delta_b\}}\}\\ &= \eta_{\mathsf{PD}Y}(\{\delta_a, \delta_b\})\end{aligned}$$

and thus by naturality and Eq. (9), any distribution in $\mu_X(\{\frac{1}{2}\delta_{\{\delta_a,\delta_b\}} + \frac{1}{2}\delta_{\{\delta_c,\delta_d\}}\})$ must also belong to the preimage of $\{\delta_a, \delta_b\}$ under $\mathsf{PD}g$, i.e. to

$$\{p\delta_a + (1-p)\delta_d \mid p \in [0,1]\} \cup \{p\delta_b + (1-p)\delta_c \mid p \in [0,1]\}$$

It follows that $\mu_X(\{\frac{1}{2}\delta_{\{\delta_a,\delta_b\}} + \frac{1}{2}\delta_{\{\delta_c,\delta_d\}}\})$ contains at most the elements

$$\{\delta_a, \delta_b, \delta_c, \delta_d\}$$

Now consider the map $h : X \to Z := \{a, c\}$ defined by $h(a) = h(b) = a, h(c) = h(d) = c$. By definition we have

$$\begin{aligned}\mathsf{PDPD}h\left(\left\{\frac{1}{2}\delta_{\{\delta_a,\delta_b\}} + \frac{1}{2}\delta_{\{\delta_c,\delta_d\}}\right\}\right) =& \left\{\frac{1}{2}\delta_{\{\delta_a\}} + \frac{1}{2}\delta_{\{\delta_c\}}\right\} \\ =& \mathsf{PD}\eta_Y(\{\frac{1}{2}\delta_a + \frac{1}{2}\delta_c\})\end{aligned}$$

It now follows by naturality and Eq. (10) that

$$\begin{array}{ccc} \{\frac{1}{2}\delta_{\{\delta_a,\delta_b\}} + \frac{1}{2}\delta_{\{\delta_c,\delta_d\}}\} & \longmapsto & \mu_X(\{\frac{1}{2}\delta_{\{\delta_a,\delta_b\}} + \frac{1}{2}\delta_{\{\delta_c,\delta_d\}}\}) \\ \downarrow & & \downarrow \\ \{\frac{1}{2}\delta_{\{\delta_a\}} + \frac{1}{2}\delta_{\{\delta_c\}}\} & \longmapsto & \mu_Y(\mathsf{PD}\eta_Y(\{\frac{1}{2}\delta_a + \frac{1}{2}\delta_c\})) = \\ & & \{\frac{1}{2}\delta_a + \frac{1}{2}\delta_c\} \end{array}$$

and we immediately get a contradiction since $\mathsf{PD}h(\{\delta_a, \delta_b, \delta_c, \delta_d\}) = \{\delta_a, \delta_c\}$ and $\{\frac{1}{2}\delta_a + \frac{1}{2}\delta_c\} \notin \{\delta_a, \delta_c\}$. ☐

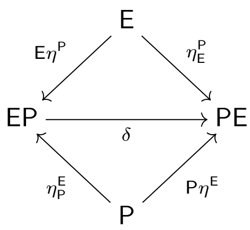
Proof of Theorem 26. First observe that $\delta : \mathsf{EQ} \to \mathsf{QE}$ can be generalised to $\delta : \mathsf{EP} \to \mathsf{PE}$. Moreover, the latter is natural because it is a composite of natural transformations

$$\begin{aligned}&\mathsf{EP} \to (\mathsf{P}\ \_\ )^{\mathbb{R}^+} \times [0, \infty] \to \mathsf{P}\left((\ \_\ )^{\mathbb{R}^+}\right) \times [0, \infty] \to \\ &\mathsf{P}\left((\ \_\ )^{\mathbb{R}^+} \times [0, \infty]\right) \to \mathsf{PE}\end{aligned}$$

the first and the last being, respectively, the obvious inclusion and quotient maps. The middle one corresponds to the map

$$(A_i)_{i\in I} \mapsto \prod_{i\in I} A_i.$$

We will now show that the natural transformation $\delta : \mathsf{EP} \to \mathsf{PE}$ makes the following diagram to commute.



Start with the upper triangle.

$$\begin{aligned} & \delta_X \circ \mathsf{E}\eta^{\mathsf{P}}_X \\ = \ & \delta_X \circ (\eta^{\mathsf{P}} \circ \times \mathrm{id}) \\ = \ & \eta^{\mathsf{P}}_{\mathsf{E}X} \end{aligned}$$

Then proceed with the lower one.

$$\begin{aligned} & \delta_X \circ \eta^{\mathsf{E}}_{\mathsf{P}X} \\ = \ & \{\ \eta_X(a) \in \mathsf{E}X \mid a \in \_\ \} \\ = \ & \mathsf{P}\eta^{\mathsf{E}}_X \end{aligned}$$

Finally, we will show that the natural transformation $\delta : \mathsf{EQ} \to \mathsf{QE}$ makes the following diagram to commute.

$$\begin{array}{ccccc}
\mathsf{EQQ} & \xrightarrow{\delta_{\mathsf{Q}}} & \mathsf{QEQ} & \xrightarrow{\mathsf{Q}\delta} & \mathsf{QQE} \\
\downarrow{\scriptstyle \mathsf{E}\mu} & & & & \downarrow{\scriptstyle \mu_{\mathsf{Q}}} \\
\mathsf{EQ} & & \xrightarrow[\delta]{\hspace{4em}} & & \mathsf{QE} \\
\uparrow{\scriptstyle \mu_{\mathsf{Q}}} & & & & \uparrow{\scriptstyle \mathsf{Q}\mu} \\
\mathsf{EEQ} & \xrightarrow[\mathsf{E}\delta]{} & \mathsf{EQE} & \xrightarrow[\delta_{\mathsf{E}}]{} & \mathsf{QEE}
\end{array}$$

Start with the upper square. Consider an element $(f, d) \in \mathsf{EQQ}X$. A straightforward calculation provides the following equations.

$$\begin{aligned}
&\delta_X \circ \mathsf{E}\mu_X(f, d) \\
&= \ \{ \ (g, d) \in \mathsf{E}X \mid g \in \cup \circ f \ \}
\end{aligned}$$

$$\begin{aligned}
&\mu_{\mathsf{Q}X} \circ \mathsf{Q}\delta_X \circ \delta_{\mathsf{Q}X}(f, d) \\
&= \ \bigcup \{ \ \delta_X \ (h, d) \mid (h, d) \in \mathsf{EQ}X \wedge h \in f \ \}
\end{aligned}$$

We will show that both sets are indeed the same. For this, start with an element $(g, d) \in \mathsf{E}X$, and reason in the following manner.

$$\begin{aligned}
& \quad g \in \cup \circ f \\
\Leftrightarrow & \quad \forall t \in \mathbb{R}^+ .\ g\,(t) \in \cup \circ f \\
\Leftrightarrow & \quad \forall t \in \mathbb{R}^+ .\ \exists Z_t \in f\,(t)\,.\ g\,(t) \in Z_t \\
\Leftrightarrow & \quad \exists (h, d) \in \mathsf{EQ}X .\ g \in h \wedge h \in f \\
\Leftrightarrow & \quad \exists (h, d) \in \mathsf{EQ}X .\ (g, d) \in \delta_X\,(h, d) \wedge h \in f
\end{aligned}$$

In order to keep the notation unburduned, and whenever no ambiguities arise, we will often use a pair $(f, d) \in \mathsf{E}X$ as if it were simply the map $f \in X^{\mathbb{R}^+}$.

Now concentrate on the lower square. Consider an element $(f, d) \in \mathsf{EEQ}X$ with finite duration, and let $e = \pi_2(f\ d)$. Then, a straightforward calculation shows that the following equations hold.

$$\begin{aligned}
&\delta_X \circ \mu_{\mathsf{Q}X}\ (f, d) \\
&= \ \{ \ (g, d + e) \in \mathsf{E}X \mid g \in (\theta_{\mathsf{Q}X} \circ f, d) +\!\!+ (f\ d) \ \}
\end{aligned}$$

$$\begin{aligned}
&\mathsf{Q}\mu_X \circ \delta_{\mathsf{E}X} \circ \mathsf{E}\delta_X\ (f, d) \\
&= \ \{ \ (\theta_X \circ g, d) +\!\!+ (g\ d) \mid (g, d) \in \mathsf{EE}X \wedge g \in \delta_X \circ f \ \}
\end{aligned}$$

We will show that both sets are actually the same. Start with an element $(h, d + e) \in \mathsf{E}X$, and reason as follows.

$$
\begin{aligned}
& h \in (\theta_{\mathsf{Q}X} \circ f, d) \mathbin{+\!\!+} (f\ d) \\
\Leftrightarrow\ & \forall t \le d\ .\ h\,(t) \in \theta_{\mathsf{Q}X} \circ f\,(t)\ \wedge \\
& \quad \forall t > d\ .\ h\,(t) \in (f\ d)\,(t-d) \\
\overset{(\star)}{\Leftrightarrow}\ & \forall t \le d\ .\ h\,(t) \in \mathsf{Q}\theta_X \circ \delta_X \circ f\,(t)\ \wedge \\
& \quad \exists (g,e) \in \delta_X\,(f\ d)\ .\forall t > d\ .\ h\,(t) = g\,(t-d) \\
\Leftrightarrow\ & \exists (g,d) \in \mathsf{EE}X\ .\ \forall t \le d\ .\ g\,(t) \in \delta_X\,(f\ t)\ \wedge \\
& \quad \theta_X \circ g\,(t) = h\,(t)\ \wedge\ \forall t > d\ .\ g\,(t) \in \delta_X\,(f\ t)\ \wedge \\
& \quad h\,(t) = (g\ d)\,(t-d) \\
\Leftrightarrow\ & \exists (g,d) \in \mathsf{EE}X\ .\ (h, d+e) = (\theta_X \circ g, d) \mathbin{+\!\!+} (g\ d) \wedge \\
& \quad g \in \delta_X \circ f
\end{aligned}
$$

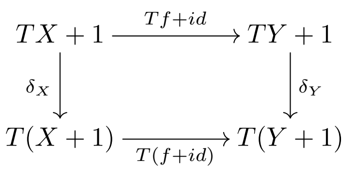
The case in which $(f,d) \in \mathsf{EEQ}X$ has infinite duration follows by similar arguments. □

Note that if instead of the functor Q one would consider the powerset, then the equivalence $(\star)$ would not hold. In particular, the equation

$$\theta_{\mathsf{P}X} \circ f = \mathsf{P}\theta_X \circ \delta_X \circ f$$

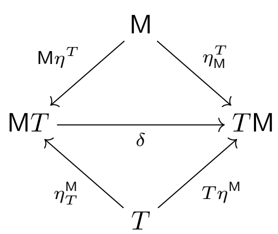
would not necessarily hold.

Proof of Theorem 27. We will first tackle naturality, by showing that the diagram below commutes for any map $f : X \to Y$.



$$
\begin{aligned}
& T(f+id) \circ [Ti_1, \eta_{X+1} \circ i_2] \\
=\ & T(f+id) \circ [Ti_1, Ti_2 \circ \eta_1] \\
=\ & [T(f+id) \circ Ti_1, T(f+id) \circ Ti_2 \circ \eta_1] \\
=\ & [Ti_1 \circ Tf, Ti_2 \circ \eta_1] \\
=\ & [Ti_1, \eta_{X+1} \circ i_2] \circ (Tf + id)
\end{aligned}
$$

Then we will show that natural transformation $\delta : \mathsf{M}T \to T\mathsf{M}$ makes the following diagram to commute.



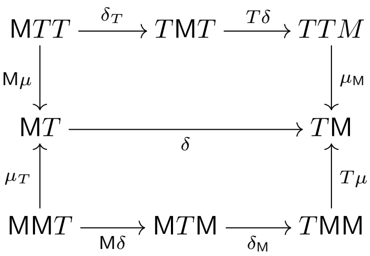
Start with the upper triangle.

$$
\begin{aligned}
& [Ti_1, \eta^T_{X+1} \circ i_2] \circ (\eta^T_X + id) \\
= \; & [Ti_1 \circ \eta^T_X, \eta^T_{X+1} \circ i_2] \\
= \; & [\eta^T_{X+1} \circ i_1, \eta^T_{X+1} \circ i_2] \\
= \; & \eta^T_{X+1}
\end{aligned}
$$

For the lower triangle reason in the following manner.

$$
\begin{aligned}
& [Ti_1, \eta^T_{X+1} \circ i_2] \circ \eta^{\mathsf{M}}_{TX} \\
= \; & [Ti_1, \eta^T_{X+1} \circ i_2] \circ i_1 \\
= \; & Ti_1 \\
= \; & T\eta^{\mathsf{M}}_X
\end{aligned}
$$

Finally, we will show that the natural transformation $\delta : \mathsf{M}T \to T\mathsf{M}$ makes the following diagram to commute.



For this we will take advantage of a well-known property of coproducts: consider two maps $f, g : X + Y \to Z$. If the equations $f \circ i_1 = g \circ i_1$, $f \circ i_2 = g \circ i_2$ hold the equation $f = g$ holds as well. We start with the upper square:

$$
\begin{aligned}
& \mu^T_{X+1} \circ T\delta_X \circ \delta_{TX} \circ i_1 \\
= \; & \mu^T_{X+1} \circ T\delta_X \circ Ti_1 \\
= \; & \mu^T_{X+1} \circ TTi_1 \\
= \; & Ti_1 \circ \mu^T_X \\
= \; & \delta_X \circ i_1 \circ \mu^T_X \\
= \; & \delta_X \circ (\mu^T_X + id) \circ i_1
\end{aligned}
$$

and similarly

$$
\begin{aligned}
& \mu^T_{X+1} \circ T\delta_X \circ \delta_{TX} \circ i_2 \\
= \; & \mu^T_{X+1} \circ T\delta_X \circ Ti_2 \circ \eta^T_1 \\
= \; & \mu^T_{X+1} \circ T(\delta_X \circ i_2) \circ \eta^T_1 \\
= \; & \mu^T_{X+1} \circ T(\eta^T_{X+1}) \circ Ti_2 \circ \eta^T_1 \\
= \; & Ti_2 \circ \eta^T_1 \\
= \; & \eta^T_{X+1} \circ i_2 \\
= \; & \delta_X \circ i_2 \\
= \; & \delta_X \circ (\mu^T_X + id) \circ i_2
\end{aligned}
$$

For the lower square we reason in the following manner.

$$\begin{aligned}
& T\mu^{\mathsf{M}}_X \circ \delta_{X+1} \circ (\delta_X + id) \circ i_1 \\
= \;& T\mu^{\mathsf{M}}_X \circ \delta_{X+1} \circ i_1 \circ \delta_X \\
= \;& T\mu^{\mathsf{M}}_X \circ T i_1 \circ \delta_X \\
= \;& T(\mu^{\mathsf{M}}_X \circ i_1) \circ \delta_X \\
= \;& T\mathrm{id} \circ \delta_X \\
= \;& \delta_X \circ \mu^{\mathsf{M}}_X \circ i_1
\end{aligned}$$

and similarly

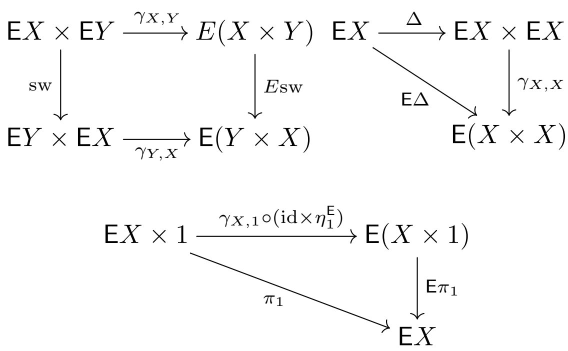
$$\begin{aligned}
& T\mu^{\mathsf{M}}_X \circ \delta_{X+1} \circ (\delta_X + id) \circ i_2 \\
= \;& T\mu^{\mathsf{M}}_X \circ \delta_X \circ i_2 \\
= \;& T\mu^{\mathsf{M}}_X \circ \eta^T_{X+1} \circ i_2 \\
= \;& T\mu^{\mathsf{M}}_X \circ T i_2 \circ \eta^T_X \\
= \;& T i_2 \circ \eta^T_X \\
= \;& \eta^T_{X+1} \circ i_2 \\
= \;& \delta_X \circ i_2 \\
= \;& \delta_X \circ \mu^{\mathsf{M}}_{TX} \circ i_2
\end{aligned}$$

□

Proof for Theorem 29. We start by presenting some properties of the natural transformation

$$\gamma : \mathsf{E} \times \mathsf{E} \to \mathsf{E}(- \times -)$$

which are proven in [NBHM16]: the following diagrams commute.



Then we make the following calculations for axioms K4-K6, assuming for each case that there is a natural transformation that respects the associated axiom (Corollary 9 ensures that this is a valid assumption).

$$\begin{aligned}
& \mathsf{E}\alpha_X \circ \gamma_{\mathsf{M}X,\mathsf{M}X} \circ \mathrm{sw} \\
= \;& \mathsf{E}\alpha_X \circ \mathsf{E}\mathrm{sw} \circ \gamma_{\mathsf{M}X,\mathsf{M}X} \\
= \;& \mathsf{E}\alpha_X \circ \gamma_{\mathsf{M}X,\mathsf{M}X}
\end{aligned}$$

$$
\begin{aligned}
& \mathsf{E}\alpha_X \circ \gamma_{\mathsf{M}X,\mathsf{M}X} \circ (\mathrm{id} \times (\mathsf{E}0_X \circ \eta_1^{\mathsf{E}})) \\
= \;& \mathsf{E}\alpha_X \circ \gamma_{\mathsf{M}X,\mathsf{M}X} \circ (\mathsf{E}\mathrm{id} \times \mathsf{E}0_X) \circ (\mathrm{id} \times \eta_1^{\mathsf{E}}) \\
= \;& \mathsf{E}\alpha_X \circ \mathsf{E}(\mathrm{id} \times 0_X) \circ \gamma_{\mathsf{M}X,1} \circ (\mathrm{id} \times \eta_1^{\mathsf{E}}) \\
= \;& \mathsf{E}\pi_1 \circ \gamma_{\mathsf{M}X,1} \circ (\mathrm{id} \times \eta_1^{\mathsf{E}}) \\
= \;& \pi_1
\end{aligned}
$$

$$
\begin{aligned}
& \mathsf{E}\alpha_X \circ \gamma_{\mathsf{M}X,\mathsf{M}X} \circ \Delta \\
= \;& \mathsf{E}\alpha_X \circ \mathsf{E}\Delta \\
= \;& \mathsf{E}\mathrm{id}
\end{aligned}
$$

Associativity (axiom K3) follows in similar terms. □