Affine Monads and Side-Effect-Freeness

Bart Jacobs

HAL Id: hal-01446033

https://inria.hal.science/hal-01446033

Submitted on 25 Jan 2017

Affine Monads and Side-Effect-Freeness

Bart Jacobs[⋆]

Institute for Computing and Information Sciences,
Radboud Universiteit, Nijmegen, The Netherlands
`bart@cs.ru.nl`

Abstract. The notions of side-effect-freeness and commutativity are typical for probabilistic models, as subclass of quantum models. This paper connects these notions to properties in the theory of monads. A new property of a monad ('strongly affine') is introduced. It is shown that for such strongly affine monads predicates are in bijective correspondence with side-effect-free instruments. Also it is shown that these instruments are commutative, in a suitable sense, for monads which are commutative (monoidal).

1 Introduction

In a recent line of work in categorical quantum foundations [6,13,16,5,3,4] the notion of effectus has been proposed. Within that context one associates an instrument with each predicate, which performs measurement. These instruments are coalgebras, of a particular form, which may change the state. Indeed, it is one of the key features of the quantum world that measurement can change the object under observation. Thus, observation may have a side-effect.

In [6] a subclass of *commutative* effectuses is defined where there is a one-to-one correspondence between predicates and side-effect-free instruments. These commutative effectuses capture the probabilistic models, as special case of quantum models. Examples of commutative effectuses are the Kleisli categories $\mathcal{K}\ell(\mathcal{D})$ and $\mathcal{K}\ell(\mathcal{G})$ of the distribution monad $\mathcal{D}$ and the Giry monad $\mathcal{G}$, and the category of commutative von Neumann algebras.

The starting point for the work presented here is: can we translate these notions of side-effect-freeness and commutativity from effectus theory to the theory of monads — and coalgebras of monads — since they are instrumental in the semantics of programming languages? Especially, is there a connection between:

1. side-effect-freeness of measurment-instruments and the property that a monad is affine (that is, preserves the final object);
2. commutativity as in effectus theory and commutativity of a monad?

The main point of the paper is that these questions can be answered positively.

The first question makes sense because both the distribution and the Giry monad are affine, and it seems that this property is typical for monads that are relevant in probability theory. We shall see below that we actually need a slightly stronger property than 'affine', namely what we call 'strongly affine'.

Given the terminological coincidence, the second question 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 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 Theorem 1 below). It is the basis for a novel logic and type theory for probabilism in [4] (see also [17]).

The background of this work is effectus theory [6] in which logic (in terms of effect modules) and instruments play an important role. Here we concentrate on these instruments, and show that they can be studied in the theory of monads, independent of the logic of effect modules. Including these effect modules in the theory (for special monads) is part of ongoing and future work [14].

2 Preliminaries

We assume that the reader is familiar with the notion of monad. We recall that a monad $T = (T, \eta, \mu)$ on a category with finite products $(\times, 1)$ 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.

$$
\begin{array}{ccc}
T(X) \times Y \xrightarrow{\ \mathrm{st}_1\ } T(X \times Y) & \quad & (T(X) \times Y) \times Z \xrightarrow{\ \cong\ } T(X) \times (Y \times Z) \\
\end{array}
\tag{1}
$$

$$
\begin{array}{cc}
X \times Y =\!=\!= X \times Y & \quad T^2(X) \times Y \xrightarrow{\ \mathrm{st}_1\ } T(T(X) \times Y) \xrightarrow{\ T(\mathrm{st}_1)\ } T^2(X \times Y) \\
\eta \times \mathrm{id} \downarrow \qquad \downarrow \eta & \quad \mu \times \mathrm{id} \downarrow \qquad\qquad\qquad\qquad\qquad \downarrow \mu \\
T(X) \times Y \xrightarrow[\ \mathrm{st}_1\ ]{} T(X \times Y) & \quad T(X) \times Y \xrightarrow{\qquad\qquad \mathrm{st}_1 \qquad\qquad} T(X \times Y)
\end{array}
\tag{2}
$$

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:

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

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

The monad T is called *commutative* (following [19]) when the order of applying strength in two coordinates does not matter, as expressed by commutation of the following diagram.

$$
\begin{array}{ccc}
& T(X \times T(Y)) \xrightarrow{T(\mathrm{st}_2)} T^2(X \times Y) \searrow^{\mu} & \\
T(X) \times T(Y) \nearrow^{\mathrm{st}_1} \searrow_{\mathrm{st}_2} & & T(X \times Y) \quad (3) \\
& T(T(X) \times Y) \xrightarrow[T(\mathrm{st}_1)]{} T^2(X \times Y) \nearrow_{\mu} &
\end{array}
$$

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. Notice that $\mathrm{dst} \circ \gamma = T(\gamma) \circ \mathrm{dst}$.

Below we shall use distributive categories. They have finite products $(\times, 1)$ and coproducts $(+, 0)$, where products distribute over coproducts, in the sense that the following maps are isomorphisms.

$$
0 \xrightarrow{\ !\ } 0 \times X \qquad (A \times X) + (B \times X) \xrightarrow{\mathrm{dis}_1 = [\kappa_1 \times \mathrm{id}, \kappa_2 \times \mathrm{id}]} (A + B) \times X
$$
$$(4)$$

Swapping yields an associated distributivity map:

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

It is an easy exercise to show that dis_1 and dis_2 interact in the following way.

$$
\begin{array}{ccc}
\big((A{\times}X){+}(B{\times}X)\big){+}\big((A{\times}Y){+}(B{\times}Y)\big) & \xrightarrow{\mathrm{dis}_1 + \mathrm{dis}_1} & ((A{+}B){\times}X){+}((A{+}B){\times}Y) \\
\big\uparrow{\scriptstyle [\kappa_1 + \kappa_1, \kappa_2 + \kappa_2]}\ \cong & & \big\downarrow{\scriptstyle \mathrm{dis}_2} \\
& & (A{+}B){\times}(X{+}Y) \qquad (5) \\
& & \big\uparrow{\scriptstyle \mathrm{dis}_1} \\
\big((A{\times}X){+}(A{\times}Y)\big){+}\big((B{\times}X){+}(B{\times}Y)\big) & \xrightarrow[\mathrm{dis}_2 + \mathrm{dis}_2]{} & (A{\times}(X{+}Y)){+}(B{\times}(X{+}Y))
\end{array}
$$

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

$$
\begin{array}{ccc}
A \times T(X) + B \times T(X) & \xrightarrow{\mathrm{dis}_1} & (A + B) \times T(X) \\
{\scriptstyle \mathrm{st}_2 + \mathrm{st}_2}\big\downarrow & & \big\downarrow \\
T(A \times X) + T(B \times X) & & \Big\downarrow{\scriptstyle \mathrm{st}_2} \qquad (6) \\
{\scriptstyle [T(\kappa_1), T(\kappa_2)]}\big\downarrow & & \big\downarrow \\
T((A \times X) + (B \times X)) & \xrightarrow{T(\mathrm{dis}_1)} & T((A + B) \times X)
\end{array}
$$

The object $2 = 1+1$ will play a special role below. In a distributive category we have two 'separation' isomorphisms written as:

$$2 \times X \xrightarrow[\cong]{\text{sep}_1} X + X \xleftarrow[\cong]{\text{sep}_2} X \times 2 \qquad (7)$$

Explicitly, they are defined as:

$$2 \times X \xrightarrow{\text{sep}_1} X + X \xleftarrow{\text{sep}_2} X \times 2$$

with dis_1^{-1}, $1 \times X + 1 \times X$, $\pi_2 + \pi_2$, $\pi_1 + \pi_1$, $X \times 1 + X \times 1$, dis_2^{-1}

These separation maps are natural in X and satisfy for instance:

$$
\begin{aligned}
\nabla \circ \text{sep}_1 &= \pi_2 & \nabla \circ \text{sep}_2 &= \pi_1 \\
(!+!) \circ \text{sep}_1 &= \pi_1 & (!+!) \circ \text{sep}_2 &= \pi_2 \\
[\kappa_2, \kappa_1] \circ \text{sep}_1 &= \text{sep}_1 \circ ([\kappa_2, \kappa_1] \times \text{id}) & [\kappa_2, \kappa_1] \circ \text{sep}_1 &= \text{sep}_2 \circ (\text{id} \times [\kappa_2, \kappa_1])
\end{aligned}
\qquad (8)
$$

These two maps are related via: $\text{sep}_1 \circ \gamma = \text{sep}_2$, for $\gamma = \langle \pi_2, \pi_1 \rangle$.

In the special case where $X = 2$ we have inverses

$$2 \times 2 \xrightarrow[{[[\langle \kappa_1, \kappa_1 \rangle, \langle \kappa_1, \kappa_2 \rangle], [\langle \kappa_2, \kappa_1 \rangle, \langle \kappa_2, \kappa_2 \rangle]]]}]{\text{sep}_1 \;\cong\;} 2 + 2 \xleftarrow[{[[\langle \kappa_1, \kappa_1 \rangle, \langle \kappa_2, \kappa_1 \rangle], [\langle \kappa_1, \kappa_2 \rangle, \langle \kappa_2, \kappa_2 \rangle]]]}]{\text{sep}_2 \;\cong\;} 2 \times 2$$

It is not hard to see that:

$$\text{sep}_1 \circ \gamma = \text{sep}_2 = [\kappa_1 + \kappa_1, \kappa_2 + \kappa_2] \circ \text{sep}_1 \qquad (9)$$

The isomorphism $[\kappa_1 + \kappa_1, \kappa_2 + \kappa_2]\colon 2 + 2 \xrightarrow{\cong} 2 + 2$ can be illustrated as:

$$
\begin{array}{ccccccc}
2 + 2 & = & (1 & + & 1) & + & (1 & + & 1) \\
 & & & \downarrow & & \times & & \downarrow & \\
2 + 2 & = & (1 & + & 1) & + & (1 & + & 1)
\end{array}
$$

3 Affine and strongly affine monads

In this section we recall what it means for a monad to be affine (see [20,21,11]), and introduce a slightly stronger notion. We describe basic properties and examples.

Definition 1. *Let* $\mathbf{C}$ *be a category with a monad* $T\colon \mathbf{C} \to \mathbf{C}$.

1. *Assuming that* $\mathbf{C}$ *has a final object* 1, *one calls* T affine *if the map* $T(1) \to 1$ *is an isomorphism, or simply, if* $T(1) \cong 1$.

2. Assuming that $\mathbf{C}$ *has binary products* $\times$ *and* T *is a strong monad, we call* T *strongly affine* *if the squares below are pullbacks.*

$$
\begin{array}{ccc}
T(X) \times Y & \xrightarrow{\ \pi_2\ } & Y \\
{\scriptstyle \mathrm{st}_1}\downarrow & & \downarrow{\scriptstyle \eta_Y} \\
T(X \times Y) & \xrightarrow[T(\pi_2)]{} & T(Y)
\end{array}
\tag{10}
$$

The notion of an 'affine monad' is well-known. What we call 'strongly affine' is new. The relationship with ordinary affine monads is quite subtle. Example 2 below show that 'strongly affine' is really stronger than 'affine'. But first we describe some properties and examples.

Lemma 1. *Let* T *be a strong monad on a category* $\mathbf{C}$ *with finite products* $(\times, 1)$.

1. *The monad* T *is affine iff the diagrams* (10) *commute. As a result, a strongly affine monad is affine.*
2. *There is at most one mediating (pullback) map for the diagram* (10).

The first point gives an alternative formulation of affiness. An older alternative formulation is: $\langle T(\pi_1), T(\pi_2)\rangle \circ \mathrm{dst} = \mathrm{id}$, see [20, Thm. 2.1], where dst is the double strength map from (3), for a commutative monad T.

The second point 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 1.

Proof. For the first point, let T be affine. We stretch Diagram (10) as follows.

$$
\begin{array}{ccccc}
 & & \xrightarrow{\ \ \pi_2\ \ } & & \\
T(X) \times Y & \xrightarrow[T(!)\times Y]{} & T(1) \times Y & \xrightarrow[\pi_2]{\ \cong\ } & Y \\
{\scriptstyle \mathrm{st}_1}\downarrow & & {\scriptstyle \mathrm{st}_1}\downarrow & & \downarrow{\scriptstyle \eta_Y} \\
T(X \times Y) & \xrightarrow[T(!\times \mathrm{id})]{} & T(1 \times Y) & \xrightarrow[\cong]{T(\pi_2)} & T(Y) \\
 & & \xrightarrow{\ \ T(\pi_2)\ \ } & &
\end{array}
$$

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:

$$
\begin{aligned}
T(\pi_2) \circ \mathrm{st}_1 &= T(\pi_2) \circ \mathrm{st}_1 \circ \langle \eta_1 \circ !_Y, \mathrm{id}\rangle \circ \pi_2 \\
&= T(\pi_2) \circ \mathrm{st}_1 \circ (\eta_1 \times \mathrm{id}) \circ \langle !_Y, \mathrm{id}\rangle \circ \pi_2 \\
&\overset{(2)}{=} T(\pi_2) \circ \eta_{1\times Y} \circ \langle !_Y, \mathrm{id}\rangle \circ \pi_2 \\
&= \eta_Y \circ \pi_2 \circ \langle !_Y, \mathrm{id}\rangle \circ \pi_2 \\
&= \eta_Y \circ \pi_2.
\end{aligned}
$$

In the other direction, assume that diagrams (10) commute. We consider the special case $X = Y = 1$.

$$
\begin{array}{ccc}
T(1) \times 1 & \xrightarrow{\;\;\pi_2\;\;} & 1 \\
{\scriptstyle \mathrm{st}_1}\downarrow & \searrow^{\pi_1} & \downarrow{\scriptstyle \eta_1} \\
T(1 \times 1) & \xrightarrow[\;T(\pi_2)=T(\pi_1)\;]{\cong} & T(1)
\end{array}
$$

The lower triangle commutes by (1). We need to prove that $T(1)$ is final. It suffices to prove that the composite $\eta_1 \circ \,!\colon T(1) \to 1 \to T(1)$ is the identity. This is obtained from the upper triangle:

$$
\eta_1 \circ \,! \;=\; \eta_1 \circ \pi_2 \circ \langle \mathrm{id}, !\rangle \;=\; \pi_1 \circ \langle \mathrm{id}, !\rangle \;=\; \mathrm{id}.
$$

For the second point in the lemma 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:

$$
\pi_1 \circ f \;\overset{(1)}{=}\; T(\pi_1) \circ \mathrm{st}_1 \circ f \;=\; T(\pi_1) \circ \mathrm{st}_1 \circ g \;\overset{(1)}{=}\; \pi_1 \circ g. \qquad \square
$$

Example 1. Three examples of affine monads are the distribution monad $\mathcal{D}$ on **Sets** for discrete probability, the Giry monad $\mathcal{G}$ on the category **Meas** of measurable spaces, for continuous probability, and the expectation monad $\mathcal{E}$ on **Sets**. We show that all of them are strongly affine.

(1) The elements of $\mathcal{D}(X)$ are the finite formal convex combinations $\sum_i r_i |x_i\rangle$ with elements $x_i \in X$ and probabilities $r_i \in [0,1]$ satisfying $\sum_i r_i = 1$. We can identify such a convex sum with a function $\varphi\colon X \to [0,1]$ whose support $\mathrm{supp}(\varphi) = \{x \mid \varphi(x) \neq 0\}$ is finite and satisfies $\sum_x \varphi(x) = 1$. We can thus write $\varphi = \sum_x \varphi(x)|x\rangle$.

We have $\mathcal{D}(1) \cong 1$, since the sole element of $\mathcal{D}(1)$ is the distribution $1|*\rangle$, where we write $*$ for the element of the singleton set $1 = \{*\}$.

We show that this monad is also strongly affine. So let in Diagram (10) $\varphi \in \mathcal{D}(X \times Y)$ be a given distribution with $\mathcal{D}(\pi_2)(\varphi) = 1|z\rangle$ for a given element $z \in Y$. Let's write $\varphi = \sum_{x,y} \varphi(x,y)|x,y\rangle$, so that $\mathcal{D}(\pi_2)(\varphi)$ is the marginal distribution:
$$
\mathcal{D}(\pi_2)(\varphi) \;=\; \sum_y \big(\textstyle\sum_x \varphi(x,y)\big)|y\rangle.
$$

If this is the trivial distribution $1|z\rangle$, then $\varphi(x,y) = 0$ for all x and $y \neq z$. We obtain a new distribution $\psi = \mathcal{D}(\pi_1)(\varphi) \in \mathcal{D}(X)$, which takes the simple form $\psi(x) = \varphi(x,z)$. The pair $(\psi, z) \in \mathcal{D}(X) \times Y$ is the unique element giving us the pullback (10), since:

$$
\mathrm{st}_1(\psi, z) \;=\; \sum_x \psi(x)|x,z\rangle \;=\; \sum_x \varphi(x,z)|x,z\rangle \;=\; \sum_{x,y} \varphi(x,y)|x,y\rangle \;=\; \psi.
$$

(2) Next we consider the Giry monad $\mathcal{G}$ on the category **Meas** of measurable spaces. The elements of $\mathcal{G}(X)$ are probability measures $\omega\colon \Sigma_X \to [0,1]$. The

unit $\eta \colon X \to \mathcal{G}(X)$ is given by $\eta(x)(M) = 1$ 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 $M \times N \mapsto \omega(M) \cdot \eta(y)(N)$, see also [10,22,12].

So let's consider the situation (10) for $T = \mathcal{G}$, with a joint probability measure $\omega \in \mathcal{G}(X \times Y)$ and an element $z \in Y$ with

$$\mathcal{G}(\pi_2)(\omega)(N) = \omega(X \times N) = \eta(z)(N), \tag{11}$$

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

$$\omega(M \times N) = \omega(M \times Y) \cdot \omega(X \times N), \tag{12}$$

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

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

$$\omega(M \times N) \leq \omega(X \times N) \overset{(11)}{=} \eta(z)(N) = 0.$$

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

$$\omega(M \times Y) \cdot \omega(X \times N) \overset{(11)}{=} \omega(M \times Y) \cdot \eta(z)(N) = \omega(M \times Y) \cdot 0 = 0.$$

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

$$
\begin{aligned}
\omega(M \times N) &= \omega(M \times N) + 0 \\
&= \omega(M \times N) + \omega(M \times \neg N) &&\text{as just shown} \\
&= \omega\big((M \times N) \cup (M \times \neg N)\big) &&\text{by additivity} \\
&= \omega(M \times Y) \\
&= \omega(M \times Y) \cdot \eta(z)(N) \\
&\overset{(11)}{=} \omega(M \times Y) \cdot \omega(X \times N).
\end{aligned}
$$

We now take $\phi \in \mathcal{G}(X)$ defined by $\phi(M) = \mathcal{G}(\pi_1)(\omega)(M) = \omega(M \times Y)$. The pair $(\phi, z) \in \mathcal{G}(X) \times Y$ is then mediating in (10):

$$
\begin{aligned}
\mathrm{st}_1(\phi, z)(M \times N) = \phi(M) \cdot \eta(z)(N) &= \omega(M \times Y) \cdot \eta(z)(N) \\
&\overset{(11)}{=} \omega(M \times Y) \cdot \omega(X \times N) \\
&\overset{(12)}{=} \omega(M \times N).
\end{aligned}
$$

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

(3) We turn to the expectation monad $\mathcal{E}(X) = \mathbf{EMod}([0,1]^X, [0,1])$ on $\mathbf{Sets}$, where $\mathbf{EMod}$ is the category of effect modules, see [15] for details. Let $h \in \mathcal{E}(X \times Y)$ satisfy $\mathcal{E}(\pi_2)(h) = \eta(z)$, for some $z \in Y$. This means that for each predicate $q \in [0,1]^Y$ we have $h(q \circ \pi_2) = q(z)$.

Our first aim is to prove the analogue of the non-entwinedness equation (12) for $\mathcal{E}$, namely:

$$h(\mathbf{1}_{U \times V}) = h(\mathbf{1}_{U \times Y}) \cdot h(\mathbf{1}_{X \times V}), \tag{13}$$

for arbitrary subsets $U \subseteq X$ and $V \subseteq Y$. Here we write $\mathbf{1}_{U \times V} \colon X \times Y \to [0, 1]$ for the obvious indicator function. We distinguish:

- if $z \notin V$, then $h(\mathbf{1}_{U \times V}) \leq h(\mathbf{1}_{X \times V}) = h(\mathbf{1}_V \circ \pi_2) = \mathbf{1}_V(z) = 0$. Hence (13) holds since both sides are 0.
- if $z \in V$, then $h(\mathbf{1}_{U \times V}) = h(\mathbf{1}_{U \times V}) + h(\mathbf{1}_{U \times \neg V}) = h(\mathbf{1}_{U \times Y}) = h(\mathbf{1}_{U \times Y}) \cdot h(\mathbf{1}_{X \times V})$.

By [15, Lemma 12] each predicate can be written as limit of step functions. It suffices to prove the result for such step functions, since by [15, Lemma 10] the map of effect modules h 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 $h(p) = \mathrm{st}_1\big(\mathcal{E}(\pi_1)(h), z\big)(p)$, so that we can take $\mathcal{E}(\pi_1)(h) \in \mathcal{E}(X)$ to obtain a pullback in (10).

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:

$$
\begin{aligned}
h(p) \;=\; h\big(\textstyle\sum_{i,j} r_{i,j} \mathbf{1}_{U_i \times V_j}\big) \;&=\; \textstyle\sum_{i,j} r_{i,j} h\big(\mathbf{1}_{U_i \times V_j}\big) \\
&\overset{(13)}{=}\; \textstyle\sum_{i,j} r_{i,j} h\big(\mathbf{1}_{U_i \times Y}\big) \cdot h\big(\mathbf{1}_{X \times V_j}\big) \\
&=\; \textstyle\sum_{i,j} r_{i,j} h\big(\mathbf{1}_{U_i \times Y}\big) \cdot \mathbf{1}_{V_j}(z) \\
&=\; \textstyle\sum_i r_{i,j_0} h\big(\mathbf{1}_{U_i \times Y}\big) \\
&=\; h\big(\textstyle\sum_i r_{i,j_0} \mathbf{1}_{U_i \times Y}\big) \\
&=\; h\big(\lambda(x, y). \, p(x, z)\big) \\
&=\; \mathrm{st}_1\big(\mathcal{E}(\pi_1)(h), z\big)(p).
\end{aligned}
$$

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

Example 2. 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 (10) with $X = \{x_1, x_2\}$ and $Y = \{y_1, y_2\}$. Define:

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

We have $\mathcal{D}_{\pm}(\pi_2)(\varphi) = 1|y_1\rangle = \eta(y_1)$, since the terms with y_2 cancel each other out. But there is no element $\psi \in \mathcal{D}_{\pm}(X)$ such that $\mathrm{st}_1(\psi, y_1) = \varphi$. Hence the square (10) is not a pullback.

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.

For the record we recall from [23,2] that each endofunctor on **Sets** can be written as a coproduct of affine functors.

The following result gives a 'graph' construction that is useful in conditional constructions in probability, see the subsequent discussion.

Proposition 1. *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. The if-part of the statement is obvious, since the correspondence is a reformulation of the pullback property of the diagram 10. In the other direction, let T be strongly affine. As stated, the mapping downwards is given by $\overline{f} = \mathrm{st}_1 \circ \langle f, \mathrm{id} \rangle$. Then:

$$T(\pi_2) \circ \overline{f} \;=\; T(\pi_2) \circ \mathrm{st}_1 \circ \langle f, \mathrm{id} \rangle \;\overset{(10)}{=}\; \eta \circ \pi_2 \circ \langle f, \mathrm{id} \rangle \;=\; \eta.$$

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

$$\overline{\overline{f}} \;=\; T(\pi_1) \circ \mathrm{st}_1 \circ \langle f, \mathrm{id} \rangle \;\overset{(1)}{=}\; \pi_1 \circ \langle f, \mathrm{id} \rangle \;=\; f.$$

In order to prove $\overline{\overline{g}} = g$ we notice that by the pullback property of diagram (10) 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:

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

The correspondence in this proposition is used (for the distribution monad $\mathcal{D}$) as Lemma 1 in [9]. 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 [8, §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 $\mathcal{G}(X \times Y)$.

4 Affine parts of monads, and causal maps

It is known for a long time that the 'affine part' of a monad can be extracted via pullbacks, see [21] (or also [11]). Here we shall relate this affine part to 'causal' maps in Kleisli categories of monads.

Proposition 2. *Let T be a monad on a category $\mathbf{C}$ with a final object 1. Assume that the following pullbacks exist in $\mathbf{C}$, for each object X.*

$$
\begin{array}{ccc}
T_a(X) & \xrightarrow{\;\;!\;\;} & 1 \\
{\scriptstyle \iota_X}\downarrow & & \downarrow{\scriptstyle \eta} \\
T(X) & \xrightarrow[T(!)]{} & T(1)
\end{array}
\qquad (14)
$$

Then:

1. *the mapping $X \mapsto T_a(X)$ is a monad on $\mathbf{C}$;*
2. *the mappings $\iota_X\colon T_a(X) \to T(X)$ are monic, and form a map of monads $T_a \Rightarrow T$;*
3. *T_a is an affine monad, and in fact the universal (greatest) affine submonad of T;*
4. *if T is a strong resp. commutative monad, then so is T_a.*

Proof. These results are standard. We shall illustrate point (3). If we take $X = 1$ in Diagram (14), then the bottom arrow $T(!_X)\colon T(X) \to T(1)$ 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 $\bar{\sigma}_X$ in:

The outer diagram commutes since S is affine, so that $\eta_1^S \circ {!}_{S(1)} = \mathrm{id}_{S(1)}$; then:

$$
T(!_X) \circ \sigma_X = \sigma_1 \circ S(!) = \sigma_1 \circ \eta_1^S \circ {!}_{S(1)} \circ S(!_X) = \eta_1^T \circ {!}_{S(X)}. \qquad \square
$$

Example 3. We list several examples of affine parts of monads.

1. 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 $1|*\rangle = \mathcal{M}(!)(\sum_i r_i |x_i\rangle) = (\sum_i r_i)|*\rangle$ iff $\sum_i r_i = 1$. Thus $\mathcal{D}(X) = \mathcal{M}_a(X)$ yields a pullback in Diagram (14).
   The monad $\mathcal{D}_\pm$ used in Example 2 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.

2. 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:

$$\mathcal{P}(!)(U) = \begin{cases} \{*\} & \text{if } U \neq \emptyset \\ \emptyset & \text{if } U = \emptyset \end{cases}$$

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

3. Let $T(X) = (S \times X)^S$ be the state monad on **Sets**, for a fixed set of states S. The unit $\eta \colon X \to T(X)$ is defined as $\eta(x) = \lambda s \in S.\,(s, x)$ so that the pullback (14) 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.

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 affine submonad.

4. 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 (14) 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 $\lambda x.\,c \colon X \to C$.

We write $\mathcal{K\!\ell}(T)$ for the Kleisli category of a monad T, and we write a fat bullet $\bullet$ for Kleisli composition $g \bullet f = \mu \circ T(g) \circ f$. For each object X there is a special 'ground' map:

$$\bar{\top}_X = \left(X \xrightarrow{\ !_X\ } 1 \xrightarrow{\ \eta_1\ } T(1) \right) \tag{15}$$

This is the result of applying the standard functor $\mathbf{C} \to \mathcal{K\!\ell}(T)$ to the map $! \colon X \to 1$ in the underlying category $\mathbf{C}$.

Causal maps have been introduced in the context of CP*-categories, see [7], where they express that measurements in the future, given by $\bar{\top}$, cannot influence the past.

Definition 2. *A Kleisli map $f\colon X \to T(Y)$ will be called* causal *or* unital *if it preserves ground, in the sense that:*

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

Causal maps are used in [6] to construct effectuses. Here we define them quite generally, for an arbitrary monad. Notice that each map $f\colon X \to T(Y)$ is causal when T is an affine monad. The following elementary observation gives a more precise description.

Lemma 2. *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 $\iota_Y \circ f' = f$.*

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

$$
\begin{array}{ccc}
X & \xrightarrow{\quad ! \quad} & \\
& \searrow f' & \\
& T_a(Y) \xrightarrow{\;!\;} 1 & \\
f\;\Big\downarrow & \iota_Y\Big\downarrow & \Big\downarrow \eta \\
& T(Y) \xrightarrow[T(!)]{} T(1) &
\end{array}
\qquad\qquad \square
$$

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 3 (1). This gives a more systematic approach than the "constrained" description from [18], which restricts multisets to a certain subset.

5 Predicates and instruments

In a very general sense we can define a *predicate* on an object X in the Kleisli category $\mathcal{K\ell}(T)$ of a monad T as a map $p\colon X \to 2$, where $2 = 1 + 1$, that is as a map $p\colon X \to T(2)$ in the underlying category. There is always a 'truth' predicate $\mathbf{1} = T(\kappa_1) \circ \bar{\top} = \eta \circ \kappa_1 \circ\,!\colon X \to 1 \to 1 + 1 \to T(1 + 1)$. Similarly there is falsity predicate $\mathbf{0} = \eta \circ \kappa_2 \circ\,!$, and a negation operation $p^\perp = T([\kappa_2, \kappa_1]) \circ p$ obtained by swapping. Clearly, $p^{\perp\perp} = p$ and $\mathbf{1}^\perp = \mathbf{0}$. In certain cases there is more algebraic structure, see [6], where predicates form effect modules.

At this stage we informally describe an *instrument* associated with a predicate $p\colon X \to T(1+1)$ as a map $\mathrm{instr}_p\colon X \to T(X+X)$ with $T(!+!) \circ \mathrm{instr}_p = p$. Such an instrument is called *side-effect-free* if the following diagram commutes in $\mathcal{K\ell}(T)$.

$$
\begin{array}{ccc}
X & \xrightarrow{\;\mathrm{instr}_p\;} & X + X \\
& \searrow\!\!\!\Big\| & \Big\downarrow \nabla=[\mathrm{id},\mathrm{id}] \\
& & X
\end{array}
$$

Equivalently, if $T(\nabla) \circ \mathrm{instr}_p = \eta$ in the underlying category.

This instrument terminology comes from [13] (see also [6]), 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 conditional statement as composite:

$$\text{if } p \text{ then } f \text{ else } g = \left(X \xrightarrow{\mathrm{instr}_p} X + X \xrightarrow{[f,g]} Y \right).$$

For example, for the distribution monad $\mathcal{D}$ a predicate on a set X is a function $p\colon X \to \mathcal{D}(1+1) \cong [0,1]$. For such a 'fuzzy' predicate there is an instrument map $\mathrm{instr}_p\colon X \to \mathcal{D}(X+X)$ given by the convex sum:

$$\mathrm{instr}_p(x) \;=\; p(x)|\kappa_1 x\rangle + (1 - p(x))|\kappa_2 x\rangle.$$

The associated if-then-else statement gives a weighted combination of the two options, where the weights are determined by the probability $p(x) \in [0,1]$.

Next we describe how such instruments can be obtained via a general construction in distributive categories.

Definition 3. *Let T be a strong monad on a distributive category $\mathbf{C}$. For a predicate $p\colon X \to T(1+1)$ we define an instrument $\mathrm{instr}_p\colon X \to T(X+X)$ as composite:*

$$\mathit{instr}_p = \left(X \xrightarrow{\langle p, \mathrm{id}\rangle} T(2) \times X \xrightarrow{\mathrm{st}_1} T(2 \times X) \xrightarrow[\cong]{T(\mathrm{sep}_1)} T(X+X) \right)$$

where sep_1 is the separation isomorphism from (7).

We collect some basic results about instrument maps.

Lemma 3. *In the context of the previous definition we have:*

1. *$T(! + !) \circ \mathit{instr}_p = p$; in particular, $\mathit{instr}_p = p$ for each $p\colon 1 \to T(2)$;*
2. *if p is causal, then instr_p is side-effect-free and causal;*
3. *$\mathit{instr}_{\mathbf{1}} = \eta \circ \kappa_1$ and $\mathit{instr}_{\mathbf{0}} = \eta \circ \kappa_2$, and $\mathit{instr}_{p^\perp} = T([\kappa_2, \kappa_1]) \circ \mathit{instr}_p$;*
4. *for a map $f\colon Y \to X$ in the underlying category,*

$$T(f + f) \circ \mathit{instr}_{p \circ f} \;=\; \mathit{instr}_p \circ f.$$

5. *for predicates $p\colon X \to T(2)$ and $q\colon Y \to T(2)$,*

$$\mathit{instr}_{[p,q]} \;=\; [T(\kappa_1 + \kappa_1), T(\kappa_2 + \kappa_2)] \circ (\mathit{instr}_p + \mathit{instr}_q).$$

Proof. We handle these points one by one.

1. We have:

$$
\begin{aligned}
T(! + !) \circ \mathrm{instr}_p &= T(! + !) \circ T(\mathrm{sep}_1) \circ \mathrm{st}_1 \circ \langle p, \mathrm{id}\rangle \\
&\overset{(8)}{=} T(\pi_1) \circ \mathrm{st}_1 \circ \langle p, \mathrm{id}\rangle \\
&\overset{(1)}{=} \pi_1 \circ \langle p, \mathrm{id}\rangle \\
&= p.
\end{aligned}
$$

2. Assume that the predicate p is causal, that is $T(!) \circ p = \bar{\top}$. We first show that the instrument instr_p is side-effect-free:

$$
\begin{aligned}
T(\nabla) \circ \text{instr}_p &= T(\nabla) \circ T(\text{sep}_1) \circ \text{st}_1 \circ \langle p, \text{id} \rangle \\
&= T(\nabla) \circ T(\pi_2 + \pi_2) \circ T(\text{dis}_1^{-1}) \circ \text{st}_1 \circ \langle p, \text{id} \rangle \\
&= T(\pi_2) \circ T(\nabla) \circ T(\text{dis}_1^{-1}) \circ \text{st}_1 \circ \langle p, \text{id} \rangle \\
&= T(\pi_2) \circ T(\nabla \times \text{id}) \circ \text{st}_1 \circ \langle p, \text{id} \rangle \\
&= T(\pi_2) \circ \text{st}_1 \circ (T(\nabla) \times \text{id}) \circ \langle p, \text{id} \rangle \\
&= T(\pi_2) \circ \text{st}_1 \circ \langle T(!) \circ p, \text{id} \rangle \\
&= T(\pi_2) \circ \text{st}_1 \circ \langle \eta \circ !, \text{id} \rangle \qquad \text{since } p \text{ is causal} \\
&\overset{(2)}{=} T(\pi_2) \circ \eta \circ \langle !, \text{id} \rangle \\
&= \eta \circ \pi_2 \circ \langle !, \text{id} \rangle \\
&= \eta.
\end{aligned}
$$

The instrument instr_p is causal too:

$$
\begin{aligned}
\bar{\top} \bullet \text{instr}_p &= T(!) \circ \text{instr}_p \\
&= T(!) \circ T(\text{sep}_1) \circ \text{st}_1 \circ \langle p, \text{id} \rangle \\
&= T(!) \circ T(\pi_1) \circ \text{st}_1 \circ \langle p, \text{id} \rangle \\
&\overset{(1)}{=} T(!) \circ \pi_1 \circ \langle p, \text{id} \rangle \\
&= T(!) \circ p \\
&= \bar{\top}.
\end{aligned}
$$

3. For the truth predicate $\mathbf{1} = \eta \circ \kappa_1 \circ {!}$ we have:

$$
\begin{aligned}
\text{instr}_\mathbf{1} &= T(\text{sep}_1) \circ \text{st}_1 \circ \langle \eta \circ \kappa_1 \circ !, \text{id} \rangle \\
&\overset{(2)}{=} T(\text{sep}_1) \circ \eta \circ \langle \kappa_1 \circ !, \text{id} \rangle \\
&= \eta \circ \text{sep}_1 \circ \langle \kappa_1 \circ !, \text{id} \rangle \\
&= \eta \circ \kappa_1.
\end{aligned}
$$

Similarly one obtains $\text{instr}_\mathbf{0} = \eta \circ \kappa_2$. Next,

$$
\begin{aligned}
&T([\kappa_2, \kappa_1]) \circ \text{instr}_p \\
&= T([\kappa_2, \kappa_1]) \circ T(\text{sep}_1) \circ \text{st}_1 \circ \langle p, \text{id} \rangle \\
&\overset{(8)}{=} T(\text{sep}_1) \circ T([\kappa_2, \kappa_1] \times \text{id}) \circ \text{st}_1 \circ \langle p, \text{id} \rangle \\
&= T(\text{sep}_1) \circ \text{st}_1 \circ (T([\kappa_2, \kappa_1]) \times \text{id}) \circ \langle p, \text{id} \rangle \\
&= T(\text{sep}_1) \circ \text{st}_1 \circ \langle p^\perp, \text{id} \rangle \\
&= \text{instr}_{p^\perp}.
\end{aligned}
$$

4. In a straightforward manner we obtain for a map f in the underlying category:

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

5. Via point (4) we get:

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

The main result of this section gives, for strongly affine monads, a bijective correspondence between predicates and side-effect-free instruments.

Theorem 1. *Let T be a strongly affine monad on a distributive category. Then there is a bijective correspondence between:*

$$
\frac{\textit{predicates } X \xrightarrow{\ p\ } T(1 + 1)}{X \xrightarrow[\ f\ ]{} T(X + X) \textit{ with } T(\nabla) \circ f = \eta}
$$

Proof. The mapping downwards is $p \mapsto \mathrm{instr}_p$, and upwards is $f \mapsto T(! + !) \circ f$. Point (2) in Lemma 3 says that $T(\nabla) \circ \mathrm{instr}_p = \eta$, since p is causal (because T is affine); point (1) tells that going down-up is the identity. For the up-down part we need to show that $f = \mathrm{instr}_p$, for $p = T(! + !) \circ f$. We use the 'strongly affine' pullback (10) to get a predicate q in:

The outer diagram commutes by (8). By construction we have $f = \mathrm{instr}_q$, see Definition 3. We thus need to prove that $q = p$. But this follows from Lemma 3 (1):

$$
p \stackrel{\mathrm{def}}{=} T(! + !) \circ f = T(! + !) \circ \mathrm{instr}_q = q. \qquad\qquad \square
$$

Example 4. We shall illustrate the situation for the powerset monad $\mathcal{P}$ on the (distributive) category **Sets**. We write $1 + 1 = 2 = \{0, 1\}$, where we identify the element $0 \in 2$ with κ_2* and $1 \in 2$ with κ_1*. Hence $\mathcal{P}(2) = \{\emptyset, \{0\}, \{1\}, \{0, 1\}\}$ and $\mathcal{P}_a(2) = \{\{0\}, \{1\}, \{0, 1\}\}$, where $\mathcal{P}_a \rightarrowtail \mathcal{P}$ is the affine submonad of non-empty subsets, see Example 3 (2).

For a predicate $p \colon X \to \mathcal{P}(2)$ the associated instrument $\mathrm{instr}_p \colon X \to \mathcal{P}(X + X)$ is, according to Definition 3, given by:

$$\mathrm{instr}_p(x) = \{\kappa_1 x \mid 1 \in p(x)\} \cup \{\kappa_2 x \mid 0 \in p(x)\}$$
$$= \begin{cases} \emptyset & \text{if } p(x) = \emptyset \\ \{\kappa_1 x\} & \text{if } p(x) = \{1\} \\ \{\kappa_2 x\} & \text{if } p(x) = \{0\} \\ \{\kappa_1 x, \kappa_2 x\} & \text{if } p(x) = \{0, 1\}. \end{cases}$$

We thus see:

$$\big(\mathcal{P}(\nabla) \circ \mathrm{instr}_p\big)(x) = \{x \mid 0 \in p(x) \text{ or } 1 \in p(x)\} = \begin{cases} \{x\} & \text{if } p(x) \neq \emptyset \\ \emptyset & \text{if } p(x) = \emptyset. \end{cases}$$

Hence these instruments are not side-effect-free, in general. But if we restrict ourselves to the (strongly affine) submonad $\mathcal{P}_a$ of non-emptyset subsets, then we do have side-effect-freeness — as shown in general in Lemma 3 (2).

In that case we have a bijective correspondence between maps $f \colon X \to \mathcal{P}_a(X + X)$ with $\mathcal{P}_a(\nabla) \circ f = \{-\}$ and predicates $p \colon X \to \mathcal{P}_a(2)$ — as shown in general in Theorem 1.

6   Commutativity

In this section we assume that T is a strong monad on a distributive category $\mathbf{C}$, so that we can associate an instrument $\mathrm{instr}_p \colon X \to T(X + X)$ with a predicate $p \colon X \to T(2)$, like in Definition 3.

Given such a predicate p we define the *assert* map $\mathrm{asrt}_p \colon X \to T(X + 1)$ as:

$$\mathrm{asrt}_p = T(\mathrm{id} + !) \circ \mathrm{instr}_p = T(\pi_2 + \pi_1) \circ T(\mathrm{dis}_1^{-1}) \circ \mathrm{st}_1 \circ \langle p, \mathrm{id} \rangle.$$

These assert maps play an important role to define conditional probabilities (after normalisation), see [4]. Here we illustrate how one can define, via these assert maps, a sequential composition operation — called 'andthen' — on predicates $p, q \colon X \to T(2)$ as:

$$p \mathrel{\&} q = [q, \kappa_2] \bullet \mathrm{asrt}_p \qquad \text{in } \mathcal{K}\ell(T)$$
$$= \mu \circ T([q, T(\kappa_2) \circ \eta]) \circ \mathrm{asrt}_p \qquad \text{in } \mathbf{C}.$$

This operation incorporates the side-effect of p, if any. Hence, in principle, this is not a commutative operation.

Example 5. We elaborate the situation described above for the state monad $T(X) = (S \times X)^S$ from Example 3 (3). A predicate on X can be identified with a map $p \colon X \to (S + S)^S$, since:

$$T(2) \,=\, \bigl(S \times 2\bigr)^S \,\cong\, \bigl(S + S\bigr)^S.$$

For $x \in X$ and $s \in S$ the value $p(x)(s) \in S + S$ describes the 'true' case via the left component, and the 'false' case via the right component. Clearly, the predicate can also change the state, and thus have a side-effect.

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:

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

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

The theorem below plays a central role for commutativity of the andthen operation $\mathbin{\&}$. It establishes a connection between commutativity of sequential composition and commutativity of the monad, as described in Diagram (3).

Theorem 2. *If T is a commutative monad, then instruments commute: for predicates $p, q \colon X \to T(2)$, the following diagram commutes in $\mathcal{K}\ell(T)$.*

$$\begin{array}{ccccc} X & \xrightarrow{\ \mathrm{instr}_p\ } & X + X & \xrightarrow{\ q+q\ } & 2 + 2 \\[2pt] \Big\| & & & & \Big\downarrow{\scriptstyle \cong}\ {\scriptstyle [\kappa_1+\kappa_1,\kappa_2+\kappa_2]} \\[2pt] X & \xrightarrow[\ \mathrm{instr}_q\ ]{} & X + X & \xrightarrow[\ p+p\ ]{} & 2 + 2 \end{array} \qquad (16)$$

Proof. The structure of the proof is given by the following diagram in the underlying category.
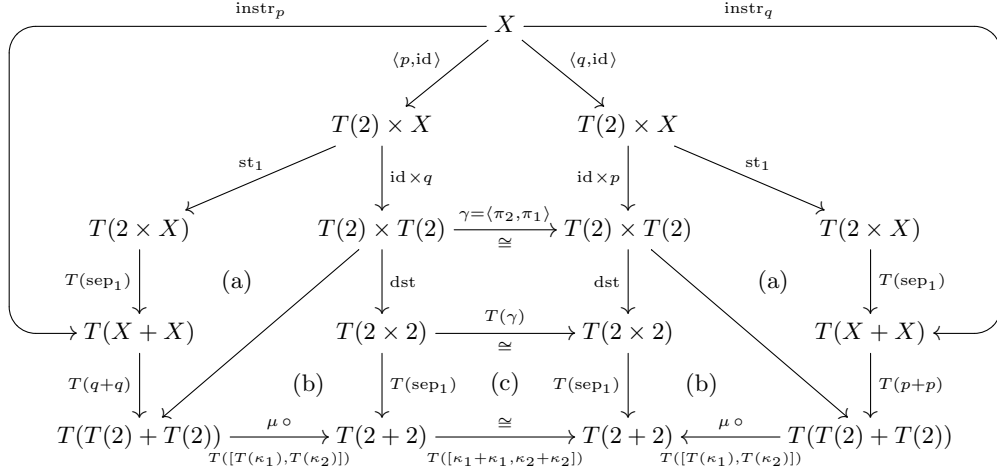
The sub-diagrams (a) commute by naturality, and sub-diagrams (b) by (6); commutation of (c) is equation (9), and the square in the middle is commutativity of the monad T, see (3). Details are left to the interested reader. □

Corollary 1. *For a commutative monad (on a distributive category), sequential composition & is commutative on causal predicates.*

Proof. We first note that in $\mathcal{K}\ell(T)$ we can write $\text{asrt}_p = (\text{id} + \bar{\top}) \bullet \text{instr}_p$. Hence if p, q are both causal, then:

$$
\begin{aligned}
p \,\&\, q &= [q, \kappa_2] \bullet (\text{id} + \bar{\top}) \bullet \text{instr}_p \\
&= [\text{id}, \kappa_2 \bullet \bar{\top}] \bullet (q + q) \bullet \text{instr}_p && \text{since } q \text{ is causal} \\
&= [\text{id}, \kappa_2 \bullet \bar{\top}] \bullet (p + p) \bullet \text{instr}_q && \text{by Theorem 2} \\
&= [p, \kappa_2] \bullet (\text{id} + \bar{\top}) \bullet \text{instr}_q && \text{since } p \text{ is causal} \\
&= q \,\&\, p. && \square
\end{aligned}
$$

7 Conclusions

We have translated the notions of side-effect-freeness and commutativity from quantum foundations (in the form of effectus theory) to monad theory, and proven some elementary results. This is only a starting point. Expecially, connections between (strong) affineness and non-locality need to be clarified.

Further, the current work forms the basis for a categorical description (that is in the making) of probability theory using strongly affine monads.

We should point out that the setting of the current work is given by distributive categories, with finite cartesian products, and not tensor products. They form in themselves already a classical setting.

Acknowledgements

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. QPEL: Quantum program and effect language. In B. Coecke, I. Hasuo, and P. Panangaden, editors, *Quantum Physics and Logic (QPL) 2014*, number 172 in Elect. Proc. in Theor. Comp. Sci., pages 133–153, 2014.
4. R. Adams and B. Jacobs. A type theory for probabilistic and Bayesian reasoning. See `arxiv.org/abs/1511.09230`, 2015.
5. 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.
6. K. Cho, B. Jacobs, A. Westerbaan, and B. Westerbaan. An introduction to effectus theory. see `arxiv.org/abs/1512.05813`, 2015.
7. B. Coecke, C. Heunen, and A. Kissinger. Categories of quantum and classical channels. *Quantum Information Processing*, pages 1–31, 2014.
8. B. Fong. Causal theories: A categorical perspective on Bayesian networks. Master's thesis, Univ. of Oxford, 2012. see `arxiv.org/abs/1301.6201`.
9. 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.
10. 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.
11. B. Jacobs. Semantics of weakening and contraction. *Ann. Pure & Appl. Logic*, 69(1):73–106, 1994.
12. B. Jacobs. Measurable spaces and their effect logic. In *Logic in Computer Science*. IEEE, Computer Science Press, 2013.
13. B. Jacobs. New directions in categorical logic, for classical, probabilistic and quantum logic. *Logical Methods in Comp. Sci.*, 11(3):1–76, 2015.
14. B. Jacobs. Effectuses from monads. MFPS'16, to appear, 2016.
15. 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.
16. 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.
17. B. Jacobs and F. Zanasi. A predicate/state transformer semantics for Bayesian learning. MFPS'16, to appear, 2016.

18. 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.

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

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

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

22. P. Panangaden. *Labelled Markov Processes*. Imperial College Press, 2009.

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