Product Rules and Distributive Laws

Joost Winter[1*]

University of Warsaw, Faculty of Mathematics, Informatics, and Mechanics
jwinter@mimuw.edu.pl

Abstract. We give a categorical perspective on various product rules, including Brzozowski's product rule $((st)_a = s_a t + o(s) t_a)$ and the familiar rule of calculus $((st)_a = s_a t + s t_a)$. It is already known that these product rules can be represented using distributive laws, e.g. via a suitable quotient of a GSOS law. In this paper, we cast these product rules into a general setting where we have two monads S and T, a (possibly copointed) behavioural functor F, a distributive law of T over S, a distributive law of S over F, and a suitably defined distributive law $TF \Rightarrow FST$. We introduce a coherence axiom giving a sufficient and necessary condition for such triples of distributive laws to yield a new distributive law of the composite monad ST over F, allowing us to determinize FST-coalgebras into lifted F coalgebras via a two step process whenever this axiom holds.

1 Introduction

In [Brz64], Brzozowski introduced a calculus of derivatives for regular expressions, including amongst other rules the *product rule*

$$(xy)_a = x_a y + o(x) y_a$$

which can be contrasted with the familiar (Leibniz) product rule from calculus (here rephrased with a somewhat uncommon nomenclature, to highlight the correspondences and differences with Brzozowski's rule)

$$(xy)_a = x_a y + x y_a.$$

In the work of Rutten, starting with [Rut98], a picture started to emerge in which Brzozowski's calculus of derivatives was absorbed into the framework of universal coalgebra. Notably, in [Rut03] and [Rut05] both product rules were considered as behavioural differential equations, or coinductive definitions, possessing unique solutions, and *defining* respectively the convolution product (Brzozowski's rule) and the shuffle product (Leibniz' rule) on the final coalgebra.

In the work of Bonsangue and Rutten together with the present author, starting with [WBR11], a coalgebraic presentation of the context-free languages

was given, in which Brzozowski's derivative rules (excluding the star rule) again play a crucial role. This work was based on a concretely presented extension of $2 \times \mathcal{P}_\omega(-^*)^A$-coalgebras into $2 \times -^A$-coalgebras, as in the following diagram:

$$
\begin{array}{ccccc}
X & \xrightarrow{\eta_X} & \mathcal{P}_\omega(X^*) & \xrightarrow{[\![-]\!]} & \mathcal{P}(A^*) \\
{\scriptstyle (o,\delta)}\downarrow & {\scriptstyle (\hat{o},\hat{\delta})} & & & \downarrow{\scriptstyle (O,\Delta)} \\
2 \times \mathcal{P}_\omega(X^*)^A & \xrightarrow{\quad 2 \times [\![-]\!]^A \quad} & & & 2 \times \mathcal{P}_\omega(X^*)^A
\end{array}
$$

In [BHKR13], it was shown that the above diagram could be understood in terms of a distributive law of the monad $\mathcal{P}_\omega(-^*)$ over the *cofree copointed functor* on $2 \times -^A$, that is, over $(- \times (2 \times -^A), \pi_1)$ by means of a suitable quotient of distributive laws. This firmly established the connection between the work of [WBR11] and the general categorical framework of *bialgebras* and *distributive laws*, considered in e.g. [LPW00], [Bar04], [Jac06a], [Jac06b], [Kli11], and [JSS12], as well as to the (closely related) *generalized powerset construction* presented in [SBBR10] and [SBBR13].

Moreover, in [BRW12], the approach was generalized from formal languages to formal power series, thus giving a coalgebraic characterization of the *constructively algebraic power series*.

In [Win14], the author's PhD thesis, this extension was broken up into a two step process, as in the diagram,

$$
\begin{array}{ccccccc}
X & \xrightarrow{\eta_X^0} & X^* & \xrightarrow{\eta_{X^*}^1} & \mathcal{P}_\omega(X^*) & \xrightarrow{[\![-]\!]} & \mathcal{P}(A^*) \\
{\scriptstyle (o,\delta)}\downarrow & {\scriptstyle (o^\#,\delta^\#)} & {\scriptstyle (\hat{o},\hat{\delta})} & & & & \downarrow{\scriptstyle (O,\Delta)} \\
2 \times \mathcal{P}_\omega(X^*)^A & & \xrightarrow{\qquad 2 \times [\![-]\!]^A \qquad} & & & & 2 \times \mathcal{P}_\omega(X^*)^A
\end{array}
$$

and it was noted that the first extension corresponded directly to the product rule (combined with the unit rule), whereas the second extension was simply an instance of ordinary determinization, applied to a coalgebra which is by itself infinite. This two-step process again easily generalizes from the setting of languages to the setting of noncommuting power series over commutative semirings; however, no general categorical presentation of this two-step process was given in the thesis.

The main aim of this paper is to give such a categorical presentation of this two-step process. This involves a setting including both distributive laws between monads $\lambda^0 : TS \Rightarrow ST$, distributive laws of a monad over a functor (or copointed functor) $\lambda^2 : SF \Rightarrow FS$, as well as a, suitably defined, distributive law of type $\lambda^1 : TF \Rightarrow FST$ (this definition will be given in Section 3).

In Section 4, we will present our main result, which states that the three laws as suggested above can be combined to form a new distributive law $\hat{\lambda}$ of the composite monad ST over F, if and only if a certain *coherence axiom* involving the three laws is satisfied. This construction works regardless whether F is an ordinary or copointed functor. The proof that satisfaction of the coherence axiom

implies the multiplicative law of the distributive law of the composite monad ST over F was automatically derived using a Prolog program written for this purpose.

We then show that both the Brzozowski and Leibniz product rules, as well as the (simpler) pointwise rule defining the Hadamard product, can be cast into this setting, and that the coherence axiom is satisfied in these cases, and additionally present a counterexample where the coherence axiom is not satisfied. We conclude with a presentation of an extension of the generalized powerset construction to incorporate these two-step extensions.

Related work. All of the aforementioned references can be regarded as related work. In particular, [Che07] gives a similar coherence condition, for the existence of composite distributive laws in the context of distributive laws between monads. The relationships between [Che07] and the present paper involve some subtleties, and in Section 6, we will present a more detailed investigation of the relationship.

Additionally, [MMS13] and [Sch14] also give a compositional approach to distributive laws, however, in a setting different from our approach, not including the combination of distributive laws between monads and distributive laws of monads over functors into new distributive laws. In [Jac06b], a coinductive presentation of the shuffle product using a two-step process is also given, however, the coinductive definition of the multiplication there is not given using a distributive law.

Acknowledgements. For various discussions, comments, and criticism (hopefully in all cases taken as constructively as possible), I would like to thank Henning Basold, Marcello Bonsangue, Helle Hvid Hansen, Bart Jacobs, Bartek Klin, Jurriaan Rot, Jan Rutten, and Alexandra Silva, as well as to Charles Paperman for suggesting automation of the proofs (which in the end was carried out for only a part of just one of the propositions). Additionally I would like to thank the anonymous referees (as well as the anonymous refees of the earlier, rejected, submission) for a large amount of corrections and constructive comments.

2 Preliminaries

2.1 General preliminaries

We assume familiar the basic notions of category theory (functors, natural transformations, monads, (Eilenberg-Moore) algebras), of (commutative) monoids and semirings, and of coalgebras for a functor. These can be found in e.g. [Awo10] or [Mac71] for general categorical notions, and in [Rut00] for the basics of coalgebra.

An important endofunctor we will use is the functor $K \times -^A$ on **Set**, representing automata with output in K over the input alphabet A, often called *Moore automata* but in this paper simply K-*automata*. The transition of coalgebras for this functor is represented as a pair (o, δ), called an *output-derivative pair*, with $\delta(x)(a)$ denoted by x_a. A final coalgebra for this functor exists, and

can be given by $(K\langle\!\langle A\rangle\!\rangle, O, \Delta)$, where $K\langle\!\langle A\rangle\!\rangle$ is the function space $A^* \to K$, $O(\sigma) = \sigma(1)$ and $\sigma_a(w) = \sigma(aw)$ for all $\sigma \in K\langle\!\langle A\rangle\!\rangle$, $a \in A$, and $w \in A^*$. Here, and elsewhere in this paper, 1 denotes the empty word.

A *copointed endofunctor* is a pair (F, ϵ), where F is a **C**-endofunctor, and ϵ is a natural transformation $\epsilon : F \Rightarrow 1_{\mathbf{C}}$. A coalgebra for a copointed endofunctor (F, ϵ) consists of a coalgebra (X, δ) for the endofunctor F, satisfying the condition that $\epsilon_X \circ \delta = 1_X$.

Of special interest to us are *cofree copointed endofunctors*, which can be constructed in any category that has binary products. Given an ordinary endofunctor F, the cofree copointed endofunctor on F is the pair $(\mathrm{Id} \times F, \pi_1)$. For any functor F such that a cofree copointed endofunctor on F exists, there is an isomorphism between coalgebras for the ordinary endofunctor F, and coalgebras for the copointed endofunctor $(\mathrm{Id} \times F, \pi_1)$.

Given a monad T (we will frequently refer to a monad simply using the name of its functorial part) and an endofunctor F, a (T, F)-bialgebra is a triple (X, α, δ) such that (X, α) is an algebra for the monad T and (X, δ) is a F-coalgebra. If F is a copointed functor (here again, we often refer to the copointed functor simply using the name of the functor), we additionally require the coalgebraic part of a (T, F)-bialgebra to be a coalgebra for the copointed functor. (We will turn to λ-bialgebras involving a distributive law soon!)

Given a commutative semiring K (in this paper, we will essentially only be concerned with commutative semirings), a K-semimodule $(M, 0, +, \times)$ consists of a commutative monoid $(M, 0, +)$ and an operation $\times : K \times M \to M$ (called *scalar product*) such that, for all $k, l \in k$ and $m, n \in M$:

$$k \times (m + n) = (k \times m) + (k \times n) \quad 1_K \times m = m$$
$$(k + l) \times m = (k \times m) + (l \times m) \quad 0_K \times m = 0$$
$$(kl) \times m = k \times (l \times m) \quad\quad k \times 0 = 0$$

For any commutative semiring K^1, the K-semimodules are exactly the algebras for the monad $(\mathrm{Lin}_K(-), \eta, \mu)$, where $\mathrm{Lin}_K(X)$ is the set of finitely supported functions on X, and, given a function $f : X \to Y$,

$$\mathrm{Lin}_K(f) = y \mapsto \sum_{x \in f^{-1}(y)} f(x).$$

The multiplication of the monad, for any X and any $f \in \mathrm{Lin}_K(\mathrm{Lin}_K(X))$, can be given by

$$\mu_X(f) = x \mapsto \sum_{g \in \mathrm{supp}(f)} f(g) \cdot g(x)$$

and the unit of the monad, for any $x \in X$, by:

$$\eta_X(x) = y \mapsto \textbf{if } x = y \textbf{ then } 1 \textbf{ else } 0$$

---

[1] For a semiring K that is not commutative, one can define two distinct monad structures on $\mathrm{Lin}_K(-)$, with a different multiplication, one corresponding to left-K-semimodules, and one corresponding to right K-semimodules. When K is commutative, these two monads are identical.

Given some $\sigma \in \mathrm{Lin}_K(X)$ and some $x \in X$, we furthermore write $[\sigma \Downarrow x]$ for the application of σ at x.

Given a commutative semiring K, a K-algebra[2] (see e.g. [Eil76]) is a tuple $(X, 0, 1, +, \cdot, \times)$ such that $(X, 0, 1, +, \cdot)$ is a semiring, $(X, 0, +, \times)$ is a K-semimodule, such that for all $k \in K$ and $x \in X$:

$$(k \times 1) \cdot x = k \times x = x \cdot (k \times 1)$$

An equivalent, and as it turns out for us more convenient, characterization is the following: a K-algebra is a tuple $(X, 0, 1, +, \cdot, \times)$ such that $(X, 1, \cdot)$ is a monoid, $(X, 0, +, \times)$ is a K-semimodule, furthermore satisfying the following axioms:

$$x \cdot (y + z) = x \cdot y + x \cdot z$$
$$(x + y) \cdot z = x \cdot z + y \cdot z$$
$$(k \times 1) \cdot x = k \times x = x \cdot (k \times 1)$$

(Note that the pair of semiring axioms $0 \cdot x = 0 = x \cdot 0$ can be derived from the pair of axioms on the third line.)

For some semirings K, the categories of K-semimodules and K-algebras correspond to familiar categories. We summarize some of these in the following table:

K	K-semimodules	K-algebras
$\mathbb{B}$	join-semilattices	idempotent semirings
$\mathbb{N}$	commutative monoids	semirings
$\mathbb{Z}$	abelian groups	rings

Moreover, it is easily verified that the monad $\mathrm{Lin}_{\mathbb{B}}(-)$ is naturally isomorphic to the finite powerset monad $\mathcal{P}_\omega$.

2.2 Distributive laws between monads

We now summarize the definitions and some important facts about distributive laws between monads, which can be found in e.g. [Bec69] and [BW85].

Given monads (S, μ^S, η^S) and (T, μ^T, η^T), a *distributive law* of the monad T over the monad S is a natural transformation $\lambda : TS \Rightarrow ST$ such that the four diagrams of natural transformations

---

[2] In the literature often called a unital associative algebra

$$
\begin{array}{ccc}
T \xrightarrow{\;T\eta^S\;} TS & \quad & TSS \xLongrightarrow{\;\;\;\;\;T\mu^S\;\;\;\;\;} TS \\
\;\;\;\searrow_{\eta^S_T} \quad \Big\Vert\lambda & \quad & \lambda S\Big\Vert \qquad\qquad\qquad \Big\Vert\lambda \\
\qquad\quad ST & \quad & STS \xrightarrow{\;S\lambda\;} SST \xrightarrow{\;\mu^S T\;} ST
\end{array}
$$

commute.

In the presence of a distributive law of a monad T over another monad S, we obtain, amongst other things, the following:

1. A monad structure on the composite functor ST, given by:

$$\eta^{ST} = \eta^S T \circ \eta^T \qquad \text{and} \qquad \mu^{ST} = S\mu^T \circ \mu^S TT \circ S\lambda T$$

2. A lifting of the monad S to a monad $\hat{S}$ in the category $\mathbf{C}^T$ of T-algebras, with the functorial part of $\hat{S}$ given by:

$$\hat{S}(X,\alpha) = (SX, S\alpha \circ \lambda_X)$$

3. An isomorphism of categories between the categories of algebras for the monad ST and algebras for the monad $\hat{S}$. Given an $\hat{S}$-algebra (X, α^T, α^S), the corresponding ST-algebra can be given by $(X, \alpha^S \circ S\alpha^T)$, and given a ST-algebra (X, α), the corresponding $\hat{S}$-algebra can be given by $(X, \alpha \circ \eta^S_{TX}, \alpha \circ S\eta^T_X)$. These constructions are mutually inverse, as shown in [Bec69, Section 2].

In the case where $T = -^*$ and $S = \mathrm{Lin}_K(-)$ (where K is a commutative semiring), there is a distributive law $\lambda : (\mathrm{Lin}_K(-))^* \Rightarrow \mathrm{Lin}_K(-^*)$ of T over S, given by:

$$
\lambda_X\left(\prod_{i=1}^{n} \sum_{j=1}^{m_i} k_{ij} \times x_{ij} \right) = \sum_{j_1=1}^{m_1} \cdots \sum_{j_n=1}^{m_n} \left(\prod_{i=1}^{n} k_{ij_i} \times \prod_{i=1}^{n} x_{ij_i} \right)
$$

(recall that $\times$ is the scalar product, and thus, the k_{ij} here are scalars of K!) Here the product symbol $\prod$ denotes the operation on the monoid structure, and the summation symbol $\sum$ and times $\times$ denote the (sum and scalar product) on the semimodule structure, respectively.

It is well-known, at least[3] in the cases of the semirings $\mathbb{B}$, $\mathbb{N}$, and $\mathbb{Z}$ (and probably in the remaining cases as well—nevertheless we give a full proof in the appendix), that this is a distributive law of $-^*$ over $\mathrm{Lin}_K(-)$:

Proposition 1. *The natural transformation λ as given above is a distributive law of the monad $\mathrm{Lin}_K(-)$ over the monad $-^*$.*

We let $K\langle-\rangle = \mathrm{Lin}_K(-^*)$ denote the resulting composite monad. Elements of $K\langle X\rangle$ can be identified with noncommuting polynomials in variables from X.

---

[3] See e.g. [Bec69] for $\mathbb{Z}$ and [Jac06a] for $\mathbb{N}$ and $\mathbb{B}$

Moreover, the algebras for this monad are precisely the K-algebras: by definition, a K-algebra is both a monoid and a semimodule additionally satisfying

$$x \cdot (y + z) = x \cdot y + x \cdot z$$
$$(x + y) \cdot z = x \cdot z + y \cdot z$$
$$(k \times 1) \cdot x = k \times x = x \cdot (k \times 1)$$

for elements x, y, z and scalars k and an algebra for the monad $K\langle - \rangle$ is both a monoid and a semimodule additionally satisfying

$$\prod_{i=1}^{n} \sum_{j=1}^{m_i} k_{ij} \times x_{ij} = \sum_{j_1=1}^{m_1} \cdots \sum_{j_n=1}^{m_n} \left(\prod_{i=1}^{n} k_{ij_i} \times \prod_{i=1}^{n} x_{ij_i} \right)$$

It is easily verified that these two conditions are equivalent.

2.3   Distributive laws of a monad over a (copointed) endofunctor

We now turn to a brief summary of distributive laws of monads over endofunctors and copointed endofunctors. The material which now follows can all be found in either [Bar04] or [JSS12]; supplementary presentations can be found in e.g. [Jac06a], [Jac06b], [Kli11], and [LPW00] (in which the notions of distributive laws of monads over endofunctors and copointed endofunctors were first considered).

Given a monad (T, μ, η) and an endofunctor F on any category $\mathbf{C}$, a *distributive law* of the monad T over F is a natural transformation $\lambda : TF \Rightarrow FT$ such that the two diagrams of natural transformations

commute.

Given a distributive law λ of T over F, a λ-bialgebra is a (T, F)-bialgebra for which the following diagram commutes:

A distributive law of a monad (T, η, μ) over a copointed endofunctor (F, ϵ) is a distributive law of the monad over the (ordinary) endofunctor F, such that additionally the diagram

commutes.

For distributive laws of a monad over a copointed endofunctor, bialgebras can again be defined: the only modification needed here is that we additionally require (X, δ) to be a coalgebra for the copointed endofunctor F. Note that any distributive law λ of a monad T over an ordinary functor F extends to a distributive law λ' of T over the cofree copointed functor $(\mathrm{Id} \times F)$, such that the λ-bialgebras and λ'-bialgebras are isomorphic as categories.

In the presence of a distributive law λ of T over F, we obtain, amongst other things, the following (both for ordinary and copointed endofunctors, see e.g. [Bar04] for proofs):

1. A lifting of F to a functor $\hat{F}$ in the category of algebras for the monad T, and of T to a functor $\hat{T}$ the category of F-coalgebras, together with an isomorphic correspondence between the categories of λ-bialgebras, $\hat{F}$-coalgebras, and algebras for the monad $\hat{T}$.
2. If F has a final coalgebra (Ω, ω), there is a unique algebra ξ for the monad T on Ω such that (Ω, ξ, ω) is a λ-bialgebra, which moreover is a final λ-bialgebra.

An important instance of a distributive law, considered in e.g. [JSS12], is the law of the functor $K \times -^A$ over the monad $\mathrm{Lin}_K(-)$ where K is any semiring[4], given by:

$$
\lambda_X \left(\sum_{i=1}^n k_i \times (o_i, a \mapsto d_{ia}) \right) = \left(\sum_{i=1}^n k_i o_i, a \mapsto \sum_{i=1}^n k_i \times d_{ia} \right)
$$

This distributive law can be obtained canonically by means of the strength operation (see e.g. [JSS12]), using the obvious K-semimodule structure on K itself. The bialgebras for this distributive law are precisely the structures that are both K-automata and K-semimodules, such that the output and derivative (with respect to an arbitrary alphabet symbol) are both linear mappings. A concrete formulation is the following: a *K-linear automaton*[5] is a tuple $(X, 0, +, \times, o, \delta)$ such that $(X, 0, +, \times)$ is a K-semimodule, (X, o, δ) is a K-automaton, and o and $\delta(-)(a)$ (for each $a \in A$) are linear mappings.

2.4 The generalized powerset construction

We now turn to a summary of some of the main observations from [SBBR10] and [SBBR13] pertaining to the *generalized powerset construction* (many instances of which can be found in these two papers as examples). Beyond that, we also sketch how the 'generalized powerset construction' framework can be extended to distributive laws over (cofree) copointed functors, and highlight some of the (additional) subtleties that arise in this case.

---

[4] If K is not commutative, we can consider left K-semimodules, but for this paper, this is not relevant.

[5] called *linear weighted automata* in e.g. [BBB$^+$12]

Given a (possibly copointed) endofunctor F such that a final F-coalgebra exists, a monad T (on the same category), and a distributive law of T over F, any FT-coalgebra (X, δ) (if F is copointed, we additionally require that $\epsilon_{TX} \circ \delta = \eta_X$) can be extended canonically to an F-coalgebra $(TX, \hat{\delta})$, where $\hat{\delta}$ can be specified as:

$$\hat{\delta} = F\mu_X \circ \lambda_{TX} \circ T\delta$$

In fact, it follows that $(TX, \mu_X, \hat{\delta})$ is a λ-bialgebra (and, hence, that FTX has the structure of an algebra for the monad T). If F has a final coalgebra (Ω, ω), we thus obtain

$$
\begin{array}{ccccc}
X & \xrightarrow{\eta_X} & TX & \xrightarrow{[\![-]\!]} & \Omega \\
{\scriptstyle\delta}\downarrow & {}_{\hat{\delta}}\nearrow & & & \downarrow{\scriptstyle\omega} \\
FTX & \xrightarrow{F[\![-]\!]} & & & F\Omega
\end{array}
$$

where $[\![-]\!]$ is a morphism of λ-bialgebras to the unique λ-bialgebra on the final F-coalgebra. If F is copointed, note that

$$\epsilon_{TX} \circ \hat{\delta} = \epsilon_{TX} \circ F\mu_X \circ \lambda_{TX} \circ T\delta = \mu_X \circ \epsilon_{TTX} \circ \lambda_{TX} \circ T\delta$$

$$= \mu_X \circ T\epsilon_X \circ T\delta = \mu_X \circ T\eta_X = 1_{TX}$$

so that $\hat{\delta}$ is indeed a coalgebra for the copointed functor.

When the functor F is a cofree copointed functor of the form $(\mathrm{Id} \times G, \pi_1)$, note that δ has to be of the form (η_X, ζ), where ζ is an (ordinary) GT-coalgebra, and because $\hat{\delta}$ is a coalgebra for the cofree copointed functor, it has to be of the form $(1_{TX}, \zeta^\flat)$ for some $\zeta^\flat$. Similarly, ω has to be of the form $(1_\Omega, \psi)$, where (Ω, ψ) is the final G-coalgebra. Now note that there is a correspondence of commuting diagrams of the form

$$
\begin{array}{ccccc}
X & \xrightarrow{\eta_X} & TX & \xrightarrow{[\![-]\!]} & \Omega \\
{\scriptstyle(\eta_X,\zeta)}\downarrow & {}_{(1_{TX},\zeta^\flat)}\nearrow & & & \downarrow{\scriptstyle(1_\Omega,\psi)} \\
TX \times GTX & \xrightarrow{[\![-]\!] \times G[\![-]\!]} & & & \Omega \times G\Omega
\end{array}
$$

and:

$$
\begin{array}{ccccc}
X & \xrightarrow{\eta_X} & TX & \xrightarrow{[\![-]\!]} & \Omega \\
{\scriptstyle\zeta}\downarrow & {}_{\zeta^\flat}\nearrow & & & \downarrow{\scriptstyle\psi} \\
GTX & \xrightarrow{G[\![-]\!]} & & & G\Omega
\end{array}
\qquad (1)
$$

However, in general we are *not* guaranteed that the second diagram occurs as the result of an ordinary distributive law of T over G. Hence, in general, although we are guaranteed that $TX \times GTX$ has the structure of an algebra for the monad T, we do not have this guarantee of GTX itself (this may in fact fail, e.g. for GSOS rules that cannot be presented using the simple SOS format). On the other hand, we are still guaranteed that $[\![-]\!]$ is a T-algebra morphism, because this follows from the upper diagram.

3 Distributive laws of a monad over a (copointed) endofunctor into a composite monad

In this section, we present a definition of a new (as far as the author is aware) type of distributive law, which allows us to formulate the 'two-step' generalized powerset construction. We start by giving the general formulation, before turning to the main examples.

Given two monads (S, μ^S, η^S), (T, μ^T, η^T) such that a distributive law $\lambda^0 : TS \Rightarrow ST$ exists, and an endofunctor F, on any category $\mathbf{C}$, a *distributive law* of the monad T over F into the composite monad ST is a natural transformation $\lambda : TF \Rightarrow FST$ such that the two diagrams of natural transformations

$$
\begin{array}{ccc}
F \xRightarrow{\eta^T F} TF & & \\
\underset{F\eta^{ST}}{\searrow} \quad \Big\Vert \lambda & \text{and} & \\
FST & &
\end{array}
\qquad
\begin{array}{ccc}
TTF \xRightarrow{\mu^T F} TF \\
T\lambda \Big\Vert \qquad \Big\Vert \lambda \\
TFST \xRightarrow{\lambda ST} FSTST \xRightarrow{F\mu^{ST}} FST
\end{array}
$$

commute.

We can again define a suitable notion of a λ-bialgebra for a distributive law of T over F into the composite monad ST. For such a λ, a λ-bialgebra (X, α, δ) is a (ST, F)-bialgebra such that the diagram

$$
\begin{array}{ccccccc}
TX & \xrightarrow{\eta^S_{TX}} & STX & \xrightarrow{\alpha} & X & \xrightarrow{\delta} & FX \\
{\scriptstyle T\delta}\Big\downarrow & & & & & & \Big\uparrow{\scriptstyle F\alpha} \\
TFX & & & \xrightarrow{\lambda_X} & & & FSTX
\end{array}
$$

commutes (note that we can here replace $\alpha \circ \eta^S_{TX}$ with α^T).

A distributive law λ of a monad (T, η, μ) over a copointed endofunctor (F, ϵ) into a composite monad is a distributive law of the monad over the (ordinary) endofunctor F into the same composite monad, such that additionally the diagram

$$
\begin{array}{ccc}
TF & \xRightarrow{\lambda} & FST \\
{\scriptstyle T\epsilon}\Big\Vert & & \Big\Vert{\scriptstyle \epsilon ST} \\
T & \xRightarrow{\eta_S} & ST
\end{array}
$$

commutes. Again, the corresponding notion of a λ-bialgebra is obtained by adding the additional requirement that (X, δ) is a coalgebra for the copointed functor F.

The following result is unspectacular, but will be useful for us later on:

Lemma 2. *Given monads (S, η^S, μ^S) and (T, η^T, μ^T), and a (possibly copointed) endofunctor F, and a distributive law $\lambda^0 : TS \Rightarrow ST$, if a natural transformation $\lambda : TF \Rightarrow FT$ is a distributive law of T over F, then $\hat{\lambda} = F\eta^S T \circ \lambda$ is a distributive law of T over F into the composite monad ST. If η is pointwise monic and F preserves monos, the converse also holds.*

3.1 Product rules as distributive laws

Instantiating $S = \mathrm{Lin}_K(-)$, $T = -^*$, and $F = K \times -^A$, we can model each of the three product rules as distributive laws of the monad (T, η^T, μ^T) over the copointed functor $(\mathrm{Id} \times F, \pi_1)$ into the composite monad $(ST, \eta^{ST}, \mu^{ST})$ as follows:

1. The Brzozowski product rule can be specified as

$$\lambda_X \left(\prod_{i=1}^{n} (x_i, o_i, a \mapsto d_{ia}) \right)$$

$$= \left(\prod_{i=1}^{n} x_i, \prod_{i=1}^{n} o_i, a \mapsto \sum_{i=1}^{n} \left(\prod_{k=1}^{i-1} o_i \right) \times \left(d_{ia} \cdot \prod_{k=i+1}^{n} x_i \right) \right)$$

or equivalently inductively as:

$$\lambda_X(1) = (1, 1, a \mapsto 0)$$
$$\lambda_X(x, o, a \mapsto d_a)w = (x\pi_0(\lambda_X(w)),$$
$$o\pi_1(\lambda_X(w)), a \mapsto d_a \pi_0(\lambda_X(w)) + o\pi_2(\lambda_X(w))(a))$$

The latter pair of equations can be more conveniently represented using the output-derivative notation as:

$$\begin{array}{ll} o(1) = 1 & 1_a = 0 \\ o(xw) = o(x)o(w) & (xw)_a = x_a w + o(x) w_a \end{array}$$

2. The Leibniz product rule can be given as

$$\lambda_X \left(\prod_{i=1}^{n} (x_i, o_i, a \mapsto d_{ia}) \right)$$

$$= \left(\prod_{i=1}^{n} x_i, \prod_{i=1}^{n} o_i, a \mapsto \sum_{i=1}^{n} \left(\left(\prod_{k=1}^{i-1} x_i \right) \cdot d_{ia} \cdot \prod_{k=i+1}^{n} x_i \right) \right)$$

or inductively and using the output-derivative notation as:

$$\begin{array}{ll} o(1) = 1 & 1_a = 0 \\ o(xw) = o(x)o(w) & (xw)_a = x_a w + x w_a \end{array}$$

3. Finally, the pointwise (Hadamard) product rule is given as

$$\lambda_X \left(\prod_{i=1}^{n} (o_i, a \mapsto d_{ia}) \right) = \left(\prod_{i=1}^{n} o_i, a \mapsto \prod_{k=1}^{n} d_{ia} \right)$$

or inductively and using the output-derivative notation as:

$$\begin{array}{ll} o(1) = 1 & 1_a = 1 \\ o(xw) = o(x)o(w) & (xw)_a = x_a w_a \end{array}$$

In the case of the Brzozowski and Leibniz product rules, it can be verified that they yield distributive laws using a lengthy verification.

Proposition 3. *The Brzozowski and Leibniz product rule both are distributive laws of $-^*$ over the copointed endofunctor $(\mathrm{Id} \times F, \pi_1)$ into the composite monad $K\langle - \rangle$.*

In the case of the Hadamard product, everything is a bit easier.

Proposition 4. *The pointwise product rule is a distributive law of $-^*$ over the endofunctor F into the composite monad $K\langle - \rangle$.*

Proof. This follows from the fact that the Hadamard product can be given as a pointwise distributive law of $-^*$ over $K \times -^A$ obtained using the strength operation and multiplication over K giving its monoid structure (see e.g. [Jac06b] for a detailed discussion of distributive laws obtained using strength). By Lemma 2, this distributive law yields another distributive law into the composite monad. $\square$

We will henceforth refer to the three above distributive laws as λ^{brz}, λ^{lei}, and λ^{had}, respectively.

From the above propositions and the definition of λ-bialgebras for distributive laws of a monad over a functor into another monad, it follows directly that any λ-bialgebra for any of the above distributive laws satisfies the corresponding product rule for arbitrary elements. For example, it is easy to see from the definitions that the λ-bialgebras for the Brzozowski product rule are precisely the (T, F)-bialgebras that satisfy the unit and product rule

$$
\begin{aligned}
o(1) &= 1 & o(xy) &= o(x)o(y) \\
1_a &= 0 & (xy)_a &= x_a y + o(x) y_a
\end{aligned}
$$

for arbitrary elements x and y of the bialgebra and alphabet symbols a. *Mutatis mutandis*, the same holds for the two other product rules.

4   Composite distributive laws

In this section we will introduce a *coherence axiom*, comparable to but different from the Yang-Baxter condition presented in [Che07]. The relation with the work in [Che07] will be discussed in Section 6. We consider a setting, where we are given monads S and T, a (possibly copointed) endofunctor F, on an arbitrary category $\mathbf{C}$, together with three distributive laws, as follows:

- $\lambda^0 : TS \Rightarrow ST$, a distributive law of T over S.
- $\lambda^1 : TF \Rightarrow FST$, a distributive law of T over F into the composite monad ST.
- $\lambda^2 : SF \Rightarrow FS$, a distributive law of S over F.

For such triples[6], we can compose the distributive laws as follows, yielding a new natural transformation $\hat{\lambda} : (ST)F \Rightarrow F(ST)$ given by:

$$STF \overset{S\lambda^1}{\Longrightarrow} SFST \overset{\lambda^2 ST}{\Longrightarrow} FSST \overset{F\mu^S T}{\Longrightarrow} FST$$

A triple of distributive laws $(\lambda^0, \lambda^1, \lambda^2)$ as above is said to satisfy the *coherence axiom* whenever the following diagram commutes:

$$
\begin{array}{ccccccc}
TSF & \overset{T\lambda^2}{\Longrightarrow} & TFS & \overset{\lambda^1 S}{\Longrightarrow} & FSTS & \overset{FS\lambda^0}{\Longrightarrow} & FSST \\
{\scriptstyle \lambda^0 F} \big\Vert & & & & & & \big\Vert {\scriptstyle F\mu^S T} \\
STF & \overset{S\lambda^1}{\Longrightarrow} & SFST & \overset{\lambda^2 ST}{\Longrightarrow} & FSST & \overset{F\mu^S T}{\Longrightarrow} & FST
\end{array}
\qquad (2)
$$

As it turns out, the coherence axiom is a sufficient and necessary condition for $\hat{\lambda}$ to be a distributive law:

Theorem 5. *Given three distributive laws as above for monads S and T and a functor (or copointed functor) F, the natural transformation $\hat{\lambda}$ is a distributive law of the composite monad ST over the functor (or copointed functor) F if and only if the coherence axiom is satisfied.*

Perhaps interesting to note, after initially finding myself unable to prove one of the directions of the above theorem by hand (the initial aim was to prove that a triple of distributive laws *always* yields a composite distributive law), I wrote a Prolog program generating all the possible ways to go from $STSTF$ to STF, together with equivalences determined by naturality squares, the multiplicative laws for the monad, and the various distributive laws (however excluding the unit laws), as well as all of these equivalences with any functor applied to it. Using this program, the source of which can be found at

$$\texttt{http://www.mimuw.edu.pl/~jwinter/distlaws.prolog}$$

the following facts were found:

1. There are 784 different ways of going from $STSTF$ to FST, using compositions of arbitrary natural transformations satisfying the regular expression:

$$(S|T|F)^* (\mu^S|\mu^T|\lambda^0|\lambda^1|\lambda^2)(S|T|F)^*$$

2. Without the coherence axiom, and using the equivalence relation named above, these divide up into two different equivalence classes (of sizes 218 and 566, respectively).
3. With the coherence axiom added, the two equivalence classes collapse into one.

---

[6] In some of the literature, monads are called triples, but in this paper it simply means a list of three elements.

After the proof was *found* automatically, it was, however, easily *verified* by hand, and transformed into a (rather large) commuting diagram.

Later, it was found (by hand) that the coherence axiom can also be derived from the existence of a composite distributive law, turning the theorem into one giving a sufficient and necessary condition, and indeed a counterexample (soon to be presented) to the coherence axiom was found. (Note that neither the theorem, nor the falling apart of the paths into two equivalence classes is *by itself* a proof that there are counterexamples to the coherence axiom, but only a proof that the coherence axiom cannot be derived using one particular method.)

Furthermore, it turns out that each of the distributive laws from Section 3.1 satisfies the coherence axiom, when combined with the distributive laws presented in Sections 2.2 and 2.3:

Proposition 6. *The three distributive laws for the Brzozowski, Leibniz, and Hadamard product rules satisfy the coherence axiom, when instantiating λ^0 with the distributive law between monads presented in Section 2.2, and instantiating λ^2 with the distributive law of S over F from Section 2.3.*

We will henceforth refer to the extensions of the three distributive laws earlier named as $\hat{\lambda}^{\mathrm{brz}}$, $\hat{\lambda}^{\mathrm{lei}}$, and $\hat{\lambda}^{\mathrm{had}}$, respectively.

We now turn to the counterexample, consisting of a distributive law of T over F into the composite monad ST, which, together with the same λ^2 and λ^0 as before, does not satisfy the coherence axiom.

Counterexample 1. Consider the natural transformation given by $\lambda^1 : (\mathbb{N} \times -^A)^* \Rightarrow \mathbb{N} \times \mathbb{N}\langle - \rangle^A$ given by

$$\lambda^1_X \left(\prod_{i=1}^n (x_i, o_i, a \mapsto d_{ia}) \right) = \left(\prod_{i=1}^n x_i, \sum_{i=1}^n o_i, a \mapsto \prod_{i=1}^n d_{ia} \right)$$

This distributive law can again be given by means of a (pointwise) distributive law from $-^*$ over $\mathbb{N} \times -^A$ using the strength operator, this time by considering the additive monoid structure on $\mathbb{N}$ (as opposed to the multiplicative monoid structure, which yields the Hadamard product).

Now consider

$$((x, 1, a \mapsto x) + (x, 1, a \mapsto x)) \cdot (x, 1, a \mapsto x)$$

which is an element of $TSF\{x\}$. It now follows that by applying

$$F\mu^S T \circ FS\lambda^0 \circ \lambda^1 S \circ T\lambda^2$$

we obtain an element $(y, o, d) \in FST\{x\}$ such that $o = 3$, whereas by applying

$$F\mu^S T \circ \lambda^2 ST \circ S\lambda^1 \circ \lambda^0 F$$

we obtain an element $(y', o', d') \in FST\{x\}$ such that $o' = 4$. This proves that the coherence axiom is not satisfied, and hence, by Theorem 5, also that $\hat{\lambda}$, although a natural transformation, is not a distributive law of the composite monad ST over F.

Proposition 7. *Given three distributive laws $\lambda^0, \lambda^1, \lambda^2$ as above, such that $\hat{\lambda}$ is a distributive law over ST over F, a (ST,F)-bialgebra (X,α,δ) is a $\hat{\lambda}$-bialgebra if and only if (X,α^S,δ) is a λ^2-bialgebra and (X,α^T,δ) is a λ^1-bialgebra.*

Proof. We will establish the result by establishing the following implications (in the diagrams, each of the (internal) faces commutes by either one of the assumed algebras or bialgebras, one of the axioms for monads or distributive laws, a naturality square, or a functor applied to any of these types of diagrams):

– If (X,α,δ) is a $\hat{\lambda}$-bialgebra, then (X,α^T,δ) is a λ^1-bialgebra. This is established by the following diagram:

$$
\begin{array}{ccccccc}
TX & \xrightarrow{\eta^S_{TX}} & STX & \xrightarrow{\alpha} & X & \xrightarrow{\delta} & FX \\
\downarrow{\scriptstyle T\delta} & & \downarrow{\scriptstyle ST\delta} & & & & \uparrow{\scriptstyle F\alpha} \\
TFX & \xrightarrow{\eta^S_{TFX}} & STFX & \xrightarrow{S\lambda^1_X} & SFSTX & \xrightarrow{\lambda^0_{STX}} & FSSTX \xrightarrow{F\mu^S_{TX}} FSTX
\end{array}
$$

with λ^1_X, η^S_{FSTX}, $F\eta^S_{STX}$, 1_{FSTX}, $FSTX$.

– If (X,α,δ) is a $\hat{\lambda}$-bialgebra, then (X,α^S,δ) is a λ^2-bialgebra. This is established by the following diagram:

$$
\begin{array}{ccccccc}
SX & \xrightarrow{S\eta^T} & STX & \xrightarrow{\alpha} & X & \xrightarrow{\delta} & FX \\
\end{array}
$$

with $S\delta$, $ST\delta$, $STFX$, $FSSTX$, $F\mu^S_{TX}$, $S\eta^T_{FX}$, $S\lambda^1_X$, λ^0_{STX}, $FS\eta^S_{TX}$, $F\alpha^S$, $SFSTX$, $SF\eta^{ST}_X$, $FS\eta^{ST}_X$, $FSTX$, 1_{FSTX}, $FSTX$, SFX, λ^0_X, FSX, $FS\eta^T_X$, $FS\alpha^T$, 1_{FSX}, FSX.

– If (X, α^T, δ) is a λ^1-bialgebra and (X, α^S, δ) is a λ^2-bialgebra, then (X, α, δ) is a $\hat{\lambda}$-bialgebra. This is established by the following diagram:

From the preceding proposition, we can now directly conclude that, given a commutative semiring K, the corresponding $\hat{\lambda}^{\mathrm{brz}}$-bialgebras are precisely the structures $(X, 0, 1, +, \cdot, \times, o, \delta)$ such that $(X, 0, 1, +, \cdot, \times)$ is a K-algebra, that $(X, 0, +, \times, o, \delta)$ is a K-linear automaton, and such that additionally the unit and product rule are satisfied for arbitrary elements. Again, *mutatis mutandis*, the same holds for the Leibniz and pointwise product rules. It may be worthwhile to note that (in the context of a singleton alphabet A and a field as underlying semiring) the bialgebras for $\hat{\lambda}^{\mathrm{lei}}$, are precisely *differential algebras* (as defined in e.g. [RR08]) together with output functions that are K-algebra morphisms.

Another corollary of this proposition is that the final $\hat{\lambda}$-bialgebra has a structure compatible with the final λ^2-bialgebra, because the final $\hat{\lambda}$-bialgebra is a λ^2-bialgebra (by the preceding proposition) compatible with the final F-coalgebra (as a consequence of [Bar04, Corollary 3.4.19]), and again by [Bar04, Corollary 3.4.19], there is a *unique* λ^2-bialgebra compatible with the final F-coalgebra. By similar reasoning, the final $\hat{\lambda}$-bialgebra also has a structure with the final λ^1-bialgebra. So, the final $\hat{\lambda}$-bialgebra can now be seen to naturally extend both the final λ^1-bialgebra and the final λ^2-bialgebra, inheriting the coalgebraic structure from the final F-coalgebra, and with algebraic structure compatible to that of the final bialgebras for λ^1 and λ^2.

5 The generalized powerset construction for composite distributive laws

We can now formulate a two-step version of the generalized powerset construction, in the setting where we are given three distributive laws $\lambda^0, \lambda^1, \lambda^2$ as in the previous section. Here the extensions can be specified as

$$\delta^\sharp \;=\; F\mu^{ST}_X \circ \lambda^1_{STX} \circ T\delta \qquad \text{and} \qquad \hat{\delta} \;=\; F\mu^S_{TX} \circ \lambda^2_{STX} \circ S\delta^\sharp.$$

giving the following diagram if a final F-coalgebra exists:

$$
\begin{array}{ccccccc}
X & \xrightarrow{\eta_X^T} & TX & \xrightarrow{\eta_{TX}^S} & STX & \xrightarrow{\llbracket - \rrbracket} & \Omega \\
\downarrow{\scriptstyle\delta} & {\scriptstyle\delta^{\sharp}}\swarrow & \downarrow{\scriptstyle\hat{\delta}} & & & & \downarrow{\scriptstyle\omega} \\
FSTX & & \xrightarrow{\hspace{2cm}F\llbracket - \rrbracket\hspace{2cm}} & & & & F\Omega
\end{array}
$$

We derive the following specification of $\hat{\delta}$ directly in terms of δ:

$$\hat{\delta} = F\mu_{TX}^S \circ \lambda_{STX}^2 \circ SF\mu_X^{ST} \circ S\lambda_{STX}^1 \circ ST\delta$$

Proposition 8. *The extension $\hat{\delta}$ obtained by the above two step extension coincides with the extension obtained in one step using the distributive law $\hat{\lambda} : (ST)F \Rightarrow F(ST)$. In other words:*

$$F\mu_{TX}^S \circ \lambda_{STX}^2 \circ SF\mu_X^{ST} \circ S\lambda_{STX}^1 \circ ST\delta = F\mu_X^{ST} \circ \hat{\lambda}_{STX} \circ ST\delta$$

Proof. The following commuting diagram

$$
\begin{array}{ccccccc}
FSSTSTX & \xrightarrow{FSS\lambda_{TX}^0} & FSSSTTX & \xrightarrow{FS\mu_{TTX}^S} & FSSTTX & \xrightarrow{FSS\mu_X^T} & FSSTX \\
\downarrow{\scriptstyle F\mu_{TSTX}^S} & & \downarrow{\scriptstyle F\mu_{STTX}^S} & & \downarrow{\scriptstyle F\mu_{TTX}^S} & & \downarrow{\scriptstyle F\mu_{TX}^S} \\
FSTSTX & \xrightarrow{FS\lambda_{TX}^0} & FSSTTX & \xrightarrow{F\mu_{TTX}^S} & FSTTX & \xrightarrow{FS\mu_X^T} & FSTX
\end{array}
$$

establishes that $F\mu_{TX}^S \circ FS\mu_X^{ST} = F\mu_X^{ST} \circ F\mu_{TSTX}^S$.

Next, observe that

$$
\begin{aligned}
& F\mu_{TX}^S \circ \lambda_{STX}^2 \circ SF\mu_X^{ST} \circ S\lambda_{STX}^1 \circ ST\delta \\
= {}& F\mu_{TX}^S \circ FS\mu_X^{ST} \circ \lambda_{STSTX}^2 \circ S\lambda_{STX}^1 \circ ST\delta \quad \textit{(naturality)} \\
= {}& F\mu_X^{ST} \circ F\mu_{TSTX}^S \circ \lambda_{STSTX}^2 \circ S\lambda_{STX}^1 \circ ST\delta \quad \textit{(result of above diagram)} \\
= {}& F\mu_X^{ST} \circ \hat{\lambda}_{STX} \circ ST\delta
\end{aligned}
$$

completing the proof. □

As a result of this proposition, we can directly conclude that, in the two-step diagram above, $FSTX$ can again be assigned the structure of an algebra for the monad ST, and the mappings $\hat{\delta}$ and $\llbracket - \rrbracket$ are ST-algebra morphisms.

Finally note that, if F is a cofree copointed functor $(\mathrm{Id} \times G)$, the same considerations as those surrounding the diagrams in (1) again apply.

Remark 9. Note that, when instantiating the above construction with the laws λ^0 from Section 2.2, λ^{brz}, and λ^2 from 2.3, the resulting composite distributive law is a law between the monad $S\langle - \rangle$ and the cofree copointed endofunctor on $K \times -^A$. This law differs from (but is closely related to) the laws obtained in [WBR13] and [MPW16], in which a distributive law is defined between a suitably defined monad on $K\langle A + - \rangle$ and the endofunctor $K \times -^A$ (without copoint). However, the resulting law $\hat{\lambda}$ still has the property, that a language is context-free (resp. a power series is constructively algebraic) iff it is the final coalgebra semantics of the determinization of a finite $K \times K\langle - \rangle^A$-coalgebra.

6 Comparison to the coherence condition from [Che07]

In [Che07], a similar coherence condition to the one from this paper is presented, providing a way of iterating distributive laws between which a coherence condition similar to (but simpler than) the one presented in this paper is used. We will concentrate on the specific case of just three monads with interacting distributive laws, as the more general iterated cases are less directly relevant to the present paper. In this case, the main result from [Che07] instantiates as:

Proposition 10. *Given monads (S, η^S, μ^S), (T, η^T, μ^T), and (U, η^U, μ^U), together with distributive laws $\lambda^0 : UT \Rightarrow TU$, $\lambda^1 : US \Rightarrow SU$, and $\lambda^2 : TS \Rightarrow ST$, if the diagram*

$$
\begin{array}{ccccc}
UTS & \overset{U\lambda^2}{\Longrightarrow} & UST & \overset{\lambda^1 T}{\Longrightarrow} & SUT \\
\lambda^0 S \Big\| & & & & \Big\| S\lambda^0 \\
TUS & \overset{T\lambda^1}{\Longrightarrow} & TSU & \overset{\lambda^2 U}{\Longrightarrow} & STU
\end{array}
$$

(called the coherence condition or Yang-Baxter condition*) commutes, the natural transformations*

$$\lambda^2 U \circ T\lambda^1 : (TU)S \Rightarrow S(TU)$$

and

$$S\lambda^0 \circ \lambda^1 T : U(ST) \Rightarrow (ST)U$$

are distributive laws between monads (of the composite monad TU over S, and of U over the composite monad ST), respectively). Moreover both of these laws yield the same composite monad on STU.

A difference between this result and the corresponding result from our paper is that the result in [Che07] is only presented in one direction, instead of as an equivalence. Thus, one can wonder if in this case, too, the result can be turned into an equivalence. In fact, it can be stated in the form of the following three-way equivalence:

Proposition 11. *Given monads S, T, and U and distributive laws λ^0, λ^1, and λ^2 as in the previous proposition, the following are equivalent:*

1. *The Yang-Baxter condition holds.*
2. *$\lambda^2 U \circ T\lambda^1$ is a distributive law of the composite monad TU over S.*
3. *$S\lambda^0 \circ \lambda^1 T$ is a distributive law of U over the composite monad ST.*

As a variant on this result, bridging the gap between it and our result, we will move to a setting where, rather than three monads, we have two monads, (S, η^S, μ^S) and (T, η^T, μ^T) and an endofunctor F, together with a distributive law between monads $\lambda^0 : TS \Rightarrow ST$, and two distributive laws of the monad-over-endofunctor type, $\lambda^1 : TF \Rightarrow FT$, and $\lambda^2 : SF \Rightarrow FS$. We now obtain the following result, in much the same manner as before:

Proposition 12. *Given monads (S, η^S, μ^S), (T, η^T, μ^T), and an endofunctor F, together with distributive laws $\lambda^0 : TS \Rightarrow ST$, $\lambda^1 : TF \Rightarrow FT$, and $\lambda^2 : SF \Rightarrow FS$, the following are equivalent:*

1. *The Yang-Baxter condition holds, as follows:*

$$
\begin{array}{ccccc}
TSF & \overset{T\lambda^2}{\Longrightarrow} & TFS & \overset{\lambda^1 S}{\Longrightarrow} & FTS \\
\lambda^0 F \big\Downarrow & & & & \big\Downarrow F\lambda^0 \\
STF & \overset{S\lambda^1}{\Longrightarrow} & SFT & \overset{\lambda^2 T}{\Longrightarrow} & FST
\end{array}
\tag{3}
$$

2. *$\lambda^2 T \circ S\lambda^1$ is a distributive law of the composite monad ST over F.*

(Note that this result is very similar to, but different from, one direction of [BMSZ15, Proposition 7.1], which is concerned with two monads, an endofunctor, and one distributve law between these monads, and two Kleisli-type laws, i.e. laws of the endofunctor over both of the monads.)

In fact, the result can be related to the results from Section 4 by means of Lemma 2 and the following proposition:

Proposition 13. *Given monads (S, η^S, μ^S) and (T, η^T, μ^T), an endofunctor F, and distributive laws $\lambda^0 : TS \Rightarrow ST$, $\lambda^1 : TF \Rightarrow FT$, and $\lambda^2 : SF \Rightarrow FS$, the Yang-Baxter condition (3) holds if and only if the law $\hat{\lambda}^1 = F\eta^S T \circ \lambda^1 : TF \Rightarrow FST$ makes diagram (2) (Section 4) commute.*

Connecting this result to the three product rules mentioned in this paper, we recall that the Hadamard product rule can also be given directly as a distributive law of type $TF \Rightarrow FT$ (with $T = -^*$, $F = K \times -^A$) using the strength operator. This law can be given concretely by

$$
\lambda_X \left(\prod_{i=1}^n (o_i, a \mapsto d_{ia}) \right) = \left(\prod_{i=1}^n o_i, a \mapsto \prod_{k=1}^n d_{ia} \right)
$$

with the difference that the right hand side is now regarded as an element of FT, instead of an element of FST (with $S = \mathrm{Lin}_K(-)$). The Leibniz and Brzozowski product rules, however, do not fit into this simpler scheme, as a direct result of the occurrence of a summation on the right hand side.

It thus turns out that, while the results from [Che07] are easily modified to a coalgebraic setting with laws $\lambda^0 : TS \Rightarrow ST$, $\lambda^1 : TF \Rightarrow FT$, and $\lambda^2 : SF \Rightarrow FS$, the usage of distributive laws into a composite monad (where λ^1 is of the form $TF \Rightarrow FST$) allows us to specify product rules of the type $(xy)_a = x_y a + o(x) y_a$, where a summation occurs in the right hand side, that are not possible to specify in a two-step mechanism using the simple variant of the result in [Che07].

Additionally, in the case of distributive laws into composite monads, we require an octagonal coherence condition, rather than the hexagonal condition from [Che07].

7 Further directions

One interesting direction for future work consists of making a further generalization from copointed functors to *comonads*. Another possible direction, owing inspiration to [Wor09] (where it is shown that K-algebras are precisely the monoids in the monoidal category of K-semimodules), consists of investigating whether we can reverse the order of the two-step process in the concrete setting of product rules, going from X via $\text{Lin}_K(X)$ to $K\langle X \rangle$ instead of via X^*. Finally, it may be very worthwhile to further develop the Prolog-program, developed here in an *ad hoc* manner to prove a part of Theorem 5, into a more general tool able to (at least in certain contexts) prove commutativity of diagrams.

References

[Awo10]   Steve Awodey. *Category Theory.* Oxford University Press, 2010.

[Bar04]   Falk Bartels. *On Generalized Coinduction and Probabilistic Specification Formats.* PhD thesis, Vrije Universiteit Amsterdam, 2004.

[BBB⁺12] Filippo Bonchi, Marcello M. Bonsangue, Michele Boreale, Jan Rutten, and Alexandra Silva. A coalgebraic perspective on linear weighted automata. *Information and Computation*, 211:77–105, 2012.

[Bec69]   Jon Beck. Distributive laws. In *Seminar on triples and categorical homology theory*, pages 119–140. Springer, 1969.

[BHKR13] Marcello M. Bonsangue, Helle H. Hansen, Alexander Kurz, and Jurriaan Rot. Presenting distributive laws. In Reiko Heckel and Stefan Milius, editors, *CALCO*, volume 8089 of *Lecture Notes in Computer Science*, pages 95–109. Springer, 2013.

[BMSZ15] Filippo Bonchi, Stefan Milius, Alexandra Silva, and Fabio Zanasi. Killing epsilons with a dagger: A coalgebraic study of systems with algebraic label structure. *Theoret. Comput. Sci.*, 604:102–126, 2015. `http://dx.doi.org/10.1016/j.tcs.2015.03.024`.

[BRW12]   Marcello M. Bonsangue, Jan Rutten, and Joost Winter. Defining context-free power series coalgebraically. In Pattinson and Schröder [PS12], pages 20–39.

[Brz64]   Janusz A. Brzozowski. Derivatives of regular expressions. *Journal of the ACM*, 11:481–494, 1964.

[BW85]   Michael Barr and Charles Wells. *Toposes, triples, and theories.* Grundlehren der mathematischen Wissenschaften. Springer-Verlag, New York, 1985.

[Che07]   Eugenia Cheng. Iterated distributive laws. `http://arxiv.org/pdf/0710.1120v1.pdf`, 2007.

[Eil76]   Samuel Eilenberg. *Automata, Languages, and Machines.* Academic Press, Inc., 1976.

[Jac06a] Bart Jacobs. A bialgebraic review of deterministic automata, regular expressions and languages. In *Essays Dedicated to Joseph A. Goguen*, pages 375–404, 2006.

[Jac06b] Bart Jacobs. Distributive laws for the coinductive solution of recursive equations. *Inf. Comput.*, 204(4):561–587, 2006.

[JSS12]   Bart Jacobs, Alexandra Silva, and Ana Sokolova. Trace semantics via determinization. In Pattinson and Schröder [PS12], pages 109–129.

[Kli11]   Bartek Klin. Bialgebras for structural operational semantics: An introduction. *Theoretical Computer Science*, 412(38):5043–5069, 2011.

[LPW00]   Marina Lenisa, John Power, and Hiroshi Watanabe. Distributivity for endofunctors, pointed and co-pointed endofunctors, monads and comonads. *Electr. Notes Theor. Comput. Sci.*, 33:230–260, 2000.

[Mac71]   Saunders MacLane. *Categories for the Working Mathematician*. Springer-Verlag, 1971.

[MMS13]   Stefan Milius, Lawrence S. Moss, and Daniel Schwencke. Abstract GSOS rules and a modular treatment of recursive definitions. *Logical Methods in Computer Science*, 9(3), 2013.

[MPW16]   Stefan Milius, Dirk Pattinson, and Thorsten Wißmann. A new foundation for finitary corecursion. *CoRR*, abs/1601.01532, 2016.

[PS12]   Dirk Pattinson and Lutz Schröder, editors. *Coalgebraic Methods in Computer Science - 11th International Workshop, CMCS 2012, Colocated with ETAPS 2012, Tallinn, Estonia, March/April 2012, Revised Selected Papers*, volume 7399 of *Lecture Notes in Computer Science*. Springer, 2012.

[RR08]   Markus Rosenkranz and Georg Regensburger. Integro-differential polynomials and operators. In *Proceedings of the Twenty-first International Symposium on Symbolic and Algebraic Computation*, ISSAC '08, pages 261–268, New York, NY, USA, 2008. ACM.

[Rut98]   Jan Rutten. Automata and coinduction (an exercise in coalgebra). In Davide Sangiorgi and Robert de Simone, editors, *CONCUR*, volume 1466 of *Lecture Notes in Computer Science*, pages 194–218. Springer, 1998.

[Rut00]   Jan Rutten. Universal coalgebra: a theory of systems. *Theoretical Computer Science*, 249(1):3–80, 2000.

[Rut03]   Jan Rutten. Behavioural differential equations: a coinductive calculus of streams, automata, and power series. *Theoretical Computer Science*, 308(1-3):1–53, 2003.

[Rut05]   Jan Rutten. A coinductive calculus of streams. *Mathematical Structures in Computer Science*, 15(1):93–147, 2005.

[SBBR10]   Alexandra Silva, Filippo Bonchi, Marcello M. Bonsangue, and Jan Rutten. Generalizing the powerset construction, coalgebraically. In Kamal Lodaya and Meena Mahajan, editors, *FSTTCS*, volume 8 of *LIPIcs*, pages 272–283. Schloss Dagstuhl—Leibniz-Zentrum für Informatik, 2010.

[SBBR13]   Alexandra Silva, Filippo Bonchi, Marcello Bonsangue, and Jan Rutten. Generalizing determinization from automata to coalgebras. *Logical Methods in Computer Science*, 9(1), 2013.

[Sch14]   Daniel Schwencke. *Compositional and Effectful Recursive Specification Formats*. PhD thesis, Technische Universität Braunschweig, 2014.

[WBR11]   Joost Winter, Marcello M. Bonsangue, and Jan Rutten. Context-free languages, coalgebraically. In Andrea Corradini, Bartek Klin, and Corina Cîrstea, editors, *CALCO*, volume 6859 of *Lecture Notes in Computer Science*, pages 359–376. Springer, 2011.

[WBR13]   Joost Winter, Marcello M. Bonsangue, and Jan J. M. M. Rutten. Coalgebraic characterizations of context-free languages. *Logical Methods in Computer Science*, 9 (3:14), 2013.

[Win14]   Joost Winter. *Coalgebraic Characterizations of Automata-Theoretic Classes*. PhD thesis, Radboud Universiteit Nijmegen, 2014.

[Wor09]   James Worthington. *Automata, Representations, and Proofs*. PhD thesis, Cornell University, 2009.

A   Appendix

In this appendix we give the proofs which were omitted in the main text.

Out of the five proofs, three involve concrete verifications where the S, T and F are instantiated, respectively, as $\mathrm{Lin}_K(-)$, $-^*$, and $- \times K \times -^A$ (or, in the case of the proof of the coherence axiom for the Hadamard product, $K \times -^A$). To make it a bit clearer to see what is going on, in these proofs we use the following color-coding reflecting the types of subexpression:

- black—for subexpressions of type X
- magenta—for subexpressions of type K, i.e. for scalars
- blue—for subexpressions of type S (finite (formal) linear combinations), e.g.:

$$\sum_{i=1}^{n} k_i \times x_i \qquad \text{and} \qquad (x) \text{ for } \eta^S(x)$$

- red—for subexpressions of type T (words), e.g.:

$$\prod_{i=1}^{n} x_i \qquad \text{and} \qquad (x) \text{ for } \eta^T(x)$$

- green—for subexpressions of type F (identity-output-derivative triples), e.g.:

$$(x, o, a \mapsto d_a)$$

For expressions of type TT and SS, we use square brackets to clarify the border between the inner and outer expression, e.g.:

$$\prod_{i=1}^{n} \left[\prod_{j=1}^{n_i} x_{ij} \right]$$

Proof. (of Proposition 1) We check each of the four axioms:

1. $\lambda \circ \mu^T S = S\mu^T \circ \lambda T \circ T\lambda$:

$$(S\mu_X^T \circ \lambda_{TX} \circ T\lambda_X)\left(\prod_{i=1}^{m}\left[\prod_{j_i=1}^{n_i}\left(\sum_{l_{ij_i}=1}^{o_{ij_i}} k_{ij_i l_{ij_i}} \times x_{ij_i l_{ij_i}}\right)\right]\right)$$

$$= (S\mu_X^T \circ \lambda_{TX})\left(\prod_{i=1}^{m}\left(\sum_{l_{i1}=1}^{o_{i1}} \cdots \sum_{l_{in_i}=1}^{o_{in_i}} \prod_{j_i=1}^{n_i} k_{ij_i l_{ij_i}} \times \prod_{j_i=1}^{n_i} x_{ij_i l_{ij_i}}\right)\right)$$

$$= S\mu_X^T\left(\sum_{l_{11}=1}^{o_{11}} \cdots \sum_{l_{1n_1}=1}^{o_{1n_1}} \cdots\cdots \sum_{l_{m1}=1}^{o_{m1}} \cdots \sum_{l_{mn_m}=1}^{o_{mn_m}} \prod_{i=1}^{m}\prod_{j_i=1}^{n_i} k_{ij_i l_{ij_i}} \times \prod_{i=1}^{m}\left[\prod_{j_i=1}^{n_i} x_{ij_i l_{ij_i}}\right]\right)$$

$$= \sum_{l_{11}=1}^{o_{11}} \cdots \sum_{l_{1n_1}=1}^{o_{1n_1}} \cdots\cdots \sum_{l_{m1}=1}^{o_{m1}} \cdots \sum_{l_{mn_m}=1}^{o_{mn_m}} \prod_{i=1}^{m}\prod_{j_i=1}^{n_i} k_{ij_i l_{ij_i}} \times \prod_{i=1}^{m}\prod_{j_i=1}^{n_i} x_{ij_i l_{ij_i}}$$

$$= \lambda_X\left(\prod_{i=1}^{m}\prod_{j_i=1}^{n_i}\left(\sum_{l_{ij_i}=1}^{o_{ij_i}} k_{ij_i l_{ij_i}} \times x_{ij_i l_{ij_i}}\right)\right)$$

$$= (\lambda_X \circ \mu_{SX}^T)\left(\prod_{i=1}^{m}\left[\prod_{j_i=1}^{n_i}\left(\sum_{l_{ij_i}=1}^{o_{ij_i}} k_{ij_i l_{ij_i}} \times x_{ij_i l_{ij_i}}\right)\right]\right)$$

2. $\lambda \circ T\mu^S = \mu^S T \circ S\lambda \circ \lambda S$.

$$(\mu_{TX}^S \circ S\lambda_X \circ \lambda_{SX})\left(\prod_{i=1}^{m}\left(\sum_{j_i=1}^{n_i} p_{ij_i} \times \left[\sum_{l_{ij_i}=1}^{o_{ij_i}} k_{ij_i l_{ij_i}} \times x_{ij_i l_{ij_i}}\right]\right)\right)$$

$$= (\mu_{TX}^S \circ S\lambda_X)\left(\sum_{j_1=1}^{n_1} \cdots \sum_{j_m=1}^{n_m} \prod_{i=1}^{m} p_{ij_i} \times \left(\prod_{i=1}^{m}\left(\sum_{l_{ij_i}=1}^{o_{ij_i}} k_{ij_i l_{ij_i}} \times x_{ij_i l_{ij_i}}\right)\right)\right)$$

$$= \mu_{TX}^S\left(\sum_{j_1=1}^{n_1} \cdots \sum_{j_m=1}^{n_m} \prod_{i=1}^{m} p_{ij_i} \times \left[\sum_{l_{1j_1}=1}^{o_{1j_1}} \cdots \sum_{l_{mj_m}=1}^{o_{mj_m}} \prod_{i=1}^{m} k_{ij_i l_{ij_i}} \times \left(\prod_{i=1}^{m} x_{ij_i l_{ij_i}}\right)\right]\right)$$

$$= \left(\sum_{j_1=1}^{n_1}\sum_{l_{1j_1}=1}^{o_{1j_1}} \cdots \sum_{j_m=1}^{n_m}\sum_{l_{mj_m}=1}^{o_{mj_m}} \prod_{i=1}^{m} p_{ij_i} k_{ij_i l_{ij_i}} \times \left(\prod_{i=1}^{m} x_{ij_i l_{ij_i}}\right)\right)$$

$$= \lambda_X\left(\prod_{i=1}^{m}\left(\sum_{j_i=1}^{n_i}\sum_{l_{ij_i}=1}^{o_{ij_i}} p_{ij_i} k_{ij_i l_{ij_i}} \times x_{ij_i l_{ij_i}}\right)\right)$$

$$= (\lambda_X \circ T\mu_X^S)\left(\prod_{i=1}^{m}\left(\sum_{j_i=1}^{n_i} p_{ij_i} \times \left[\sum_{l_{ij_i}=1}^{o_{ij_i}} k_{ij_i l_{ij_i}} \times x_{ij_i l_{ij_i}}\right]\right)\right)$$

3. $\lambda \circ \eta^T S = S \eta^T$.

$$(\lambda_X \circ \eta_{SX}^T)\left(\sum_{i=1}^{n} k_i \times x_i\right)$$

$$= \lambda_X\left(\left(\sum_{i=1}^{n} k_i \times x_i\right)\right)$$

$$= \left(\sum_{i=1}^{n} k_i \times (x_i)\right)$$

$$= S\eta_X^T\left(\sum_{i=1}^{n} k_i \times x_i\right)$$

4. $\lambda \circ T \eta^S = \eta^S T$.

$$(\lambda_X \circ T \eta_X^S)\left(\prod_{i=1}^{n} x_i\right)$$

$$= \lambda_X\left(\prod_{i=1}^{n} (x_i)\right)$$

$$= \left(\left(\prod_{i=1}^{n} x_i\right)\right)$$

$$= \eta_{TX}^S\left(\prod_{i=1}^{n} x_i\right)$$

Proof. (of Lemma 2) The multiplication diagram is satisfied because of the commuting diagram:

$$
\begin{array}{c}
TTF \xrightarrow{\ \mu^T F\ } TF \\
T\lambda^1 \Big\downarrow \qquad\qquad \Big\downarrow \lambda^1 \\
TFT \xrightarrow{\lambda^1 T} FTT \xrightarrow{\ F\mu^T\ } FT \\
TF\eta^S T \Big\downarrow \quad FT\eta^S T \Big\downarrow \quad {\small F\eta^S TT} \searrow \quad F\eta^S T \Big\downarrow \\
TFST \xrightarrow{\lambda^1 ST} FTST \xrightarrow{F\lambda^0 T} FSTT \xrightarrow{FS\mu^T} FST \\
F\eta^S TST \Big\downarrow \quad F\eta^S STT \Big\downarrow \quad F\eta^S ST \Big\downarrow \quad 1_{FST}\searrow \\
FSTST \xrightarrow{FS\lambda^0 T} FSSTT \xrightarrow{FSS\mu^T} FSST \xrightarrow{F\mu^S T} FST \\
\qquad F\mu^S TT \searrow \qquad FS\mu^T \nearrow \\
FSTT
\end{array}
$$

The unit diagram is satisfied because of the commuting diagram:

$$
\begin{array}{c}
F \xrightarrow{\ \eta^T F\ } TF \\
F\eta^T \searrow \quad \Big\downarrow \lambda^1 \\
FT \xrightarrow{\ F\eta^S T\ } FST
\end{array}
$$
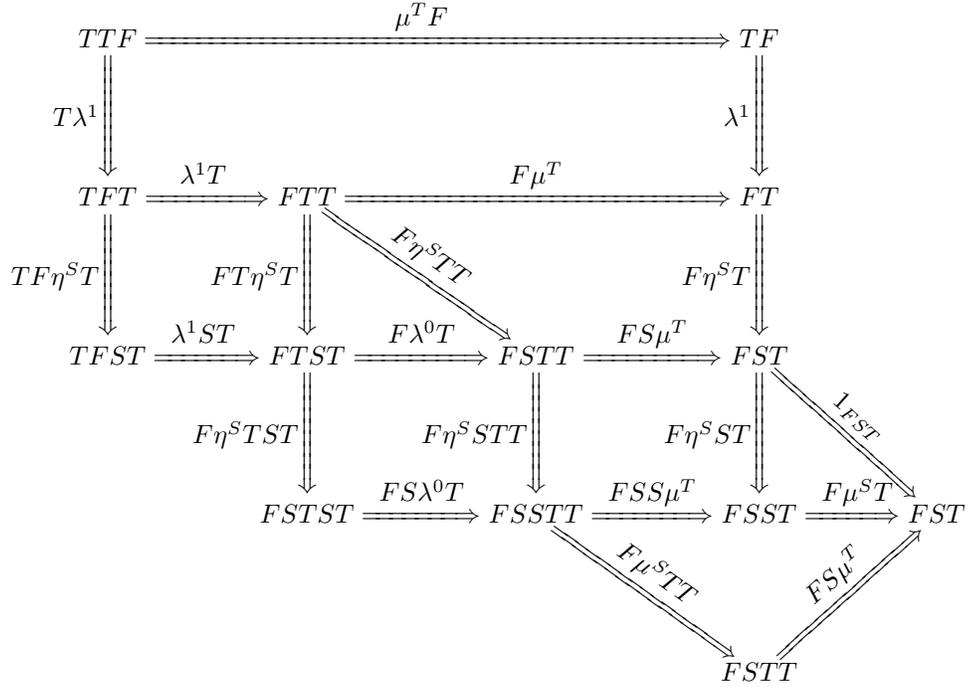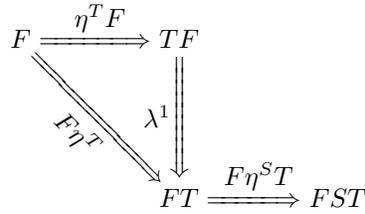
If F is copointed, the counit diagram is satisfied because of the commuting diagram:

$$
\begin{array}{c}
TF \xrightarrow{\ \lambda^1\ } FT \xrightarrow{\ F\eta^S T\ } FST \\
T\epsilon \searrow \quad \epsilon T\Big\downarrow \qquad \epsilon ST \Big\downarrow \\
T \xrightarrow{\ \eta^S T\ } ST
\end{array}
$$

For the converse, assume that η is pointwise monic, and that F preserves monos. As a result of

$$
\begin{array}{ccccccc}
TTF & \xrightarrow{\quad \mu^T F \quad} & TF & \xrightarrow{\quad \lambda \quad} & FT \\
\end{array}
$$

The diagram contains:

- $TTF \xrightarrow{\mu^T F} TF \xrightarrow{\lambda} FT$ (top row)
- $T\lambda$ (left vertical)
- $TFT \xrightarrow{TF\eta^S T} TFST$, $FSTST \xrightarrow{FS\lambda^0 T} FSSTT$, $F\eta^S T$
- λST, $F\eta^S TST$, $F\eta^S STT$, $F\mu^S TT$
- λT, $FTST \xrightarrow{F\lambda^0 T} FSTT \xrightarrow{1_{FSTT}} FSTT \xrightarrow{FS\mu^T} FST$
- $FT\eta^S T$, $F\eta^S TT$, $F\eta^S TT$
- $F\eta^S T$
- $FTT \xrightarrow{F\mu^T} FT$ (bottom row)

and the fact that $F\eta^S T$ is pointwise monic, it follows that $\lambda \circ \mu^T F = F\mu^T \circ \lambda T \circ T\lambda$.

As a result of

$$
\begin{array}{ccc}
F & \xrightarrow{\eta^T F} & TF \\
& & \\
& FT & FT \\
& & \\
& & FST \\
\end{array}
$$

The diagram contains:

- $F \xrightarrow{\eta^T F} TF$ (top)
- $F\eta^T$ (left diagonal to FT)
- λ (right vertical, TF to FT)
- $F\eta^S T$ (diagonal from FT to FST)
- $F\eta^S T$ (right vertical, FT to FST)

and the fact that $F\eta^S T$ is pointwise monic, it follows that $\lambda \circ \eta^T F = F\eta^T$.

Finally, if λ is a law for a copointed endofunctor (F, ϵ), from

and the fact that $\eta^S T$ is pointwise monic, it follows that $\epsilon T \circ \lambda = T \epsilon$.

Proof. (of Proposition 3)

In the first two parts of this proof, we use the following convention: given a doubly indexed expression such as

$$\prod_{i=1}^{m} \prod_{j=1}^{n_i} x_{ij},$$

let $\hat{n}$ denote $\sum_{i=1}^{m} n_i$, and let f be the bijection between $\{1, \ldots \hat{n}\}$ and

$$\{(i,j) \,|\, 1 \le i \le m, 1 \le j \le n_i\}$$

defined in order to make the equality

$$\prod_{i=1}^{m} \prod_{j=1}^{n_i} x_{ij} = \prod_{i=1}^{\hat{n}} x_{f(i)}$$

hold, or equivalently:

$$\mu_X^T \left(\prod_{i=1}^{m} \left[\prod_{j=1}^{n_i} x_{ij} \right] \right) = \prod_{i=1}^{\hat{n}} x_{f(i)}$$

(Note that the corresponding equality for summations will also hold!)

We now check the various axioms for the Brzozowski and Leibniz product rules:

1. The multiplication axiom for the Brzozowski product rule:

$$(F\mu_X^{ST} \circ \lambda_{TX} \circ T\lambda_X)\left(\prod_{i=1}^{m}\left[\prod_{j=1}^{n_i}(x_{ij}, o_{ij}, a \mapsto d_{ija})\right]\right)$$

$$= (F\mu_X^{ST} \circ \lambda_{TX})\left(\prod_{i=1}^{m}\left(\prod_{j=1}^{n_i}x_{ij}, \prod_{j=1}^{n_i}o_{ij}, a \mapsto \sum_{j=1}^{n_i}\left(\prod_{h=1}^{j-1}o_{ih}\right) \times \left(d_{ija} \cdot \prod_{h=j+1}^{n_i}x_{ih}\right)\right)\right)$$

$$= F\mu_X^{ST}\left(\prod_{i=1}^{m}\left(\prod_{j=1}^{n_i}x_{ij}\right), \prod_{i=1}^{m}\prod_{j=1}^{n_i}o_{ij}, a \mapsto \sum_{i=1}^{m}\left(\prod_{k=1}^{i-1}\prod_{l=1}^{n_k}o_{kl}\right) \times\right.$$

$$\left.\left(\left(\sum_{j=1}^{n_i}\left(\prod_{h=1}^{j-1}o_{ih}\right) \times \left(d_{ija} \cdot \prod_{h=j+1}^{n_i}x_{ih}\right)\right) \cdot \prod_{k=i+1}^{m}\left(\prod_{l=1}^{n_k}x_{kl}\right)\right)\right)$$

$$= \left(\prod_{i=1}^{m}\prod_{j=1}^{n_i}x_{ij}, \prod_{i=1}^{m}\prod_{j=1}^{n_i}o_{ij}, a \mapsto \sum_{i=1}^{m}\sum_{j=1}^{n_i}\left(\left(\prod_{k=1}^{i-1}\prod_{l=1}^{n_k}o_{kl}\prod_{h=1}^{j-1}o_{ih}\right) \times\right.\right.$$

$$\left.\left.\left(d_{ija} \cdot \left(\prod_{h=j+1}^{n_i}x_{ih}\right)\prod_{k=i+1}^{m}\prod_{l=1}^{n_k}x_{kl}\right)\right)\right)$$

$$= \left(\prod_{i=1}^{\hat{n}}x_{f(i)}, \prod_{i=1}^{\hat{n}}o_{f(i)}, a \mapsto \sum_{i=1}^{\hat{n}}\left(\left(\prod_{k=1}^{i-1}o_{f(k)}\right) \times \left(d_{f(k)a} \cdot \prod_{k=i+1}^{\hat{n}}x_{f(k)}\right)\right)\right)$$

$$= \lambda_X\left(\prod_{i=1}^{\hat{n}}(x_{f(i)}, o_{f(i)}, a \mapsto d_{f(i)a})\right)$$

$$= (\lambda_X \circ \mu_{FX}^{T})\left(\prod_{i=1}^{m}\left[\prod_{j=1}^{n_i}(x_{ij}, o_{ij}, a \mapsto d_{ija})\right]\right)$$

2. The multiplication axiom for the Leibniz product rule:

$$(F\mu_X^{ST} \circ \lambda_{TX} \circ T\lambda_X)\left(\prod_{i=1}^{m}\left[\prod_{j=1}^{n_i}(x_{ij}, o_{ij}, a \mapsto d_{ija})\right]\right)$$

$$= (F\mu_X^{ST} \circ \lambda_{TX})\left(\prod_{i=1}^{m}\left(\prod_{j=1}^{n_i}x_{ij}, \prod_{j=1}^{n_i}o_{ij}, a \mapsto \sum_{j=1}^{n_i}\left(\left(\prod_{h=1}^{j-1}x_{ih}\right)\cdot d_{ija}\cdot\prod_{h=j+1}^{n_i}x_{ih}\right)\right)\right)$$

$$= F\mu_X^{ST}\left(\prod_{i=1}^{m}\left(\prod_{j=1}^{n_i}x_{ij}\right), \prod_{i=1}^{m}\prod_{j=1}^{n_i}o_{ij}, a \mapsto \sum_{i=1}^{m}\left(\prod_{k=1}^{i-1}\left(\prod_{l=1}^{n_k}x_{kl}\right)\right)\cdot\right.$$
$$\left.\left(\left(\sum_{j=1}^{n_i}\left(\left(\prod_{h=1}^{j-1}x_{ih}\right)\cdot d_{ija}\cdot\prod_{h=j+1}^{n_i}x_{ih}\right)\right)\cdot\prod_{k=i+1}^{m}\left(\prod_{l=1}^{n_k}x_{kl}\right)\right)\right)$$

$$= \left(\prod_{i=1}^{m}\prod_{j=1}^{n_i}x_{ij}, \prod_{i=1}^{m}\prod_{j=1}^{n_i}o_{ij}, a \mapsto \sum_{i=1}^{m}\sum_{j=1}^{n_i}\left(\left(\left(\prod_{k=1}^{i-1}\prod_{l=1}^{n_k}x_{kl}\prod_{h=1}^{j-1}x_{ih}\right)\cdot\right.\right.\right.$$
$$\left.\left.\left.d_{ija}\cdot\left(\prod_{h=j+1}^{n_i}x_{ih}\right)\prod_{k=i+1}^{m}\prod_{l=1}^{n_k}x_{kl}\right)\right)\right)$$

$$= \left(\prod_{i=1}^{\hat{n}}x_{f(i)}, \prod_{i=1}^{\hat{n}}o_{f(i)}, a \mapsto \sum_{i=1}^{\hat{n}}\left(\left(\left(\prod_{k=1}^{i-1}x_{f(k)}\right)\cdot d_{f(k)a}\cdot\prod_{k=i+1}^{\hat{n}}x_{f(k)}\right)\right)\right)$$

$$= \lambda_X\left(\prod_{i=1}^{\hat{n}}(x_{f(i)}, o_{f(i)}, a \mapsto d_{f(i)a})\right)$$

$$= (\lambda_X \circ \mu_{FX}^{T})\left(\prod_{i=1}^{m}\left[\prod_{j=1}^{n_i}(x_{ij}, o_{ij}, a \mapsto d_{ija})\right]\right)$$

3. The unit axiom for both the Brzozowski and Leibniz product rules are established as follows:

$$(\lambda_X \circ \eta_{FX}^{T})(x, o, a \mapsto d_a)$$
$$= \lambda_X\left((x, o, a \mapsto d_a)\right)$$
$$= (((x)), o, a \mapsto ((d_a)))$$
$$= \eta_{TX}^{S}(x, o, a \mapsto d_a)$$

4. The counit axiom for the Brzozowski product rule:

$$(\epsilon_{STX} \circ \lambda_X)\left(\prod_{i=1}^{n} (x_i, o_i, a \mapsto d_i)\right)$$

$$= \epsilon_{STX}\left(\left(\prod_{i=1}^{n} x_i\right), \prod_{i=1}^{n} o_i, a \mapsto \sum_{i=1}^{n}\left(\prod_{k=1}^{i-1} o_i\right) \times \left(d_{ia} \cdot \prod_{k=i+1}^{n} x_i\right)\right)$$

$$= \left(\prod_{i=1}^{n} x_i\right)$$

$$= \eta_{TX}^S\left(\prod_{i=1}^{n} x_i\right)$$

$$= (\eta_{TX}^S \circ \epsilon_{TX})\left(\prod_{i=1}^{n} (x_i, o_i, a \mapsto d_i)\right)$$

5. The counit axiom for the Leibniz product rule:

$$(\epsilon_{STX} \circ \lambda_X)\left(\prod_{i=1}^{n} (x_i, o_i, a \mapsto d_i)\right)$$

$$= \epsilon_{STX}\left(\left(\prod_{i=1}^{n} x_i\right), \prod_{i=1}^{n} o_i, a \mapsto \sum_{i=1}^{n}\left(\left(\prod_{k=1}^{i-1} o_i\right) \cdot d_{ia} \cdot \prod_{k=i+1}^{n} x_i\right)\right)$$

$$= \left(\prod_{i=1}^{n} x_i\right)$$

$$= \eta_{TX}^S\left(\prod_{i=1}^{n} x_i\right)$$

$$= (\eta_{TX}^S \circ \epsilon_{TX})\left(\prod_{i=1}^{n} (x_i, o_i, a \mapsto d_i)\right)$$

Proof. (of Theorem 5)

First, assume that λ^1, λ^2 and λ^0 satisfy the coherence axiom. Then the following diagram commutes (the commutation of the combination of this and the next

diagram was proved by the Prolog program):

$$
\begin{array}{l}
STSTF \xrightarrow{STS\lambda^1} STSFST \quad SSFSTST \xrightarrow{S\lambda^2STST} SFSSTST \xrightarrow{SF\mu^STST} SFSTST \\[2pt]
\Big\downarrow{\scriptstyle S\lambda^0TF\ S\lambda^0FST} \quad \nearrow{\scriptstyle SS\lambda^1ST} \quad \Big\downarrow{\scriptstyle SSFS\lambda^0T} \quad \Big\downarrow{\scriptstyle SFSS\lambda^0T}\ \ {\scriptstyle SFS\lambda^0T} \\[2pt]
SSTTF \xrightarrow{SST\lambda^1} SSTFST \quad SSFSSTT \xrightarrow{S\lambda^2SSTT} SFSSSTT \xrightarrow{SF\mu^SSTT} SFSSTT \\[2pt]
\qquad\qquad\qquad {\scriptstyle SSF\mu^STT}\ \Big\downarrow \qquad {\scriptstyle SFS\mu^STT}\ \Big\downarrow \qquad {\scriptstyle SF\mu^STT}\ \Big\downarrow \\[2pt]
{\scriptstyle SS\mu^TF}\ \Big\downarrow \qquad SSFSTT \xrightarrow{S\lambda^2STT} SFSSTT \xrightarrow{SF\mu^STT} SFSTT \\[2pt]
\qquad \searrow{\scriptstyle SSFS\mu^T} \qquad\qquad {\scriptstyle \lambda^2SSTT}\ \Big\downarrow \qquad {\scriptstyle \lambda^2STT}\ \Big\downarrow \\[2pt]
SSTF \xrightarrow{SS\lambda^1} SSFST \quad {\scriptstyle \mu^SFSTT}\ \Big\downarrow \qquad FSSSTT \xrightarrow{FS\mu^STT} FSSTT \\[2pt]
\qquad\qquad \searrow{\scriptstyle \mu^SFST} \qquad\qquad {\scriptstyle F\mu^SSTT}\ \Big\downarrow \qquad {\scriptstyle F\mu^STT}\ \Big\downarrow \\[2pt]
{\scriptstyle \mu^STF}\ \Big\downarrow \qquad SFSTT \xrightarrow{\lambda^2STT} FSSTT \xrightarrow{F\mu^STT} FSTT \\[2pt]
\qquad\qquad {\scriptstyle SFS\mu^T}\ \Big\downarrow \quad {\scriptstyle FSS\mu^T}\ \Big\downarrow \qquad {\scriptstyle FS\mu^T}\ \Big\downarrow \\[2pt]
STF \xrightarrow{S\lambda^1} SFST \xrightarrow{\lambda^2ST} FSST \xrightarrow{F\mu^ST} FST
\end{array}
$$
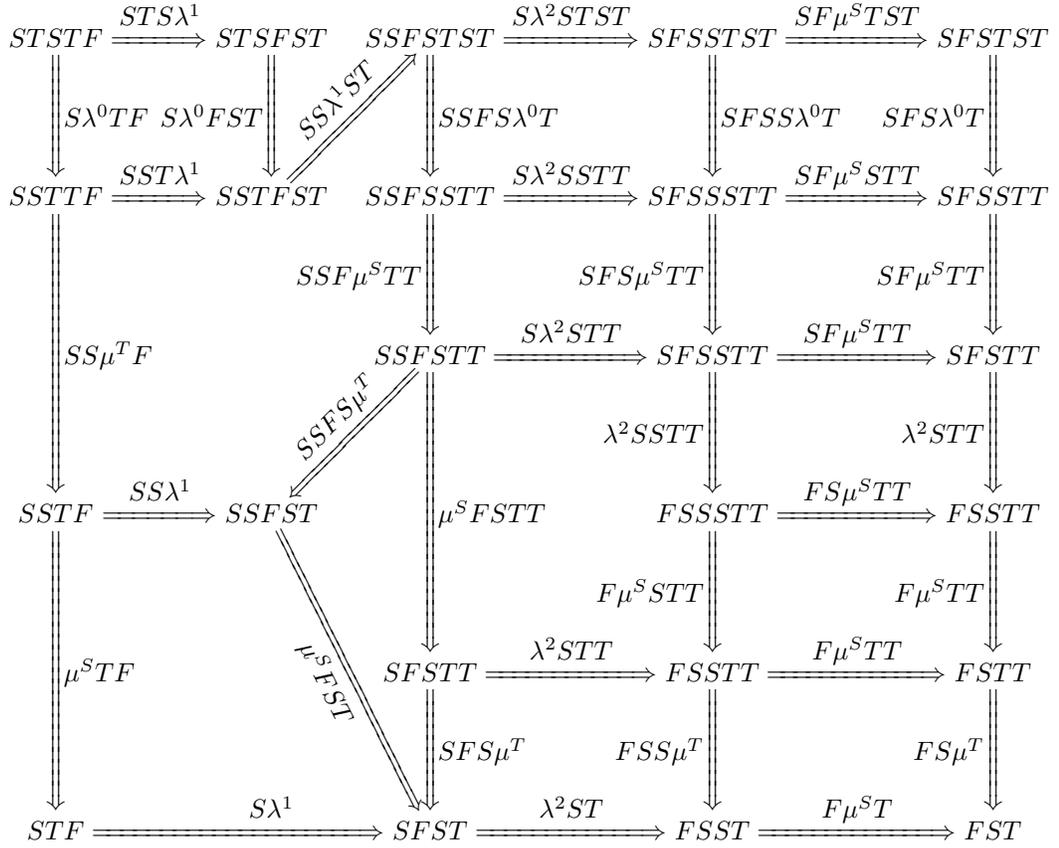
and thus we have:

$$
FS\mu^T \circ F\mu^STT \circ \lambda^2STT \circ SF\mu^STT \circ SFS\lambda^0T \circ SF\mu^STST\circ
$$
$$
S\lambda^2STST \circ SS\lambda^1ST \circ S\lambda^0FST \circ STS\lambda^1
$$
$$
= F\mu^ST \circ \lambda^2ST \circ S\lambda^1 \circ \mu^STF \circ SS\mu^TF \circ S\lambda^0TF
$$

Next, observe the following diagram also commutes,

$$
\begin{array}{ccccc}
STSFST & \xrightarrow{ST\lambda^2 ST} & STFSST & \xrightarrow{STF\mu^S T} & STFST \\
\Big\downarrow{\scriptstyle S\lambda^0 FST} & & \Big\downarrow{\scriptstyle S\lambda^1 SST} & & \Big\downarrow{\scriptstyle S\lambda^1 ST} \\
SSTFST & & SFSTSST & & \\
\Big\downarrow{\scriptstyle SS\lambda^1 ST} & & & \searrow{\scriptstyle SFST\mu^S T} & \\
SSFSTST & & & & \\
\Big\downarrow{\scriptstyle S\lambda^2 STST} & & {\scriptstyle SFS\lambda^0 ST} & & \\
SFSSTST & & & & \\
\Big\downarrow{\scriptstyle SF\mu^S TST} & & & & SFSTST \\
SFSTST & \xleftarrow{SF\mu^S TST} & SFSSTST & & \\
\Big\downarrow{\scriptstyle SFS\lambda^0 T} & & \Big\downarrow{\scriptstyle SFSS\lambda^0 T} & \searrow{\scriptstyle SFS\lambda^0 T} & \\
SFSSTT & \xleftarrow{SF\mu^S STT} & SFSSSTT & & {\scriptstyle \lambda^2 STST} \\
\Big\downarrow{\scriptstyle SF\mu^S TT} & & \Big\downarrow{\scriptstyle SFS\mu^S TT} & & \\
SFSTT & \xleftarrow{SF\mu^S TT} & SFSSTT & & \\
\Big\downarrow{\scriptstyle \lambda^2 STT} & & \Big\downarrow{\scriptstyle \lambda^2 SSTT} & & \\
FSSTT & \xleftarrow{FS\mu^S TT} & FSSSTT & \xleftarrow{FSS\lambda^0 T} & FSSTST \\
\Big\downarrow{\scriptstyle F\mu^S TT} & & \Big\downarrow{\scriptstyle F\mu^S STT} & & \Big\downarrow{\scriptstyle F\mu^S TST} \\
FSTT & \xleftarrow{F\mu^S TT} & FSSTT & \xleftarrow{FS\lambda^0 T} & FSTST \\
\Big\downarrow{\scriptstyle FS\mu^T} & & \Big\downarrow{\scriptstyle FSS\mu^T} & & \\
FST & \xleftarrow{F\mu^S T} & FSST & &
\end{array}
$$
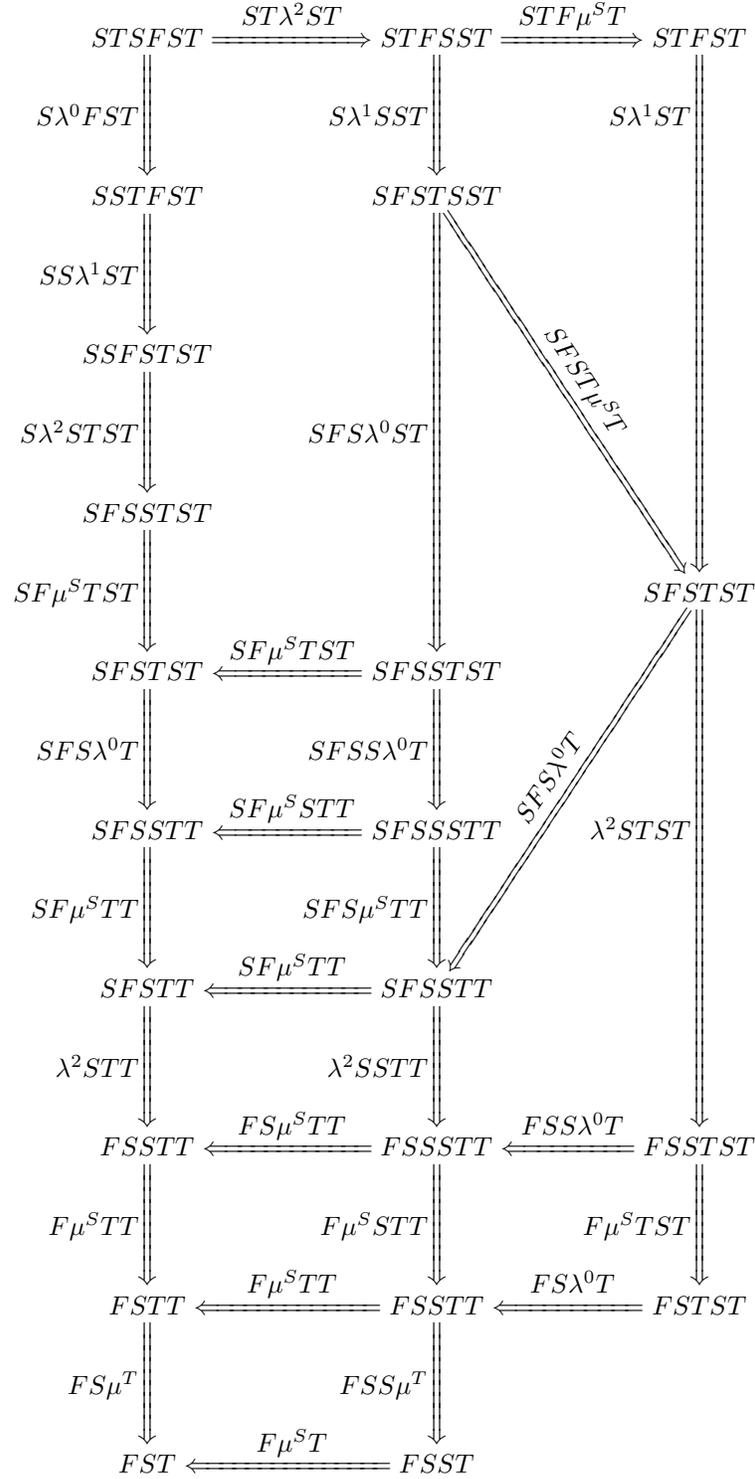
giving:

$$FS\mu^T \circ F\mu^S TT \circ \lambda^2 STT \circ SF\mu^S TT \circ SFS\lambda^0 T \circ SF\mu^S TST \circ$$
$$S\lambda^2 STST \circ SS\lambda^1 ST \circ S\lambda^0 FST$$
$$= F\mu^S T \circ FSS\mu^T \circ FS\lambda^0 T \circ F\mu^S TST \circ \lambda^2 STST \circ S\lambda^1 ST \circ STF\mu^S T \circ ST\lambda^2 ST$$

Combining this, we get

$$\hat{\lambda} \circ \mu^{ST} F$$
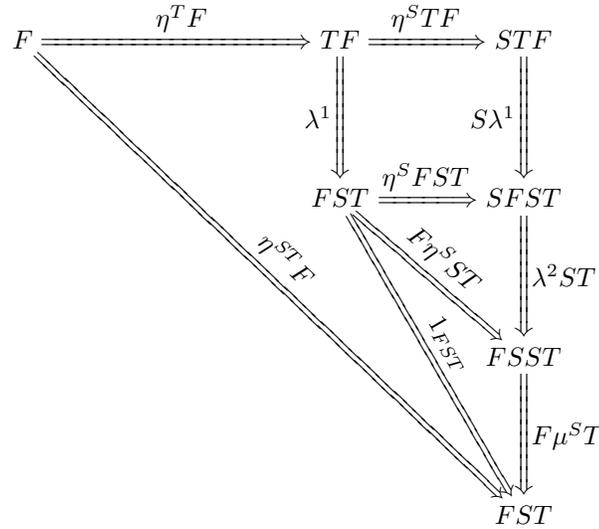$$= F\mu^S T \circ \lambda^2 ST \circ S\lambda^1 \circ \mu^S TF \circ SS\mu^T F \circ S\lambda^0 TF$$
$$= FS\mu^T \circ F\mu^S TT \circ \lambda^2 STT \circ SF\mu^S TT \circ SFS\lambda^0 T \circ SF\mu^S TST \circ$$
$$S\lambda^2 STST \circ SS\lambda^1 ST \circ S\lambda^0 FST \circ STS\lambda^1$$
$$= F\mu^S T \circ FSS\mu^T \circ FS\lambda^0 T \circ F\mu^S TST \circ \lambda^2 STST \circ S\lambda^1 ST \circ STF\mu^S T \circ ST\lambda^2 ST$$
$$= F\mu^{ST} \circ \hat{\lambda} ST \circ ST\hat{\lambda}$$

so the main diagram for the distributive law $\hat{\lambda}$ commutes.

Next, we check that the unit diagram commutes. This follows from the following diagram:

Finally, if (F, ϵ) is a copointed functor and λ^1 and λ^2 are copointed laws, we check that $\epsilon T \circ \lambda = T\epsilon$:

$$
\begin{array}{ccccccc}
STF & \xRightarrow{S\lambda^1} & SFST & \xRightarrow{\lambda^2 ST} & FSST & \xRightarrow{F\mu^S T} & FST \\
& \searrow{\scriptstyle ST\epsilon} & & \searrow{\scriptstyle S\epsilon ST} & \Downarrow{\scriptstyle \epsilon SST} & & \Vert{\scriptstyle \epsilon ST} \\
& & ST & \xRightarrow{S\eta^S T} & SST & & FST \\
& & & \searrow{\scriptstyle 1_{ST}} & \searrow{\scriptstyle \mu^S T} & & \Downarrow{\scriptstyle \epsilon ST} \\
& & & & & ST &
\end{array}
$$

This concludes one direction of the proof. For the converse, consider the following diagram,

$$
\begin{array}{ccc}
TSF \xrightarrow{\eta^S TSF} STSF \xrightarrow{STS\eta^T F} STSTF \\
\end{array}
$$

The diagram contains the following objects and morphisms:

- $TSF \xrightarrow{\eta^S TSF} STSF \xrightarrow{STS\eta^T F} STSTF$
- $STSF \xrightarrow{STSF\eta^{ST}} STSFST$
- $STSTF \xrightarrow{STS\lambda^1} STSFST$
- $TSF \xrightarrow{T\lambda^2} TFS$
- $STSF \xrightarrow{ST\lambda^2} STFS$
- $STSFST \xrightarrow{ST\lambda^2 ST} STFSST$
- $TFS \xrightarrow{\eta^S TFS} STFS \xrightarrow{STFS\eta^{ST}} STFSST$
- $STFS \xrightarrow{STFS\eta^T} STFST$
- $STFST \xrightarrow{STFS\eta^{ST}} STFSST$
- $STFSST \xrightarrow{STF\mu^S T} STFST$
- $STFST \xrightarrow{1_{STFST}} STFST$
- $TFS \xrightarrow{\lambda^1 S} FSTS$
- $STFS \xrightarrow{S\lambda^1 S} SFSTS$
- $STFST \xrightarrow{S\lambda^1 ST} SFSTST$
- $FSTS \xrightarrow{\eta^S FSTS} SFSTS \xrightarrow{SFSTS\eta^T} SFSTST$
- $FSTS \xrightarrow{F\eta^S STS} FSSTS$
- $SFSTS \xrightarrow{\lambda^2 STS} FSSTS$
- $SFSTST \xrightarrow{\lambda^2 STST} FSSTST$
- $FSSTS \xrightarrow{FSSTS\eta^T} FSSTST$
- $FSTS \xrightarrow{1_{FSTS}} FSTS$
- $FSSTS \xrightarrow{F\mu^S TS} FSTS$
- $FSSTST \xrightarrow{F\mu^S TST} FSTST$
- $FSTS \xrightarrow{FSTS\eta^T} FSTST$
- $FSTS \xrightarrow{FS\lambda^0} FSST$
- $FSTST \xrightarrow{FS\lambda^0 T} FSSTT$
- $FSST \xrightarrow{FSST\eta^T} FSSTT$
- $FSST \xrightarrow{F\mu^S T} FST$
- $FSSTT \xrightarrow{F\mu^S TT} FSTT$
- $FST \xrightarrow{FST\eta^T} FSTT$
- $FST \xrightarrow{1_{FST}} FST$
- $FSTT \xrightarrow{FS\mu^T} FST$

which establishes that

$$
FS\mu^{ST} \circ \hat{\lambda}ST \circ ST\hat{\lambda} \circ STS\eta^{T}F \circ \eta^{S}TSF
$$
$$
= FS\mu^{T} \circ F\mu^{S}TT \circ FS\lambda^{0}T \circ F\mu^{S}TST \circ \lambda^{2}STST \circ S\lambda^{1}ST\circ
$$
$$
STF\mu^{S}T \circ ST\lambda^{2}ST \circ STS\lambda^{1} \circ STS\eta^{T}F \circ \eta^{S}TSF
$$
$$
= F\mu^{S}T \circ FS\lambda^{0} \circ \lambda^{1}S \circ T\lambda^{2}
$$

and the following diagram,

which establishes that:

$$
S\mu^{ST}F \circ STS\eta^{T}F \circ \eta^{S}TSF
$$
$$
= S\mu^{T}F \circ \mu^{S}TTF \circ S\lambda^{0}TF \circ STS\eta^{T}F \circ \eta^{S}TSF
$$
$$
= \lambda^{0}F
$$

Combining this, we obtain

$$
F\mu^{S}T \circ FS\lambda^{0} \circ \lambda^{1}S \circ T\lambda^{2}
$$
$$
= FS\mu^{ST} \circ \hat{\lambda}ST \circ ST\hat{\lambda} \circ STS\eta^{T}F \circ \eta^{S}TSF
$$
$$
= \hat{\lambda} \circ FS\mu^{ST} \circ STS\eta^{T}F \circ \eta^{S}TSF
$$
$$
= \hat{\lambda} \circ \lambda^{0}F
$$

completing the proof.

Proof. (of Proposition 6) We verify the coherence axiom for each of the product rules.

1. Brzozowski product rule:

$$(F\mu_{TX}^S \circ FS\lambda_X^0 \circ \lambda_{SX}^1 \circ T\lambda_X^2)\left(\prod_{i=1}^{n}\left(\sum_{j=1}^{m_i} p_{ij} \times \left(x_{ij}, o_{ij}, a \mapsto d_{ija}\right)\right)\right)$$

$$= (F\mu_{TX}^S \circ FS\lambda_X^0 \circ \lambda_{SX}^1)\left(\prod_{i=1}^{n}\left(\left(\sum_{j=1}^{m_i} p_{ij} \times x_{ij}\right), \sum_{j=1}^{m_i} p_{ij}o_{ij}, a \mapsto \left(\sum_{j=1}^{m_i} p_{ij} \times d_{ija}\right)\right)\right)$$

$$= (F\mu_{TX}^S \circ FS\lambda_X^0)\left(\left(\left(\prod_{i=1}^{n}\left(\sum_{j=1}^{m_i} p_{ij} \times x_{ij}\right)\right)\right), \prod_{i=1}^{n}\sum_{j=1}^{m_i} p_{ij}o_{ij}, a \mapsto \right.$$
$$\left. \sum_{i=1}^{n}\left(\prod_{k=1}^{i-1}\sum_{j=1}^{m_k} p_{kj}o_{kj} \times \left(\left(\sum_{j=1}^{m_i} p_{ij} \times d_{ija}\right) \cdot \prod_{k=i+1}^{n}\left(\sum_{j=1}^{m_k} p_{kj} \times x_{kj}\right)\right)\right)\right)$$

$$= F\mu_{TX}^S\left(\left[\sum_{j_1=1}^{m_1}\cdots\sum_{j_n=1}^{m_n}\left(\prod_{i=1}^{n} p_{ij_i}\right) \times \left(\prod_{i=1}^{n} x_{ij_i}\right)\right], \prod_{i=1}^{n}\sum_{j=1}^{m_i} p_{ij}o_{ij}, a \mapsto \right.$$
$$\left. \sum_{i=1}^{n}\left(\prod_{k=1}^{i-1}\sum_{j=1}^{m_k} p_{kj}o_{kj} \times \left[\sum_{j_i=1}^{m_i}\cdots\sum_{j_n=1}^{m_n}\left(\prod_{k=i}^{n} p_{kj_k}\right) \times \left(d_{ij_ia}\prod_{k=i+1}^{n} x_{kj_k}\right)\right]\right)\right)$$

$$= \left(\sum_{j_1=1}^{m_1}\cdots\sum_{j_n=1}^{m_n}\left(\prod_{i=1}^{n} p_{ij_i}\right) \times \left(\prod_{i=1}^{n} x_{ij_i}\right), \prod_{i=1}^{n}\sum_{j=1}^{m_i} p_{ij}o_{ij}, a \mapsto \right.$$
$$\left. \sum_{i=1}^{n}\sum_{j_1=1}^{m_i}\cdots\sum_{j_n=1}^{m_n}\left(\prod_{k=1}^{i-1}\sum_{j=1}^{m_k} p_{kj}o_{kj}\prod_{k=i}^{n} p_{kj_k}\right) \times \left(d_{ij_ia}\prod_{k=i+1}^{n} x_{kj_k}\right)\right)$$

$$= \left(\sum_{j_1=1}^{m_1}\cdots\sum_{j_n=1}^{m_n}\left(\prod_{i=1}^{n} p_{ij_i}\right) \times \left(\prod_{i=1}^{n} x_{ij_i}\right), \prod_{i=1}^{n}\sum_{j=1}^{m_i} p_{ij}o_{ij}, a \mapsto \right.$$
$$\left. \sum_{i=1}^{n}\sum_{j_i=1}^{m_i}\cdots\sum_{j_n=1}^{m_n}\left(\sum_{j_1=1}^{m_1}\cdots\sum_{j_{i-1}=1}^{m_{i-1}}\prod_{k=1}^{i-1} p_{kj_k}o_{kj_k}\prod_{k=i}^{n} p_{kj_k}\right) \times \left(d_{ij_ia}\prod_{k=i+1}^{n} x_{kj_k}\right)\right)$$

$$= \left(\sum_{j_1=1}^{m_1}\cdots\sum_{j_n=1}^{m_n}\left(\prod_{i=1}^{n} p_{ij_i}\right) \times \left(\prod_{i=1}^{n} x_{ij_i}\right), \sum_{j_1=1}^{m_1}\cdots\sum_{j_n=1}^{m_n}\prod_{i=1}^{n} p_{ij_i}o_{ij_i}, a \mapsto \right.$$
$$\left. \sum_{j_1=1}^{m_1}\cdots\sum_{j_n=1}^{m_n}\sum_{i=1}^{n}\left(\prod_{k=1}^{n} p_{kj_k}\prod_{k=1}^{i-1} o_{kj_k}\right) \times \left(d_{ij_ia}\prod_{k=l+1}^{n} x_{kj_k}\right)\right)$$

$$= F\mu_{TX}^S\left(\sum_{j_1=1}^{m_1}\cdots\sum_{j_n=1}^{m_n}\left(\prod_{i=1}^{n} p_{ij_i}\right) \times \left[\left(\prod_{i=1}^{n} x_{ij_i}\right)\right], \sum_{j_1=1}^{m_1}\cdots\sum_{j_n=1}^{m_n}\prod_{i=1}^{n} p_{ij_i}\prod_{i=1}^{n} o_{ij_i}, a \mapsto \right.$$
$$\left. \sum_{j_1=1}^{m_1}\cdots\sum_{j_n=1}^{m_n}\left(\prod_{i=1}^{n} p_{ij_i}\right) \times \left[\sum_{i=1}^{n}\left(\prod_{k=1}^{i-1} o_{kj_k}\right) \times \left(d_{ij_ia}\prod_{k=i+1}^{n} x_{kj_k}\right)\right]\right)$$

$$= (F\mu_{TX}^S \circ \lambda_{STX}^2)\left(\sum_{j_1=1}^{m_1}\cdots\sum_{j_n=1}^{m_n}\left(\prod_{i=1}^{n} p_{ij_i}\right) \times \left(\left(\left(\prod_{i=1}^{n} x_{ij_i}\right)\right), \prod_{i=1}^{n} o_{ij_i}, a \mapsto \right.\right.$$
$$\left.\left. \sum_{i=1}^{n}\left(\prod_{k=1}^{i-1} o_{kj_k}\right) \times \left(d_{ij_ia} \cdot \prod_{k=i+1}^{n} x_{kj_k}\right)\right)\right)$$

$$= (F\mu_{TX}^S \circ \lambda_{STX}^2 \circ S\lambda_X^1)\left(\sum_{j_1=1}^{m_1}\cdots\sum_{j_n=1}^{m_n}\left(\prod_{i=1}^{n} p_{ij_i}\right) \times \left(\prod_{i=1}^{n}\left(x_{ij_i}, o_{ij_i}, a \mapsto d_{ij_ia}\right)\right)\right)$$

$$= (F\mu_{TX}^S \circ \lambda_{STX}^2 \circ S\lambda_X^1 \circ \lambda_{FX}^0)\left(\prod_{i=1}^{n}\left(\sum_{j=1}^{m_i} p_{ij} \times \left(x_{ij}, o_{ij}, a \mapsto d_{ija}\right)\right)\right)$$

2. Leibniz product rule:
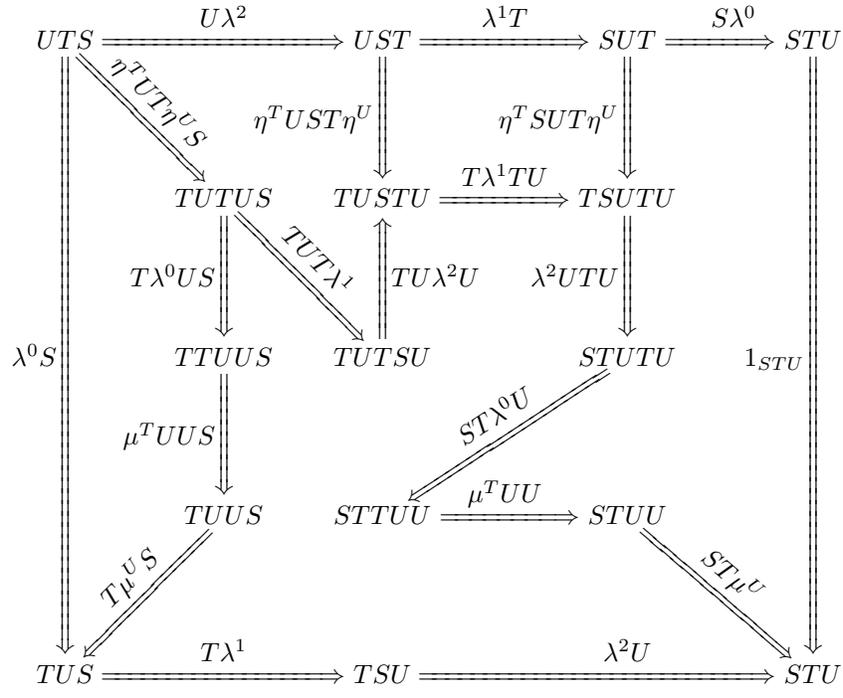
$$(F\mu_{TX}^S \circ FS\lambda_X^0 \circ \lambda_{SX}^1 \circ T\lambda_X^2)\left(\prod_{i=1}^n \left(\sum_{j=1}^{m_i} p_{ij} \times \left(x_{ij}, o_{ij}, a \mapsto d_{ija}\right)\right)\right)$$

$$= (F\mu_{TX}^S \circ FS\lambda_X^0 \circ \lambda_{SX}^1)\left(\prod_{i=1}^n \left(\left(\sum_{j=1}^{m_i} p_{ij} \times x_{ij}\right), \sum_{j=1}^{m_i} p_{ij}o_{ij}, a \mapsto \left(\sum_{j=1}^{m_i} p_{ij} \times d_{ija}\right)\right)\right)$$

$$= (F\mu_{TX}^S \circ FS\lambda_X^0)\left(\left(\left(\prod_{i=1}^n \left(\sum_{j=1}^{m_i} p_{ij} \times x_{ij}\right)\right)\right), \prod_{i=1}^n \sum_{j=1}^{m_i} p_{ij}o_{ij}, a \mapsto \right.$$

$$\left. \sum_{i=1}^n \left(\left(\left(\prod_{k=1}^{i-1}\left(\sum_{j=1}^{m_k} p_{kj} \times x_{kj}\right)\right) \cdot \left(\sum_{j=1}^{m_i} p_{ij} \times d_{ija}\right) \cdot \prod_{k=i+1}^n \left(\sum_{j=1}^{m_k} p_{kj} \times x_{kj}\right)\right)\right)\right)$$

$$= F\mu_{TX}^S\left(\left[\sum_{j_1=1}^{m_1} \cdots \sum_{j_n=1}^{m_n} \left(\prod_{i=1}^n p_{ij_i}\right) \times \left(\prod_{i=1}^n x_{ij_i}\right)\right], \prod_{i=1}^n \sum_{j=1}^{m_i} p_{ij}o_{ij}, a \mapsto \right.$$

$$\left. \sum_{i=1}^n \left[\sum_{j_i=1}^{m_i} \cdots \sum_{j_n=1}^{m_n} \left(\prod_{k=i}^n p_{kj_k}\right) \times \left(\left(\prod_{k=1}^{i-1} x_{kj_k}\right) \cdot d_{ij_i a} \cdot \prod_{k=i+1}^n x_{kj_k}\right)\right]\right)$$

$$= \left(\sum_{j_1=1}^{m_1} \cdots \sum_{j_n=1}^{m_n} \left(\prod_{i=1}^n p_{ij_i}\right) \times \left(\prod_{i=1}^n x_{ij_i}\right), \prod_{i=1}^n \sum_{j=1}^{m_i} p_{ij}o_{ij}, a \mapsto \right.$$

$$\left. \sum_{i=1}^n \sum_{j_i=1}^{m_i} \cdots \sum_{j_n=1}^{m_n} \left(\prod_{k=i}^n p_{kj_k}\right) \times \left(\left(\prod_{k=1}^{i-1} x_{kj_k}\right) \cdot d_{ij_i a} \cdot \prod_{k=i+1}^n x_{kj_k}\right)\right)$$

$$= F\mu_{TX}^S\left(\sum_{j_1=1}^{m_1} \cdots \sum_{j_n=1}^{m_n} \left(\prod_{i=1}^n p_{ij_i}\right) \times \left[\left(\prod_{i=1}^n x_{ij_i}\right)\right], \sum_{j_1=1}^{m_1} \cdots \sum_{j_n=1}^{m_n} \prod_{i=1}^n p_{ij_i} \prod_{i=1}^n o_{ij_i}, a \mapsto \right.$$

$$\left. \sum_{j_1=1}^{m_1} \cdots \sum_{j_n=1}^{m_n} \left(\prod_{i=1}^n p_{ij_i}\right) \times \left[\sum_{i=1}^n \left(\left(\prod_{k=1}^{i-1} x_{kj_k}\right) \cdot d_{ij_i a} \cdot \prod_{k=i+1}^n x_{kj_k}\right)\right]\right)$$

$$= (F\mu_{TX}^S \circ \lambda_{STX}^2)\left(\sum_{j_1=1}^{m_1} \cdots \sum_{j_n=1}^{m_n} \left(\prod_{i=1}^n p_{ij_i}\right) \times \left(\left(\left(\prod_{i=1}^n x_{ij_i}\right)\right), \prod_{i=1}^n o_{ij_i}, a \mapsto \right.\right.$$

$$\left.\left. \sum_{i=1}^n \left(\left(\prod_{k=1}^{i-1} x_{kj_k}\right) \cdot d_{ij_i a} \cdot \prod_{k=i+1}^n x_{kj_k}\right)\right)\right)$$

$$= (F\mu_{TX}^S \circ \lambda_{STX}^2 \circ S\lambda_X^1)\left(\sum_{j_1=1}^{m_1} \cdots \sum_{j_n=1}^{m_n} \left(\prod_{i=1}^n p_{ij_i}\right) \times \left(\prod_{i=1}^n \left(x_{ij_i}, o_{ij_i}, a \mapsto d_{ij_i a}\right)\right)\right)$$

$$= (F\mu_{TX}^S \circ \lambda_{STX}^2 \circ S\lambda_X^1 \circ \lambda_{FX}^0)\left(\prod_{i=1}^n \left(\sum_{j=1}^{m_i} p_{ij} \times \left(x_{ij}, o_{ij}, a \mapsto d_{ija}\right)\right)\right)$$
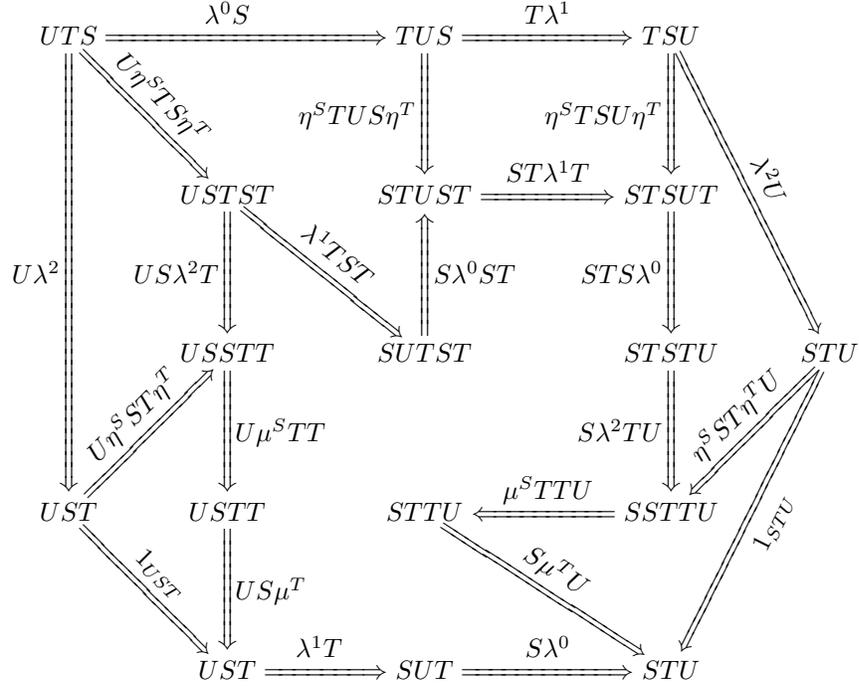
3. Hadamard product rule:

$$(F\mu_{TX}^S \circ FS\lambda_X^0 \circ \lambda_{SX}^1 \circ T\lambda_X^2)\left(\prod_{i=1}^{n}\left(\sum_{j=1}^{m_i} p_{ij} \times \left(o_{ij}, a \mapsto d_{ija}\right)\right)\right)$$

$$= (F\mu_{TX}^S \circ FS\lambda_X^0 \circ \lambda_{SX}^1)\left(\prod_{i=1}^{n}\left(\sum_{j=1}^{m_i} p_{ij}o_{ij}, a \mapsto \left(\sum_{j=1}^{m_i} p_{ij} \times d_{ija}\right)\right)\right)$$

$$= (F\mu_{TX}^S \circ FS\lambda_X^0)\left(\prod_{i=1}^{n}\sum_{j=1}^{m_i} p_{ij}o_{ij}, a \mapsto \left(\left(\prod_{i=1}^{n}\left(\sum_{j=1}^{m_i} p_{ij} \times d_{ija}\right)\right)\right)\right)$$

$$= F\mu_{TX}^S\left(\prod_{i=1}^{n}\sum_{j=1}^{m_i} p_{ij}o_{ij}, a \mapsto \left(\left[\sum_{j_1=1}^{m_1}\cdots\sum_{j_n=1}^{m_n}\prod_{i=1}^{n} p_{ij_i} \times \left(\prod_{i=1}^{n} d_{ij_ia}\right)\right]\right)\right)$$

$$= \left(\prod_{i=1}^{n}\sum_{j=1}^{m_i} p_{ij}o_{ij}, a \mapsto \left(\sum_{j_1=1}^{m_1}\cdots\sum_{j_n=1}^{m_n}\prod_{i=1}^{n} p_{ij_i} \times \left(\prod_{i=1}^{n} d_{ij_ia}\right)\right)\right)$$

$$= F\mu_{TX}^S\left(\sum_{j_1=1}^{m_1}\cdots\sum_{j=i_n}^{m_n}\prod_{i=1}^{n} p_{ij_i}o_{ij_i}, a \mapsto \left(\sum_{j_1=1}^{m_1}\cdots\sum_{j_n=1}^{m_n}\prod_{i=1}^{n} p_{ij_i} \times \left[\left(\prod_{i=1}^{n} d_{ij_ia}\right)\right]\right)\right)$$

$$= (F\mu_{TX}^S \circ \lambda_{STX}^2)\left(\sum_{j_1=1}^{m_1}\cdots\sum_{j_n=1}^{m_n}\left(\prod_{i=1}^{n} p_{ij_i}\right) \times \left(\prod_{i=1}^{n} o_{ij_i}, a \mapsto \left(\left(\prod_{i=1}^{n} d_{ij_ia}\right)\right)\right)\right)$$

$$= (F\mu_{TX}^S \circ \lambda_{STX}^2 \circ S\lambda_X^1)\left(\sum_{j_1=1}^{m_1}\cdots\sum_{j_n=1}^{m_n}\left(\prod_{i=1}^{n} p_{ij_i}\right) \times \left(\prod_{i=1}^{n}\left(o_{ij_i}, a \mapsto d_{ij_ia}\right)\right)\right)$$

$$= (F\mu_{TX}^S \circ \lambda_{STX}^2 \circ S\lambda_X^1 \circ \lambda_{FX}^0)\left(\prod_{i=1}^{n}\left(\sum_{j=1}^{m_i} p_{ij} \times \left(o_{ij}, a \mapsto d_{ija}\right)\right)\right)$$

Proof. (of Proposition 11) The implications $1 \Rightarrow 2$ and $1 \Rightarrow 3$ follow directly from Proposition 10.

The implication $2 \Rightarrow 1$ is proved by the following diagram:

$$
\begin{array}{ccccccc}
UTS & \xrightarrow{U\lambda^2} & UST & \xrightarrow{\lambda^1 T} & SUT & \xrightarrow{S\lambda^0} & STU
\end{array}
$$

$\eta^T UT\eta^U S$ $\qquad$ $\eta^T UST\eta^U$ $\qquad$ $\eta^T SUT\eta^U$

$TUTUS$ $\qquad$ $TUSTU \xrightarrow{T\lambda^1 TU} TSUTU$

$TUT\eta^U$

$T\lambda^0 US$ $\qquad$ $TU\lambda^2 U$ $\qquad$ $\lambda^2 UTU$

$\lambda^0 S$ $\qquad$ $TTUUS$ $\qquad$ $TUTSU$ $\qquad$ $STUTU$ $\qquad$ 1_{STU}

$\mu^T UUS$ $\qquad$ $ST\lambda^0 U$

$TUUS$ $\qquad$ $STTUU \xrightarrow{\mu^T UU} STUU$

$T\mu^U S$ $\qquad\qquad\qquad\qquad\qquad\qquad$ $ST\mu^U$

$$
\begin{array}{ccccc}
TUS & \xrightarrow{T\lambda^1} & TSU & \xrightarrow{\lambda^2 U} & STU
\end{array}
$$

The implication $3 \Rightarrow 1$ is proved by the following diagram:

$$
\begin{array}{ccccc}
UTS & \xrightarrow{\lambda^0 S} & TUS & \xrightarrow{T\lambda^1} & TSU
\end{array}
$$

$U\eta^S TS\eta^T$ $\qquad$ $\eta^S TUS\eta^T$ $\qquad$ $\eta^S TSU\eta^T$ $\qquad$ $\lambda^2 U$

$USTST$ $\qquad$ $STUST \xrightarrow{ST\lambda^1 T} STSUT$

$\lambda^1 TST$

$U\lambda^2$ $\quad$ $US\lambda^2 T$ $\qquad$ $S\lambda^0 ST$ $\qquad$ $STS\lambda^0$

$USSTT$ $\qquad$ $SUTST$ $\qquad$ $STSTU$ $\qquad$ STU

$U\eta^S ST\eta^T$ $\qquad$ $U\mu^S TT$ $\qquad$ $S\lambda^2 TU$ $\qquad$ $\eta^S ST\eta^T U$

UST $\quad$ $USTT$ $\qquad$ $STTU \xleftarrow{\mu^S TTU} SSTTU$

1_{UST} $\qquad$ $US\mu^T$ $\qquad$ $S\mu^T U$ $\qquad$ 1_{STU}

$$
\begin{array}{ccccc}
UST & \xrightarrow{\lambda^1 T} & SUT & \xrightarrow{S\lambda^0} & STU
\end{array}
$$

Proof. (*of Proposition 12*) The implication $1 \Rightarrow 2$ follows from the diagram

$$
\begin{array}{ccccccccc}
STSTF & \xrightarrow{STS\lambda^1} & STSFT & \xrightarrow{ST\lambda^2 T} & STFST & \xrightarrow{S\lambda^1 ST} & SFTST & \xrightarrow{\lambda^2 TST} & FSTST \\
\Big\| {\scriptstyle S\lambda^0 TF} && \Big\downarrow {\scriptstyle S\lambda^0 FT} &&&& \Big\downarrow {\scriptstyle SF\lambda^0 T} && \Big\downarrow {\scriptstyle FS\lambda^0 T} \\
SSTTF & \xrightarrow{SST\lambda^1} & SSTFT & \xrightarrow{SS\lambda^1 T} & SSFTT & \xrightarrow{S\lambda^2 TT} & SFSTT & \xrightarrow{\lambda^2 STT} & FSSTT \\
\Big\| {\scriptstyle \mu^S TTF} && \Big\| {\scriptstyle \mu^S TFT} && \Big\| {\scriptstyle \mu^S FTT} &&&& \Big\| {\scriptstyle F\mu^S TT} \\
STTF & \xrightarrow{ST\lambda^1} & STFT & \xrightarrow{S\lambda^1 T} & SFTT & \xrightarrow{\quad\lambda^2 TT\quad} && & FSTT \\
\Big\| {\scriptstyle S\mu^T F} &&&& \Big\| {\scriptstyle SF\mu^T} &&&& \Big\| {\scriptstyle FS\mu^T} \\
STF & \xrightarrow{\qquad S\lambda^1 \qquad} && & SFT & \xrightarrow{\qquad \lambda^2 T \qquad} && & FST
\end{array}
$$

and the diagram:

$$
\begin{array}{ccc}
F & \xrightarrow{\eta^T F} TF \xrightarrow{\eta^S TF} & STF \\
& {\scriptstyle F\eta^T}\searrow \quad \Big\| {\scriptstyle \lambda^1} & \Big\| {\scriptstyle S\lambda^1} \\
& FT \xrightarrow{\eta^S FT} & SFT \\
& {\scriptstyle F\eta^S T}\searrow \quad & \Big\| {\scriptstyle \lambda^2 T} \\
& & FST
\end{array}
$$

If (F, ϵ) is a copointed endofunctor, and λ^1 and λ^2 are copointed laws, the $\lambda^2 T \circ S\lambda^1$ again is a copointed law, because of the diagram:

$$
\begin{array}{ccc}
STF & \xrightarrow{S\lambda^1} SFT \xrightarrow{\lambda^2 T} & FST \\
{\scriptstyle ST\epsilon}\searrow & \Big\| {\scriptstyle S\epsilon T} & \swarrow {\scriptstyle \epsilon ST} \\
& ST &
\end{array}
$$

The implication $2 \Rightarrow 1$ follows from the same argument as the direction $2 \Rightarrow 1$ from Proposition 11, with T replaced by S, with U replaced by T, and with S replaced by F.

Proof. (of Proposition 13) If λ^0, λ^1 and λ^2 make the hexagonal condition commute, then λ^0, $\hat{\lambda}^1 = F\eta^S T \circ \lambda^1$, and λ^2 make the octagonal condition commute, as a result of the following commuting diagram:

$$
\begin{array}{ccccccc}
TSF & \overset{T\lambda^2}{\Longrightarrow} & TFS & \overset{\lambda^1 S}{\Longrightarrow} & FTS & \overset{F\eta^S TS}{\Longrightarrow} & FSTS & \overset{FS\lambda^0}{\Longrightarrow} & FSST \\
\end{array}
$$

with diagonal arrows labelled $F\lambda^0$, $F\eta^S ST$, $F\mu^S T$, 1_{FST}; middle row:

$$
STF \overset{S\lambda^1}{\Longrightarrow} SFT \overset{\lambda^2 T}{\Longrightarrow} FST \overset{1_{FST}}{\Longrightarrow} FST
$$

and vertical maps $\lambda^0 F$, $S\lambda^1$, $SF\eta^S T$, $FS\eta^S T$; bottom row:

$$
SFT \overset{SF\eta^S T}{\Longrightarrow} SFST \overset{\lambda^2 ST}{\Longrightarrow} FSST \overset{F\mu^S T}{\Longrightarrow} FST
$$

Conversely, if λ^0, $\hat{\lambda}^1 = F\eta^S T \circ \lambda^1$, and λ^2 make the octagonal condition commute, then λ^0, λ^1 and λ^2 make the hexagonal condition commute, as a result of the following commuting diagram:

$$
\begin{array}{ccccc}
TSF & \overset{T\lambda^2}{\Longrightarrow} & TFS & \overset{\lambda^1 S}{\Longrightarrow} & FTS & \overset{F\lambda^0}{\Longrightarrow} & FST \\
\end{array}
$$

with $F\eta^S TS$, $FS\lambda^0$, $F\eta^S ST$, $FSTS$, $FSST$; left vertical $\lambda^0 F$ and right 1_{FST}; interior maps $\lambda^2 ST$, $SF\eta^S T$, $FS\eta^S T$, $F\mu^S T$, 1_{FST}; bottom row:

$$
STF \overset{S\lambda^1}{\Longrightarrow} SFT \overset{\lambda^2 T}{\Longrightarrow} FST \overset{1_{FST}}{\Longrightarrow} FST
$$