Adaptives:

Brandon Shapiro   David I. Spivak

Abstract

1   Introduction

2   The Monoidal Double Category $\mathbb{O}\mathbf{rg}$

In [Spi21], the second author defined a category-enriched multicategory $\mathbb{O}\mathbf{rg}$, whose objects are polynomials and whose morphisms are polynomial coalgebras. In this chapter, we describe how $\mathbb{O}\mathbf{rg}$ in fact more naturally takes the form of a monoidal double category, with coalgebras as horizontal morphisms, maps of polynomials as vertical morphisms, and the Dirichlet tensor product $\otimes$ providing the monoidal structure.

2.1 $[p, q]$-coalgebras

We first recall the definitions of the internal hom polynomials $[p, q]$ and concretely describe the category of $[p, q]$-coalgebras, which forms the category of morphisms from p to q in the underlying bicategory of $\mathbb{O}\mathbf{rg}$.

Definition 2.1. For polynomials

$$p = \sum_{I \in p(1)} y^{p[I]} \qquad \text{and} \qquad q = \sum_{J \in q(1)} y^{q[J]},$$

their *internal hom* with respect to the tensor product $\otimes$ is the polynomial

$$[p, q] = \sum_{\phi : p \to q} y^{\sum_{I \in p(1)} q[\phi_1(I)]}.$$

$\diamondsuit$

Using [Spi21, Definition 2.10] applied to $[p, q]$, we get the following definition of $[p, q]$-coalgebras.

Definition 2.2. The category $[p, q]$-**Coalg** has as objects pairs $\mathbf{S} = (S, \beta)$ where S is a set and $\beta : S \to [p, q] \triangleleft S$, and morphisms from $\mathbf{S}$ to $\mathbf{S}'$ given by functions $f : S \to S'$ making the diagram below commute.

$$
\begin{array}{ccc}
S & \xrightarrow{\;\;\beta\;\;} & [p, q] \triangleleft S \\
{\scriptstyle f}\downarrow & & \downarrow{\scriptstyle [p,q]\triangleleft f} \\
S' & \xrightarrow{\;\;\beta'\;\;} & [p, q] \triangleleft S'
\end{array}
$$

$\diamond$

Unwinding this definition, we get that β is determined by a function

$$\beta_0 : S \to \mathbf{Poly}(p, q) = [p, q](1)$$

and for each $s \in S$ a function

$$\beta_s : \sum_{I \in p(1)} q[\beta_0(s)_1(I)] \to S.$$

For each $s \in S$, called a *state*, we call $\beta_0(s)$ the *action of s* and β_s the *update function of s*. A coalgebra map is then a function between the state sets that preserves actions and updates.

When, for each $s \in S$, β_s is the constant function sending everything to s, we say the coalgebra $\mathbf{S}$ has *no updates*, as no output $I \in p(1)$ nor input from $q[\beta_0(s)_1(I)]$ can produce a different state from s.

Example 2.3. A special case of $[p, q]$-coalgebras with no updates are given by maps $\phi \in \mathbf{Poly}(p, q)$. For each such ϕ, there is a coalgebra $\mathbf{S}_\phi$ with a singleton set of states, β_0 sending the point to ϕ, and (by necessity) no updates. $\diamond$

2.2 Composition of coalgebras

We now describe how $[p, q]$-coalgebras behave like morphisms from p to q.

Proposition 2.4. *The categories $[p, q]$-**Coalg** form the categories of morphisms in a bicategory $\mathbb{O}$**rg** with polynomials as objects.*

It is no accident that $\mathbb{O}$**rg** denotes both this bicategory and the categorical operad in [Spi21, Definition 2.19], as both are derived from the monoidal double category $\mathbb{O}$**rg** described in the following sections. For now, we merely present the horizontal identities and composites in this bicategory.

The identity coalgebra in $[p, p]$-**Coalg** is given by the one-state coalgebra $\mathbf{S}_{\mathrm{id}_p}$, which we will also denote by id_p when its status as a coalgebra is clear from context.

Composites are defined as

$$[p, q]\text{-}\mathbf{Coalg} \times [q, r]\text{-}\mathbf{Coalg} \to \Big([p, q] \otimes [q, r]\Big)\text{-}\mathbf{Coalg} \to [p, r]\text{-}\mathbf{Coalg},$$

where the first functor is described in [Spi21, Proposition 2.13], and the second is given by the fact that $(-)$-**Coalg** is a functor **Poly** $\to$ **Cat**, using a map of polynomials $[p, q] \otimes [q, r] \to [p, r]$. On positions, this map takes the form

$$\Big([p, q] \otimes [q, r]\Big)(1) = \mathbf{Poly}(p, q) \times \mathbf{Poly}(q, r) \xrightarrow{\circ} \mathbf{Poly}(p, r) = [p, r](1)$$

and on directions it is given for $\phi : p \to q$ and $\psi : q \to r$ by the function

$$\Big(\sum_{I \in p(1)} q[\phi_1(I)]\Big) \times \Big(\sum_{J \in q(1)} r[\psi_1(J)]\Big) \leftarrow \sum_{I \in p(1)} r[\psi_1(\phi_1(I))]$$

which sends (I, j) to $(I, \psi_{\phi_1(I)}(j), \phi_1(I), j)$.

Concretely, the composite of a $[p, q]$-coalgebra $\mathbf{S}$ and a $[q, r]$-coalgebra $\mathbf{S}'$ is a $[p, r]$-coalgebra $\mathbf{S}; \mathbf{S}'$ is described as follows:

- its states are given by $S \times S'$
- the action of the pair (s, s') is given by the composite

$$p \xrightarrow{\beta_0(s)} q \xrightarrow{\beta_0'(s')} r$$

- the update function of (s, s') is induced by the functions

$$\sum_{I \in p(1)} r[\beta_0'(s')_1(\beta_0(s)_1(I))] \xrightarrow{\beta_0(s)_1} \sum_{J \in q(1)} r[\beta_0'(s')_1(J)] \xrightarrow{\beta_{s'}'} S',$$

$$\sum_{I \in p(1)} r[\beta_0'(s')_1(\beta_0(s)_1(I))] \xrightarrow{\beta_0'(s')_{\beta_0(s)_1(I)}} \sum_{I \in p(1)} q[\beta_0(s)_1(I)] \xrightarrow{\beta_s} S.$$

Horizontal composition of morphisms of coalgebras, the 2-cells of the bicategory, is given simply by the cartesian product. The coherence isomorphisms and axioms for a bicategory then follow from the essential uniqueness of finite products of sets, and the unitality and associativity of composition of polynomial maps.

2.3   Product of coalgebras

[Spi21, Proposition 2.13] shows that the tensor product $\otimes$ of polynomials extends to make $\mathbb{O}\mathbf{rg}$ a monoidal bicategory. In particular, for polynomials p, q, p', q' there is a functor

$$[p, q]\text{-}\mathbf{Coalg} \times [p', q']\text{-}\mathbf{Coalg} \to \Big([p, q] \otimes [p', q']\Big)\text{-}\mathbf{Coalg} \to [p \otimes p', q \otimes q']\text{-}\mathbf{Coalg}$$

derived from the map of polynomials $[p, q] \otimes [p', q'] \to [p \otimes p', q \otimes q']$ given on positions by

$$\mathbf{Poly}(p, q) \times \mathbf{Poly}(p', q') \xrightarrow{\otimes} \mathbf{Poly}(p \otimes p', q \otimes q')$$

and on directions by, for $\phi : p \to q$ and $\phi' : p' \to q'$,

$$\Big(\sum_{I \in p(1)} q[\phi_1(I)]\Big) \times \Big(\sum_{I' \in p'(1)} q'[\phi_1'(I')]\Big) \leftarrow \sum_{(I, I') \in p(1) \times p'(1)} q[\phi_1(I)] \times q'[\phi_1'(I')]$$

3

sending (I, I', j, j') to (I, j, I', j').

Concretely, this tensor product takes a $[p, q]$-coalgebra **S** and a $[p', q']$-coalgebra **S'** to the $[p \otimes p', q \otimes q']$-coalgebra with states $S \times S'$, action

$$S \times S' \to \mathbf{Poly}(p, q) \times \mathbf{Poly}(p', q') \to \mathbf{Poly}(p \otimes p', q \otimes q'),$$

and update described similarly componentwise. The tensor product of morphisms of coalgebras is also given by the cartesian product of functions, and it is (very) tedious but ultimately straightforward to check that the essential uniqueness of products guarantees that $\otimes$ gives a monoidal structure on $\mathbb{O}$**rg**.

2.4 $\mathbb{O}$rg as a double category

While defining $\mathbb{O}$**rg** as a monoidal bicategory is sufficient for most of the constructions of $\mathbb{O}$**rg**-enriched structures in the next chapter, distinguishing the maps of polynomials as a special class of coalgebras allows $\mathbb{O}$**rg** to encode all of **Poly** and also facilitates a good definition of morphism between adaptives.

Specifically, the definition of $\mathbb{O}$**rg** as a monoidal bicategory extends to a monoidal (pseudo-)double category with coalgebras as horizontal morphisms, maps in **Poly** as vertical morphisms, and squares as below given by maps of coalgebras from $\mathbf{S}_\psi; \mathbf{S}$ to $\mathbf{S}'; \mathbf{S}_\phi$.

$$
\begin{array}{ccc}
p & \xrightarrow{\;\mathbf{S}\;} & q \\
{\scriptstyle \phi}\downarrow & \Downarrow & \downarrow{\scriptstyle \psi} \\
p' & \xrightarrow[\;\mathbf{S}'\;]{} & q'
\end{array}
$$

As $\mathbf{S}_\phi$ and $\mathbf{S}_\psi$ have only one state and composition of coalgebras acts as the cartesian product on states, such a square amounts to a function $S \to S'$ making the diagram below commute.

$$
\begin{array}{ccccc}
S & \xrightarrow{\;\beta\;} & [p, q] \triangleleft S & \xrightarrow{\;\psi_*\;} & [p, q'] \triangleleft S \\
{\scriptstyle f}\downarrow & & & & \downarrow{\scriptstyle [p,q']\triangleleft f} \\
S' & \xrightarrow[\;\beta'\;]{} & [p', q'] \triangleleft S' & \xrightarrow[\;\phi^*\;]{} & [p, q'] \triangleleft S'
\end{array}
$$

Identities and composites for these squares are determined by the bicategory structure, as this double category is a restriction in the vertical direction of the double category of lax-commuting squares in a bicategory.[1]

We now proceed to discuss various categorical structures enriched in $\mathbb{O}$**rg**, which describe dynamical systems equipped with extra algebraic structure that allows us to remove abstraction barriers when considering nested layers and complicated arrangements of the components of the system.

---

[1]It should be noted however that the vertical arrows in $\mathbb{O}$**rg** are regarded as polynomial maps rather than coalgebras, so that they compose strictly unitally and associatively.

3 $\mathbb{O}$rg-Enrichment as Nested Dynamical Structure

A monoidal double category is a viable setting for enriching various categorical structures (using the notions of enrichment in [Lei99] and [Sha22]). Generally speaking, enrichment in $\mathbb{O}$**rg** replaces the usual set of arrows between two objects in a category or similar structure with a $[p, q]$-coalgebra for some choice of polynomials p, q. Therefore not only can the arrows be realized as maps of polynomials $p \to q$, but these maps carry dynamics that encode how a position in p and a direction in q determine a transition to one arrow to another.

3.1 $\mathbb{O}$rg-enriched categories

Enrichment of categories only makes use of the double category structure of $\mathbb{O}$**rg**, as any double category forms an fc-multicategory (also known as a virtual double category) in the sense of [Lei99]. The following definition of enrichment in $\mathbb{O}$**rg** is an unwinded version of [Lei99], which defines categories enriched in any fc-multicategory.

Definition 3.1. An $\mathbb{O}$**rg**-enriched category A consists of
- A set A_0 of objects
- For each $a \in A_0$, a polynomial p_a
- For each $a, b \in A_0$, a $[p_a, p_b]$-coalgebra $\mathbf{S}_{a,b}$
- For each $a \in A_0$, an "identitor" square in $\mathbb{O}$**rg** as below left
- For each $a, b, c \in A_0$, a "compositor" square in $\mathbb{O}$**rg** as below right

$$
\begin{array}{ccc}
p_a & \xrightarrow{\;\mathrm{id}_{p_a}\;} & p_a \\
\Big\| & \Downarrow & \Big\| \\
p_a & \xrightarrow[\mathbf{S}_{a,a}]{} & p_a
\end{array}
\qquad\qquad
\begin{array}{ccccc}
p_a & \xrightarrow{\;\mathbf{S}_{a,b}\;} & p_b & \xrightarrow{\;\mathbf{S}_{b,c}\;} & p_c \\
\Big\| & & \Downarrow & & \Big\| \\
p_a & & \xrightarrow[\mathbf{S}_{a,c}]{} & & p_c
\end{array}
$$

such that these squares satisfy unit and associativity equations. ◇

The sets $S_{a,b}$ form an underlying ordinary category of A. In fact, an $\mathbb{O}$**rg**-enriched category could be equivalently defined as an ordinary category such that each object a is equipped with a polynomial p_a and each set of arrows $\mathrm{Hom}(a, b)$ is the set of states for a $[p_a, p_b]$-coalgebra $\mathbf{S}_{a,b}$, with composition and identities respecting the coalgebra structure. This means that the arrow id_a in $\mathrm{Hom}(a, a)$ acts as the identity map on p_a and is unchanged by updates, while for f in $\mathrm{Hom}(a, b)$ and g in $\mathrm{Hom}(b, c)$ the composite $g \circ f$ acts as the composite $p_a \to p_b \to p_c$ of the actions of f and g, and updates as the composite of the updates of f, g.

3.2 $\mathbb{O}$rg-enriched operads

A monoidal double category also gives rise to an fm-multicategory in the sense of [Lei99], so it makes sense to talk about multicategories enriched in $\mathbb{O}$**rg** as in [Lei99].

Definition 3.2. An $\mathbb{O}$**rg**-enriched multicategory A consists of

- A set A_0 of objects
- For each $a \in A_0$, a polynomial p_a
- For each $a_1, ..., a_n, b \in A_0$, a $[p_{a_1} \otimes \cdots \otimes p_{a_n}, p_b]$-coalgebra $\mathbf{S}_{a_1,...,a_n;b}$
- For each $a \in A_0$, an "identitor" square in $\mathbb{O}$**rg** as below left
- For each $a_{1,1}, ..., a_{1,m_1}, a_{2,1}, ..., a_{n,m_n}, b_1, ..., b_n, c \in A_0$, a "compositor" square in $\mathbb{O}$**rg** as below right

$$
\begin{array}{ccc}
p_a & \xrightarrow{\mathrm{id}_{p_a}} & p_a \\
\| & \Downarrow & \| \\
p_a & \xrightarrow[\mathbf{S}_{a;a}]{} & p_a
\end{array}
\qquad
\begin{array}{ccccc}
p_{a_{1,1}} \otimes \cdots \otimes p_{a_{n,m_n}} & \xrightarrow{\otimes_i \mathbf{S}_{a_{i,1},...,a_{i,m_i};b_i}} & p_{b_1} \otimes \cdots \otimes p_{b_n} & \xrightarrow{\mathbf{S}_{b_1,...,b_n;c}} & p_c \\
\| & & \Downarrow & & \| \\
p_{a_{1,1}} \otimes \cdots \otimes p_{a_{n,m_n}} & & \xrightarrow[\mathbf{S}_{a_{1,1},...,a_{n,m_n};c}]{} & & p_c
\end{array}
$$

such that these squares satisfy unit and associativity equations. $\qquad \Diamond$

There is an underlying multicategory of A given by the sets $\mathbf{S}_{a_1,...,a_n;b}$.

We will mostly be interested in when A has only one object, in which case we call it an $\mathbb{O}$**rg**-enriched operad. An $\mathbb{O}$**rg**-enriched operad amounts to precisely a polynomial p and for each $n \in \mathbb{N}$ a $[p^{\otimes n}, p]$-coalgebra S_n equipped with maps $id_p \to S_1$ and $\otimes_i S_{n_i} \to S_{\sum_i n_i}$ satisfying the usual equations.

Example 3.3. A *collective* (as defined in [NS21]) is a $\otimes$-monoid in **Poly**, meaning a polynomial p equipped with a monoid structure on its positions $p(1)$ and co-unital co-associative "distribution" functions $p[I \cdot J] \to p[I] \times p[J]$ for each $I, J \in p(1)$. This can be viewed as an $\mathbb{O}$**rg**-enriched operad where $\mathbf{S}_n$ is given by $\mathbf{S}_{\cdot_n}$, the single-state coalgebra on the n-ary monoidal product $\cdot_n : p^{\otimes n} \to p$ and the above maps of coalgebras are isomorphisms deduced from the equations for a monoid. $\qquad \Diamond$

Example 3.4. In the previous example, the coalgebras $\mathbf{S}_n$ are determined by a single map of polynomials, with no possible updates as there is only one state. This can be generalized to an intermediate notion between collectives and $\mathbb{O}$**rg**-enriched multicategories, where there are multiple states and maps but still no updates.

Given a multicategory M and a multifunctor $F : M \to$ **Poly**, there is an $\mathbb{O}$**rg**-enriched multicategory A_F with

- object set $\mathrm{Ob}(M)$
- for each $a \in \mathrm{Ob}(M)$, the polynomial $p_a = F(a)$
- for each tuple $(a_1, ..., a_n; b)$ in $\mathrm{Ob}(M)$, $S_{a_1,...,a_n;b} = M(a_1, ..., a_n; b)$
- the map $\beta_0 : M(a_1, ..., a_n; b) \to \textbf{Poly}(p_{a_1} \otimes \cdots \otimes p_{a_n}, p_b)$ is given by F
- no updates, in that for any state s in $M(a_1, ..., a_n; b)$, β_s sends every input and output to s

$\qquad \Diamond$

Example 3.5. Let $\mathbf{S}$ be any p-coalgebra for a polynomial p. There is an $\mathbb{O}$**rg**-enriched operad with the polynomial p, $\mathbf{S}_0 = \mathbf{S}$, $\mathbf{S}_1 = \mathbf{S}_{\mathrm{id}_p}$, and all other $\mathbf{S}_n$ the empty coalgebras. $\qquad \Diamond$

3.3 Adaptives

A monoidal double category is precisely a representable MC-multicategory in the sense of [Sha22], so we can also enrich strict monoidal categories in $\mathbb{O}\mathbf{rg}$.[2] These will be similar to $\mathbb{O}\mathbf{rg}$-enriched multicategories but allow for many-to-many coalgebras rather than just many-to-1.

Definition 3.6. An $\mathbb{O}\mathbf{rg}$-enriched strict monoidal category A consists of
- A monoid A_0 of objects
- For each $a \in A_0$, a polynomial p_a
- An isomorphism of polynomials $y \cong p_e$ for e the unit of A_0
- For each $a, a' \in A_0$, an isomorphism of polynomials $p_a \otimes p_{a'} \cong p_{aa'}$
- For each $a, b \in A_0$, a $[p_a, p_b]$-coalgebra $\mathbf{S}_{a,b}$
- For each $a \in A_0$, an "identitor" square in $\mathbb{O}\mathbf{rg}$ as below left
- For each $a, b, c \in A_0$, a "compositor" square in $\mathbb{O}\mathbf{rg}$ as below center
- For each $a, a', b, b' \in A_0$, a "productor" square in $\mathbb{O}\mathbf{rg}$ as below right

$$
\begin{array}{ccc}
p_a \xrightarrow{\mathrm{id}_{p_a}} p_a & \qquad p_a \xrightarrow{\mathbf{S}_{a,b}} p_b \xrightarrow{\mathbf{S}_{b,c}} p_c & \qquad p_a \otimes p_{a'} \xrightarrow{\mathbf{S}_{a,b} \otimes \mathbf{S}_{a',b'}} p_b \otimes p_{b'} \\
\| \quad \Downarrow \quad \| & \| \qquad \Downarrow & \| \quad \Downarrow \quad \| \\
p_a \xrightarrow{\mathbf{S}_{a,a}} p_a & \qquad p_a \xrightarrow{\mathbf{S}_{a,c}} p_c & \qquad p_{aa'} \xrightarrow{\mathbf{S}_{aa',bb'}} p_{bb'}
\end{array}
$$

such that these isomorphisms and squares satisfy unit, associativity, and interchange equations. ◊

Here the sets $S_{a,b}$ for the arrows in the underlying strict monoidal category of A.

Definition 3.7. An *adaptive* is an $\mathbb{O}\mathbf{rg}$-enriched strict monoidal category with object monoid $\mathbb{N}$. ◊

Concretely, an adaptive consists of a polynomial p (so that in the notation above $p_n := p^{\otimes n}$ for $n \in \mathbb{N}$) along with a $[p^{\otimes m}, p^{\otimes n}]$-coalgebra $S_{m,n}$ for each $m, n \in \mathbb{N}$, equipped with the maps of coalgebras as in the squares above. We denote such an adaptive "on p" as $(p, \mathbf{S})$, where $\mathbf{S}$ now encodes all of the coalgebras $\mathbf{S}_{m,n}$ that constitute the adaptive and the structure maps are implicit.

The definition of morphisms of adaptives comes from morphisms of enriched monoidal categories as defined in [Sha22].

Definition 3.8. A morphism of adaptives from $(p, \mathbf{S})$ to $(p', \mathbf{S}')$ is given by a map of polynomials $\phi : p \to p'$ and, for each $m, n \in \mathbb{N}$, "commutor" squares as below in $\mathbb{O}\mathbf{rg}$ which commute with the identitor, compositor, and productor squares.

$$
\begin{array}{ccc}
p^{\otimes m} & \xrightarrow{\mathbf{S}_{m,n}} & p^{\otimes n} \\
\phi^{\otimes m} \downarrow & \Downarrow & \downarrow \phi^{\otimes n} \\
p'^{\otimes m} & \xrightarrow{\mathbf{S}'_{m,n}} & p'^{\otimes n}
\end{array}
$$

◊

---

[2]We use throughout the notion *strong* enrichment in a monoidal double category from [Sha22].

This definition of morphism is the direct benefit for the theory of adaptives of treating $\mathbb{O}\mathbf{rg}$ as a monoidal double category rather than a monoidal bicategory. Otherwise morphisms could either only be easily defined between adaptives on the same polynomial, which is rather restrictive, or take the form of a $[p, p']$-coalgebra, which we believe to be too general to be of much use.

Example 3.9. For a fixed polynomial p, there is a terminal adaptive on p where $\mathbf{S}^!_{m,n}$ is the terminal $[p^{\otimes m}, p^{\otimes n}]$-coalgebra.

Its states are infinite $[p^{\otimes m}, p^{\otimes n}]$-trees: trees inductively defined by a root labeled with a polynomial map $\phi : p^{\otimes m} \to p^{\otimes n}$ and for each pair

$$\Big((I_1, ..., I_m) \in p(1)^m, (i_1, ..., i_n) \in p[\phi(I_1, ..., I_m)_1] \times \cdots \times p[\phi(I_1, ..., I_m)_n]\Big)$$

an edge out of the root to another $[p^{\otimes m}, p^{\otimes n}]$-tree. The action of such a tree is simply the root map ϕ, and the update sends a pair as above to the new tree at the end of the corresponding edge.

The idea is that the states encode all possible trajectories along different actions, and this coalgebra is terminal because from any other coalgebra there is a map to $\mathbf{S}^!_{m,n}$ sending each state to the tree whose root is labeled by the action of the state and whose edges from the root go to the trees for each of the state's possible updates. This map is (uniquely) a coalgebra map as in order to preserve actions and updates each state must be sent to the tree rooted by its action and branching according to its updates.

The adaptive structure on these terminal coalgebras is then each to define, as for any identity, composition, or product $[p^{\otimes m}, p^{\otimes n}]$-coalgebra there is a unique map to the terminal one, and this uniqueness automatically ensures all of the equations are satisfied.

This is the terminal adaptive on p as for any other adaptive on p there is a morphism of adaptives given by the identity map on p and with commutor squares to $\mathbf{S}^!_{m,n}$ simply the unique map of coalgebras from any $[p^{\otimes m}, p^{\otimes n}]$-coalgebra to the terminal one. ◊

4 Adaptives in Nature

4.1 The Gradient Descent Adaptive

Deep learning uses the algorithm of gradient descent to optimize a choice of function based on external feedback on its output. This naturally fits into the paradigm of adaptives, as functions $\mathbb{R}^m \to \mathbb{R}^n$ can form the states of a polynomial coalgebra, with the feedback providing the information needed to update the choice of function. These functions can be composed and juxtaposed to provide compositions of networks of such gradient descenders in a way that respects the updates.

Definition 4.1. For the rest of this section, let

$$S_{m,n} = \{M \in \mathbb{N}, f : \mathbb{R}^{M+m}, p \in \mathbb{R}^M | f \text{ is differentiable}\}.$$

The idea is that these states are the possible parameters among which a gradient descender is meant to find the optimal choice, while f dictates how the parameter affects the resulting function $f(p, -)$. In the dynamics of these states described below, only the value of the parameter p is updated while the dimension M of the parameter space and the parameterized function f remain fixed, though network composition of descenders will involve combining these data in nontrivial ways.[3]

Let $t = \sum_{x \in \mathbb{R}} y^{T_x \mathbb{R}}$. Of course for all practical purposes $T_x \mathbb{R}$ can be regarded as simply $\mathbb{R}$, but in both the description of polynomials as bundles and the intuition for how we use the directions in this example it will make sense to think of the directions at x as the tangent space. We proceed to define an adaptive on t with $S_{m,n}$ the states of the coalgebra $\mathbf{S}_{m,n}$ which updates using gradient descent. The structure maps in the adaptive encode how certain networks of gradient descenders can be composed into a single descender with a larger parameter space.

Definition 4.2. The $[t^{\otimes m}, t^{\otimes n}]$-coalgebra structure on $S_{m,n}$ is given by
- $\beta_0(M, f, p)_1 : \mathbb{R}^m \to \mathbb{R}^n$ given by $f(p, -)$
- For $x \in \mathbb{R}^m$, $\beta_0(M, f, p)_x : T_{f(p,x)} \to T_x$ sends y to $\pi_m Df^\top y$
- The update function $\beta_{M,f,p}$ sends $x \in \mathbb{R}^m$ and $y \in T_{f(p,x)}$ to $(M, f, p + \epsilon \pi_M Df^\top y)$ for some fixed value of ϵ

The action of a state as a map $t^{\otimes m} \to t^{\otimes n}$ is given by applying the parameterized function f with the parameter p resulting in a function $\mathbb{R}^m \to \mathbb{R}^n$ as desired. The transpose of the derivative of f sends a feedback vector $y \in T_{f(p,x)} \mathbb{R}^n$, which can be interpreted as the difference in $\mathbb{R}^n$ between the "correct" result for x and the current approximation $f(p, x)$, to the corresponding "correction" to (p, x) in $\mathbb{R}^{M+m}$. The projection of this correction to $T_x \mathbb{R}^m$ provides the action of the state on directions, which in a network can then be further propogated back down to another gradient descender that output x (more on this below). The projection to $T_p \mathbb{R}^M$ provides the direction and magnitude with which to update the parameters (scaled by ϵ according to the desired responsiveness of the learning).

Thus far, we have provided the data of the polynomial t and the $[t^{\otimes m}, t^{\otimes n}]$-coalgebra $\mathbf{S}_{m,n}$ needed to define an adaptive. We now define the identitor, compositor, and producer morphisms of coalgebras presented by the squares in Definition 3.6.
- The identitors $\mathrm{id}_{t^{\otimes n}} \to \mathbf{S}_{n,n}$ send the unique state in the domain to the state $(0, \mathrm{id}_{\mathbb{R}^n}, 0) \in S_{n,n}$.
- The compositors $\mathbf{S}_{\ell,m}; \mathbf{S}_{m,n} \to \mathbf{S}_{\ell,n}$ send the pair $((L, f, p), (M, g, q))$ to

$$\left(M + L, g(-, f(-, -)) : \mathbb{R}^{M+L+\ell} \xrightarrow{\mathrm{id} \times f} \mathbb{R}^{M+m} \xrightarrow{g} \mathbb{R}^n, (q, p) \in \mathbb{R}^{M+L} \right).$$

---

[3]Other versions of a gradient descent adaptive might have the states be simply some space of functions from $\mathbb{R}^m$ to $\mathbb{R}^n$, but this type of model is more difficult to describe in a computation-friendly manner as the space of functions that can arise from network compositions of the types of functions used in common practice is not finite dimensional.

- The productors $\mathbf{S}_{m,n} \otimes \mathbf{S}_{m',n'} \to \mathbf{S}_{m+m',n+n'}$ send the pair $((M, f, p), (M', f', p'))$ to $(M + M', (f, f'), (p, p'))$

These structure maps ensure that whenever two gradient descenders are combined in series or parallel, the resulting composite descender retains the parameter spaces of both. Likewise when the input or output of a descender is wired past some other descender in the network as below, it does not contribute any new parameters and its effect is merely to preserve its input/output until plugged into a descender elsewhere in the network.

Theorem 4.3. *These structure maps are maps of coalgebras and satisfy the coherence equations of an adaptive described in Definition A.1.*

This is proven in Appendix A.

4.2 The Gravity Adaptive

A Coherence equations and proofs

We now present the equations that must be satisfied by the identitors, compositors, and productors of an adaptive. While we only provide these in the narrow setting of adaptives, the equations governing $\mathbb{O}\mathbf{rg}$-enriched categories, multicategories, and monoidal categories are all quite similarly defined.

Definition A.1. The equations between the structure maps in an adaptive are as follows:

- The identitor interchange law

$$
\begin{array}{ccc}
p^{\otimes n} \otimes p^{\otimes n'} & \xrightarrow{\mathrm{id}_{p^{\otimes n}} \otimes \mathrm{id}_{p^{\otimes n'}}} & p^{\otimes n} \otimes p^{\otimes n'} \\
\| & \Downarrow & \| \\
p^{\otimes n} \otimes p^{\otimes n'} & \xrightarrow{\mathbf{S}_{n,n} \otimes \mathbf{S}_{n',n'}} & p^{\otimes n} \otimes p^{\otimes n'} \\
\wr\| & \Downarrow & \wr\| \\
p^{\otimes(n+n')} & \xrightarrow{\mathbf{S}_{n+n',n+n'}} & p^{\otimes(n+n')}
\end{array}
\quad = \quad
\begin{array}{ccc}
p^{\otimes n} \otimes p^{\otimes n'} & \xrightarrow{\mathrm{id}_{p^{\otimes n}} \otimes \mathrm{id}_{p^{\otimes n'}}} & p^{\otimes n} \otimes p^{\otimes n'} \\
\wr\| & \wr\| & \wr\| \\
p^{\otimes(n+n')} & \xrightarrow{\mathrm{id}_{p^{\otimes(n+n')}}} & p^{\otimes(n+n')} \\
\| & \Downarrow & \| \\
p^{\otimes(n+n')} & \xrightarrow{\mathbf{S}_{n+n',n+n'}} & p^{\otimes(n+n')}
\end{array}
$$

- The compositor interchange law

$$
\begin{array}{ccccc}
p^{\otimes \ell} \otimes p^{\otimes \ell'} & \xrightarrow{\mathbf{S}_{\ell,m} \otimes \mathbf{S}_{\ell',m'}} & p^{\otimes m} \otimes p^{\otimes m'} & \xrightarrow{\mathbf{S}_{m,n} \otimes \mathbf{S}_{m',n'}} & p^{\otimes n} \otimes p^{\otimes n'} \\
\wr\| & \Downarrow & \wr\| & \Downarrow & \wr\| \\
p^{\otimes(\ell+\ell')} & \xrightarrow{\mathbf{S}_{\ell+\ell',m+m'}} & p^{\otimes(m+m')} & \xrightarrow{\mathbf{S}_{m+m',n+n'}} & p^{\otimes(n+n')} \\
\| & \Downarrow & \| & & \| \\
p^{\otimes(\ell+\ell')} & & \xrightarrow{\mathbf{S}_{\ell+\ell',n+n'}} & & p^{\otimes(n+n')}
\end{array}
\quad = \quad
\begin{array}{ccccc}
p^{\otimes \ell} \otimes p^{\otimes \ell'} & \xrightarrow{\mathbf{S}_{\ell,m} \otimes \mathbf{S}_{\ell',m'}} & p^{\otimes m} \otimes p^{\otimes m'} & \xrightarrow{\mathbf{S}_{m,n} \otimes \mathbf{S}_{m',n'}} & p^{\otimes n} \otimes p^{\otimes n'} \\
\| & & \Downarrow & & \| \\
p^{\otimes \ell} \otimes p^{\otimes \ell'} & & \xrightarrow{\mathbf{S}_{\ell,n} \otimes \mathbf{S}_{\ell',n'}} & & p^{\otimes n} \otimes p^{\otimes n'} \\
\wr\| & & \Downarrow & & \wr\| \\
p^{\otimes(\ell+\ell')} & & \xrightarrow{\mathbf{S}_{\ell+\ell',n+n'}} & & p^{\otimes(n+n')}
\end{array}
$$

- The compositor associativity law

$$
\begin{array}{ccccccc}
p^{\otimes k} & \xrightarrow{\ \mathbf{s}_{k,\ell}\ } & p^{\otimes \ell} & \xrightarrow{\ \mathbf{s}_{\ell,m}\ } & p^{\otimes m} & \xrightarrow{\ \mathbf{s}_{m,n}\ } & p^{\otimes n} \\
\| & & \Downarrow & & \| \quad \| & & \| \\
p^{\otimes k} & \xrightarrow[\ \mathbf{s}_{k,m}\ ]{} & & & p^{\otimes m} & \xrightarrow[\ \mathbf{s}_{m,n}\ ]{} & p^{\otimes n} \\
\| & & & \Downarrow & & & \| \\
p^{\otimes k} & \xrightarrow[\qquad\quad \mathbf{s}_{k,n}\qquad\quad]{} & & & & & p^{\otimes n}
\end{array}
\qquad = \qquad
\begin{array}{ccccccc}
p^{\otimes k} & \xrightarrow{\ \mathbf{s}_{k,\ell}\ } & p^{\otimes \ell} & \xrightarrow{\ \mathbf{s}_{\ell,m}\ } & p^{\otimes m} & \xrightarrow{\ \mathbf{s}_{m,n}\ } & p^{\otimes n} \\
\| \quad \| & & \| & & \Downarrow & & \| \\
p^{\otimes k} & \xrightarrow[\ \mathbf{s}_{k,\ell}\ ]{} & p^{\otimes \ell} & \xrightarrow[\ \mathbf{s}_{\ell,n}\ ]{} & & & p^{\otimes n} \\
\| & & & \Downarrow & & & \| \\
p^{\otimes k} & \xrightarrow[\qquad\quad \mathbf{s}_{k,n}\qquad\quad]{} & & & & & p^{\otimes n}
\end{array}
$$

- The compositor unit laws

$$
\begin{array}{ccccccc}
p^{\otimes m} & \xrightarrow{\ \mathrm{id}_{p^{\otimes m}}\ } & p^{\otimes m} & \xrightarrow{\ \mathbf{s}_{m,n}\ } & p^{\otimes n} \\
\| & \Downarrow & \| \quad \| & & \| \\
p^{\otimes m} & \xrightarrow[\ \mathbf{s}_{m,m}\ ]{} & p^{\otimes m} & \xrightarrow[\ \mathbf{s}_{m,n}\ ]{} & p^{\otimes n} \\
\| & & \Downarrow & & \| \\
p^{\otimes m} & \xrightarrow[\qquad \mathbf{s}_{m,n}\qquad]{} & & & p^{\otimes n}
\end{array}
\ = \
\begin{array}{ccc}
p^{\otimes m} & \xrightarrow{\ \mathbf{s}_{m,n}\ } & p^{\otimes n} \\
\| & \| & \| \\
p^{\otimes m} & \xrightarrow[\ \mathbf{s}_{m,n}\ ]{} & p^{\otimes n}
\end{array}
\ = \
\begin{array}{ccccccc}
p^{\otimes m} & \xrightarrow{\ \mathbf{s}_{m,n}\ } & p^{\otimes n} & \xrightarrow{\ \mathrm{id}_{p^{\otimes n}}\ } & p^{\otimes n} \\
\| & \| & \| & \Downarrow & \| \\
p^{\otimes m} & \xrightarrow[\ \mathbf{s}_{m,n}\ ]{} & p^{\otimes n} & \xrightarrow[\ \mathbf{s}_{n,n}\ ]{} & p^{\otimes n} \\
\| & & \Downarrow & & \| \\
p^{\otimes m} & \xrightarrow[\qquad \mathbf{s}_{m,n}\qquad]{} & & & p^{\otimes n}
\end{array}
$$

- The productor associativity law

$$
\begin{array}{ccc}
p^{\otimes m} \otimes p^{\otimes m'} \otimes p^{\otimes m''} & \xrightarrow{\ \mathbf{s}_{m,n}\otimes\mathbf{s}_{m',n'}\otimes\mathbf{s}_{m'',n''}\ } & p^{\otimes n} \otimes p^{\otimes n'} \otimes p^{\otimes n''} \\
\wr\| & \Downarrow & \wr\| \\
p^{\otimes(m+m')} \otimes p^{\otimes m''} & \xrightarrow{\ \mathbf{s}_{m+m',n+n'}\otimes\mathbf{s}_{m'',n''}\ } & p^{\otimes(n+n')} \otimes p^{\otimes n''} \\
\wr\| & \Downarrow & \wr\| \\
p^{\otimes(m+m'+m'')} & \xrightarrow[\ \mathbf{s}_{m+m'+m'',n+n'+n''}\ ]{} & p^{\otimes(n+n'+n'')}
\end{array}
\ = \
\begin{array}{ccc}
p^{\otimes m} \otimes p^{\otimes m'} \otimes p^{\otimes m''} & \xrightarrow{\ \mathbf{s}_{m,n}\otimes\mathbf{s}_{m',n'}\otimes\mathbf{s}_{m'',n''}\ } & p^{\otimes n} \otimes p^{\otimes n'} \otimes p^{\otimes n''} \\
\wr\| & \Downarrow & \wr\| \\
p^{\otimes m} \otimes p^{\otimes(m+m'')} & \xrightarrow{\ \mathbf{s}_{m,n}\otimes\mathbf{s}_{m'+m'',n'+n''}\ } & p^{\otimes n} \otimes p^{\otimes(n'+n'')} \\
\wr\| & \Downarrow & \wr\| \\
p^{\otimes(m+m'+m'')} & \xrightarrow[\ \mathbf{s}_{m+m'+m'',n+n'+n''}\ ]{} & p^{\otimes(n+n'+n'')}
\end{array}
$$

- The productor unit laws

$$
\begin{array}{ccc}
p^{\otimes m} & \xrightarrow{\ \mathbf{s}_{m,n}\ } & p^{\otimes n} \\
\wr\| & \wr\| & \wr\| \\
p^{\otimes 0} \otimes p^{\otimes m} & \xrightarrow{\ \mathrm{id}_{p^{\otimes 0}}\otimes\mathbf{s}_{m,n}\ } & p^{\otimes 0} \otimes p^{\otimes n} \\
\| & \Downarrow & \| \\
p^{\otimes 0} \otimes p^{\otimes m} & \xrightarrow{\ \mathbf{s}_{0,0}\otimes\mathbf{s}_{m,n}\ } & p^{\otimes 0} \otimes p^{\otimes n} \\
\wr\| & \Downarrow & \wr\| \\
p^{\otimes m} & \xrightarrow[\ \mathbf{s}_{m,n}\ ]{} & p^{\otimes n}
\end{array}
\ = \
\begin{array}{ccc}
p^{\otimes m} & \xrightarrow{\ \mathbf{s}_{m,n}\ } & p^{\otimes n} \\
\| & \| & \| \\
p^{\otimes m} & \xrightarrow[\ \mathbf{s}_{m,n}\ ]{} & p^{\otimes n}
\end{array}
\ = \
\begin{array}{ccc}
p^{\otimes m} & \xrightarrow{\ \mathbf{s}_{m,n}\ } & p^{\otimes n} \\
\wr\| & \wr\| & \wr\| \\
p^{\otimes m} \otimes p^{\otimes 0} & \xrightarrow{\ \mathbf{s}_{m,n}\otimes\mathrm{id}_{p^{\otimes 0}}\ } & p^{\otimes n} \otimes p^{\otimes 0} \\
\| & \Downarrow & \| \\
p^{\otimes m} \otimes p^{\otimes 0} & \xrightarrow{\ \mathbf{s}_{m,n}\otimes\mathbf{s}_{0,0}\ } & p^{\otimes n} \otimes p^{\otimes 0} \\
\wr\| & \Downarrow & \wr\| \\
p^{\otimes m} & \xrightarrow[\ \mathbf{s}_{m,n}\ ]{} & p^{\otimes n}
\end{array}
$$

$$\Diamond$$

We now proceed to show that the adaptives described above for gradient descent and gravitational motion satisfy these equations, and that their structure maps as defined on states are indeed morphisms of coalgebras.

11

Proof. For the identitor the state $(0, \mathrm{id}_{\mathbb{R}^n}, 0)$ acts as the identity function on $\mathbb{R}^n$ and on directions by the transpose of its derivative, which is also the identity. The updates in the coalgebras $\mathbf{S}_{m,n}$ only modify the parameter p, so as the parameter here is 0-dimensional this state is never changed by the update function, as is the case in the coalgebra $\mathrm{id}_{t^{\otimes n}}$. Therefore this function is a map of coalgebras.

<div align="right">□</div>

References

[Lei99]   Tom Leinster. *Generalized enrichment for categories and multicategories.* arXiv:9901139. 1999. arXiv: `9901139 [math.CT]` (cit. on p. 5).

[NS21]   Nelson Niu and David I. Spivak. *Collectives: Compositional protocols for contributions and returns.* 2021. DOI: `10.48550/ARXIV.2112.11518`. URL: `https://arxiv.org/abs/2112.11518` (cit. on p. 6).

[Sha22]   Brandon Shapiro. *Enrichment of Algebraic Higher Categories. In Preparation.* 2022 (cit. on pp. 5, 7).

[Spi21]   David I. Spivak. *Learners' languages.* 2021. DOI: `10.48550/ARXIV.2103.01189`. URL: `https://arxiv.org/abs/2103.01189` (cit. on pp. 1–3).