Adaptives:
Self-similar hierarchical dynamical systems

Brandon Shapiro David I. Spivak

Abstract

1 Introduction

Brainstorm:
Self-similarity in nature
Polynomial ecosystem
Collectives
Deep learning, gradient descenders
Open dynamical systems
Hierarchy
Organization doesn't defy physics

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.[1]

[1] In fact, $\mathbb{O}\mathbf{rg}$ is a duoidal double double category, with a second monoidal structure given by $\triangleleft$, but we will not use that here.

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

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

◊

For intuition, a $[p, q]$-coalgebra (denoted $p \nrightarrow q$) is a machine that outputs maps $p \to q$ and inputs what flows out of p and what flows in to q. More precisely, 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 (1) commute.

$$\begin{tikzcd} S \arrow[r, "\beta"] \arrow[d, "f"'] & {[p,q] \triangleleft S} \arrow[d, "{[p,q] \triangleleft f}"] \\ S' \arrow[r, "\beta'"'] & {[p,q] \triangleleft S'} \end{tikzcd} \tag{1}$$

◊

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 $s \in S$, the update β_s is the constant function sending everything to s, we say the coalgebra $\mathbf{S}$ is *static*, as it remains constantly at s regardless of the output $I \in p(1)$ and input $j \in q[\beta_0(s)_1(I)]$.

Example 2.3. A special case of a static $[p, q]$-coalgebra is given by a map $\phi \in \mathbf{Poly}(p, q)$. For each such ϕ, there is a coalgebra $\{\phi\}$ with a singleton set of states, β_0 sending the point to ϕ; it is by necessity static.

A coalgebra is static iff it is the coproduct of one-element coalgebras. ◊

2.2 Composition of morphism 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 hom-categories in a bicategory $\mathbb{O}$**rg**, which has 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 $\{\mathrm{id}_p\}$.

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 which we denote $\mathbf{S} \mathbin{⨟} \mathbf{S}'$ and define as follows:

- its state set is 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 $\phi' : p' \to q'$,

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

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

Concretely, this tensor product takes a $[p,q]$-coalgebra $\mathbf{S}$ and a $[p',q']$-coalgebra $\mathbf{S}'$ 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 $\mathbb{O}\mathbf{rg}$.

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

Defining $\mathbb{O}\mathbf{rg}$ as a monoidal bicategory is sufficient for most of the constructions of $\mathbb{O}\mathbf{rg}$-enriched structures in the next chapter. However, using a double category structure to cast morphisms $\phi \colon p \to q$ in $\mathbf{Poly}$ as one-state coalgebras $\mathbf{S}_\phi \in [p,q]\text{-}\mathbf{Coalg}$ (see Example 2.3) facilitates our eventual definition of adaptives, in particular the maps between them.

Specifically, the definition of $\mathbb{O}\mathbf{rg}$ as a monoidal bicategory extends to a monoidal (pseudo-)double category with coalgebras as horizontal morphisms, maps in $\mathbf{Poly}$ as vertical morphisms, and squares as in (2) given by maps of coalgebras from $\{\psi\} \fatsemi \mathbf{S}$ to $\mathbf{S}' \fatsemi \{\phi\}$.

$$\begin{tikzcd} p \arrow[r, "\mathbf{S}"] \arrow[d, "\phi"'] \arrow[dr, phantom, "\Downarrow"] & q \arrow[d, "\psi"] \\ p' \arrow[r, "\mathbf{S}'"'] & q' \end{tikzcd} \tag{2}$$

As $\{\phi\}$ and $\{\psi\}$ 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 (3) commute.

$$
\begin{tikzcd}
S \arrow[r, "\beta"] \arrow[d, "f"'] & {[p,q] \triangleleft S} \arrow[r, "\psi_*"] & {[p,q'] \triangleleft S} \arrow[d, "{[p,q'] \triangleleft f}"] \\
S' \arrow[r, "\beta'"'] & {[p',q'] \triangleleft S'} \arrow[r, "\phi^*"] & {[p,q'] \triangleleft S'}
\end{tikzcd}
\tag{3}
$$

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.[2]

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

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]). Intuitively, enrichment in $\mathbb{O}\mathbf{rg}$ 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 each arrow be realized as a map of polynomials $p \to q$, but this map carries dynamics that encode how a position in p and a direction in q determine a transition from one arrow to another.

In Section 3.1, we explain enrichment of a category; in Section 3.2, we explain enrichment of a multicategory.

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

Enrichment of categories only makes use of the double category structure of $\mathbb{O}\mathbf{rg}$, 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 $\mathbb{O}\mathbf{rg}$ is an unwinded version of [Lei99], which defines categories enriched in any fc-multicategory.

Definition 3.1. An $\mathbb{O}\mathbf{rg}$-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 $\mathbb{O}\mathbf{rg}$ as in (4) left; and

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

- for each $a, b, c \in A_0$, a "compositor" square in $\mathbb{O}\mathbf{rg}$ as in (4) right:

$$\begin{tikzcd} p_a \arrow[r, "\{\mathrm{id}_{p_a}\}"] \arrow[d, equal] \arrow[dr, phantom, "\Downarrow"] & p_a \arrow[d, equal] \\ p_a \arrow[r, "\mathbf{S}_{a,a}"'] & p_a \end{tikzcd} \qquad \begin{tikzcd} p_a \arrow[r, "\mathbf{S}_{a,b}"] \arrow[d, equal] & p_b \arrow[r, "\mathbf{S}_{b,c}"] \arrow[d, phantom, "\Downarrow"] & p_c \arrow[d, equal] \\ p_a \arrow[rr, "\mathbf{S}_{a,c}"'] & & p_c \end{tikzcd} \tag{4}$$

such that these squares satisfy unit and associativity equations. ◊

The sets $S_{a,b}$ form an underlying ordinary category of A. In fact, an $\mathbb{O}\mathbf{rg}$-enriched category could be equivalently defined as an ordinary category such that each object a is assigned a polynomial p_a and each set of arrows $\mathrm{Hom}(a, b)$ is the set of states for an assigned $[p_a, p_b]$-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 it updates as the composite of the updates of f, g.

3.2 $\mathbb{O}\mathbf{rg}$-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 $\mathbb{O}\mathbf{rg}$ as in [Lei99].

Definition 3.2. An $\mathbb{O}\mathbf{rg}$-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 $\mathbb{O}\mathbf{rg}$ as in (5) left; and
- for each $a_{1,1}, \ldots, a_{1,m_1}, \ \ldots, \ a_{n,1}, \ldots, a_{n,m_n}, \ b_1, \ldots, b_n$, and $c \in A_0$, a "compositor" square in $\mathbb{O}\mathbf{rg}$ as in (5) right

$$\begin{tikzcd} p_a \arrow[r, "\{\mathrm{id}_{p_a}\}"] \arrow[d, equal] \arrow[dr, phantom, "\Downarrow"] & p_a \arrow[d, equal] \\ p_a \arrow[r, "\mathbf{S}_{a;a}"'] & p_a \end{tikzcd} \qquad \begin{tikzcd} p_{a_{1,1}} \otimes \cdots \otimes p_{a_{n,m_n}} \arrow[r, "\otimes_i \mathbf{S}_{a_{i,1},\ldots,a_{i,m_i};b_i}"] \arrow[d, equal] & p_{b_1} \otimes \cdots \otimes p_{b_n} \arrow[r, "\mathbf{S}_{b_1,\ldots,b_n;c}"] \arrow[d, phantom, "\Downarrow"] & p_c \arrow[d, equal] \\ p_{a_{1,1}} \otimes \cdots \otimes p_{a_{n,m_n}} \arrow[rr, "\mathbf{S}_{a_{1,1},\ldots,a_{n,m_n};c}"'] & & p_c \end{tikzcd} \tag{5}$$

such that these squares satisfy unit and associativity equations. ◊

The sets $S_{a,b}$ form an underlying multicategory of A, with a description similar to that below Definition 3.1.

We will mostly be interested in what we call an $\mathbb{O}\mathbf{rg}$-enriched operad on p, the case when A has only one object which is assigned the polynomial p. It consists simply of a $[p^{\otimes n}, p]$-coalgebra S_n for each $n \in \mathbb{N}$ equipped with maps

$$\{\mathrm{id}_p\} \to S_1 \qquad \text{and} \qquad \bigotimes_i S_{n_i} \to S_{\sum_i n_i} \tag{6}$$

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 $\{\cdot_n\}$, the one-element coalgebra on the n-ary monoidal product $\cdot_n : p^{\otimes n} \to p$, and where the maps of coalgebras in (6) are isomorphisms deduced from the equations for a monoid. $\Diamond$

Example 3.4. In Example 3.3, 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 the coalgebras are still static but may have multiple states.

Given a multicategory M and a multifunctor $F : M \to \mathbf{Poly}$, 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 $(a_1, \ldots, a_n; b)$ in $\mathrm{Ob}(M)$, $S_{a_1,\ldots,a_n;b} = M(a_1, \ldots, a_n; b)$;
- the map $\beta_0 : M(a_1, \ldots, a_n; b) \to \mathbf{Poly}(p_{a_1} \otimes \cdots \otimes p_{a_n}, p_b)$ is given by F; and
- this coalgebra is static, in that for any state s in $M(a_1, \ldots, a_n; b)$, the update function β_s is the constant function at s. $\Diamond$

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

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 $\mathbb{O}$**rg**.[3] These will be similar to $\mathbb{O}$**rg**-enriched multicategories but allow for many-to-many coalgebras rather than just many-to-1.

Definition 3.6. An $\mathbb{O}$**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 $\mathbb{O}$**rg** as in Eq. (7) left;
- for each $a, b, c \in A_0$, a "compositor" square in $\mathbb{O}$**rg** as in Eq. (7) center; and
- for each $a, a', b, b' \in A_0$, a "productor" square in $\mathbb{O}$**rg** as in Eq. (7) right:

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

$$
\begin{tikzcd}
p_a \arrow[r, "\{\mathrm{id}_{p_a}\}"] \arrow[d, equal] \arrow[dr, phantom, "\Downarrow"] & p_a \arrow[d, equal] \\
p_a \arrow[r, "\mathbf{S}_{a,a}"'] & p_a
\end{tikzcd}
\qquad
\begin{tikzcd}
p_a \arrow[r, "\mathbf{S}_{a,b}"] \arrow[d, equal] & p_b \arrow[r, "\mathbf{S}_{b,c}"] \arrow[d, phantom, "\Downarrow"] & p_c \arrow[d, equal] \\
p_a \arrow[rr, "\mathbf{S}_{a,c}"'] & & p_c
\end{tikzcd}
\qquad
\begin{tikzcd}
p_a \otimes p_{a'} \arrow[r, "\mathbf{S}_{a,b} \otimes \mathbf{S}_{a',b'}"] \arrow[d, equal, "\wr"'] \arrow[dr, phantom, "\Downarrow"] & p_b \otimes p_{b'} \arrow[d, equal, "\wr"'] \\
p_{aa'} \arrow[r, "\mathbf{S}_{aa',bb'}"'] & p_{bb'}
\end{tikzcd}
\tag{7}
$$

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

Here the sets $S_{a,b}$ form the arrows in a strict monoidal category underlying 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 (7). 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.

Morphisms of adaptives are special cases of the general notion of morphisms for 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 in (8) in $\mathbb{O}\mathbf{rg}$ which commute with the identitor, compositor, and productor squares.

$$
\begin{tikzcd}
p^{\otimes m} \arrow[r, "\mathbf{S}_{m,n}"] \arrow[d, "\phi^{\otimes m}"'] \arrow[dr, phantom, "\Downarrow"] & p^{\otimes n} \arrow[d, "\phi^{\otimes n}"] \\
p'^{\otimes m} \arrow[r, "\mathbf{S}'_{m,n}"'] & p'^{\otimes n}
\end{tikzcd}
\tag{8}
$$

◊

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 as a monoidal bicategory (as done in [Spi21]). 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 $\mathbf{S}^{p!}$ on p where $\mathbf{S}^{p!}_{m,n}$ is the terminal $[p^{\otimes m}, p^{\otimes n}]$-coalgebra.

Its states are (not necessarily finite) $[p^{\otimes m}, p^{\otimes n}]$-trees: trees inductively defined by a root labeled with a polynomial map $\phi : p^{\otimes m} \to p^{\otimes n}$ and for each pair

$$
\Big((I_1, \ldots, I_m) \in p(1)^m, (i_1, \ldots, i_n) \in p[\phi(I_1, \ldots, I_m)_1] \times \cdots \times p[\phi(I_1, \ldots, I_m)_n]\Big) \tag{9}
$$

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 in (9) to the new tree at the end of the corresponding edge.

The idea is that the state-set of the terminal adaptive encodes all possible trajectories along different actions, and this coalgebra is terminal because from any other coalgebra there is a map to $\mathbf{S}^{p!}_{m,n}$ 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 because, 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 the terminal coalgebra $\mathbf{S}^{p!}$ 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 $\mathbf{S}^{p!}_{m,n}$ simply the unique map of coalgebras from any $[p^{\otimes m}, p^{\otimes n}]$-coalgebra to the terminal one. ◊

4 Adaptives in Nature

Our main results are that adaptives describe phenomena in the natural world and human technologies. In this paper, we focus on deep learning and gravity.

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.[4]

[4]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.

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 $x \in \mathbb{R}^m$, $\beta_0(M, f, p)_x : T_{f(p,x)}\mathbb{R}^n \to T_x\mathbb{R}^m$ sends y to $\pi_m Df^\top \cdot y$
- The update function $\beta_{M,f,p}$ sends $x \in \mathbb{R}^m$ and $y \in T_{f(p,x)}$ to $(M, f, p + \epsilon\pi_M Df^\top \cdot y)$ 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 productor 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 $\mathbf{S}_{\ell,m} \mathbin{⨟} \mathbf{S}_{m,n} \to \mathbf{S}_{\ell,n}$ send the pair $((L, f, p), (M, g, q))$ to

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

- 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

In pursuit of the idea of using adaptives to describe the dynamics of the world, we describe an adaptive that models how bodies move through space under the force of gravity. For simplicity, however, these bodies will not be prevented from overlapping so as to avoid encoding the complicated dynamics of potential collisions. This adaptive will be modeled as an $\mathbb{O}\mathbf{rg}$-enriched operad.

Definition 4.4. For the rest of this section, let C denote the unit cube in $\mathbb{R}^3$ and

$$S_1 = \{(s : C \to C, v \in T_0C) |\text{ s is a rectilinear embedding}\}.$$

An embedding $s : C \to C$ can be interpreted as a rectangular prism inside the unit cube with faces parallel to those of the outer cube. Let $S_n = S_1^n$, whose elements are made up of n such prisms $\vec{s} = (s_1, ..., s_n)$ and velocities $\vec{v} = (v_1, ..., v_n)$. ◊

These prisms inside the unit cube are meant to describe the positions and sizes of n bodies moving through space. In physical models they typically wouldn't be allowed to overlap, but for simplicity we will in effect only consider their centers of mass when computing gravitational effects.

Consider the polynomial

$$g = \sum_{m\in\mathbb{R}_{>0} c\in C} y^{t\in\mathbb{R}_{\geq 0}},$$

whose positions we interpret as the mass m of a body and its center of mass c, and whose directions we interpret as lengths of time t. We now describe an adaptive on g with state sets S_n, where the action of a state computes its center of mass and the update modifies the positions and velocities of the bodies according to their current velocities and the force of gravity.

Definition 4.5. For $m_1, ..., m_n \in \mathbb{R}_{>0}$ and $c_1, ..., c_n \in C$, representing n positions in the unit cube C and their masses, the vector pointing from the ith body to the center of mass of the others is given by

$$\Delta c_i = \frac{m_1 c_1 + \cdots + \widehat{m_i c_i} + \cdots m_n c_n}{m_1 + \cdots + \hat{m}_i + \cdots + m_n} - c_i.$$

The gravitational pull of the ith body in the system is given by

$$G_i(\vec{m}, \vec{c}) = \frac{m_i(m_1 + \cdots + \hat{m}_i + \cdots m_n)G}{|\Delta c_i|^3} \Delta c_i,$$

where G is the gravitational constant. We will write $\vec{G}(\vec{m}, \vec{c})$ for the tuple of these n different vectors. ◊

Definition 4.6. The $[g^{\otimes n}, g]$-coalgebra structure on S_n is given by

- $\beta_0(\vec{s},\vec{v})_1 : \mathbb{R}^n_{>0}C^n \longrightarrow C$ sends $(m_1, \ldots, m_n, c_1, \ldots, c_n)$ to

$$c_{\vec{s},\vec{m}} = \frac{m_1 s_1(c_1) + \cdots + m_n s_n(c_n)}{m_1 + \cdots + m_n}$$

- $\beta_0(\vec{s},\vec{v})_{\vec{m},\vec{c}} : \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}^n_{\geq 0}$ sends t to $(t, \ldots, t)$[5]
- The update function $\beta_{\vec{s},\vec{v}}$ sends $(\vec{m} \in \mathbb{R}^n_{>0}, \vec{c} \in C^n)$ and $t \in \mathbb{R}_{\geq 0}$ to

$$\left(\vec{s} + t\vec{v} + \frac{t^2}{2}\vec{G}(\vec{m}, \vec{s}(\vec{c})), \vec{v} + t\vec{G}(\vec{m}, \vec{s}(\vec{c}))\right)$$

◊

This coalgebra structure is designed so that n moving bodies are agglomerated into a single body by juxtaposition, with the center of mass of the new body computed as the center of mass of those that constitute it. Then for a given length of time, the positions and velocities of the original bodies are updated appropriately according to their original velocities and the force of gravity. The operadic composition of these states describe how to rescale these velocities according to the perspective of fitting their outer cube frame into a prism inside yet another outer cube.

- The identitor state in S_1 is given by $(\mathrm{id}_C, 0)$
- The operadic compositor sends

$$(\vec{s},\vec{v}) \in S_n, (\vec{s}^1, \vec{v}^1) \in S_{\ell_1}, \ldots, (\vec{s}^n, \vec{v}^n) \in S_{\ell_n}$$

to

$$\left(s_1 \circ s^1_1, \ldots, s_1 \circ s^1_{\ell_1}, s_2 \circ s^2_1, \ldots, s_n \circ s^n_{\ell_n}, v_1 + Ds_1 \cdot v^1_1, \ldots, v_1 + Ds_1 \cdot v^1_{\ell_1}, v_2 + Ds_2 \cdot v^2_1, \ldots, v_n + Ds_n \cdot v^n_{\ell_n}\right)$$

Intuitively, the identitor describes a body filling the entire cube and not moving, while the operadic compositor describes plugging in n systems into the prisms of another system. When doing so, the resulting prisms are precisely the composite embeddings of a prism of an inner system into the cube that is then again embedded into the cube as a prism in the outer system. The resulting velocities combine the velocity of the inner body in the cube of the inner system, scaled according to how that cube embeds into the cube of the outer system, then added to the velocity of the entire prism in the outer system. In other words, it is the velocity of a body moving inside a box which is itself moving.

Theorem 4.7. *These structure maps are maps of coalgebras and satisfy the coherence equations of an adaptive described in Definition A.1.*

This is discussed in more detail in Appendix A, though the main thrust of the proof is entirely given by physical principles. The identitor picks out a static state in S_1, as it has zero velocity, and hence is a coalgebra morphism. It also clearly acts as an identity with respect to the compositor, as plugging an inner system into an outer system which is just one prism occupying the entire cube recovers the inner system.

[5] This simply says that when the entire system is observed for a length of time t, so is each component. Perhaps this could be varied to incorporate some version of relativity.

The compositor is a coalgebra because the center of mass construction is "associative" in the sense of being computable in stages according to any (even nested) partition of the contributing masses, and the update is the physical evolution of the bodies in the system under their velocities and gravity, which is not affected by abstraction barriers such as drawing or removing boxes around some of the masses.

Brandon says: Is this true? when the systems are rescaled it will change the distances and hence the force. This could be fine if the scaling of the velocities and distance traveled depends in the same way as the force does... but applying s on the outside scales G(m,c) once whereas applying s on the inside scales distances quadratically... this seems worrying. $\diamond$

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

$$
\begin{tikzcd}
p^{\otimes n} \otimes p^{\otimes n'} \arrow[r, "\{\mathrm{id}_{p^{\otimes n}}\} \otimes \{\mathrm{id}_{p^{\otimes n'}}\}"] \arrow[d, equal] \arrow[dr, phantom, "\Downarrow"] & p^{\otimes n} \otimes p^{\otimes n'} \arrow[d, equal] \\
p^{\otimes n} \otimes p^{\otimes n'} \arrow[r, "\mathbf{s}_{n,n} \otimes \mathbf{s}_{n',n'}"] \arrow[d, equal, "\wr"'] \arrow[dr, phantom, "\Downarrow"] & p^{\otimes n} \otimes p^{\otimes n'} \arrow[d, equal, "\wr"'] \\
p^{\otimes (n+n')} \arrow[r, "\mathbf{s}_{n+n',n+n'}"'] & p^{\otimes (n+n')}
\end{tikzcd}
=
\begin{tikzcd}
p^{\otimes n} \otimes p^{\otimes n'} \arrow[r, "\{\mathrm{id}_{p^{\otimes n}}\} \otimes \{\mathrm{id}_{p^{\otimes n'}}\}"] \arrow[d, equal, "\wr"'] \arrow[dr, phantom, "\wr\Vert"] & p^{\otimes n} \otimes p^{\otimes n'} \arrow[d, equal, "\wr"'] \\
p^{\otimes (n+n')} \arrow[r, "\{\mathrm{id}_{p^{\otimes (n+n')}}\}"] \arrow[d, equal] \arrow[dr, phantom, "\Downarrow"] & p^{\otimes (n+n')} \arrow[d, equal] \\
p^{\otimes (n+n')} \arrow[r, "\mathbf{s}_{n+n',n+n'}"'] & p^{\otimes (n+n')}
\end{tikzcd}
\tag{10}
$$

- The compositor interchange law

$$
\begin{tikzcd}
p^{\otimes \ell} \otimes p^{\otimes \ell'} \arrow[r, "\mathbf{s}_{\ell,m} \otimes \mathbf{s}_{\ell',m'}"] \arrow[d, equal, "\wr"'] \arrow[dr, phantom, "\Downarrow"] & p^{\otimes m} \otimes p^{\otimes m'} \arrow[r, "\mathbf{s}_{m,n} \otimes \mathbf{s}_{m',n'}"] \arrow[d, equal, "\wr"'] \arrow[dr, phantom, "\Downarrow"] & p^{\otimes n} \otimes p^{\otimes n'} \arrow[d, equal, "\wr"'] \\
p^{\otimes (\ell+\ell')} \arrow[r, "\mathbf{s}_{\ell+\ell',m+m'}"'] \arrow[d, equal] & p^{\otimes (m+m')} \arrow[r, "\mathbf{s}_{m+m',n+n'}"'] \arrow[d, phantom, "\Downarrow"] & p^{\otimes (n+n')} \arrow[d, equal] \\
p^{\otimes (\ell+\ell')} \arrow[rr, "\mathbf{s}_{\ell+\ell',n+n'}"'] & & p^{\otimes (n+n')}
\end{tikzcd}
=
\begin{tikzcd}
p^{\otimes \ell} \otimes p^{\otimes \ell'} \arrow[r, "\mathbf{s}_{\ell,m} \otimes \mathbf{s}_{\ell',m'}"] \arrow[d, equal] & p^{\otimes m} \otimes p^{\otimes m'} \arrow[r, "\mathbf{s}_{m,n} \otimes \mathbf{s}_{m',n'}"] \arrow[d, phantom, "\Downarrow"] & p^{\otimes n} \otimes p^{\otimes n'} \arrow[d, equal] \\
p^{\otimes \ell} \otimes p^{\otimes \ell'} \arrow[rr, "\mathbf{s}_{\ell,n} \otimes \mathbf{s}_{\ell',n'}"] \arrow[d, equal, "\wr"'] & \arrow[d, phantom, "\Downarrow"] & p^{\otimes n} \otimes p^{\otimes n'} \arrow[d, equal, "\wr"'] \\
p^{\otimes (\ell+\ell')} \arrow[rr, "\mathbf{s}_{\ell+\ell',n+n'}"'] & & p^{\otimes (n+n')}
\end{tikzcd}
\tag{11}
$$

- The compositor associativity law

$$
\begin{tikzcd}
p^{\otimes k} \arrow[r, "\mathbf{S}_{k,\ell}"] \arrow[d, equal] & p^{\otimes \ell} \arrow[r, "\mathbf{S}_{\ell,m}"] \arrow[d, phantom, "\Downarrow"] & p^{\otimes m} \arrow[r, "\mathbf{S}_{m,n}"] \arrow[d, equal] & p^{\otimes n} \arrow[d, equal] \\
p^{\otimes k} \arrow[rr, "\mathbf{S}_{k,m}"'] \arrow[d, equal] & {} \arrow[d, phantom, "\Downarrow"] & p^{\otimes m} \arrow[r, "\mathbf{S}_{m,n}"'] & p^{\otimes n} \arrow[d, equal] \\
p^{\otimes k} \arrow[rrr, "\mathbf{S}_{k,n}"'] & & & p^{\otimes n}
\end{tikzcd}
=
\begin{tikzcd}
p^{\otimes k} \arrow[r, "\mathbf{S}_{k,\ell}"] \arrow[d, equal] & p^{\otimes \ell} \arrow[r, "\mathbf{S}_{\ell,m}"] \arrow[d, equal] & p^{\otimes m} \arrow[r, "\mathbf{S}_{m,n}"] \arrow[d, phantom, "\Downarrow"] & p^{\otimes n} \arrow[d, equal] \\
p^{\otimes k} \arrow[r, "\mathbf{S}_{k,\ell}"'] \arrow[d, equal] & p^{\otimes \ell} \arrow[rr, "\mathbf{S}_{\ell,n}"'] \arrow[d, phantom, "\Downarrow"] & & p^{\otimes n} \arrow[d, equal] \\
p^{\otimes k} \arrow[rrr, "\mathbf{S}_{k,n}"'] & & & p^{\otimes n}
\end{tikzcd}
\tag{12}
$$

- The compositor unit laws

$$
\begin{tikzcd}
p^{\otimes m} \arrow[r, "\{\mathrm{id}_{p^{\otimes m}}\}"] \arrow[d, equal] & p^{\otimes m} \arrow[r, "\mathbf{S}_{m,n}"] \arrow[d, equal] & p^{\otimes n} \arrow[d, equal] \\
p^{\otimes m} \arrow[r, "\mathbf{S}_{m,m}"'] \arrow[d, equal] & p^{\otimes m} \arrow[r, "\mathbf{S}_{m,n}"'] \arrow[d, phantom, "\Downarrow"] & p^{\otimes n} \arrow[d, equal] \\
p^{\otimes m} \arrow[rr, "\mathbf{S}_{m,n}"'] & & p^{\otimes n}
\end{tikzcd}
=
\begin{tikzcd}
p^{\otimes m} \arrow[r, "\mathbf{S}_{m,n}"] \arrow[d, equal] & p^{\otimes n} \arrow[d, equal] \\
p^{\otimes m} \arrow[r, "\mathbf{S}_{m,n}"'] & p^{\otimes n}
\end{tikzcd}
=
\begin{tikzcd}
p^{\otimes m} \arrow[r, "\mathbf{S}_{m,n}"] \arrow[d, equal] & p^{\otimes n} \arrow[r, "\{\mathrm{id}_{p^{\otimes m}}\}"] \arrow[d, equal] & p^{\otimes n} \arrow[d, equal] \\
p^{\otimes m} \arrow[r, "\mathbf{S}_{m,n}"'] \arrow[d, equal] & p^{\otimes n} \arrow[r, "\mathbf{S}_{n,n}"'] \arrow[d, phantom, "\Downarrow"] & p^{\otimes n} \arrow[d, equal] \\
p^{\otimes m} \arrow[rr, "\mathbf{S}_{m,n}"'] & & p^{\otimes n}
\end{tikzcd}
\tag{13}
$$

- The productor associativity law

$$
\begin{tikzcd}
p^{\otimes m} \otimes p^{\otimes m'} \otimes p^{\otimes m''} \arrow[rr, "\mathbf{S}_{m,n} \otimes \mathbf{S}_{m',n'} \otimes \mathbf{S}_{m'',n''}"] \arrow[d, equal, "\wr"'] & {} \arrow[d, phantom, "\Downarrow"] & p^{\otimes n} \otimes p^{\otimes n'} \otimes p^{\otimes n''} \arrow[d, equal, "\wr"'] \\
p^{\otimes (m+m')} \otimes p^{\otimes m''} \arrow[rr, "\mathbf{S}_{m+m',n+n'} \otimes \mathbf{S}_{m'',n''}" description] \arrow[d, equal, "\wr"'] & {} \arrow[d, phantom, "\Downarrow"] & p^{\otimes (n+n')} \otimes p^{\otimes n''} \arrow[d, equal, "\wr"'] \\
p^{\otimes (m+m'+m'')} \arrow[rr, "\mathbf{S}_{m+m'+m'',n+n'+n''}"'] & & p^{\otimes (n+n'+n'')}
\end{tikzcd}
=
\begin{tikzcd}
p^{\otimes m} \otimes p^{\otimes m'} \otimes p^{\otimes m''} \arrow[rr, "\mathbf{S}_{m,n} \otimes \mathbf{S}_{m',n'} \otimes \mathbf{S}_{m'',n''}"] \arrow[d, equal, "\wr"'] & {} \arrow[d, phantom, "\Downarrow"] & p^{\otimes n} \otimes p^{\otimes n'} \otimes p^{\otimes n''} \arrow[d, equal, "\wr"'] \\
p^{\otimes m} \otimes p^{\otimes (m+m'')} \arrow[rr, "\mathbf{S}_{m,n} \otimes \mathbf{S}_{m'+m'',n'+n''}" description] \arrow[d, equal, "\wr"'] & {} \arrow[d, phantom, "\Downarrow"] & p^{\otimes n} \otimes p^{\otimes (n'+n'')} \arrow[d, equal, "\wr"'] \\
p^{\otimes (m+m'+m'')} \arrow[rr, "\mathbf{S}_{m+m'+m'',n+n'+n''}"'] & & p^{\otimes (n+n'+n'')}
\end{tikzcd}
\tag{14}
$$

- The productor unit laws

$$
\begin{tikzcd}
p^{\otimes m} \arrow[rr, "\mathbf{S}_{m,n}"] \arrow[d, equal, "\wr"'] & {} \arrow[d, equal, "\wr"'] & p^{\otimes n} \arrow[d, equal, "\wr"'] \\
p^{\otimes 0} \otimes p^{\otimes m} \arrow[rr, "\{\mathrm{id}_{p^{\otimes 0}}\} \otimes \mathbf{S}_{m,n}" description] \arrow[d, equal] & {} \arrow[d, phantom, "\Downarrow"] & p^{\otimes 0} \otimes p^{\otimes n} \arrow[d, equal] \\
p^{\otimes 0} \otimes p^{\otimes m} \arrow[rr, "\mathbf{S}_{0,0} \otimes \mathbf{S}_{m,n}" description] \arrow[d, equal, "\wr"'] & {} \arrow[d, phantom, "\Downarrow"] & p^{\otimes 0} \otimes p^{\otimes n} \arrow[d, equal, "\wr"'] \\
p^{\otimes m} \arrow[rr, "\mathbf{S}_{m,n}"'] & {} & p^{\otimes n}
\end{tikzcd}
=
\begin{tikzcd}
p^{\otimes m} \arrow[r, "\mathbf{S}_{m,n}"] \arrow[d, equal] & p^{\otimes n} \arrow[d, equal] \\
p^{\otimes m} \arrow[r, "\mathbf{S}_{m,n}"'] & p^{\otimes n}
\end{tikzcd}
=
\begin{tikzcd}
p^{\otimes m} \arrow[rr, "\mathbf{S}_{m,n}"] \arrow[d, equal, "\wr"'] & {} \arrow[d, equal, "\wr"'] & p^{\otimes n} \arrow[d, equal, "\wr"'] \\
p^{\otimes m} \otimes p^{\otimes 0} \arrow[rr, "\mathbf{S}_{m,n} \otimes \{\mathrm{id}_{p^{\otimes 0}}\}" description] \arrow[d, equal] & {} \arrow[d, phantom, "\Downarrow"] & p^{\otimes n} \otimes p^{\otimes 0} \arrow[d, equal] \\
p^{\otimes m} \otimes p^{\otimes 0} \arrow[rr, "\mathbf{S}_{m,n} \otimes \mathbf{S}_{0,0}" description] \arrow[d, equal, "\wr"'] & {} \arrow[d, phantom, "\Downarrow"] & p^{\otimes n} \otimes p^{\otimes 0} \arrow[d, equal, "\wr"'] \\
p^{\otimes m} \arrow[rr, "\mathbf{S}_{m,n}"'] & {} & p^{\otimes n}
\end{tikzcd}
\tag{15}
$$

◊

We now proceed to prove that the coalgebras and structure maps defined above for gradient descent and gravitational motion form adaptives. In each case, it suffices to show that the structure maps on states preserve coalgebra structure, and that equations (10) through (15) are satisfied.

Proof of Theorem 4.3. The unit and associativity equations follow immediately from associativity and unitality of addition, cartesian products, and function composition. The interchange equations follow from the preservation of 0 under addition and identity functions under cartesian products, the analogous interchange property of function composition and cartesian products of functions, and the fact that the compositors and productors act the same way on the parameters and their dimension.

It then remains only to show that the identitors, compositors, and productors are morphisms of coalgebras. This is immediate for the productors, as each component of the action and update functions respects the cartesian products of functions and parameters that define them, so we proceed only for the identitors and compositors.

For the identitors, 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 $\mathbf{S}_{n,n}$ 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.

The compositors preserve the component of the action on positions as, for states

$$(L \in \mathbb{N}, f : \mathbb{R}^{L+\ell}, p \in \mathbb{R}^L) \quad \text{and} \quad (M \in \mathbb{N}, g : \mathbb{R}^{M+m}, q \in \mathbb{R}^M),$$

we have

$$g(q, -) \circ f(p, -) = g(-, f(-, -))(q, p, -).$$

This may seem like a trivial rewriting, but it illustrates how the compositor was defined in order for the action to be preserved, as on the left we have the composite of the actions on positions as in $\mathbf{S}_{\ell,m}; \mathbf{S}_{m,n}$, and on the right we apply the compositor and take the action of the resulting state in $\mathbf{S}_{\ell,n}$.

To show that the compositor preserves both the action on directions and the update we note that by the chain rule, for $x \in \mathbb{R}^\ell$ and $z \in T_{g(q, f(p,x))}$,

$$D\Big(g(-, f(-, -))\Big)^\top z = Df^\top \cdot \pi_m(Dg^\top \cdot z) \in T_{(p,x)}\mathbb{R}^{L+\ell}.$$

Applying π_ℓ to both sides above shows that the compositor preserves the action on directions, as on the left we have the action on directions after applying the compositor and on the right we have the composition of the actions of (L, f, p) and (M, g, q) on directions as in $\mathbf{S}_{\ell,m}; \mathbf{S}_{m,n}$.

Finally for updates, we observe by the chain rule that the update rule in $\mathbf{S}_{\ell,n}$ agrees with that in $\mathbf{S}_{\ell,m}; \mathbf{S}_{m,n}$ under the compositor, as either way for x, z as above the composite state of (L, f, p) and (M, g, q) updates to

$$\Big(M + L, g(-, f(-, -)), \Big(p + \epsilon \pi_L(Df^\top \cdot \pi_m(Dg^\top \cdot z)), q + \epsilon \pi_M(Dg^\top \cdot z)\Big)\Big).$$

□

References

[Lei99] Tom Leinster. *Generalized enrichment for categories and multicategories*. arXiv:9901139. 1999. arXiv: 9901139 [math.CT] (cit. on pp. 5, 6).

[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. 7).

[Sha22] Brandon Shapiro. *Enrichment of Algebraic Higher Categories*. *In Preparation*. 2022 (cit. on pp. 5, 7, 8).

[Spi21] David I. Spivak. "Learners' languages". In: *Proceedings of the 4th Annual Conference on Applied Category Theory*. ACT. Cambridge, UK: EPTCS, 2021 (cit. on pp. 1–4, 8).