From Probability Monads to Commutative Effectuses ⋆

Bart Jacobs

Institute for Computing and Information Sciences,
Radboud University Nijmegen
P.O. Box 9010, 6500 GL Nijmegen, The Netherlands
Email: B.Jacobs@cs.ru.nl *URL:* http://www.cs.ru.nl/B.Jacobs

Abstract

Effectuses have recently been introduced as categorical models for quantum computation, with probabilistic and Boolean (classical) computation as special cases. These 'probabilistic' models are called commutative effectuses, and are the focus of attention here. The paper describes the main known 'probability' monads: the monad of discrete probability measures, the Giry monad, the expectation monad, the probabilistic power domain monad, the Radon monad, and the Kantorovich monad. It also introduces successive properties that a monad should satisfy so that its Kleisli category is a commutative effectus. The main properties are: partial additivity, strong affineness, and commutativity. It is shown that the resulting commutative effectus provides a categorical model of probability theory, including a logic using effect modules with parallel and sequential conjunction, predicate- and state-transformers, normalisation and conditioning of states.

Contents

⋆ The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement nr. 320571

1 Introduction

An effectus is a relatively simple category, with finite coproducts and a final object, satisfying some elementary properties: certain squares have to be pullbacks and certain parallel maps have to be jointly monic, see (25) and (13) below. These effectuses have been introduced in [30], and give rise to a rich theory that includes quantum computation, see the overview paper [10]. Subclasses of 'commutative' effectuses and 'Boolean' effectuses have been identified. These Boolean effectuses capture classical (deterministic) computation, and can be characterised as extensive categories, see [10, Sec. 13] for details. This is a non-trivial result. A similar result for commutative effectuses is still missing. It should lead to a characterisation of (categorical) models of probabilistic computation.

This paper combines two earlier conference publications [33,34] into a single integrated account. It introduces in a step-by-step manner successive properties of monads that ensure that their Kleisli categories are commutative effectuses. This is applied to six known examples of 'probability' monads, namely: the discrete probability distributions monad $\mathcal{D}$ on sets, the Giry monad $\mathcal{G}$ on measurable spaces, the expectation monad $\mathcal{E}$ on sets, the probabilistic powerdomain monad $\mathcal{V}$ on (continuous) directed complete partial orders, the Radon monad $\mathcal{R}$ on compact Hausdorff spaces, and the Kantorovich monad $\mathcal{K}$ on (1-bounded) metric spaces. One way to read this paper is as an instantiation of the general theory of effectuses to the special case of Kleisli categories of a monad. It goes beyond [52], which focuses on commutativity of the monad and ignores the (partially) additive structure of the monad; the latter leads to partial sums $\ovee$ of partial maps, and to effect module structure on predicates that play an important role here.

The paper establishes one half of a conjectured characterisation of these commutative effectuses as Kleisli categories of certain monads. The main result of this paper says that if the monad is partially additive, strongly affine, and commutative then its Kleisli category is a (monoidal) commutative effectus. Affineness of a monad T means that it preserves the final object: $T(1) \cong 1$. The property 'strong affineness' comes from [33], where it is used to prove a bijective correspondence between predicates and side-effect-free instruments. The relation between predicates and associated actions (instruments / coalgebras) comes from quantum theory in general, and effectus theory in particular. This relationship is complicated in the quantum case, but quite simple in the probabilistic case (see Proposition 27 below). It is the basis for a novel logic and type theory for probability in [3] (see also [42]). Partial additivity of a monad has been introduced in [27] where it is used to obtain partially additive structure on homsets of a Kleisli category. This result is re-used here, as a step towards constructing effectuses, following [9].

In future work we hope to find a construction in the other direction, turning a commutative effectus, possibly satisfying some additional properties, into a 'probabilistic' monad. Until such a characterisation result exists, we use the term 'probabilistic' monad in an intuitive sense, without providing a precise definition.

This work is organised as follows. It starts with two preliminary sections 2 and 3 explaining effect modules, and the basic categorical setting in which we will be working: distributive categories with a strong monad. Section 3 includes the six running monad examples. Section 4 investigates the structure of partial maps and predicates for monads which are partially additive. Using the additional requirement of strong affiness, introduced in Section 5, we prove in Section 6 that the Kleisli category is an effectus. Subsequently, Section 7 shows how certain actions/coalgebras, namely instrument and assert maps, can be associated with predicates. These actions are useful for 'if-then-else' style constructions, taking probabilities as weights into account, but also for conditioning of states. Section 8 shows that the additional requirement of commutativity of a monad makes the associated Kleisli category not only monoidal, but also 'commutative' in the effectus-theoretic sense. The latter means, essentially, that the sequential conjunction ('andthen') operation & is commutative. This commutativity and side-effect-freeness of actions is essential for the probabilistic world, in contrast to the quantum world, where observation instruments have side-effects, and consequently & is not commutative. The final section 9 reaps the fruits of all these efforts: it exploits a known result that effectuses with the unit interval $[0, 1]$ as scalars automatically admit normalisation of non-zero partial states. This normalisation is used to define the conditioning $\omega|_p$ of a state ω by a predicate p, forming the updated state 'ω given p'. It forms the basis of Bayesian reasoning. Several basic results about conditioning are proven at an abstract level, in what may be called 'categorical probability theory'.

2 Preliminaries

We assume that the reader is familiar with the basics of probability theory and also with the basics of category theory. The common structures in algebraic logic, like Boolean algebras or Heyting algebras, are not appropriate for the logic of probabilistic models. Instead, we need to use effect modules. Since these structures are relatively unfamiliar, we introduce them here. We shall see several examples in the sequel.

Before reading the definition of a partial commutative monoid (PCM), think of the unit interval $[0, 1]$ with addition $+$. This $+$ is obviously only a partial operation, which is commutative and associative in a suitable sense. This will be formalised next.

A partial commutative monoid (PCM) consists of a set M with a zero element $0 \in M$ and a partial binary operation $\ovee\colon M \times M \to M$ satisfying the three requirements below. They involve the notation $x \perp y$ for: $x \ovee y$ is defined; in that case x, y are called orthogonal.

(i) Commutativity: $x \perp y$ implies $y \perp x$ and $x \ovee y = y \ovee x$;
(ii) Associativity: $y \perp z$ and $x \perp (y \ovee z)$ implies $x \perp y$ and $(x \ovee y) \perp z$ and also $x \ovee (y \ovee z) = (x \ovee y) \ovee z$;
(iii) Zero: $0 \perp x$ and $0 \ovee x = x$;

The notion of effect algebra is due to [20], see also [16] for an overview. An effect algebra is a PCM $(E, 0, \ovee)$ with an orthosupplement. The latter is a total unary 'negation' operation $(-)^\perp\colon E \to E$ satisfying:

(iv) $x^\perp \in E$ is the unique element in E with $x \ovee x^\perp = 1$, where $1 = 0^\perp$;
(v) $x \perp 1 \Rightarrow x = 0$.

A homomorphism $E \to D$ of effect algebras is given by a function $f\colon E \to D$ between the underlying sets satisfying $f(1) = 1$, and if $x \perp x'$ in E then both $f(x) \perp f(x')$ in D and $f(x \ovee x') = f(x) \ovee f(x')$. Effect algebras and their homomorphisms form a category, denoted by **EA**.

The unit interval $[0,1]$ is a PCM with sum of $r,s \in [0,1]$ defined if $r+s \leq 1$, and in that case $r \ovee s = r+s$. The unit interval is also an effect algebra with $r^\perp = 1-r$. In [37] it is shown that the category **EA** is symmetric monoidal, and that this unit interval $[0,1]$, with its (total) multiplication is a monoid in **EA**. An effect module E is an action $[0,1] \otimes E \to E$ wrt. this monoid.

More concretely, an *effect module* is an effect algebra E with a scalar multiplication $s \cdot x$, for $s \in [0,1]$ and $x \in E$ forming an action:

$$1 \cdot x = x \qquad (r \cdot s) \cdot x = r \cdot (s \cdot x),$$

and preserving sums (that exist) in both arguments:

$$\begin{aligned} 0 \cdot x &= 0 \qquad & (r+s) \cdot x &= r \cdot x \ovee s \cdot x \\ s \cdot 0 &= 0 \qquad & s \cdot (x \ovee y) &= s \cdot x \ovee s \cdot y. \end{aligned}$$

We write **EMod** for the category of effect modules, where morphisms are maps of effect algebras that preserve scalar multiplication (*i.e.* are 'equivariant'). A simple example of an effect module is a set $[0,1]^X$ of fuzzy predicates on a set X. It inherits effect algebra structure from $[0,1]$, pointwise. Its scalar multiplication $s \cdot p \in [0,1]^X$, for $p \in [0,1]^X$ and $s \in [0,1]$ is given by $(s \cdot p)(x) = s \cdot p(x)$. It is not hard to see that for a function f we get a map of effect modules $(-) \circ f \colon [0,1]^Y \to [0,1]^X$. This yields a functor $\mathbf{Sets} \to \mathbf{EMod}^{\mathrm{op}}$.

3 Running monad examples

This section describes the six monad examples $T \colon \mathbf{C} \to \mathbf{C}$ that serve as our main illustrations of 'probability monads'. We first briefly review the underlying categories $\mathbf{C}$ in these examples. What they have in common is that they are distributive categories with disjoint coprojections. We recall from [11] that coprojections $\kappa_i \colon X_i \to X_1 + X_2$ in a distributive category are automatically monic, and that the initial object 0 is strict — that is, each map $X \to 0$ is an isomorphism.

Definition 1 *A category is called* distributive *if it has finite products* $(\times, 1)$ *and coproducts* $(+, 0)$*, where products distribute over coproducts, in the sense that the following maps are isomorphisms.*

$$0 \xrightarrow{\;!\;} 0 \times X \qquad (A \times X) + (B \times X) \xrightarrow{\mathrm{dis}_1 = [\kappa_1 \times \mathrm{id}, \kappa_2 \times \mathrm{id}]} (A+B) \times X \tag{1}$$

We say that the coprojections are disjoint *if the diagrams below are pullbacks.*

$$\begin{tikzcd} 0 \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & X_2 \arrow[d, tail, "\kappa_2"] \\ X_1 \arrow[r, tail, "\kappa_1"'] & X_1 + X_2 \end{tikzcd} \tag{2}$$

We call such a category non-trivial *if it additionally satisfies: for each object X we have: $X \not\cong 0$ iff there is a map $x\colon 1 \to X$. This implies $1 \not\cong 0$.*

Swapping the distributivity map dis_1 in (1) yields an associated distributivity map:

$$(X \times A) + (X \times B) \xrightarrow[= \gamma \circ \mathrm{dis}_1 \circ (\gamma + \gamma)]{\mathrm{dis}_2 \,=\, [\mathrm{id} \times \kappa_1, \mathrm{id} \times \kappa_2]} X \times (A + B)$$

where $\gamma = \langle \pi_2, \pi_1 \rangle$ is the (product) swap isomorphism.

In a distributive category we sometimes write $n = 1 + \cdots + 1$ for the n-fold sum (copower) of the final object 1. There is an associated isomorphism $\mathrm{sep}_n \colon n \times X \to X + \cdots + X$ obtained as:

$$\mathrm{sep}_n \stackrel{\text{def}}{=} \Big(n \times X \xrightarrow{\cong} 1 \times X + \cdots + 1 \times X \xrightarrow[\cong]{\pi_2 + \cdots + \pi_2} X + \cdots + X \Big). \tag{3}$$

We write **Sets** for the category of sets and functions. Finite products in **Sets** are given by the singleton final set $1 = \{*\}$ and by the usual cartesian products $X \times Y$. Finite coproducts involve the empty set 0 and the disjoint union $X + Y$. Notice that $n = 1 + \cdots + n \cong \{0, 1, \ldots, n-1\}$ is an n-element set, with $2 = 1 + 1 \cong \{0, 1\}$ as special case.

The category **Meas** contains as objects measurable spaces $X = (X, \Sigma_X)$, consisting of a set X together with a σ-algebra $\Sigma_X \subseteq \mathcal{P}(X)$. A morphism $X \to Y$ in **Meas**, from (X, Σ_X) to (Y, Σ_Y), is a measurable function $f\colon X \to Y$, *i.e.* a function satisfying $f^{-1}(M) \in \Sigma_X$ for each measurable subset $M \in \Sigma_Y$. With each topological space X with opens $\mathcal{O}(X)$ one associates the least σ-algebra containing $\mathcal{O}(X)$. This is the Borel algebra/space on X, written as $\mathcal{B}(X)$. In particular the unit interval $[0, 1]$ forms a measurable space. Its measurable subsets are generated by the intervals $[q, 1]$, where q is a rational number in $[0, 1]$. The (categorical) product $X_1 \times X_2$ of two measurable spaces X_i carries the least σ-algebra making both projections $\pi_i \colon X_1 \times X_2 \to X_i$ measurable functions; equivalently, this σ-algebra is generated by the rectangles $M_1 \times M_2$ with $M_i \in \Sigma_{X_i}$ The coproduct $X_1 + X_2$ involves the disjoint union of the underlying sets with the σ-algebra given by the direct images $\kappa_i M = \{\kappa_i x \mid x \in M\}$ for $M \in \Sigma_{X_i}$, where $\kappa_i \colon X_i \to X_1 + X_2$ is the coprojection map. The empty set 0 and the final set 1 are initial and final measurable spaces, with the trivial (discrete) σ-algebra.

We write **CH** for the category of compact Hausdorff topological spaces, with continuous functions between them. Finite products $(1, \times)$ are given as in

Sets, with the standard product topology: the coarsest (least) one making the projections π_i continuous. Finite coproducts are also as in **Sets**, with the finest (greatest) topology making the coprojections κ_i continuous.

We write **Dcpo** for the category of directed complete partial orders (dcpo's), with (Scott) continuous functions between them. For a dcpo X we write $\mathcal{O}(X)$ for the complete lattice of Scott open subsets: upward closed subsets $U \subseteq X$ with: if $\bigvee_i x_i \in U$, then $x_i \in U$ for some index i. Finite products and coproducts are as in **Sets**, with the obvious orders. The full subcategory **CDcpo** $\hookrightarrow$ **Dcpo** contains *continuous* dcpo's where each element is a directed join of elements way below it.

A metric space is a pair $X = (X, d_X)$ where d_X is a distance function on X. The category $\mathbf{Met}_1$ contains the '1-bounded' metric spaces, with distance function $d_X \colon X \times X \to [0,1]$ taking values in the unit interval $[0,1]$. A map $f \colon (X, d_X) \to (Y, d_Y)$ in $\mathbf{Met}_1$ is a function $f \colon X \to Y$ which is non-expansive: $d_Y(f(x), f(x')) \leq d_X(x, x')$. Products in the category $\mathbf{Met}_1$ use the cartesian product of the underlying sets, with distance function $d((x,y),(x',y')) = \max\Big(d(x,x'), d(y,y')\Big)$. Coproducts are also as in sets, where $d(\kappa_i z, \kappa_j z')$ equals $d(z,z')$ if $i = j$ and 1 otherwise. Here we use 1-boundedness.

We continue to describe the monads that we will use on these categories. In general, for a monad $T \colon \mathbf{C} \to \mathbf{C}$, we write η for its unit and μ for its multiplication. The associated Kleisli category is denoted by $\mathcal{K}\ell(T)$. We typically write $\bullet$ for composition in this Kleisli category, in order to distinguish it from ordinary composition $\circ$ in the underlying category $\mathbf{C}$. Recall that $g \bullet f = \mu \circ T(g) \circ f$. Each map $f \colon X \to Y$ in $\mathbf{C}$ gives a map $\langle f \rangle = \eta \circ f \colon X \to Y$ in $\mathcal{K}\ell(T)$. This gives a functor $\mathbf{C} \to \mathcal{K}\ell(T)$, since $\langle g \circ f \rangle = \langle g \rangle \bullet \langle f \rangle$. A standard fact is that $\mathcal{K}\ell(T)$ inherits coproducts from $\mathbf{C}$, with coprojections $\langle \kappa_i \rangle$.

The monad is called *strong* if there is a 'strength' natural transformation st_1 with components $(\mathrm{st}_1)_{X,Y} \colon T(X) \times Y \to T(X \times Y)$ making the following diagrams commute — in which we omit indices, for convenience.

$$
\begin{tikzcd}
T(X) \times Y \arrow[r, "\mathrm{st}_1"] \arrow[dr, "\pi_1"'] & T(X \times Y) \arrow[d, "T(\pi_1)"] \\
& T(X)
\end{tikzcd}
\qquad
\begin{tikzcd}
(T(X) \times Y) \times Z \arrow[r, "\cong"] \arrow[d, "\mathrm{st}_1 \times \mathrm{id}"'] & T(X) \times (Y \times Z) \arrow[dd, "\mathrm{st}_1"] \\
T(X \times Y) \times Z \arrow[d, "\mathrm{st}_1"'] & \\
T((X \times Y) \times Z) \arrow[r, "\cong"] & T(X \times (Y \times Z))
\end{tikzcd}
\tag{4}
$$

$$
\begin{tikzcd}
X \times Y \arrow[r, equal] \arrow[d, "\eta \times \mathrm{id}"'] & X \times Y \arrow[d, "\eta"] \\
T(X) \times Y \arrow[r, "\mathrm{st}_1"'] & T(X \times Y)
\end{tikzcd}
\qquad
\begin{tikzcd}
T^2(X) \times Y \arrow[r, "\mathrm{st}_1"] \arrow[d, "\mu \times \mathrm{id}"'] & T(T(X) \times Y) \arrow[r, "T(\mathrm{st}_1)"] & T^2(X \times Y) \arrow[d, "\mu"] \\
T(X) \times Y \arrow[rr, "\mathrm{st}_1"'] & & T(X \times Y)
\end{tikzcd}
\tag{5}
$$

Each monad on the category **Sets** of sets and functions is automatically strong, via the definition $\mathrm{st}_1(u, y) = T(\lambda x.\, \langle x, y\rangle)(u)$.

Given a strength map $\mathrm{st}_1 \colon T(X) \times Y \to T(X \times Y)$ we define an associated version st_2 via swapping:

$$\mathrm{st}_2 = \Big(X \times T(Y) \xrightarrow[\cong]{\gamma} T(Y) \times X \xrightarrow{\mathrm{st}_1} T(Y \times X) \xrightarrow[\cong]{T(\gamma)} T(X \times Y)\Big)$$

where $\gamma = \langle \pi_2, \pi_1 \rangle$ is the swap map.

The strength and distributivity maps also interact in the obvious way. There are two formulations, with st_1 and dis_2 and with st_2 and dis_1, which are both derivable. We describe the version that we actually need later on — and leave the verification to the meticulous reader.

$$\begin{tikzcd}
A \times T(X) + B \times T(X) \arrow[r, "\mathrm{dis}_1"] \arrow[d, "\mathrm{st}_2 + \mathrm{st}_2"'] & (A+B) \times T(X) \arrow[dd, "\mathrm{st}_2"] \\
T(A \times X) + T(B \times X) \arrow[d, "{[T(\kappa_1), T(\kappa_2)]}"'] & \\
T((A \times X) + (B \times X)) \arrow[r, "T(\mathrm{dis}_1)"] & T((A+B) \times X)
\end{tikzcd} \tag{6}$$

The monad T is called *commutative* (following [49]) when the order of applying strength in two coordinates does not matter, as expressed by commutation of the following diagram.

$$\begin{tikzcd}
 & T(X \times T(Y)) \arrow[r, "T(\mathrm{st}_2)"] & T^2(X \times Y) \arrow[dr, "\mu"] & \\
T(X) \times T(Y) \arrow[ur, "\mathrm{st}_1"] \arrow[dr, "\mathrm{st}_2"'] & & & T(X \times Y) \\
 & T(T(X) \times Y) \arrow[r, "T(\mathrm{st}_1)"'] & T^2(X \times Y) \arrow[ur, "\mu"'] &
\end{tikzcd} \tag{7}$$

We then write $\mathrm{dst} \colon T(X) \times T(Y) \to T(X \times Y)$ for 'double strength', to indicate the resulting single map, from left to right. This dst is called a 'Fubini' map in [52]. Notice that $\mathrm{dst} \circ \gamma = T(\gamma) \circ \mathrm{dst}$. The Kleisli category $\mathcal{K}\ell(T)$ of a commutative monad T is symmetric monoidal, with tensor $X_1 \otimes X_2 = X_1 \times X_2$ on objects. For Kleisli maps $f_i \colon X_i \to T(Y_i)$ we get $f_1 \otimes f_2 \colon X_1 \otimes X_2 \to Y_1 \otimes Y_2$ in $\mathcal{K}\ell(T)$ given by:

$$f_1 \otimes f_2 = \Big(X_1 \times X_2 \xrightarrow{f_1 \times f_2} T(Y_1) \times T(Y_2) \xrightarrow{\mathrm{dst}} T(Y_1 \times Y_2)\Big).$$

The tensor unit is the final object 1 from $\mathbf{C}$ — which is, in general, not final in $\mathcal{K}\ell(T)$.

Below we fix the terminology that we will be using for predicates and states. It is formulated quite generally, for an arbitrary monad. But as we shall see it really makes sense for the 'probability' monads that we will consider in this paper. The terminology of predicates, states, scalars, and validity comes from effectus theory [10].

Definition 2 *Let T be a monad on a distributive category $\mathbf{C}$.*

(i) *Maps in $\mathbf{C}$ of the form $p\colon X \to T(2)$, where $2 = 1 + 1$, will be called* predicates *on X. We write* Pred(X) *for the set of predicates on X.*

(ii) *The truth and falsity predicates are defined as:*

$$\mathbf{1} \stackrel{\text{def}}{=} \Big(X \xrightarrow{\;!\;} 1 \xrightarrow{\langle\kappa_1\rangle} T(2) \Big) \qquad\qquad \mathbf{0} \stackrel{\text{def}}{=} \Big(X \xrightarrow{\;!\;} 1 \xrightarrow{\langle\kappa_2\rangle} T(2) \Big)$$

The orthosupplement *$p^\perp$ of a predicate p is obtained by swapping the outcomes:*

$$p^\perp \stackrel{\text{def}}{=} \Big(X \xrightarrow{\;p\;} T(2) \xrightarrow[\cong]{T([\kappa_2,\kappa_1])} T(2) \Big)$$

Clearly, $p^{\perp\perp} = p$ and $\mathbf{1}^\perp = \mathbf{0}$ and $\mathbf{0}^\perp = \mathbf{1}$.

(iii) *Maps of the form $1 \to T(2)$, that is, predicates on 1, are called* scalars. *They act on predicates via Kleisli composition:*

$$s \cdot p \stackrel{\text{def}}{=} \Big(X \xrightarrow{\;p\;} 2 \xrightarrow{\quad [s,\mathbf{0}] \quad} 2 \Big)$$

That is, $s \cdot p = [s, \mathbf{0}] \bullet p = \mu \circ T([s, \mathbf{0}]) \circ p$. In particular, this yields a monoid structure on the set Pred(1) *of scalars, and a monoid action* Pred$(1) \times$ Pred$(X) \to$ Pred(X).

(iv) *Maps $\omega\colon 1 \to T(X)$ in $\mathbf{C}$ are called* states *of X. We write* Stat(X) *for the set of states of X.*

(v) *The* validity, *or* expected value, *of a predicate $p\colon X \to T(2)$ on X in a state $\omega\colon 1 \to T(X)$ of X is the scalar $\omega \models p$ that is obtained by Kleisli composition:*

$$\omega \models p \stackrel{\text{def}}{=} p \bullet \omega.$$

This is an abstract version of the Born rule from quantum theory, and a variation on the 'integration pairing' from [52].

(vi) *For a Kleisli map $f\colon X \to T(Y)$, a state ω of X and a predicate q on Y we define predicate and state transformer functions f^* and f_* via Kleisli pre- and post-composition:*

$$\begin{array}{ccc} \mathrm{Pred}(Y) \xrightarrow{\;f^*\;} \mathrm{Pred}(X) & \qquad\qquad & \mathrm{Stat}(X) \xrightarrow{\;f_*\;} \mathrm{Stat}(Y) \\ q \longmapsto q \bullet f & & \omega \longmapsto f \bullet \omega. \end{array}$$

Then $f_*(\omega) \models q \;=\; q \bullet f \bullet \omega \;=\; \omega \models f^*(q)$. Moreover, $f^*(\mathbf{1}) = \mathbf{1}$, $f^*(\mathbf{0}) = \mathbf{0}$, and $f^*(q^\perp) = f^*(q)^\perp$.

For the points below we assume that T is a commutative monad, so that its Kleisli category has tensors $\otimes$.

(vii) *For states* $\omega_i \colon 1 \to T(X_i)$ *we write* $\omega_1 \otimes \omega_2 = \mathrm{dst} \circ \langle \omega_1, \omega_2 \rangle \colon 1 \to T(X_1 \otimes X_2)$ *for their 'product' state. In the other direction, given a 'joint' state* $\rho \colon 1 \to T(X_1 \otimes X_2)$ *we can form their 'marginals'* $\rho_i \colon 1 \to T(X_i)$ *as* $\rho_i = \langle \pi_i \rangle_*(\rho)$, *that is, as composites in* $\mathbf{C}$*:*

$$\rho_i = \Big(1 \xrightarrow{\rho} T(X_1 \times X_2) \xrightarrow{T(\pi_i)} T(X_i) \Big).$$

The joint state ρ *will be called* non-entwined *if* $\rho = \rho_1 \otimes \rho_2$.

(viii) *For predicates* p *on* X *and* q *on* Y *we write* $p \odot q$ *for the* parallel conjunction *predicate on* $X \otimes Y$ *defined as composite:*

$$X \times Y \xrightarrow{p \times q} T(2) \times T(2) \xrightarrow{\mathrm{dst}} T(2 \times 2) \xrightarrow{T(\mathrm{sep}_2)} T(2+2) \xrightarrow{T([\mathrm{id}, \mathbf{0}])} T(2)$$

The terminology 'non-entwined' in point (vii) is similar to what is called 'non-entangled' in quantum theory and 'independent' in probability theory, where it is often used for random variables instead of for distributions (states). We have chosen to use the new expression 'non-entwinedness' since we describe the property that a state is the product of its marginals at a high level of generality, for an arbitrary commutative monad.

Remark 3 *It can be shown that if the monad* T *is commutative, then the multiplication* $s \cdot r = [s, \mathbf{0}] \bullet r$ *of scalars in point* (iii) *is commutative. It coincides with the parallel conjunction* $\odot$ *from point* (viii)*, see [30, Prop. 10.2] for details.*

Below we list the six monads that will be our running examples in this paper. We describe the essentials, and refer to the relevant literature for further details. We briefly describe what predicates and states are in each case.

3.1 The discrete probability distribution monad $\mathcal{D}$ *on* **Sets**

The elements of $\mathcal{D}(X)$ are the finite formal convex combinations of elements of a set X, written as $\sum_i r_i |x_i\rangle$, where elements $x_i \in X$ and probabilities $r_i \in [0,1]$ satisfy $\sum_i r_i = 1$. The ket notation $|-\rangle$ is meaningless syntactic sugar, that serves to distinguish elements $x \in X$ from their occurrences $|x\rangle$ in such formal sums. We can identify such a convex sum with a 'mass' function $\omega \colon X \to [0,1]$ whose support $\mathrm{supp}(\omega) = \{x \mid \omega(x) \neq 0\}$ is finite and satisfies $\sum_x \omega(x) = 1$. We can thus write $\omega = \sum_x \omega(x)|x\rangle$.

For a function $f \colon X \to Y$ one writes $\mathcal{D}(f) \colon \mathcal{D}(X) \to \mathcal{D}(Y)$ for the function defined by:

$$\mathcal{D}(f)\Big(\textstyle\sum_i r_i |x_i\rangle\Big) = \sum_i r_i |f(x_i)\rangle \qquad \text{that is:} \qquad \mathcal{D}(f)(\omega)(y) = \sum_{x \in f^{-1}(y)} \omega(x).$$

The unit $\eta\colon X \to \mathcal{D}(X)$ and multiplication $\mu\colon \mathcal{D}^2(X) \to \mathcal{D}(X)$ are given by:

$$\eta(x) = 1|x\rangle \qquad \mu(\Omega)(x) = \sum_{\varphi} \Omega(\varphi) \cdot \varphi(x).$$

Hence Kleisli composition is: $(g \bullet f)(x)(z) = \sum_y g(y)(z) \cdot f(x)(y)$. The monad $\mathcal{D}$ is commutative, with $\mathrm{st}_1\colon \mathcal{D}(X) \times Y \to \mathcal{D}(X \times Y)$ and $\mathrm{dst}\colon \mathcal{D}(X) \times \mathcal{D}(Y) \to \mathcal{D}(X \times Y)$ given by:

$$\mathrm{st}_1(\omega, y) = \sum_x \omega(x)|x, y\rangle \qquad \text{and} \qquad \mathrm{dst}(\omega, \rho)(x, y) = \omega(x) \cdot \rho(y).$$

It is easy to see that $\mathcal{D}(0) \cong 0$, $\mathcal{D}(1) \cong 1$, and $\mathcal{D}(2) \cong [0,1]$. The latter tells us that predicates $X \to \mathcal{D}(2)$ can be identified with fuzzy predicates $p\colon X \to [0,1]$. We have $p^{\perp}(x) = 1 - p(x)$, $\mathbf{1}(x) = 1$, and $\mathbf{0}(x) = 0$. The scalars are the elements of the unit interval $[0,1]$. States of X are maps $1 \to \mathcal{D}(X)$, which can be identified with probability distributions $\omega \in \mathcal{D}(X)$. The validity $\omega \models p$ is the expected value, in discrete probability theory:

$$\omega \models p = \sum_x \omega(x) \cdot p(x).$$

Notice that this is a finite sum, since the support of ω is finite.

For a Kleisli map $f\colon X \to \mathcal{D}(Y)$, a state $\omega \in \mathcal{D}(X)$, and a predicate $q \in [0,1]^Y$ we have $f^*(q) \in [0,1]^X$ and $f_*(\omega) \in \mathcal{D}(Y)$ given by:

$$f^*(q)(x) = f(x) \models q = \sum_y f(x)(y) \cdot q(y) \qquad f_*(\omega)(y) = \sum_x f(x)(y) \cdot \omega(x).$$

For more information, see *e.g.* [36].

3.2 The continuous probability distributions monad $\mathcal{G}$ on **Meas**

Next we consider the Giry monad $\mathcal{G}$ on the category **Meas** of measurable spaces. For a measurable space $X \in$ **Meas**, the elements of $\mathcal{G}(X)$ are probability measures $\omega\colon \Sigma_X \to [0,1]$. Each measurable subset $M \in \Sigma_X$ yields a function $\mathrm{ev}_M\colon \mathcal{G}(X) \to [0,1]$, namely $\mathrm{ev}_M(\omega) = \omega(M)$. Thus one can equip the set $\mathcal{G}(X)$ with the least σ-algebra making all these maps ev_M measurable. We obtain a functor **Meas** $\to$ **Meas** since for a map $f\colon X \to Y$ in **Meas** we get a measurable function $\mathcal{G}(f)\colon \mathcal{G}(X) \to \mathcal{G}(Y)$ given by:

$$\mathcal{G}(f)\Big(\Sigma_X \xrightarrow{\phi} [0,1]\Big) = \Big(\Sigma_Y \xrightarrow{f^{-1}} \Sigma_X \xrightarrow{\phi} [0,1]\Big).$$

We have $\mathcal{G}(0) \cong 0$, $\mathcal{G}(1) \cong 1$, and $\mathcal{G}(2) \cong [0,1]$. Hence predicates on $X \in$ **Meas** are now measurable functions/predicates $X \to [0,1]$, and scalars are probabilities, in $[0,1]$. A state $1 \to X$ in $\mathcal{K}\ell(\mathcal{G})$ is a probability measure $\omega \in$

$\mathcal{G}(X)$. The theory of Lebesgue integration tells us how to obtain for a predicate $p\colon X \to [0,1]$, and a state $\omega \in \mathcal{G}(X)$, the value:

$$\int p \, \mathrm{d}\omega \in [0,1] \qquad \text{which turns out to be the validity } \omega \models p.$$

Interpreting this integral as validity goes back to [53]. It allows us to describe Kleisli composition $\bullet$, and thus, implicitly, the multiplication μ of the monad. For $f\colon X \to Y$ and $g\colon Y \to Z$ in $\mathcal{K}\ell(\mathcal{G})$ we have for $x \in X$ and $M \in \Sigma_Z$,

$$(g \bullet f)(x)(M) = \int g(-)(M) \, \mathrm{d}f(x) = f(x) \models g(-)(M),$$

where $g(-)(M)\colon Y \to [0,1]$ is the predicate sending $y \in Y$ to $g(y)(M) \in [0,1]$. Its validity is computed in the state $f(x) \in \mathcal{G}(Y)$.

The unit $\eta\colon X \to \mathcal{G}(X)$ is given by $\eta(x)(M) = \mathbf{1}_M(x)$, where $\mathbf{1}_M\colon X \to [0,1]$ is the indicator function. Thus $\eta(x)(M) = 1$ if $x \in M$ and $\eta(x)(M) = 0$ if $x \notin M$, for each $M \in \Sigma_X$. The strength map $\mathrm{st}_1\colon \mathcal{G}(X) \times Y \to \mathcal{G}(X \times Y)$ is defined as the probability measure $\mathrm{st}_1(\omega, y)\colon \Sigma_{X\times Y} \to [0,1]$ determined by $M \times N \mapsto \omega(M) \cdot \eta(y)(N)$. The double strength map dst is determined by $\mathrm{dst}(\omega,\rho)(M \times N) = \omega(M) \times \rho(N)$. It makes the monad $\mathcal{G}$ commutative, via Fubini's theorem.

For a Kleisli map $f\colon X \to \mathcal{G}(Y)$, a predicate q on Y together with a state ω of X, we have:

$$\begin{aligned} f^*(q)(x) &= f(x) \models q \;=\; \int q \, \mathrm{d}f(x) \\ f_*(\omega)(N) &= \omega \models f(-)(N) \;=\; \int f(-)(N) \, \mathrm{d}\omega. \end{aligned}$$

More information can be found in [24,53,59,29].

3.3 *The expectation monad* $\mathcal{E}$ *on* **Sets**

There are two equivalent ways to define the expectation monad $\mathcal{E}$, using maps of effect modules (as in the original description from [38]), or using maps of C^*-algebras, see [22]. Here we shall follow the first approach, mainly because the second approach is very similar to the one used below for the Radon monad.

At the end of Section 2 we have seen the functor $[0,1]^{(-)}\colon \mathbf{Sets} \to \mathbf{EMod}^{\mathrm{op}}$. It is the basis for the expectation monad $\mathcal{E}$ on **Sets**, defined as the homset of effect module maps:

$$\mathcal{E}(X) = \mathbf{EMod}\big([0,1]^X, \, [0,1]\big).$$

For a function $f\colon X \to Y$ and an element $\omega \in \mathcal{E}(X)$ we define $\mathcal{E}(f)(\omega) \in \mathcal{E}(Y)$ as $\mathcal{E}(f)(\omega)(q) = \omega(q \circ f)$. That is, functoriality of $\mathcal{E}$ is given by:

$$\mathcal{E}\Big(X \xrightarrow{f} Y\Big)(\omega) = \Big([0,1]^Y \xrightarrow{(-)\circ f} [0,1]^X \xrightarrow{\omega} [0,1]\Big).$$

The unit $\eta\colon X \to \mathcal{E}(X)$ is $\eta(x)(p) = p(x)$. Kleisli composition of $f\colon X \to \mathcal{E}(Y)$ and $g\colon Y \to \mathcal{E}(Z)$ is defined for $x \in X$ and $q \in [0,1]^Z$ as:

$$(g \bullet f)(x)(q) = f(x)(\lambda y.\, g(y)(q)).$$

The monad $\mathcal{E}$ is strong, like any monad on **Sets**, but it does not seem to be commutative.

We have isomorphisms $\mathcal{E}(0) \cong 0$, and $\mathcal{E}(1) \cong \mathbf{EMod}([0,1],[0,1]) \cong 1$, and $\mathcal{E}(2) \cong \mathbf{EMod}([0,1]^2,[0,1]) \cong [0,1]$. Hence, predicates on a set X are fuzzy predicates $p \in [0,1]^X$, and scalars are probabilities, like for the monad $\mathcal{D}$ in Subsection 3.1. A state is a map $1 \to \mathcal{E}(X)$, and thus a map of effect algebras $\omega\colon [0,1]^X \to [0,1]$, as described above. The validity $\omega \models p$ is obtained simply by function application $\omega(p)$. For a Kleisli map $f\colon X \to \mathcal{E}(Y)$ we have predicate and state transformers:

$$f^*(q)(x) = f(x) \models q = f(x)(q) \qquad f_*(\omega)(q) = \omega \models f(-)(q) = \omega(f(-)(q)).$$

More information can be found in [38,22,39]. In [22, Lemma 4.1] it is shown that $\mathcal{E}(X)$ can equivalently be described as the set of states $\mathrm{Stat}(\ell^\infty(X))$ on the commutative C^*-algebra $\ell^\infty(X)$ of bounded functions $X \to \mathbb{C}$. This gives a clear similarity with the Radon monad described below, in Subsection 3.5.

3.4 The probabilistic powerdomain monad $\mathcal{V}$ on **Dcpo**

A (continuous) *valuation* on a dcpo X is a Scott continuous map $\omega\colon \mathcal{O}(X) \to [0,1]$ which satisfies $\omega(\emptyset) = 0$, $\omega(X) = 1$, and $\omega(U \cup V) = \omega(U) + \omega(V) - \omega(U \cap V)$ for all opens U, V. The requirement $\omega(X) = 1$ means that valuations as used here are normalised. Without this requirement we speak of 'sub-valuations'; they are standardly used in the theory of probabilistic powerdomains. We prefer to use proper, normalised valuations to obtain affineness, see Section 5. We write $\mathcal{V}(X)$ for the set of valuations on a dcpo X, ordered pointwise, with pointwise directed joins. This yields a dcpo again, and an endofunctor $\mathcal{V}\colon \mathbf{Dcpo} \to \mathbf{Dcpo}$, where $\mathcal{V}(f)(\omega)(U) = \omega\Big(f^{-1}(U)\Big)$, for $f\colon X \to Y$, $\omega \in \mathcal{V}(X)$ and $U \in \mathcal{O}(Y)$. This functor restricts to the category **CDcpo** of continuous dcpo's, see [44, Thm. 8.2].

It is not hard to see that $\mathcal{V}(1) \cong 1$ and $\mathcal{V}(2) \cong [0,1]$. A predicate on X thus corresponds to a continuous function $p\colon X \to [0,1]$. The unit interval $[0,1]$ is the set of scalars. A state $1 \to \mathcal{V}(X)$ is a valuation $\omega \colon \mathcal{O}(X) \to [0,1]$. Also in this domain-theoretic case one can define an integral $\int p \, \mathrm{d}\omega \in [0,1]$ as join of integrals of simple functions, see [43,44] for details. As we shall see, it is the validity $\omega \models p$.

This $\mathcal{V}$ forms a monad on (continuous) dcpo's, that is, on both the categories **Dcpo** and **CDcpo**. The unit $\eta \colon X \to \mathcal{V}(X)$ is given by $\eta(x)(U) = \mathbf{1}_U(x)$, where $\mathbf{1}_U \colon X \to [0,1]$ is the indicator function for U. For maps $f \colon X \to Y$ and $g \colon Y \to Z$ in $\mathcal{K}\ell(\mathcal{V})$ we have:

$$(g \bullet f)(x)(U) = f(x) \models g(-)(U) = \int g(-)(U) \, \mathrm{d}f(x).$$

This monad $\mathcal{V}$ is strong, with strength map $\mathrm{st}_1 \colon \mathcal{V}(X) \times Y \to \mathcal{V}(X \times Y)$ given by $\mathrm{st}_1(\omega, y)(U \times V) = \omega(U) \cdot \mathbf{1}_V(y)$. The induced 'double' strength $\mathrm{dst} \colon \mathcal{V}(X) \times \mathcal{V}(Y) \to \mathcal{V}(X \times Y)$ is given by $\mathrm{dst}(\omega, \psi)(U \times V) = \omega(U) \cdot \psi(V)$. This $\mathcal{V}$ is a commutative monad, by Fubini for $\mathcal{V}$.

Finally, the predicate and state transformers associated with $f \colon X \to \mathcal{V}(Y)$ are:

$$\begin{aligned} f^*(q)(x) &= f(x) \models q \;=\; \int q \, \mathrm{d}f(x) \\ f_*(\omega)(U) &= \omega \models f(-)(U) \;=\; \int f(-)(U) \, \mathrm{d}\omega. \end{aligned}$$

For more information, see *e.g.* [43,44,17,45,62].

3.5 The Radon monad $\mathcal{R}$ on **CH**

In order to describe the Radon monad $\mathcal{R}$ one starts from a compact Hausdorff space X, and forms the commutative C^*-algebra $\mathrm{C}(X)$ of continuous functions $\phi \colon X \to \mathbb{C}$, which are automatically bounded. It is basic result in the theory of C^*-algebras that the set of states $\mathrm{Stat}(\mathscr{A})$ on a C^*-algebra $\mathscr{A}$ is a compact Hausdorff space. These states are linear functions $\omega \colon \mathscr{A} \to \mathbb{C}$ which are positive and unital. Hence we define:

$$\mathcal{R}(X) = \mathrm{Stat}\big(\mathrm{C}(X)\big) \qquad \mathcal{R}\Big(X \xrightarrow{f} Y\Big)(\omega) = \Big(\mathrm{C}(Y) \xrightarrow{\mathrm{C}(f) = (-) \circ f} \mathrm{C}(X) \xrightarrow{\omega} \mathbb{C}\Big).$$

The unit is $\eta(x)(\phi) = \phi(x)$, and Kleisli composition is $(g \bullet f)(x)(\phi) = g(\lambda y.\, g(y)(\phi))$, like for the expectation monad. Again we have $\mathcal{R}(0) \cong 0$, $\mathcal{R}(1) \cong \mathrm{Stat}(\mathbb{C}) \cong 1$, and $\mathcal{R}(2) \cong \mathrm{Stat}(\mathbb{C} \times \mathbb{C}) \cong [0,1]$. The latter means that scalars are probabilities, and that predicates are continuous functions $X \to [0,1]$, forming maps in **CH**. A state is a map $1 \to \mathcal{R}(X)$, and thus a state $\omega \colon \mathrm{C}(X) \to \mathbb{C}$, as described above. The validity $\omega \models p$ is again simply function application $\omega(p)$. For a Kleisli map $f \colon X \to \mathcal{R}(Y)$ we have:

$$f^*(q)(x) = f(x) \models q = f(x)(q) \qquad f_*(\omega)(\psi) = \omega \models f(-)(\psi) = \omega(f(-)(\psi)).$$

The Radon monad occurs in [57,21,22]. The main result of [21], presented as a probabilistic version of Gelfand duality, states that the Kleisli category $\mathcal{K}\ell(\mathcal{R})$ of the Radon monad is the opposite $(\mathbf{CCstar}_{\mathrm{PU}})^{\mathrm{op}}$ of the category of commutative C^*-algebras, with positive unital maps between them. There is no (published) proof of commutativity of the Radon monad [1] . States/elements of $\mathcal{R}(X)$ correspond to 'Radon' (aka. 'inner regular') probability measures ω on the Borel sets $\mathcal{B}(X)$, see [60, Thm. 2.14]; they satisfy $\omega(S) = \sup_{K \subseteq S} \omega(K)$ where K ranges over compact sets.

3.6 The Kantorovich monad $\mathcal{K}$ on $\mathbf{Met}_1$

What we call the Kantorovich monad $\mathcal{K}$ looks like the earlier monads $\mathcal{G}$ and $\mathcal{V}$, but it acts on the category $\mathbf{Met}_1$ of 1-bounded metric spaces. The key ingredient of $\mathcal{K}$ is the metric that is defined on probability measures, which is commonly called the Kantorovich metric. This monad has been introduced in [8]. It assigns to a metric space X the set of probability measures $\mathcal{B}(X) \to [0,1]$, where $\mathcal{B}(X)$ is the Borel σ-algebra generated by the (metric) topology of X. As before we have $\mathcal{K}(0) \cong 0$, $\mathcal{K}(1) \cong 1$, and $\mathcal{K}(2) \cong [0,1]$, so that scalars are probabilities, and predicates on X are non-expansive functions $p \colon X \to [0,1]$. For a state $\omega \in \mathcal{K}(X)$ one can define an integral $\int p \, \mathrm{d}\omega \in [0,1]$, which amounts to validity $\omega \models p$.

In order to obtain a functor $\mathcal{K} \colon \mathbf{Met}_1 \to \mathbf{Met}_1$ one uses an appropriate distance function on $\mathcal{K}(X)$, going back to Kantorovich: for probability measures $\omega, \rho \colon \mathcal{B}(X) \to [0,1]$ take as distance:

$$d(\omega, \rho) = \bigvee \{ \, |\omega \models p - \rho \models p| \mid p \in \mathrm{Pred}(X) \}. \tag{8}$$

A proof that this definition makes $\mathcal{K}(X)$ a metric space can be found in [18, Prop. 2.5.14]. Each non-expansive function $f \colon X \to Y$ is continuous, and yields an inverse image function $f^{-1} \colon \mathcal{B}(Y) \to \mathcal{B}(X)$. Hence we can define $\mathcal{K}(f)(\omega) = \omega \circ f^{-1}$.

The unit, Kleisli composition, strength and commutativity, and predicate and state transformers for $\mathcal{K}$ are as for the Giry monad $\mathcal{G}$ and the probabilistic powerdomain monad $\mathcal{V}$.

There are several variations of the six monads that we have described above. For instance, instead of the monad $\mathcal{D}$ which captures discrete probability distributions with *finite* support, one can use the monad $\mathcal{D}_\infty$ having functions

[1] Robert Furber claims to have a proof and intends to publish it at some stage (private communication).

$\omega\colon X \to [0,1]$ with arbitrary support and sole requirement $\sum_x \omega(x) = 1$. In that case one can prove that the support is at most countable.

The expectation monad $\mathcal{E}(X) = \mathbf{EMod}([0,1]^X, [0,1])$ is of 'double dual' or 'continuation' form (see also [50]). It is isomorphic to the set of finitely additive measures on $\mathcal{P}(X)$, that is, to the homset $\mathbf{EA}(\mathcal{P}(X), [0,1])$. Abstractly the isomorphism arises from the fact that the set of predicates $[0,1]^X$ is isomorphic to the tensor product of effect algebras $[0,1] \otimes \mathcal{P}(X)$, see [37]. In [40] it is shown that the Giry monad $\mathcal{G}$ can equivalently be described in double dual form, as $\mathcal{G}(X) \cong \omega\text{-}\mathbf{EMod}(\mathbf{Meas}(X,[0,1]),[0,1])$, where ω-$\mathbf{EMod}$ is the category of ω-complete effect modules, with joins of ascending ω-chains (and maps preserving them); again this arises from a tensor product $\mathbf{Meas}(X,[0,1]) \cong [0,1] \otimes_\omega \Sigma_X$, in the category of ω-complete effect algebras. A similar Riesz-Markov-Kakutani style representation theorem exists for the double dual Radon monad, relating it to Radon measures (as already mentioned above). In [46,47] similar monads are studied on the category of ordered compact spaces. There is also a double dual monad $\mathbf{DcEMod}(\mathrm{C}(X,[0,1]),[0,1])$ on $\mathbf{Dcpo}$, see [31], that uses directed complete effect modules. In one direction, integration gives an injection $\mathcal{V}(X) \to \mathbf{DcEMod}(\mathrm{C}(X,[0,1]),[0,1])$. It is unclear whether this map is an isomorphism.

In the end, looking back at this series of examples, we see many similarities. For instance, all monad examples T are 'affine', in the sense that $T(1) \cong 1$; this makes 1 a final object in the Kleisli category $\mathcal{K}\ell(T)$. Further, in all case the set of scalars $1 \to T(2)$ is the unit interval $[0,1]$ of probabilities. We see that the Kantorovich metric (8) can be defined for all examples, since it only involves validity and structure of the unit interval (norm $|-|$ and join $\bigvee$). These similarities will be investigated further in subsequent sections, at a more abstract level. But we should be aware that this structure is rather special. For instance, the non-empty powerset monad $\mathcal{P}_+$ is affine, but has the three element set $\{0, 1, \top\}$ as scalars.

4 Partial maps and predicates

In Subsection 3.1 we have described a discrete probability distribution as a formal convex sum $\sum_i r_i|x_i\rangle$ with $\sum_i r_i = 1$. A *subdistribution* is such a formal sum with $\sum_i r_i \leq 1$. These subdistributions are used to handle partiality in the context of probabilistic computation, where the 'one deficit' $1-(\sum_i r_i)$ is used as probability of non-termination. These subdistributions can be captured as elements of the set $\mathcal{D}(X+1)$, where the 'lift' operation $X+1$ is used inside the monad $\mathcal{D}$. This is a technique that works more generally, and will be exploited in this section.

Let T now be a monad on a distributive category $\mathbf{C}$. The *lift* monad $(-)+1$ exists not only on the category $\mathbf{C}$, but also on $\mathcal{K}\ell(T)$, with unit and multiplication of the latter described in $\mathbf{C}$ as:

$$X \xrightarrow{\langle\kappa_1\rangle} T(X+1) \qquad (X+1)+1 \xrightarrow{\langle[\mathrm{id},\kappa_2]\rangle} T(X+1)$$

These maps are obtained via the functor $\langle - \rangle \colon \mathbf{C} \to \mathcal{K}\ell(T)$ from the unit and multiplication of the lift monad $(-)+1$ on $\mathbf{C}$. It is not hard to see that the Kleisli category of the lift monad $(-)+1$ on $\mathcal{K}\ell(T)$ is the Kleisli category of the monad $T' = T((-)+1)$ on $\mathbf{C}$. Hence we consider the category $\mathcal{K}\ell(T')$ as the category of *partial maps* in $\mathcal{K}\ell(T)$. The special maps in $\mathcal{K}\ell(T')$ of the form $\langle\kappa_1\rangle \bullet f = T(\kappa_1) \circ f$, for f in $\mathcal{K}\ell(T)$, are called *total*. Hence we consider $\mathcal{K}\ell(T)$ as a subcategory of total maps in the category $\mathcal{K}\ell(T')$ of partial maps; this is justified when, later on, under additional assumptions, the mapping $f \mapsto \langle\kappa_1\rangle \bullet f$ gives a faithful functor $\mathcal{K}\ell(T) \to \mathcal{K}\ell(T')$, see Lemma 21.

The unit η' and multiplication μ' of the monad T' are given by:

$$X \xrightarrow[\langle\kappa_1\rangle]{\eta' =} T(X+1) \qquad T\Big(T(X+1)+1\Big) \xrightarrow[\mu \circ T([\mathrm{id},\langle\kappa_2\rangle])]{\mu' =} T(X+1)$$

Abstractly, this T' is a monad since there is always a distributive law of monads $T(-)+1 \Rightarrow T((-)+1)$. In general, given such a law $ST \Rightarrow TS$, the composite TS is a monad again. Moreover, the monad S can be lifted to a monad $\overline{S}$ on $\mathcal{K}\ell(T)$, and its Kleisli category $\mathcal{K}\ell(\overline{S})$ is the same as the Kleisli category $\mathcal{K}\ell(TS)$ of the composite monad.

Kleisli composition in $\mathcal{K}\ell(T')$, written as $\bullet'$, is related to composition $\circ$ in $\mathbf{C}$ and to composition $\bullet$ in $\mathcal{K}\ell(T)$ via:

$$\begin{aligned} g \bullet' f \;=\; \mu' \circ T'(g) \circ f &= \mu \circ T([\mathrm{id}, \langle\kappa_2\rangle]) \circ T(g + \mathrm{id}) \circ f \\ &= \mu \circ T([g, \langle\kappa_2\rangle]) \circ f \;=\; [g, \langle\kappa_2\rangle] \bullet f. \end{aligned} \tag{9}$$

Moreover, if $h \colon X \to Y$ is map in $\mathcal{K}\ell(T)$, the corresponding total map $\langle\kappa_1\rangle \bullet h \colon X \to Y$ in $\mathcal{K}\ell(T')$ satisfies:

$$g \bullet' (\langle\kappa_1\rangle \bullet h) = [g, \langle\kappa_2\rangle] \bullet \langle\kappa_1\rangle \bullet h = g \bullet h.$$

We summarise and fix notation.

Definition 4 *For a monad T on a category $\mathbf{C}$ we write $T' = T((-)+1)$ for the associated 'partial map' monad. We write $\bullet'$ as in (9) for its Kleisli composition, with identity/unit $\eta' = \langle\kappa_1\rangle = \eta \circ \kappa_1$. Thus we will be working with three different categories with identity and composition notation as described below.*

$$(\mathbf{C}, \mathrm{id}, \circ) \qquad (\mathcal{K}\ell(T), \eta, \bullet) \qquad (\mathcal{K}\ell(T'), \langle\kappa_1\rangle, \bullet').$$

A map $f\colon X \to Y$ in $\mathcal{K}\ell(T')$ will be called a partial map from X to Y. We define its kernel predicate $\ker(f)$ *and ortho-kernel predicate* $\ker^{\perp}(f)$ *on X as:*

$$\ker^{\perp}(f) \stackrel{\text{def}}{=} \mathbf{1} \bullet' f = T(! + \mathrm{id}) \circ f \qquad \ker(f) = \Big(\ker^{\perp}(f)\Big)^{\perp}.$$

The monad T will be called affine *if $T(1) \cong 1$. The final object 1 in $\mathbf{C}$ is then also final $\mathcal{K}\ell(T)$, and the initial object 0 in $\mathbf{C}$ is a* zero object *in $\mathcal{K}\ell(T')$: it is both initial and final.*

With this notation we can describe a partial map from X to Y equivalently as:

$$X \longrightarrow T(Y+1) \text{ in } \mathbf{C} \qquad X \longrightarrow Y+1 \text{ in } \mathcal{K}\ell(T) \qquad X \longrightarrow Y \text{ in } \mathcal{K}\ell(T').$$

As a special case, predicates on X can be described equivalently as:

$$X \longrightarrow T(1+1) \text{ in } \mathbf{C} \qquad X \longrightarrow 1+1 \text{ in } \mathcal{K}\ell(T) \qquad X \longrightarrow 1 \text{ in } \mathcal{K}\ell(T').$$

We see that the description in the category $\mathcal{K}\ell(T')$ of partial maps is easiest.

Lemma 5 *Let T be a strong monad on a distributive category $\mathbf{C}$. The monad $T' = T((-)+1)$ is then also strong, with strength maps:*

$$\mathrm{st}_1' = \left(\begin{array}{c} T(X+1) \times Y \\ \big\downarrow{\scriptstyle \mathrm{st}_1} \\ T((X+1) \times Y) \\ \big\downarrow{\scriptstyle T(\mathrm{dis}_1^{-1})} \\ T((X \times Y) + (1 \times Y)) \\ \big\downarrow{\scriptstyle T(\mathrm{id}+\pi_1)} \\ T((X \times Y) + 1) \end{array}\right) \qquad \mathrm{st}_2' = \left(\begin{array}{c} X \times T(Y+1) \\ \big\downarrow{\scriptstyle \mathrm{st}_2} \\ T(X \times (Y+1)) \\ \big\downarrow{\scriptstyle T(\mathrm{dis}_2^{-1})} \\ T((X \times Y) + (X \times 1)) \\ \big\downarrow{\scriptstyle T(\mathrm{id}+\pi_2)} \\ T((X \times Y) + 1) \end{array}\right) \tag{10}$$

Proof. Via some elementary categorical reasoning one verifies that the above map st_1' makes Diagrams (4) and (5) commute, and yields the map st_2', via twisting both input and output with swap isomorphism $\langle \pi_2, \pi_1 \rangle$. $\square$

The monad T' is not automatically commutative if T is commutative, as the following counterexample shows. This implication 'T commutative $\Rightarrow T'$ commutative' requires an additional 'affiness' assumption, see Lemma 10 later on.

Example 6 *For the powerset monad $\mathcal{P}$ on* **Sets** *the strength maps $\mathrm{st}_1'\colon \mathcal{P}(X+1) \times Y \to \mathcal{P}((X \times Y)+1)$ and $\mathrm{st}_2'\colon X \times \mathcal{P}(Y+1) \to \mathcal{P}((X \times Y)+1)$ from Lemma 5 are described by:*

$$\begin{aligned} \mathrm{st}_1'(U, y) &= \{(x,y) \mid x \in U\} \cup \{* \mid * \in U\} \\ \mathrm{st}_2'(x, V) &= \{(x,y) \mid y \in V\} \cup \{* \mid * \in V\}. \end{aligned}$$

Moreover, the multiplication $\mu' \colon \mathcal{P}(\mathcal{P}(A+1)+1) \to \mathcal{P}(A+1)$ *is:*

$$\mu'(W) = \{a \in A \mid \exists U \in W.\, a \in U\} \cup \{* \mid \exists U \in W.\, * \in W\} \cup \{* \mid * \in W\}.$$

The two paths in (7) *are different on* $(\emptyset, \{*\}) \in \mathcal{P}(X+1) \times \mathcal{P}(Y+1)$ *since:*

$$\begin{array}{l} \Big(\mu' \circ \mathcal{P}(\mathrm{st}_2') \circ \mathrm{st}_1'\Big)(\emptyset, \{*\}) = \Big(\mu' \circ \mathcal{P}(\mathrm{st}_2')\Big)(\emptyset) \;=\; \mu'(\emptyset) \;=\; \emptyset \\ \Big(\mu' \circ \mathcal{P}(\mathrm{st}_1') \circ \mathrm{st}_2'\Big)(\emptyset, \{*\}) = \Big(\mu' \circ \mathcal{P}(\mathrm{st}_1')\Big)(\{*\}) \;=\; \mu'(\{*\}) \;=\; (\{*\}). \end{array}$$

In general, a monad S is called *additive* if it sends finite coproducts to products, via (canonical) isomorphisms $S(0) \cong 1$ and $S(X+Y) \cong S(X) \times S(Y)$. For instance, the powerset monad is additive. In [13] it is shown that a monad S is additive iff the coproducts $(0,+)$ of its Kleisli category $\mathcal{K}\ell(S)$ are biproducts iff the products $(1, \times)$ of its category $\mathcal{EM}(S)$ of Eilenberg-Moore algebras are biproducts.

If T is an affine monad, then $T'(0) = T(0+1) \cong T(1) \cong 1$. But the additivity requirement $T'(X+Y) \cong T'(X) \times T'(Y)$ does not hold for our examples. Instead, a weaker property holds, called 'partial additivity', see Definition 7 below. This means that coproducts $+$ in $\mathcal{K}\ell(T')$ are not biproducts. But as we shall see, they behave a bit like products, and do have 'partial projections', written as $\rhd_i$.

Let T thus be an affine monad, on a distributive category $\mathbf{C}$. As mentioned, the initial object $0 \in \mathbf{C}$ is a zero object in the Kleisli category $\mathcal{K}\ell(T')$ of partial maps. Explicitly, for each pair of objects $X, Y \in \mathbf{C}$ there is a *zero map* $\mathbf{0} = \mathbf{0}_{X,Y} \colon X \to T'(Y)$ satisfying:

$$\mathbf{0}_{X,Y} = \Big(X \xrightarrow{!} 1 \cong T'(0) \xrightarrow{T'(!)} T'(Y)\Big) = \Big(X \xrightarrow{!} 1 \xrightarrow{\langle \kappa_2 \rangle} T(Y+1)\Big).$$

We have $\mathbf{0} \bullet' f = \mathbf{0} = g \bullet' \mathbf{0}$ for all maps f, g in $\mathcal{K}\ell(T')$. We can now define 'partial projections' $\rhd_1 \colon X+Y \to X$ and $\rhd_2 \colon X+Y \to Y$ in $\mathcal{K}\ell(T')$ via cotuples:

$$\rhd_1 \stackrel{\text{def}}{=} \Big(X + Y \xrightarrow[= \langle \mathrm{id}+! \rangle]{[\langle \kappa_1 \rangle, \mathbf{0}]} T(X+1)\Big) \qquad \rhd_2 \stackrel{\text{def}}{=} \Big(X + Y \xrightarrow[= \langle [\kappa_2 \circ !, \kappa_1] \rangle]{[\mathbf{0}, \langle \kappa_1 \rangle]} T(Y+1)\Big).$$

These maps are natural in X, Y, in the category $\mathcal{K}\ell(T')$, and satisfy $\rhd_1 \circ [\kappa_2, \kappa_1] = \rhd_2$. Notice that on $1+1 = 2$ the first projection $\rhd_1 \colon 1+1 \to T(1+1)$ is the unit/identity and second projection $\rhd_2 \colon 1+1 \to T(1+1)$ is the swap map $\langle [\kappa_2, \kappa_1] \rangle$.

We can then form 'bicartesian' maps $\mathrm{bc} = \mathrm{bc}_{X,Y} \colon T'(X+Y) \to T'(X) \times T'(Y)$, as a tuple of the Kleisli liftings of $\rhd_1, \rhd_2$ connecting coproducts and products.

Explicitly:

$$\mathrm{bc} \stackrel{\text{def}}{=} \langle\, \mu' \circ T'(\rhd_1),\ \mu' \circ T'(\rhd_2)\,\rangle. \tag{11}$$

For an additive monad these maps bc are isomorphisms, see [13]. We'll use a weaker requirement.

Definition 7 (After [27]) *An affine monad T on a distributive category $\mathbf{C}$ is* partially additive *if it is affine and if these maps* bc *from* (11) *are monic in $\mathbf{C}$, and the naturality squares below are pullbacks in $\mathbf{C}$, for all $f\colon X \to A$, $g\colon Y \to B$ in $\mathbf{C}$ — where, recall, $T' = T((-)+1)$.*

$$\begin{tikzcd}
T'(X+Y) \arrow[r, "T'(f+g)"] \arrow[d, tail, "\mathrm{bc}"'] & T'(A+B) \arrow[d, tail, "\mathrm{bc}"] \\
T'(X) \times T'(Y) \arrow[r, "T'(f)\times T'(g)"'] & T'(A) \times T'(B)
\end{tikzcd} \tag{12}$$

We observe that the requirement that the map bc is monic means that the two partial projections $\rhd_1\colon X+Y \to X, \rhd_2\colon X+Y \to Y$ are jointly monic in $\mathcal{K}\ell(T')$. In particular, the following two maps in $\mathcal{K}\ell(T)$ are jointly monic (see [30, Assump. 1]).

$$\begin{tikzcd}
(1+1)+1 \arrow[r, bend left, "\mathsf{IV} = {[\rhd_1, \kappa_2]} = {[\mathrm{id}, \kappa_2]}"] \arrow[r, bend right, "\mathsf{XV} = {[\rhd_2, \kappa_2]} = {[[\kappa_2, \kappa_1], \kappa_2]}"'] & 1+1
\end{tikzcd} \tag{13}$$

For a partially additive monad T we can define a partial sum operation $\ovee$ on the homsets of the Kleisli category $\mathcal{K}\ell(T')$ of partial maps, as in [5,27], and in [30,9,10]. This sum $\ovee$ then exists in particular for predicates. We recall the construction and prove some basic results. Stronger results will be obtained later on, in Lemma 22, under additional assumptions.

- First, two partial maps $f, g\colon X \to T'(Y)$ are called *orthogonal*, written as $f \perp g$, if there is a (necessarily unique) bound $b\colon X \to T'(Y+Y)$ such that $\mathrm{bc} \circ b = \langle f, g\rangle$, *i.e.* such that $\rhd_1 \bullet' b = f$ and $\rhd_2 \bullet' b = g$.
- Next, if $f \perp g$ via bound b, then we define their sum $\ovee$ by $f \ovee g = \nabla \bullet' b = T(\nabla + \mathrm{id}) \circ b\colon X \to T'(Y)$, where $\nabla = [\mathrm{id}, \mathrm{id}]\colon Y+Y \to Y$ is the codiagonal.

Lemma 8 *For a partially additive monad T on a distributive category $\mathbf{C}$,*

(i) the set of maps $X \to T(Y+1)$ in $\mathbf{C}$, that is, the homset of maps $X \to Y$ in $\mathcal{K}\ell(T')$, is a 'partial commutative monoid' (PCM) via $(\mathbf{0}, \ovee)$;
(ii) this structure $(\mathbf{0}, \ovee)$ is preserved by pre- and post-composition in $\mathcal{K}\ell(T')$;
(iii) scalar multiplication satisfies $s \cdot \mathbf{0} = \mathbf{0}$ and $s \cdot (p \ovee q) = (s \cdot p) \ovee (s \cdot q)$; this scalar multiplication is preserved pre-composition;

(iv) the 'untying axiom' of [5,56,9] holds: if $f \perp g$ *then* $(\kappa_1 \bullet' f) \perp (\kappa_2 \bullet' g)$*;*
(v) if $(\mathbf{1} \bullet' f) \perp (\mathbf{1} \bullet' g)$ *then* $f \perp g$*;*
(vi) $p \ovee p^{\perp} = \mathbf{1}$*, for each predicate* p*.*

The three points (i), (ii) and (iv) say that the category $\mathcal{K}\ell(T')$ is a finitely partially additive category (a FinPAC, for short, see [5,9]).

Proof. (i) The operation $\ovee$ is obviously commutative: if $b \colon X \to Y + Y$ in $\mathcal{K}\ell(T')$ is a bound for f, g, then $[\kappa_2, \kappa_1] \bullet' b$ is a bound g, f. Next, the zero map $\mathbf{0} \colon X \to Y$ in $\mathcal{K}\ell(T')$ is a unit for $\ovee$: the equation $f \ovee \mathbf{0} = f$ is obtained via the bound $b = \kappa_1 \bullet' f$. What requires more care is (partial) associativity: Let $f, g, h \colon X \to Y$ be given in $\mathcal{K}\ell(T')$ with $f \perp g$ via bound b, and $(f \ovee g) \perp h$ via bound c. We thus have $\rhd_1 \bullet' b = f$, $\rhd_2 \bullet' b = g$ and $\rhd_1 \bullet' c = f \ovee g = \nabla \bullet' b$, $\rhd_2 \bullet' c = h$. Consider the following pullback in the underlying category $\mathbf{C}$.

$$\begin{tikzcd}
X \arrow[dr, dotted, "d"] \arrow[drr, bend left, "c"] \arrow[ddr, bend right, "\langle b,h\rangle"'] & & \\
& T'((Y+Y)+Y) \arrow[r, "T'(\nabla+\mathrm{id})"] \arrow[d, tail, "\mathrm{bc}"'] \arrow[dr, phantom, "\lrcorner", very near start] & T'(Y+Y) \arrow[d, tail, "\mathrm{bc}"] \\
& T'(Y+Y) \times T'(Y) \arrow[r, "T'(\nabla)\times\mathrm{id}"'] & T'(Y) \times T'(Y)
\end{tikzcd}$$

Take $d' = T([[[\kappa_2 \circ !, \kappa_1 \circ \kappa_1], \kappa_1 \circ \kappa_2], \kappa_2]) \circ d \colon X \to T'((Y+Y)+Y) = T(((Y+Y)+Y)+1) \to T((Y+Y)+1) = T'(Y+Y)$. We leave it to the reader to check $\rhd_1 \bullet' d' = g$ and $\rhd_2 \bullet' d' = h$, so that d' proves $g \perp h$.

Next we take $d'' = T'([\mathrm{id}, \kappa_2]) \circ d \colon X \to T'(Y+Y)$. One can prove $\rhd_1 \bullet' d'' = f$ and $\rhd_2 \bullet' d'' = g \ovee h$, so that d'' proves $f \perp (g \ovee h)$. We now obtain associativity:

$$\begin{aligned}
f \ovee (g \ovee h) &= \nabla \bullet' d'' \\
&= T([\kappa_1 \circ \nabla, \kappa_2]) \circ T([\mathrm{id}, \kappa_2] + \mathrm{id}) \circ d \\
&= T([\kappa_1 \circ \nabla \circ [\mathrm{id}, \kappa_2], \kappa_2]) \circ d \\
&= T([\kappa_1 \circ [[\mathrm{id}, \mathrm{id}], \mathrm{id}], \kappa_2]) \circ d \\
&= T([\kappa_1 \circ [\mathrm{id}, \mathrm{id}] \circ [\kappa_1 \circ [\mathrm{id}, \mathrm{id}], \kappa_2], \kappa_2]) \circ d \\
&= T((\nabla \circ (\nabla + \mathrm{id})) + \mathrm{id}) \circ d \\
&= T(\nabla + \mathrm{id}) \circ T((\nabla + \mathrm{id}) + \mathrm{id}) \circ d \\
&= \nabla \bullet \big(T'(\nabla + \mathrm{id}) \circ d\big) \\
&= \nabla \bullet c \\
&= (f \ovee g) \ovee h.
\end{aligned}$$

(ii) Sums $\ovee$ are preserved by pre-composition in $\mathcal{K}\ell(T')$, that is $(f \bullet' h) \ovee (g \bullet'$

$h) = (f \ovee g) \bullet' h$. Indeed, if b is a bound for f, g, then obviously $b \bullet' h$ is a bound for $f \bullet' h$ and $g \bullet' h$, proving preservation of sums.

Sums $\ovee$ are also preserved by post-composition in $\mathcal{K}\ell(T')$, that is: $(h \bullet' f) \ovee (h \bullet' g) = h \bullet' (f \ovee g)$. If b is a bound for f, g, then $(h + h) \bullet' b$ is a bound for $h \bullet' f$ and $h \bullet' g$, and thus:

$$\begin{aligned} (h \bullet' f) \ovee (h \bullet' g) &= \nabla \bullet' (h + h) \bullet' b \\ &= h \bullet' \nabla \bullet' b \\ &= h \bullet' (f \ovee g). \end{aligned}$$

(iii) This follows directly from the previous point, since scalar multiplication $s \cdot p$ equals $s \bullet' p$, see Definition 2 (iii) and the description of $\bullet'$ in (9).

(iv) For the untying axiom, let $f \perp g$, for $f, g \colon X \to T'(Y)$, via bound $b \colon X \to T'(Y + Y)$. One can take as new bound $b' = T'(\kappa_1 + \kappa_2) \circ b \colon X \to T'((Y+Y)+(Y+Y))$. It is easy to see that b' proves $(\kappa_1 \bullet' f) \perp (\kappa_2 \bullet' g)$.

(v) Let $(\mathbf{1} \bullet' f) \perp (\mathbf{1} \bullet' g)$, for $f, g \colon X \to T'(Y)$, via bound $b \colon X \to T'(1+1)$. Then we use the following pullback instance of (12).

$$\begin{tikzcd}
X \arrow[drr, bend left, "b"] \arrow[dr, dotted, "c"] \arrow[ddr, bend right, "\langle f,g \rangle"'] & & \\
& T'(Y+Y) \arrow[r, "T'(!+!)"'] \arrow[d, tail, "\mathrm{bc}"'] \arrow[dr, phantom, very near start, "\lrcorner"] & T'(1+1) \arrow[d, tail, "\mathrm{bc}"] \\
& T'(Y) \times T'(Y) \arrow[r, "T'(!)\times T'(!)"'] & T'(1) \times T'(1)
\end{tikzcd}$$

The map c is by construction a bound for f, g, showing $f \perp g$.

(vi) Let $p \colon X \to T(2) = T'(1)$ be a predicate. We take as bound $b = T(\kappa_1) \circ p \colon X \to T'(1 + 1) = T((1 + 1) + 1)$. One easily checks that $\rhd_1 \bullet' b = p$ and $\rhd_2 \bullet' b = p^\perp$, and also that $p \ovee p^\perp = \nabla \bullet' b = \mathbf{1}$. $\square$

At the end of this section we return to our running monad examples from the previous section. All these monads are partially additive. Showing this is not so interesting, and so we concentrates on partial maps and predicates, and on their partially additive structure $(\mathbf{0}, \ovee)$.

Example 9 *In all the monad examples in Subsection 3.1 – 3.6 the predicates on an object X are maps of the form $X \to [0,1]$, of some sort (measurable, continuous, non-expansive, ...). In each of these cases the partial sum $p \ovee q$ of $p, q \colon X \to [0,1]$ exists — that is, p and q are orthogonal: $p \perp q$ — iff $p(x) + q(x) \leq 1$ for all $x \in X$. In that case their sum $p \ovee q \colon X \to [0,1]$ is defined as $(p \ovee q)(x) = p(x) + q(x)$. This $\ovee$ is obviously commutative and associative, with unit element $\mathbf{0}$, given by $\mathbf{0}(x) = 0$. Moreover, the orthosupplement $p^\perp \colon X \to [0,1]$ is given by $p^\perp(x) = 1 - p(x)$, so that indeed $p \ovee p^\perp = \mathbf{1}$.*

We briefly look at partial maps and their partial sum $\ovee$. These partial maps *correspond in each case to 'sub' distribution/measures, where the total probability is not equal to one, but less than one.*

(i) *For the distribution monad* $\mathcal{D}$ *a partial map* $X \to Y$ *is a function* $f\colon X \to \mathcal{D}(Y+1)$. *Its kernel predicate* $\ker(f) \in [0,1]^X$ *is* $\ker(f)(x) = f(x)(*) = 1-(\sum_y f(x)(y))$. *This is the 'one-deficit' predicate that captures the probability of non-termination. Two parallel partial maps* f, g *are orthogonal if for each* x,

$$\textstyle\sum_y f(x)(y) + g(x)(y) \leq 1 \qquad \text{that is} \qquad f(x)(*) + g(x)(*) \geq 1.$$

In that case $(f \ovee g)(x)(y) = f(x)(y) + g(x)(y)$.

(ii) *For the Giry monad* $\mathcal{G}$, *a partial map* $X \to Y$ *is a measurable function* $f\colon X \to \mathcal{G}(Y+1)$. *Its kernel* $\ker(f)\colon X \to [0,1]$ *is given by* $\ker(f)(x) = f(x)(\{*\}) = 1 - f(x)(Y)$. *We now have* $f \perp g$ *iff* $f(x)(Y) + g(x)(Y) \leq 1$ *for each* x*; in that case* $(f \ovee g)(x)(N) = f(x)(N) + g(x)(N)$ *for* $N \in \Sigma_Y$. *A similar description applies to the probabilistic powerdomain monad* $\mathcal{V}$ *and to the Kantorovich monad* $\mathcal{K}$. *We note that the monad* $\mathcal{V}'(X) = \mathcal{V}(X+1)$ *contains sub-valuations* $\rho\colon \mathcal{O}(X) \to [0,1]$, *which need not satisfy* $\rho(X) = 1$. *They are commonly used in probabilistic domain theory.*

(iii) *For the expectation monad* $\mathcal{E}$ *we first notice that:*

$$\mathcal{E}(Y+1) = \mathbf{EMod}([0,1]^{Y+1}, [0,1]) \cong \mathbf{EMod}([0,1]^Y \times [0,1], [0,1]).$$

These effect module maps $\omega\colon [0,1]^Y \times [0,1] \to [0,1]$ *can be identified with 'substate' functions* $[0,1]^Y \to [0,1]$ *that preserve* $0, \ovee$ *and scalar multiplication, but not the unit* 1. *For a partial map* $X \to Y$ *for* $\mathcal{E}$, *that is, for a function* f *from* X *to such substates, the kernel* $\ker(f) \in [0,1]^X$ *captures non-termination, via* $\ker(f)(x) = 1 - f(x)(\mathbf{1})$. *Two parallel maps* f, g *are orthogonal iff* $f(x)(\mathbf{1}) + g(x)(\mathbf{1}) \leq 1$ *for all* x, *and in that case one has* $(f \ovee g)(x)(p) = f(x)(p) + g(x)(p)$, *where* $p \in [0,1]^Y$.

(iv) *For the Radon monad* $\mathcal{R}$ *we have, similarly to the previous point:*

$$\mathcal{R}(Y+1) = \mathrm{Stat}(\mathrm{C}(Y+1)) \cong \mathrm{Stat}(\mathrm{C}(Y) \times \mathbb{C}) \cong \mathrm{SubStat}(\mathrm{C}(Y)),$$

where, in general, $\mathrm{SubStat}(\mathscr{A})$ *is the set of positive and* subunital *maps* $\omega\colon \mathscr{A} \to \mathbb{C}$, *satisfying* $\omega(1) \leq 1$, *instead of* $\omega(1) = 1$. *A partial map* $X \to Y$ *for the Radon monad* $\mathcal{R}$ *can thus be identified with a function* $f\colon X \to \mathrm{SubStat}(\mathrm{C}(Y))$. *Its kernel* $\ker(f)\colon X \to [0,1]$ *is the continuous function given by* $\ker(f)(x) = 1 - f(x)(\mathbf{1})$. *As before, two parallel maps* f, g *are orthogonal iff* $f(x)(\mathbf{1}) + g(x)(\mathbf{1}) \leq 1$ *for all* x, *and in that case one has* $(f \ovee g)(x)(\phi) = f(x)(\phi) + g(x)(\phi)$, *where* $\phi \in \mathrm{C}(Y)$.

5 Affineness and strong affiness of monads

This section first recalls the basic theory of affine monads — which preserve the final object 1 — following [51,55,25]. It then digs deeper into affineness and introduces a slightly stronger notion, called 'strong affiness', following [33]. We describe basic properties and examples. Strong affineness will have two roles in the sequel of the paper:

- it allows us to prove stronger properties about the partial monoid structure $(\mathbf{0}, \otimes\!\!\!\!\bigcirc)$ from the previous section, see Lemma 22 in the next section;
- it implies that instruments that will be associated with predicates are side-effect-free, and gives a bijective correspondence between predicates and such instruments, see Proposition 27.

The first point where we need affiness is for an extension of Lemma 5.

Lemma 10 *Let T be an affine commutative monad on a distributive category. The associated monad $T' = T((-)+1)$ is then also commutative, with 'double strength' map:*

$$\mathrm{dst}' = \left(\begin{array}{c} T(X+1) \times T(Y+1) \\ \big\downarrow{\scriptstyle \mathrm{dst}} \\ T((X+1) \times (Y+1)) \\ \big\downarrow{\scriptstyle T(\mathrm{ddis})} \\ T((X \times Y) + (1 \times Y) + (X \times 1) + (1 \times 1)) \\ \big\downarrow{\scriptstyle T([\kappa_1, \kappa_2 \circ !, \kappa_2 \circ !, \kappa_2 \circ !])} \\ T((X \times Y) + 1) \end{array} \right) \tag{14}$$

The map ddis *is the obvious 'double distributivity' isomorphism, combining* dis_1 *and* dis_2 *from the beginning of Section 3.*

As a result, not only $\mathcal{K}\ell(T)$ is a symmetric monoidal category, but also the category $\mathcal{K}\ell(T')$ of partial maps. The tensor in the latter category will be written as $\otimes'$.

Proof. We know from Lemma 5 that the monad T' is also strong, with strength maps $\mathrm{st}_1', \mathrm{st}_2'$ as described there. We have to verify that the two paths $\mu' \circ T'(\mathrm{st}_2') \circ \mathrm{st}_1'$ and $\mu' \circ T'(\mathrm{st}_1') \circ \mathrm{st}_2'$ in Diagram (7) are the same. This involves a lengthy computation, where we indicate via marked equations $\stackrel{\text{(aff)}}{=}$

and $\stackrel{\text{(com)}}{=}$ the use of the assumptions that T is affine resp. commutative.

$$\begin{array}{cl}
\multicolumn{2}{l}{\mu' \circ T'(\mathrm{st}_2') \circ \mathrm{st}_1'} \\
\stackrel{(10)}{=} & \mu \circ T([\mathrm{id}, T(\kappa_2) \circ \eta]) \circ T(\mathrm{st}_2' + \mathrm{id}) \circ T(\mathrm{id} + \pi_1) \circ T(\mathrm{dis}_1^{-1}) \circ \mathrm{st}_1 \\
= & \mu \circ T([\mathrm{id}, T(\kappa_2)]) \circ T(\mathrm{st}_2' + (\eta \circ \pi_1)) \circ T(\mathrm{dis}_1^{-1}) \circ \mathrm{st}_1 \\
\stackrel{\text{(aff)}}{=} & \mu \circ T([\mathrm{id}, T(\kappa_2)]) \circ T(\mathrm{st}_2' + (T(\pi_1) \circ \mathrm{st}_2)) \circ T(\mathrm{dis}_1^{-1}) \circ \mathrm{st}_1 \\
\stackrel{(10)}{=} & \mu \circ T([\mathrm{id}, T(\kappa_2)]) \circ T(T(\mathrm{id} + \pi_2) + T([\pi_1, \pi_1])) \\
& \qquad \circ\ T(T(\mathrm{dis}_2^{-1}) + T(\mathrm{dis}_2^{-1})) \circ T(\mathrm{st}_2 + \mathrm{st}_2) \circ T(\mathrm{dis}_1^{-1}) \circ \mathrm{st}_1 \\
= & \mu \circ T^2([\mathrm{id} + !, !]) \circ T([T(\kappa_1), T(\kappa_2)]) \circ T(T(\mathrm{dis}_2^{-1}) + T(\mathrm{dis}_2^{-1})) \\
& \qquad \circ\ T(\mathrm{st}_2 + \mathrm{st}_2) \circ T(\mathrm{dis}_1^{-1}) \circ \mathrm{st}_1 \\
= & \mu \circ T^2([\mathrm{id} + !, !]) \circ T(\mathrm{dis}_2^{-1} + \mathrm{dis}_2^{-1}) \circ T([T(\kappa_1), T(\kappa_2)]) \\
& \qquad \circ\ T(\mathrm{st}_2 + \mathrm{st}_2) \circ T(\mathrm{dis}_1^{-1}) \circ \mathrm{st}_1 \\
\stackrel{(6)}{=} & \mu \circ T^2([\mathrm{id} + !, !]) \circ T^2(\mathrm{dis}_2^{-1} + \mathrm{dis}_2^{-1}) \circ T^2(\mathrm{dis}_1^{-1}) \circ T(\mathrm{st}_2) \circ \mathrm{st}_1 \\
= & T([\mathrm{id} + !, !]) \circ T(\mathrm{dis}_2^{-1} + \mathrm{dis}_2^{-1}) \circ T(\mathrm{dis}_1^{-1}) \circ \mu \circ T(\mathrm{st}_2) \circ \mathrm{st}_1 \\
\stackrel{\text{(com)}}{=} & T([\mathrm{id} + !, !]) \circ T(\mathrm{dis}_1^{-1} + \mathrm{dis}_1^{-1}) \circ T(\mathrm{dis}_2^{-1}) \circ \mu \circ T(\mathrm{st}_1) \circ \mathrm{st}_2 \\
= & \cdots \quad \text{(as before)} \\
= & \mu' \circ T'(\mathrm{st}_1') \circ \mathrm{st}_2'.
\end{array}$$

□

It is known for a long time that the 'affine part' of a monad can be extracted via pullbacks, see [55] (or also [25]). Here we shall relate this affine part to 'causal' maps in Kleisli categories of monads.

Proposition 11 *Let T be a monad on a category $\mathbf{C}$ with a final object 1. Assume that the pullbacks below exist in $\mathbf{C}$, for each object X. This defines a mapping $X \mapsto T_a(X)$.*

$$\begin{tikzcd}
T_a(X) \arrow[r, "!"] \arrow[d, "\iota_X"'] \arrow[dr, phantom, "\lrcorner", very near start] & 1 \arrow[d, "\eta"] \\
T(X) \arrow[r, "T(!)"'] & T(1)
\end{tikzcd} \tag{15}$$

Then:

(i) this mapping $X \mapsto T_a(X)$ is a monad on $\mathbf{C}$;
(ii) the mappings $\iota_X : T_a(X) \to T(X)$ are monic, and form a map of monads $T_a \Rightarrow T$;
(iii) T_a is an affine monad, and in fact the universal (greatest) affine sub-monad of T;
(iv) if T is a strong resp. commutative monad, then so is T_a.

Proof. These results are standard. We shall illustrate point (iii). If we take $X = 1$ in Diagram (15), then the bottom arrow $T(!_X) : T(X) \to T(1)$ is the identity. Hence top arrow $T_a(1) \to 1$ is an isomorphism, since isomorphisms are preserved under pullback.

To see that $T_a \Rightarrow T$ is universal, let $\sigma\colon S \Rightarrow T$ be a map of monads, where S is affine, then we obtain a map $\overline{\sigma}_X$ in:

$$\begin{tikzcd}
S(X) \arrow[dr, dotted, "\overline{\sigma}_X"] \arrow[drr, bend left, "!_{S(X)}"] \arrow[ddr, bend right, "\sigma_X"'] & & \\
& T_a(X) \arrow[r, "!"] \arrow[d, tail, "\iota_X"'] \arrow[dr, phantom, "\lrcorner", very near start] & 1 \arrow[d, "\eta^T"] \\
& T(X) \arrow[r, "T(!)"'] & T(1)
\end{tikzcd}$$

The outer diagram commutes since S is affine, so that $\eta^S_1 \circ !_{S(1)} = \mathrm{id}_{S(1)}$; then:

$$T(!_X) \circ \sigma_X = \sigma_1 \circ S(!_X) = \sigma_1 \circ \eta^S_1 \circ !_{S(1)} \circ S(!_X) = \eta^T_1 \circ !_{S(X)}. \qquad \square$$

For the record we recall from [63,2] that each endofunctor on **Sets** can be written as a coproduct of affine functors.

Example 12 *We list several examples of affine parts of monads.*

(i) *Let $\mathcal{M} = \mathcal{M}_{\mathbb{R}_{\geq 0}}$ be the multiset monad on* **Sets** *with the non-negative real numbers $\mathbb{R}_{\geq 0}$ as scalars. Elements of $\mathcal{M}(X)$ are thus finite formal sums $\sum_i r_i|x_i\rangle$ with $r_i \in \mathbb{R}_{\geq 0}$ and $x_i \in X$. The affine part $\mathcal{M}_a$ of this monad is the distribution monad $\mathcal{D}$ since $1|\ast\rangle = \mathcal{M}(!)(\sum_i r_i|x_i\rangle) = (\sum_i r_i)|\ast\rangle$ iff $\sum_i r_i = 1$. Thus $\mathcal{D}(X) = \mathcal{M}_a(X)$ yields a pullback in Diagram* (15).

The monad $\mathcal{D}_\pm$ used in Example 18 can be obtained in a similar manner as an affine part, not of the multiset monad $\mathcal{M}_{\mathbb{R}_{\geq 0}}$ with non-negative coefficients, but from the multiset monad $\mathcal{M}_{\mathbb{R}}$ with arbitrary coefficients: its multisets are formal sums $\sum_i r_i|x_i\rangle$ where the r_i are arbitrary real numbers.

(ii) *For the powerset monad $\mathcal{P}$ on* **Sets** *the affine submonad $\mathcal{P}_a \rightarrowtail \mathcal{P}$ is given by the* non-empty *powerset monad. Indeed, for a subset $U \subseteq X$ we have:*

$$\mathcal{P}(!)(U) = \{!(x) \mid x \in U\} = \{\ast \mid x \in U\} = \begin{cases} \{\ast\} & \textit{if } U \neq \emptyset \\ \emptyset & \textit{if } U = \emptyset \end{cases}$$

Hence $\mathcal{P}(!)(U) = \{\ast\} = \eta(\ast)$ iff U is non-empty. It is not hard to see that the non-empty powerset monad $\mathcal{P}_a$ is strongly affine.

(iii) *Let $T(X) = (S \times X)^S$ be the state monad on* **Sets**, *for a fixed set of states S. In this example the word 'state' refers to all the information stored in memory, to which a program has access, via reading and writing. The unit $\eta\colon X \to T(X)$ is defined as $\eta(x) = \lambda s \in S.\,(s, x)$ so that the pullback* (15)

is given by:

$$\begin{aligned} T_a(X) &= \{f \in (S \times X)^S \mid T(!)(f) = \eta(*)\} \\ &= \{f \in (S \times X)^S \mid \forall s.\, (\mathrm{id} \times !)(f(s)) = (s, *)\} \\ &= \{f \in (S \times X)^S \mid \forall s.\, \pi_1 f(s) = s\} \\ &\cong X^S. \end{aligned}$$

Thus, Kleisli maps $Y \to T_a(X) = X^S$ may use states $s \in S$ to compute the output in X, but they cannot change states: they are side-effect-free. This theme will be elaborated in Section 7.

In a similar way one shows that the list monad $X \mapsto X^\star$ and the lift monad $X \mapsto X + 1$ have the identity monad as their affine submonad.

(iv) Fix a set C and consider the continuation, (or double-dual) monad $\mathcal{C}$ on **Sets** *given by $\mathcal{C}(X) = C^{(C^X)}$, with unit $\eta\colon X \to \mathcal{C}(X)$ given by $\eta(x)(f) = f(x)$. The pullback* (15) *is then:*

$$\begin{aligned} \mathcal{C}_a(X) &= \{h \in C^{(C^X)} \mid \mathcal{C}(!)(h) = \eta(*)\} \\ &= \{h \in C^{(C^X)} \mid \forall f \in C^1.\, h(f \circ !) = f(*)\} \\ &= \{h \in C^{(C^X)} \mid \forall c \in C.\, h(\lambda x.\, c) = c\}. \end{aligned}$$

This is the submonad of functions $h\colon C^X \to C$ which have output $c \in C$ on the constant function $\lambda x.\, c\colon X \to C$.

For each monad T and each object X there is a special 'ground' map:

$$\bar{\dagger}_X = \langle !_X \rangle = \Big(X \xrightarrow{!_X} 1 \xrightarrow{\eta_1} T(1)\Big) = \Big(X \xrightarrow{\eta_X} T(X) \xrightarrow{T(!_X)} T(1)\Big) \tag{16}$$

Thus, $\bar{\dagger}_X$ is a map $X \to 1$ in $\mathcal{K}\ell(T)$, or equivalently, a map $X \to 0$ in $\mathcal{K}\ell(T')$.

Below we use these ground maps $\bar{\dagger}$ to define 'causal' maps. They have been introduced in the context of CP*-categories, see [12], where they express the property that measurements in the future, given by $\bar{\dagger}$, cannot influence the past.

Definition 13 *A Kleisli map $f\colon X \to T(Y)$ will be called* causal *or* unital *if it preserves ground, in the sense that:*

$$\bar{\dagger}_Y \bullet f = \bar{\dagger}_X \qquad \text{that is} \qquad T(!_Y) \circ f = T(!_X) \circ \eta_X.$$

Causal maps are used in [10] to construct effectuses. Here we define them quite generally, for an arbitrary monad. Notice that each map $f\colon X \to T(Y)$ is automatically causal when T is an affine monad. The following elementary observation gives a more precise description.

Lemma 14 *A Kleisli map $f\colon X \to T(Y)$ is causal if and only if it restricts to a (necessarily unique) map $f'\colon X \to T_a(Y)$ for the affine submonad $\iota\colon T_a \rightarrowtail T$, where $\iota_Y \circ f' = f$. Hence there is an isomorphism of categories:*

$$\mathit{Caus}\Big(\mathcal{K}\ell(T)\Big) \cong \mathcal{K}\ell(T_a),$$

where $\mathit{Caus}\Big(\mathcal{K}\ell(T)\Big) \hookrightarrow \mathcal{K}\ell(T)$ is the subcategory with causal maps only.

Proof. Obviously, the causality requirement $\bar{\dagger}_Y \bullet f = T(!) \circ f = \eta_1 \circ ! = \bar{\dagger}_X$ means that the outer diagram commutes in:

$$\begin{tikzcd}
X \arrow[drr, bend left, "!"] \arrow[dr, dotted, "f'"] \arrow[ddr, bend right, "f"'] & & \\
& T_a(Y) \arrow[r, "!"] \arrow[d, tail, "\iota_Y"'] \arrow[dr, phantom, "\lrcorner", very near start] & 1 \arrow[d, "\eta"] \\
& T(Y) \arrow[r, "T(!)"'] & T(1)
\end{tikzcd}$$

□

As a result, a Kleisli map $X \to \mathcal{D}(X)$ for the distribution monad $\mathcal{D}$ can equivalently be described as a causal map $X \to \mathcal{M}(X)$ for the multiset monad $\mathcal{M}$, see Example 12 (i). This gives a more systematic approach than the "constrained" description from [48], which restricts multisets to a certain subset.

There is more to say about affine parts of monads, especially in relation to additive monads and their affine parts being partially additive, as part of longer story about the relation between a linear and a probabilistic world. But instead we turn to a stronger version of 'affineness', which we call 'strong affiness', not only because it implies ordinary affiness but also because it involves strength.

Definition 15 *Let $\mathbf{C}$ be a category with finite products $(1, \times)$ and let $T\colon \mathbf{C} \to \mathbf{C}$ be a strong monad. This T will be called* strongly affine *if the squares below are pullbacks in $\mathbf{C}$.*

$$\begin{tikzcd}
T(X) \times Y \arrow[r, "\pi_2"] \arrow[d, "\mathrm{st}_1"'] & Y \arrow[d, "\eta_Y"] \\
T(X \times Y) \arrow[r, "T(\pi_2)"'] & T(Y)
\end{tikzcd} \tag{17}$$

Of course, the corresponding diagrams with st_2 instead of st_1 and π_1 instead of π_2 are then also pullbacks, via the swap isomorphism $\langle \pi_2, \pi_1 \rangle$.

The notion of an 'affine monad' is well-known. What we call 'strongly affine' is new. The relationship with ordinary affine monads is a bit subtle. Example 18 below show that 'strongly affine' is really stronger than 'affine'. But first we describe some properties and examples.

Lemma 16 *Let T be a strong monad on a category $\mathbf{C}$ with finite products.*

(i) The following three points are equivalent:
  (a) T is affine, that is, $T(1) \cong 1$;
  (b) the diagrams (17) commute;
  (c) $\langle T(\pi_1), T(\pi_2)\rangle \circ \mathrm{dst} = \mathrm{id}$ for both paths of dst in (7).
(ii) There is at most one mediating (pullback) map for the diagram (17).

By the implication (ib) ⇒ (ia), a strongly affine monad is affine. The equivalence (ia) ⇔ (ic) is from [51, Thm. 2.1]. Point (ii) is useful when we wish to prove that a particular monad is strongly affine: we only need to prove existence of a mediating map, since uniqueness holds in general, see Example 17.

Proof. For the implication (ia) ⇒ (ib), let T be affine. We stretch Diagram (17) as follows.

$$\begin{tikzcd}
T(X)\times Y \arrow[r, "T(!)\times Y"'] \arrow[d, "\mathrm{st}_1"'] \arrow[rr, bend left, "\pi_2"] & T(1)\times Y \arrow[r, "\cong", "\pi_2"'] \arrow[d, "\mathrm{st}_1"'] & Y \arrow[d, "\eta_Y"] \\
T(X\times Y) \arrow[r, "T(!\times \mathrm{id})"] \arrow[rr, bend right, "T(\pi_2)"'] & T(1\times Y) \arrow[r, "T(\pi_2)", "\cong"'] & T(Y)
\end{tikzcd}$$

The square on the left commutes by naturality of strength. For the one on the right we use that $T(1)$ is final, so that $\pi_2 \colon T(1)\times Y \to Y$ is an isomorphism, with inverse $\langle \eta_1 \circ !_Y, \mathrm{id}\rangle$. Hence:

$$\begin{aligned}
T(\pi_2)\circ \mathrm{st}_1 &= T(\pi_2)\circ \mathrm{st}_1 \circ \langle \eta_1 \circ !_Y, \mathrm{id}\rangle \circ \pi_2 \\
&= T(\pi_2)\circ \mathrm{st}_1 \circ (\eta_1 \times \mathrm{id}) \circ \langle !_Y, \mathrm{id}\rangle \circ \pi_2 \\
&\overset{(5)}{=} T(\pi_2)\circ \eta_{1\times Y} \circ \langle !_Y, \mathrm{id}\rangle \circ \pi_2 \\
&= \eta_Y \circ \pi_2 \circ \langle !_Y, \mathrm{id}\rangle \circ \pi_2 \\
&= \eta_Y \circ \pi_2.
\end{aligned}$$

For the implication (ib) ⇒ (ic) assume that diagrams (17) commute, that is:

$$T(\pi_2)\circ \mathrm{st}_1 = \eta \circ \pi_2 \qquad \text{and} \qquad T(\pi_1)\circ \mathrm{st}_2 = \eta \circ \pi_1. \tag{18}$$

This second equation follows from the first one by pre-composing it with the swap map $\gamma = \langle \pi_2, \pi_1\rangle$. We now prove $T(\pi_i) \circ \mathrm{dst} = \pi_i$ for the upper path in (7) with $\mathrm{dst} = \mu \circ T(\mathrm{st}_2) \circ \mathrm{st}_1$.

$$\begin{tikzcd}
T(X) \arrow[r, equal] & T(X) \arrow[r, "\eta"] \arrow[rr, equal, bend left] & T^2(X) \arrow[r, "\mu"] & T(X) \\
T(X) \times T(Y) \arrow[u, "\pi_1"] \arrow[d, "\pi_2"'] \arrow[r, "\mathrm{st}_1"] \arrow[ur, phantom, "(4)"] \arrow[dr, phantom, "(18)"] & T(X \times T(Y)) \arrow[u, "T(\pi_1)"'] \arrow[d, "T(\pi_2)"'] \arrow[r, "T(\mathrm{st}_2)"] \arrow[ur, phantom, "(18)"] \arrow[dr, phantom, "(4)"] & T^2(X \times Y) \arrow[u, "T^2(\pi_1)"'] \arrow[d, "T^2(\pi_2)"'] \arrow[r, "\mu"] & T(X \times Y) \arrow[u, "T(\pi_1)"'] \arrow[d, "T(\pi_2)"'] \\
T(Y) \arrow[r, "\eta"'] \arrow[rrr, equal, bend right] & T^2(Y) \arrow[r, equal] & T^2(Y) \arrow[r, "\mu"'] & T(Y)
\end{tikzcd}$$

In a similar way one proves these equations $T(\pi_i) \circ \mathrm{dst} = \pi_i$ for the lower path for dst in (7).

For the implication (ic) $\Rightarrow$ (ia) let $\langle T(\pi_1), T(\pi_2)\rangle \circ \mathrm{dst} = \mathrm{id}$. Then $\pi_1 = \pi_2 \colon T(1) \times T(1) \to T(1)$, since $\pi_1 = \pi_2 \colon 1 \times 1 \to 1$ and thus:

$$\pi_1 = T(\pi_1) \circ \mathrm{dst} = T(\pi_2) \circ \mathrm{dst} = \pi_2$$

in:

$$\begin{tikzcd}
T(1) \times T(1) \arrow[d, "\mathrm{dst}"'] \arrow[dr, bend left, "\pi_i"] & \\
T(1 \times 1) \arrow[r, "\cong", "T(\pi_i)"'] & T(1)
\end{tikzcd}$$

For each object X there is always a map $X \to T(1)$, namely $\bar{\dagger} = \eta \circ\, !$. If we have two map $f, g \colon X \to T(1)$, then we form the tuple $\langle f, g\rangle \colon X \to T(1) \times T(1)$ and get:

$$f = \pi_1 \circ \langle f, g\rangle = \pi_2 \circ \langle f, g\rangle = g.$$

Finally, for point (ii) in Lemma 16 we prove uniqueness of mediating maps. Assume we have two maps $f, g \colon Z \to T(X) \times Y$ with $\pi_2 \circ f = \pi_2 \circ g$ and $\mathrm{st}_1 \circ f = \mathrm{st}_1 \circ g$. We then obtain $\pi_1 \circ f = \pi_1 \circ g$ from:

$$\pi_1 \circ f \overset{(4)}{=} T(\pi_1) \circ \mathrm{st}_1 \circ f = T(\pi_1) \circ \mathrm{st}_1 \circ g \overset{(4)}{=} \pi_1 \circ g. \qquad \square$$

All the example monads in Subsection 3.1 – 3.6 are not just affine, but strongly affine. The proofs are not entirely trivial, and subtly different each time. So we include all these verifications.

Example 17 *(i) In order to see that the distribution monad $\mathcal{D}$ is strongly affine, let in Diagram (17) a joint distribution $\omega \in \mathcal{D}(X \times Y)$ be given with $\mathcal{D}(\pi_2)(\omega) = 1|z\rangle$ for some element $z \in Y$. Write $\omega = \sum_{x,y} \omega(x, y)|x, y\rangle$, so that $\mathcal{D}(\pi_2)(\omega)$ is the marginal distribution:*

$$\mathcal{D}(\pi_2)(\omega) = \textstyle\sum_y \Big(\sum_x \omega(x, y)\Big)\Big| y \Big\rangle.$$

If this is the trivial distribution $1|z\rangle$*, then* $\omega(x,y) = 0$ *for all* x *and* $y \neq z$*. We obtain a new distribution* $\rho = \mathcal{D}(\pi_1)(\omega) \in \mathcal{D}(X)$*, which takes the simple form* $\rho(x) = \omega(x,z)$*. The pair* $(\rho, z) \in \mathcal{D}(X) \times Y$ *is the unique element giving us the pullback* (17)*, since:*

$$\mathrm{st}_1(\rho, z) = \textstyle\sum_x \rho(x)|x,z\rangle = \sum_x \omega(x,z)|x,z\rangle = \sum_{x,y} \omega(x,y)|x,y\rangle = \omega.$$

(ii) Next let's consider the situation (17) *for* $T = \mathcal{G}$ *the Giry monad, with a joint probability measure* $\omega \in \mathcal{G}(X \times Y)$ *and an element* $z \in Y$ *which mapped to the same element in* $\mathcal{G}(Y)$*, via the outer maps in Diagram* (17)*. Thus, for each* $N \in \Sigma_Y$*,*

$$\begin{aligned} \eta(z)(N) &= \mathcal{G}(\pi_2)(\omega)(N) \\ &= \omega(\pi_2^{-1}(N)) \\ &= \omega(X \times N). \end{aligned} \tag{19}$$

for all $N \in \Sigma_Y$*. We prove 'non-entwinedness' of* ω*, that is,* ω *is the product of its marginals, see Definition 2* (vii)*. Abstractly this means* $\omega = \mathrm{dst}\big(\mathcal{G}(\pi_1)(\omega), \mathcal{G}(\pi_2)(\omega)\big)$*, and concretely:*

$$\omega(M \times N) = \omega(M \times Y) \cdot \omega(X \times N), \tag{20}$$

for all $M \in \Sigma_X$ *and* $N \in \Sigma_Y$*. We distinguish two cases.*

- *If* $z \notin N$*, then, by monotonicity of the probability measure* ω*,*

$$\omega(M \times N) \leq \omega(X \times N) \stackrel{(19)}{=} \eta(z)(N) = \mathbf{1}_N(z) = 0.$$

Hence $\omega(M \times N) = 0$*. But also:*

$$\omega(M \times Y) \cdot \omega(X \times N) \stackrel{(19)}{=} \omega(M \times Y) \cdot \eta(z)(N) = \omega(M \times Y) \cdot 0 = 0.$$

- *If* $z \in N$*, then* $z \notin \neg N$*, so that:*

$$\begin{array}{rcll} \omega(M \times N) & = & \omega(M \times N) + 0 & \\ & = & \omega(M \times N) + \omega(M \times \neg N) & \textit{as just shown} \\ & = & \omega\big((M \times N) \cup (M \times \neg N)\big) & \textit{by additivity} \\ & = & \omega(M \times Y) & \\ & = & \omega(M \times Y) \cdot \eta(z)(N) & \\ & \stackrel{(19)}{=} & \omega(M \times Y) \cdot \omega(X \times N). & \end{array}$$

We now take $\rho \in \mathcal{G}(X)$ *defined by* $\rho(M) = \mathcal{G}(\pi_1)(\omega)(M) = \omega(M \times Y)$*.*

The pair $(\rho, z) \in \mathcal{G}(X) \times Y$ *is then mediating in* (17)*:*

$$\begin{aligned}\mathrm{st}_1(\rho, z)(M \times N) \;=\; \rho(M) \cdot \eta(z)(N) \;&=\; \omega(M \times Y) \cdot \eta(z)(N) \\ &\stackrel{(19)}{=} \omega(M \times Y) \cdot \omega(X \times N) \\ &\stackrel{(20)}{=} \omega(M \times N).\end{aligned}$$

Hence the Giry monad $\mathcal{G}$ *is strongly affine.*

(iii) We turn to the expectation monad $\mathcal{E}(X) = \mathbf{EMod}([0,1]^X, [0,1])$ *on* **Sets**, *where* **EMod** *is the category of effect modules, see Section 2. Let* $\omega \in \mathcal{E}(X \times Y)$ *satisfy* $\mathcal{E}(\pi_2)(\omega) = \eta(z)$*, for some* $z \in Y$*. This means that for each predicate* $q \in [0,1]^Y$ *we have* $\omega(q \circ \pi_2) = q(z)$*.*

Our first aim is to prove the analogue of the non-entwinedness equation (20) *for* $\mathcal{E}$*, namely:*

$$\omega(\mathbf{1}_{U \times V}) = \omega(\mathbf{1}_{U \times Y}) \cdot \omega(\mathbf{1}_{X \times V}), \tag{21}$$

for arbitrary subsets $U \subseteq X$ *and* $V \subseteq Y$*.*

- *If* $z \notin V$*, then* $\omega(\mathbf{1}_{U \times V}) \leq \omega(\mathbf{1}_{X \times V}) = \omega(\mathbf{1}_V \circ \pi_2) = \mathbf{1}_V(z) = 0$*. Hence* (21) *holds since both sides are* 0*.*
- *If* $z \in V$*, then* $\omega(\mathbf{1}_{U \times V}) = \omega(\mathbf{1}_{U \times V}) + \omega(\mathbf{1}_{U \times \neg V}) = \omega(\mathbf{1}_{U \times Y}) = \omega(\mathbf{1}_{U \times Y}) \cdot \omega(\mathbf{1}_{X \times V})$*.*

By [38, Lemma 12] each fuzzy predicate can be written as limit of step functions. It suffices to prove the result for such step functions, since by [38, Lemma 10] the map of effect modules ω *is automatically continuous.*

Hence we concentrate on an arbitrary step function $p \in [0,1]^{X \times Y}$ *of the form* $p = \sum_{i,j} r_{i,j} \mathbf{1}_{U_i \times V_j}$*, where the* $U_i \subseteq X$ *and* $V_j \subseteq Y$ *form disjoint covers, and* $r_{i,j} \in [0,1]$*. We prove that* $\omega(p) = \mathrm{st}_1\Big(\mathcal{E}(\pi_1)(\omega), z\Big)(p)$*, so that we can take* $\mathcal{E}(\pi_1)(\omega) \in \mathcal{E}(X)$ *to obtain a pullback in* (17)*.*

Let j_0 *be the (unique) index with* $z \in V_{j_0}$*, so that* $p(x,z) = \sum_i r_{i,j_0} \mathbf{1}_{U_i}(x)$*. Then:*

$$\begin{aligned}\omega(p) \;=\; \omega\Big(\textstyle\sum_{i,j} r_{i,j} \mathbf{1}_{U_i \times V_j}\Big) \;&=\; \textstyle\sum_{i,j} r_{i,j} \omega\Big(\mathbf{1}_{U_i \times V_j}\Big) \\ &\stackrel{(21)}{=} \textstyle\sum_{i,j} r_{i,j} \omega\Big(\mathbf{1}_{U_i \times Y}\Big) \cdot \omega\Big(\mathbf{1}_{X \times V_j}\Big) \\ &=\; \textstyle\sum_{i,j} r_{i,j} \omega\Big(\mathbf{1}_{U_i \times Y}\Big) \cdot \mathbf{1}_{V_j}(z) \\ &=\; \textstyle\sum_i r_{i,j_0} \omega\Big(\mathbf{1}_{U_i \times Y}\Big) \\ &=\; \omega\Big(\textstyle\sum_i r_{i,j_0} \mathbf{1}_{U_i \times Y}\Big) \\ &=\; \omega\Big(\lambda(x,y).\, p(x,z)\Big) \\ &=\; \mathrm{st}_1\Big(\mathcal{E}(\pi_1)(\omega), z\Big)(p).\end{aligned}$$

(iv) We show that the probabilistic powerdomain $\mathcal{V}$ *on the category* **CDcpo** *of* continuous *dcpo's is strongly affine. We use the result, due to Lawson,*

that a valuation on the opens $\mathcal{O}(X)$ *of a continuous dcpo* X *can be extended in a unique way to a measure on the Borel sets* $\mathcal{B}(X)$, *see [44,4]. We recall that* $\mathcal{B}(X)$ *is the least* σ*-algebra that contains* $\mathcal{O}(X)$.

We show that Diagram (17) *is a pullback, for* $T = \mathcal{V}$. *The proof is similar to the one for the Giry monad in point* (ii), *but uses the unique extension to Borel sets. Let* $\omega \in \mathcal{V}(X \times Y)$ *satisfy* $\mathcal{V}(\pi_2)(\omega) = \eta(z)$, *for a given element* $z \in Y$. *This means* $\omega(X \times V) = \omega(\pi_2^{-1}(V)) = \mathcal{V}(\pi_2)(\omega)(V) = \eta(z)(V) = \mathbf{1}_V(z)$, *for each* $V \in \mathcal{O}(Y)$. *We write* $\widehat{\omega}\colon \mathcal{B}(X) \to [0,1]$ *for the unique extension of* $\omega\colon \mathcal{O}(X) \to [0,1]$. *Since* η *extends to a measure on* $\mathcal{B}(X)$, *and* $\widehat{\omega}(X \times -)$ *is also a measure that extends* $\omega(X \times -)$ *we get:*

$$\widehat{\omega}(X \times V) = \mathbf{1}_V(z), \qquad \textit{for each } V \in \mathcal{B}(X). \tag{22}$$

Our first aim is to show that $\widehat{\omega}$ *is non-entwined, that is, satisfies* $\widehat{\omega}(U \times V) = \widehat{\omega}(U \times Y) \cdot \widehat{\omega}(X \times V)$ *for all* $U, V \in \mathcal{B}(X)$. *We distinguish two cases, like in the previous verifications.*

- *If* $z \notin V$, *then by monotonicity:*

$$\widehat{\omega}(U \times V) \leq \widehat{\omega}(X \times V) \overset{(22)}{=} \mathbf{1}_V(z) = 0.$$

  Hence $\widehat{\omega}(U \times V) = 0 = \widehat{\omega}(U \times Y) \cdot \widehat{\omega}(X \times V)$.
- *If* $z \in V$, *then* $z \notin \neg V$. *We note that Borel sets (but not open sets) are closed under negation/complement. Hence with the extension* $\widehat{\omega}$ *to Borel sets we can reason as follows.*

$$\begin{aligned}
\widehat{\omega}(U \times V) &= \widehat{\omega}(U \times V) + 0 \\
&= \widehat{\omega}(U \times V) + \widehat{\omega}(U \times \neg V) && \textit{as just shown} \\
&= \widehat{\omega}\big((U \times V) \cup (U \times \neg V)\big) && \textit{by additivity} \\
&= \widehat{\omega}(U \times Y) \\
&= \widehat{\omega}(U \times Y) \cdot \mathbf{1}_V(z) \\
&\overset{(22)}{=} \widehat{\omega}(U \times Y) \cdot \widehat{\omega}(X \times V).
\end{aligned}$$

But now we are done since we can take $\rho = \mathcal{V}(\pi_1)(\omega) = \omega(- \times Y) \in \mathcal{V}(X)$, *satisfying:*

$$\mathrm{st}_1(\rho, z)(U \times V) = \rho(U) \cdot \mathbf{1}_V(z) \overset{(22)}{=} \omega(U \times Y) \cdot \omega(X \times V) = \omega(U \times V).$$

(v) *The proof that the Radon monad* $\mathcal{R}$ *on compact Hausdorff spaces is strongly affine that is presented below is due to Robert Furber; it uses the Cauchy-Schwartz inequality for positive maps on* C^**-algebras. We first note that the strength map* $\mathrm{st}_1\colon \mathcal{R}(X) \times Y \to \mathcal{R}(X \times Y)$ *is determined by* $\mathrm{st}_1(\omega, z)(\phi \otimes \psi) = \omega(\phi) \cdot \psi(z)$. *These tensors* $\phi \otimes \psi = \lambda(x,y).\, \phi(x) \cdot \psi(y) \in \mathrm{C}(X \times Y) \cong \mathrm{C}(X) \otimes \mathrm{C}(Y)$ *form a dense subset. Hence the above description of* st_1 *suffices.*

We turn to Diagram (17)*. Let $\omega \in \mathcal{R}(X \times Y)$ and $z \in Y$ be given with $\mathcal{R}(\pi_2)(\omega) = \eta(z)$. This means that $\omega(\mathbf{1} \otimes \psi) = \psi(z)$, for each $\psi \in \mathrm{C}(Y)$, where $\mathbf{1} \in \mathrm{C}(X)$ is the function that is constantly* 1*. The Cauchy-Schwartz inequality for the positive map ω yields:*

$$\begin{aligned}
\left|\omega(\phi \otimes \psi)\right|^2 &= \omega\Big((\phi \otimes \mathbf{1}) \cdot (\mathbf{1} \otimes \psi)\Big)^* \cdot \omega\Big((\phi \otimes \mathbf{1}) \cdot (\mathbf{1} \otimes \psi)\Big) \\
&\leq \omega\Big((\phi \otimes \mathbf{1}) \cdot (\phi \otimes \mathbf{1})^*\Big) \cdot \omega\Big((\mathbf{1} \otimes \psi)^* \cdot (\mathbf{1} \otimes \psi)\Big) \\
&= \omega\Big((\phi \cdot \phi^*) \otimes \mathbf{1}\Big) \cdot \omega\Big(\mathbf{1} \otimes (\psi^* \cdot \psi)\Big) \\
&= \omega\Big((\phi \cdot \phi^*) \otimes \mathbf{1}\Big) \cdot (\psi^* \cdot \psi)(z) \\
&= \omega\Big((\phi \cdot \phi^*) \otimes \mathbf{1}\Big) \cdot \psi(z)^* \cdot \psi(z).
\end{aligned}$$

Hence if $\psi(z) = 0$, then $\omega(\phi \otimes \psi) = 0$. Consider the function $\psi' \in \mathrm{C}(Y)$ given by $\psi'(y) = \psi(z) - \psi(y)$. Since $\psi'(z) = 0$, we get $\omega(\phi \otimes \psi') = 0$, as just shown, and thus by linearity of ω:

$$\begin{aligned}
\omega(\phi \otimes \psi) \;=\; \omega(\phi \otimes \psi) + \omega(\phi \otimes \psi') &= \omega(\phi \otimes (\psi - \psi')) \\
&= \omega(\phi \otimes \psi(z)) \\
&= \omega(\phi \otimes \mathbf{1}) \cdot \psi(z) \\
&= \omega(\phi \otimes \mathbf{1}) \cdot \omega(\mathbf{1} \otimes \psi).
\end{aligned}$$

We can now take as state $\rho = \mathcal{R}(\pi_1)(\omega) \in \mathcal{R}(X)$ given by $\rho(\phi) = \omega(\phi \otimes \mathbf{1})$. This gives the mediating element that we seek, since:

$$\mathrm{st}_1(\rho, z)(\phi \otimes \psi) = \rho(\phi) \cdot \psi(z) = \omega(\phi \otimes \mathbf{1}) \cdot \omega(\mathbf{1} \otimes \psi) = \omega(\phi \otimes \psi).$$

The following (counter) example is due to Kenta Cho.

Example 18 *An example of an affine but not strongly affine monad is the 'generalised distribution' monad $\mathcal{D}_\pm$ on* **Sets***. Elements of $\mathcal{D}_\pm(X)$ are finite formal sums $\sum_i r_i|x_i\rangle$ with $r_i \in \mathbb{R}$ and $x_i \in X$ satisfying $\sum_i r_i = 1$. The other data of a (strong) monad are similar to the ordinary distribution monad $\mathcal{D}$. Clearly $\mathcal{D}_\pm(1) \cong 1$, i.e. $\mathcal{D}_\pm$ is affine.*

Now consider the square (17) *with $X = \{x_1, x_2\}$ and $Y = \{y_1, y_2\}$. Define:*

$$\omega = 1|x_1, y_1\rangle + 1|x_1, y_2\rangle + (-1)|x_2, y_2\rangle \in \mathcal{D}_\pm(X \times Y).$$

We have $\mathcal{D}_\pm(\pi_2)(\omega) = 1|y_1\rangle = \eta(y_1)$, since the terms with y_2 cancel each other out. But there is no element $\psi \in \mathcal{D}_\pm(X)$ such that $\mathrm{st}_1(\psi, y_1) = \omega$. Hence the square (17) *is not a pullback.*

Since $\mathcal{D}_\pm$ is the affine part of a monad, namely of $\mathcal{M}_\mathbb{R}$, see Example 12 (i)*, we see that the affine part of a monad need not be strongly affine.*

The fact that the terms in this example cancel each other out is known as 'interference' in the quantum world. It already happens with negative coefficients. This same monad $\mathcal{D}_{\pm}$ is used in [1]. How the notions of non-locality and contextuality that are studied there relate to strong affineness requires further investigation.

The following result gives a 'graph' construction that is useful in conditional constructions in probability, see the subsequent discussion. It will play a crucial role for side-effect-freeness in Proposition 27.

Proposition 19 *For a strongly affine monad T there is a canonical bijective correspondence:*

$$\frac{Y \xrightarrow{f} T(X)}{Y \xrightarrow[g]{} T(X \times Y) \text{ with } T(\pi_2) \circ g = \eta}$$

What we mean by 'canonical' is that the mapping downwards is given by $f \mapsto \mathrm{st}_1 \circ \langle f, \mathrm{id} \rangle$.

Proof. As stated, the mapping downwards is given by $\overline{f} = \mathrm{st}_1 \circ \langle f, \mathrm{id}\rangle$. This map $\overline{f}$ satisfies the side-condition below the double lines:

$$T(\pi_2) \circ \overline{f} = T(\pi_2) \circ \mathrm{st}_1 \circ \langle f, \mathrm{id}\rangle \stackrel{(17)}{=} \eta \circ \pi_2 \circ \langle f, \mathrm{id}\rangle = \eta.$$

In the other direction we map $g \colon Y \to T(X \times Y)$ to $\overline{g} = T(\pi_1) \circ g$. Then:

$$\overline{\overline{f}} = T(\pi_1) \circ \mathrm{st}_1 \circ \langle f, \mathrm{id}\rangle \stackrel{(4)}{=} \pi_1 \circ \langle f, \mathrm{id}\rangle = f.$$

In order to prove $\overline{\overline{g}} = g$ we notice that by the pullback property of diagram (17) we know that there is a unique $h \colon Y \to T(X)$ with $g = \mathrm{st}_1 \circ \langle h, \mathrm{id}\rangle = \overline{h}$. But then $\overline{\overline{h}} = h$, by what we have just shown, so that:

$$\overline{\overline{g}} = \overline{\overline{\overline{h}}} = \overline{h} = g. \qquad \square$$

The correspondence in this proposition is used (for the distribution monad $\mathcal{D}$) as Lemma 1 in [23]. There, the map $\mathrm{st}_1 \circ \langle f, \mathrm{id}\rangle$ is written as $\mathrm{gr}(f)$, and called the graph of f. It is used in the description of conditional probability. It is also used (implicitly) in [19, §3.1], where a measure/state $\omega \in \mathcal{G}(X)$ and a Kleisli map $f \colon X \to \mathcal{G}(Y)$ give rise to a joint probability measure $\mathrm{gr}(f) \bullet \omega$ in $\mathcal{G}(X \times Y)$.

6 Strongly affine monads and effectuses

We will use the property of strong affiness to obtain stronger results about the structure $(\mathbf{0}, \ovee)$ on homsets from Lemma 8. For this the following additional property is useful. It transfers the disjointness of coprojections in Definition 1 to a setting with a monad.

Definition 20 *A monad T on a distributive category $\mathbf{C}$ with disjoint coprojections is called* not-trivialising *if the following rectangles are pullbacks in $\mathbf{C}$.*

$$\begin{tikzcd} 0 \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & T(X_2) \arrow[d, tail, "T(\kappa_2)"] \\ T(X_1) \arrow[r, tail, "T(\kappa_1)"'] & T(X_1 + X_2) \end{tikzcd} \quad \textit{or, equivalently, just} \quad \begin{tikzcd} 0 \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & T(1) \arrow[d, tail, "T(\kappa_2)"] \\ T(1) \arrow[r, tail, "T(\kappa_1)"'] & T(1 + 1) \end{tikzcd} \tag{23}$$

The fact that commutation of the rectangle on the right suffices follows from the following diagram (and the fact that the initial object 0 is strict).

$$\begin{tikzcd} Y \arrow[dr, dotted] \arrow[drr, bend left] \arrow[ddd, bend right] & & & \\ & 0 \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner", very near start] & T(1) \arrow[d, "T(\kappa_2)"] & T(X_2) \arrow[l, "T(!)"'] \arrow[dd, "T(\kappa_2)"] \\ & T(1) \arrow[r, tail, "T(\kappa_1)"'] & T(1+1) & \\ & T(X_1) \arrow[u, "T(!)"] \arrow[rr, "T(\kappa_1)"'] & & T(X_1 + X_2) \arrow[ul, "T(!+!)"'] \end{tikzcd}$$

The second rectangle in (23) is easy to check, certainly when the monad is affine, so that $T(1) \cong 1$. It then says that the intersection of the zero and one scalars $\mathbf{0}, \mathbf{1} \colon 1 \to T(2)$ is empty. This is obviously the case in all our examples in Section 3, where the set of scalars is each time the unit interval $[0, 1]$, in which the zero and one scalars are obviously disjoint.

The following technical result is important for a better understanding of the structure of partial maps and predicates, see Lemma 22 below.

Lemma 21 *Let T be a strongly affine monad on a non-trivial distributive category $\mathbf{C}$. The following diagrams are then pullbacks in the Kleisli category $\mathcal{K}\ell(T)$.*

$$\begin{tikzcd} X \arrow[r, "!"] \arrow[d, "\langle\kappa_i\rangle"'] & 1 \arrow[d, "\langle\kappa_i\rangle"] \\ X + X \arrow[r, "!+!"'] & 1 + 1 \end{tikzcd} \qquad \begin{tikzcd} 1 \arrow[r, equal] \arrow[d, "\langle\kappa_2\rangle"'] & 1 \arrow[d, "\langle\kappa_2\rangle"] \\ X + 1 \arrow[r, "!+\mathrm{id}"'] & 1 + 1 \end{tikzcd} \qquad \begin{tikzcd} X \arrow[r, "!"] \arrow[d, "\langle\kappa_1\rangle"'] & 1 \arrow[d, "\langle\kappa_1\rangle"] \\ X + 1 \arrow[r, "!+\mathrm{id}"'] & 1 + 1 \end{tikzcd} \tag{24}$$

For this last (third) pullback we need to assume that the monad T is not-trivialising. We can then prove that maps $T(\kappa_i)$ are monic in $\mathbf{C}$ — making coprojections $\langle\kappa_i\rangle$ monic in $\mathcal{K}\ell(T)$.

Proof. The proof that the diagram on the left in (24) is a pullback is obtained by taking $Y = 2 = 1 + 1$ in Diagram (17) and using the distributivity isomorphism $\mathrm{sep}_2\colon X \times 2 \to X + X$ from (3). We leave it to the meticulous reader to check that the following two diagrams commute.

$$
\begin{tikzcd}
X \times 1 \arrow[r, "\mathrm{id}\times\kappa_1"] \arrow[d, "\pi_1"'] & X \times 2 \arrow[d, "\mathrm{sep}_2"'] \arrow[dr, "\pi_2"] & \\
X \arrow[r, "\kappa_1"'] & X + X \arrow[r, "!+!"'] & 2
\end{tikzcd}
\qquad
\begin{tikzcd}
T(X) \times 2 \arrow[r, "\mathrm{st}_1"] \arrow[d, "\mathrm{sep}_2"'] & T(X \times 2) \arrow[d, "T(\mathrm{sep}_2)"] \\
T(X) + T(X) \arrow[r, "{[T(\kappa_1), T(\kappa_2)]}"'] & T(X + X)
\end{tikzcd}
\quad (*)
$$

We now show that the left diagram in (24) is a pullback in $\mathcal{K}\ell(T)$, for $i = 1$. Let $f\colon Y \to T(X + X)$ satisfy $(! + !) \bullet f = \langle\kappa_1\rangle \bullet !$, that is, $T(! + !) \circ f = T(\kappa_1) \circ \eta \circ !$. Take $f' = T(\mathrm{sep}_2^{-1}) \circ f\colon Y \to T(X \times 2)$, and consider the pullback (17). We get:

$$T(\pi_2) \circ f' = T(\pi_2) \circ T(\mathrm{sep}_2^{-1}) \circ f \overset{(*)}{=} T(! + !) \circ f = \eta \circ \kappa_1 \circ !.$$

Hence there is a unique map $g\colon Y \to T(X)$ in (17) with $\mathrm{st}_1 \circ \langle g, \kappa_1 \circ !\rangle = f'$. This g is the mediating map that we want, since:

$$
\begin{aligned}
f \;=\; T(\mathrm{sep}_2) \circ f' \;&=\; T(\mathrm{sep}_2) \circ \mathrm{st}_1 \circ \langle g, \kappa_1 \circ !\rangle \\
&\overset{(*)}{=} [T(\kappa_1), T(\kappa_2)] \circ \mathrm{sep}_2 \circ (\mathrm{id} \times \kappa_1) \circ \langle g, !\rangle \\
&\overset{(*)}{=} [T(\kappa_1), T(\kappa_2)] \circ \kappa_1 \circ \pi_1 \circ \langle g, !\rangle \\
&= T(\kappa_1) \circ g \\
&= \langle\kappa_1\rangle \bullet g.
\end{aligned}
$$

Uniqueness is left to the reader.

We continue with the diagram in the middle in (24). The case $X \cong 0$ trivially holds. If $X \not\cong 0$, then we may assume a map $x\colon 1 \to X$, since the underlying category is non-trivial, see Definition 1. Now let $f\colon Y \to T(X + 1)$ satisfy $T(! + \mathrm{id}) \circ f = \langle\kappa_2\rangle \circ !$. Then $f' = T(\mathrm{id} + x) \circ f\colon Y \to T(X + X)$ satisfies $T(! + !) \circ f' = T(! + \mathrm{id}) \circ f = \langle\kappa_2\rangle \circ !$. Using the pullback on the left in (24) we get a $g\colon Y \to T(X)$ with $T(\kappa_2) \circ g = f'$. But then:

$$
\begin{aligned}
f \;=\; T(\mathrm{id} + !) \circ f' \;=\; T(\mathrm{id} + !) \circ T(\kappa_2) \circ g \;&=\; T(\kappa_2) \circ T(!) \circ g \\
&\overset{(**)}{=} T(\kappa_2) \circ \eta \circ ! \;=\; \langle\kappa_2\rangle \bullet !.
\end{aligned}
$$

The equation $\overset{(**)}{=}$ holds because $T(1)$ is final. This finality also yields uniqueness of the mediating map $!$.

For the third rectangle in (24) the case $X \cong 0$ is covered by the requirement that T is not-trivialising: if $f\colon Y \to T(0+1)$ satisfies $T(!+\mathrm{id}) \circ f = T(\kappa_1) \circ \eta \circ !$, then $f = T(\kappa_2) \circ \eta \circ !$, since $T(0+1) \cong T(1) \cong 1$. We thus have $T(\kappa_1) \circ \eta \circ ! = T(\kappa_2) \circ \eta \circ !$, so that $Y \to T(1)$ factors through 0, via the pullback (23). This implies $Y \cong 0$, since the initial object in a distributive category is strict [11]. But then we are done.

When $X \ncong 0$ we can use a map $x\colon 1 \to X$ and proceed like for the middle rectangle. Finally, we show that the maps $T(\kappa_1)\colon T(X) \to T(X+Y)$ are monic in $\mathbf{C}$. If $f, g\colon Y \to T(X)$ satisfy $T(\kappa_1) \circ f = T(\kappa_1) \circ g$, then $f = g$ by uniqueness of the mediating map in the pullback on the right in (24). Obviously, $! \bullet f = ! \bullet g$, but also:

$$\begin{aligned} \langle \kappa_1 \rangle \bullet f \;=\; T(\kappa_1) \circ f = T(\mathrm{id}+!) \circ T(\kappa_1) \circ f \\ = T(\mathrm{id}+!) \circ T(\kappa_1) \circ g \;=\; T(\kappa_1) \circ f \;=\; \langle \kappa_1 \rangle \bullet g. \quad \square \end{aligned}$$

These pullbacks play an important role in the following results extending Lemma 8.

Lemma 22 *Let T be a partially additive, strongly affine, not-trivialising monad on a non-trivial distributive category $\mathbf{C}$. Then:*

(i) $\mathbf{1} \bullet' f = \mathbf{0}$ *iff* $f = \mathbf{0}$*;*
(ii) $\mathbf{1} \bullet' f = \mathbf{1}$ *iff f is total iff* $\ker(f) = \mathbf{0}$*;*
(iii) for each object $X \in \mathbf{C}$, the set of predicates:

$$\mathrm{Pred}(X) = \mathbf{C}(X, T(2)) = \mathcal{K}\ell(T)(X, 2) = \mathcal{K}\ell(T')(X, 1).$$

is an effect module;
(iv) for each (total) map $f\colon X \to Y$ in $\mathcal{K}\ell(T)$ the associated predicate transformer $f^* = (-) \bullet f\colon \mathrm{Pred}(Y) \to \mathrm{Pred}(X)$ is a map of effect modules. As a result, there is a predicate functor $\mathrm{Pred}\colon \mathcal{K}\ell(T) \to \mathbf{EMod}^{\mathrm{op}}$.
(v) the partial sum $\ovee$ on partial maps is positive: $f \ovee g = \mathbf{0} \Rightarrow f = g = \mathbf{0}$, and $f \ovee g = f \Rightarrow g = \mathbf{0}$;
(vi) these partial maps carry a partial order defined by $f \leq g$ iff $f \ovee h = g$ for some h.

Proof. (i) Obviously, if $f = \mathbf{0}$, then $\mathbf{1} \bullet' f = \mathbf{1} \bullet' \mathbf{0} = \mathbf{0}$. In the other direction, the assumption $\mathbf{1} \bullet' f = \mathbf{0}$ means $T(!+\mathrm{id}) \circ f = \eta \circ \kappa_2 \circ !$. Using the pullback in the middle of (24) we obtain $f = \langle \kappa_2 \rangle \bullet ! = \mathbf{0}$.

(ii) Since $\ker(f) = (\mathbf{1} \bullet' f)^{\perp}$, see Definition 4, one obviously has $\mathbf{1} \bullet' f = \mathbf{1}$ iff $\ker(f) = \mathbf{0}$.

If $f\colon X \to Y$ in $\mathcal{K}\ell(T')$ is total, say $f = \langle \kappa_1 \rangle \bullet g = T(\kappa_1) \circ g$ for a necessarily unique $g\colon X \to Y$ in $\mathcal{K}\ell(T)$, then:

$$\begin{aligned}
\mathbf{1} \bullet' f &= T(! + \mathrm{id}) \circ T(\kappa_1) \circ g \\
&= T(\kappa_1) \circ T(!) \circ g \\
&= T(\kappa_1) \circ \eta_1 \circ ! && \text{since } T(1) \text{ is final} \\
&= \mathbf{1}.
\end{aligned}$$

In the other direction, let $\mathbf{1} \bullet' f = \mathbf{1}$. This means $T(!+\mathrm{id}) \circ f = \langle\kappa_1\rangle \bullet !$. The pullback in $\mathcal{K}\ell(T)$ on the right in (24) then yields a necessarily unique map g with $f = \langle\kappa_1\rangle \bullet g$. This makes f total.

(iii) From Lemma 8 we already know that predicates on an object X form a partial commutative monoid with scalar multiplication and an orthosupplement. The only to points that remain are (iv) and (v) from Section 2

For point (iv), let $p \owedge q = \mathbf{1}$, say via bound $b\colon X \to T'(1 + 1)$; we need to prove $q = p^{\perp}$. We have $p \owedge q = \mathbf{1} = \nabla \bullet' b = T(\nabla + \mathrm{id}) \circ b = T(!+\mathrm{id}) \circ b = \mathbf{1} \bullet' b$. By the previous point we obtain that b is total, say $b = \langle\kappa_1\rangle \bullet c = T(\kappa_1) \circ c$. But then we are done:

$$\begin{aligned}
p^{\perp} \;=\; T([\kappa_2, \kappa_1]) \circ p &= T([\kappa_2, \kappa_1]) \circ (\rhd_1 \bullet' b) \\
&= T([\kappa_2, \kappa_1]) \circ T([\mathrm{id}, \kappa_2]) \circ T(\kappa_1) \circ c \\
&= T([[\kappa_2, \kappa_1], \kappa_2]) \circ T(\kappa_1) \circ c \\
&= \rhd_2 \bullet' b \\
&= q.
\end{aligned}$$

Finally, for point (v) let $\mathbf{1} \perp p$, say via $b\colon X \to T'(1 + 1)$; we have to prove $p = \mathbf{0}$. The assumption translates to: $T([\mathrm{id}, \kappa_2]) \circ b = \rhd_1 \bullet' b = \mathbf{1} = \langle\kappa_1\rangle \circ !$, as in the above diagram on the right. Consider the isomorphism $\sigma = \text{\large ⤫} = [[\kappa_2, \kappa_1 \circ \kappa_1], \kappa_1 \circ \kappa_2]\colon 2 + 1 \xrightarrow{\cong} 2 + 1$, so that the outer diagram below on the right commutes:

$$\begin{aligned}
&T(! + \mathrm{id}) \circ T(\sigma) \circ b \\
&\quad = T([[\kappa_2, \kappa_1 \circ ! \circ \kappa_1], \kappa_1 \circ ! \circ \kappa_2]) \circ b \\
&\quad = T([[\kappa_2, \kappa_1], \kappa_1]) \circ b \\
&\quad = T([\kappa_2, \kappa_1]) \circ T([\mathrm{id}, \kappa_2]) \circ b \\
&\quad = T([\kappa_2, \kappa_1]) \circ \langle\kappa_1\rangle \circ ! \\
&\quad = \langle\kappa_2\rangle \circ !.
\end{aligned}$$

$$\begin{tikzcd}
X \arrow[drr, bend left, "!"] \arrow[dr, dotted, "!"] \arrow[ddr, bend right, "T(\sigma)\circ b"'] & & \\
& 1 \arrow[r, equal] \arrow[d, "\langle\kappa_2\rangle"'] \arrow[dr, phantom, "\lrcorner", very near start] & 1 \arrow[d, "\langle\kappa_2\rangle"] \\
& 2+1 \arrow[r, "!+\mathrm{id}"'] & 1+1
\end{tikzcd}$$

Hence $T(\sigma) \circ b = \langle \kappa_2 \rangle \circ !$ by the middle pullback in (24). But then:

$$\begin{aligned} p \;=\; \rhd_2 \bullet' b &= T([[\kappa_2, \kappa_1], \kappa_2]) \circ T(\sigma^{-1}) \circ \langle \kappa_2 \rangle \circ ! \\ &= T([[\kappa_2, \kappa_1], \kappa_2]) \circ T([\kappa_2 + \mathrm{id}, \kappa_1 \circ \kappa_1]) \circ T(\kappa_2) \circ \eta \circ ! \\ &= T(\kappa_2) \circ \eta \circ ! \\ &= \mathbf{0}. \end{aligned}$$

(iv) Let $f \colon X \to T(Y)$ be a (total) Kleisli map, giving $f' = \langle \kappa_1 \rangle \bullet f = T(\kappa_1) \circ f \colon X \to T(Y+1)$ as associated partial map in $\mathcal{K}\ell(T')$. The predicate transformer $f^* = (-) \bullet f = (-) \bullet' f'$ preserves the effect module structure by Lemma 8 (ii),(iii):

$$\begin{aligned} f^*(\mathbf{0}) &= \mathbf{0} \bullet f \;=\; \mathbf{0} \bullet' f' \;=\; \mathbf{0} \\ f^*(p \ovee q) &= (p \ovee q) \bullet' f' \;=\; (p \bullet' f') \ovee (q \bullet' f') \;=\; f^*(p) \ovee f^*(q) \\ f^*(s \cdot p) &= (s \cdot p) \bullet' f' \;=\; s \cdot (p \bullet' f') \;=\; s \cdot f^*(p) \\ f^*(\mathbf{1}) &= \mathbf{1} \bullet f \;=\; \mathbf{1} \qquad \text{since } f \text{ is total, see point (ii) above.} \end{aligned}$$

(v) We use the standard result that effect algebras are positive. Hence if $f \ovee g = \mathbf{0}$, then

$$(\mathbf{1} \bullet' f) \ovee (\mathbf{1} \bullet' g) = \mathbf{1} \bullet' (f \ovee g) = \mathbf{1} \bullet' \mathbf{0} = \mathbf{0}.$$

Therefore $\mathbf{1} \circ' f = \mathbf{1} \circ' g = \mathbf{0}$, so that $f = g = \mathbf{0}$ by point (i).
Similarly, if $f \ovee g = f$, then $(\mathbf{1} \bullet' f) \ovee (\mathbf{1} \bullet' g) = \mathbf{1} \bullet' f$, so that $\mathbf{1} \bullet' g = \mathbf{0}$, and thus $g = \mathbf{0}$.

(vi) Obviously, the order $\leq$, given by $f \leq g$ iff $f \ovee h = g$ for some h, is reflexive and transitive. It is also anti-symmetric: if $f \leq g$ via $f \ovee h = g$ and $g \leq f$ via $g \ovee k = f$, then $f \ovee (h \ovee k) = f$, so that $h \ovee k = \mathbf{0}$ and thus $h = k = \mathbf{0}$ by the previous point. But then $f = g$. $\square$

The main result of this section gives conditions that ensure that a Kleisli category is an effectus, see [30,10]. Briefly, an effectus is a category with finite coproducts and a final object in which the two maps $\text{IV}, \text{XV} \colon (1+1)+1 \rightrightarrows 1+1$ in (13) are jointly monic, and in which the following diagrams are pullbacks.

$$\begin{tikzcd} X + Y \arrow[r, "\mathrm{id}+!"] \arrow[d, "!+\mathrm{id}"'] & X + 1 \arrow[d, "!+\mathrm{id}"] \\ 1 + Y \arrow[r, "\mathrm{id}+!"'] & 1 + 1 \end{tikzcd} \qquad \begin{tikzcd} X \arrow[r, "!"] \arrow[d, "\kappa_1"'] & 1 \arrow[d, "\kappa_1"] \\ X + Y \arrow[r, "!+!"'] & 1 + 1 \end{tikzcd} \tag{25}$$

Theorem 23 *Let T be a partially additive, strongly affine, not-trivialising monad on a non-trivial distributive category* $\mathbf{C}$. *Its Kleisli category $\mathcal{K}\ell(T)$ is then an effectus.*

One possible way to prove that $\mathcal{K}\ell(T)$ is an effectus works by showing that the category of partial maps $\mathcal{K}\ell(T')$ is an 'effectus in partial form'. This is an application of [9, Thm. 4.10], which re-appears as [10, Thm. 53 (2)], using (our) Lemmas 8 and Lemma 22. The category of total maps in $\mathcal{K}\ell(T')$ is then $\mathcal{K}\ell(T)$, by Lemma 22 (ii). Instead we provide a direct proof.

Proof. Since partial additivity of the monad T implies that we have jointly monic maps $(1+1)+1 \rightrightarrows 1+1$ in (13), one only has to show that the commuting diagrams in (25) are pullbacks in $\mathcal{K}\ell(T)$.

- Let $f\colon Z \to T(X+1)$ and $g\colon Z \to T(1+Y)$ satisfy $(!+\mathrm{id}) \bullet f = (\mathrm{id}+!) \bullet g$. We transform these f, g into parallel maps $f', g'\colon Y \to T((X+Y)+1) = T'(X+Y)$ via:

  $$f' = T(\kappa_1 + \mathrm{id}) \circ f \qquad g' = T([\kappa_2, \kappa_1 \circ \kappa_2]) \circ g.$$

  Then:

  $$\begin{aligned} \mathbf{1} \bullet' f' &= T(!+\mathrm{id}) \circ T(\kappa_1 + \mathrm{id}) \circ f \\ &= T(!+\mathrm{id}) \circ f \\ &= T(\mathrm{id}+!) \circ g \\ &= T([\kappa_2, \kappa_1]) \circ T(!+\mathrm{id}) \circ T([\kappa_2, \kappa_1]) \circ g \\ &= T([\kappa_2, \kappa_1]) \circ T(!+\mathrm{id}) \circ T([\kappa_2, \kappa_1 \circ \kappa_2]) \circ g \\ &= (\mathbf{1} \bullet' g')^{\perp}. \end{aligned}$$

  Therefore:
  $$\mathbf{1} = (\mathbf{1} \bullet' f') \ovee (\mathbf{1} \bullet' g') = \mathbf{1} \bullet' (f' \ovee g').$$

  This means that the sum $f' \ovee g'$ exists and is total. Hence, by Lemma 22 (ii) there is a unique map $k\colon Z \to T(X+Y)$ with $T(\kappa_1) \circ k = f' \ovee g'$. This k is the mediating map that we seek.
- In order to show that the diagram on the right in (25) is a pullback in $\mathcal{K}\ell(T)$, let $f\colon Z \to T(X+Y)$ satisfy $(!+!) \bullet f = \langle\kappa_1\rangle \bullet\, !$. This means that $T(!+!) \circ f = \mathbf{1}$. Take $f' = T(\mathrm{id}+!) \circ f\colon Z \to T(X+1)$. It satisfies:

  $$\mathbf{1} \bullet' f' = T(!+\mathrm{id}) \circ T(\mathrm{id}+!) \circ f = T(!+!) \circ f = \mathbf{1}.$$

  Hence f' is total by Lemma 22 (ii), and thus of the form $f' = \langle\kappa_1\rangle \bullet g$, for a unique map $g\colon Z \to T(X)$. This g is the required mediating map. □

We may conclude that the Kleisli categories of all the monad examples $\mathcal{D}$, $\mathcal{G}$, $\mathcal{E}$, $\mathcal{V}$, $\mathcal{R}$, $\mathcal{K}$ in Section 3 are effectuses. This is a first important step. Further structure will be uncovered in later sections.

Lemma 22 shows that taking predicates forms a functor $\mathrm{Pred}\colon \mathcal{K}\ell(T) \to \mathbf{EMod}^{\mathrm{op}}$. This can be shown in general, for every effectus, see [30,10] for details. In fact, one can also show in general that states in an effectus form convex sets. A summary of these results is given in the 'state-and-effect' triangle below, specialised to a Kleisli category effectus.

$$
\begin{tikzcd}
\mathbf{EMod}^{\mathrm{op}} \arrow[rr, bend left, "\mathrm{Hom}(-,[0,1])"] \arrow[rr, phantom, "\top"] & & \mathbf{Conv} \arrow[ll, bend left, "\mathrm{Hom}(-,[0,1])"] \\
& \mathcal{K}\ell(T) \arrow[ul, "\mathrm{Hom}(-,1+1)=\mathrm{Pred}"] \arrow[ur, "\mathrm{Stat}=\mathrm{Hom}(1,-)"'] &
\end{tikzcd}
$$

The category **Conv** is the category of convex sets, in which formal convex sums have an actual sum. This is the category of Eilenberg-Moore algebras of the distribution monad $\mathcal{D}$, see [26,28] for details. Such state-and-effect triangles are studied systematically in [32,31] as formalisation of the fundamental (dual adjoint) relationship between states and predicates/effects, and between state transformers and predicate transformers in programming semantics and logic [15].

7 Predicates, tests, and instruments

So far we have used, for a monad T, maps $X \to T(2)$ as predicates on X, where $2 = 1 + 1$. There is a more general notion of *test*, or *n-test* to be more precise, as a map $X \to T(n)$. For the trivial identity monad on **Sets**, an n-test $t\colon X \to n$ gives a partition of the set X, consisting of n disjoint subsets $t^{-1}(i) \subseteq X$ which together cover X. These subsets correspond to the different outcomes of the test.

The idea extends to monads. For instance, a test $t\colon X \to \mathcal{D}(n)$ for the distribution monad can be identified with n predicates $t_i \in [0,1]^X$, given by $t_i(x) = t(x)(i)$, with $t_0 \ovee \cdots \ovee t_{n-1} = \mathbf{1}$. This works more generally.

We shall associate a certain 'instrument' map with a test — and thus in particular with predicates. These instruments can be used for a 'case' construct, as will be explained below. But they can also be used for conditional probabilities, see Section 9.

Definition 24 *Let T be a strong monad on a distributive category* $\mathbf{C}$. *With each n-test $t\colon X \to n$ in $\mathcal{K}\ell(T)$ we associate an instrument map* $\mathrm{instr}_t\colon X \to X + \cdots + X$ *in $\mathcal{K}\ell(T)$ in the following manner.*

$$\mathrm{instr}_t \stackrel{\text{def}}{=} \Big(X \xrightarrow{\langle t, \mathrm{id}\rangle} T(n) \times X \xrightarrow{\mathrm{st}_1} T(n \times X) \xrightarrow[\cong]{\mathrm{sep}_n} T(X + \cdots + X)\Big),$$

where the distributivity isomorphism sep_n *comes from* (3)*. This instrument is called* side-effect-free *if the following diagram commutes in* $\mathcal{K}\ell(T)$.

$$\begin{tikzcd}
X \arrow[r, "\mathrm{instr}_t"] \arrow[dr, equal] & X + \cdots + X \arrow[d, "\nabla = [\mathrm{id}, \ldots, \mathrm{id}]"] \\
& X
\end{tikzcd}$$

Especially, for each predicate (2-test) $p\colon X \to 2$ *we define an 'assert' map as:*

$$\mathrm{asrt}_p \stackrel{\text{def}}{=} \Big(X \xrightarrow{\mathrm{instr}_p} T(X + X) \xrightarrow{T(\mathrm{id}+!)} T(X + 1) \Big).$$

Hence asrt_p *is an endomap* $X \to X$ *in* $\mathcal{K}\ell(T')$*, for* $T' = T((-) + 1)$.

We then define for another predicate q *on* X *the* sequential conjunction *predicate* $p \,\&\, q$*, pronounced as 'p and then q', as:*

$$p \,\&\, q \stackrel{\text{def}}{=} q \bullet' \mathrm{asrt}_p = \mu \circ T([q, \mathbf{0}]) \circ \mathrm{instr}_p.$$

This instrument terminology comes from [30] (see also [10,58]), where it is used in a setting for quantum computation. Here we adapt the terminology to a monad setting. The instrument is used to interpret, for instance, a 'case' statement as composite:

$$\mathsf{case}\ t\ \mathsf{of}\ (f_0, \ldots, f_{n-1}) = \Big(X \xrightarrow{\mathrm{instr}_t} X + \cdots + X \xrightarrow{[f_0, \ldots, f_{n-1}]} Y \Big).$$

It works as a generalised if-then-else, taking probabilities into account.

Example 25 *We shall review instruments, assert maps, and sequential conjunction in the standard examples from Subsections 3.1 – 3.6. They all have side-effect-free instruments — because they are affine, see Lemma 26* (iii) *later on. We add another monad example where instruments do have side-effects.*

(i) *For the distribution monad* $\mathcal{D}$ *we already mentioned that an* n*-test* $t\colon X \to \mathcal{D}(n)$ *corresponds to* n *predicates* $t_i \in [0,1]^X$*, with* $t_i(x) = t(x)(i)$*, satisfying* $\bigovee_i t_i = \mathbf{1}$*. The associated instrument map* $\mathrm{instr}_t\colon X \to \mathcal{D}(X + \cdots + X)$ *gives a weighted combination of the different coproduct outcomes:*

$$\mathrm{instr}_t(x) = t_0(x)|\kappa_0 x\rangle + \cdots + t_{n-1}(x)|\kappa_{n-1} x\rangle.$$

This instrument is side-effect-free since $t(x) \in \mathcal{D}(n)$ *is a distribution:*

$$\begin{aligned}\Big(\nabla \bullet \mathrm{instr}_t\Big)(x) \;=\; \mathcal{D}(\nabla)\Big(\mathrm{instr}_t(x)\Big) &= (\textstyle\sum_i t_i(x))|x\rangle \\ &= (\textstyle\sum_i t(x)(i))|x\rangle \;=\; 1|x\rangle \;=\; \eta(x).\end{aligned}$$

For predicates $p, q \in [0,1]^X$ *we get:*

$$\mathrm{asrt}_p(x) = p(x)|x\rangle + p^\perp(x)|*\rangle \qquad (p \,\&\, q)(x) = p(x) \cdot q(x).$$

(ii) *For the Giry monad* $\mathcal{G}$ *an* n*-test* $t\colon X \to \mathcal{G}(n) \cong \mathcal{D}(n)$ *can also be identified with* n *predicates* t_i *on* X *with* $\bigovee_i t_i = \mathbf{1}$. *The associated instrument* $\mathrm{instr}_t\colon X \to \mathcal{G}(X + \cdots + X)$ *is given on* $x \in X$ *and* $N \in \Sigma_X$,

$$\mathrm{instr}_t(x)(\kappa_i N) = t_i(x) \cdot \mathbf{1}_N(x).$$

It is side-effect-free since:

$$\begin{aligned}
\big(\nabla \bullet \mathrm{instr}_t\big)(x)(M) &= \mathcal{G}(\nabla)\big(\mathrm{instr}_t(x)\big)(M) \\
&= \mathrm{instr}_t(x)(\nabla^{-1}(M)) \\
&= \mathrm{instr}_t(x)(\kappa_0 M \cup \cdots \cup \kappa_{n-1} M) \\
&= \textstyle\sum_i \mathrm{instr}_t(x)(\kappa_i M) \\
&= \textstyle\sum_i t_i(x) \cdot \mathbf{1}_M(x) \\
&= (\textstyle\sum_i t(x)(i)) \cdot \mathbf{1}_M(x) \;=\; \mathbf{1}_M(x) \;=\; \eta(x)(M).
\end{aligned}$$

Next, for a predicate (measurable function) $p\colon X \to [0,1]$ *we get:*

$$\begin{cases} \mathrm{asrt}_p(x)(M) = p(x) \cdot \mathbf{1}_M \\ \mathrm{asrt}_p(x)(\{*\}) = p^\perp(x) \end{cases} \quad \text{and} \quad (p \,\&\, q)(x) = p(x) \cdot q(x).$$

The situation is similar for the probabilistic powerdomain monad $\mathcal{V}$ *and the Kantorovich monad* $\mathcal{K}$.

(iii) *For the expectation monad* $\mathcal{E}$ *on* **Sets** *tests* $t\colon X \to \mathcal{E}(n) \cong \mathcal{D}(n)$ *correspond to predicates* t_i *with* $\bigovee_i t_i = \mathbf{1}$. *The associated instrument map* $\mathrm{instr}_t\colon X \to \mathcal{E}(X + \cdots + X)$ *is given on* $x \in X$ *and* $q \in [0,1]^{X+\cdots+X}$ *by:*

$$\mathrm{instr}_t(x)(q) = \textstyle\sum_i t_i(x) \cdot q(\kappa_i x).$$

Again the instrument is side-effect-free, since for $p \in [0,1]^X$,

$$\begin{aligned}
\big(\nabla \bullet \mathrm{instr}_t\big)(x)(p) &= \mathcal{E}(\nabla)\big(\mathrm{instr}_t(x)\big)(p) \\
&= \mathrm{instr}_t(x)(p \circ \nabla) \\
&= \textstyle\sum_i t_i(x) \cdot p(x) \;=\; p(x) \;=\; \eta(x)(p).
\end{aligned}$$

Next we have for $p, q \in [0,1]^X$ *and* $r \in [0,1]^{X+1}$

$$\mathrm{asrt}_p(x)(r) = p(x) \cdot r(x) + p^\perp(x) \cdot r(*) \qquad (p \,\&\, q)(x) = p(x) \cdot q(x).$$

(iv) *In all of the above examples instruments are side-effect-free and sequential conjunction & is commutative. This is not always the case, as will be illustrated via the state monad* $T(X) = (S \times X)^S$ *on* **Sets**, *for a fixed set of states* S, *see Example 12* (iii). *A predicate on a set* X *can be identified with a map* $p\colon X \to (S+S)^S$, *since:*

$$T(2) = \Big(S \times 2\Big)^S \cong \Big(S + S\Big)^S.$$

For $x \in X$ *and* $s \in S$ *the value* $p(x)(s) \in S + S$ *describes the 'true' case via the left component of the coproduct* $S+S$, *and the 'false' case via the right component. Clearly, the predicate can also change the state (have a side-effect): the output state* s' *in* $p(x)(s) = \kappa_i s'$ *can be different from the input state* s. *The idea that predicates can have a side-effect is quite natural in imperative languages: consider for instance the equality predicate* `i == j++`. *It returns a Boolean but changes the value of the variable* `j`.

The associated instrument $\mathrm{instr}_p\colon X \to (S \times (X+X))^S \cong (S \times X + S \times X)^S$ *is described by:*

$$\mathrm{instr}_p(x)(s) = \begin{cases} \kappa_1(s', x) & \text{if } p(x)(s) = \kappa_1 s' \\ \kappa_2(s', x) & \text{if } p(x)(s) = \kappa_2 s' \end{cases}$$

Similarly, $\mathrm{asrt}_p\colon X \to (S \times (X+1))^S \cong (S \times X + S)^S$ *is:*

$$\mathrm{asrt}_p(x)(s) = \begin{cases} \kappa_1(s', x) & \text{if } p(x)(s) = \kappa_1 s' \\ \kappa_2 s' & \text{if } p(x)(s) = \kappa_2 s' \end{cases}$$

Hence for predicates $p, q\colon X \to (S+S)^S$ *we have* $p \mathbin{\&} q\colon X \to (S+S)^S$ *described by:*

$$\Big(p \mathbin{\&} q\Big)(x)(s) = \begin{cases} q(x)(s') & \text{if } p(x)(s) = \kappa_1 s' \\ \kappa_2 s' & \text{if } p(x)(s) = \kappa_2 s' \end{cases}$$

The side-effect s' *of* p *is passed on to* q, *if* p *holds. Clearly, & is not commutative for the state monad.*

We collect some basic results about instrument maps.

Lemma 26 *A strong monad* T *on a distributive category* **C** *satisfies the following properties.*

(i) *A test* $t\colon X \to n$ *can be recovered from its instrument, via the following* *diagram in* $\mathcal{K}\ell(T)$.

$$\begin{tikzcd} X \arrow[r, "\mathrm{instr}_t"] \arrow[dr, "t"'] & X + \cdots + X \arrow[d, "! + \cdots + !"] \\ & 1 + \cdots + 1 = n \end{tikzcd}$$

In particular, $\mathrm{instr}_s = s$ *for each scalar* $s\colon 1 \to T(2)$.

(ii) Similarly, a predicate p *can be recovered from its assert map as* $p = \mathbf{1} \bullet' \mathrm{asrt}_p$.

(iii) If t *is causal, then* instr_t *is side-effect-free and causal.*

(iv) For the truth predicate $\mathbf{1}$ *and the falsity predicate* $\mathbf{0}$ *we have:*

$$\begin{cases} \mathrm{instr}_{\mathbf{1}} = \langle \kappa_1 \rangle \\ \mathrm{instr}_{\mathbf{0}} = \langle \kappa_2 \rangle \end{cases} \quad \textit{and} \quad \begin{cases} \mathrm{asrt}_{\mathbf{1}} = \langle \kappa_1 \rangle \\ \mathrm{asrt}_{\mathbf{0}} = \mathbf{0}. \end{cases}$$

(v) $\mathrm{instr}_{p^\perp} = T([\kappa_2, \kappa_1]) \circ \mathrm{instr}_p$, *and* $\mathrm{asrt}_{p^\perp} = T([\kappa_2, \kappa_1]) \circ T(!+\mathrm{id}) \circ \mathrm{instr}_p$.

(vi) $\mathrm{asrt}_p \ovee \mathrm{asrt}_{p^\perp} = T(\kappa_1) \circ T(\nabla) \circ \mathrm{instr}_p$ *in the homset of partial endomaps; as a result, if* instr_p *is side-effect-free, then* $\mathrm{asrt}_p \ovee \mathrm{asrt}_{p^\perp} = \mathrm{id}$ *in* $\mathcal{K}\ell(T')$.

(vii) For a map $f\colon Y \to X$ *in the underlying category* $\mathbf{C}$*, the following diagrams commute in* $\mathbf{C}$.

$$\begin{tikzcd} Y \arrow[r, "\mathrm{instr}_{t\circ f}"] \arrow[d, "f"'] & T(Y + \cdots + Y) \arrow[d, "T(f+\cdots+f)"] \\ X \arrow[r, "\mathrm{instr}_t"'] & T(X + \cdots + X) \end{tikzcd} \qquad \begin{tikzcd} Y \arrow[r, "\mathrm{asrt}_{p\circ f}"] \arrow[d, "f"'] & T(Y + 1) \arrow[d, "T(f+\mathrm{id})"] \\ X \arrow[r, "\mathrm{asrt}_p"'] & T(X + 1) \end{tikzcd}$$

(viii) For predicates p *on* X *and* q *and* Y*, the following diagram commutes in* $\mathbf{C}$.

$$\begin{tikzcd} X + Y \arrow[r, "\mathrm{instr}_p + \mathrm{instr}_q"] \arrow[dr, bend right, "\mathrm{instr}_{[p,q]}"'] & T(X+X) + T(Y+Y) \arrow[d, "{[T(\kappa_1+\kappa_1), T(\kappa_2+\kappa_2)]}"] \\ & T((X+Y) + (X+Y)) \end{tikzcd}$$

We can then write:

$$\mathrm{asrt}_{[p,q]} = \mathrm{asrt}_p + \mathrm{asrt}_q,$$

where the $+$ *on the right-hand-side is used for maps in* $\mathcal{K}\ell(T')$.

Let T *now be commutative, so that* $\mathcal{K}\ell(T)$ *has tensors* $\otimes$ *and parallel conjunction* $\odot$ *as in Definition 2* (viii).

(ix) For a test t *on* X *the following diagram commutes in* $\mathcal{K}\ell(T)$.

$$\begin{tikzcd} X \otimes Y \arrow[r, "\mathrm{instr}_t \otimes \mathrm{id}"] \arrow[dr, bend right, "\mathrm{instr}_{t\circ\pi_1}"'] & (X + \cdots + X) \otimes Y \arrow[d, equal, "\wr"] \\ & X \otimes Y + \cdots + X \otimes Y \end{tikzcd}$$

(x) *For predicates p on X and q on Y one has in $\mathcal{K}\ell(T)$,*

$$\begin{tikzcd}
X \otimes Y \arrow[rr, "\mathrm{instr}_p \otimes \mathrm{instr}_q"] \arrow[ddrr, bend right, "\mathrm{instr}_{p \odot q}"'] & & (X+X) \otimes (Y+Y) \arrow[d, "\cong"', "\mathrm{ddis}"] \\
& & (X \otimes Y) + (X \otimes Y) + (X \otimes Y) + (X \otimes Y) \arrow[d, "T({[\kappa_1, \kappa_2, \kappa_2, \kappa_2]})"] \\
& & (X \otimes Y) + (X \otimes Y)
\end{tikzcd}$$

The map ddis *is the 'double distributivity' map from* (14)*. We can now prove for assert maps:*

$$\begin{aligned}
\mathrm{asrt}_{p \odot q} &= T([\kappa_1, \kappa_2 \circ !, \kappa_2 \circ !, \kappa_2 \circ !]]) \circ T(\mathrm{ddis}) \circ (\mathrm{asrt}_p \otimes \mathrm{asrt}_q) \\
&= \mathrm{dst}' \circ (\mathrm{asrt}_p \otimes \mathrm{asrt}_q) \;=\; \mathrm{asrt}_p \otimes' \mathrm{asrt}_q,
\end{aligned}$$

where dst' *and* $\otimes'$ *are the double strength and the tensor of the category* $\mathcal{K}\ell(T')$ *of partial maps from Lemma 10.*

Proof. (i) We have:

$$\begin{aligned}
(! + \cdots + !) \bullet \mathrm{instr}_t &= T(! + \cdots + !) \circ T(\mathrm{sep}_n) \circ \mathrm{st}_1 \circ \langle t, \mathrm{id} \rangle \\
&= T(\pi_1) \circ \mathrm{st}_1 \circ \langle t, \mathrm{id} \rangle \\
&\overset{(4)}{=} \pi_1 \circ \langle t, \mathrm{id} \rangle \\
&= t.
\end{aligned}$$

(ii) For a predicate p,

$$\mathbf{1} \bullet' \mathrm{asrt}_p = T(! + \mathrm{id}) \circ T(\mathrm{id} + !) \circ \mathrm{instr}_p = T(! + !) \circ \mathrm{instr}_p = p.$$

(iii) Assume that the test t is causal, that is $T(!) \circ t = \ddagger$. We first show that the instrument instr_t is side-effect-free:

$$\begin{aligned}
\nabla \bullet \mathrm{instr}_t &= T(\nabla) \circ T(\mathrm{sep}_n) \circ \mathrm{st}_1 \circ \langle t, \mathrm{id} \rangle \\
&= T(\pi_2) \circ \mathrm{st}_1 \circ \langle t, \mathrm{id} \rangle \\
&= T(\pi_2) \circ T(! \times \mathrm{id}) \circ \mathrm{st}_1 \circ \langle t, \mathrm{id} \rangle \\
&= T(\pi_2) \circ \mathrm{st}_1 \circ (T(!) \times \mathrm{id}) \circ \langle t, \mathrm{id} \rangle \\
&= T(\pi_2) \circ \mathrm{st}_1 \circ \langle T(!) \circ t, \mathrm{id} \rangle \\
&= T(\pi_2) \circ \mathrm{st}_1 \circ \langle \eta \circ !, \mathrm{id} \rangle \qquad \text{since } t \text{ is causal} \\
&\overset{(5)}{=} T(\pi_2) \circ \eta \circ \langle !, \mathrm{id} \rangle \\
&= \eta \circ \pi_2 \circ \langle !, \mathrm{id} \rangle \\
&= \eta.
\end{aligned}$$

The instrument instr_t is causal too:

$$\begin{aligned}
\bar{\top} \bullet \text{instr}_t &= T(!) \circ \text{instr}_t \\
&= T(!) \circ T(\text{sep}_1) \circ \text{st}_1 \circ \langle t, \text{id} \rangle \\
&= T(!) \circ T(\pi_1) \circ \text{st}_1 \circ \langle t, \text{id} \rangle \\
&\overset{(4)}{=} T(!) \circ \pi_1 \circ \langle t, \text{id} \rangle \\
&= T(!) \circ t \\
&= \bar{\top}.
\end{aligned}$$

(iv) For the truth predicate $\mathbf{1} = \eta \circ \kappa_1 \circ !$ we have:

$$\begin{aligned}
\text{instr}_{\mathbf{1}} &= T(\text{sep}_2) \circ \text{st}_1 \circ \langle \eta \circ \kappa_1 \circ !, \text{id} \rangle \\
&\overset{(5)}{=} T(\text{sep}_2) \circ \eta \circ \langle \kappa_1 \circ !, \text{id} \rangle \\
&= \eta \circ \text{sep}_2 \circ \langle \kappa_1 \circ !, \text{id} \rangle \\
&= \eta \circ \kappa_1.
\end{aligned}$$

Next,

$$\text{asrt}_{\mathbf{1}} = T(\text{id} + !) \circ \text{instr}_{\mathbf{1}} = T(\text{id} + !) \circ \langle\kappa_1\rangle = \langle\kappa_1\rangle.$$

The proofs for the falsity predicate $\mathbf{0}$ are similar.

(v) For a predicate p,

$$\begin{aligned}
T([\kappa_2, \kappa_1]) \circ \text{instr}_p &= T([\kappa_2, \kappa_1]) \circ T(\text{sep}_2) \circ \text{st}_1 \circ \langle p, \text{id} \rangle \\
&= T(\text{sep}_2) \circ T([\kappa_2, \kappa_1] \times \text{id}) \circ \text{st}_1 \circ \langle p, \text{id} \rangle \\
&= T(\text{sep}_2) \circ \text{st}_1 \circ (T([\kappa_2, \kappa_1]) \times \text{id}) \circ \langle p, \text{id} \rangle \\
&= T(\text{sep}_2) \circ \text{st}_1 \circ \langle p^\perp, \text{id} \rangle \\
&= \text{instr}_{p^\perp}.
\end{aligned}$$

Hence:

$$\begin{aligned}
\text{asrt}_{p^\perp} &= T(\text{id} + !) \circ \text{instr}_{p^\perp} && \text{by definition} \\
&= T(\text{id} + !) \circ T([\kappa_2, \kappa_1]) \circ \text{instr}_p && \text{as just shown} \\
&= T([\kappa_2, \kappa_1]) \circ T(! + \text{id}) \circ \text{instr}_p.
\end{aligned}$$

(vi) Take as bound $b = \langle\kappa_1\rangle \bullet \text{instr}_p \colon X \to T'(X + X)$. Then:

$$\begin{aligned}
\rhd_1 \bullet' b &= T(\text{id} + !) \circ \text{instr}_p = \text{asrt}_p \\
\rhd_2 \bullet' b &= T([\kappa_2 \circ !, \kappa_1]) \circ \text{instr}_p = T(\text{id} + !) \circ T([\kappa_2, \kappa_1]) \circ \text{instr}_p \\
&= T(\text{id} + !) \circ \text{instr}_{p^\perp} \\
&= \text{asrt}_{p^\perp}.
\end{aligned}$$

Hence:

$$\mathrm{asrt}_p \owedge \mathrm{asrt}_{p^\perp} \;=\; \nabla \bullet' b = T(\nabla + \mathrm{id}) \circ T(\kappa_1) \circ \mathrm{instr}_p$$
$$= T(\kappa_1) \circ T(\nabla) \circ \mathrm{instr}_p.$$

If the instr_p is side-effect-free, that is, if $T(\nabla) \circ \mathrm{instr}_p = \eta$, then $\mathrm{asrt}_p \owedge \mathrm{asrt}_{p^\perp} \leq T(\kappa_1) \circ \eta = \langle\kappa_1\rangle$, where $\langle\kappa_1\rangle \colon X \to T(X+1)$ is the identity $X \to X$ in $\mathcal{K}\ell(T')$. In particular, both asrt_p and $\mathrm{asrt}_{p^\perp}$ are below the identity.

(vii) In a straightforward manner we obtain for a map f in the underlying category:

$$\begin{aligned}
&T(f + \cdots + f) \circ \mathrm{instr}_{p \circ f} \\
&\quad = T(f + \cdots + f) \circ T(\mathrm{sep}_n) \circ \mathrm{st}_1 \circ \langle t \circ f, \mathrm{id}\rangle \\
&\quad = T(\mathrm{sep}_n) \circ T(\mathrm{id} \times f) \circ \mathrm{st}_1 \circ \langle t \circ f, \mathrm{id}\rangle \qquad \text{by naturality of } \mathrm{sep}_n \\
&\quad = T(\mathrm{sep}_n) \circ \mathrm{st}_1 \circ (\mathrm{id} \times f) \circ \langle t \circ f, \mathrm{id}\rangle \\
&\quad = T(\mathrm{sep}_n) \circ \mathrm{st}_1 \circ \langle t, \mathrm{id}\rangle \circ f \\
&\quad = \mathrm{instr}_t \circ f.
\end{aligned}$$

The corresponding result for assert now follows easily.

(viii) Via point (vii) we get:

$$\begin{aligned}
&[T(\kappa_1 + \kappa_1), T(\kappa_2 + \kappa_2)] \circ (\mathrm{instr}_p + \mathrm{instr}_q) \\
&\quad = [T(\kappa_1 + \kappa_1) \circ \mathrm{instr}_{[p,q] \circ \kappa_1}, T(\kappa_2 + \kappa_2) \circ \mathrm{instr}_{[p,q] \circ \kappa_2}] \\
&\quad = [\mathrm{instr}_{[p,q]} \circ \kappa_1, \mathrm{instr}_{[p,q]} \circ \kappa_2] \\
&\quad = \mathrm{instr}_{[p,q]}.
\end{aligned}$$

Then:

$$\begin{aligned}
&\mathrm{asrt}_{[p,q]} \\
&\quad = T(\mathrm{id} + !) \circ \mathrm{instr}_{[p,q]} \\
&\quad = T(\mathrm{id} + !) \circ [T(\kappa_1 + \kappa_1), T(\kappa_2 + \kappa_2)] \circ (\mathrm{instr}_p + \mathrm{instr}_q) \\
&\quad = [T(\kappa_1 + !), T(\kappa_2 + !)] \circ (\mathrm{instr}_p + \mathrm{instr}_q) \\
&\quad = [T(\kappa_1 + \mathrm{id}), T(\kappa_2 + \mathrm{id})] \circ (T(\mathrm{id} + !) + T(\mathrm{id} + !)) \circ (\mathrm{instr}_p + \mathrm{instr}_q) \\
&\quad = [T(\kappa_1 + \mathrm{id}), T(\kappa_2 + \mathrm{id})] \circ (\mathrm{asrt}_p + \mathrm{asrt}_q).
\end{aligned}$$

(ix) We use that the following diagram commutes in **C**.

$$
\begin{tikzcd}
X \times Y \arrow[d, "\langle t,\mathrm{id}\rangle\times\mathrm{id}"'] \arrow[dr, "\langle t\circ\pi_1,\mathrm{id}\rangle"] \arrow[dddd, bend right=60, "\mathrm{instr}_t\otimes\mathrm{id}"'] & & \\
(T(n)\times X)\times Y \arrow[r, "\cong"] \arrow[d, "\mathrm{st}_1\times\mathrm{id}"'] & T(n)\times(X\times Y) \arrow[dr, "\mathrm{st}_1"] & \\
T(n\times X)\times Y \arrow[r, "\mathrm{st}_1"] \arrow[d, "T(\mathrm{sep}_n)\times\mathrm{id}"'] & T((n\times X)\times Y) \arrow[r, "\cong"] \arrow[ddl, "T(\mathrm{sep}_n\times\mathrm{id})"] & T(n\times(X\times Y)) \arrow[dd, "T(\mathrm{sep}_n)"] \\
T(X+\cdots+X)\times Y \arrow[d, "\mathrm{st}_1"'] & & \\
T((X+\cdots+X)\times Y) \arrow[rr, "\cong"'] & & T((X\times Y)+\cdots+(X\times Y))
\end{tikzcd}
$$

(x) This involves a complicated diagram chase. The left path below first describes $\mathrm{instr}_p \otimes \mathrm{instr}_q$, and then the connecting map $T([\kappa_1, \kappa_2, \kappa_2, \kappa_2]) \circ T(\mathrm{ddis})$. The right path describes $\mathrm{instr}_{p \odot q}$.

$$
\begin{tikzcd}
 & X\times Y \arrow[d, "\langle p,\mathrm{id}\rangle\times\langle q,\mathrm{id}\rangle"'] \arrow[dr, "\langle p\times q,\mathrm{id}\rangle"] & & \\
 & (T(2)\times X)\times(T(2)\times Y) \arrow[r, "\beta", "\cong"'] \arrow[d, "\mathrm{st}_1\times\mathrm{st}_1"] & (T(2)\times T(2))\times(X\times Y) \arrow[d, "\mathrm{dst}\times\mathrm{id}"'] & \\
 & T(2\times X)\times T(2\times Y) \arrow[dl, "T(\mathrm{sep}_2)\times T(\mathrm{sep}_2)"'] \arrow[d, "\mathrm{dst}"] & T(2\times 2)\times(X\times Y) \arrow[d, "\mathrm{st}_1"'] \arrow[dr, "T(\mathrm{sep}_2)\times\mathrm{id}"] & \\
T(X+X)\times T(Y+Y) \arrow[dr, "\mathrm{dst}"'] & T((2\times X)\times(2\times Y)) \arrow[r, "T(\beta)", "\cong"'] \arrow[d, "T(\mathrm{sep}_2\times\mathrm{sep}_2)"] & T((2\times 2)\times(X\times Y)) \arrow[d, "T(\mathrm{sep}_2\times\mathrm{id})"] & T(2+2)\times(X\times Y) \arrow[d, "T([\mathrm{id},\mathbf{0}])\times\mathrm{id}"] \\
 & T((X+X)\times(Y+Y)) \arrow[d, "T(\mathrm{ddis})"'] & T((2+2)\times(X\times Y)) \arrow[dl, "T(\mathrm{sep}_4)"'] \arrow[d, "T([\mathrm{id},\mathbf{0}]\times\mathrm{id})"] & T(2)\times(X\times Y) \arrow[dl, "\mathrm{st}_1"] \\
 & T((X\times Y)+(X\times Y)+(X\times Y)+(X\times Y)) \arrow[d, "T([\kappa_1,\kappa_2,\kappa_2,\kappa_2])"'] & T(2\times(X\times Y)) \arrow[d, "T(\mathrm{sep}_2)"] & \\
 & T((X\times Y)+(X\times Y)) \arrow[r, equal] & T((X\times Y)+(X\times Y)) &
\end{tikzcd}
$$

The isomorphism in the middle is:

$$
(A\times X)\times(B\times Y) \xrightarrow[\cong]{\beta=\langle\pi_1\times\pi_1,\pi_2\times\pi_2\rangle} (A\times B)\times(X\times Y)
$$

This β can also be described as composite of associativity and symmetry isomorphisms for $\otimes$. This allows us to prove commutation of the upper-middle rectangle, using (4). Commutation of the lower-middle rectangle invovles some elementary bookkeeping. □

We conclude with a re-interpretation of an earlier result, namely Proposition 19, as a bijective correspondence between tests and side-effect-free instruments. The correspondence extends to predicates and assert maps. The main point here is that strong affineness connects predicates to side-effect-free instruments and assert maps.

Proposition 27 *Let T be a strongly affine monad on a distributive category.*

(i) Each instrument is then side-effect-free, and the mapping $t \mapsto \mathrm{instr}_t$ gives a bijective correspondence:

$$\frac{\textit{tests } X \xrightarrow{t} T(n)}{\overline{X \xrightarrow[f]{} T(X + \cdots + X) \textit{ with } T(\nabla) \circ f = \eta}}$$

(ii) Assume next that T satisfies the conditions from Theorem 23, making $\mathcal{K}\ell(T)$ an effectus. Then, each partial map $f\colon X \to T(X+1)$ that is below the identity on X in $\mathcal{K}\ell(T')$ satisfies $f = \mathrm{asrt}_p$, for the predicate $p = \mathbf{1} \bullet' f$. This gives a bijective correspondence:

$$\frac{\textit{predicates } X \xrightarrow{p} T(2)}{\overline{X \xrightarrow[f]{} X \textit{ in } \mathcal{K}\ell(T') \textit{ with } f \leq \mathrm{id}}}$$

(iii) The equation $\mathrm{asrt}_{p\&q} = \mathrm{asrt}_q \bullet' \mathrm{asrt}_p$ holds.

The condition $f \leq \mathrm{id}$ in point (ii) expresses side-effect-freeness for partial endomaps f.

Proof. (i) If T is strongly affine, then instruments are side-effect-free, since, as in the proof of Lemma 26 (iii):

$$\begin{aligned} \nabla \bullet \mathrm{instr}_p \;=\; T(\nabla) \circ T(\mathrm{sep}_n) \circ \mathrm{st}_1 \circ \langle t, \mathrm{id}\rangle \;&=\; T(\pi_2) \circ \mathrm{st}_1 \circ \langle t, \mathrm{id}\rangle \\ &\overset{(17)}{=}\; \eta \circ \pi_2 \circ \langle t, \mathrm{id}\rangle \;=\; \eta. \end{aligned}$$

The bijective correspondence in point (i) is a minor reformulation of the one from Proposition 19.

(ii) In the downward direction we send a predicate p to the partial map asrt_p. It is below the identity, by Lemma 26 (vi), using that instr_p is side-effect-free, by point (i). We recover p via: $T(! + \mathrm{id}) \circ \mathrm{asrt}_p = \mathbf{1} \bullet' \mathrm{asrt}_p = p$.

The upward direction requires more work. Let $f\colon X \to T(X+1)$ be a map below the identity $\langle\kappa_1\rangle\colon X \to X$ in $\mathcal{K}\ell(T')$. We write $p = \mathbf{1} \bullet' f$. Then $f \ovee g = \langle\kappa_1\rangle$ for some map g, say with bound $b\colon X \to T'(X+X)$. Hence:

$$\mathbf{1} = \mathbf{1} \bullet' \langle\kappa_1\rangle = \mathbf{1} \bullet' (f \ovee g) = \mathbf{1} \bullet' \nabla \bullet' b = \mathbf{1} \bullet' b.$$

Lemma 22 (ii) yields that b is a total map, say of the form $b = T(\kappa_1) \circ c$, for a unique map $c\colon X \to T(X+X)$. This map c satisfies:

$$\begin{aligned}
T(\mathrm{id}+!) \circ c &= T([\mathrm{id}+!, \kappa_2]) \circ T(\kappa_1) \circ c \\
&= \rhd_1 \bullet' b \;=\; f \\
T(!+!) \circ c &= T(\mathrm{id}+!) \circ f \;=\; \mathbf{1} \bullet' f \;=\; p \\
T(\kappa_1) \circ T(\nabla) \circ c &= T(\nabla + \mathrm{id}) \circ T(\kappa_1) \circ c \\
&= \nabla \bullet' b \;=\; f \owedge g \;=\; \langle \kappa_1 \rangle \;=\; T(\kappa_1) \circ \eta.
\end{aligned}$$

Then $T(\nabla) \circ c = \eta$, since $T(\kappa_1)$ is monic, see Lemma 21. The bijective correspondence in point (i) yields $c = \mathrm{instr}_p$, and thus:

$$f \;=\; T(\mathrm{id}+!) \circ c = T(\mathrm{id}+!) \circ \mathrm{instr}_p \;=\; \mathrm{asrt}_p.$$

(iii) Let p, q be predicates on the same object X. As in the previous point, the associated assert maps are below the identity. Hence their composite satisfies $\mathrm{asrt}_q \bullet' \mathrm{asrt}_p \leq \mathrm{id} \bullet' \mathrm{id} = \mathrm{id}$, using Lemma 8 (ii). Moreover,

$$\mathbf{1} \bullet' \mathrm{asrt}_q \bullet' \mathrm{asrt}_p = q \bullet' \mathrm{asrt}_p = p \,\&\, q = \mathbf{1} \bullet' \mathrm{asrt}_{p \& q}.$$

Since $\mathrm{asrt}_{p\&q} \leq \mathrm{id}$, we get $\mathrm{asrt}_q \bullet' \mathrm{asrt}_p = \mathrm{asrt}_{p\&q}$ from the bijective correspondence in the previous point. □

8 Commutativity, of monads and of sequential conjunction

In this section we assume that T is a strong monad on a distributive category $\mathbf{C}$, so that we can define instrument and assert maps, and sequential conjunction &, see Definition 24. In Example 25 we have seen that for all our ‘probability’ monads andthen & is commutative. But this does not hold in general, see in particular the state monad in Example 25 (iv). It is also a key feature of the quantum world that sequential conjunction is not commutative, see [10] for details.

The main result of this section, Corollary 29, says that if a monad T is commutative, then the sequential conjunction (‘andthen’) operation & on predicates in $\mathcal{K}\ell(T)$ is also commutative. Given the terminological coincidence, this may seem natural, but the settings are quite different and *a priori* unrelated. Here we do establish a connection, via a non-trivial calculation.The theorem below plays a central role.

Theorem 28 *If T is a commutative monad on a distributive category, then instruments commute: for predicates p, q on an object X, the following diagram* *commutes in* $\mathcal{K}\ell(T)$*.*

$$
\begin{tikzcd}
X \arrow[r, "\mathrm{instr}_p"] \arrow[d, equal] & X+X \arrow[r, "q+q"] & 2+2 \arrow[d, "{[\kappa_1+\kappa_1,\kappa_2+\kappa_2]}", "\cong"'] \\
X \arrow[r, "\mathrm{instr}_q"'] & X+X \arrow[r, "p+p"'] & 2+2
\end{tikzcd}
\tag{26}
$$

Proof. The structure of the proof is given by the following diagram in the underlying category.

$$
\begin{tikzcd}
& & X \arrow[dl, "\langle p,\mathrm{id}\rangle"'] \arrow[dr, "\langle q,\mathrm{id}\rangle"] \arrow[dddll, bend right, "\mathrm{instr}_p"'] \arrow[dddrr, bend left, "\mathrm{instr}_q"] & & \\
& T(2)\times X \arrow[dl, "\mathrm{st}_1"'] \arrow[d, "\mathrm{id}\times q"] & & T(2)\times X \arrow[d, "\mathrm{id}\times p"'] \arrow[dr, "\mathrm{st}_1"] & \\
T(2\times X) \arrow[d, "T(\mathrm{sep}_2)"'] \arrow[dr, phantom, "(a)"] & T(2)\times T(2) \arrow[rr, "\gamma=\langle\pi_2,\pi_1\rangle", "\cong"'] \arrow[d, "\mathrm{dst}"] \arrow[ddl] & & T(2)\times T(2) \arrow[d, "\mathrm{dst}"'] \arrow[ddr] \arrow[dr, phantom, "(a)"] & T(2\times X) \arrow[d, "T(\mathrm{sep}_2)"] \\
T(X+X) \arrow[d, "T(q+q)"'] \arrow[dr, phantom, "(b)"] & T(2\times 2) \arrow[rr, "T(\gamma)", "\cong"'] \arrow[d, "T(\mathrm{sep}_2)"] \arrow[drr, phantom, "(c)"] & & T(2\times 2) \arrow[d, "T(\mathrm{sep}_2)"'] \arrow[dr, phantom, "(b)"] & T(X+X) \arrow[d, "T(p+p)"] \\
T(T(2)+T(2)) \arrow[r, "\mu\circ", "T({[T(\kappa_1),T(\kappa_2)]})"'] & T(2+2) \arrow[rr, "\cong", "T({[\kappa_1+\kappa_1,\kappa_2+\kappa_2]})"'] & & T(2+2) & T(T(2)+T(2)) \arrow[l, "\mu\circ"', "T({[T(\kappa_1),T(\kappa_2)]})"]
\end{tikzcd}
$$

Sub-diagrams (a) commute by naturality, and sub-diagrams (b) by (6); commutation of (c) is easy, and the square in the middle is commutativity of the monad T, see (7). Details are left to the interested reader. □

Our next result combines all our previous requirements on a monad. It guarantees that the resulting monad is a commutative effectus. It is an open question to what extend these conditions are also necessary.

Corollary 29 *Let* T *be a commutative, partially additive, strongly affine, not-trivialising monad on a non-trivial distributive category. Its Kleisli category* $\mathcal{K}\ell(T)$ *is then a* monoidal, commutative *effectus.*

Proof. Theorem 23 tells that the Kleisli category $\mathcal{K}\ell(T)$ is an effectus. We now have the additional assumption that the monad is commutative. We show that this implies that sequential conjunction & is commutative. For this we first note that in $\mathcal{K}\ell(T)$ we can write $\mathrm{asrt}_p = (\mathrm{id} + \bar{\top}) \bullet \mathrm{instr}_p$ for the ground map $\bar{\top} = \eta \circ !\colon X \to T(1)$ from (16). Moreover, each predicate p is causal, satisfying $\bar{\top} \bullet p = \bar{\top}$, since the monad T is affine, see Lemma 14. Hence in

$\mathcal{K}\ell(T)$,

$$\begin{aligned}
p \,\&\, q &= [q, \kappa_2] \bullet (\mathrm{id} + \bar{\bar{\top}}) \bullet \mathrm{instr}_p \\
&= [\mathrm{id}, \kappa_2 \bullet \bar{\bar{\top}}] \bullet (q + q) \bullet \mathrm{instr}_p && \text{since } q \text{ is causal} \\
&= [\mathrm{id}, \kappa_2 \bullet \bar{\bar{\top}}] \bullet (p + p) \bullet \mathrm{instr}_q && \text{by Theorem 28} \\
&= [p, \kappa_2] \bullet (\mathrm{id} + \bar{\bar{\top}}) \bullet \mathrm{instr}_q && \text{since } p \text{ is causal} \\
&= q \,\&\, p.
\end{aligned}$$

The Kleisli category $\mathcal{K}\ell(T)$ now satisfies the two requirements for a *commutative* effectus (from [10]):

- there is correspondence between predicates on X and partial maps $X \to X$ which are below the identity ('side-effect-free'), see Proposition 27 (ii);
- sequential conjunction & is commutative.

The effectus $\mathcal{K}\ell(T)$ is also *monoidal*, which according to [10] means three things:

- the category $\mathcal{K}\ell(T)$ is symmetric monoidal; this follows from the fact that the monad T is commutative, as mentioned in Section 3;
- the tensor unit 1 is final in $\mathcal{K}\ell(T)$; this follows from the fact that T is affine — *i.e.* satisfies $T(1) \cong 1$ — and means that tensors $\otimes$ come with projections $X \leftarrow X \otimes Y \to Y$ in $\mathcal{K}\ell(T)$; these projections are used for weakening of predicates and marginalisation of states;
- the tensor $\otimes$ of $\mathcal{K}\ell(T)$ distributes over coproducts $(0, \times)$; this is guaranteed by the fact that the underlying category is distributive. □

This result can be applied, in principle, to our monad examples in Subsection 3.1 – 3.6. A problem is that it is not known for all of the monads if they are commutative. For instance, this is unclear for the expectation monad $\mathcal{E}$, but nevertheless its sequential conjunction is commutative, see Example 25 (iii).

9 Normalisation and conditioning

In this final section we illustrate how the abstract and uniform effectus-theoretic look at probability monads can be used, in particular to give a systematic description of normalisation and conditioning of states, as in [41] and [42]. We start with a first observation (of Sean Tull); we copy the details from [10] and adapt them to the current context, with Kleisli categories as effectus. The observation applies to our running examples because they all have the unit interval as set of scalars (see Section 3).

Lemma 30 *Let T be a monad whose Kleisli category $\mathcal{K}\ell(T)$ is an effectus, as in Theorem 23, with the unit interval $[0,1]$ as its set of scalar $1 \to T(2)$. For each non-zero 'partial' state $\omega\colon 1 \to X+1$ in $\mathcal{K}\ell(T)$ there is a unique 'total' state $\mathrm{nrm}(\omega)\colon 1 \to X$ making, for the scalar $r = \mathbf{1} \bullet' \omega$, the following diagram in $\mathcal{K}\ell(T)$ commute.*

$$\begin{tikzcd} 1 \arrow[r, "r"] \arrow[dr, bend right, "\omega"'] & 1+1 \arrow[d, "\mathrm{nrm}(\omega)+\mathrm{id}"] \\ & X+1 \end{tikzcd} \tag{27}$$

Proof. Let $\omega\colon 1 \to T(X+1)$ be a non-zero partial state, with corresponding scalar $r = \mathbf{1} \bullet' \omega = T(!+\mathrm{id}) \circ \omega \in [0,1]$. The assumption $\omega \neq \mathbf{0}$ translates to $r \neq 0$ — by Lemma 22 (i). Thus we can find an $n \in \mathbb{N}$ and $r' \in [0,1]$ with $r' < r$ and $n \cdot r + r' = 1$. More abstractly, we can find scalars $s_1, \ldots, s_m \in [0,1]$ with $\bigovee_i s_i \cdot r = \sum_i s_i \cdot r = 1$. We now form the scaled partial states $\omega \bullet' s_i\colon 1 \to T(X+1)$. Their scalars $\mathbf{1} \bullet' \omega \bullet' s_i = r \bullet s_i = r \cdot s_i \in [0,1]$ are orthogonal (summable), so the maps $\omega \bullet' s_i$ are orthogonal too, since $\mathbf{1} \bullet' (-)$ reflects orthogonality, by Lemma 8 (v). Consider the partial state $\bigovee_i (\omega \bullet' s_i)$. It is actually total, by Lemma 22 (ii), since:

$$\mathbf{1} \bullet' \bigovee_i (\omega \bullet' s_i) = \bigovee_i \mathbf{1} \bullet' \omega \bullet' s_i = \bigovee_i r \bullet' s_i = \textstyle\sum_i s_i \cdot r = \mathbf{1}.$$

Hence we define $\mathrm{nrm}(\omega)\colon 1 \to T(X)$ to be the unique map with $T(\kappa_1) \circ \mathrm{nrm}(\omega) = \bigovee_i (\omega \bullet' s_i)$. By construction Diagram (27) commutes:

$$\begin{aligned} (\mathrm{nrm}(\omega) + \mathrm{id}) \bullet r &= [T(\kappa_1) \circ \mathrm{nrm}(\omega), \langle \kappa_2 \rangle] \bullet r \\ &= \big(T(\kappa_1) \circ \mathrm{nrm}(\omega)\big) \bullet' r \\ &= \big(\bigovee_i \omega \bullet' s_i\big) \bullet' r \\ &= \bigovee_i \omega \bullet' s_i \bullet' r \\ &= \omega \bullet' \bigovee_i s_i \bullet' r \\ &= \omega \bullet' \mathbf{1} \\ &= \omega. \end{aligned}$$

We still have to prove uniqueness. If $\rho\colon 1 \to T(X)$ also satisfies $(T(\kappa_1) \circ \rho) \bullet' r = \omega = (T(\kappa_1) \circ \mathrm{nrm}(\omega)) \bullet' r$, then we obtain $\rho = \mathrm{nrm}(\omega)$ from the fact that $T(\kappa_1)$ is monic, see Lemma 21.

$$\begin{aligned} T(\kappa_1) \circ \rho &= (T(\kappa_1) \circ \rho) \bullet' \bigovee_i r \bullet' s_i \\ &= \bigovee_i (T(\kappa_1) \circ \rho) \bullet' r \bullet' s_i \\ &= \bigovee_i \omega \bullet' s_i \\ &= \bigovee_i (T(\kappa_1) \circ \mathrm{nrm}(\omega)) \bullet' r \bullet' s_i \\ &= (T(\kappa_1) \circ \mathrm{nrm}(\omega)) \bullet' \bigovee_i r \bullet' s_i \\ &= T(\kappa_1) \circ \mathrm{nrm}(\omega). \end{aligned}$$

□

We briefly illustrate how normalisation works in the running examples.

Example 31 *Let $\omega \in \mathcal{D}(X+1)$ be a non-zero partial state (subdistribution), say $\omega = \sum_{i<n} r_i|x_i\rangle + r_n|*\rangle$. The non-zero requirement means that $r_n \neq 1$. The normalised state $\mathrm{nrm}(\omega) \in \mathcal{D}(X)$ is the (proper) distribution:*

$$\mathrm{nrm}(\omega) = \sum_i \frac{r_i}{1-r_n}|x_i\rangle \quad \textit{with} \quad \sum_i \frac{r_i}{1-r_n} = \frac{\sum_i r_i}{1-r} = \frac{1-r_n}{1-r_n} = 1.$$

Let $\omega \in \mathcal{G}(X+1)$ now be a non-zero partial state for the Giry monad $\mathcal{G}$. The probability measure $\omega\colon \Sigma_{X+1} \to [0,1]$ then satisfies $\omega(\{*\}) \neq 1$, or equivalently, $\omega(X) \neq 0$. Its normalised probability measure $\mathrm{nrm}(\omega)\colon \Sigma_X \to [0,1]$ is given by $\mathrm{nrm}(\omega)(M) = \frac{\omega(M)}{\omega(X)}$. This works in the same way for the probabilistic powerdomain $\mathcal{V}$ and the Kantorovich monad $\mathcal{K}$.

Similarly, let $\omega \in \mathcal{E}(X+1) = \mathbf{EMod}([0,1]^{X+1},[0,1])$ be non-zero. Then $\omega(\mathbf{1}_X) \neq 0$, so that $\mathrm{nrm}(\omega)(p) = \frac{\omega(p)}{\omega(\mathbf{1}_X)}$ for $p \in [0,1]^X$. A similar construction works for the Radon monad.

The normalisation operation nrm on partial states is (algebraically) not well-behaved, especially because of the non-zeroness precondition. A better behaved alternative, for discrete probability, is described in [35]. In the current setting we can prove the following two points.

Lemma 32 *In the context of Lemma 30,*

(i) $\mathrm{nrm}(T(\kappa_1) \circ \omega) = \omega$, *for a (total) state* $\omega\colon 1 \to T(X)$;
(ii) $T(f) \circ \mathrm{nrm}(\omega) = \mathrm{nrm}(T(f+\mathrm{id}) \circ \omega)$, *for a map f in the underlying category.*

Proof. The first point follows from uniqueness in Diagram (27) since the scalar associated with the artificially partial state $T(\kappa_1) \circ \omega$ is the scalar $\mathbf{1}$:

$$\begin{aligned}
\mathbf{1} \bullet' (T(\kappa_1) \circ \omega) &= T(! + \mathrm{id}) \circ T(\kappa_1) \circ \omega \\
&= T(\kappa_1) \circ T(!) \circ \omega \\
&= T(\kappa_1) \circ \eta \qquad\qquad \text{since } T(1) \text{ is final} \\
&= \mathbf{1}.
\end{aligned}$$

We are then done by uniqueness in (27), since:

$$(\omega + \mathrm{id}) \bullet \mathbf{1} = \mu \circ T([T(\kappa_1) \circ \omega, \langle\kappa_2\rangle]) \circ \eta \circ \kappa_1 = T(\kappa_1) \circ \omega.$$

For the second point we first note that the scalars associated with the partial states ω and $T(f+\mathrm{id}) \circ \omega$ are the same, since:

$$\mathbf{1} \bullet' (T(f+\mathrm{id}) \circ \omega) = T(!+\mathrm{id}) \circ T(f+\mathrm{id}) \circ \omega = T(!+\mathrm{id}) \circ \omega = \mathbf{1} \bullet' \omega$$

Next we are done by uniqueness in Diagram (27):

$$
\begin{aligned}
&\Big((T(f)\circ \mathrm{nrm}(\omega)) + \mathrm{id}\Big) \bullet \Big(\mathbf{1} \bullet' (T(f+\mathrm{id})\circ\omega)\Big) \\
&\quad= \; ((\langle f\rangle \bullet \mathrm{nrm}(\omega)) + \mathrm{id}) \bullet (\mathbf{1}\bullet' \omega) \\
&\quad= \; (\langle f\rangle + \mathrm{id}) \bullet (\mathrm{nrm}(\omega) + \mathrm{id}) \bullet (\mathbf{1}\bullet' \omega) \\
&\quad\overset{(27)}{=} \; T(f+\mathrm{id})\circ\omega. \qquad\square
\end{aligned}
$$

Normalisation forms the basis for conditioning. It uses assert maps from Definition 24. For a state ω of X and a predicate p on X we obtain a partial state $\rho = \mathrm{asrt}_p \bullet \omega \colon 1 \to X+1$. In this situation one can say that p is a density function for ρ, see Example 34 (ii) below. We obtain a conditional state $\omega|_p$ by normalising the partial state ρ to a total state. This conditional state $\omega|_p$ is introduced in [41] and used in [42]. In slightly different form it occurs in [52, Sect. 16].

Definition 33 *Let $\mathcal{K}\ell(T)$ be an effectus with scalars $[0,1]$. Let ω be a state of X and p a predicate on the same object X for which the validity $\omega \models p \in [0,1]$ is non-zero. Then we define the conditional state $\omega|_p$, pronounced as "ω given p", as normalisation of $\mathrm{asrt}_p \bullet \omega$ in $\mathcal{K}\ell(T)$, see:*

$$\omega|_p = \mathrm{nrm}\Big(1 \xrightarrow{\omega} X \xrightarrow{\mathrm{asrt}_p} X+1\Big).$$

Using Diagram (27) we see that $\omega|_p$ is the unique state satisfying, in $\mathcal{K}\ell(T)$,

$$
\begin{tikzcd}
1 \arrow[rr, "\omega\models p"] \arrow[dr, "\omega"'] & & 1+1 \arrow[dd, "\omega|_p+\mathrm{id}"] \\
& X \arrow[dr, "\mathrm{asrt}_p"'] & \\
& & X+1
\end{tikzcd}
\qquad (28)
$$

The normalisation used above, in the definition of $\omega|_p$, exists because by Lemma 26 (ii):

$$
\begin{aligned}
\mathbf{1}\bullet' (\mathrm{asrt}_p \bullet \omega) \;=\; \mathbf{1}\bullet' \mathrm{asrt}_p \bullet' (\langle\kappa_1\rangle \bullet \omega) &= p \bullet' (\langle\kappa_1\rangle \bullet \omega) \\
&= p\bullet\omega \;=\; \omega \models p \;\neq\; \mathbf{0}.
\end{aligned}
$$

We illustrate conditioning for the distribution monad and the Giry monad and show that standard conditional probability forms a special case, using 'sharp' predicates.

Example 34 *(i) For a predicate $p \in [0,1]^X$ and a state/distribution $\omega \in \mathcal{D}(X)$ with $\omega \models p \neq 0$ the conditional distribution $\omega|_p \in \mathcal{D}(X)$ is given* *by:*

$$\omega|_p = \sum_{x \in X} \frac{\omega(x) \cdot p(x)}{\omega \models p} \big| x \big\rangle.$$

For another predicate $q \in [0,1]^X$ we use the validity $\omega|_p \models q$ as "q, given p" wrt. distribution ω. This specialises to the usual form of conditional probability[2] *. For an event/subset $E \subseteq X$ we write $\mathbf{1}_E \colon X \to [0,1]$ for the associated 'sharp' predicate. The probability of E is commonly written as $P(E)$, or as $P_\omega(E)$, with the state/distribution ω explicit. Notice that $P_\omega(E)$ is a special case of our validity notation:*

$$\omega \models \mathbf{1}_E = \textstyle\sum_x \omega(x) \cdot \mathbf{1}_E(x) = \sum_{x \in E} \omega(x) = P_\omega(E).$$

The conditional probability $P_\omega(D \mid E)$ also arises as special case:

$$\begin{aligned} \Big(\omega|_{\mathbf{1}_E} \models \mathbf{1}_D\Big) &= \sum_{x \in X} \omega|_{\mathbf{1}_E}(x) \cdot \mathbf{1}_D(x) \;=\; \sum_{x \in X} \frac{\omega(x) \cdot \mathbf{1}_E(x) \cdot \mathbf{1}_D(x)}{\omega \models \mathbf{1}_E} \\ &= \frac{\sum_{x \in X} \omega(x) \cdot \mathbf{1}_{E \cap D}(x)}{P_\omega(E)} \;=\; \frac{P_\omega(E \cap D)}{P_\omega(E)} \;=\; P_\omega(D \mid E). \end{aligned}$$

As illustration, consider a distribution $\omega = \frac{1}{4}|a\rangle + \frac{1}{3}|b\rangle + \frac{5}{12}|c\rangle$ on a set $A = \{a, b, c\}$, a subset $E = \{a, c\} \subseteq A$ and a predicate $p \in [0,1]^A$ with $p(a) = \frac{1}{2}, p(b) = \frac{1}{4}, p(c) = 1$. Then one can check:

$$\begin{array}{lll} \omega \models \mathbf{1}_E = \frac{2}{3} & \omega|_{\mathbf{1}_E} = \frac{3}{8}|a\rangle + \frac{5}{8}|c\rangle & \omega|_{\mathbf{1}_E} \models p = \frac{13}{16} \\ \omega \models p = \frac{5}{8} & \omega|_p = \frac{1}{5}|a\rangle + \frac{2}{15}|b\rangle + \frac{2}{3}|c\rangle & \omega|_p \models \mathbf{1}_E = \frac{13}{15}. \end{array}$$

Conditional states can also be used to define the (regular) conditional associated with a joint state $\omega \in \mathcal{D}(X \times Y)$. First, each element $y \in Y$ gives rise to a singleton (sharp) predicate $\mathbf{1}_{\{y\}} \in [0,1]^Y$, and thus to $p_y = \pi_2^*(\mathbf{1}_{\{y\}}) \in [0,1]^{X \times Y}$. This predicate p_y satisfies $p_y(x, z) = 1$ iff $y = z$. Under suitable side-conditions we can define a conditional *map $f \colon Y \to \mathcal{D}(X)$ for $\omega \in \mathcal{D}(X \times Y)$ as:*

$$f(y) = (\pi_1)_*\Big(\omega|_{p_y}\Big) = \sum_x \frac{\omega(x, y)}{\omega_2(y)} \big| x \big\rangle,$$

where $\omega_2 \in \mathcal{D}(Y)$ is the marginal, given by $\omega_2(y) = \omega \models p_y = \sum_x \omega(x, y)$. Then we can reconstruct ω from f and this marginal ω_2 as $\omega = \mathrm{gr}(f)_*(\omega_2)$, where $\mathrm{gr}(f) \colon Y \to \mathcal{D}(X \times Y)$ is $\mathrm{gr}(f) = \mathrm{st}_1 \circ \langle f, \mathrm{id}\rangle$, as in Proposition 19.

(ii) We turn to the Giry monad $\mathcal{G}$ and recall that for a state/measure $\omega \in \mathcal{G}(X)$ and a predicate (measurable function) $p \colon X \to [0,1]$ we have $\omega \models$

[2] We avoid the notation $D \mid E$, or $q \mid p$, for conditional probability, because it wrongly suggests that '$|$' is an operation on predicates. Instead, we use '$|$' in $\omega|_p$ describing it as a (right) action of predicates p on distributions ω, see Theorem 35 (ii).

$p = p \bullet \omega = \int p \,\mathrm{d}\omega$. Each measurable subset $M \in \Sigma_X$ gives a 'sharp' predicate $\mathbf{1}_M \colon X \to [0,1]$ with $\omega \models \mathbf{1}_M = \int \mathbf{1}_M \,\mathrm{d}\omega = \omega(M) = P_\omega(M)$. Following the descriptions from Subsection 3.2 and Example 25 (ii) *we get:*

$$\Big(\mathrm{asrt}_p \bullet \omega\Big)(M) = \int \mathrm{asrt}_p(-)(M)\,\mathrm{d}\omega = \int p(-)\cdot \mathbf{1}_M \,\mathrm{d}\omega = \int_M p\,\mathrm{d}\omega.$$

Now we see that the conditional state/measure $\omega|_p \colon \Sigma_X \to [0,1]$ is given by:

$$\omega|_p(M) = \frac{\int_M p\,\mathrm{d}\omega}{\omega \models p} = \frac{\int_M p\,\mathrm{d}\omega}{\int p\,\mathrm{d}\omega}.$$

If we specialise to sharp predicates given by measurable subsets $M, N \subseteq X$ we obtain the usual formulation of conditional probability $P_\omega(M \mid N)$:

$$\omega|_{\mathbf{1}_N} \models \mathbf{1}_M = \omega|_{\mathbf{1}_N}(M) = \frac{\int_M \mathbf{1}_N\,\mathrm{d}\omega}{\int \mathbf{1}_N\,\mathrm{d}\omega} = \frac{\int \mathbf{1}_{M\cap N}\,\mathrm{d}\omega}{\omega(N)} = \frac{\omega(M\cap N)}{\omega(N)}.$$

Regular conditionals are more complicated for the Giry monad than for the distribution monad, essentially since the singleton predicate $\mathbf{1}_{\{y\}}$ that we used in the previous point may not exist, see e.g. [14] for more information.

(iii) For the expectation monad $\mathcal{E}$ on **Sets**, *let $\omega \in \mathcal{E}(X)$ be a set and $p \in [0,1]^X$ be a predicate with $\omega \models p = \omega(p) \neq 0$. Following Example 25* (iii), *the partial state $\mathrm{asrt}_p \bullet \omega \in \mathcal{E}(X+1)$ is given on $r \in [0,1]^{X+1}$ by:*

$$(\mathrm{asrt}_p \bullet \omega)(r) = \omega\Big(\lambda x.\, \mathrm{asrt}_p(x)(r)\Big) = \omega\Big(\lambda x.\, p(x)\cdot r(x) + p^\perp(x)\cdot r(*)\Big).$$

Hence the conditional state $\omega|_p \in \mathcal{E}(X)$ is defined on $q \in [0,1]^X$ as:

$$\omega|_p(q) = \frac{\omega(\lambda x.\, p(x)\cdot q(x))}{\omega(p)} = \frac{\omega(p \,\&\, q)}{\omega(p)}.$$

We conclude with a number of fundamental properties of conditional states that can be proved abstractly, independent of the monad involved.

Theorem 35 *Let $\mathcal{K}\ell(T)$ be an effectus with scalars* $[0,1]$.

(i) Bayes' rule holds, in multiplicative form:

$$\Big(\omega|_p \models q\Big)\cdot\Big(\omega \models p\Big) = \Big(\omega \models p \,\&\, q\Big).$$

(ii) Conditioning behaves like an action:

$$\omega|_{\mathbf{1}} = \omega \qquad \textit{and} \qquad (\omega|_p)|_q = \omega|_{p\&q}.$$

(iii) For a map f in the underlying category:

$$T(f) \circ (\omega|_{q \circ f}) = (T(f) \circ \omega)|_q.$$

For the remaining points we assume that the monad T is commutative, so that the effectus $\mathcal{K}\ell(T)$ is monoidal, see Corollary 29.

(iv) Fubini holds: for states $\omega \colon 1 \to X$ and $\rho \colon 1 \to Y$, yielding a product state $\omega \otimes \rho \colon 1 \to X \otimes Y$ as in Definition 2 (vii), *and for a predicate p on the product $X \otimes Y$,*

$$\omega \models (\mathrm{id} \otimes \rho)^*(p) \;=\; \omega \otimes \rho \models p \;=\; \rho \models (\omega \otimes \mathrm{id})^*(p).$$

(v) Let ω, ρ be states on X, Y and p, q be predicates on X, Y. Conditioning can be done in parallel, since:

$$(\omega \otimes \rho)|_{(p \odot q)} = (\omega|_p) \otimes (\rho|_q).$$

(vi) Let ω and ρ be two states on X and Y and p a predicate on $X \otimes Y$. Marginalising after conditioning $\omega \otimes \rho$ with p is the same as conditioning with a reindexed version of the predicate p:

$$(\pi_1)_*\Big((\omega \otimes \rho)|_p\Big) = \omega|_{(\mathrm{id} \otimes \rho)^*(p)} \qquad \text{and} \qquad (\pi_2)_*\Big((\omega \otimes \rho)|_p\Big) = \rho|_{(\omega \otimes \mathrm{id})^*(p)}.$$

Proof. (i) Recall from (28) that the conditional state $\omega|_p = \mathrm{nrm}(\mathrm{asrt}_p \bullet \omega)$ satisfies $(\omega|_p + \mathrm{id}) \bullet (\omega \models p) = \mathrm{asrt}_p \bullet \omega$. Hence:

$$\begin{aligned}
(\omega|_p \models q) \cdot (\omega \models p) &= [\omega|_p \models q, \mathbf{0}] \bullet (\omega \models p) && \text{see Definition 2 (iii)} \\
&= [q \bullet \omega|_p, \mathbf{0}] \bullet (\omega \models p) \\
&= [q, \mathbf{0}] \bullet (\omega|_p + \mathrm{id}) \bullet (\omega \models p) \\
&= [q, \mathbf{0}] \bullet \mathrm{asrt}_p \bullet \omega \\
&= (p \,\&\, q) \bullet \omega \\
&= \omega \models (p \,\&\, q).
\end{aligned}$$

(ii) The first equation follows directly from the equation $\mathrm{asrt}_{\mathbf{1}} = \langle \kappa_1 \rangle \colon X \to X+1$ in Lemma 26 (iv). Then $\omega|_{\mathbf{1}} = \mathrm{nrm}(\mathrm{asrt}_\omega \bullet \omega) = \mathrm{nrm}(\langle \kappa_1 \rangle \bullet \omega) = \omega$ by Lemma 32 (i). For the second equation we use $\mathrm{asrt}_q \bullet \mathrm{asrt}_p = \mathrm{asrt}_{p \& q}$, see Proposition 27 (iii), together with the uniqueness of normalisations from (27):

$$\begin{aligned}
&((\omega|_p)|_q + \mathrm{id}) \bullet (\omega \models p \,\&\, q) \\
&\quad = ((\omega|_p)|_q + \mathrm{id}) \bullet \Big((\omega|_p \models q) \cdot (\omega \models p)\Big) && \text{by the previous point} \\
&\quad = ((\omega|_p)|_q + \mathrm{id}) \bullet [\omega|_p \models q, \mathbf{0}] \bullet (\omega \models p) \\
&\quad = [(\omega|_p)|_q + \mathrm{id}) \bullet \omega|_p \models q, \mathbf{0}] \bullet (\omega \models p) \\
&\quad = [\mathrm{asrt}_q \bullet \omega|_p, \mathbf{0}] \bullet (\omega \models p) && \text{by (28)} \\
&\quad = [\mathrm{asrt}_q, \mathbf{0}] \bullet (\omega|_p + \mathrm{id}) \bullet (\omega \models p) \\
&\quad = [\mathrm{asrt}_q, \mathbf{0}] \bullet \mathrm{asrt}_p \bullet \omega && \text{by (28) again} \\
&\quad = \Big(\mathrm{asrt}_q \bullet' \mathrm{asrt}_p\Big) \bullet \omega \\
&\quad = \mathrm{asrt}_{p\&q} \bullet \omega.
\end{aligned}$$

(iii) For a map $f\colon X \to Y$ and a predicate q on Y we have:

$$\begin{aligned}
f \bullet \omega|_{q\circ f} &= T(f) \circ \mathrm{nrm}\Big(\mathrm{asrt}_{q\circ f} \bullet \omega\Big) \\
&= \mathrm{nrm}((\langle f \rangle + \mathrm{id}) \bullet \mathrm{asrt}_{q\circ f} \bullet \omega) && \text{by Lemma 32 (ii)} \\
&= \mathrm{nrm}(\mathrm{asrt}_q \bullet \langle f \rangle \bullet \omega) && \text{by Lemma 26 (vii)} \\
&= (T(f) \circ \omega)|_q.
\end{aligned}$$

(iv) As in Definition 2 (vii) we supress the isomorphism $1 \cong 1 \otimes 1$, $X \cong X \otimes 1$ and $Y \cong 1 \otimes Y$ in writing $\omega \otimes \rho\colon 1 \to X \otimes Y$, $\mathrm{id} \otimes \rho\colon X \to X \otimes Y$ and $\omega \otimes \mathrm{id}\colon Y \to X \otimes Y$ in $\mathcal{K}\ell(T)$. The result then follows from simple equations in $\mathcal{K}\ell(T)$.

$$\begin{aligned}
\omega \models (\mathrm{id} \otimes \rho)^*(p) &= p \bullet (\mathrm{id} \otimes \rho) \bullet \omega \\
&= p \bullet (\omega \otimes \rho) \\
&= (\omega \otimes \rho) \models p \\
&= p \bullet (\omega \otimes \mathrm{id}) \bullet \rho \\
&= \rho \models (\omega \otimes \mathrm{id})^*(p).
\end{aligned}$$

(v) Let's abbreviate as $m\colon (X+1) \times (Y+1) \to (X \times Y) + 1$ the map that is used in Lemma 26 (x) for the definition of the tensor $\otimes'$ in $\mathcal{K}\ell(T')$. We are done by uniqueness if we can show:

$$((\omega|_p \otimes \rho|_q) + \mathrm{id}) \bullet (\omega \otimes \rho \models p \odot q) = \mathrm{asrt}_{p\odot q} \bullet (\omega \otimes \rho).$$

First we use Remark 3 to see:

$$\begin{aligned}
\omega \otimes \rho \models p \odot q &= (p \odot q) \bullet (\omega \otimes \rho) \\
&= \langle m \rangle \bullet (p \otimes q) \bullet (\omega \otimes \rho) \\
&= \langle m \rangle \bullet ((\omega \models p) \otimes (\rho \models q)) \\
&= (\omega \models p) \cdot (\rho \models q).
\end{aligned}$$

We now use the following diagram chase in $\mathcal{K}\ell(T)$.

$$\begin{tikzcd}
1 \arrow[rr, "\omega\otimes\rho"] \arrow[dr, "\omega"'] \arrow[ddd, "\omega\models p"'] & & X \otimes Y \arrow[rr, "\mathrm{asrt}_{p\odot q}"] \arrow[dr, "\mathrm{asrt}_p\otimes\mathrm{asrt}_q"] & & (X \otimes Y) + 1 \\
& X \arrow[ur, "\mathrm{id}\otimes\rho"'] \arrow[dr, "\mathrm{asrt}_p"] & & (X+1)\otimes(Y+1) \arrow[ur, "\langle m \rangle"'] & \\
& & X + 1 \arrow[r, "\mathrm{id}\otimes(\rho\models q)"] & (X+1)\otimes 2 \arrow[u, "\mathrm{id}\otimes(\rho|_q+\mathrm{id})"'] & \\
2 \arrow[urr, "\omega|_p+\mathrm{id}"] \arrow[rrr, "\mathrm{id}\otimes(\rho\models q)"'] & & & 2 \otimes 2 \arrow[u, "(\omega|_p+\mathrm{id})\otimes\mathrm{id}"'] \arrow[r, "\langle m \rangle"'] & 2 \arrow[uuu, "(\omega|_p\otimes\rho|_q)+\mathrm{id}"']
\end{tikzcd}$$

(vi) We apply uniqueness of normalisation to get the required result, via Fubini:

$$\begin{aligned}
&\Big((\pi_1)_*\Big((\omega \otimes \rho)|_p\Big) + \mathrm{id}\Big) \bullet \Big(\omega \models (\mathrm{id} \otimes \rho)^*(p)\Big) \\
&\quad= (\pi_1 + \mathrm{id}) \bullet ((\omega \otimes \rho)|_p + \mathrm{id}) \bullet (\omega \otimes \rho \models p) \\
&\quad\overset{(28)}{=} (\pi_1 + \mathrm{id}) \bullet \mathrm{asrt}_p \bullet (\mathrm{id} \otimes \rho) \bullet \omega \\
&\quad\overset{(*)}{=} \mathrm{asrt}_{(\mathrm{id}\otimes\rho)^*(p)} \bullet \omega \\
&\quad\overset{(28)}{=} \Big(\omega|_{(\mathrm{id}\otimes\rho)^*(p)} + \mathrm{id}\Big) \bullet (\omega \models (\mathrm{id} \otimes \rho)^*(p)).
\end{aligned}$$

The marked equation $\overset{(*)}{=}$ is obtained by unraveling the definition of assert maps. □

At this stage we have reached a level of abstraction where we can use the logic and structure of states and effect (predicates forming effect modules) to reason about probability. This is used for instance in [42] to precisely describe:

- backwards inference $\omega|_{f^*(p)}$ as first pulling back predicate p to $f^*(p)$ via the predicate transformer f^*, and then forming the conditional state;
- forward inference $f_*(\omega|_p)$ as first conditoning and then moving the resulting state forward by the state transformer f_*.

Such abstract descriptions hopefully simplify probabilistic reasoning, see for instance [7,61].

10 Conclusions

This paper describes in step-by-step manner how certain properties of monads lead to the structure of a monoidal commutative effectus. This is a basic categorical universe for probability theory. The approach applies to the standard monads used for probability: $\mathcal{D}$, $\mathcal{G}$, $\mathcal{E}$, $\mathcal{V}$, $\mathcal{R}$, $\mathcal{K}$.

There is ample room for future work. For instance, other properties from effectus theory could be added to the present framework, such as images, comprehension, quotients, see [10], or subcategories of pure maps with daggers [64]. Also, more probability theory can be lifted to the abstract categorical level, where regular conditionals are of immediate interest, see [14]. On a different note, it would be nice to have a characterisation result in the opposite direction of the paper: each monoidal commutative effectus is the Kleisli category of a suitable 'probability' monad. Possibly such a monad can be obtained via the 'codensity' construction [54,6].

Acknowledgements

Thanks are due to Kenta Cho, Robert Furber, Mathys Rennela, Bram Westerbaan, and Fabio Zanasi for helpful discussions on the topic of the paper, and to the anonymous referees of [33,34], and of this paper, for suggesting several improvements.

References

[1] S. Abramsky and A. Brandenburger. The sheaf-theoretic structure of non-locality and contextuality. *New Journ. of Physics*, 13:113036, 2011.

[2] J. Adámek and J. Velebil. Analytic functors and weak pullbacks. *Theory and Appl. of Categories*, 21(11):191–209, 2008.

[3] R. Adams and B. Jacobs. A type theory for probabilistic and Bayesian reasoning. See `arxiv.org/abs/1511.09230`, 2015.

[4] M. Alvarez-Manilla, A. Edalat, and N. Saheb-Djahromi. An extension result for continuous valuations. *Journ. London Math. Soc.*, 61(02):629–640, 2000.

[5] M. Arbib and E. Manes. Partially additive categories and flow-diagram semantics. *Journ. Algebra*, 62(1):203–227, 1980.

[6] T. Avery. Codensity and the Giry monad. *Ann. Pure & Appl. Logic*, 220:1229–1251, 2016.

[7] J. Borgström, A.D. Gordon, M. Greenberg, J. Margetson, and J. Van Gael. Measure transformer semantics for Bayesian machine learning. *Logical Methods in Comp. Sci.*, 9(3):1–39, 2013.

[8] F. van Breugel. The metric monad for probabilistic nondeterminism. Unpublished note, available from `http://www.cse.yorku.ca/~franck/research/drafts/monad.pdf`, 2005.

[9] K. Cho. Total and partial computation in categorical quantum foundations. In C. Heunen, P. Selinger, and J. Vicary, editors, *Quantum Physics and Logic (QPL) 2015*, number 195 in Elect. Proc. in Theor. Comp. Sci., pages 116–135, 2015.

[10] K. Cho, B. Jacobs, A. Westerbaan, and B. Westerbaan. An introduction to effectus theory. see `arxiv.org/abs/1512.05813`, 2015.

[11] R. Cockett. Introduction to distributive categories. *Math. Struct. in Comp. Sci.*, 3:277–307, 1993.

[12] B. Coecke, C. Heunen, and A. Kissinger. Categories of quantum and classical channels. *Quantum Information Processing*, pages 1—-31, 2014.

[13] D. Coumans and B. Jacobs. Scalars, monads and categories. In C. Heunen, M. Sadrzadeh, and E. Grefenstette, editors, *Quantum Physics and Linguistics. A Compositional, Diagrammatic Discourse*, pages 184–216. Oxford Univ. Press, 2013.

[14] J. Culbertson and K. Sturtz. A categorical foundation for bayesian probability. *Appl. Categorical Struct.*, 22(4):647–662, 2014.

[15] E. Dijkstra and C. Scholten. *Predicate Calculus and Program Semantics*. Springer, Berlin, 1990.

[16] A. Dvurečenskij and S. Pulmannová. *New Trends in Quantum Structures*. Kluwer Acad. Publ., Dordrecht, 2000.

[17] A. Edalat. Domain theory and integration. *Theor. Comp. Sci.*, 151:163–193, 1995.

[18] G. Edgar. *Integral, Probability, and Fractal Measures*. Springer-Verlag, New York, 1998.

[19] B. Fong. Causal theories: A categorical perspective on Bayesian networks. Master's thesis, Univ. of Oxford, 2012. see `arxiv.org/abs/1301.6201`.

[20] D. J. Foulis and M.K. Bennett. Effect algebras and unsharp quantum logics. *Found. Physics*, 24(10):1331–1352, 1994.

[21] R. Furber and B. Jacobs. From Kleisli categories to commutative C^*-algebras: Probabilistic Gelfand duality. In R. Heckel and S. Milius, editors, *Conference on Algebra and Coalgebra in Computer Science (CALCO 2013)*, number 8089 in Lect. Notes Comp. Sci., pages 141–157. Springer, Berlin, 2013.

[22] R. Furber and B. Jacobs. From Kleisli categories to commutative C^*-algebras: Probabilistic Gelfand duality. *Logical Methods in Comp. Sci.*, 11(2):1–28, 2015.

[23] R. Furber and B. Jacobs. Towards a categorical account of conditional probability. In C. Heunen, P. Selinger, and J. Vicary, editors, *Quantum Physics and Logic (QPL) 2015*, number 195 in Elect. Proc. in Theor. Comp. Sci., pages 179–195, 2015.

[24] M. Giry. A categorical approach to probability theory. In B. Banaschewski, editor, *Categorical Aspects of Topology and Analysis*, number 915 in Lect. Notes Math., pages 68–85. Springer, Berlin, 1982.

[25] B. Jacobs. Semantics of weakening and contraction. *Ann. Pure & Appl. Logic*, 69(1):73–106, 1994.

[26] B. Jacobs. Convexity, duality, and effects. In C. Calude and V. Sassone, editors, *IFIP Theoretical Computer Science 2010*, number 82(1) in IFIP Adv. in Inf. and Comm. Techn., pages 1–19. Springer, Boston, 2010.

[27] B. Jacobs. From coalgebraic to monoidal traces. In B. Jacobs, M. Niqui, J. Rutten, and A. Silva, editors, *Coalgebraic Methods in Computer Science*, volume 264 of *Elect. Notes in Theor. Comp. Sci.*, pages 125–140. Elsevier, Amsterdam, 2010.

[28] B. Jacobs. Probabilities, distribution monads, and convex categories. *Theor. Comp. Sci.*, 412(28):3323–3336, 2011.

[29] B. Jacobs. Measurable spaces and their effect logic. In *Logic in Computer Science*. IEEE, Computer Science Press, 2013.

[30] B. Jacobs. New directions in categorical logic, for classical, probabilistic and quantum logic. *Logical Methods in Comp. Sci.*, 11(3):1–76, 2015.

[31] B. Jacobs. A recipe for state and effect triangles. Extended version of [32], 2015.

[32] B. Jacobs. A recipe for state and effect triangles. In L. Moss and P. Sobocinski, editors, *Conference on Algebra and Coalgebra in Computer Science (CALCO 2015)*, volume 35 of *LIPIcs*, pages 116–129. Schloss Dagstuhl, 2015.

[33] B. Jacobs. Affine monads and side-effect-freeness. In I. Hasuo, editor, *Coalgebraic Methods in Computer Science (CMCS 2016)*, number 9608 in Lect. Notes Comp. Sci., pages 53–72. Springer, Berlin, 2016.

[34] B. Jacobs. Effectuses from monads. In L. Birkedal, editor, *Math. Found. of Programming Semantics*, number 325 in Elect. Notes in Theor. Comp. Sci., pages 169–183. Elsevier, Amsterdam, 2016.

[35] B. Jacobs. Hyper normalisation and conditioning for discrete probability distributions. See `arxiv.org/abs/1607.02790`, 2016.

[36] B. Jacobs. Introduction to coalgebra. Towards mathematics of states and observations. Cambridge Univ. Press, to appear, 2016.

[37] B. Jacobs and J. Mandemaker. Coreflections in algebraic quantum logic. *Found. of Physics*, 42(7):932–958, 2012.

[38] B. Jacobs and J. Mandemaker. The expectation monad in quantum foundations. In B. Jacobs, P. Selinger, and B. Spitters, editors, *Quantum Physics and Logic (QPL) 2011*, number 95 in Elect. Proc. in Theor. Comp. Sci., pages 143–182, 2012.

[39] B. Jacobs, J. Mandemaker, and R. Furber. The expectation monad in quantum foundations. Inf. & Comp., 2016.

[40] B. Jacobs and A. Westerbaan. An effect-theoretic account of Lebesgue integration. In D. Ghica, editor, *Math. Found. of Programming Semantics*, number 319 in Elect. Notes in Theor. Comp. Sci., pages 239–253. Elsevier, Amsterdam, 2015.

[41] B. Jacobs, B. Westerbaan, and A. Westerbaan. States of convex sets. In A. Pitts, editor, *Foundations of Software Science and Computation Structures*, number 9034 in Lect. Notes Comp. Sci., pages 87–101. Springer, Berlin, 2015.

[42] B. Jacobs and F. Zanasi. A predicate/state transformer semantics for Bayesian learning. In L. Birkedal, editor, *Math. Found. of Programming Semantics*, number 325 in Elect. Notes in Theor. Comp. Sci., pages 185–200. Elsevier, Amsterdam, 2016.

[43] C. Jones. *Probabilistic Non-determinism*. PhD thesis, Edinburgh Univ., 1989.

[44] C. Jones and G. Plotkin. A probabilistic powerdomain of evaluations. In *Logic in Computer Science*, pages 186–195. IEEE, Computer Science Press, 1989.

[45] A. Jung and R. Tix. The troublesome probabilistic powerdomain. In A. Edalat, A. Jung, K. Keimel, and M. Kwiatkowska, editors, *Comprox III, Third Workshop on Computation and Approximation*, number 13 in Elect. Notes in Theor. Comp. Sci., pages 70–91. Elsevier, Amsterdam, 1998.

[46] K. Keimel. The monad of probability measures over compact ordered spaces and its Eilenberg-Moore algebras. *Topology and its Applications*, 156:227–239, 2008.

[47] K. Keimel. Abstract ordered compact convex sets and algebras of the (sub)probabilistic power domain monad over ordered compact spaces. *Algebra an Logic*, 48(5):330–343, 2009.

[48] B. Klin. Structural operational semantics for weighted transition systems. In J. Palsberg, editor, *Semantics and Algebraic Specification. Essays Dedicated to Peter D. Mosses on the Occasion of His 60th Birthday*, number 5700 in Lect. Notes Comp. Sci., pages 121–139. Springer, Berlin, 2009.

[49] A. Kock. Monads on symmetric monoidal closed categories. *Arch. Math.*, XXI:1–10, 1970.

[50] A. Kock. On double dualization monads. *Math. Scand.*, 27:151–165, 1970.

[51] A. Kock. Bilinearity and cartesian closed monads. *Math. Scand.*, 29:161–174, 1971.

[52] A. Kock. Commutative monads as a theory of distributions. *Theory and Appl. of Categories*, 26(4):97–131, 2012.

[53] D. Kozen. Semantics of probabilistic programs. *Journ. Comp. Syst. Sci*, 22(3):328–350, 1981.

[54] T. Leinster. Codensity and the ultrafilter monad. *Theory and Appl. of Categories*, 28(13):332–370, 2013.

[55] H. Lindner. Affine parts of monads. *Arch. Math.*, XXXIII:437–443, 1979.

[56] E. Manes and M. Arbib. *Algebraic Appoaches to Program Semantics*. Texts and Monogr. in Comp. Sci.,. Springer, Berlin, 1986.

[57] M. Mislove. Probabilistic monads, domains and classical information. In E. Kashefi, J. Krivine, and F. van Raamsdonk, editors, *Developments of Computational Methods (DCM 2011)*, number 88 in Elect. Proc. in Theor. Comp. Sci., pages 87–100, 2012.

[58] M. Ozawa. Quantum measuring processes of continuous observables. *Journ. Math. Physics*, 25:79–87, 1984.

[59] P. Panangaden. *Labelled Markov Processes*. Imperial College Press, London, 2009.

[60] W. Rudin. *Functional Analysis*. McGraw-Hill Book Company, 1987. Third, intern. edition.

[61] S. Staton, H. Yang, C. Heunen, O. Kammar, and F. Wood. Semantics for probabilistic programming: higher-order functions, continuous distributions, and soft constraints. Logic in Computer Science, 2016.

[62] R. Tix, K. Keimel, and G. Plotkin. *Semantic Domains for Combining Probability and Non-Determinism*. Number 129 in Elect. Notes in Theor. Comp. Sci. Elsevier, Amsterdam, 2005.

[63] V. Trnková. Some properties of set functors. *Comment. Math. Univ. Carolinae*, 10:323–352, 1969.

[64] A. Westerbaan and B. Westerbaan. Paschke dilations. QPL 2016, 2016.