## 0. Families, Limits & Colimits in Set

families of subsets
families of sets
total set, set of sections
families as functions
families with transitivity, intersection, unions
example: formal power series, 2-adic numbers, Prüfer group
diagram, limits, colimits

# Part 1

✓ 1. Theories
monoids
abbreviating global uniform operation
theory

✓ 2. Models
example: $\mathrm{Vec}_{\mathbb{R}}$
variable notation $\omega(x, y)$

✓ Extended Example: Smooth Algebras

Limits & Colimits of Models
internal models (submodels)
internal equations (congruences)
factorization, isomorphism reflection
products
equalizers
intro to coequalizers
Todo: forgetful symmetry → reflexive coequalizers
directed unions

✓ 3. Theory Morphisms
left adjoints (only an unmotivated proof sketch)
! shoddy sketch

# Part 2

✓ 4. Effects & Monads
example: partial functions
example: multivalued functions
functional effect
category of effects

2nd axiomatization of functional effect
example: finitary non-determinism with multiplicities
example: finitary non-determinism without multiplicities
example: unbounded non-determinism without multiplicities
example: linear combinations

reduction operation
motivating neutrality & associativity
monads on Set
monads ≃ functional effect
example: $\mathbb{E}$ linear combinations, Bool linear combinations
Bool relations, $\mathbb{E}$ relations

✓ 5. Monoidal Cats
bifunctor
monoids and their actions in Set
monoids and their actions in $\mathbb{E}, \otimes$
monoidal functor, monoidal action
monoid of $\mathbb{E}, \otimes$ acting on $\mathcal{X}$ through $\otimes$
endofunctors, endofunctors
monads on any category
monad models
monoid effects
set monoids are embedded into monoids in Set

✓ 6. Monads vs. Theories
Eilenberg-Moore
Monad Model, Functional Effect Model
examples
monad operations
monads as theories
! A little study → category of effects is category of operations and arities

canonical presentation
coproducts

✓ 7. Monadicity
split coequalizers
U-split adjunctions
failure of canonical presentation
monadicity

# Part 3

✓ 8. Extended Example: Possibilistic Spaces

✓ Commutative Theories
commuting operations
commutation of operations $\omega \circ \eta$
central operations
commutative models

commutative theories

Hom & Bihomomorphisms
Tensor Product
Symmetric Closed Monoidal
Models with Monoid Structure

## Keywords

Adjoint Functor
Algebraic Theory
Lawvere Theory
Eilenberg-Moore Category
Kleisli Category
Triples
Monads
Monadicity
Monoidal Categories
Giry Monad
Tensor Product
Eckmann-Hilton

## References


Category Theory in Context, Emily Riehl

Algebraic Theories: a categorical introduction to general algebra, J. Adámek, J. Rosický, E.M. Vitale
Functorial Semantics of Algebraic Theories, F. William Lawvere

Basic Concepts of Synthetic Differential Geometry, Lavendhomme
Synthetic Differential Geometry, Anders Kock
Synthetic Differential Geometry of Manifolds, Anders Kock
Models for Smooth Infinitesimal Analysis, Moerdijk, Reyes
Smooth Manifolds and Observables, Jet Nestruev

New Approach to Arakelov Geometry, Nikolai Durov

How to lose your fear of Tensor Products, Timothy Gowers
Codensity and the Giry Monad, Tom Avery

# Families

Now we'll develop the often neglected concept of a **family** of objects varying over some object to better understand various notions of **limits** and **colimits**, and create basic building blocks for construction of **multisorted algebraic theories**.

## Subset valued Families

**define** Let $X : \mathbf{Set}$, then an $X$-indexed **family of subsets** of $T : \mathbf{Set}$ is a map

$$A : X \longrightarrow \mathcal{P}[T] : \mathbf{Set}$$

i.e. to each $x : X$ we associate a **subset** $A_x \subseteq T$. ■

Given such a **family**, we have its **union** & **intersection**

**define** For $A : X \to \mathcal{P}[T] : \mathbf{Set}$,

$$\bigcup_{x:X} A_x := \left\{ t : T \;\middle|\; \exists x : X,\ t : A_x \right\}$$

$$\bigcap_{x:X} A_x := \left\{ t : T \;\middle|\; \forall x : X,\ t : A_x \right\} ■$$

These two **operations** are characterized by **universal properties** of **join** and **meet**.

**prop** Given $A : X \to \mathcal{P}[T] : \mathbf{Set}$, then

(join) $\bigcup_{x:X} A_x$ is the **least upper bound** of $A$, that is

(1) $\forall x : X$,
$$A_x \subseteq \bigcup_{y:X} A_y$$

(2) $\forall B \subseteq T$,
if $\forall x : X$,
$$A_x \subseteq B$$
then $\bigcup_{x:X} A_x \subseteq B$.

(meet) $\bigcap_{x:X} A_x$ is the **greatest lower bound** of $A$, that is

(1) $\forall x : X$,
$$\bigcap_{y:X} A_y \subseteq A_x$$

(2) $\forall B \subseteq T$,
if $\forall x : X$,
$$B \subseteq A_x$$
then $B \subseteq \bigcap_{x:X} A_x$ ■

By changing our conception of **family**, we change the corresponding notions of **limits** and **colimits**; we may even call **set-indexed, subset valued families** as **families internal** to $T$, — we know precisely how each $A_x$ is **embedded** into $T$.

**example** Given $T : \mathbf{Set}$, for any two **subsets** $A, B \subseteq T$, it makes sense to talk about their **union** $A \cup B \subseteq T$ — it is the **least subset containing** $A$ and $B$.

This is possible because $T$ provides us with canonical comparison of two **subsets** $\subseteq$, this additional structure allows any two **subsets** to know their mutual relationships to each other.

On the other hand given $A, B : \mathbf{Set}$, does it really make sense to talk about their **union**?
Given only two bare sets we do not know how they stand in relation to each other; it is meaningless to write $A \subseteq B$ for arbitrary **sets**. The best we can do is assume that they have no non-trivial **canonical** relationship. So we can think of them as **disjoint**.
For this reason, without additional structure, we can only talk about their **disjoint union** $A + B : \mathbf{Set}$.
Then $A, B$ are **embedded** into $A + B$ via $A \xrightarrow{i} A + B \xleftarrow{j} B$.

As an example take **symbols** $x, y, z, w$, $a, b, c$ and consider **sets**

$$A := \{x, y, z, w\}, \quad B := \{a, b, c\}$$

Then the **inclusions** $i, j$ generate new **symbols** so that

$$A + B := \{ i(x),\ i(y),\ i(z),\ i(w),\ j(a),\ j(b),\ j(c) \};$$

this works as intended, since we have not related $A$ & $B$ in any way.

Now consider $B + B$. Even though $B$ is related to itself in a **canonical** way via **identity**, that is even though $b = b$ we **can't** assert $i(b) = j(b)$, so that $B + B = \{ i(a),\ i(b),\ i(c),\ j(a),\ j(b),\ j(c) \}$ is a 6-element **set**, not a 3-element **set**.

**Disjoint union** should work in the same **uniform** way for all possible pairs of **sets**; it can't look for possible relationships between $A$ & $B$ and always choose the right, canonical one, since there may not be the canonical relationship — that's the job of the mathematician who uses $A$ & $B$.

To talk about **union** of $A, B : \mathbf{Set}$ we require additional structure. For example we can choose an **intersection** of $A$ & $B$ by **identifying** some elements of $A$ with some elements of $B$.

This **identification** can be given with a **set** $S$ and a pair of maps

$$f : S \longrightarrow A,\ g : S \longrightarrow B : \mathbf{Set},$$

then we define

$$A \sqcup_S B := \frac{A + B}{\{ f(s) = g(s) \mid s : S \}}$$

This construction is usually called the **pushout** of $A$ and $B$ **along** $S$ ■

Now we can feel the additional complexity of going from

$A, B : \mathcal{P}[T]$ a **poset** to $A, B : \mathbf{Set}$ a **category**.

We remind the reader about the **universal properties** of **disjoint union (coproduct)** & **cartesian product**; you may even take them as **definitions** of $A + B$ & $A \times B$.

$$\mathcal{P}[T],\ \cup,\ \cap \quad \rightsquigarrow \quad \mathbf{Set},\ +,\ \times$$

They can be formulated over arbitrary **category** $\mathbb{E}$

**define** (Product) Let $A, B : \mathbb{E}$, then **cartesian product** of $A$ & $B$ consists of

$$A \times B : \mathbb{E} \quad \text{with} \quad \pi_A : A \times B \to A,\ \pi_B : A \times B \to B : \mathbb{E} \quad \text{(projections)}$$

where $\forall Z : \mathbb{E}$,

$$\forall f : Z \to A,\ g : Z \to B : \mathbb{E},$$
$$\exists ! \langle f, g \rangle : Z \to A \times B : \mathbb{E},$$
$$\pi_A \langle f, g \rangle = f,$$
$$\pi_B \langle f, g \rangle = g \ ■$$

We can think of the **unique pairing** $\langle f, g \rangle$ of $f$ & $g$ in $\mathbf{Set}$ as

$$\langle f, g \rangle(z) := \langle f(z), g(z) \rangle$$

**define** (Coproduct) Let $A, B : \mathbb{E}$, then the **disjoint union** of $A$ & $B$ consists of

$$A + B : \mathbb{E} \quad \text{with} \quad i_A : A \to A + B,\ i_B : B \to A + B : \mathbb{E} \quad \text{(inclusions)}$$

where $\forall Z : \mathbb{E}$,

$$\forall f : A \to Z,\ g : B \to Z : \mathbb{E},$$
$$\exists ! (f | g) : A + B \to Z : \mathbb{E},$$
$$i_A (f | g) = f,$$
$$i_B (f | g) = g \ ■$$

We can think of the **unique conditional (switch)** $(f | g)$ of $f$ & $g$ in $\mathbf{Set}$ as

$$(f | g)(t) := \begin{cases} f(a) & \text{if } t = i_A(a) \\ g(b) & \text{if } t = i_B(b) \end{cases}$$

Read "$f$ or $g$" for $(f | g)$.
Diagrammatically we have

$$\begin{tikzcd} & Z & \\ A \times B \arrow[ur, "\langle f, g \rangle"] \arrow[r, "\pi_B"'] & B \arrow[u, "g"'] \arrow[l, "\pi_A"] & \\ A \arrow[u, "f"] \arrow[ur, "i_A"'] & & \end{tikzcd} \qquad \begin{tikzcd} A \arrow[r, "i_A"] \arrow[dr, "f"'] & A + B \arrow[d, "(f|g)"] & B \arrow[l, "i_B"'] \arrow[dl, "g"] \\ & Z & \end{tikzcd}$$

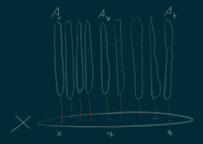
# Set valued families

**define** Given $X:Set$, by an $X$-indexed family of sets we mean a function

$$A_\cdot : X \longrightarrow Set\,;$$

that is to each $x:X$, we associate a set $A_x$ ■

It is often advantageous to visualize such a family as a sort of sequence or a bundle of sets varying over elements of $X$




For this reason we'll use the words bundle and family interchangeably.

We call individual set $A_x$ at $x:X$ the fiber of the bundle at $x$.

From each such family we can construct two sets.

## Total Set - Dependent Pairs - Disjoint Union

First one is bundling up of fibers into one set

$$\sum_{x:X} A_x : Set \quad \left[\text{or } \sum x:X.\, A_x\right]$$

We'll call it the total set of $A_\cdot : X \longrightarrow Set$

Its elements are pairs $\langle x, a\rangle$ where $x:X$ and $a:A_x$.

We call $\langle x, a\rangle : \sum_{x:X} A_x$ a dependently typed pair since

the type of the second component depends on the first component (not its type).

Since each element knows on which fiber it lies we get a projection map

$$p : \sum_{x:X} A_x \longrightarrow X$$
$$\langle x, a\rangle \longmapsto x$$

This need not be surjective - we may choose empty sets for some fibers

**example** Let $A:Set$, then the total space of the family

$$\mathbb{N} \longrightarrow Set$$
$$n \longmapsto A^n$$

is the set of finite sequences (lists) over $A$

$$\sum_{n:\mathbb{N}} A^n = List[A].$$

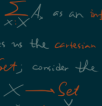
$$\langle n, \langle a_1, \dots, a_n\rangle\rangle \longleftrightarrow [a_1, \dots, a_n]$$

The bundle projection corresponds to the

$$length : List[A] \longrightarrow \mathbb{N}$$ ■

**example** A family indexed by $[2] = \{0, 1\}$ is simply a choice of two sets $A, B$.

The total space construction gives the disjoint union $A+B$.




■

We can think of $\sum_{x:X} A_x$ as an infinitary generalization of disjoint union.

**example** But it also gives us the cartesian product of two sets.

Let $X, Y:Set$, consider the constant family (trivial bundle)

$$X \longrightarrow Set$$
$$x \longmapsto Y$$

Its total set is the cartesian product

$$\sum_{x:X} Y = X \times Y,$$

since the type of the second component $y$ of the pair $\langle x, y\rangle : \sum_{x:X} Y$ does not depend (depends trivially) on the first component $x$.

Here we even get a nice projection on the second component, which we do not have in the general case

■

## Sections - Dependent Functions - Cartesian Product

A dependently typed function $s$ of a family $A_\cdot : X \to Set$ associates to each $x:X$ an element $s_x : A_x$, that is the type of the output depends on the input.

Equivalently, it is a map

$$s : X \longrightarrow \sum_{x:X} A_x$$

such that $\forall x:X.$

$$p(s(x)) = x\,;$$

we also call such a map a bundle section since it is a section of the bundle projection $p : \sum_{x:X} A_x \longrightarrow X$.

$$p\,s = id_X$$

If some of the fibers are empty, then the family can't have any sections.

We denote the set of all sections by

$$\forall_{x:X} A_x \quad \text{or just } \forall x:X.\, A_x$$

The choice of notation might appear to be strange at first, but it is very convenient.

Constructing

$$\text{family } A_\cdot : X \to Set \rightsquigarrow \begin{array}{l} \text{bundle } \sum_{x:X} A_x : Set \\ \text{sections } \forall_{x:X} A_x : Set \end{array}$$

are analogues of classical existential & universal quantifiers of predicates on sets.

$$\text{predicate } P : X \to Bool \rightsquigarrow \begin{array}{l} \exists x:X.\, P(x) : Bool \\ \forall x:X.\, P(x) : Bool\,; \end{array}$$

Note the analogy:

to give an element of the set $\sum_{x:X} A_x$ $\simeq$ to prove the proposition $\exists x:X.\, P(x)$

is to give $x:X$ together with some element of $A_x$ is to give $x:X$ together with with a proof of the proposition $P(x)$

to give an element of the set $\forall_{x:X} A_x$ $\simeq$ to prove the proposition $\forall x:X.\, P(x)$

is to give for each $x:X$ an element of $A_x$ is to give for each $x:X$ a proof of the proposition $P(x)$

Much more can be said about this analogy, but we have no use for it here - search Curry-Howard Correspondence if you are interested.

**example** Addition of real numbers is a binary operation

$$+ : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}.$$

But we can add finitely many real numbers, so we define

$$+_3 : \mathbb{R}^3 \longrightarrow \mathbb{R}, \quad +_4 : \mathbb{R}^4 \longrightarrow \mathbb{R}, \quad \dots$$
$$\langle x, y, z\rangle \longmapsto x+y+z$$

Special cases

$$+_0 : \mathbb{R}^0 \longrightarrow \mathbb{R}, \quad +_1 : \mathbb{R}^1 \longrightarrow \mathbb{R}, \quad +_2 := + : \mathbb{R}^2 \to \mathbb{R}.$$
$$\langle\rangle \longmapsto 0 \qquad x \longmapsto x$$

Then for each $n:\mathbb{N}$ we have $+_n : \mathbb{R}^n \to \mathbb{R}$.

Hence we get a section

$$+ : \forall n:\mathbb{N}.\, \mathbb{R}^n \to \mathbb{R}$$

of an $\mathbb{N}$-indexed family that assigns to $n:\mathbb{N}$ the set of all functions $\mathbb{R}^n \to \mathbb{R}$. ■

**example** Consider the family $A_\cdot : [2] \longrightarrow Set$ for some fixed $A, B : Set$,

$$0 \longmapsto A_0 := A$$
$$1 \longmapsto A_1 := B,$$

then $\forall_{i:[2]} A_i = A \times B$

We can think of $\forall_{x:X} A_x$ as an infinitary generalization of cartesian product.

Thinking of it in this way one often uses the notation $\prod_{x:X} A_x$ in $\forall_{x:X} A_x$ ■

**example** Let $X, Y:Set$, then for sections of the constant family $x:X \to Y:Set$ we have

$$\forall_{x:X} Y = X \to Y,$$

so $\forall_{x:X} A_x$ is a kind of generalized function space ■

## Universal Characterization

**define** Given $A_\cdot, B_\cdot : X \longrightarrow Set$ two families,

Then a family map $f : A_\cdot \longrightarrow B_\cdot$ is given by

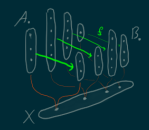
for each $x:X$,

a function $f_x : A_x \longrightarrow B_x : Set$. ■

But this is just a section of the family $x:X \mapsto A_x \to B_x$

$$f : \forall x:X.\, A_x \to B_x$$

Here we can really see the niceness of $\forall$ notation.




Now we can organize families over fixed indexing set $X$ as a category.

**define** Let $Set(X)$ be the category whose objects are $X$-indexed families, and morphisms are family maps. ■

We can easily see that we have functors

$$\sum_X : Set(X) \longrightarrow Set$$
$$A_\cdot \longmapsto \sum_{x:X} A_x$$

$$\forall_X : Set(X) \longrightarrow Set$$
$$A_\cdot \longmapsto \forall_{x:X} A_x$$

$$const : Set \longrightarrow Set(X)$$
$$A \longmapsto x:X \to A$$

Now note that given $A_\cdot : X \to Set$, $Z : Set$, we have that

(1) family map $f_\cdot : \forall x:X.\, A_x \longrightarrow Z$, that is $f_\cdot : A_\cdot \longrightarrow const(Z) : Set(X)$

correspond to

$$f : \sum_{x:X} A_x \longrightarrow Z : Set$$

(2) family map $f_\cdot : \forall x:X.\, Z \longrightarrow A_x$, that is $f_\cdot : const(Z) \longrightarrow A_\cdot : Set(X)$

correspond to $f : Z \longrightarrow \forall_{x:X} A_x : Set$.

These are the universal properties of $\sum_{x:X} A_x$ and $\forall_{x:X} A_x$; and since

$$\frac{A_\cdot \to const(Z) : Set(X)}{\sum_{x:X} A_x \to Z : Set} \quad \& \quad \frac{const(Z) \longrightarrow A_\cdot : Set(X)}{Z \longrightarrow \forall_{x:X} A_x : Set},$$

we get adjunctions $\sum_X \dashv const_X \dashv \forall_X$

We can formulate the universal properties as generalized disjoint sums and cartesian products.

**define** The set of sections (cartesian product) of the family $A_\cdot : X \longrightarrow Set$ consists of

$\forall_{x:X} A_x : Set$ with a family of maps for each $x:X$,

$$\pi_x : \forall_{x:X} A_x \longrightarrow A_x$$

where $\forall Z : Set$,

$\forall$ families of maps for each $x:X$

$$f_x : Z \to A_x,$$

$$\exists!\, f : Z \longrightarrow \forall_{x:X} A_x$$

$$\forall x:X.\quad \pi_x f = f_x$$ ■

Note that

$$\pi_\cdot : \forall x:X.\left[\forall_{y:Y} A_y \longrightarrow A_x\right]$$
$$\pi_x(s) := s_x$$

And for $f_\cdot : \forall x:X.\left[Z \to A_x\right]$ denote the unique $f : Z \to \forall_{x:X} A_x$ by

$$\langle f_\cdot \rangle_{x:X}$$

**define** The total space of the family $A_\cdot : X \longrightarrow Set$ consists of

$\sum_{x:X} A_x : Set$ with a family of maps for each $x:X$,

$$i_x : A_x \longrightarrow \sum_{x:X} A_x$$

where $\forall Z : Set$,

$\forall$ families of maps for each $x:X$

$$f_x : A_x \to Z,$$

$$\exists!\, f : \sum_{x:X} A_x \longrightarrow Z.$$

$$\forall x:X.\quad f\, i_x = f_x$$ ■

Note $i_\cdot : \forall x:X.\left[A_x \longrightarrow \sum_{y:Y} A_y\right]$

$$i_x(a) := \langle x, a\rangle$$

Todo How to denote the unique $f$?

We generalize from $\mathbf{Set}$ to $\mathbb{E}$ any category.

**define** Let $X:\mathbf{Set}$, then
$$A_\cdot : X \longrightarrow \mathbb{E} \text{ an } X\text{-indexed family of objects in } \mathbb{E}$$
is a map that to each $x:X$ associates some object $A_x:\mathbb{E}$. ■

**define** Let $X:\mathbf{Set}$, $A_\cdot, B_\cdot : X \longrightarrow \mathbb{E}$ be families,
then $f_\cdot : A_\cdot \longrightarrow B_\cdot$ a family of morphisms from $A_\cdot$ to $B_\cdot$
is a map that to each $x:X$ associates some morphism $f_x : A_x \longrightarrow B_x : \mathbb{E}$.
We can write $f_\cdot : \forall x:X.\ A_x \longrightarrow B_x : \mathbb{E}$
since morphism between any two objects in $\mathbb{E}$ form a set. ■

This defines category $\mathbb{E}(X)$ of $\mathbb{E}$-valued families over $X$.

**define** We have the constant family functor
$$\mathrm{const} : \mathbb{E} \longrightarrow \mathbb{E}(X)$$
$$A \rightsquigarrow _:X \to A$$ ■

**define** The total object of $A_\cdot : X \to \mathbb{E}$ consists of
$$\sum_{x:X} A_x : \mathbb{E} \text{ with } i_\cdot : \forall x:X.\ A_x \longrightarrow \sum_{x:X} A_x : \mathbb{E}$$
where $\forall Z:\mathbb{E}$
the map that to each
$$f_\cdot : \forall x:X.\ A_x \longrightarrow Z : \mathbb{E}$$
associates the family of maps
$$i_\cdot f_\cdot : \forall x:X.\ A_x \longrightarrow Z : \mathbb{E}$$
is invertible. ■

**define** The object of sections of $A_\cdot : X \to \mathbb{E}$ consists of
$$\forall_{x:X} A_x : \mathbb{E} \text{ with } \pi_\cdot : \forall x:X.\ \forall_{y:X} A_y \longrightarrow A_x : \mathbb{E}$$
where $\forall Z:\mathbb{E}$
the map that to each
$$f_\cdot : Z \longrightarrow \forall_{x:X} A_x : \mathbb{E}$$
associates the family of maps
$$\pi_\cdot f_\cdot : \forall x:X.\ Z \longrightarrow A_x : \mathbb{E}$$
is invertible. ■

## families as functions

An important example of families comes from set functions.

**example** Let $p : E \longrightarrow X : \mathbf{Set}$. Then we have the family
$$p^{-1} : X \longrightarrow \mathbf{Set}$$
$$x \longrightarrow p^{-1}(x) \coloneqq \{e:E \mid p(e) = x\}$$
Then $\sum_{x:X} p^{-1}(x) \cong E$ via
$$\langle x, e\rangle \longrightarrow e$$
$$\langle p(e), e\rangle \longleftarrow e.$$
And sections of the family $s : \forall_{x:X} p^{-1}(x)$ correspond to sections of $p$, that is
maps $s : X \longrightarrow E : \mathbf{Set}$ s.t. $p \circ s = \mathrm{id}_X$.

Starting from a family $A_\cdot : X \longrightarrow \mathbf{Set}$,
we have the projection $p : \sum_{x:X} A_x \longrightarrow X$,
and its family $p^{-1} : \mathbf{Set}(X)$ is isomorphic to $A_\cdot : \mathbf{Set}(A)$.
Recall that the slice category $\mathbf{Set}/X$
has as objects maps $p : E \to X$,
and morphism from $\begin{tikzcd} E \arrow[dr, "p"'] & \\ & X \end{tikzcd}$ to $\begin{tikzcd} F \arrow[dl, "q"] \\ X \end{tikzcd}$
is a map $f : E \to F$ s.t. $\begin{tikzcd} E \arrow[rr, "f"] \arrow[dr, "p"'] & & F \arrow[dl, "q"] \\ & X & \end{tikzcd}$

Such morphisms correspond to family maps $f : p^{-1} \longrightarrow q^{-1} : \mathbf{Set}(X)$,
so that $\mathbf{Set}/X \simeq \mathbf{Set}(X)$.

This is a nice simplification, but it is misleading:
if we were to develop families indexed by
something less discrete than sets, such an equivalence
may not be possible. ■

For development of multisorted algebraic theories we
will need the category of finitely supported families.

**define** Let $X:\mathbf{Set}$, then
$$\mathbf{Set}_{\mathrm{fin}}(X) \text{ is the full subcategory of } \mathbf{Set}(X)$$
whose objects are families $A_\cdot : X \longrightarrow \mathbf{Set}$ such that
$$\sum_{x:X} A_x \text{ is a finite set.}$$ ■

Note $\mathbf{Set}(1) = \mathbf{Set}$, and
$$\mathbf{Set}_{\mathrm{fin}}(1) = \mathbf{Set}_{\mathrm{fin}}$$ the full subcategory of finite sets.

Given $X : \mathbf{Set}_{\mathrm{fin}}$, then $\mathbf{Set}_{\mathrm{fin}}(X)$ contains all fibrewise finite families.

**define** Given $A_\cdot, B_\cdot : X \longrightarrow \mathbf{Set}$, define the cartesian product
of two families to be fibrewise product
$$A_\cdot \times_X B_\cdot : X \longrightarrow \mathbf{Set}$$
$$x \longrightarrow A_x \times_X B_x$$ ■

We have two projections
$$\pi_1 : \forall x:X.\ A_x \times B_x \longrightarrow A_x$$
$$\langle a, b\rangle \longrightarrow a$$
similarly $\pi_2 : A_\cdot \times_X B_\cdot \longrightarrow B_\cdot$.
It is not hard to see that
$$\forall Z_\cdot : \mathbf{Set}(X),\quad \underline{\underline{f_\cdot : Z_\cdot \longrightarrow A_\cdot \times_X B_\cdot : \mathbf{Set}(X)}}$$
$$\langle f_1 : Z_\cdot \longrightarrow A_\cdot,\ f_2 : Z_\cdot \longrightarrow B_\cdot\rangle : \mathbf{Set}(X) \times \mathbf{Set}(X)$$
so that this fibrewise product implements the universal property
of cartesian product in the category $\mathbf{Set}(X)$.

When instead we view families as objects in the slice category $\mathbf{Set}/X$,
the fibrewise product gives us the pullback:
let $p : E \to X$, $q : F \to X$, then the corresponding map to the family $p^{-1} \times_X q^{-1} : X \to \mathbf{Set}$
is given by $E \times_X F \coloneqq \{\langle a, b\rangle : E \times F \mid p(a) = p(b)\}$
$$p \times_X q : E \times_X F \longrightarrow X$$
$$\langle a, b\rangle \longrightarrow p(a).$$

Diagrammatically we have
$\forall Z:\mathbf{Set},\ f : Z \to E,\ g : Z \to F,$
$\exists!\ f \times_X g : Z \longrightarrow E \times_X F.$

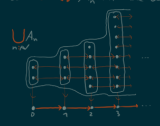
$$\begin{tikzcd}
Z \arrow[dr, "f\times_X g"] \arrow[drr, bend left, "g"] \arrow[ddr, bend right, "f"'] & & \\
& E \times_X F \arrow[r, "\pi_2"] \arrow[d, "\pi_1"] \arrow[dr, "p\circ\pi_1"] & F \arrow[d, "q"] \\
& E \arrow[r, "p"'] & X
\end{tikzcd}$$ ■

## Diagrams, Limits, Colimits

Now we'll consider families whose indexing set has extra structure.
In mathematics one often considers monotonic sequences of objects.
For example in real analysis we make use of non-decreasing sequences
of numbers $a_0 \leq a_1 \leq a_2 \leq \ldots \mathbb{R}$.
This can be specified by a map
$$a_\cdot : \mathbb{N} \longrightarrow \mathbb{R} \text{ s.t.}$$
$$\forall n \leq m.\quad a_n \leq a_m$$
In probability theory we consider non-decreasing sequences of subsets of some set $T$
$$A_0 \subseteq A_1 \subseteq A_2 \subseteq \ldots T$$
Then the sequence can be given by a map
$$A_\cdot : \mathbb{N} \longrightarrow \mathcal{P}[T] \text{ such that}$$
$$\forall n \leq m.\quad A_n \subseteq A_m : T$$
We can see this as a new notion of a family, its indexing object is not just a simple set,
but a poset and the family takes advantage of this structure.

Let $A_\cdot : \mathbb{N} \longrightarrow \mathbf{Set}$, then we can construct the total set $\sum_{n:\mathbb{N}} A_n$.
Now suppose we have injections $A_0 \xrightarrow{i_{0,1}} A_1 \xrightarrow{i_{1,2}} A_2 \longrightarrow \ldots$
With such injections it should be possible to construct a kind of union.
If $A_0 \subseteq A_1 \subseteq \ldots T : \mathbf{Set}$ were subsets, i.e. $i_{n,n+1}$ were corresponding inclusions
then this union could easily be specified by a subset of $T$.
But how to construct such a union without an enclosing set in which $A_0 \subseteq A_1 \subseteq \ldots$?
We have the total set $\sum_{n:\mathbb{N}} A_n$ but in it $A_\cdot$'s are disjoint.
Supposing we had such an external union $\bigcup_{n:\mathbb{N}} A_n : \mathbf{Set}$,
then we would want that atoms for $a : A_n$, $b : A_{n+1}$ s.t. $i_{n,n+1}(a) = b$,
then $a = b : \bigcup_{n:\mathbb{N}} A_n : \mathbf{Set}$

The solution is to identify $A_n$ with the corresponding part of $A_{n+1}$ in $\sum_{n:\mathbb{N}} A_n$.

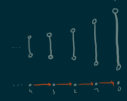



This we can do.

**define**
$$\bigcup_{n:\mathbb{N}} A_n \coloneqq \frac{\sum_{n:\mathbb{N}} A_n}{\{\langle n, a\rangle = \langle m, i_{m,n}(a)\rangle \mid n, m : \mathbb{N},\ a : A_n\}} : \mathbf{Set}$$ ■

Similarly starting with injections $\ldots \longrightarrow A_{-2} \xrightarrow{i_{-2,-1}} A_{-1} \xrightarrow{i_{-1,0}} A_0 : \mathbf{Set}$ –
for the sake of non-triviality imagine the sets to be infinite –
we can build their intersection with the help of the sections

$$\bigcap_{n:\mathbb{N}} A_n = \{\ \}$$



Given $a : \bigcap_{n:\mathbb{N}} A_n$, it should be semialternously in $A_0, A_1, \ldots$
We can represent such an element with $a : \forall n:\mathbb{N}.\ A_n$ s.t.
$$a_0 = i_{-1,0}(a_{-1}),\quad a_1 = i_{-1,1}(a_{-1}),\ \ldots$$

**define**
$$\bigcap_{n:\mathbb{N}} A_n \coloneqq \left\{ a : \forall_{n:\mathbb{N}} A_n \;\middle|\; \begin{array}{l} \forall n:\mathbb{N} \\ a_n = i_{n+1,n}(a_{n+1}) \end{array} \right\} : \mathbf{Set}$$ ■

In all these examples it didn't really matter that the maps were injections or that the order on $\mathbb{N}$ was linear.

Given a sequence of sets and transition maps

$$A_0 \xrightarrow{t_{01}} A_1 \xrightarrow{t_{12}} A_2 \longrightarrow \cdots$$

we get new composite transitions such as

$$A_1 \xrightarrow{t_{12}\circ t_{23}} A_3$$

Now we can generalise set-indexed families to poset-indexed families

**define** $\mathcal{A}$ $P$: Poset indexed family of sets consists of

$$\mathcal{A}: P \longrightarrow \mathbf{Set} \quad \text{a set indexed family}$$

with a selection of transition maps

$$\forall x \leq y : P. \quad \underset{\text{a map}}{\alpha_{x,y}} : A_x \longrightarrow A_y : \mathbf{Set}$$

where on $\forall x : P$

$$\alpha_{x,x} = id_{A_x} : A_x \longrightarrow A_x$$

and $\forall x \leq y \leq z : P$.

$$\begin{tikzcd} A_x \arrow[rr, "\alpha_{x,z}"] \arrow[dr, "\alpha_{x,y}"'] & = & A_z \\ & A_y \arrow[ur, "\alpha_{y,z}"'] & \end{tikzcd}$$

■

Define $[n..] : \mathbf{Poset}$ to be $\{n, n+1, \ldots\}$ considered with the standard linear order

$$[n..] \simeq \underset{n}{\bullet} \longrightarrow \underset{n+1}{\bullet} \longrightarrow \underset{n+2}{\bullet} \longrightarrow \cdots$$

Then a sequence $A_0 \xrightarrow{t_{01}} A_1 \xrightarrow{t_{12}} A_2 \longrightarrow \cdots : \mathbf{Set}$ precisely corresponds to families $\mathcal{A} : [0..] \longrightarrow \mathbf{Set}$.

By $[n..]^{op}$ we mean $[n..]$ with the opposite order

$$[n..] := \cdots \longrightarrow \underset{n+1}{\bullet} \longrightarrow \underset{n}{\bullet} \longrightarrow \underset{n-1}{\bullet}$$

We give further more concrete examples

**example** [Formal Power Series]

We wish to adjoin to $\mathbb{R}$ a number $\varepsilon$ so small that $\varepsilon^2 = 0$, yet not so small as to be zero.

We will construct the resulting number system (an $\mathbb{R}$-algebra) by mimicking the construction of complex numbers where we adjoin number $i$ such that $i^2 = -1$.

**define** The algebra of dual numbers is defined to be

$$\mathbb{D}[\mathbb{R}] := \mathbb{R} \times \mathbb{R};$$

denote the pair $\langle a, b \rangle : \mathbb{R} \times \mathbb{R}$ by $a + b\varepsilon : \mathbb{D}[\mathbb{R}]$.

Define addition by

$$(a + b\varepsilon) + (c + d\varepsilon) := (a+c) + (b+d)\varepsilon;$$

since we wish $\varepsilon^2 = 0$, we would want

$$(a+b\varepsilon)(c+d\varepsilon) = ac + (ac + bd)\varepsilon + bd\,\varepsilon^{\cancel{2}\,0};$$

so we define

$$(a+b\varepsilon)(c+d\varepsilon) := (ac) + (ac+bd)\varepsilon. \quad ■$$

In addition we can embed $\mathbb{R}$ into $\mathbb{D}[\mathbb{R}]$ via $r : \mathbb{R} \longrightarrow r + 0\varepsilon : \mathbb{D}[\mathbb{R}]$.

This makes $\mathbb{D}[\mathbb{R}]$ a commutative $\mathbb{R}$-algebra.

Dual numbers are not a field since $\varepsilon$ is a zero divisor, but they are still fairly nice (they are a local ring)

In fact $a + b\varepsilon : \mathbb{D}[\mathbb{R}]$ is invertible iff $a \neq 0$;

this can be seen from $(a+b\varepsilon)(a-b\varepsilon) = a^2$,

so that $(a+b\varepsilon)\dfrac{1}{a^2}(a-b\varepsilon) = 1$,

so $\dfrac{1}{(a+b\varepsilon)} = \dfrac{1}{a^2}(a-b\varepsilon)$;

this makes sense iff $a^2 \neq 0$ iff $a \neq 0$.

We think of $\mathbb{D}[\mathbb{R}] := \mathbb{D}[\mathbb{R}]$ as the result of adjoining a first order infinitesimal, we can generalise to algebras $\mathbb{D}_n[\mathbb{R}]$ of dimension $n+1$, where we adjoin an $n$th order infinitesimal $\varepsilon$ s.t. $\varepsilon^{n+1} = 0$.

For example $\mathbb{D}_2[\mathbb{R}] := \left\{ a + b\varepsilon + c\varepsilon^2 \,\middle|\, a, b, c : \mathbb{R} \right\}$,

where $(a_1 + b_1\varepsilon + c_1\varepsilon^2)(a_2 + b_2\varepsilon + c_2\varepsilon^2) = a_1a_2 + (a_1b_2 + b_1a_2)\varepsilon + (a_1c_2 + b_1b_2 + c_1a_2)\varepsilon^2 : \mathbb{D}_2[\mathbb{R}]$

since $\varepsilon^3 = 0 : \mathbb{D}_2[\mathbb{R}]$.

Set $\mathbb{D}_1[\mathbb{R}] := \mathbb{R}$. Given any such higher-order number $3 - 2\varepsilon + \frac{7}{2}\varepsilon^2 + 2\varepsilon^3 : \mathbb{D}_3[\mathbb{R}]$, we can truncate it to $3 - 2\varepsilon + \frac{7}{2}\varepsilon^2 : \mathbb{D}_2[\mathbb{R}]$, $3 - 2\varepsilon : \mathbb{D}_1[\mathbb{R}]$, $3 : \mathbb{D}_0[\mathbb{R}]$.

But $3 - 2\varepsilon + \frac{7}{2}\varepsilon^2 + 2\varepsilon^3 : \mathbb{D}_3[\mathbb{R}]$ is a truncation of many even higher-order numbers.

Note that such truncations are algebra homomorphisms.

Is there an algebra $\mathbb{D}_\omega[\mathbb{R}]$ whose numbers can be truncated to numbers in $\mathbb{D}_n[\mathbb{R}]$ for any $n$, a kind of power series $a_0 + a_1\varepsilon + a_2\varepsilon^2 + \ldots$ ?

We can implement it by applying the intersection construction from the previous example onto the family of algebras with truncations

$$\begin{tikzcd} \cdots \arrow[r] & \mathbb{D}_2[\mathbb{R}] \arrow[r, "t"] \arrow[d] & \mathbb{D}_1[\mathbb{R}] \arrow[r, "t"] \arrow[d] & \mathbb{R} \arrow[d] \\ \cdots \arrow[r] & 2 \arrow[r] & 1 \arrow[r] & 0 \end{tikzcd}$$

A section $z : \forall_{n:[0..]^{op}} \mathbb{D}_n[\mathbb{R}]$ is simply a sequence $\langle \ldots, c_0 + c_1\varepsilon + c_2\varepsilon^2 : \mathbb{D}_2[\mathbb{R}], b_0 + b_1\varepsilon : \mathbb{D}_1[\mathbb{R}], a_0 : \mathbb{R} \rangle$

we wish to get a convenient number so that $b_0 + b_1\varepsilon : \mathbb{D}_1[\mathbb{R}]$ truncates to $a_0$, and so on

**define** $\mathbb{D}_\omega[\mathbb{R}] := \left\{ z : \forall_{n:[0..]^{op}} \mathbb{D}_n[\mathbb{R}] \,\middle|\, \forall n. \sim t(z_{n+1}) = z_n \right\}$ ■

Hence to define $z : \mathbb{D}_\omega[\mathbb{R}]$ is to define its truncations to all $\mathbb{D}_n[\mathbb{R}]$.

We can make it into an $\mathbb{R}$-algebra via

**define** $\forall z, w : \mathbb{D}_\omega[\mathbb{R}]$,

$$(z + w)_n := z_n + w_n : \mathbb{D}_n[\mathbb{R}],$$

$$(z \cdot w)_n := z_n \cdot w_n : \mathbb{D}_n[\mathbb{R}];$$

Then $z + w : \mathbb{D}_\omega[\mathbb{R}]$ since truncations are algebra homomorphism, and so preserve addition,

$$\begin{aligned} t((z+w)_{n+1}) &= t(z_{n+1} + w_{n+1}) \\ &= t(z_{n+1}) + t(w_{n+1}) \\ &= z_n + w_n \\ &= (z+w)_n. \end{aligned}$$

Similarly for multiplication ■

**define** To each $r : \mathbb{R}$, there corresponds $\langle \ldots, r + 0\cdot\varepsilon + 0\cdot\varepsilon^2, r + 0\cdot\varepsilon, r \rangle : \mathbb{D}_\omega[\mathbb{R}]$.

This makes $\mathbb{R}$ a subring of $\mathbb{D}_\omega[\mathbb{R}]$ ■

Beware that $\mathbb{D}[\mathbb{R}]$ is not a subring of $\mathbb{D}_\omega[\mathbb{R}]$ – $\mathbb{D}[\mathbb{R}]$ contains nilpotent elements, but $\mathbb{D}_\omega[\mathbb{R}]$ is actually an integral domain.

Algebra $\mathbb{D}_\omega[\mathbb{R}]$ is precisely the algebra of formal power series $\mathbb{R}[[x]]$:

$$a_0 + a_1x + a_2x^2 + \ldots : \mathbb{R}[[x]] \cong \langle \ldots, a_0 + a_1\varepsilon + a_2\varepsilon^2, a_0 + a_1\varepsilon, a_0 \rangle : \mathbb{D}_\omega[\varepsilon].$$

See for yourself that the pointwise multiplication on $\mathbb{D}_\omega[\mathbb{R}]$ really corresponds to multiplication of formal power series. ■

**example** [2-adic numbers]

Let $\mathbb{Z}_{2^n}$ be the ring of integers modulo $2^n$.

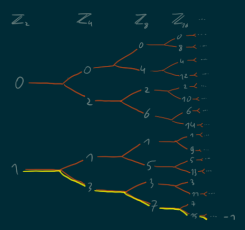
Given any number in $\mathbb{Z}_{2^{n+1}}$ we can truncate it to one in $\mathbb{Z}_{2^n}$.

for example take $30 : \mathbb{Z}_{32}$, then
$$30 = 14 : \mathbb{Z}_{16}$$
$$14 = 6 : \mathbb{Z}_8$$
$$6 = 2 : \mathbb{Z}_4$$
$$2 = 0 : \mathbb{Z}_2.$$

The number $30 : \mathbb{Z}_{32}$ is a truncation of two different numbers $30, 62 : \mathbb{Z}_{64}$.

We say that $30 : \mathbb{Z}_{32}$ lifts to $30, 62 : \mathbb{Z}_{64}$. Any $a : \mathbb{Z}_{2^n}$ lifts to only $a, a+2^n : \mathbb{Z}_{2^{n+1}}$.

We can see all possible lifts starting with $0, 1 : \mathbb{Z}_2$ as an infinite binary tree




In a certain sense the infinite paths one can take starting from 0 or 1 are a new type of numbers — the 2-adic integers; note that there are uncountable number of such paths.

we can construct them by taking a limit
$$\mathbb{Z}_{2^\infty} := \lim\left[\cdots \to \mathbb{Z}_8 \to \mathbb{Z}_4 \to \mathbb{Z}_2\right] : \mathbf{Set}.$$

Since any $z : \mathbb{Z}_{2^\infty}$ is just a sequence $\langle \ldots, z_3, z_2, z_1 \rangle$ such that $z_1 = z_1 : \mathbb{Z}_2$, $z_3 = z_1 : \mathbb{Z}_4$, $z_2 = z_2 : \mathbb{Z}_2$,

we can define componentwise ring operations.

This ring includes $\mathbb{Z}$ as a subring via
$$m : \mathbb{Z} \to \langle \ldots, m : \mathbb{Z}_8, m : \mathbb{Z}_4, m : \mathbb{Z}_2 \rangle,$$
for example
$$3 : \mathbb{Z} \to \langle \ldots, 3, 3, 3 \rangle$$
$$-1 : \mathbb{Z} \to \langle \ldots, 15, 7, 3, 1 \rangle$$
$$-16 : \mathbb{Z} \to \langle \ldots, 64-16, 32-16, 0, 0, 0, 0 \rangle$$

It contains some inverses to integers, like $3 : \mathbb{Z}$.

Note that $3 : \mathbb{Z}_{2^n}$ always has an inverse since 3 is coprime to $2^n$.

Let $3 \cdot a_1 = 1 : \mathbb{Z}_2$, $3 \cdot a_2 = 1 : \mathbb{Z}_4$, $3 \cdot a_3 = 1 : \mathbb{Z}_8$, ...

Then $a_n$ is uniquely determined inverse of $3 : \mathbb{Z}_{2^n}$.

Note that $a_{n+1} = a_n : \mathbb{Z}_{2^n}$, since reduction is a homomorphism, so it preserves inverses.

Since $a_2 : \mathbb{Z}_4$ is a lift of $a_1 = 1 : \mathbb{Z}_2$ it must be $1 : \mathbb{Z}_4$ or $1+2=3 : \mathbb{Z}_4$. If it $a_2 = 3 : \mathbb{Z}_4$

Since $a_3 : \mathbb{Z}_8$ is a lift of $a_2 = 3 : \mathbb{Z}_4$ it must be $3 : \mathbb{Z}_8$ or $3+4=7 : \mathbb{Z}_8$. Its $a_3 = 3 : \mathbb{Z}_8$

In this way we can easily compute a bunch of lifts even by hand

$$1/3 : \mathbb{Q} \simeq 1 : \mathbb{Z}_2,\ 3 : \mathbb{Z}_4,\ 3 : \mathbb{Z}_8,\ 3+8=11 : \mathbb{Z}_{16},\ 11 : \mathbb{Z}_{32},\ 11+32=43 : \mathbb{Z}_{64},\ 43 : \mathbb{Z}_{128},\ 43+128=171 : \mathbb{Z}_{256}$$

But we won't find $1/2$ in $\mathbb{Z}_{2^\infty}$, so its not a field, but its still nice (integral domain)

Since $\mathbb{Z}_2$ is a field, we still can construct analogues of dual number $D_1[\mathbb{Z}_2]$ and their higher order versions $D_n[\mathbb{Z}_2]$.

For example, $D_2[\mathbb{Z}_2] := \{0, 1, \varepsilon, 1+\varepsilon, \varepsilon^2, 1+\varepsilon^2, \varepsilon+\varepsilon^2, 1+\varepsilon+\varepsilon^2\}$.

Note that $D_n[\mathbb{Z}_2]$ has $2^{n+1}$ elements, but it is not so to $\mathbb{Z}_{2^{n+1}}$, as a ring — $D_n[\mathbb{Z}_2]$ has characteristic 2, while $\mathrm{char}\ \mathbb{Z}_{2^{n+1}} = 2^{n+1}$.

The limit $\lim\left[\cdots \to D_2[\mathbb{Z}_2] \to D_1[\mathbb{Z}_2] \to \mathbb{Z}_2\right]$ is again the ring of formal power series $\mathbb{Z}_2[x]$

The 2-adic integers have characteristic 0, while $\mathrm{char}\ \mathbb{Z}_2[x] = 2$ ■

Now we turn to colimits.

**example** [Prüfer Group]

We considered the family of rings & ring homomorphisms
$$\cdots \to \mathbb{Z}_8 \to \mathbb{Z}_4 \to \mathbb{Z}_2.$$

Now will denote the underlying Abelian Group of $\mathbb{Z}_{2^n}$ by $\mathbb{Z}/_{2^n}\mathbb{Z}$.

Then we get a family of Abelian Groups & group homomorphisms in the opposite direction
$$\mathbb{Z}/_{2\mathbb{Z}} \hookrightarrow \mathbb{Z}/_{4\mathbb{Z}} \hookrightarrow \mathbb{Z}/_{8\mathbb{Z}} \hookrightarrow \cdots$$

The colimit of this family is the quotient of the total set
$$\mathbb{Z}/_{2^\infty}\mathbb{Z} := \frac{\sum_{n:[1..]} \mathbb{Z}/_{2^n}\mathbb{Z}}{\{\langle n, a\rangle = \langle m, i(a)\rangle \mid n, m, a : \mathbb{Z}/_{2^n}\mathbb{Z}\}}$$

This has the structure of an Abelian group — the 2-Prüfer group.

We'll illustrate how to define addition on it.

First we define a binary operation $+$ on the total set of the family:
$$+ : \sum_{n:[1..]} \mathbb{Z}/_{2^n}\mathbb{Z} \times \sum_{n:[1..]} \mathbb{Z}/_{2^n}\mathbb{Z} \to \sum_{n:[1..]} \mathbb{Z}/_{2^n}\mathbb{Z}$$
$$\forall n, m : [1..].$$
case $n \leq m$.
$$\langle n, a\rangle + \langle m, b\rangle := \langle m, i^{m-n}(a) + b\rangle$$
case $n \geq m$.
$$\langle n, a\rangle + \langle m, b\rangle := \langle n, a + i^{n-m}(b)\rangle.$$

To get $+ : \mathbb{Z}/_{2^n}\mathbb{Z} \times \mathbb{Z}/_{2^n}\mathbb{Z} \to \mathbb{Z}/_{2^n}\mathbb{Z}$,

we need to check the preserving relations.

Let $n \leq m$, $a : \mathbb{Z}/_{2^n}\mathbb{Z}$, $b : \mathbb{Z}/_{2^m}\mathbb{Z}$. Then
$$\langle n, a\rangle = \langle n+1, i(a)\rangle : \mathbb{Z}/_{2^n}\mathbb{Z}, \quad \langle m, b\rangle = \langle m+1, i(b)\rangle : \mathbb{Z}/_{2^n}\mathbb{Z}.$$

We check that $\langle n, a\rangle + \langle m, b\rangle = \langle n+1, i(a)\rangle + \langle m+1, i(b)\rangle : \mathbb{Z}/_{2^n}\mathbb{Z}$;

note $\langle n, a\rangle + \langle m, b\rangle = \langle m, i^{m-n}(a) + b\rangle$
$$= \langle m+1, i(i^{m-n}(a) + b)\rangle$$
$$= \langle m+1, i^{m-n+1}(a) + i(b)\rangle$$
$$= \langle m+1, i^{m-n}(i(a)) + i(b)\rangle$$
$$= \langle m+1, i(a)\rangle + \langle m+1, i(b)\rangle.$$

Symmetrically for $n \geq m$.

Now this may look complicated at first sight, you can check your intuition against a much simpler internal definition — yet the point is to understand the technology of external colimits, not this particular group.

**define** The cyclic group of $n$-elements can be realised as the subgroup of $\mathbb{C}^\times$ the group of non-zero complex numbers
$$C_n := \{z : \mathbb{C} \mid z^n = 1\}.$$

Define the 2-Prüfer group as
$$C_{2^\infty} := \bigcup_{n:[1..\infty)} C_{2^n} = \left\{z : \mathbb{C} \;\middle|\; \exists^{n:[1..\infty)}_{z^{2^n} = 1}\right\}$$ ■

Denote by $e_{2^n} : C_{2^n}$ the generator of $C_{2^n}$ that is closest on the upper half-circle to 1.

Then we can visualize $C_{2^\infty}$ as a countable, but dense subset of the circle.

■

**example** Another use of colimits is to build polynomials in infinite number of variables.

More precisely, given $X : \mathbf{Set}$

we can construct a ring of polynomials $\mathbb{R}[X]$,

s.t. any $f : \mathbb{R}[X]$ can be thought of as being

in $f : \mathbb{R}[\{x_1, \ldots, x_n\}]$ for some finite subset $\{x_1, \ldots, x_n\} \subseteq X$.

Simply note that given $S \subseteq T \subseteq X$ finite subsets

we have a natural inclusion $\mathbb{R}[S] \hookrightarrow \mathbb{R}[T]$.

This gives us a family over a poset $\mathcal{P}_{fin}[X]$ of finite subsets of $X$.

Then we can set
$$\mathbb{R}[X] := \mathrm{colim}_{S : \mathcal{P}_{fin}[X]} \mathbb{R}[S].$$

The reader may try to workout the details of the ring structure on $\mathbb{R}[X]$.

This example will be very important to our exploration of algebraic theories. ■

So far we've considered families over posets we can go further.

Note that a family $A : 1 \to \mathbf{Set}$ is just one set $A$.

We can think of $1$ as a point

**example** Let $A : \mathbf{Set}$, think of $A$ as a family $A : 1 \to \mathbf{Set}$, $A_* := A$.

Then both its total set $\sum_{x:1} A_x$

and the set of sections $\bigvee_{x:1} A_x$ are $A$.

To the family $A$ we attach a transition map $f : A \to A$.

Then we still can perform the limit and colimit:

the limit is a subset of the set of sections
$$\{x : A \mid f(x) = x\};$$

the colimit is a quotient of the total set
$$\frac{A}{\{f(x) = x \mid x : A\}}$$ ■

**example** Let $M$ be a monoid, then an $M$-set consists of a set $A$ with $\star : M \times A \to A$

where $m \star (m' \star a) = (m m') \star a$

$1 \star a = a$.

We call $\star$ an action of $M$ on $A$.

Then again we can think of $A$ as a family indexed by $1$;

The monoid action gives us for each $m : M$ a transition map
$$A \to A$$
$$a \mapsto m \star a$$

Then the limit of the family with these transition maps is
$$\{a : A \mid \forall^{m:M}_{m \star a = a}\},$$
and the colimit
$$\frac{A}{\{m \star a = a \mid m : M, a : A\}}$$

When $M := G$ is a group, this limit is usually called the invariants of the $G$-set $A$ and denoted by $A^G$;

the colimit is the quotient of $A$ by $G$ denoted $A/G$. ■

**example** We can consider a more sophisticated example.

Let $A : \mathbf{Set}$, then consider $\mathbb{P}_{fin}[A]$ the set of finite subsets.

Let $\forall n : \mathbb{N}$ $A^{\underline{n}}$ denote the set of all permutations of size $n$ of elements of $A$. Formally $A^{\underline{n}} := [n] \hookrightarrow A$.

**example** $\{a, b, c\}^{\underline{2}} = \left\{ \begin{array}{l} \langle a, b\rangle, \langle b, a\rangle, \\ \langle a, c\rangle, \langle c, a\rangle, \\ \langle b, c\rangle, \langle c, b\rangle \end{array} \right\}$ ■

Now $A^{\underline{0}} = \{\langle\rangle\}$, $A^{\underline{1}} = A$.

Note, assuming $m := \#A$,

$$\#\left[A^{\underline{n}}\right] = m^{\underline{n}} := \begin{cases} m \cdot (m-1) \cdot \ldots \cdot (m-(n-1)) & \text{if } n \leq m \\ 0 & \text{otherwise.} \end{cases}$$

Let $[n]!$ denote the group of permutations on $n$-letters.

Then $[n]!$ acts on $A^{\underline{n}}$ on the right.

Consider the family

$$\begin{tikzcd} \mathbb{N} \arrow[r, rightsquigarrow] & \mathbf{Set} \\ n \arrow[r, rightsquigarrow] & A^{\underline{n}} \end{tikzcd}$$

with transition maps

$$\begin{tikzcd} \forall n \mathbb{N} \ni \sigma : [n]! : \\ A^{\underline{n}} \arrow[r] & A^{\underline{n}} \\ p \arrow[r, rightsquigarrow] & p \circ \sigma \end{tikzcd}$$

Then the colimit of this family is

$$\mathbb{P}_{fin}[A] \simeq \{\langle\rangle\} + A + \frac{A^{\underline{2}}}{[2]!} + \frac{A^{\underline{3}}}{[3]!} + \ldots \quad ■$$

$$\begin{array}{ccccc} A^{\underline{0}} & A^{\underline{1}} & A^{\underline{2}} & A^{\underline{3}} & A^{\underline{4}} \ldots \\ | & | & | & | & | \\ \circ & \circ & \circ\circ & \circ\circ\circ & \circ\circ\circ\circ \\ \circ & \circ & \circ\circ & \circ\circ\circ & \circ\circ\circ\circ \\ \circ & \circ & \circ & \circ & \circ \ldots \\ & & \circ & \circ & \circ \\ & & & \circ & \circ \\ & & & & \circ \end{array}$$

**example** A pair of sets $A, B$ can be thought of as $[2]$-indexed diagram.

Then the set of sections is $A \times B$, the limit set is $A + B$.

(1) Let us consider a single transition arrow $A \xrightarrow{f} B$.

Then the limit is

$$\left\{ \langle a, b\rangle : A \times B \;\middle|\; f(a) = b \right\} \simeq A \quad \text{the graph of } f.$$

The colimit is

$$\frac{A + B}{\left\{ a = f(a) \;\middle|\; a : A \right\}} \simeq B$$

(2) Now suppose we have two transition arrows $A \overset{f}{\underset{g}{\rightrightarrows}} B$

The limit is

$$\left\{ \langle a, b\rangle : A \times B \;\middle|\; f(a) = b,\ g(a) = b \right\} \simeq \left\{ a : A \;\middle|\; f(a) = f(b) \right\}$$

the equalizer set of $f$ & $g$.

The colimit is

$$\frac{A + B}{\left\{ a = f(a) \,|\, a : A \right\} \cup \left\{ a = g(a) \,|\, a : A \right\}} \simeq \frac{B}{\left\{ b = b' \;\middle|\; \begin{array}{l} \exists a : A \\ f(a) = b,\ g(a) = b' \end{array} \right\}}$$

the coequalizer of $f$ & $g$. ■

We can now unify all our examples

**define** Let $D$ be a small category.

Then an $D$-indexed diagram of sets is a functor

$$A : D \longrightarrow \mathbf{Set} \quad ■$$

**define** Let $A, B : D \longrightarrow \mathbf{Set}$ be diagrams,

a diagram morphism from $A$ to $B$,

consists of $f : \forall x : D.\ A_x \longrightarrow B_x : \mathbf{Set}$

where $\forall \alpha : x \to y : D,\ a : A_x.$

$$f(\langle x, a\rangle) = \alpha(f(a)).$$

This is simply a natural transformation

$$f : A \Rightarrow B : D \to \mathbf{Set}. \quad ■$$

**define** This gives us in the obvious way the category $\mathbf{Set}(D)$ of $D$-diagrams of sets ■

Given a diagram $A : D \longrightarrow \mathbf{Set}$, since $D$ is small category, we can talk about its underlying set, so the total set $\sum_{x:D} A_x$ and the set of sections $\forall_{x:D} A_x$ still make sense.

**define** For $A : D \longrightarrow \mathbf{Set}$ define

(1) the limit

$$\lim_{x:D} A_x := \left\{ s : \forall_{x:D} A_x \;\middle|\; \begin{array}{l} \forall \alpha : x \to y : D \\ \alpha(s) = s_y : A_y \end{array} \right\}$$

(2) the colimit

$$\operatorname{colim}_{x:D} A_x := \frac{\sum_{x:D} A_x}{\left\{ \langle x, a\rangle = \langle y, \alpha(a)\rangle \;\middle|\; \alpha : x \to y : D,\ a : A_x \right\}} \quad ■$$

**define** We have the functor

$$\begin{array}{l} \lim_D : \mathbf{Set}(D) \longrightarrow \mathbf{Set} \\ A \rightsquigarrow \lim_{x:D} A_x, \end{array}$$

where for $f : A \to B : \mathbf{Set}(D)$ we restrict $f : \forall_{x:D} A_x \longrightarrow \forall_{x:D} B_x$ to

$$f : \lim_{x:D} A_x \longrightarrow \lim_{x:D} B_x;$$

we check $\forall s : \lim_{x:D} A_x,\ \alpha : x \to y : D$

$$\begin{aligned} \alpha(f(s)_x) &= \alpha(f(s_x)) \\ &= f(\alpha(s_x)) \\ &= f(s_y) \\ &= f(s)_y \quad ■ \end{aligned}$$

**define** We have the functor

$$\begin{array}{l} \operatorname{colim}_D : \mathbf{Set}(D) \longrightarrow \mathbf{Set} \\ A \rightsquigarrow \operatorname{colim}_{x:D} A_x, \end{array}$$

where for $f : A \to B : \mathbf{Set}(D)$ we restrict $f : \sum_{x:D} A_x \longrightarrow \sum_{x:D} B_x$ to the quotient

$$f : \operatorname{colim}_{x:D} A_x \longrightarrow \operatorname{colim}_{x:D} B_x;$$

we check for $\alpha : x \to y : X,\ a : A_x$, that

since $\langle x, a\rangle = \langle y, \alpha(a)\rangle : \operatorname{colim}_{x:D} A_x$,

we must have

$$\begin{aligned} f(\langle x, a\rangle) &= \langle x, f(a)\rangle : \operatorname{colim}_{x:D} B_x \\ &= \langle y, \alpha(f(a))\rangle \\ &= \langle y, f(\alpha(a))\rangle \\ &= f(\langle y, \alpha(a)\rangle). \quad ■ \end{aligned}$$

**define** For $A : \mathbf{Set}$, we have the constant diagram

$$\begin{array}{l} \operatorname{const} A : D \longrightarrow \mathbf{Set} \\ \operatorname{const} A_x := A \\ \forall \alpha : x \to y : D,\ a : A. \\ \qquad \alpha(a) = a \end{array}$$

This gives us a functor

$$\operatorname{const} : \mathbf{Set} \longrightarrow \mathbf{Set}(D) \quad ■$$

**prop** Let $A : D \longrightarrow \mathbf{Set}$ a diagram, $Z : \mathbf{Set}$.

Then

$$\begin{array}{ll} & f : Z \longrightarrow \lim_{x:D} A_x : \mathbf{Set} \\ (1) & \overline{\overline{\quad f : \operatorname{const}(Z) \longrightarrow A : \mathbf{Set}(D) \quad}} \\ & f : \operatorname{colim}_{x:D} A_x \longrightarrow Z : \mathbf{Set} \\ (2) & \overline{\overline{\quad f : A \longrightarrow \operatorname{const}(Z) : \mathbf{Set}(D) \quad}} \end{array}$$

That is

$$\begin{tikzcd} \mathbf{Set} \arrow[r, shift left, "\lim_D"] \arrow[r, shift right, "\operatorname{const}", "\perp"'] & \mathbf{Set}(D) \arrow[l, shift left, "\operatorname{colim}_D"] \arrow[l, shift right, "\operatorname{const}"'] \end{tikzcd} \quad ■$$

# The case of monoids

To see how an algebraic theory should be defined, first we'll consider what we mean by an operation of an algebraic structure of a given theory.

To do this, we'll examine monoids – but you can replace them with your favourite algebraic structure.

**define** A monoid consists of

$$M : \mathbf{Set} \text{ with } \underset{1 : M}{\star : M \times M \to M} : \mathbf{Set}$$

where $\forall a, b, c : M.$

$$(a \star b) \star c = a \star (b \star c) \quad \text{(associativity)}$$

$\forall a : M.$

$$a \star 1 = a = 1 \star a. \quad \text{(neutrality)} \blacksquare$$

**define** Let $M := M_{\star_M, 1_M}$ & $N := N_{\star_N, 1_N}$ be monoids. Then a morphism from $M$ to $N$ of monoids consists of

$$f : M \longrightarrow N : \mathbf{Set}$$

where (i) $f(1) = 1$,

(ii) $\forall a, b : M.$

$$f(a \star b) = f(a) \star f(b) \blacksquare$$

This defines in the obvious way the category of monoids denoted

$$\mathbf{Monoid}.$$

## & What is a Monoid operation?

Informally we know that $\star$, the Monoid multiplication, is a Monoid operation.

Given $M : \mathbf{Monoid}$, we have

$$\underset{M}{\star} : M \times M \longrightarrow M : \mathbf{Set}$$

Note that in general $\underset{M}{\star}$ is not a morphism of monoids $[M \times M$ has monoid product structure$]$, so it doesn't live in the category $\mathbf{Monoid}$.

Its just a set function.

But given another $N : \mathbf{Monoid}$, we have a different function

$$\underset{N}{\star} : N \times N \longrightarrow N : \mathbf{Set},$$

and even though we know that

$\underset{M}{\star}$ & $\underset{N}{\star}$ are completely different functions,

they are somehow a part of the same operations $\star$.

Intuitively $\star$ multiplication is defined uniformly for all monoids.

By this we mean that first of all the type of multiplication $\star$

$$\star : \forall M : \mathbf{Monoid}.\ M \times M \longrightarrow M : \mathbf{Set},$$

and for related monoids, the corresponding components of $\star$ are also related, formally, given

$$M \xrightarrow{f} N : \mathbf{Monoid}$$

we have

$$\begin{tikzcd} M \times M \arrow[r, "f \times f"] \arrow[d, "\underset{M}{\star}"'] \arrow[dr, phantom, "\circ"] & N \times N : \mathbf{Set} \arrow[d, "\underset{N}{\star}"] \\ M \arrow[r, "f"'] & N \end{tikzcd}$$

This condition precisely states that $f$ preserves multiplication

$$f(a \star b) = f(a) \star f(b)$$

In this way $\star$ becomes a global uniform operation on all monoids.

Beside multiplication we can construct derived monoid operations.

**example** Let $\omega : \forall M : \mathbf{Monoid}.$

$$M^{\times 4} \longrightarrow M$$

$$\langle x, y, z, w \rangle \rightsquigarrow w \star (z \star (y \star (x \star x))).$$

Then $\forall f : M \to N$ $\mathbf{Monoid}$.

$$\begin{tikzcd} M^{\times 4} \arrow[r, "f^{\times 4}"] \arrow[d, "\omega_M"'] \arrow[dr, phantom, "\circ"] & N^{\times 4} : \mathbf{Set} \arrow[d, "\omega_N"] \\ M \arrow[r] & N \end{tikzcd}$$

i.e.

$$\langle x, y, z, w \rangle : M^{\times 4}$$

$$w \star (z \star (y \star (x \star x))) \qquad \langle f(x), f(y), f(z), f(w) \rangle : N^{\times 4}$$

$$f(w \star (z \star (y \star (x \star x))) = f(w) \star (f(z) \star (f(y) \star (f(x) \star f(x))).$$

We say that $\omega$ has arity 4. $\blacksquare$

We can also see $1$ as such an operation.

**example** $1 : \forall M : \mathbf{Monoid}.\ M$,

i.e. $1 : \forall M : \mathbf{Monoid}.$

$$M^0 \to M.$$

So we say that $1$ has arity 0. $\blacksquare$

On the other hand, for the particular monoid of real numbers

$$\mathbb{R}\ [\text{considered with addition \& } 0]$$

we have a function

$$\cos : \mathbb{R} \longrightarrow \mathbb{R},$$

but it can't be considered as a part of a global Monoid operation, since $\cos$ takes advantage of extra structure of $\mathbb{R}$ that's not common to all monoids.

We can generalize to $n : \mathbb{N}$-ary operations

$$\omega : \forall M : \mathbf{Monoid}.\ M^{\times n} \to M.$$

We can generalize further.

**example** Let $\eta : \forall M : \mathbf{Monoid}.$

$$M^{\times 3} \longrightarrow M^{\times 4}$$

$$\langle x, y, z \rangle \rightsquigarrow \langle x \star y,\ x \star z,\ y \star z,\ 1 \rangle$$

This satisfies

$\forall f : M \to N$ $\mathbf{Monoid}$.

$$\begin{tikzcd} M^{\times 3} \arrow[r, "f^{\times 3}"] \arrow[d, "\eta_M"'] \arrow[dr, phantom, "\circ"] & N^{\times 3} : \mathbf{Set} \arrow[d, "\eta_N"] \\ M^{\times 4} \arrow[r, "f^{\times 4}"'] & N^{\times 4} \end{tikzcd}$$

We say that $\eta$ has arity $\bullet^3 \to \bullet^4$.

Of course $\eta$ is determined by 4 operations of arity $\bullet^3 \to \bullet^1$, i.e. 3-ary operations, so we don't really get anything new.

But it will be convenient to have such operations at our disposal. $\blacksquare$

We make one last generalization. So far our arities used natural numbers $n : \mathbb{N}$.

**define** $\forall n : \mathbb{N}$

$$[n] := \{0, 1, \ldots, n-1\} \blacksquare$$

Note, given $X : \mathbf{Set}$, we have a natural iso

$$X^n \cong X^{[n]} : \mathbf{Set}$$

with

$$\langle x_0, \ldots, x_n \rangle \rightsquigarrow \lambda k.\, x_k$$

$$\langle X(0), \ldots, X(n-1) \rangle \leftarrow\!\shortmid X$$

i.e. $n$-tuples can be thought of as functions in $[[n] \to X] = X^{[n]}$

Our final notion of arity is arbitrary finite set;

Let $\mathbf{Set}_{fin}$ denote the full-subcategory of $\mathbf{Set}$ of finite sets.

**define** Let $a, b : \mathbf{Set}_{fin}$. Then a monoid operation of arity $\bullet^a \to \bullet^b$ is a map. It is

$$\omega : \forall M : \mathbf{Monoid}.\ M^a \to M^b : \mathbf{Set}$$

where $\forall M \xrightarrow{f} N$ $\mathbf{Monoid}$.

$$\begin{tikzcd} M^a \arrow[r, "f^a"] \arrow[d, "\omega_M"'] \arrow[dr, phantom, "\circ"] & N^a \arrow[d, "\omega_N"] \\ M^b \arrow[r, "f^b"'] & N^b \end{tikzcd} \blacksquare$$

Note that by $f^a : M^a \to N^a$ we mean

$$f^a(m) := \lambda k.\, f(m(k))$$
$$= f \circ m$$

So such an operation is precisely a natural transformation

$$\omega : \|\cdot\|^a \Rightarrow \|\cdot\|^b$$

where $\|\cdot\| : \mathbf{Mon} \to \mathbf{Set}$ is the forgetful functor.

In more linear notation given $m : M^a$, we'll write

$$f(m) := f^a(m) = f \circ m : M^b$$

$$\omega(m) := \omega_M\ m : M,$$

hence the condition becomes

$$\forall m : M^a.$$

$$f(\omega(m)) = \omega(f(m))$$

**example** Let $a := \{a, b, c\}$, $b := \{x, y\} : \mathbf{Set}_{fin}$, then we have

$$\tau : \forall M : \mathbf{Monoid}.$$

$$M^a \to M^b$$

$$m \rightsquigarrow \lambda k.\, b.$$

case $k$ is

$$x \rightsquigarrow m(a) \star m(b)$$

$$y \rightsquigarrow m(b) \star m(c)$$

The operation $\tau$ of arity $a \to b$ is basically the same operation as $\tau$ of arity $3 \to 2$

$$\langle a, b, c \rangle \rightsquigarrow \langle \underset{x}{a \star b},\ \underset{y}{b \star c} \rangle$$

so we don't really get new operations. Usefulness of this will become more apparent when we get to monads;

such a view of arities is more amenable to further, more radical generalization later on.

Now we have a formal definition of a monoid operation; from the basic operations that we have started with

$$\star \text{ of arity } \overset{2}{\bullet} \to \overset{1}{\bullet},$$

$$1 \qquad \overset{0}{\bullet} \to \overset{1}{\bullet}$$

we can construct many more derived operations [such as $\omega$ & $\eta$].

We need to ask the following questions:

(1) What is formally meant by a derived operations?

(2) Is every monoid operation given by the previous definition a derived operation?

(3) If we fix $a, b : \mathbf{Set}_{fin}$, does the collection of all operations of arity $a \to b$ form a small set?

These questions are easily answered with basic category theory & the construction of a free monoid. This we will do later.

We define our first algebraic theory mon.

**define** Let $a, b : \mathbf{Set}_{fin}$; define

$$\bullet^a \longrightarrow \bullet^b : \mathrm{mon} := \left\{ \omega : \bullet^a \xrightarrow{\text{Monoid}} \bullet^b \;\middle|\; \omega \text{ is a monoid operation} \right\}$$ ■

Later we'll see that $\bullet^a \longrightarrow \bullet^b : \mathrm{mon}$ is a set.

Now we'll abstract out relevant features of the structure of all operations.

(composition) Given $\omega : \bullet^a \longrightarrow \bullet^b : \mathrm{mon}$, $\eta : \bullet^b \longrightarrow \bullet^c : \mathrm{mon}$, we have obvious composition

$$\eta\omega : \bullet^a \longrightarrow \bullet^c : \mathrm{mon}$$

since $\forall M \xrightarrow{f} N : \mathrm{mon}$

$$\begin{tikzcd} M^a \arrow[r, "f^a"] \arrow[d, "\omega_M"'] & N^a \arrow[d, "\omega_N"] \\ M^b \arrow[r, "f^b"] \arrow[d, "\eta_M"'] & N^b \arrow[d, "\eta_N"] \\ M^c \arrow[r, "f^c"] & N^c \end{tikzcd}$$

i.e. $f(\eta\omega(m)) = f(\eta(\omega(m)))$
$= \eta(f(\omega(m)))$
$= \eta(\omega(f(m)))$
$= \eta\omega(f(m))$

(trivial operation) We have the trivial operations such as

$$\forall M. \quad M^1 \longrightarrow M$$
$$\langle x, y, z \rangle \rightsquigarrow x$$
$$\langle x, y, z \rangle \rightsquigarrow y$$
$$\langle x, y, z \rangle \rightsquigarrow z$$

or $\forall M.$

$$M^{\times 2} \longrightarrow M^{\times 3}$$
$$\langle x, y \rangle \longrightarrow \langle y, x, x \rangle$$

These are available in all algebraic structures, not just monoids.

Such trivial operations correspond to finite set functions.
Consider the last example; it corresponds to

$$[2] \longleftarrow [3]$$
$$1 \leftsquigarrow 0$$
$$0 \leftsquigarrow 1$$
$$0 \leftsquigarrow 2.$$

Note the contravariance.
In general given $a \xleftarrow{\alpha} b : \mathbf{Set}_{fin}$, we get

$$(-)^*\alpha : \forall M. \quad M^a \longrightarrow M^b$$
$$m \rightsquigarrow m\alpha$$

Then ${}^*\alpha : (-)^*\alpha : \bullet^a \longrightarrow \bullet^b : \mathrm{mon}$ is a monoid operation since

$$f(m)^*\alpha = fm\alpha$$
$$= (fm)^*\alpha$$

As a special case we get the identity

$$\mathrm{id} : \bullet \longrightarrow \bullet : \mathrm{mon} \quad \left[ \mathrm{id} : \bullet^1 \longrightarrow \bullet^1 : \mathrm{mon} \right]$$
$$x \rightsquigarrow x$$

Note that all the trivial operations of arity $\bullet^n \to \bullet^i$ are projections

$$\pi_i^n : \bullet^n \longrightarrow \bullet : \mathrm{mon}$$
$${}^*\langle x_1, \dots, x_n \rangle \rightsquigarrow x_i$$

(cartesianness) Another important feature is that we can recover mon from the knowledge of operations of arity $\bullet^a \to \bullet^1$.
So on an example the operation

$$\omega : \bullet^2 \longrightarrow \bullet^3 : \mathrm{mon}$$
$$\langle x, y \rangle \rightsquigarrow \langle x \star y, e, x \star y \star x \rangle$$

has 3 components, at which we can get at with trivial projections

$$\pi_1\omega, \pi_1\omega, \pi_3\omega : \bullet^2 \longrightarrow \bullet : \mathrm{mon}$$
$$\pi_1\omega(x, y) = x \star y$$
$$\pi_1\omega(x, y) = e$$
$$\pi_3\omega(x, y) = x \star y \star x.$$

To put these components back together we have the pairing operation.

Let $\omega : \bullet^a \longrightarrow \bullet^b$, $\eta : \bullet^a \longrightarrow \bullet^c$, then we can pair them up

$$\langle \omega, \eta \rangle : \bullet^a \longrightarrow \bullet^{b+c}$$
$$x \rightsquigarrow \langle \omega(x), \eta(x) \rangle$$

Then $\langle \pi_1\omega, \pi_2\omega \rangle : \bullet^2 \longrightarrow \bullet^2$
$\langle x, y \rangle \rightsquigarrow \langle x \star y, x \star y \star x \rangle$

So we have the iso

$$\bullet^a \longrightarrow \bullet^{b_1 + \dots + b_n} : \mathrm{mon} \underset{\langle \bullet, \dots, \bullet \rangle}{\overset{\text{decomposition}}{\underset{\sim}{\rightleftarrows}}} \left( \bullet^a \longrightarrow \bullet^{b_1} : \mathrm{mon} \right) \times \dots \times \left( \bullet^a \longrightarrow \bullet^{b_n} : \mathrm{mon} \right)$$

As a special case

$$\bullet^a \longrightarrow \bullet^0 : \mathrm{mon} \simeq 1;$$

we'll denote the unique operation $\perp : \bullet^a \longrightarrow \bullet^0$.

This leads us to the following definition

**define** A single-sorted algebraic theory $\mathbb{T}$ consists of

(op) $\forall a, b : \mathbf{Set}_{fin}$.
a set
$$\bullet^a \longrightarrow \bullet^b : \mathbb{T}$$
We call elements of this set the $\mathbb{T}$-operations of arity $\bullet^a \to \bullet^b$.

(comp) $\forall a, b, c : \mathbf{Set}_{fin}$.
a set map
$$\therefore : \left[ \bullet^b \longrightarrow \bullet^c : \mathbb{T} \right] \times \left[ \bullet^a \longrightarrow \bullet^b : \mathbb{T} \right] \longrightarrow \left[ \bullet^a \longrightarrow \bullet^c : \mathbb{T} \right]$$
called composition

(triv) $\forall a, b : \mathbf{Set}_{fin}$.
a set map
$$(\bullet^a \longrightarrow \bullet^b : \mathbf{Set}_{fin}) \longrightarrow (\bullet^a \longrightarrow \bullet^b : \mathbb{T}),$$
we denote
$\forall a : \mathbf{Set}_{fin}$.
$$\mathrm{id}_a := {}^*\mathrm{id}_a : \bullet^a \longrightarrow \bullet^a : \mathbb{T} \quad \text{where } a \xleftarrow{\mathrm{id}_a} a : \mathbf{Set}_{fin}$$
$\forall a, b : \mathbf{Set}_{fin}$.
$$\pi_a := {}^*\iota_a : \bullet^{a+b} \longrightarrow \bullet^a : \mathbb{T} \quad \text{where } a \xrightarrow{\iota_a} a + b \xleftarrow{\iota_b} b : \mathbf{Set}_{fin},$$
$$\pi_b := {}^*\iota_b : \bullet^{a+b} \longrightarrow \bullet^b : \mathbb{T}.$$
are the coproduct inclusions.

(product) $\forall x, a, b : \mathbf{Set}_{fin}$.
a set map
$$\langle \cdot, \cdot \rangle : \left( \bullet^x \longrightarrow \bullet^a : \mathbb{T} \right) \times \left( \bullet^x \longrightarrow \bullet^b : \mathbb{T} \right) \longrightarrow \left( \bullet^x \longrightarrow \bullet^{a+b} : \mathbb{T} \right)$$
$\forall x : \mathbf{Set}_{fin}$.
an element
$$\langle \rangle : \bullet^x \longrightarrow \bullet^0 : \mathbb{T}$$

**where** (category) $\forall \bullet^a \xrightarrow{f} \bullet^b \xrightarrow{g} \bullet^c \xrightarrow{h} \bullet^d : \mathbb{T}$,
$$h(gf) = (hg)f : \bullet^a \longrightarrow \bullet^d : \mathbb{T}$$
$\forall \bullet^a \xrightarrow{f} \bullet^b : \mathbb{T}$.
$$f\,\mathrm{id}_a = f = \mathrm{id}_b f : \bullet^a \longrightarrow \bullet^b : \mathbb{T}$$

(functoriality) $\forall a \xleftarrow{\alpha} b \xleftarrow{\beta} c : \mathbf{Set}_{fin}$.
$${}^*(\alpha\beta) = {}^*\beta\,{}^*\alpha : \bullet^a \longrightarrow \bullet^c : \mathbb{T}$$
$\forall a : \mathbf{Set}_{fin}$.
$${}^*\mathrm{id}_a = \mathrm{id}_a : \bullet^a \longrightarrow \bullet^a : \mathbb{T}$$

(cartesianness) $\forall x, a, b : \mathbf{Set}_{fin}$.
$$\forall \omega : \bullet^x \longrightarrow \bullet^a : \mathbb{T},\ \eta : \bullet^x \longrightarrow \bullet^b : \mathbb{T}.$$
the operation
$$\langle \omega, \eta \rangle : \bullet^x \longrightarrow \bullet^{a+b} : \mathbb{T}$$
is unique such that
$$\pi_a \langle \omega, \eta \rangle = \omega$$
$$\pi_b \langle \omega, \eta \rangle = \eta$$
$\forall x : \mathbf{Set}_{fin}$.
$\langle \rangle$ is the only operation in $\bullet^x \longrightarrow \bullet^0 : \mathbb{T}$ ■

Cartesianness basically states that we have

$$\bullet^a \longrightarrow \bullet^{b_1 + \dots + b_n} : \mathbb{T} \underset{\text{tuple}}{\overset{\text{component}}{\underset{\sim}{\rightleftarrows}}} \left( \bullet^a \longrightarrow \bullet^{b_1} : \mathbb{T} \right) \times \dots \times \left( \bullet^a \longrightarrow \bullet^{b_n} : \mathbb{T} \right)$$

Informally this gives us a finite product preserving functor

$$\mathbb{T} : \mathbf{Set}_{fin}^{op} \longrightarrow \mathbb{T}$$

where $\mathbb{T}$ has finite sets as objects,
& for $a, b : \mathbb{T}$ morphisms from $a$ to $b$ are operations of arity $\bullet^a \to \bullet^b$
& $\iota$ is identity on objects.
Especially the second axiom in cartesianness states that $\bullet^0$ is the terminal object.

So the definition of single-sorted algebraic theory simply becomes

**define** A single-sorted algebraic theory is a Cartesian category on $\mathbf{Set}_{fin}^{op}$ ■

To clarify, $\bullet^0 \longrightarrow \bullet : \mathbb{T}$ contains the constants of the theory.
We'll often denote constants by

$$\bullet : \mathbb{T} := 1 \longrightarrow \bullet : \mathbb{T} := \bullet^0 \longrightarrow \bullet : \mathbb{T}.$$

Given $\omega_1 : \bullet^a \longrightarrow \bullet^{b_1} : \mathbb{T}, \dots, \omega_n : \bullet^a \longrightarrow \bullet^{b_n} : \mathbb{T}$
we can tuple them up, which we denote by

$$\langle \omega_1, \dots, \omega_n \rangle : \bullet^a \longrightarrow \bullet^{b_1 + \dots + b_n} : \mathbb{T}$$

Given $\omega : \bullet^a \longrightarrow \bullet^b : \mathbb{T}$, and $i : b$, we have $b = \{i\} + b \setminus \{i\}$;
denote $\pi_i := \pi_{\{i\}} : \bullet^a \longrightarrow \bullet : \mathbb{T}$

Cartesianness implies

**prop** $\forall a, b : \mathbf{Set}_{fin}$.
$$\left[ \bullet^a \longrightarrow \bullet^b : \mathbb{T} \right] \xrightarrow{\pi_i} b \longrightarrow \left[ \bullet^a \longrightarrow \bullet : \mathbb{T} \right] : \mathbf{Set}$$
$$\omega \rightsquigarrow \lambda i : b.\ \pi_i b$$ ■

An element of $b \longrightarrow \left[ \bullet^a \longrightarrow \bullet : \mathbb{T} \right]$ is just a $b$'s worth of operations in $\bullet^a \longrightarrow \bullet : \mathbb{T}$

An important property to note is

**prop** $\forall \mathcal{X} : \bullet^a \longrightarrow \bullet^{b+c} : \mathbb{T},\ \omega : \bullet^a \longrightarrow \bullet^b : \mathbb{T},\ \eta : \bullet^a \longrightarrow \bullet^c : \mathbb{T}.$
$$\mathcal{X} = \langle \omega, \eta \rangle$$
iff
$$\pi_a \mathcal{X} = \omega,\ \pi_b \mathcal{X} = \eta$$ ■

The most important trivial operations are permutations,
e.g.
$$\delta : \bullet^4 \longrightarrow \bullet^4 : \mathbb{T}$$
$$\langle u, v, s, t \rangle \rightsquigarrow \langle v, s, u, t \rangle$$

projections $\pi_1, \pi_2 : \bullet^2 \longrightarrow \bullet : \mathbb{T}$
$$\langle x, y \rangle \rightsquigarrow x$$
$$\langle x, y \rangle \rightsquigarrow y$$

duplication $\Delta : \bullet \longrightarrow \bullet^2 : \mathbb{T}$
$$x \rightsquigarrow \langle x, x \rangle$$

From these we can generate any other trivial operation with the help of composition and pairing.

Finally we can combine arbitrary operations with independent product

**define** [independent product]
$$\forall \omega : \bullet^a \longrightarrow \bullet^b : \mathbb{T},\ \eta : \bullet^c \longrightarrow \bullet^d : \mathbb{T}.$$
$$\omega \times \eta := \langle \omega \pi_a, \eta \pi_c \rangle : \bullet^{a+c} \longrightarrow \bullet^{b+d} : \mathbb{T}$$ ■

Which is to be thought of as

$$\langle a, c \rangle \rightsquigarrow \langle \omega(a), \eta(c) \rangle$$

as opposed to the pairing

$$x \rightsquigarrow \langle \omega(x), \eta(x) \rangle.$$

Now that we have a definition of algebraic theories, we need to

(1) recover the notion of its models
(2) define morphisms of algebraic theories.

We extract two notions from the previous definition.

**define** Let $\mathbb{A}$ be a category. Then $\mathbb{E}$ a category on $\mathbb{A}$ consists of

(1) a category $\mathbb{E}$ whose objects are those of $\mathbb{A}$

(2) $\forall a, b : \mathbb{A}$.
we have a map
$$\overline{(\ )} : (a \to b : \mathbb{A}) \to (a \to b : \mathbb{E})$$

**where**
$\forall a \xrightarrow{\alpha} b \xrightarrow{\beta} c : \mathbb{A}$,
$$\overline{(\alpha\beta)} = \overline{\alpha}\,\overline{\beta}$$
$\forall a : \mathbb{A}$.
$$\overline{\mathrm{id}_a} = \mathrm{id}_a$$ ■

**define** Let $\mathbb{A}$ be a category with products. Then
$\mathbb{E}$ a category on $\mathbb{A}$ is Cartesian
iff $\forall x, a, b : \mathbb{A}$.
the map
$$\left[ x \longrightarrow a \times b : \mathbb{E} \right] \to \left[ x \to a : \mathbb{E} \right] \times \left[ x \to b : \mathbb{E} \right]$$
$$f \rightsquigarrow \langle \overline{\pi_a} f, \overline{\pi_b} f \rangle$$
is invertible;
$\forall x : \mathbb{A}$.
$$\left[ x \xrightarrow{!} 1 : \mathbb{E} \right] \to 1$$
is invertible.

That is, the canonical functor $\mathbb{A} \to \mathbb{E}$ preserves products. ■

# Models

A **Set model** of a theory $\mathbb{T}$ is simply a set $M$ together with interpretations of all operations in $\mathbb{T}$

$$\omega : \bullet^a \longrightarrow \bullet^b \in \mathbb{T} \quad \rightsquigarrow \quad [\![\omega]\!] : M^a \longrightarrow M^b : \mathbf{Set}$$

that respects the structure of $\mathbb{T}$.

This data can be presented in a functorial way.

**define** Let $\mathbb{T}$ be a single-sorted algebraic theory.

Then $M$, a $\mathbb{T}$-model in $\mathbf{Set}$ consists of

(Fam) $M^{\bullet} : \mathbf{Set}_{fin} \longrightarrow \mathbf{Set}$ a family

(functor) $[\![\cdot]\!] : \forall a, b : \mathbf{Set}_{fin}.$

$$\left(\bullet^a \longrightarrow \bullet^b \in \mathbb{T}\right) \longrightarrow \left(M^a \longrightarrow M^b : \mathbf{Set}\right)$$

where $\forall \bullet^a \xrightarrow{\omega} \bullet^b \xrightarrow{\eta} \bullet^c \in \mathbb{T}.$

$$[\![\eta\omega]\!] = [\![\eta]\!][\![\omega]\!] : M^a \longrightarrow M^c$$

$\forall a : \mathbf{Set}_{fin}.$

$$[\![\mathrm{id}_a]\!] = \mathrm{id}_{M^a} : M^a \longrightarrow M^a$$

(proviso) $\forall a, b : \mathbf{Set}_{fin}.$

$$\forall m : M^a, m' : M^b.$$

$$\exists! \langle m, m' \rangle : M^{a+b}.$$

$$[\![\pi_a]\!] \langle m, m' \rangle = m$$

$$[\![\pi_b]\!] \langle m, m' \rangle = m'$$

$M^0$ has a unique element $\diamond$ ■

This gives us a product-preserving functor

$$M : \mathbb{T} \longrightarrow \mathbf{Set}$$

**define** Let $M, N \in \mathbb{T}$-models.

A homomorphism $f$ from $M$ to $N$ consists of

$$f : \forall a : \mathbf{Set}_{fin}.$$

$$M^a \xrightarrow{f^a} N^a$$

where $\forall \bullet^a \xrightarrow{\omega} \bullet^b \in \mathbb{T}.$

$$\begin{tikzcd} M^a \arrow[r, "f^a"] \arrow[d, "{[\![\omega]\!]}_M"'] & N^a \arrow[d, "{[\![\omega]\!]}_N"] \\ M^b \arrow[r, "f^b"'] & N^b. \end{tikzcd} ■$$

This is precisely a natural transformation

$$f : M \Rightarrow N : \mathbb{T} \longrightarrow \mathbf{Set}.$$

This is what Lawvere called **Functorial Semantics**:

Theories are Categories
Models are Functors to Set
Homomorphisms are Natural Transformations.

Introducing notation

$$\omega(m) := [\![\omega]\!] m,$$

$$f(m) = f^a(m),$$

the homomorphism condition becomes familiar

$$\forall m : M^a.$$

$$f(\omega(m)) = \omega(f(m))$$

which is precisely what we have seen for Monoids.

**prop** We get the category of $\mathbb{T}$-models in $\mathbf{Set}$ which we denote

$$\mathbb{T}\mathbb{C}\mathbf{Set}. ■$$

Obviously these definitions could be extended to models valued in arbitrary categories with finite products

$$\mathbb{T}\mathbb{C}\mathbb{E}.$$

We have the forgetful functor

**define** $\|\cdot\|_{\mathbb{T}} : \mathbb{T}\mathbb{C}\mathbf{Set} \longrightarrow \mathbf{Set}$

$\forall M : \mathbb{T} \times \mathbf{Set}.$

$$\|M\|_{\mathbb{T}} := M^1$$

$\forall f : M \longrightarrow N : \mathbb{T}\mathbb{C}\mathbf{Set}$

$$\|f\|_{\mathbb{T}} = f^1 : M^1 \longrightarrow N^1 ■$$

We often write $\|M\|_{\mathbb{T}} = M^1$.

Note that given a constant of the theory $k : \bullet^0 \longrightarrow \bullet \in \mathbb{T}$, it's interpretation is simply an element $[\![k]\!] : M$.

The following example is meant to familiarize the reader with all the trivial special details.

**example** Each field $K$ defines an algebraic theory $\mathbf{vec}_K$ whose models are vector spaces over $K$ and homomorphisms are linear maps.

The operations are all linear combinations.

To make this precise consider the following linear combinations:

$$\omega : \bullet^2 \longrightarrow \bullet : \mathbf{vec}_K$$

$$(v, w) \mapsto 3v - 6w,$$

or a pair of linear combinations

$$\eta : \bullet^3 \longrightarrow \bullet^2 : \mathbf{vec}_K$$

$$(a, b, c) \mapsto (2a - 3b, -b + 6c).$$

Composing we have

$$\omega\eta : \bullet^3 \longrightarrow \bullet : \mathbf{vec}_K$$

$$(a, b, c) \mapsto 3(2a - 3b) - 6(-b + 6c)$$

$$= 6a + (-9 + 6)b - 36c$$

$$= 6a - 3b - 36c$$

Note that to specify $\omega$ & $\eta$ is to give matrices

$$\omega = \begin{bmatrix} 3 & -6 \end{bmatrix}, \quad \eta = \begin{bmatrix} 2 & -3 & 0 \\ 0 & -1 & 6 \end{bmatrix}$$

and the composition is just matrix multiplication

$$\omega\eta = \begin{bmatrix} 3 & -6 \end{bmatrix} \begin{bmatrix} 2 & -3 & 0 \\ 0 & -1 & 6 \end{bmatrix} = \begin{bmatrix} 6 & -3 & -36 \end{bmatrix}.$$

For this reason we can define algebraic theory $K$ as

**define** $\bullet^n \longrightarrow \bullet^m : \mathbf{vec}_K := \left\{ \text{from-}n\text{-to-}m \text{ matrices over } K \right\}$

$$\simeq \left\{ f : K^n \longrightarrow K^m \text{ linear} \right\}$$

with matrix multiplication as composition. ■

The set $\bullet^0 \longrightarrow \bullet : \mathbf{vec}_K$ contains only the empty matrix $[\,]$, which should be thought of as the trivial linear combination whose value is zero.

It is more enlightening to give examples of trivial operations than to define them in general.

The projection $p : \bullet^2 \longrightarrow \bullet^1 : \mathbf{vec}_K$

$$(v, w) \mapsto (v, 0) \text{ or}$$

$$p := \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

The duplication

$$\Delta : \bullet \longrightarrow \bullet^2 : \mathbf{vec}_K$$

$$x \mapsto (x, x)$$

is $\Delta = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

We have two basic ways to combine operations, one is to pair them up:

from $\omega : \bullet^m \longrightarrow \bullet^0 \in \mathbb{T}$, $\eta : \bullet^n \longrightarrow \bullet^0 \in \mathbb{T}$

we get $\langle \omega, \eta \rangle : \bullet^n \longrightarrow \bullet^{n+m} \in \mathbb{T}$.

In $K$ this is realized as follows:

$$\begin{bmatrix} 3 & 5 & -2 \\ 3 & 0 & 0 \end{bmatrix} : \bullet^3 \longrightarrow \bullet^2 : \mathbf{vec}_K, \quad \begin{bmatrix} 0 & 0 & 2 \end{bmatrix} : \bullet^3 \longrightarrow \bullet^1 : \mathbf{vec}_K$$

pair up to

$$\left\langle \begin{bmatrix} 3 & 5 & -2 \\ 3 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 2 \end{bmatrix} \right\rangle = \begin{bmatrix} 3 & 5 & -2 \\ 3 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix} : \bullet^3 \longrightarrow \bullet^3 : \mathbf{vec}_K$$

The other is independent product:

$$\omega : \bullet^m \longrightarrow \bullet^0 \in \mathbb{T}, \quad \eta : \bullet^n \longrightarrow \bullet^0 \in \mathbb{T}$$

gets us $\omega \times \eta : \bullet^{m+n} \longrightarrow \bullet^{0+0} \in \mathbb{T}$.

In $K$,

$$\begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \times \begin{bmatrix} 3 & 0 & -5 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 3 & -3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -5 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} : \bullet^5 \longrightarrow \bullet^5 : \mathbf{vec}_K$$

It is easy to show that vector spaces give a model of $\mathbf{vec}_K$

Let $V : \mathbf{Vec}_K$. Define

$$V^{\bullet} : \mathbf{vec}_K \circlearrowleft \mathbf{Set}$$

$$V^n = V^{\bullet n} : \mathbf{Set}$$

& we interpret $[\![\bullet^n \longrightarrow \bullet^m]\!] : \bullet^n \longrightarrow \bullet : \mathbf{vec}_K$

as $[\![\bullet^n \longrightarrow \bullet^m]\!] : V^n \longrightarrow V$

$$[\![\bullet^n \longrightarrow \bullet^m]\!](v_1, \ldots, v_n) \mapsto k_1 v_1 + k_2 v_2 + \ldots + k_n v_n$$

This indeed defines a model because linear combination of linear combination is again a linear combination

Given a linear map $f : V \longrightarrow W : \mathbf{Vec}_K$ we have

$$f^{\bullet} : V \longrightarrow W : \mathbf{vec}_K$$

$$f^{\bullet}(v_1, \ldots, v_n) = (f(v_1), \ldots, f(v_n)),$$

thus a model homomorphism because $f$ is linear

$$f^{\bullet}([\![\bullet^n \longrightarrow \bullet^m]\!](v_1, \ldots, v_n)) = f^{\bullet}(k_1 v_1 + k_2 v_2 + \ldots + k_n v_n)$$

$$= k_1 f(v_1) + k_2 f(v_2) + \ldots + k_n f(v_n)$$

$$= [\![\bullet^n \longrightarrow \bullet^m]\!](f(v_1), \ldots, f(v_n))$$

Now we'll look at models of $K$ to convince ourselves that these are indeed the usual vector spaces.

**prop** Let $V : \mathbf{vec}_K \circlearrowleft \mathbf{Set}$,

then the underlying set of $V$ has a canonical vector space structure

**proof** Given $\omega : \bullet^n \longrightarrow \bullet^m : \mathbf{vec}_K$ we think of the interpretation in $V$ by

$$\omega := [\![\omega]\!] : V^n \longrightarrow V^m : \mathbf{Set}$$

The addition comes from interpretation of

$$+ := [\,1\;1\,] : \bullet^2 \longrightarrow \bullet : \mathbf{vec}_K \text{ in } V \quad (\text{intuitively } \langle v, w \rangle \rightsquigarrow 1 \cdot v + 1 \cdot w)$$

that is $\forall v, w : V$

$$v + w := [\![+]\!](v, w) := [\![1\;1]\!](v, w) : V \times V \longrightarrow V;$$

subtraction from $- := [\,-1\,] : \bullet \longrightarrow \bullet : \mathbf{vec}_K$

multiplication by $k : K$ from

$$k := [\![k]\!] : \bullet \longrightarrow \bullet : \mathbf{vec}_K \quad [v \mapsto k \cdot v]$$

zero from $0 := [\,]\,] : \bullet^0 \longrightarrow \bullet$ using the empty matrix.

Let $v, w, u : V$, $k, k' : K$.

(commutativity) We illustrate commutativity of $+$:

we need to show

$$[\,1\;1\,](v, w) = [\,1\;1\,](w, v).$$

Q. How can we get the pair $(w, v)$ from $(v, w)$?

we apply $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} : \bullet^2 \longrightarrow \bullet^2 : \mathbf{vec}_K$ $[(v, w) \mapsto (w, v)]$.

So $[\,1\;1\,](w, v) = [\,1\;1\,] \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} (v, w)$

but $[\,1\;1\,] \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = [\,1\;1\,] : \bullet^2 \longrightarrow \bullet : K$

So commutativity holds.

(associativity) We need to show

$$[\,1\;1\,]([\,1\;1\,](v, w), u) = [\,1\;1\,](v, [\,1\;1\,](w, u)) : V.$$

To get at the first expression, we need to extract from the triple $(v, w, u)$ the pair $(v, w)$ & the singleton $u$.

These projections are given by

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} : \bullet^3 \longrightarrow \bullet^2$$

& $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix} : \bullet^3 \longrightarrow \bullet$.

Note that pairing them up we have the identity

$$\left\langle \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \right\rangle = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} : \bullet^3 \longrightarrow \bullet^3$$

Then to get at $[\,1\;1\,](v, [\,1\;1\,](w, u))$, we apply

$$[\,1\;1\,]\left\langle [\,1\;1\,] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, [\,0\;0\;1\,] \right\rangle = [\,1\;1\,]\left\langle [\,1\;1\;0\,], [\,0\;0\;1\,] \right\rangle : \bullet^3 \longrightarrow \bullet$$

$$= [\,1\;1\,] \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= [\,1\;1\;1\,]$$

To get at $[\,1\;1\,](v, [\,1\;1\,](w, u))$, we apply

$$[\,1\;1\,]\left[\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, [\,1\;1\,] \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\right] = [\,1\;1\,] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$= [\,1\;1\;1\,]$$

(neutrality) $v + 0 = v$ is implied by

$$[\,1\;1\,]([\,1\,] \times [\,1\,]) = [\,1\;1\,] \begin{bmatrix} 1 \\ 0 \end{bmatrix} : \bullet \longrightarrow \bullet$$

$$= [\,1\,].$$

At first it may be surprising that

$$[\,1\,] \times [\,1\,] = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \& \quad \text{not } [\,1\,] \times [\,1\,] = [\,1\,]$$

but we need to remember that by $[\,]$ we mean

$$[\,] : \bullet^0 \longrightarrow \bullet : \mathbf{vec}_K$$

& not $\perp : \bullet^0 \longrightarrow \bullet^1 : \mathbf{vec}_K$

So $[\,1\,] \times [\,1\,]$ is in $\bullet \longrightarrow \bullet^2 : \mathbf{vec}_K$ and not in $\bullet \longrightarrow \bullet : \mathbf{vec}_K$.

(additive inverse) $v + -v = 0$ is implied by

$$[\,1\;1\,]([\,1\,] \times [\,-1\,]) = [\,1\;1\,] \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$= [\,1\;1\,] \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$= [\,0\,]$$

Note the two linear combination is not $(1 \cdot \bullet^0 \longrightarrow \bullet$, but $[\,0\,] : \bullet \longrightarrow \bullet$.

(distributivity) $k \cdot (v + w) = k \cdot v + k \cdot w$:

$$[k][\,1\;1\,] = [\,1\;1\,]\left\langle [k][\,1\;0\,], [k][\,0\;1\,] \right\rangle$$

$$= [\,1\;1\,] \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$$

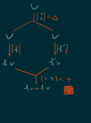
$$= [k\;k]$$

$$k \cdot v + k' \cdot v = (k + k') \cdot v:$$

$$[\,1\;1\,]([k][\,1\,] \times [k'][\,1\,]) = [\,1\;1\,] \begin{bmatrix} k & 0 \\ 0 & k' \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$= [\,1\;1\,] \begin{bmatrix} k \\ k' \end{bmatrix}$$

$$= [k + k']$$




■

**prop** Given $f : V \longrightarrow W : \mathbf{vec}_K \circlearrowleft \mathbf{Set}$,

the underlying set function is a linear map

**proof** It is enough to note that

$$\forall k, k' : K$$

$$\forall v, w : V.$$

$$f^{\bullet}([\![k \quad k']\!](v, w)) = [\![k \quad k']\!](f(v), f(w))$$

i.e. $f(k \cdot v + k' \cdot w) = k \cdot f(v) + k' \cdot f(w)$

since $k \cdot v + k' \cdot w = [\,1\;1\,]([k][\,1\,] \times [k'][\,1\,])(v, w)$

$$= [\,1\;1\,] \begin{bmatrix} k & 0 \\ 0 & k' \end{bmatrix} (v, w)$$

$$= [k \quad k'](v, w) ■$$

In this example we thought of vector spaces as sets in which we can perform arbitrary linear combinations.

The standard definition of vector spaces privileges a set of operations – a particular presentation of $\mathbf{vec}_K$.

Some theories have multiple convenient presentations.

Some theories do not admit any reasonable presentation – such theories are often overlooked in the classical framework of universal algebra which heavily emphasizes presentations.

Algebraic Theories, which consider all operations as a unified whole, are an analogous step to when people did their work only in coordinates, and later realized the invariant approach through the concept of a vector space and linear map. At first it may have been considered pointless generality, too removed from actual practice. But even the definition of matrix multiplication is hard to swallow when you do not know about composition of linear maps – ask any undergraduate in a first computational course on linear algebra.

This is the reason why an invariant study of algebraic theories is important, we gain a new perspective so we can see further and discover new examples, new explanations, we gain a new language to better express and communicate our thoughts.

I'll try to make my case by developing further examples.

This is not to say that we should abandon presentations – far from it. Even in the simple case of vector spaces its pretty cumbersome to work only with matrices without choosing the natural presentation. The point is – presentations are not the underlying stuff, they are a poor choice for conceptual understanding. ■

**example** Actually, in the last example we did not take advantage of all of the field structure.

This works for arbitrary ring $R$ [unital, not necessarily commutative], we denote the theory by $\mathbf{mod}_R$.

Then $\mathbf{mod}_R \circlearrowleft \mathbf{Set}$ is the category of left modules of $R$. ■

It is convenient to develop variable-like notation for the abstract operations in an algebraic theory.

Given algebraic theory $\mathbb{T}$, we'll often denote the projections

$$\pi_1, \pi_2, \pi_3 : \bullet^3 \longrightarrow \bullet : \mathbb{T}$$

by $x := \pi_1$, $y := \pi_2$, $z := \pi_3$

as if they were variable names.

Then $id = \langle x, y, z \rangle : \bullet^3 \longrightarrow \bullet^3 : \mathbb{T}$.

Given $\omega : \bullet^3 \longrightarrow \bullet : \mathbb{T}$, we can without harm write

$$\omega \langle x, y, z \rangle \quad \text{for } \omega,$$

since $\omega \langle x, y, z \rangle = \omega\, id = \omega$.

Trivially,

$$\omega = \eta \quad \# \quad \omega \langle x, y, z \rangle = \eta \langle x, y, z \rangle.$$

We can easily denote permutations such as

$$\langle x, z, y \rangle \;\; \text{or} \;\; \langle z, x, y \rangle : \bullet^3 \longrightarrow \bullet$$

Assuming $x := id : \bullet \longrightarrow \bullet : \mathbb{T}$, we can express duplication as

$$\Delta = \langle x, x \rangle$$

Assuming $\langle x, y \rangle := id : \bullet^2 \longrightarrow \bullet^2$, for $\omega : \bullet^m \longrightarrow \bullet$, $\eta : \bullet^k \longrightarrow \bullet$, we have

$$x \langle \omega, \eta \rangle = \omega$$

$$y \langle \omega, \eta \rangle = \eta$$

Such usage heavily overloads names $x, y, z \ldots$

It requires us to remember the arities of projections – e.g. $\langle x, x \rangle$ means different things depending on whether $x = id : \bullet \longrightarrow \bullet$ or $\langle x, y \rangle := id : \bullet^2 \longrightarrow \bullet^2$;

in the first case $\langle x, x \rangle : \bullet \longrightarrow \bullet^2$, $x \longmapsto \langle x, x \rangle$; in the second $\langle x, x \rangle : \bullet^2 \longrightarrow \bullet^2$, $\langle x, y \rangle \longmapsto \langle x, x \rangle$.

In the example of $\mathbf{vec}_K$, setting $+ := [1 \;\; 1] : \bullet^2 \longrightarrow \bullet$, $\langle u, v, w \rangle := id : \bullet^3 \longrightarrow \bullet^3$, we can write the linear combination

$$[3 \;\; 4 \;\; 2] : \bullet^3 \longrightarrow \bullet \quad \big[ \langle u, v, w \rangle \longmapsto 3 \cdot u + 4 \cdot v + 2 \cdot w \big]$$

as $\big( [3] u + [4] v \big) + [2] w$, where $\omega + \eta := + \langle \omega, \eta \rangle$

Or

$$\langle w, v \rangle = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} : \bullet^3 \longrightarrow \bullet^2$$

How does this notation work in models?

Let $M : \mathbb{T} \xrightarrow{\mathcal{C}^2} \mathbf{Set}$, then assuming $\langle x, y, z \rangle := id : \bullet^3 \longrightarrow \bullet^3$, we have

$$[\![ x ]\!] : M^3 \longrightarrow M$$

$$\forall a, b, c : M. \quad [\![ x ]\!](a, b, c) = a.$$

Stating it once more

$$\forall x, y, z : M. \quad [\![ x ]\!](z, y, x) = z.$$

Or $[\![ \langle x, x \rangle ]\!](x, y, z) = \langle x, x \rangle$

**example** [Boolean Algebras]

Let $Bool := \{ \bot, \top \} : \mathbf{Set}$ be a two element set.

**define** We define an algebraic theory $bool$

$$\bullet^a \longrightarrow \bullet^b : bool := Bool^a \longrightarrow Bool^b$$

composition is composition of functions ■

Trivial operations are the obvious ones like

$$\pi_2 : \bullet^3 \longrightarrow \bullet : bool$$
$$\langle x, y, z \rangle \mapsto y.$$

We can think of operations in $\bullet^n \longrightarrow \bullet : bool$ as truth tables.

For example

$$\wedge : \bullet^2 \longrightarrow \bullet : bool$$
$$\langle \top, \top \rangle \mapsto \top$$
$$\langle \bot, \top \rangle \mapsto \bot$$
$$\langle \top, \bot \rangle \mapsto \bot$$
$$\langle \bot, \bot \rangle \mapsto \bot$$

is the familiar truth table for AND

$$\left[\begin{array}{c|c|c} \top & \top & \top \\ \top & \bot & \bot \\ \bot & \top & \bot \\ \bot & \bot & \bot \end{array}\right]$$

Similarly,

negation $\neg := \left[\begin{array}{c|c} \top & \bot \\ \bot & \top \end{array}\right] : \bullet \longrightarrow \bullet : bool$

xor (exclusive disjunction), or, $\vee$ read

$$\langle \oplus, \vee, NAND \rangle := \left[\begin{array}{cc|ccc} \top & \top & \bot & \top & \top \\ \top & \bot & \top & \top & \top \\ \bot & \top & \top & \top & \top \\ \bot & \bot & \bot & \top & \top \end{array}\right] : \bullet^2 \longrightarrow \bullet^3 : bool$$

Note $NAND = \neg \wedge$.

We have two special constants

$$\bot, \top : 1 \longrightarrow \bullet : bool.$$

Another, more unusual operation is

$$\text{if} \cdot \text{then} \cdot \text{else} \cdot : \bullet^3 \longrightarrow \bullet : bool$$
$$\text{if } \top \text{ then } x \text{ else } y := x$$
$$\text{if } \bot \text{ then } x \text{ else } y := y.$$

This is not to be confused with implication

$$\Rightarrow := \left[\begin{array}{c|c|c} \top & \top & \top \\ \top & \bot & \bot \\ \bot & \top & \top \\ \bot & \bot & \top \end{array}\right] : \bullet^2 \longrightarrow \bullet : bool$$

The collection of operations

$$\{ \bot, \top : 1 \longrightarrow \bullet; \;\; \neg : \bullet \longrightarrow \bullet; \;\; \vee, \wedge : \bullet^2 \longrightarrow \bullet \}$$

satisfy

$$(x \vee y) \vee z = x \vee (y \vee z) \qquad (x \wedge y) \wedge z = x \wedge (y \wedge z)$$
$$x \vee y = y \vee x \qquad x \wedge y = y \wedge x$$
$$x \vee \bot = x \qquad x \wedge \top = x$$
$$x \vee x = x \qquad x \wedge x = x$$
$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$$
$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$$
$$\neg\neg x = x$$
$$x \vee \top \vee x = \top$$
$$x \wedge \neg x = \bot.$$

So models of $bool$ are Boolean Algebras.

It's not hard to see that $bool \circlearrowleft Set = Bool$.

We'll show that the above operations generate the whole theory. By this we mean that every operation $\omega : \bullet^n \longrightarrow \bullet^m : bool$ is expressible in terms of trivial operations, composition, pairing, and the operations in the generating subset.

Consider

$$\omega := \left[\begin{array}{ccc|c} \top & \top & \top & \bot \\ \top & \top & \bot & \top \\ \top & \bot & \top & \bot \\ \top & \bot & \bot & \top \\ \bot & \top & \top & \bot \\ \bot & \top & \bot & \bot \\ \bot & \bot & \top & \top \\ \bot & \bot & \bot & \top \end{array}\right] : \bullet^3 \longrightarrow \bullet : bool$$

First we'll figure out how to present atomic operations – those that are $\top$ precisely for one input, and $\bot$ otherwise.

Consider $\eta : \bullet^3 \longrightarrow \bullet : bool$ s.t. $\eta$ $\top$ precisely on $\langle \top, \top, \bot \rangle$

Supposing $\langle x, y, z \rangle := id : \bullet^3 \longrightarrow \bullet^3$,

it is not hard to see that

$$\eta = x \wedge y \wedge \neg z.$$

Obviously we can express each atomic part in terms of $\{ \wedge : \bullet^2 \longrightarrow \bullet, \; \neg : \bullet \longrightarrow \bullet \}$.

Then to express generic operation we OR its atomic parts, e.g.

$$\omega = (x \wedge y \wedge \neg z) \vee (x \wedge \neg y \wedge \neg z) \vee (\neg x \wedge \neg y \wedge \neg z).$$

A different generating set is

$$\{ \bot, \top : 1 \longrightarrow \bullet; \;\; \oplus, \wedge : \bullet^2 \longrightarrow \bullet \}.$$

These satisfy ring axioms with $\oplus$ as addition, $\wedge$ as multiplication, $\bot$ as zero, $\top$ as one, & $id : \bullet \longrightarrow \bullet$ as minus (since $x \oplus x = \bot \Rightarrow x + x = 0$).

Such rings are called boolean rings, because every element is idempotent, i.e. $x \wedge x = x$.

To see that we have a generating set, it is enough to express $\vee : \bullet^2 \longrightarrow \bullet$ & $\neg : \bullet \longrightarrow \bullet$ as

$$x \vee y = x \oplus y \oplus x \wedge y.$$
$$\neg x = \top \oplus x$$

We have even smaller generating set

$$\{ \top : 1 \longrightarrow \bullet; \; NAND : \bullet^2 \longrightarrow \bullet \},$$

to see this note

$$\neg x = NAND(x, x)$$
$$\bot = \neg \top$$
$$x \wedge y = \neg NAND(x, y)$$
$$x \vee y = \neg((\neg x) \wedge (\neg y)). \; ■$$

From the last example we abstract the idea of a generating set of operations.

**define** Let $\mathbb{T}$ be an algebraic theory.

Then an internal subtheory

$$S \subseteq \mathbb{T} \text{ is}$$

a family of subsets $\forall a, b : Set_{\mathbb{T}}$,

$$S_{a,b} \subseteq \bullet^a \longrightarrow \bullet^b \mathbb{T}$$

where (1) $\forall$ trivial $\omega : \bullet^a \longrightarrow \bullet^b : \mathbb{T}$,

$$\omega : S_{a,b}$$

(2) $\forall \omega : S_{a,b}, \eta : S_{b,c}$

$$\eta \omega : S_{a,c}$$

(3) $\forall \omega : S_{a,b}, \eta : S_{a,c}$

$$\langle \omega, \eta \rangle : S_{a,b+c} \; ■$$

Denote the collection of internal subtheories of $\mathbb{T}$ by $Sub(\mathbb{T})$.

It is naturally a poset.

We always have the largest subtheory $1_{\mathbb{T}} : Sub(\mathbb{T})$ and the smallest $\emptyset$ which contains only the trivial operations.

**prop** Let $S : I \longrightarrow Sub(\mathbb{T})$ be a family of subtheories indexed by $I : Set$.

Then $\bigcap_{i:I} S^i$ defined by

$$\Big[ \bigcap_{i:I} S^i \Big]_{a,b} := \bigcap_{i:I} S^i_{a,b}$$

is again a subtheory. ■

This makes $Sub(\mathbb{T})$ into a complete lattice. So it makes sense to generate the least subtheory containing a chosen subset of operations.

**define** Given a family of subsets $\forall a, b : Set_{\mathbb{T}}$, $S_{a,b} \subseteq \bullet^a \longrightarrow \bullet^b : \mathbb{T}$,

$S$ is a generating subset

$:\Leftrightarrow$ the least internal subtheory containing $S$ is $1_{\mathbb{T}} : Sub(\mathbb{T})$ ■

Later we'll develop an external point of view of generating algebraic theories from operations & equations without reference to some encompassing theory – which is traditionally the starting point of classical universal algebra.

**example** Given a field $K$, the algebraic theory $\mathbf{vec}_K$ is generated by

$$\{ + := [1 \;\; 1] : \bullet^2 \longrightarrow \bullet; \;\; [k] : \bullet \longrightarrow \bullet \mid k : K; \; 0 : 1 \longrightarrow \bullet \}$$

It's enough to show that we can generate any $\omega : \bullet^n \longrightarrow \bullet : K$.

We'll show how we can express

$$\omega(u, v, w) \longmapsto a \cdot u + b \cdot v + c \cdot w$$

i.e.

$$[a \;\; b \;\; c] : \bullet^3 \longrightarrow \bullet : K \quad \text{for } a, b, c : K$$

Just note, assuming $\langle u, v, w \rangle := id : \bullet^3 \longrightarrow \bullet^3$,

$$[a \;\; b \;\; c] = [a] u + [b] v + [c] w$$
$$= + \langle + \langle [a] u, [b] v \rangle, [c] w \rangle$$

since

$$+ \langle [a] u, [b] v \rangle = [1 \;\; 1] \langle [1] [1 \;\; 0 \;\; 0], [b] [0 \;\; 1 \;\; 0] \rangle$$
$$= [1 \;\; 1] \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \end{bmatrix}$$
$$= [a \;\; b \;\; 0],$$

&

$$+ \langle [a \;\; b \;\; 0], [c] w \rangle = [1 \;\; 1] \begin{bmatrix} a & b & 0 \\ 0 & 0 & c \end{bmatrix}$$
$$= [a \;\; b \;\; c] \; ■$$

**example** [Sets]

The simplest algebraic theory is set

$$\bullet^{a} \longrightarrow \bullet^{b} : \mathrm{set} := a \longleftarrow b : \mathrm{Set}_{op}.$$

Note the reversal of arrows.

Thus set contains only the trivial operation,

obviously $\mathrm{set} \otimes \mathrm{Set} \cong \mathrm{Set}$

As we've already seen, assuming $\langle x, y, z \rangle := \mathrm{id} : \bullet^{3} \to \bullet^{3} : \mathrm{set}$,

we can represent $\langle y, x, x, z \rangle : \bullet^{3} \to \bullet^{4}$

as

$$\begin{tikzcd} 0 & 0 \arrow[l] \\ 1 \arrow[r] & 1 \\ 2 \arrow[r] & 2 \end{tikzcd}$$

This theory plays a special role among all the algebraic theories—

it is the initial theory; this we'll discuss later.

For this reason we'll often write $0 := \mathrm{set}$. ■

**example** [Commutative Rings]

As we've seen before, each ring gives rise to the theory of its left modules.

But we also have the theory of rings, whose models are all rings.

We'll consider the case of commutative rings.

The operations are integer polynomial combinations like

$$\langle r, s \rangle \mapsto 3r^{2} - 7r^{8}s^{2} + 5.$$

define $\bullet^{n} \longrightarrow \bullet^{m} : \mathrm{cring} := \{ \omega : \mathbb{Z}^{n} \to \mathbb{Z}^{m} \text{ polynomial functions} \}$ ■

We will do non-commutative case later ■

**example** [Commutative $\mathbb{R}$-Algebras]

We can change the integer coefficients from the previous case

to real coefficients.

define $\bullet^{n} \longrightarrow \bullet^{m} : \mathrm{calg}_{\mathbb{R}} := \{ \omega : \mathbb{R}^{n} \to \mathbb{R}^{m} \text{ polynomial functions} \}$ ■

Models of $\mathrm{calg}_{\mathbb{R}}$ are unital, commutative, associative $\mathbb{R}$-algebras.

Of course there is nothing special about $\mathbb{R}$ or $\mathbb{Z}$;

we can have $\mathrm{calg}_{R}$ for $R$ commutative ring or even a commutative algebra ■

**example** [$G$-sets]

Let $G$ be a group. A $G$-set $X$ is a set s.t.

we can multiply its elements by elements of $G$,

$$g : G,\ x : X \mapsto g \star x : X \quad \text{s.t.}$$

$$(gh) \star x = g \star (h \star x)$$

$$1 \star x = x.$$

Such $G$-sets are models of an algebraic theory

which we again denote by $G$.

It contains, for fixed $g : G$, operations such as

$$\bullet^{2} \longrightarrow \bullet : G$$

$$\langle x, y \rangle \mapsto g \star x$$

Note that we should have

$$\bullet^{a} \longrightarrow \bullet : G \cong G \times a : \mathrm{Set},$$

then, assuming $\langle x_1, \ldots, x_n \rangle := \mathrm{id} : \bullet^{n} \longrightarrow \bullet^{n} : G$,

a generic operation in $\bullet^{n} \longrightarrow \bullet : G$ is $g \star x_i$.

Then we are forced to have

$$\bullet^{a} \longrightarrow \bullet : G \cong b \longrightarrow G \times a : \mathrm{Set}$$

$$\cong (b \longrightarrow a) \times (b \longrightarrow G)$$

$$\cong (\bullet^{a} \longrightarrow \bullet^{b} : \mathrm{set}) \times (b \longrightarrow G).$$

We can see that any operation is

a trivial operation together with a choice of $b$'s worth of $G$-elements.

So for example any operation in $\bullet^{2} \longrightarrow \bullet^{3} : G$ is just $\langle g_2 \star x,\ g_3 \star y,\ g_1 \star x \rangle$.

The composition of

$\omega := \langle g_2 \star x,\ g_3 \star y,\ g_1 \star x \rangle : \bullet^{2} \to \bullet^{3} : G$ with

$\eta := \langle h \star u,\ h' \star u \rangle : \bullet^{3} \to \bullet : G$ where $\langle u, v, w \rangle := \mathrm{id} : \bullet^{3} \to \bullet^{3} : G$

is

$\eta\omega = \langle h g_2 \star y,\ h' g_1 \star x \rangle : \bullet^{2} \to \bullet^{2} : G$. ■

# Smooth Algebras – an Extended Example

The next example requires some familiarity with smooth manifolds.
Denote the category of smooth manifolds by $\mathrm{Man}^\infty$.
Given $X : \mathrm{Man}^\infty$, we have a commutative ring

$$\mathcal{C}^\infty(X) := \{ f : X \to \mathbb{R} \text{ smooth} \}$$

with

$$(f+g)(x) := f(x) + g(x)$$
$$(f \times g)(x) := f(x)\, g(x)$$
$$0(x) := 0$$
$$1(x) := 1.$$

Given a smooth map $H : X \longrightarrow Y : \mathrm{Man}^\infty$, we get a ring homomorphism by pullback of functions

$$(-)^* H : \mathcal{C}^\infty(Y) \longrightarrow \mathcal{C}^\infty(X)$$
$$(f)^* H := f \circ H$$

Ring $\mathcal{C}^\infty(X)$ has more structure – it is a commutative $\mathbb{R}$-algebra, since we can define

$$(r \cdot f)(x) := r f(x).$$

But we can do much more. Consider the function $\cos : \mathbb{R} \to \mathbb{R}$.
We can extend it to an operation

$$\cos : \mathcal{C}^\infty(X) \longrightarrow \mathcal{C}^\infty(X)$$
$$\cos(f)(x) := \cos(f(x))$$

Similarly, given a smooth $\omega : \mathbb{R}^n \to \mathbb{R}$, we get

$$\omega : \mathcal{C}^\infty(X)^{\times n} \longrightarrow \mathcal{C}^\infty(X)$$
$$\omega(f_1, \dots, f_n)(x) := \omega(f_1(x), \dots, f_n(x))$$

We can check that $(-)^* H : \mathcal{C}^\infty(Y) \longrightarrow \mathcal{C}^\infty(X)$ preserves such operations,

$$(\omega(g_1, \dots, g_n))^* H(x) = \omega(g_1(H(x)), \dots, g_n(H(x)))$$
$$= \omega((g_1, \dots, g_n)^* H)(x).$$

Motivated by this we define the theory of smooth algebras

**define** $\bullet^n \longrightarrow \bullet^m : \mathcal{C}^\infty := \{ \omega : \mathbb{R}^n \to \mathbb{R}^m \text{ smooth} \}$ ■

In the framework of algebraic theories such thoughts do not seem heretical at all – we simply replace polynomial with smooth in the case of $\mathrm{alg}_{\mathbb{R}}$.

**define** A smooth algebra is a model of $\mathcal{C}^\infty$ ■

Good luck finding a reasonable presentation of $\mathcal{C}^\infty$.

**prop** The map

$$\left[ X \xrightarrow{f} Y : \mathrm{Man}^\infty \longrightarrow \mathcal{C}^\infty(Y) \xrightarrow{(-)^* f} \mathcal{C}^\infty(X) : \mathcal{C}^\infty\mathrm{Set} \right]$$

is invertible;

equivalently functor $\mathcal{C}^\infty : \mathrm{Man}^\infty \longrightarrow \mathcal{C}^\infty\mathrm{Set}^{op}$ is an embedding [fully faithful] ■

For details see Smooth Manifolds and Observables, Ch 7
or Synthetic Differential Geometry, Kock, Part 3, §5.

This theorem allows us to think of the category $\mathrm{DSpace} := [\mathcal{C}^\infty\mathrm{Set}]^{op}$ as a category of generalized spaces in which differential calculus is possible.
In this new notation we have an embedding $\mathrm{Man}^\infty \hookrightarrow \mathrm{DSpace}$; we'll think of it as $\mathrm{Man}^\infty \subseteq \mathrm{DSpace}$.

Such spaces in general are not as nice as manifolds – they may contain singularities, they may be infinite dimensional, or even infinitesimal.

Yet they contain spaces that are important to differential geometry, but are not themselves manifolds.

As a warm up consider points of $X : \mathrm{Man}^\infty$.
These are maps $p : 1 \longrightarrow X : \mathrm{Man}^\infty$
and they correspond to

$$p : \mathcal{C}^\infty(X) \longrightarrow \mathbb{R} : \mathcal{C}^\infty\mathrm{Set}.$$

Especially note the condition

$$p(f \times g) = p(f)\, p(g).$$

Turning this around

**define** Given $A : \mathcal{C}^\infty\mathrm{Set}$, denote its counterpart in $\mathrm{DSpace}$ by

$$\mathrm{Spec}[A]$$

We call it the spectrum or the space of the algebra.
We get (definitionally – only a formal notion) the equivalence

$$\mathrm{DSpace}^{op} \underset{\mathrm{Spec}}{\overset{\mathcal{C}^\infty}{\rightleftarrows}} \mathcal{C}^\infty\mathrm{Set} \quad ■$$

This is purely formal! As an objects we have

$$\mathrm{Spec}[A] := A,$$

but we think of $\mathrm{Spec}[A] : \mathrm{DSpace}$ as living in a different category than $A : \mathcal{C}^\infty\mathrm{Set}$.

First, note that the point $1$ as a manifold embedded into $\mathrm{DSpace}$ is no to $\mathrm{Spec}[\mathbb{R}] : \mathrm{DSpace}$ where $\mathbb{R} : \mathcal{C}^\infty\mathrm{Set}$ is thought of as an algebra (and not as a manifold).

Confusion may arise when we consider

$\mathbb{R}$ as a space vs $\mathbb{R}$ as a smooth algebra.

Note

$$\mathbb{R} = \mathrm{Spec}[\mathcal{C}^\infty(\mathbb{R})] : \mathrm{DSpace}$$

But

$$\mathbb{R} = \mathcal{C}^\infty(1) : \mathcal{C}^\infty\mathrm{Set}$$

i.e. $1 = \mathrm{Spec}[\mathbb{R}] : \mathrm{DSpace}$.

Note that the algebra of $\mathbb{R} : \mathrm{DSpace}$ is huge.
But the algebra of a point $1 : \mathrm{DSpace}$ is small – it's even a 1-dimensional vector space over $\mathbb{R}$.

More generally for $X : \mathrm{Man}^\infty$,

$$X = \mathrm{Spec}[\mathcal{C}^\infty(X)] : \mathrm{DSpace}.$$

**define** The points of $\mathrm{Spec}[A] : \mathrm{DSpace}$ form the set

$$|\mathrm{Spec}[A]| := \{ p : 1 \longrightarrow \mathrm{Spec}[A] : \mathrm{DSpace} \}$$
$$:= \{ p : A \longrightarrow \mathbb{R} : \mathcal{C}^\infty\mathrm{Set} \} \quad ■$$

This gives us a representable functor

$$|\mathrm{Spec}| : \mathrm{DSpace} \longrightarrow \mathrm{Set}$$

that sends a space to its underlying set of points.

This is consistent with the the classical notion of a point of a manifold.

The representing object is $1 := \mathbb{R} : \mathrm{DSpace}$;
not a big deal – we have $1$ in $\mathrm{Man}^\infty$.

To see that we get something new we consider the construction of the tangent space; first we'll recall the classical definition of a tangent vector, then we'll extend it to all spaces.

**define** Given $X : \mathrm{Man}^\infty$, a tangent vector at $p : X$ is a map

$$d : \mathcal{C}^\infty(X) \longrightarrow \mathbb{R} : \mathrm{Set}$$

where

(Linearity) $d(r \cdot f + s \cdot g) = r\, d(f) + s \cdot d(g)$

(Leibnitz) $d(f \times g) = f(p)\, d(g) + d(f) \cdot g(p)$;

Set

$$\mathrm{T}_p X := \{ \text{tangent vectors at } p : X \} \quad ■$$

**define**

$$\mathrm{T}X := \sum_{p : X} \mathrm{T}_p X \quad ■$$

Then

$$\begin{array}{c} \mathrm{T}X \\ \downarrow \\ X \end{array} \qquad \begin{array}{c} \langle p, d \rangle \\ \{ \\ p \end{array}$$

can be given the structure of a vector bundle;
we call it the tangent bundle of $X$.

We distinguish between smooth algebras & commutative $\mathbb{R}$-algebras.
For brevity denote $\mathrm{CAlg}_{\mathbb{R}} := \mathrm{alg}_{\mathbb{R}}\mathcal{C}^\infty\mathrm{Set}$, $\mathcal{C}^\infty\mathrm{Alg} := \mathcal{C}^\infty\mathrm{Set}$

Any vector $d : \mathrm{T}_p X$ is a derivation on $\mathbb{R}$-algebra $\mathcal{C}^\infty(X)$ at $p$.
To lift this notion to arbitrary commutative $\mathbb{R}$-algebra, we need to consider the point $p$ as an $\mathbb{R}$-algebra homomorphism $p : \mathcal{C}^\infty(X) \longrightarrow \mathbb{R}$

$$f \rightsquigarrow f(p)$$

**define** Let $p : A \to \mathbb{R} : \mathrm{CAlg}_{\mathbb{R}}$
Then a derivation at $p$ is an $\mathbb{R}$-linear map

$$d : A \longrightarrow \mathbb{R}$$

where

$$d(a \times b) = p(a)\, d(b) + d(a)\, p(b) \quad ■$$

## Dual Numbers

Now we'll see how we can view the tangent bundle in $\mathrm{DSpace}$.

For this purpose we define the algebra of dual numbers $\mathbb{R}[\varepsilon]$.
We can think of it as $\mathbb{R}$ adjoined with an element $\varepsilon \neq 0$ so such that $\varepsilon^2 = 0$. We call such elements first-order infinitesimals.

**define** $\mathbb{R}[\varepsilon] := \dfrac{\mathbb{R}[x]}{(x^2)}$ as $\mathbb{R}$-algebras ■

We can implement $\mathbb{R}[\varepsilon]$ more directly as a 2-dim $\mathbb{R}$-algebra.

**define** Let $\mathbb{R}[\varepsilon] := \mathbb{R} \oplus \mathbb{R}$ as a vector space;
denote the pair $\langle a, b \rangle : \mathbb{R} \oplus \mathbb{R}$ by $a + b\varepsilon : \mathbb{R}[\varepsilon]$.
Set $\varepsilon \cdot \varepsilon := 0$. This makes $\mathbb{R}[\varepsilon]$ into a commutative $\mathbb{R}$-algebra ■

Multiplication is computed as follows

$$(a + b\varepsilon)(c + d\varepsilon) = ac + (ad + bc)\varepsilon + bd\varepsilon^2 : \mathbb{R}[\varepsilon]$$
$$= ac + (ad + bc)\varepsilon.$$

Dual numbers are equipped with two maps

**define** The projection on the first component

$$(-)_0 : \mathbb{R}[\varepsilon] \longrightarrow \mathbb{R}$$
$$(a + b\varepsilon)_0 := a.$$

The projection on the second component

$$(-)_0' : \mathbb{R}[\varepsilon] \longrightarrow \mathbb{R}$$
$$(a + b\varepsilon)_0' := b \quad ■$$

The first projection is an $\mathbb{R}$-algebra homomorphism.

The second is more interesting; it is linear but it doesn't preserve multiplication, but

$$[(a + b\varepsilon)(c + d\varepsilon)]_0' = ad + bc : \mathbb{R},$$

" $\forall z, w : \mathbb{R}[\varepsilon]$,

$$(zw)_0' = z_0' w_0' + z_0' w_0 : \mathbb{R}.$$

So $(-)_0' : \mathbb{R}[\varepsilon] \to \mathbb{R}$ is a derivation at $(-)_0 : \mathbb{R}[\varepsilon] \to \mathbb{R}$.

Now given $V : A \to \mathbb{R}[\varepsilon]$ a commutative $\mathbb{R}$-algebra homomorphism, we get

$V_0 : A \longrightarrow \mathbb{R}$ an $\mathbb{R}$-algebra homomorphism,

& $V_0' : A \longrightarrow \mathbb{R}$ a derivation at $V_0$ since

$$V(a \times b)_0' = [V(a) \times V(b)]_0' : \mathbb{R}$$
$$= V(a)_0\, V(b)_0' + V(a)_0'\, V(b)_0$$

In the opposite direction, given a derivation $d : A \to \mathbb{R}$ at $p : A \to \mathbb{R}$, the map $p + d\varepsilon : A \longrightarrow \mathbb{R}[\varepsilon]$ is an $\mathbb{R}$-algebra homomorphism, since

$$(p + d\varepsilon)(a \times b) = p(a \times b) + d(a \times b)\,\varepsilon$$
$$= p(a)\,p(b) + (p(a)\,d(b) + d(a)\,p(b))\,\varepsilon$$
$$= (p + d\varepsilon)(a)\,(p + d\varepsilon)(b)$$

So derivations of $A$ correspond to homomorphisms $A \longrightarrow \mathbb{R}[\varepsilon]$;
Especially for $X : \mathrm{Man}^\infty$,

$$V : \mathrm{T}X \text{ vector} \simeq V : A \longrightarrow \mathbb{R}[\varepsilon] \ \mathbb{R}\text{-algebra homomorphism}$$

This has a much more satisfying form; it is analogous to

$$p : X \text{ point} \simeq p : A \longrightarrow \mathbb{R} \ \mathcal{C}^\infty\text{-homomorphism}.$$

Here we can see vectors as direct generalizations of points.
We can make this even more apparent.
But why do vectors work with $\mathbb{R}$-algebra homomorphisms, and points with $\mathcal{C}^\infty$-homomorphisms?

We will resolve this by seeing dual numbers as $\mathcal{C}^\infty$-algebra.

To equip $\mathbb{R}[\varepsilon]$ with $C^\infty$-algebra structure we need to extend all $n$-ary smooth functions to $\mathbb{R}[\varepsilon]$.

First, let us consider $f : \bullet \to C^\infty$.

If we informally identify $f$ with its Taylor Expansion at $0$

$$f(x) = f(0) + f'(0)x + \tfrac{1}{2}f''(0)x^2 + \dots$$

then $f(\varepsilon) = f(0) + f'(0)\varepsilon + \tfrac{1}{2}f''(0)\varepsilon^2 + \dots$

$$= f(0) + f'(0)\varepsilon.$$

More generally, to see how $f(a+b\varepsilon)$ should be defined, we can Taylor expand $f$ at $a$ and see

$$f(a+b\varepsilon) = f(a) + f'(a)b\varepsilon.$$

For $\mathbb{R}[\varepsilon]$ to be a $C^\infty$-algebra we need to check

$$[\![g \circ f]\!] = [\![g]\!][\![f]\!],$$

this amounts to the chain rule

$$\begin{aligned} g(f(a+b\varepsilon)) &= g(f(a) + f'(a)b\varepsilon) \\ &= gf(a) + g'(f(a))f'(a)b\varepsilon \\ &= gf(a) + (gf)'(a)b\varepsilon \\ &= gf(a+b\varepsilon). \end{aligned}$$

To generalize this to multivariable case, let $f : \overset{2}{\bullet} \to C^\infty$, we need to make sense out of

$$f(a+c\varepsilon, b+d\varepsilon).$$

Thinking of the tuple $(a+c\varepsilon, b+d\varepsilon)$ as a point + vector

$$(a,b) + \begin{bmatrix} c \\ d \end{bmatrix}\varepsilon$$

we can use the total derivative of $f$.

**define** $\forall \omega : \overset{n}{\bullet} \to \overset{m}{\bullet} : C^\infty$

$\forall p, v : \mathbb{R}^n$.

$$\omega(p + v\cdot\varepsilon) := \omega(p) + d_p\omega(v)\cdot\varepsilon : \mathbb{R}[\varepsilon]^{\times m} \blacksquare$$

**prop** $\mathbb{R}[\varepsilon]$ is a $C^\infty$-algebra

**proof** Let $\overset{n}{\bullet} \xrightarrow{\omega} \overset{m}{\bullet} \xrightarrow{\eta} \overset{l}{\bullet} : C^\infty$, then

$$\begin{aligned} \eta(\omega(p + v\cdot\varepsilon)) &= \eta(\omega(p) + d_p\omega(v)\cdot\varepsilon) \\ &= \eta(\omega(p)) + d_{\omega(p)}\eta(d_p\omega(v))\,\varepsilon \\ &= \eta\omega(p) + d_p\eta\omega(v)\,\varepsilon \\ &= \eta\omega(p + v\cdot\varepsilon) \blacksquare \end{aligned}$$

It can be shown that

**prop** Let $A : \overset{\infty}{C}\mathbf{Set}$

Then $\forall v : A \to \mathbb{R}[\varepsilon]$ $\mathbb{R}$-algebra homomorphism

$$v : A \to \mathbb{R}[\varepsilon] : \overset{\infty}{C}\mathbf{Set} \blacksquare$$

By analysing the second component of $v$ we get the notion of $C^\infty$-derivation which is by the previous theorem amazingly equivalent to $\mathbb{R}$-algebra derivation.

**define** A $C^\infty$-derivation at $p : A \to \mathbb{R} : \overset{\infty}{C}\mathbf{Set}$ is a map

$$d : A \to \mathbb{R} : \mathbf{Set}$$

where $\forall \omega : \overset{n}{\bullet} \to \bullet : C^\infty$.

$\forall a_1, \dots, a_n : A$.

$$d(\omega(a_1, \dots, a_n)) = d_{(a_1, \dots, a_n)}\omega \begin{bmatrix} d(a_1) \\ \vdots \\ d(a_n) \end{bmatrix} \blacksquare$$

In particular this implies for

$$f : \bullet \to C^\infty$$

$$d(g(a)) = f'(g(a))d(a).$$

We denote the space of $\mathbb{R}[\varepsilon]$ by

$$\Delta := \mathrm{Spec}[\mathbb{R}[\varepsilon]] : \mathbf{DSpace}$$

**define** A tangent vector of $X : \mathbf{DSpace}$ is a map $v : \Delta \to X : \mathbf{DSpace}$, denote the tangent space by

$$TX := \Delta \to X \blacksquare$$

Let $A := \overset{\infty}{C}(X)$, then $v : \Delta \to X : \mathbf{DSpace}$ becomes $v : A \to \mathbb{R}[\varepsilon] : \overset{\infty}{C}\mathbf{Set}$, & it is located at $v_0 : A \to \mathbb{R} : \overset{\infty}{C}\mathbf{Set}$, which is a point $v_0 : 1 \to X : \mathbf{DSpace}$.

**define** For $H : X \to Y : \mathbf{DSpace}$, we define the total derivative of $H$ by

$$dH : TX \to TY$$

$$\Delta \xrightarrow{v} X \to \Delta \xrightarrow{H \circ v} Y \blacksquare$$

So the tangent space functor is represented by $\Delta$.

This is reminiscent of the definition of a tangent vector at $x \in X$ as a curve $\gamma : \underline{1} \to X$ defined on some interval $\underline{1}$ going through $x$ consider trip to first order: $\gamma_1 \simeq \gamma_2$ iff in some chart

$$\gamma_1'(0) = \gamma_2'(0)$$

Space $\Delta$ can be thought of as an interval so small that curves $\gamma : \Delta \to X$ retain only the first order information.

We can even think of $\Delta$ as

$$\Delta \subseteq \mathbb{R} : \mathbf{DSpace}$$

since we have the canonical surjective map

$$\overset{\infty}{C}(\mathbb{R}) \twoheadrightarrow \mathbb{R}[\varepsilon]$$

$$f \rightsquigarrow f(0) + f'(0)\varepsilon.$$

Space $\Delta$ contains the point $0 : 1 \to \Delta : \mathbf{DSpace}$ that comes from $(\,)_0 : \mathbb{R}[\varepsilon] \to \mathbb{R} : \overset{\infty}{C}\mathbf{Set}$.

It deserves to be called $0$ because we have

$$\begin{tikzcd} 1 \arrow[r, "0"] \arrow[rr, bend right, "0"'] & \Delta \arrow[r, hook] & \mathbb{R} \end{tikzcd},$$

i.e. $\overset{\infty}{C}(\mathbb{R}) \to \mathbb{R}[\varepsilon] \xrightarrow{(\,)_0} \mathbb{R}$

$$f \rightsquigarrow f(0) + f'(0)\varepsilon \rightsquigarrow f(0)$$

is evaluation at $0$.

In general points of $\Delta$ are maps

$$p : \mathbb{R}[\varepsilon] \to \mathbb{R} : \overset{\infty}{C}\mathbf{Set};$$

but $0 = p(0) = p(\varepsilon^2) = p(\varepsilon)^2 : \mathbb{R}$

so $p(\varepsilon) = 0$,

hence $p(a+b\varepsilon) = p(a) + p(b)\,p(\varepsilon)$

$$= p(a \cdot 1)$$

$$= a \cdot p(1)$$

$$= a.$$

So the only point of $\Delta$ is $0$.

How can this be if $\Delta \neq \{0\}$?

That's because points $p : 1 \to X$ are so called global points – they are points in context $1$; we can change the context by replacing $1$ with some other space.

For example tangent vectors of $X$ are point in context $\Delta$ ($\Delta$-points?)

In general spaces are not determined by their global points like sets are; but they are determined by points in all possible contexts (which in this case are all the other spaces, not just $1$) this is essentially the content of Yoneda lemma.

In some sense we have

$$\Delta = \{x : \mathbb{R} \mid x^2 = 0\}$$

but $\{0\} = \{x : \mathbb{R} \mid x = 0\}$;

more precisely it can be shown that

$\forall$ context $X : \mathbf{DSpace}$.

$$X \to \Delta : \mathbf{DSpace} = \{x : X \to \mathbb{R} : \mathbf{DSpace} \mid x^2 = 0\}$$

but $X \to \{0\} : \mathbf{DSpace} = \{x : X \to \mathbb{R} : \mathbf{DSpace} \mid x = 0\}$

i.e.

$\forall A : \overset{\infty}{C}\mathbf{Set}$.

$$\mathbb{R}[\varepsilon] \to A : \overset{\infty}{C}\mathbf{Set} = \left\{ f : \overset{\infty}{C}(\mathbb{R}) \to A : \overset{\infty}{C}\mathbf{Set} \;\middle|\; \begin{array}{l} \forall x : \mathbb{R}. \\ f(x^2) = 0 \end{array} \right\}.$$

Besides $\Delta$ we have other infinitesimal spaces such as

$$\Delta_n := \mathrm{Spec}\left[\frac{\mathbb{R}[x]}{(x^{n+1})}\right],$$

or $\mathrm{Spec}\left[\dfrac{\mathbb{R}[x,y]}{(x^2, xy, y^2)}\right]$,

$\Delta_n$ can be thought of as the space of $n$-th order infinitesimals

$$\Delta_n := \{x : \mathbb{R} \mid x^{n+1} = 0\}.$$

Collectively the algebras of such spaces are called Weil Algebras and they represent various jet notions of Differential Geometry.

Definitions such as

$$\Delta := \{x : \mathbb{R} \mid x^2 = 0\}$$

don't make direct sense in classical settings.

But there is an approach of Synthetic Differential Geometry [SDG] initiated by Lawvere, Kock & others, where instead of starting off mathematics with sets, we axiomatize smooth spaces as our basic building blocks; in this world every function is differentiable, every space has its tangent space & the lateral definition of $\Delta$ makes sense out of the box.

A basic axiom assumes a ring object $\mathbb{R}$, & states for $\Delta := \{\varepsilon : \mathbb{R} \mid \varepsilon^2 = 0\}$,

$$\forall f : \Delta \to \mathbb{R}.$$

$$\exists!\, a, b : \mathbb{R}.$$

$$\forall \varepsilon : \Delta.$$

$$f(\varepsilon) = a + b\varepsilon$$

i.e. graph of every $\mathbb{R}$-valued function on $\Delta$ is a straight line.

Note $f(0) = a$; denote the slope of the line by $f'(0) := b$.

With this we can define derivative of arbitrary $f : \mathbb{R} \to \mathbb{R}$ at $x : \mathbb{R}$ as

$f'(x) :=$ the unique $d : \mathbb{R}$ s.t. $\forall \varepsilon : \Delta$.

$$f(x+\varepsilon) = f(x) + d\,\varepsilon.$$

With this it is trivial to prove the chain rule:

given $f, g : \mathbb{R} \to \mathbb{R}$, $x : \mathbb{R}$

then $\forall \varepsilon : \Delta$.

$$\begin{aligned} gf(x+\varepsilon) &= g(f(x+\varepsilon)) \\ &= g(f(x) + f'(x)\varepsilon) \quad \text{since } f'(x)\varepsilon : \Delta \\ &= g(f(x)) + g'(f(x))f'(x)\,\varepsilon \\ &= gf(x) + g'(f(x))f'(x)\,\varepsilon, \end{aligned}$$

because derivatives are unique, we know

$$(gf)'(x) = g'(f(x))f'(x).$$

This is a proof of chain rule from THE BOOK!

Such synthetic framework elucidates a lot of calculus & differential geometry

To reflect the results back to classical world one can build various topos models that satisfy the axioms and contain manifolds in a nice way.

The starting point of construction of such models is of course the algebraic theory $C^\infty$, the category $\mathbf{DSet}$ is not actually that great – for example given spaces $X, Y$, the set of all maps $X \to Y$ need not be a space – but it contains nice sites whose sheaves are models of SDG.

For introduction to SDG see

Basic Concepts of Synthetic Differential Geometry, Lavendhomme

first part of Synthetic Differential Geometry, Kock

Synthetic Differential Geometry of Moerdijk, Kock.

For models of SDG see

the rest of Synthetic Differential Geometry, Kock

Models for Smooth Infinitesimal Analysis, Moerdijk, Reyes.

For introduction to differential geometry from the point of view of commutative algebra see

Smooth Manifolds and Observables, Nestruev.

# Free models over finite sets

Every algebraic theory $\mathbb{T}$ has special models.

**example** We can make the collection of $n$-ary operations into a model; for example we can interpret $\omega : \overset{3}{\bullet} \longrightarrow \bullet : \mathbb{T}$ in $\overset{2}{\bullet} \longrightarrow \bullet : \mathbb{T}$:

$$\forall a, b, c : \overset{2}{\bullet} \longrightarrow \bullet : \mathbb{T}.$$
$$[\![\omega]\!](a, b, c) := \omega \langle a, b, c \rangle \quad \blacksquare$$

**define** Given $a : \mathbf{Set}_{fin}$, we make

$$\mathbb{T}[a] := \overset{a}{\bullet} \longrightarrow \bullet : \mathbb{T}$$

into a $\mathbb{T}$-model by interpreting

$$\omega : \overset{n}{\bullet} \longrightarrow \overset{m}{\bullet} : \mathbb{T} \text{ as}$$
$$[\![\omega]\!] : \mathbb{T}[a]^n \longrightarrow \mathbb{T}[a]^m$$
$$\forall x_1, \dots, x_n : \mathbb{T}[a].$$
$$[\![\omega]\!](x_1, \dots, x_n) := \omega \langle x_1, \dots, x_n \rangle \quad \left[\text{note } \langle x_1, \dots, x_n \rangle : \overset{a}{\bullet} \longrightarrow \overset{n}{\bullet}\right].$$

This is a model because

$$[\![\eta\omega]\!](x_1, \dots, x_n) = (\eta\omega)\langle x_1, \dots, x_n \rangle$$
$$= \eta(\omega \langle x_1, \dots, x_n \rangle)$$
$$= [\![\eta]\!]([\![\omega]\!]\langle x_1, \dots, x_n \rangle)$$

$$[\![\mathrm{id}_{\mathbb{T}[a]}]\!](x) = \mathrm{id}_{\mathbb{T}[a]}(x) \quad \blacksquare$$

Taking the set $\{x_1, \dots, x_n\}$ as generators we'll denote

$$\mathbb{T}[x_1, \dots, x_n] := \mathbb{T}[\{x_1, \dots, x_n\}].$$

Such models play a special role in the category of all models; in a certain sense they are generated in a most minimal or free way. More precisely they implement the universal property of free model over a set of generators

**define** A free model of $\mathbb{T}$ over $X : \mathbf{Set}$ [not necessarily finite] consists of

$$\mathbb{T} : \mathbb{T}\text{-}\mathbf{Set} \text{ with } \eta : X \longrightarrow \mathbb{T} : \mathbf{Set}$$

where $\forall M : \mathbb{T}\text{-}\mathbf{Set},\ g : X \longrightarrow M : \mathbf{Set}.$

$$\exists! \bar{g} : \mathbb{T} \longrightarrow M : \mathbb{T}\text{-}\mathbf{Set}.$$
$$\forall x : X.$$
$$\bar{g}(\eta(x)) = g(x) \quad \blacksquare$$

We call $X$ the generating set, $\eta$ the inclusion of generators. We will often denote it by

$$\iota_X := \eta$$

thinking of it as a restriction from model $\mathbb{T}$ to the set of generators $X$. The name inclusion can be misleading; $\eta$ need not be injective.

**example** Consider the theory $\mathbb{1}$ of a point:

$$\overset{a}{\bullet} \longrightarrow \bullet : \mathbb{1} := \{\perp_a\}$$

Any model $M : \mathbb{1}\text{-}\mathbf{Set}$ is a 1-element set since we have a constant

$$c := [\![\perp_a]\!] : M;$$

then $\perp_a : \bullet \to \bullet : \mathbb{1}$ must interpret as identity & as the constant operation returning $\perp_a$. So given $x : M$

$$c = [\![\perp_a]\!](x) = x.$$

Let $a := \{x, y\} : \mathbf{Set}_{fin}$ be a 2-element set, then

$\mathbb{1}[x, y] = \{\perp_x\}$ with the only possible (non-injective) inclusion of generators

is a free model on 2 generators.

This is so because any map $g : \{x, y\} \longrightarrow M : \mathbf{Set}$ to a model $M$ must be the unique terminal map, this uniquely extends to the only terminal map in $\mathbb{1}[x, y] \longrightarrow M$ which is a homomorphism $\blacksquare$

Given such a free model $\mathbb{T}$ over $X \xrightarrow{\eta} \mathbb{T}$, we can restrict any

$$g : \mathbb{T} \longrightarrow M : \mathbb{T}\text{-}\mathbf{Set}$$

to $g|_X : X \longrightarrow M : \mathbf{Set}$ where $\iota_X := \eta$

The universal property states that this process is reversible; that is to define any homomorphism $g : \mathbb{T} \longrightarrow M$ is the same as sending the generators to elements of $M$ without restrictions.

Now we show that for a finite set we can implement the universal property.

Since this is a universal property, all its implementations are canonically isomorphic.

**example** For $K$ a field we have the theory $\mathbf{Vec}_K$.

Consider finite set $\{x, y\}$; denote the corresponding projections on $\mathbf{Vec}_K[x, y] := \mathbf{Vec}_K[\{x, y\}]$ again by $x, y : \mathbf{Vec}_K[x, y]$.

Then elements of $\mathbf{Vec}_K[x, y]$ are vectors from $\{x, y\}$ to $\mathbb{1} := \{\bullet\}$ i.e. row vectors $[a\ b]$

so $x = [1\ 0]$, $y = [0\ 1] : \mathbf{Vec}_K[x, y]$,
& $a \cdot x + b \cdot y = [a\ b]$

So $\mathbf{Vec}_K[x, y]$ is a 2-dimensional vec space over $K$ with basis $x, y$. $\blacksquare$

Consider $\mathbb{T}[x, y] : \mathbb{T}\text{-}\mathbf{Set}$. Given $a, b : M : \mathbb{T}\text{-}\mathbf{Set}$, we have the interpretation map

$$[\![\cdot]\!]_M(a, b) : \mathbb{T}[x, y] \longrightarrow M. \qquad \mathbb{T}[$$

Especially note $[\![x]\!](a, b) = a$, $[\![y]\!](a, b) = a$,

that is generators are sent to the chosen elements in $M$.

This map is a homomorphism, we can illustrate this with $\omega : \overset{3}{\bullet} \longrightarrow \bullet : \mathbb{T}$, $u, t, s : \mathbb{T}[x, y] := \overset{2}{\bullet} \longrightarrow \bullet : \mathbb{T}$:

$$[\![\,[\![\omega]\!](u, t, s)]\!]_M(a, b) = [\![\omega\langle u, t, s \rangle]\!]_M(a, b)$$
$$= [\![\omega]\!]_M([\![\langle u, t, s \rangle]\!]_M(a, b))$$
$$= [\![\omega]\!]_M([\![u]\!]_M(a, b), [\![t]\!]_M(a, b), [\![s]\!]_M(a, b))$$

which is precisely the homomorphism condition for 3-ary operation

This interpretation map is precisely the extension of the map

$$\{x, y\} \longrightarrow M$$
$$x \rightsquigarrow a$$
$$y \rightsquigarrow b.$$

This construction generalises polynomial evaluation in a commutative ring i.e. given $a, b : R$ a ring, & a 2-variable polynomial $p : \mathbb{Z}[x, y]$, we can substitute $a, b$ for $x, y$ in $p$ and denote the result by $p(a, b) := [\![p]\!](a, b) : R$.

The $a, b$ substitution map

$$\mathbb{Z}[x, y] \longrightarrow R$$
$$p \rightsquigarrow p(a, b)$$

is a ring homomorphism.

**prop** For $a : \mathbf{Set}_{fin}$, denote the projections in $\overset{a}{\bullet} \longrightarrow \bullet : \mathbb{T}$ determined by $i : a$ as $x_i : \overset{a}{\bullet} \longrightarrow \bullet : \mathbb{T}$

Then $\mathbb{T}[a] : \mathbb{T}\text{-}\mathbf{Set}$ with $x : a \longrightarrow \mathbb{T}[a] : \mathbf{Set}$ implements the universal property of a free model of $\mathbb{T}$.

**proof** Let $M : \mathbb{T}\text{-}\mathbf{Set}$, $g : a \longrightarrow M : \mathbf{Set}$.

We wish to construct

$$\bar{g} : \mathbb{T}[a] \longrightarrow M : \mathbb{T}\text{-}\mathbf{Set}$$

such that $\bar{g}(x_i) = g(i)$.

Since $\bar{g}$ has to be a homomorphism, we need to have

$$\forall \omega : \overset{n}{\bullet} \longrightarrow \bullet : \mathbb{T}.$$
$$\bar{g}(\omega) = \bar{g}(\omega\langle x_i : i : a \rangle) \quad \text{since}$$
$$= [\![\omega]\!](\bar{g}(x_i : i : a))$$
$$= [\![\omega]\!]_M(\lambda i : a.\ \bar{g}(x_i))$$

so we are forced to define

$$\bar{g}(\omega) := [\![\omega]\!]_M(\lambda i : a.\ g(i))$$

it is not hard to see that $\bar{g}$ is a homomorphism $\blacksquare$

We can make $\mathbb{T}[\cdot]$ into a functor $\mathbb{T}[\cdot] : \mathbf{Set}_{fin} \longrightarrow \mathbf{Set}$

**define** Let $\alpha : a \longrightarrow b : \mathbf{Set}_{fin}$, $\omega : \overset{n}{\bullet} \to \bullet : \mathbb{T}$.

Then we get $\overset{n}{\alpha} : \overset{n}{\bullet} \longrightarrow \overset{n}{\bullet} : \mathbb{T}$ so we define

$$\alpha(\omega) := \omega^n \alpha : \overset{b}{\bullet} \longrightarrow \bullet : \mathbb{T}.$$

This is obviously functorial. $\blacksquare$

**example** Given $\omega : \overset{2}{\bullet} \longrightarrow \bullet : \mathbb{T}$, since we have the inclusion $\iota : \{2\} \hookrightarrow \{3\}$ we get $\iota(\omega) : \overset{3}{\bullet} \longrightarrow \bullet : \mathbb{T}$

$$\iota(\omega)(x, y, z) = \omega(x, y). \quad \blacksquare$$

Recall that models of $\mathbb{T}$ are just finite product preserving functors $M : \mathbb{T} \longrightarrow \mathbf{Set}$.
$a \rightsquigarrow M^a$

Note that for $a, b : \mathbf{Set}_{fin}$,

$$\mathbb{T}[a]^b = \overset{a}{\bullet} \longrightarrow \overset{b}{\bullet} : \mathbb{T}.$$

This makes free models on finite sets correspond to representable functors

$$\mathbb{T} \longrightarrow \mathbf{Set}$$
$$\overset{b}{\bullet} \rightsquigarrow \overset{a}{\bullet} \longrightarrow \overset{b}{\bullet} : \mathbb{T}$$

# Global Uniform Operations

Free models enable us to make progress with global uniform operations.

Let $\mathbb{T}$ be an algebraic theory.

**define** A global uniform operation of arity $\bullet^a \to \bullet^b$ is a map

$$\omega : \forall M : \mathbb{T} \circlearrowright Set,\ M^a \longrightarrow M^b : Set$$

where $\forall f : M \longrightarrow N : \mathbb{T} \circlearrowright Set,\ \forall m : M^a,\ f(\omega_M(m)) = \omega_N(f(m))$ ■

This is exactly a natural transformation

$$\omega : \|\cdot\|^a \Rightarrow \|\cdot\|^b : \mathbb{T} \circlearrowright Set \longrightarrow Set$$

**example** Each $\omega : \bullet^a \to \bullet^b : \mathbb{T}$ defines a global uniform operation – the interpretation of $\omega$ in all models

$$[\![\omega]\!]_\cdot : \forall M : \mathbb{T} \circlearrowright Set,\ M^a \longrightarrow M^b : Set$$

We can recover $\omega$ from $[\![\omega]\!]_\cdot$ by instantiation at $\mathbb{T}[a]$ –

$$[\![\omega]\!]_{\mathbb{T}[a]} : \mathbb{T}[a]^a \longrightarrow \mathbb{T}[a]^b;$$

let $x_\cdot : a \longrightarrow \mathbb{T}[a]$ be inclusion of generators,

then $[\![\omega]\!]_{\mathbb{T}[a]}(x_\cdot) : b \longrightarrow [\bullet^a \to \bullet : \mathbb{T}]$,

which is just

$$[\![\omega]\!]_{\mathbb{T}[a]}(id) : \bullet^a \to \bullet^b : \mathbb{T}$$

But interpretation of $\omega$ in $\mathbb{T}[a]$ is just composition of operations

$$[\![\omega]\!]_{\mathbb{T}[a]}(id) = \omega\, id = \omega \quad ■$$

Free models allow us to invert this process.

**prop** Let $\omega : \forall M : \mathbb{T} \circlearrowright Set,\ M^a \to M^b$ be a global uniform operation.

Then $\exists!\, \omega : \bullet^a \to \bullet^b : \mathbb{T}.$
$[\![\omega]\!] = \omega.$

**proof** Instantiate $\omega$ at $\mathbb{T}[a]$,

$$\omega_{\mathbb{T}[a]} : \mathbb{T}[a]^a \longrightarrow \mathbb{T}[a]^b;$$

applying the inclusion of generators $x_\cdot : a \longrightarrow \mathbb{T}[a]$, we get $\omega_{\mathbb{T}[a]}(x_\cdot) : b \longrightarrow \mathbb{T}[a]$

that is an operation

$$\langle \omega_{\mathbb{T}[a]}(x_\cdot) \rangle : \bullet^a \to \bullet^b : \mathbb{T}$$

If we had such an $\omega$, then as we have seen, we are forced to have

$$\omega = [\![\omega]\!]_{\mathbb{T}[a]}(x_\cdot) = \langle \omega_{\mathbb{T}[a]}(x_\cdot) \rangle$$

this shows uniqueness.

It remains to show that

$$[\![\langle \omega_{\mathbb{T}[a]}(x_\cdot) \rangle]\!] = \omega \text{ as global uniform operations}$$

Let $M : \mathbb{T} \circlearrowright Set,\ m_\cdot : M^a$

We have the following expressions

$$\omega_M : M^a \longrightarrow M^b : Set$$

$$\omega_{\mathbb{T}[a]}(x_\cdot) : \mathbb{T}[a]^b = b \longrightarrow \mathbb{T}[a] : Set$$

$$\langle \omega_{\mathbb{T}[a]}(x_\cdot) \rangle : \bullet^a \to \bullet^b : \mathbb{T}$$

$$\omega_{\mathbb{T}[a]} = [\![\langle \omega_{\mathbb{T}[a]}(x_\cdot) \rangle]\!]_{\mathbb{T}[a]} : \mathbb{T}[a]^a \longrightarrow \mathbb{T}[a]^b : Set$$

We need to show

$$[\![\langle \omega_{\mathbb{T}[a]}(x_\cdot) \rangle]\!]_M (m_\cdot) = \omega_M(m_\cdot)$$

Since $\mathbb{T}[a]$ is free with $x_\cdot : a \longrightarrow \mathbb{T}[a]$, we can lift $m_\cdot : a \longrightarrow M : Set$ to a homomorphism $\overline{m_\cdot} : \mathbb{T}[a] \longrightarrow M : \mathbb{T} \circlearrowright Set$ such that $\overline{m_\cdot}(x_\cdot) = m_\cdot$

Then

$$\begin{aligned}
[\![\langle \omega_{\mathbb{T}[a]}(x_\cdot) \rangle]\!]_M (m_\cdot) &= [\![\langle \omega_{\mathbb{T}[a]}(x_\cdot) \rangle]\!]_M (\overline{m_\cdot}(x_\cdot)) \\
&\overset{(1)}{=} \overline{m_\cdot}\left( [\![\langle \omega_{\mathbb{T}[a]}(x_\cdot) \rangle]\!]_{\mathbb{T}[a]} (x_\cdot) \right) \\
&= \overline{m_\cdot}\left( \omega_{\mathbb{T}[a]}(x_\cdot) \right) \\
&\overset{(2)}{=} \omega_M\left( \overline{m_\cdot}(x_\cdot) \right) \\
&= \omega_M(m_\cdot)
\end{aligned}$$

where (1) because $\overline{m_\cdot}$ is a homomorphism from $\mathbb{T}[a]$, so it interprets operations of $\mathbb{T}[a]$ in $M$ and evaluates them at $m_\cdot$.

(2) because $\omega$ is preserved by all homomorphisms. ■

Another way of seeing the correspondence is with Yoneda Lemma.

The Universal Property of $\mathbb{T}[a]$ & $\mathbb{T}[b]$ allows us to represent the functors

$$\|\cdot\|^a,\ \|\cdot\|^b : \mathbb{T} \circlearrowright Set \longrightarrow Set,$$

so

$$\|\cdot\|^a \Rightarrow \|\cdot\|^b \text{ global uniform operations}$$

$$\overset{(2)}{\cong} \quad \mathbb{T}[b] \longrightarrow \mathbb{T}[a] : \mathbb{T} \circlearrowright Set$$

$$\overset{(3)}{\cong} \quad b \longrightarrow \mathbb{T}[a] : Set$$

$$\bullet^a \to \bullet^b : \mathbb{T} \text{ operations}$$

where (2) is covariant Yoneda

(3) because $\mathbb{T}[b]$ is free.

Our proof basically explicitly computes the bijection above.

Composition of (2) with (1) deserves to be stated as a separate theorem.

**prop** For $a, b : Set_{fin}$, $\mathbb{T}$ an algebraic theory we have

$$\frac{\mathbb{T}[b] \longrightarrow \mathbb{T}[a] : \mathbb{T} \circlearrowright Set}{\bullet^a \to \bullet^b : \mathbb{T}} \quad ■$$

Note the contravariance; it generalizes the contravariance of trivial operations

$$\frac{b \longrightarrow a : Set}{\bullet^a \to \bullet^b : Set}$$

where Set is the trivial theory of sets.

We will encounter this again in a different context.

Note that the definition of global uniform operations makes direct use only of the category of models and the forgetful functor into Set, it doesn't invoke the algebraic theory directly.

We may be in a situation where we have a category $\mathbb{E}$ with a functor $\mathcal{U} : \mathbb{E} \longrightarrow Set$ which we suspect might be the category of models of some theory.

If that is so, then its operations must be the natural transformations

$$\omega : \forall M : \mathbb{E},\ F(M)^a \longrightarrow F(M)^b : Set,$$

in this way we can recover the algebraic theory $\mathbb{T}$ and the category $\mathbb{E}$ is canonically equivalent to $\mathbb{T} \circlearrowright Set$.

It may happen that $\mathbb{E}$ with $\mathcal{U}$ is not a category of models of any algebraic theory.

Yet, if $\mathcal{U}$ is sufficiently nice, we still could build some theory $\mathbb{T}$.

We can think of this process as extracting algebraic content of $\mathbb{E}$ with $\mathcal{U}$.

The resulting models will have some relationship to $\mathbb{E}$:

each object of $\mathbb{E}$ will give rise to a model

each morphism of $\mathbb{E}$ will give rise to a homomorphism.

This we will explore later with monadicity theorem.

# Free models over Set

So far we've implemented free models over finite sets; what about arbitrary sets?
In the case of a $\mathbf{Set}_{fin}$ we have operations of finite arity a

$$\overset{a}{\bullet}\longrightarrow\bullet:\mathbb{T}.$$

We'll attempt to develop a generalization to arbitrary arities:

$$\overset{X}{\bullet}\longrightarrow\bullet:\mathbb{T}\quad\text{for arbitrary } X:\mathbf{Set}$$

— a kind of possibly infinitary operation.

To make sense of this consider again vector spaces over $K$.

Let $V:\mathbf{vec}_K\subset\mathbf{Set}$. What would it mean to have an operation of arity $\aleph_0=\{0,1,2,\dots\}$?

Interpreting such $\omega:\overset{\aleph_0}{\bullet}\longrightarrow\bullet:\mathbf{vec}_K$, we have $[\![\omega]\!]:V^{\aleph_0}\longrightarrow V$.

Note that $V^{\aleph_0}$ is simply the space of all infinite sequences of vectors in $V$.

Consider such a sequence

$$(v_0,v_1,v_2,\dots):V^{\aleph_0}$$

Then $\omega$ might be a projection

$$(v_0,v_1,v_2,\dots)\rightsquigarrow v_3.$$

Or it can linearly combine a finite subsequence of vectors

$$(v_0,v_1,v_2,\dots)\rightsquigarrow 3\cdot v_0-v_3+2\cdot v_{40}$$

Does it make sense to consider infinite linear combinations such as

$$(v_0,v_1,v_2,\dots)\rightsquigarrow a_0\cdot v_0+a_1\cdot v_1+\dots\ ?$$

Not really; atleast not for all fields and not for all possible sequences of scalars $(a_0,a_1,\dots):K^{\aleph_0}$.

So for our notion of an algebraic theory to finitary — later we'll consider infinitary notions.

Hence in our use we'll make do with operations that take advantage only of a finite part of a sequence $(v_0,v_1,\dots)$.

We can represent $(v_0,v_1,v_2,\dots)\rightsquigarrow 3\cdot v_0-v_2+12\cdot v_3$ as a column vector $[3\ 0\ -1\ 12]:\bullet^{\aleph_0}\longrightarrow\bullet:\mathbf{vec}_K$

& $(v_0,v_1,v_2,\dots)\rightsquigarrow v_0-v_2$ as $[1\ 0\ -1]:\bullet^{\aleph_0}\longrightarrow\bullet:\mathbf{vec}_K$

So we can think of $\bullet^{\aleph_0}\longrightarrow\bullet:\mathbf{vec}_K$ as a sort of union of column vectors that are individually finite, but unbounded as a whole.

Care must be taken to not include duplicate operations such as $[1\ 0\ -1]$ & $[1\ 0\ -1\ 0]$; we can take care of this by requiring that the last entry be non-zero.

This trick happens to work for $\mathbf{vec}_K$, but for a general theory $\mathbb{T}$ & $X:\mathbf{Set}$, what we'll do is the directed union

$$\bigcup_{\substack{a\subseteq X\\ a\text{ finite}}}\mathbb{T}[a]$$

By this we mean the colimit of the diagram $\mathbb{P}_{fin}(X)\longrightarrow\mathbf{Set}$, $a\subseteq b\rightsquigarrow\mathbb{T}[a]\to\mathbb{T}[b]$.

In particular given $a\subseteq a'$, we'll identify $t:\mathbb{T}[a]$ with its image in $\mathbb{T}[b]$ which is also denoted $t:\mathbb{T}[b]$.

**define** Let $\mathbb{T}$ be an algebraic theory, $X:\mathbf{Set}$.
Define a set

$$\mathbb{T}[X]:=\bigcup_{\substack{a\subseteq X\\ a\text{ finite}}}\mathbb{T}[a]$$

Given $\omega:\overset{n}{\bullet}\longrightarrow\bullet:\mathbb{T}$, define its interpretation

$$\omega:\mathbb{T}[X]^n\longrightarrow\mathbb{T}[X]$$

$$\omega(t_1:\mathbb{T}[a_1],\dots,t_n:\mathbb{T}[a_n]):=\omega(t_1:\mathbb{T}[\bigcup_{i=1}^n a_i],\dots,t_n:\mathbb{T}[\bigcup_{i=1}^n a_i]).$$

This obviously makes $\mathbb{T}[X]$ into a $\mathbb{T}$-model. ■

We get a map $\eta:X\longrightarrow\mathbb{T}[X]:\mathbf{Set}$

$$x\rightsquigarrow x:\mathbb{T}[\{x\}]$$

**prop** Let $X:\mathbf{Set}$, then

$$\mathbb{T}[X]:\mathbb{T}\circlearrowright\mathbf{Set}\text{ with }\eta:X\longrightarrow\mathbb{T}[X]:\mathbf{Set}$$

implements the universal property of a free model of $\mathbb{T}$.

**proof** Let $M:\mathbb{T}\circlearrowright\mathbf{Set}$, $g:X\longrightarrow M:\mathbf{Set}$.
To define a map

$$\bar g:\mathbb{T}[X]\longrightarrow M:\mathbf{Set}$$

it is to define its restrictions to all a $\mathbb{P}_{fin}[X]$

$$\bar g|_{\mathbb{T}[a]}:\mathbb{T}[a]\longrightarrow M:\mathbf{Set}$$

such that they are compatible:
for $a\subseteq a':\mathbb{P}_{fin}[X]$, $t:\mathbb{T}[a]$ we have $\bar g|_{\mathbb{T}[a']}(t)=\bar g|_{\mathbb{T}[a]}(t)\cdot M$.
Fix a $\mathbb{P}_{fin}[X]$, then $g|_a:a\longrightarrow M:\mathbf{Set}$ lifts to a homomorphism

$$\bar g|_{\mathbb{T}[a]}:=\overline{g|_a}:\mathbb{T}[a]\longrightarrow M:\mathbf{Set}\circlearrowright\mathbf{Set}$$

Obviously such lifts are compatible.
Then for $\omega:\overset{n}{\bullet}\longrightarrow\bullet:\mathbb{T}$,

$$\begin{aligned}\bar g(\omega(t_1:\mathbb{T}[a_1],\dots,t_n:\mathbb{T}[a_n]))&=\bar g(\omega(t_1:\mathbb{T}[a_1],\dots,t_n:\mathbb{T}[a_n]))\text{ for some finite } a_1,\dots,a_n\subseteq X\\&=\overline{g|_{\bigcup_{i=1}^n a_i}}\left(\omega(t_1:\mathbb{T}[\bigcup_{i=1}^n a_i],\dots,t_n:\mathbb{T}[\bigcup_{i=1}^n a_i])\right)\\&=\omega\left(\overline{g|_{\bigcup_{i=1}^n a_i}}(t_1:\mathbb{T}[\bigcup_{i=1}^n a_i]),\dots,\overline{g|_{\bigcup_{i=1}^n a_i}}(t_n:\mathbb{T}[\bigcup_{i=1}^n a_i])\right)\\&=\omega(\bar g(t_1:\mathbb{T}[X]),\dots,\bar g(t_n:\mathbb{T}[X])),\end{aligned}$$

that is $\bar g$ is a $\mathbb{T}$-homomorphism

(uniqueness) Let $h:\mathbb{T}[X]\longrightarrow M:\mathbb{T}\circlearrowright\mathbf{Set}$ such that $h(\eta(x))=g(x)$.
Then $h$ and $\bar g$ agree on $\mathbb{T}[a]$ for all finite $a\subseteq X$ is $h=\bar g$:

$$\begin{tikzcd}a\arrow[r, "\eta"] \arrow[d, "g"'] & \mathbb{T}[a]\arrow[d, shift left, "h"] \arrow[d, shift right, "\bar g"'] \\ M \arrow[r, equal] & M\end{tikzcd}$$

First note that $\eta|_{\mathbb{T}[a]}:a\to\mathbb{T}[X]$ $a\subseteq X$ allowed by definition of $\mathbb{T}[X]$.

So $h|_{\mathbb{T}[a]}\circ\eta=h\circ\eta|_a$.

But $h|_{\mathbb{T}[a]}=\bar g|_{\mathbb{T}[a]}$ since $h|_{\mathbb{T}[a]}\circ\eta=h\circ\eta|_a=g|_a$. ■

We can think of going from $\mathbb{T}[a]$ to $\mathbb{T}[X]$ as a process of extension of a functor $F:\mathbf{Set}_{fin}\longrightarrow\mathbf{Set}$ to a functor $F:\mathbf{Set}\longrightarrow\mathbf{Set}$.

**define** Let $F:\mathbf{Set}_{fin}\longrightarrow\mathbf{Set}$ be a functor. Then

(1) for $X:\mathbf{Set}$ define

$$\bar F[X]:=\bigcup_{\substack{a\subseteq X\\ a\text{ finite}}}F[a]$$

(2) for $f:a\longrightarrow Y:\mathbf{Set}$, a $\mathbf{Set}_{fin}$ we obviously set $\bar F[f]:F[a]\longrightarrow\bar F[Y]$ since we can corestrict $f$ to its finite image in $Y$.

$$\begin{tikzcd}a\arrow[rr, "f"] \arrow[dr, "\iota_a"'] & & Y \\ & \operatorname{im} f \arrow[ur, hook, "\iota_Y"'] & \end{tikzcd}\rightsquigarrow\begin{tikzcd}F[a]\arrow[r, "F[f]"] \arrow[dr, "F[\iota_a]"'] & \bigcup_{\substack{b\subseteq Y\\ b\text{ finite}}}F[b] \\ & F[\operatorname{im} f] \arrow[u, hook, "\iota_Y"'] \end{tikzcd}$$

(3) for $f:X\longrightarrow Y:\mathbf{Set}$, to define $\bar F[f]:\bar F[X]\to\bar F[Y]:\mathbf{Set}$ we only need to define its restriction $\bar F[f]|_{F[a]}$ in a compatible way for all $a\subseteq X$ finite:
set $\bar F[f]|_{F[a]}:=\bar F[f|_a]$ which is obviously compatible.

(4) Given $X\xrightarrow{f}Y\xrightarrow{g}Z:\mathbf{Set}$, $a\subseteq X$ finite.
functoriality follows from $\bar F[g\circ f]|_{F[a]}=\bar F[g]\circ\bar F[f]|_{F[a]}$,
since

$$\begin{tikzcd}a\arrow[r, "f|_a"] \arrow[dr, "\iota_a"'] & Y\arrow[r, "g"] & Z \\ & \operatorname{im} f \arrow[u, hook, "\iota_Y"] \arrow[r, "f|_{\operatorname{im} f}"] & \operatorname{im} g \arrow[u, hook, "\iota_Z"'] \end{tikzcd}\rightsquigarrow\begin{tikzcd}F[a]\arrow[r, "F[f|_a]"] \arrow[dr, "F[\iota_a]"'] & \bigcup_{\substack{b\subseteq Y\\ b\text{ finite}}}F[b]\arrow[r, "F[g]"] & \bigcup_{\substack{c\subseteq Z\\ c\text{ finite}}}F[c] \\ & F[\operatorname{im} f] \arrow[u, hook, "\iota_Y"] \arrow[r, "F[f|_{\operatorname{im} f}]"] & F[\operatorname{im} g] \arrow[u, hook, "\iota_Z"'] \end{tikzcd}$$

and $F[g|_{\operatorname{im} f}]\circ F[f|_{\operatorname{im} f}]\circ F[\iota_a]=F[\iota_c]\circ F[g\circ f]$. ■

**define** In reverse, given functor $F:\mathbf{Set}\longrightarrow\mathbf{Set}$, we can restrict it to $F|_{\mathbb{P}_{fin}}:\mathbf{Set}_{fin}\longrightarrow\mathbf{Set}$, and get the finitarization of $F$

$$F_{fin}:=\bar F|_{\mathbb{P}_{fin}}:\mathbf{Set}\longrightarrow\mathbf{Set}.$$

Obviously we have a natural transformation

$$\varepsilon:F_{fin}\Longrightarrow F:\mathbf{Set}\longrightarrow\mathbf{Set}.$$

Since we have a canonical map

$$\varepsilon_X:\bigcup_{\substack{a\subseteq X\\ a\text{ finite}}}F[a]\longrightarrow F[X]:\mathbf{Set}\ ■$$

Note that given $F:\mathbf{Set}\to\mathbf{Set}$, $\mathcal{U}:\mathbf{Set}\longrightarrow\mathbf{Set}$, we have a correspondence

$$\frac{F\Longrightarrow\mathcal{U}:\mathbf{Set}\to\mathbf{Set}}{F\Longrightarrow\mathcal{U}|_{\mathbb{P}_{fin}}:\mathbf{Set}_{fin}\to\mathbf{Set}}$$

Functors that can be recovered in full from their action of finite sets are called finitary.

**define** Functor $F:\mathbf{Set}\longrightarrow\mathbf{Set}$ is finitary

if $\forall X:\mathbf{Set}$
the canonical map

$$\bigcup_{\substack{a\subseteq X\\ a\text{ finite}}}F[a]\longrightarrow F[X]:\mathbf{Set}$$

is invertible. ■

**example** $\mathbb{P}_{fin}:\mathbf{Set}\to\mathbf{Set}$ $\{X\rightsquigarrow\text{finite subsets of }X\}$ is the finitarization of the covariant powerset functor $\mathbb{P}:\mathbf{Set}\longrightarrow\mathbf{Set}$ we met. Obviously $\mathbb{P}$ is not finitary. ■

**example** For $F:\mathbf{Set}$, its finitarization is a finitary functor, so especially for $\mathbb{T}$ finitary algebraic theory, the functor $\mathbb{T}[\cdot]:\mathbf{Set}\longrightarrow\mathbf{Set}$ is finitary. ■

**prop** The map $\eta:\forall X:\mathbf{Set}$

$$X\longrightarrow\mathbb{T}[X]$$

is a natural transformation

**proof** Let $f:X\longrightarrow Y$, if $X$ & $Y$ are finite we already know that the naturality holds. Let $a\subseteq X$ finite.
Then

Naturality follows from the commut

1 Let $a\subseteq a\to b:\mathbf{Set}_{fin}$,
then

$$\begin{tikzcd}\mathbb{T}[a]\arrow[r, "\mathbb{T}[a]"] & \mathbb{T}[b] \\ a\arrow[u, "\eta"] \arrow[r, "a"] & b\arrow[u, "\eta"'] \end{tikzcd}$$

2 Let $X:\mathbf{Set}$, $a\subseteq X$ finite,
then

$$\begin{tikzcd}\mathbb{T}[a]\arrow[r, hook] & \bigcup_{\substack{a\subseteq X\\ a\text{ finite}}}\mathbb{T}[a] \\ a\arrow[u, "\eta"] \arrow[r, hook] & X\arrow[u, "\eta"'] \end{tikzcd}$$

Since given $f:X\longrightarrow Y:\mathbf{Set}$, $a\subseteq X$ finite, we have

$$\begin{tikzcd}\mathbb{T}[a]\arrow[r, "\mathbb{T}[f|_a]"] & \mathbb{T}[\operatorname{im} f]\arrow[r, hook] & \mathbb{T}[Y] \\ a\arrow[u, "\eta"] \arrow[r, "f|_a"] \arrow[rr, bend right, "f"'] & \operatorname{im} f\arrow[u, "\eta"] \arrow[r, hook] & Y\arrow[u, "\eta"'] \end{tikzcd}\quad■$$

# Limits & Colimits

Fix a finitary single sorted algebraic theory $\mathbb{T}$ and a model $M : \mathbb{T}\mathbf{Set}$.

## Internal Models (Submodels)

We have the notion of submodels (internal models) generalizing subgroups, subrings, subvectorspaces and so on.

**define** A submodel of $M$ consists of

$$S \subseteq M : \mathbf{Set}$$

where $\forall \omega : n \to_{\bullet} \mathbb{T}$,

$$\forall a_1, \dots, a_n : S.$$
$$\omega(a_1, \dots, a_n) : S. \blacksquare$$

**prop** submodels are closed under arbitrary subset intersections.

This makes $\mathrm{Sub}_{\mathbb{T}}[M] := \{\text{submodels of } M\}$ into a complete lattice. ■

**define** Since $\mathrm{Sub}_{\mathbb{T}}[M]$ is a complete lattice, we get a closure operator on subsets of $M$

$$\langle - \rangle : \mathcal{P}[M] \longrightarrow \mathcal{P}[M]$$

$\langle X \rangle$ is the least submodel containing the subset $X \subseteq M$. ■

Joins $S \vee T$ in $\mathrm{Sub}_{\mathbb{T}}[M]$ need not be subset unions.

**example** Take two differing 1-dimensional subspaces of $\mathbb{R}^2$ as a vector space over $\mathbb{R}$.
Then their union is not a subspace; the join is the least plane that contains them both. ■

This is reflected even externally – later will see that all limits of models are nice – they can be computed in $\mathbf{Set}$, but colimits are usually more complicated.

**prop** Let $f : M \to N : \mathbb{T}\mathbf{Set}$, $S \subseteq M$, $T \subseteq N$ submodels.
Then (1) $f^{-1}(T) \subseteq M$ is a submodel.
(2) $f(S) \subseteq N$ is a submodel.

**proof** Let $\omega : n \to_{\bullet} \mathbb{T}$. Then
(1) given $\{a_i : f^{-1}(T)\}_i^n$, we have
$$f(\omega(\{a_i\})) = \omega(\{f(a_i)\}) : T$$
(2) let $\{a_i : S\}_i^n$, then
$$f(\omega(\{a_i\})) = \omega(\{f(a_i)\}) : f(S)$$
and the fact that $f : S \to f(S)$ is a surjection, we get that $f(S)$ is a submodel of $N$. ■

**example** Given $f : M \to N : \mathbb{T}\mathbf{Set}$
$\mathrm{im}\, f \subseteq N$ is a submodel. ■

**prop** [Universal property of submodels]
Each internal submodel $S : \mathrm{Sub}_{\mathbb{T}}[M]$ defines the obvious external model $[S \subseteq M] : \mathbb{T}\mathbf{Set}$, together with the canonical inclusion $\iota_S : [S \subseteq M] \to M : \mathbb{T}\mathbf{Set}$ it has the following universal property.

$$\forall f : N \longrightarrow M : \mathbb{T}\mathbf{Set} \text{ where } \forall x : N.\ f(x) : S.$$
$$\exists!\ \text{construction } {}_S f : N \longrightarrow [S \subseteq M] : \mathbb{T}\mathbf{Set}.$$
$$\iota_S \circ {}_S f = f : M \to N : \mathbb{T}\mathbf{Set}$$

The homomorphisms $f : N \to M : \mathbb{T}\mathbf{Set}$ such that $\forall x : N.\ f(x) : S$ are said to lie in $S$.
The universal property states that we have a canonical correspondence

$$\frac{f : N \to M : \mathbb{T}\mathbf{Set} \text{ that lie in } S}{f : N \longrightarrow [S \subseteq M] : \mathbb{T}\mathbf{Set}} \blacksquare$$

**prop** $\mathrm{Sub}_{\mathbb{T}}[S \subseteq M] = \{T : \mathrm{Sub}_{\mathbb{T}}[M] \mid T \subseteq S\}$ as complete lattices. ■

This is quite obvious – we formulate this order so that the reader gets used to the distinction of internal vs. external.
From now on we won't distinguish notationally between $[S \subseteq M]$ and $S$.

## Internal quotients (Congruences)

As sets have the notion of an equivalence relation, so do models have congruences.

**define** A congruence on $M$ consists of an equivalence relation $\simeq$ on $M : \mathbf{Set}$ where

$$\forall \omega : n \to_{\bullet} \mathbb{T},$$
$$\forall a_1, a_1', \dots, a_n, a_n' : M.$$
$$\text{if } a_1 \simeq a_1', \dots, a_n \simeq a_n'$$
$$\text{then } \omega(a_1, \dots, a_n) \simeq \omega(a_1', \dots, a_n'). \blacksquare$$

Note that this condition is vacuously true for $\omega : 1 \to_{\bullet} \mathbb{T}$ constants.

**prop** congruences are closed under arbitrary equivalence relation intersections.

This makes $\mathrm{Cong}_{\mathbb{T}}[M] := \{\text{congruences on } M\}$ into a complete lattice. ■

We'll often denote congruences by $\simeq$, $\sim$ or $\equiv$, but sometimes it's better to think of them as internal quotients $Q, E$ – a kind of dual notion to internal models.
Denote the join of $Q, E$ in $\mathrm{Cong}(M)$ by $Q \cup E$.

**define** Since $\mathrm{Cong}_{\mathbb{T}}[M]$ is a complete lattice, we get a closure operator on relations of $M : \mathbf{Set}$

$$\langle - \rangle : \mathcal{P}[M \times M] \longrightarrow \mathcal{P}[M \times M]$$

$\langle R \rangle$ is the least congruence containing relation $R \subseteq M \times M$.
We'll often write $\langle a \sim b : M \mid a, b : M,\ aRb \rangle$ for $\langle R \rangle$. ■

**example** For the initial theory set congruences are just equivalence relations. ■

**example** For the theory $\mathbf{vec}_K$ of vector spaces over $K$ a congruence on $V : \mathbf{vec}_K\mathbf{Set}$ is an equivalence relation $\simeq$ on $V$ such that

$$v \simeq v' : V \Rightarrow k \cdot v \simeq k \cdot v' : V$$
$$v \simeq v',\ w \simeq w' : V \Rightarrow v + w \simeq v' + w' : V$$

Such a congruence defines a subspace of zero vectors according to $\star$

$$\mathrm{Zero}_{\simeq} := \{v : V \mid v \simeq 0\}.$$

This is a subspace:

$$0 : \mathrm{Zero}_{\simeq} \text{ since } 0 \simeq 0$$
$$v, w : \mathrm{Zero}_{\simeq} \Rightarrow v + w : \mathrm{Zero}_{\simeq} \text{ since } v \simeq 0,\ w \simeq 0 \Rightarrow v + w \simeq 0 + 0,$$
$$v : \mathrm{Zero}_{\simeq} \Rightarrow k \cdot v : \mathrm{Zero}_{\simeq} \text{ since } v \simeq 0 \Rightarrow k \cdot v \simeq k \cdot 0 = 0.$$

Conversely each subspace $S \subseteq V$ defines a congruence $v \simeq w$ iff $v - w : S$ so that $\mathrm{Zero}_{\simeq} = S$, that is we think of $S$ as vectors that we wish to identify with $0$.
In fact this gives a bijective correspondence between subspaces and congruences – this is the reason why in Linear Algebra people rarely speak of congruences. This works without change for the theory $\mathrm{mod}_R$ of $R$-modules over a ring $R$. ■

**example** A congruence on a ring $R$ is an equivalence relation $\simeq$ such that $r \simeq r',\ s \simeq s' : R \Rightarrow r + s \simeq r' + s' : R$, $r \cdot s \simeq r' \cdot s' : R$.
$r \simeq r' : R \Rightarrow -r \simeq -r' : R$.
It defines the zero set $\mathrm{Zero}_{\simeq} := \{r : R \mid r \simeq 0\}$.
This set $\mathrm{Zero}_{\simeq}$ is closed under $+, -, 0$, but as for multiplication we have

$$r : R,\ z : \mathrm{Zero}_{\simeq} \Rightarrow rz, zr : \mathrm{Zero}_{\simeq} \text{ since } z \simeq 0 \Rightarrow r \cdot z \simeq r \cdot 0 = 0$$
$$zr \simeq 0 \cdot r = 0.$$

This makes $\mathrm{Zero}_{\simeq}$ into an ideal of $R$, it mimicks the properties of $0$.

$$0 + 0 \simeq 0,\quad -0 = 0,$$
$$r \cdot 0 = 0 \simeq 0 \cdot r.$$

In fact again ideals are in a bijective correspondence with congruences. ■

**example** Similarly there is a canonical bijective correspondence between congruences of groups and normal subgroups. ■

**prop** Let $f : M \longrightarrow N : \mathbb{T}\mathbf{Set}$, $\simeq$ a congruence on $N$.
Then the equivalence relation $f^{-1}(\simeq)$

$$a \sim f(\simeq) a' : M \quad \text{iff} \quad f(a) \simeq f(a') : N$$

is a congruence called the inverse image (or pullback) of $\simeq : Q[N]$.

**proof** Let $\omega : n \to_{\bullet} \mathbb{T}$, and $\{a_i\}_i^n, \{a_i'\}_i^n : M$ such that $\forall i.\ f(a_i) \simeq f(a_i') : N$.
Then $\forall i.$

$$f(\omega(\{a_i\})) = \omega(\{f(a_i)\})$$
$$\simeq \omega(\{f(a_i')\})$$
$$= f(\omega(\{a_i'\})) \blacksquare$$

**example** As a special case we have the pullback of identity on $N$:

$$a \sim_{\ker f} a' : M \quad \text{iff} \quad f(a) = f(a') : N$$

called the kernel of $f$. ■

**prop** [Universal property of internal quotients]
Fix an internal quotient (congruence) $Q : \simeq : Q[M]$.
Then since $Q$ is an equivalence relation, we can form the quotient set $[M/Q] : [M/\simeq] : \mathbf{Set}$.
The congruence conditions on $\simeq$ are precisely the requirement that for all operations $\omega : n \to_{\bullet} \mathbb{T}$ the interpretation $\omega : M^n \to M$ is compatible with $\simeq$.
So we get the external quotient of $Q$ denoted $[M/Q]\ \mathbb{T}\mathbf{Set}$.
Then each $a : M$ gives the element $a : [M/Q]$, and

$$a \simeq b : M \quad \text{iff} \quad a = b : [M/Q].$$

This defines the quotient homomorphism $q_Q : M \longrightarrow [M/Q] : \mathbb{T}\mathbf{Set}$.
Then $[M/Q]\ \mathbb{T}\mathbf{Set}$ with $q_Q : M \longrightarrow [M/Q] : \mathbb{T}\mathbf{Set}$ has a universal property

$$\forall f : M \to N : \mathbb{T}\mathbf{Set} \text{ where } f \text{ respect } \simeq$$
$$\exists!\ \text{restriction } f|_Q : [M/Q] \longrightarrow N : \mathbb{T}\mathbf{Set}.$$
$$q_Q \circ f|_Q = f$$

The universal property states that we have a canonical correspondence

$$\frac{f : M \to N : \mathbb{T}\mathbf{Set} \text{ that respect } \simeq}{f : M/Q \longrightarrow N : \mathbb{T}\mathbf{Set}} \blacksquare$$

**prop** $Q_{\mathbb{T}}[M/Q] \simeq \{E : Q_{\mathbb{T}}[M] \mid Q \subseteq E\}$ as complete lattices. ■

From now on we'll just write $M/Q$.

Actually we can relate submodels and congruences.
For that we need a $\mathbb{T}\mathbf{Cat}$ structure on $M \times M$ which will be explored later. For now accept that for $\omega : n \to_{\bullet} \mathbb{T}$,

$$\omega((a_1, b_1), \dots, (a_n, b_n)) := \langle \omega(a_1, \dots, a_n), \omega(b_1, \dots, b_n) \rangle.$$

Then

**prop** $Q_{\mathbb{T}}[M] = \{E : \mathrm{Sub}_{\mathbb{T}}[M \times M] \mid E \text{ is an equivalence relation}\}$

**proof** (1) Let $Q = \simeq : Q_{\mathbb{T}}[M]$. Then $Q \subseteq M \times M$ is an equivalence relation, and $\langle a_1, a_1' \rangle, \dots, \langle a_n, a_n' \rangle : Q \Rightarrow a_1 \simeq a_1', \dots, a_n \simeq a_n'$

$$\Rightarrow \omega(a_1, \dots, a_n) \simeq \omega(a_1', \dots, a_n')$$
$$\Rightarrow \langle \omega(a_1, \dots, a_n), \omega(a_1', \dots, a_n') \rangle : Q$$

(2) I think now it is obvious. ■

## Morphisms

We have an important property of the forgetful functor $\mathbb{T}\mathbf{Set} \to \mathbf{Set}$: it reflects isomorphisms.
More directly

**prop** Let $f : M \longrightarrow N : \mathbb{T}\mathbf{Set}$, $g : N \to M : \mathbf{Set}$ where $gf = id_M$, $fg = id_N$.
Then $g$ is a homomorphism.
That is if $f$ is a bijection, then $f$ is isomorphism of models.

**proof**

$$g(\omega(\{a_j\})) = g(\omega(\{a_j\} \circ f \circ g(\{a_j\})))$$
$$= g(f(\omega(\{a_j\} \circ g(\{a_j\}))))$$
$$= \omega(\{a_j\} \circ g(\{a_j\})) \blacksquare$$

Each homomorphism has a canonical kernel/image factorization.

**prop** Let $f : M \longrightarrow N : \mathbb{T}\mathbf{Set}$.
Then we have the kernel congruence $\ker_f := \simeq_f : Q[M]$, and the image submodel $\mathrm{im}_f : \mathrm{Sub}[N]$, so we get the quotient map $q_f : M \longrightarrow [M/\ker_f] : \mathbb{T}\mathbf{Set}$, and the inclusion map $\iota_f : [\mathrm{im}_f \subseteq N] \hookrightarrow N : \mathbb{T}\mathbf{Set}$.
We have the canonical iso (since it is a bijection in $\mathbf{Set}$)

$$\mathrm{im}_f|_{\ker_f} : \mathrm{im}_f|_{\ker_f}(\{f\}_{\ker_f}) = \{f\}_{\ker_f} : [M/\ker_f] \xrightarrow{\sim} [\mathrm{im}_f \subseteq N]$$
$$a \longmapsto f(a)$$

So we have the factorization

$$\begin{tikzcd} M \arrow[r, "f"] \arrow[d, "q_f"'] & N \\ {[M/\ker_f]} \arrow[r, "\sim", "\mathrm{im}_f|_{\ker_f}"'] & {[\mathrm{im}_f \subseteq N]} \arrow[u, hook, "\iota_f"'] \end{tikzcd} \blacksquare$$

# products

**prop** Let $M, N : \mathbb{T}\text{-}\mathbf{Set}$, then the cartesian product of sets $M \times N : \mathbf{Set}$ has a canonical $\mathbb{T}$-model structure that makes it into cartesian product of $M$ and $N$ in $\mathbb{T}\text{-}\mathbf{Set}$. ■

**example** Take $\omega : \cdot^2 \longrightarrow \cdot : \mathbb{T}$, $a, a' : M$, $b, b' : N : \mathbb{T}\text{-}\mathbf{Set}$ then

$$\omega(\langle a, b\rangle, \langle a', b'\rangle) := \langle \omega(a, a'), \omega(b, b')\rangle \ ■$$

We prove it more generally for arbitrary families of models.

**prop** Let $X : \mathbf{Set}$, $M_\cdot : X \longrightarrow \mathbb{T}\text{-}\mathbf{Set}$ be a family of models.

Then the set $\forall x{:}X.\, M_x : \mathbf{Set}$ has canonical $\mathbb{T}$-model structure making it into the cartesian product of the family $M_\cdot$.

**proof** Let $\omega : \cdot^n \longrightarrow \cdot : \mathbb{T}$, $s^1, \dots, s^n : \forall x{:}X.\, M_x$.

Set $\omega(s^1, \dots, s^n) : \forall x{:}X.\, M_x$,

$$\omega(s^1, \dots, s^n)_x := \omega(s^1_x, \dots, s^n_x) : M_x.$$

This defines a $\mathbb{T}$-model structure on $\forall x{:}X.\, M_x$.

Let $f_\cdot : \forall x{:}X.$
$$N \longrightarrow M_x : \mathbb{T}\text{-}\mathbf{Set}$$

Then in $\mathbf{Set}$ we have unique extension of $f_\cdot$ to

$$\langle f\rangle : N \longrightarrow \forall x{:}X.\, M_x : \mathbf{Set}$$

To show cartesian product in $\mathbb{T}\text{-}\mathbf{Mod}$ we only need to check that $\langle f\rangle$ is a homomorphism.

And indeed $\forall \omega : \cdot^k \longrightarrow \cdot : \mathbb{T}$, $b_1, \dots, b_k : N$, $x : X$.

$$\begin{aligned}\langle f\rangle(\omega(b_1, \dots, b_k))_x &= f_x(\omega(b_1, \dots, b_k)) \\ &= \omega(f_x(b_1), \dots, f_x(b_k)) \\ &= \omega(\langle f\rangle(b_1)_x, \dots, \langle f\rangle(b_k)_x) \\ &= \omega(\langle f\rangle(b_1), \dots, \langle f\rangle(b_k))_x \ ■\end{aligned}$$

As a special case [coming from the constant family], $M^X : \mathbb{T}\text{-}\mathbf{Set}$ where $X : \mathbf{Set}$, $M : \mathbb{T}\text{-}\mathbf{Set}$.

# equalizers

**prop** Let $f, g : M \longrightarrow N : \mathbf{Set}$, then the subset

$$\{a : M \mid f(a) = g(a)\}$$ is a submodel of $M$.

Then $\{\langle a\rangle \mid f(a) = g(a)\} \subset M$ with the obvious inclusion is the equalizer of $f$ and $g$.

**proof** Let $f(a_1) = g(a_1), \dots, f(a_n) = g(a_n)$.

Then
$$\begin{aligned}f(\omega(a_1, \dots, a_n)) &= \omega(f(a_1), \dots, f(a_n)) \\ &= \omega(g(a_1), \dots, g(a_n)) \\ &= g(\omega(a_1, \dots, a_n)). \ ■\end{aligned}$$

So

**prop** $\mathbb{T}\text{-}\mathbf{Set}$ has all limits and the forgetful functor preserves them. ■

# coequalizers

**prop** Let $f, g : M \longrightarrow N : \mathbf{Set}$, consider the congruence

$$\langle f(a) = g(a), M \mid a : M\rangle : Q[N].$$

Then $\dfrac{N}{\langle f(a) = g(a) \mid a : M\rangle}$ with the obvious quotient map of $N$ is the coequalizer of $f$ and $g$.

It is important to note that the $\mathbb{T}$-congruence $\langle f(a) = g(a), M \mid a : M\rangle : Q[N : \mathbb{T}\text{-}\mathbf{Set}]$ may be bigger than the equivalence relation $\{f(a) = g(a), M \mid a : M\} : Q[N : \mathbf{Set}]$, so the forgetful functor $\mathbb{T}\text{-}\mathbf{Set} \longrightarrow \mathbf{Set}$ need not preserve all coequalizers. ■

Are there coequalizers that are preserved?

**example** Every external quotient is such a coequalizer.

Let $Q := Q[M]$. Then we have the model $[Q \subset M \times M]$ with restrictions of projections

$$\ell_Q, r_Q : [Q \subset M \times M] \longrightarrow M : \mathbb{T}\text{-}\mathbf{Set}$$
$$\ell_Q(a, b) := a, \quad r_Q(a, b) := b$$

Then the coequalizer of $\ell_Q, r_Q$ in $\mathbf{Set}$ is

$$\{a = b \mid a, b : M, a \sim b\} : Q[M : \mathbf{Set}]$$

but this is just the congruence $\sim$.

So the forgetful functor preserves coequalizers coming from congruences in this way. ■

Now we'll try to attempt a proof that all coequalizers are preserved, and see where it fails.

Recall the following about equivalence relations on $\mathbf{Set}$.

**prop** Let $X : \mathbf{Set}$, $R \in X \times X$ a relation.

Denote $\sim := \{a \sim a' \mid a, a' : X, aRa'\} : Q[X]$.

Then $\forall a, a' : X.$
$$a \sim a' \iff \exists n : \mathbb{N}.\ \exists a_1, \dots, a_n : X.\ a = a_1, a_1 R a_2, \dots, a_{n-1} R a_n, a_n = a'$$

We'll visualize the situation as

$$a = a_1 \overset{R}{\sim} a_2 \overset{R}{\sim} \cdots \overset{R}{\sim} a_{n-1} \overset{R}{\sim} a_n = a'$$

and write $a \sim a_1 R a_2 \cdots a_{n-1} R a_n \sim a'$; we call such sequence of relationships an $R$-chain. ■

Let $f, g : M \longrightarrow N : \mathbb{T}\text{-}\mathbf{Set}$, define $\simeq := \{f(a) = g(a) \mid a : M\} : Q[N : \mathbf{Set}]$.

We'll try to see if $\simeq$ is a congruence. If it is, it should respect $\omega : \cdot^2 \longrightarrow \cdot : \mathbb{T}$.

Let $x \simeq x'$, $y \simeq y' : N$. Is it the case that $\omega(x, y) \simeq \omega(x', y')$?

Since $x \simeq x'$, we have $n : \mathbb{N}$, $x = x_1, x_2, \dots, x_n = x' : N$ such that we have $a_1, \dots, a_{n-1} : M$ where $f(a_i) = x_i$, $g(a_i) = x_{i+1}$. For the sake of concreteness we assume $n = 3$, and visualize the situation as a chain

$$x = x_1 \overset{f}{\sim} a_1 \overset{g}{\sim} x_2 \overset{f}{\sim} a_2 \overset{g}{\sim} x_3 = x'$$

Similarly for $y \simeq y'$ assume we have

$$y = y_1 \overset{f}{\sim} b_1 \overset{g}{\sim} y_2 \overset{f}{\sim} b_2 \overset{g}{\sim} y_3 = y'.$$

Then since $f$ and $g$ are homomorphisms, we have

$$\omega(x, y) \overset{f}{\sim} \omega(a_1, b_1) \overset{g}{\sim} \omega(x_2, y_2) \overset{f}{\sim} \omega(a_2, b_2) \overset{g}{\sim} \omega(x', y')$$

so $\omega(x, y) \simeq \omega(x', y')$. Wait, what? So is $\simeq$ a congruence after all?

No; we have made an unwarranted assumption that the chain coming from $x \simeq x'$ has the same length as the chain from $y \simeq y'$.

This need not be the case in general; fortunately we can easily augment our assumption so that we can always find chains of the same length.

**prop** Let $R \subseteq X \times X$ be a reflexive relation on $X : \mathbf{Set}$.

Then given $a \simeq a'$, $b \simeq b' : X$,
$$\exists k : \mathbb{N}.\ \exists R\text{-chains of length } k$$
$$a = a_1 R a_2 \cdots a_{n-1} R a_n = a'$$
$$b = b_1 R b_2 \cdots b_{n-1} R b_n = b'$$

Similarly for a finite sequence $a^1 \simeq a^2$, $a^2 \simeq a^3$, $\dots$, $a^n \simeq a^n$:

**proof** Just put out the smaller chain with

$$x \overset{R}{\sim} x \overset{R}{\sim} \cdots \overset{R}{\sim} x \overset{R}{\sim} x \quad \text{where } x = a' \text{ or } x = b'$$

since $xRx$ by reflexivity of $R$. ■

**define** A relation $R \subseteq M \times M$ is a $\mathbb{T}$-relation
$$:\iff \forall \omega : \cdot^n \to \cdot : \mathbb{T},\ a_1 R a_1', \dots, a_n R a_n' \Rightarrow \omega(a_1, \dots, a_n)\, R\, \omega(a_1', \dots, a_n')$$
$\iff R$ is a submodel of $M \times M$. ■

**example** Let $f, g : M \longrightarrow N : \mathbb{T}\text{-}\mathbf{Set}$. Define a relation $R$ on $N$
$$xRx' :\iff \exists a : M.\ x = f(a),\ g(a) = x'$$

Then $R$ is a $\mathbb{T}$-relation; we'll illustrate it for operations of arity 2 $\omega : \cdot^2 \longrightarrow \cdot : \mathbb{T}$.

Let $xRx'$, $yRy'$ that is we have $a, b : M$ such that $x = f(a)$, $g(a) = x'$ and $y = f(b)$, $g(b) = y'$.

Then
$$\begin{aligned}\omega(x, y) &= \omega(f(a), f(b)) \\ &= f(\omega(a, b)),\end{aligned}$$
and similarly $\omega(x', y') = g(\omega(a, b))$.

So $\omega(x, y) R \omega(x', y')$. ■

**prop** Let $R \in M \times M$ be a reflexive $\mathbb{T}$-relation.

Then the equivalence relation
$$\sim := \{a \sim a' \mid a, a' : M, aRa'\} : Q[M : \mathbf{Set}]$$
is a congruence.

**proof** Let $a_1 \sim a_1', \dots, a_n \sim a_n' : M$, $\omega : \cdot^n \to \cdot : \mathbb{T}$.

Then we have some $k : \mathbb{N}$ and $R$-chains of length $k$
$$a_1 = a_{1,1} R a_{1,2} R \dots R\ a_{1,k} = a_1'$$
$$a_2 = a_{2,1} R a_{2,2} R \dots R\ a_{2,k} = a_2'$$
$$\vdots$$
$$a_n = a_{n,1} R a_{n,2} R \dots R\ a_{n,k} = a_n'.$$

They have the same length because $R$ is reflexive.

Then since $R$ is a $\mathbb{T}$-relation, we get an $R$-chain
$$\begin{aligned}\omega(a_1, a_2, \dots, a_n) &= \omega(a_{1,1}, a_{2,1}, \dots, a_{n,1}) \\ &R\ \omega(a_{1,2}, a_{2,2}, \dots, a_{n,2}) \\ &\vdots \\ &R\ \omega(a_{1,k}, a_{2,k}, \dots, a_{n,k}) \\ &= \omega(a_1', \dots, a_n').\end{aligned}$$

So $\omega(a_1, a_2, \dots, a_n) \sim \omega(a_1', \dots, a_n')$. ■

Let $f, g : M \longrightarrow N : \mathbb{T}\text{-}\mathbf{Set}$.

We've already seen that the generating relation of the equivalence $\simeq := \{f(a) = g(a), M \mid a : M\} : Q[N : \mathbf{Set}]$ is a $\mathbb{T}$-relation. To make it reflexive, we assume that we have a map $\mathrm{refl} : N \longrightarrow M : \mathbf{Set}$ such that
$$f \circ \mathrm{refl} = \mathrm{id}_N = g \circ \mathrm{refl}.$$

Then $x : N$ is in a relation with itself since
$$x = f(\mathrm{refl}(x)) \text{ and } g(\mathrm{refl}(x)) = x$$

So $\simeq$ is a congruence.

Note that $\mathrm{refl}$ need not even be a homomorphism.

When the pair $f, g$ have such a common section $\mathrm{refl}$, the coequalizer of $f, g$ is called reflexive coequalizer.

**prop** The forgetful functor $\mathbb{T}\text{-}\mathbf{Set} \longrightarrow \mathbf{Set}$ preserves reflexive coequalizers ■

**example** Note that given a congruence $Q : Q[M]$, the map $\ell, r : [Q \subset M \times M] \longrightarrow M : \mathbb{T}\text{-}\mathbf{Set}$ have a common section $\mathrm{refl}_Q : M \longrightarrow [Q \subset M \times M] : \mathbf{Set}$
$$\mathrm{refl}_Q(a) := \langle a, a\rangle$$

It's even a homomorphism. ■

More generally we will consider a situation where we have a functor
$$\mathcal{U} : \mathbb{E} \longrightarrow \mathbf{Set}$$
where $\mathbb{E}$ is to be thought of as a category of models of some algebraic theory and $\mathcal{U}$ as the forgetful functor.

Now since $\mathbb{E}$ is a general category, it is harder to talk about congruences and quotients. So instead we talk about coequalizers. To make $\mathcal{U}$ seem more like a forgetful functor we'll assume that given a reflexive pair over $\mathcal{U}$ — that is a pair of maps $f, g : a \to b : \mathbb{E}$, and a map $r : \mathcal{U}(b) \to \mathcal{U}(a)$ such that $\mathcal{U}(f) \circ r = \mathrm{id}_b = \mathcal{U}(g) \circ r$ — the category $\mathbb{E}$ has the coequalizer of all such reflexive pairs, and $\mathcal{U}$ preserves them.

**Todo**

you forgot symmetry

## Directed Unions

In general computing arbitrary joins of submodels is hard — it need not be a union of subsets.

As an example take two subspaces $S_1, S_2 \subseteq V$ of a vector space.

Their subset union $S_1 \cup S_2$ need not be a subspace — take for example two independent lines.

But if we have an increasing sequence of subspaces

$$S_0 \subseteq S_1 \subseteq S_2 \subseteq \dots$$

then their join is a subset union $\bigcup_{k \in \mathbb{N}} S_k$

This holds in general.

**prop** Let $S_0 \subseteq S_1 \subseteq \dots : \mathrm{Sub}_{\uparrow}[M]$.

Then $\bigcup_{k \in \mathbb{N}} S_k : \mathrm{Sub}_{\uparrow}[M]$

**proof** Let $\omega : {}^{n}\!\to\!\bullet : \mathbb{T}$, $a_1, \dots, a_n : \bigcup_{k \in \mathbb{N}} S_k$.

Then $\exists k_1, \dots, k_n : \mathbb{N}$

$$a_1 : S_{k_1}, \dots, a_n : S_{k_n}.$$

Take $k_{max} := \max(k_1, \dots, k_n)$.

Then $a_1, \dots, a_n : S_{k_{max}}$.

Hence $\omega(a_1, \dots, a_n) : S_{k_{max}} \subseteq \bigcup_{k \in \mathbb{N}} S_k$ ∎

We can abstract many things from the case of $\omega$.

First of all a monotone sequence $S_0 \subseteq S_1 \subseteq \dots : \mathrm{Sub}_{\uparrow}[M]$ is simply a monotone map $S : \omega \to \mathrm{Sub}_{\uparrow}[M]$.

Then we made use of least upper bound (maximum in this case) of finite number of elements in $\omega$. Actually we don't even need the least in the least upper bound — its enough to have some upper bound.

We abstract this into the notion of a directed set.

**define** A directed set consists of

$D : \mathrm{Set}$ with preorder $\sqsubseteq$ [reflexive & transitive]

$0 : D$

$+ : D \times D \to D$

where (1) $\forall i : D$
$0 \sqsubseteq i$

(2) $\forall i, j : D$
$i, j \sqsubseteq i + j$

This guarantees that any finite set has some upper bound.

**prop** Let $S : D \to \mathrm{Sub}_{\uparrow}[M]$ be a monotone functor, $D$ a directed set.

Then $\bigcup_{k \in D} S_k : \mathrm{Sub}_{\uparrow}[M]$

**proof** Let $\omega : {}^{n}\!\to\!\bullet : \mathbb{T}$, $a_1, \dots, a_n : \bigcup_{k \in D} S_k$.

Then $\exists k_1, \dots, k_n : D$.

$$a_1 : S_{k_1}, \dots, a_n : S_{k_n}.$$

Take $k_{upper} : D$ some upper bound of $k_1, \dots, k_n$.

Then $a_1, \dots, a_n : S_{k_{upper}}$ since $k_i \sqsubseteq k_{upper}$ so $a_i : S_{k_i} \subseteq S_{k_{upper}}$.

Hence $\omega(a_1, \dots, a_n) : S_{k_{upper}} \subseteq \bigcup_{k \in D} S_k$ ∎

Our considerations depend on finitarity of our operations.

If our notion of algebraic theory allowed for infinitary operations, directed sets would not be enough. Take $S : \mathbb{N} \to \mathrm{Sub}_{\uparrow}[M]$.

& suppose we have an infinitary operation $\omega : {}^{\mathbb{N}}\!\to\!\bullet : \mathbb{T}$.

Given $a_0, a_1, \dots : \bigcup_{k \in \mathbb{N}} S_k$, then $a_0 : S_{k_0}, a_1 : S_{k_1}, \dots$.

Then there is no gurantee that we can have an upper bound of $k_0, k_1, \dots : \mathbb{N}$.

So we can't push through our proof. We'd need to replace directed sets with something stronger.

This simplification of directed joins of submodels is phased out on the external level with special kinds of colimits.

**prop** Let $M : D \to \mathbb{T}\mathcal{C}\mathrm{Set}$ be a family, $D$ a directed set.

Then the colimit of $M$ in $\mathrm{Set}$

$$\bigcup_{i \in D} M_i$$

is a $\mathbb{T}$-model and is also a colimit in $\mathbb{T}\mathcal{C}\mathrm{Set}$.

**proof** (1) Let $\omega : {}^{n}\!\to\!\bullet : \mathbb{T}$, define the interpretation

$$\omega(a_1, M_{k_1}, \dots, a_n, M_{k_n}) := \omega(a_1, M_{k}, \dots, a_n, M_{k}) \text{ where } k := k_1 + \dots + k_n.$$

This is computable, and defines a $\mathbb{T}$-model $\bigcup_{i \in D} M_i : \mathbb{T}\mathcal{C}\mathrm{Set}$.

(2) Its enought to show that given a compatible family

$$f_i : \forall i : D \quad M_i \to N \quad \text{of homomorphisms,}$$

the colimit $\bar{f} : \bigcup_{i \in D} M_i \to N : \mathrm{Set}$ is a homomorphism.

Note

$$\begin{aligned} \bar{f}\Big(\omega\big(a_1, \bigcup_{i \in D} M_i, \dots, a_n, \bigcup_{i \in D} M_i\big)\Big) &= \bar{f}_k\big(\omega(a_1, M_k, \dots, a_n, M_k)\big) \\ &= \omega\big(\bar{f}_k(a_1), \dots, \bar{f}_k(a_n)\big) \\ &= \omega\big(\bar{f}(a_1), \dots, \bar{f}(a_n)\big) \end{aligned}$$ ∎

### Coproducts of free models

We will see later that $\mathbb{T}\mathcal{C}\mathrm{Set}$ has arbitrary coproducts, but the underlying set of a coproduct of models need not be a coproduct of the underlying sets. Now we only show the simple case of coproducts of free models — we leave the general case for later after we have developed some more sophisticated tools.

**prop** (1) Let $X, Y : \mathrm{Set}$. Then the coproduct of $\mathbb{T}[X]$ and $\mathbb{T}[Y]$ is $\mathbb{T}[X] + \mathbb{T}[Y] := \mathbb{T}[X+Y]$ with $\mathbb{T}[i_X] : \mathbb{T}[X] \to \mathbb{T}[X+Y]$
$\mathbb{T}[i_Y] : \mathbb{T}[Y] \to \mathbb{T}[X+Y]$

(2) Let $X : A \to \mathrm{Set}$ be a family of sets where $A : \mathrm{Set}$.

Then the coproduct of the family of free models $\mathbb{T}[X] : A \to \mathbb{T}\mathcal{C}\mathrm{Set}$ is just $\mathbb{T}\big[\sum_{a : A} X_a\big]$ with $\mathbb{T}[i_a] : \forall a : A$
$\mathbb{T}[X_a] \to \mathbb{T}\big[\sum_{a : A} X_a\big]$.

**proof** (1) Let $f : \mathbb{T}[X] \to S$, $g : \mathbb{T}[Y] \to S : \mathbb{T}\mathcal{C}\mathrm{Set}$.

Then we have the mappings $f\eta : X \to S$, $g\eta : Y \to S : \mathrm{Set}$,

hence $f\eta + g\eta : X + Y \to S : \mathrm{Set}$, let

$$f + g := \overline{f\eta + g\eta} : \mathbb{T}[X+Y] \to S : \mathbb{T}\mathcal{C}\mathrm{Set}.$$

Then $(f + g)\mathbb{T}[i_X] = f$, since $(f + g)\mathbb{T}[i_X]\eta = \overline{f\eta + g\eta}\, \eta\, i_X$

$$= (f\eta + g\eta)\, i_X$$

$$= f\eta$$

Similarly $(f + g)\mathbb{T}[i_Y] = g$.

For uniqueness let $h : \mathbb{T}[X+Y] \to S$ where $h\,\mathbb{T}[i_X] = f$, $h\,\mathbb{T}[i_Y] = g$.

Then $h\eta = f\eta + g\eta$ since $h\eta\, i_X = h\,\mathbb{T}[i_X]\eta$ and $h\eta\, i_Y = g\eta$.

So $h = \overline{f\eta + g\eta} = f + g$ $\qquad = f\eta$.

(2) The proof is no different from (1). ∎

## Canonical Presentation

**prop** (1) Let $X \subseteq M : \mathrm{Set}$ then we have the left of the inclusion

$$\overline{X \subseteq M} : \mathbb{T}[X] \to M : \mathbb{T}\mathcal{C}\mathrm{Set}$$

Its image is

$$\text{in } \overline{X \subseteq M} = \langle X \rangle : \mathrm{Sub}[M]$$

(2) As a special case denote the left of the identity $id_M : M \to M : \mathrm{Set}$ to

$$\varepsilon_M := \overline{id_M} : \mathbb{T}[M] \to M : \mathbb{T}\mathcal{C}\mathrm{Set}.$$

Then $M \cong \mathbb{T}[M] / \ker_{\varepsilon_M}$ that is every model is a quotient of a free model ∎

# Theories as First Class Mathematical Objects

**define** Let $\mathbb{T}, \mathbb{U}$ be single-sorted algebraic theories.

Then a morphism $f$ from $\mathbb{T}$ to $\mathbb{U}$ of algebraic theories is given by

$$h(\cdot) : \forall a, b : \mathrm{Set}_{fin}.$$
$$\left[\bullet \xrightarrow{a} \bullet : \mathbb{U}\right] \longrightarrow \left[\bullet \xrightarrow{a} \bullet : \mathbb{T}\right]$$

where (functoriality) $\forall \bullet \xrightarrow{\omega} \bullet \xrightarrow{\eta} \bullet : \mathbb{T}$,

$$h(\eta \cdot \omega) = h(\eta) \cdot h(\omega) : \bullet \xrightarrow{a} \bullet \xrightarrow{c} \bullet : \mathbb{U}$$

(cartesianness) $\forall a \xrightarrow{\alpha} b : \mathrm{Set}_{fin}$.

$$h({}^{b}\alpha) = {}^{a}\alpha : \bullet \xrightarrow{a} \bullet \xrightarrow{b} \bullet : \mathbb{U} \quad \blacksquare$$

**prop** (i) $h(\mathrm{id}_a) = \mathrm{id}_a : \bullet \xrightarrow{a} \bullet : \mathbb{U}$

(ii) $h(\mathbb{T}_a) = \mathbb{T}_a : \bullet \xrightarrow{a+b} \bullet : \mathbb{U}$

$h(\mathbb{T}_b) = \mathbb{T}_b : \bullet \xrightarrow{a+b} \bullet : \mathbb{U}$

(iii) Let $\omega : \bullet \xrightarrow{a} \bullet : \mathbb{T}$, $\eta : \bullet \xrightarrow{c} \bullet : \mathbb{T}$, then

$$h\langle \omega, \eta \rangle = \langle h(\omega), h(\eta) \rangle$$

**proof** (i) & (ii) are implied by cartesianness

(iii) Note $\mathbb{T}_a h\langle \omega, \eta \rangle = h(\mathbb{T}_a) h\langle \omega, \eta \rangle$

$= h(\mathbb{T}_a \langle \omega, \eta \rangle)$

$= h(\omega)$,

similarly $\mathbb{T}_b h\langle \omega, \eta \rangle = h(\eta)$. $\blacksquare$

By (iii) to specify $h$ it is enough to give its action on operations of arity $\bullet \xrightarrow{a} \bullet$.

Informally, morphisms of theories are functors

$$\begin{tikzcd} \mathbb{T} \arrow[rr, "h"] \arrow[dr, "\mathrm{Set}^{op}_{fin}"'] & & \mathbb{U} \\ & \mathrm{Set}^{op}_{fin} \arrow[ur, "\mathbb{T}"'] & \end{tikzcd}$$

Since they are functors, we can define 2-morphisms

**define** Let $\mathbb{T} \underset{h'}{\overset{h}{\rightrightarrows}} \mathbb{U}$ be morphisms of theories.

Then a 2-morphism $\alpha$ from $h$ to $h'$ is given by

$$\alpha(\cdot) : \forall a, b : \mathrm{Set}_{fin}.$$
$$\left[\bullet \xrightarrow{a} \bullet : \mathbb{T}\right] \longrightarrow \left[\bullet \xrightarrow{a} \bullet : \mathbb{U}\right]$$

where $\forall \omega : \bullet \xrightarrow{a} \bullet : \mathbb{T}$,

$$\begin{tikzcd} \bullet \arrow[r, "h(\omega)"] \arrow[d, "\alpha(\mathrm{id}_a)"'] & \bullet \arrow[d, "\alpha(\mathrm{id}_b)"] \\ \bullet \arrow[r, "h'(\omega)"'] \arrow[ur, phantom, "\alpha(\omega)"] & \bullet \end{tikzcd} : \mathbb{U} \quad \blacksquare$$

These 2-morphisms are precisely the natural transformations $\alpha : f \to g$.

so it is enough to specify their action on identities of $\mathbb{T}$

$$\alpha : \forall a \, \mathrm{Set}_{fin}.$$
$$(\bullet \xrightarrow{a} \bullet) : \mathbb{U}$$
$$\alpha_a = \alpha(\mathrm{id}_a).$$

We can think of $\alpha$ as a deformation of operations given by action of $h$ to operations given by action of $h'$.

In this way we have a 2-category

$$\mathbf{Law}$$

of single-sorted algebraic theories named in honor of William Lawvere.

Fix $h : \mathbb{T} \longrightarrow \mathbb{U} : \mathbf{Law}$

This defines a functor forgetful functor

**define** $\mathcal{F} : \mathbb{U}\mathbf{CSet} \longrightarrow \mathbb{T}\mathbf{CSet}$

$(\mathcal{N})\mathcal{F} = \mathcal{N}$ where $\forall \omega : \bullet \xrightarrow{a} \cdots \bullet : \mathbb{T}$, $b_1, \ldots, b_n : \mathcal{N}$

$$\omega(b_1, \ldots, b_n) = h(\omega)(b_1, \ldots, b_n)$$

This obviously defines a functor. $\blacksquare$

Given $\mathcal{N} : \mathbb{U}\mathbf{CSet}$, we'll often just write $\mathcal{N}_{\mathcal{F}} : \mathbb{T}\mathbf{CSet}$ instead of $(\mathcal{N})\mathcal{F}$.

We can formulate the universal property of free models of $\mathbb{U}$ relative to $h$.

**define** Let $M : \mathbb{T}\mathbf{CSet}$, then a free $\mathbb{U}$-model relative to $h$ consists of

$$h(M) : \mathbb{U}\mathbf{CSet} \text{ with } \eta : M \longrightarrow h(M)_{\mathcal{F}} : \mathbb{T}\mathbf{CSet}$$

where $\forall \mathcal{N} : \mathbb{U}\mathbf{CSet}$

$$\forall f : M \longrightarrow \mathcal{N}_{\mathcal{F}} : \mathbb{T}\mathbf{CSet}$$
$$\exists ! \, \bar{f} : h(M) \longrightarrow \mathcal{N} : \mathbb{U}\mathbf{CSet}$$
$$\eta \, \bar{f}_{\mathcal{F}} = f \quad \blacksquare$$

We implement the universal property for free models of $\mathbb{T}$.

**prop** Let $X : \mathbf{Set}$ then

$$h(\mathbb{T}[X]) := \mathbb{U}[X] \text{ with } \eta := \overline{\eta^{\mathbb{T}}_X} : \mathbb{T}[X] \longrightarrow \mathbb{U}[X]_{\mathcal{F}} : \mathbb{T}\mathbf{CSet}$$

**proof**

$$\frac{\dfrac{\mathbb{U}[X] \longrightarrow \mathcal{N} : \mathbb{U}\mathbf{CSet}}{X \longrightarrow \mathcal{N} : \mathbf{Set}}}{\mathbb{T}[X] \longrightarrow \mathcal{N}_{\mathcal{F}} : \mathbb{T}\mathbf{CSet}} \quad \blacksquare$$

To implement this in full we need a few preliminaries.

It is obvious that $h : \mathbb{T} \to \mathbb{U} : \mathbf{Law}$ is pretty much just a natural transformation

$$h : \forall a : \mathbf{Set}_{fin}$$
$$\mathbb{T}[a] \longrightarrow \mathbb{U}[a] : \mathbf{Set};$$

since $\mathbb{T}[a] = \bullet \xrightarrow{a} \cdots \bullet : \mathbb{T}$, $\mathbb{U}[a] = \bullet \xrightarrow{a} \cdots \bullet : \mathbb{U}$

Naturality amounts to preservation of binal operations

$$\forall \bullet \xrightarrow{a} \xrightarrow{\alpha} \xrightarrow{b} \bullet : \mathbf{Set}$$
$$h(\alpha \cdot {}^{a}\mathbb{T}) = \alpha \cdot {}^{a}\mathbb{U}$$

In reverse it is easy to see that each

$$h : \mathbb{T}[\cdot] \Rightarrow \mathbb{U}[\cdot] : \mathbf{Set}_{fin} \longrightarrow \mathbf{Set}$$

such that $h(\eta \cdot \omega) = h(\eta) \cdot h(\omega)$, $h(X_1 \xrightarrow{a} \cdots \mathbb{T}) = X_1 \xrightarrow{a} \cdots \mathbb{U}$

defines a morphism $h : \mathbb{T} \to \mathbb{U} : \mathbf{Law}$.

Since $\mathbb{T}[\cdot], \mathbb{U}[\cdot] : \mathbf{Set}_{fin} \longrightarrow \mathbf{Set}$ is the finitary part of $\mathbb{T}[\cdot], \mathbb{U}[\cdot] : \mathbf{Set} \longrightarrow \mathbf{Set}$

we have the extension

$$h : \mathbb{T}[\cdot] \Rightarrow \mathbb{U}[\cdot] : \mathbf{Set} \to \mathbf{Set}$$

It only gives a construction of $h(M)$ for $M : \mathbb{T}\mathbf{CSet}$, without the proof that it satisfies the universal property.

Let $\mathcal{U} := \overline{\eta^{\mathbb{T}}_M} : \mathbb{T}[M] \to M : \mathbb{T}\mathbf{CSet}$ be the canonical presentation.

Denote $\approx := \ker \mathcal{U} : Q_{\mathbb{T}}(\mathbb{T}[M : \mathbf{Set}])$.

Then since we have $h : \mathbb{T}[M] \to \mathbb{U}[M : \mathbf{Set}]$, we can define

$$\approx_h := \langle h(t) = h(t') \mid t \approx t' : \mathbb{T}[M] \rangle : Q_{\mathbb{U}}(\mathbb{U}[M : \mathbf{Set}]);$$

Then $\mathbb{U}[M : \mathbf{Set}] / \approx_h$ with $\eta := h_M : M \longrightarrow \left[\mathbb{U}[M : \mathbf{Set}] / \approx_h\right]_h : \mathbb{T}\mathbf{CSet}$

implements the universal property.

So

**prop** Each $h : \mathbb{T} \longrightarrow \mathbb{U} : \mathbf{Law}$ gives rise to an adjunction

$$\mathbb{T}\mathbf{CSet} \underset{\mathcal{F}}{\overset{h_!}{\rightleftarrows}} \mathbb{U}\mathbf{CSet} \quad \blacksquare$$

**example** Given any $\mathbb{T} : \mathbf{Law}$, there is unique morphism $! : \mathbf{set} \longrightarrow \mathbb{T} : \mathbf{Law}$.

Then $!^* : \mathbb{T}\mathbf{CSet} \longrightarrow \mathbf{Set}$ is just the forgetful functor,

and $!_* : \mathbf{Set} \longrightarrow \mathbb{T}\mathbf{CSet}$ the free model functor. $\blacksquare$

**example** Each monoid $M$ defines the theory of $M$-sets also denoted by $\mathbf{set}(M)$.

Then given a monoid homomorphism $h : M \longrightarrow N$,

we get $h : \mathbf{set}(M) \longrightarrow \mathbf{set}(N) : \mathbf{Law}$

Actually this is invertable so

$$\mathbf{Mon} \text{ embeds into } \mathbf{Law}.$$

Then also $\mathbf{Grp}$, $\mathbf{Ab}$ are embedded into $\mathbf{Law}$.

That is we can think of individual groups (monoids) as theories $\blacksquare$

**example** Each field $K$ defines the theory $\mathbf{va}_K$

Let $K \xrightarrow{h} L$ be a field homomorphism.

Then it defines $\mathbf{vec}_K \xrightarrow{h} \mathbf{vec}_L : \mathbf{Law}$

Thus is not a bijection; morphisms $h : \mathbf{vec}_K \longrightarrow \mathbf{vec}_L : \mathbf{Law}$

correspond to maps $h : K \longrightarrow L$ such that $h(\alpha) = \alpha$, $h(a \cdot b) = h(a) h(b)$.

This works for general rings $\blacksquare$

**example** Denote the theory of groups by $\mathbf{grp}$, and the theory of abelian groups by $\mathbf{ab}$.

Then each group operation $\omega : \bullet \xrightarrow{n} \bullet : \mathbf{grp}$ is still an abelian group operation $\omega : \bullet \xrightarrow{n} \bullet : \mathbf{ab}$,

we can say that $\mathbf{ab} : \mathbf{Law}$ is a quotient theory of $\mathbf{grp} : \mathbf{Law}$

$$q : \mathbf{grp} \longrightarrow \mathbf{ab} : \mathbf{Law}$$

Then $q_*$ is just the embedding $\mathbf{Ab} \hookrightarrow \mathbf{Grp}$;

the $q^*$ is abelianization. $\blacksquare$

**example** Let $\mathbf{mon}$ denote the theory of monoids; then we have $q : \mathbf{mon} \longrightarrow \mathbf{grp}$,

and $q_*$ is the Grothendieck Group construction $\blacksquare$

**example** Let $K$ be a field, $\mathbf{lie}_K$ the theory of Lie algebras over $K$,

and $\mathbf{alg}_K$ the theory of associative algebras over $K$.

Then we have $\ell : \mathbf{lie}_K \longrightarrow \mathbf{alg}_K : \mathbf{Law}$

that sends the commutator $[a, b]$ to $a \cdot b - b \cdot a$.

Then $\ell^*$ associates to each algebra its underlying Lie algebra.

The $\ell_*$ is the Universal Enveloping Algebra of a Lie algebra. $\blacksquare$

**example** Let $G$ be a group, $K$ a field. Denote the theory of $K$-linear $G$-representations by $\mathbf{vec}_K(G)$. Then for $f : G \longrightarrow H$ a group homomorphism,

we get $f : \mathbf{vec}_K(G) \longrightarrow \mathbf{vec}_K(H) : \mathbf{Law}$,

and $f_*$ associates to each $V$ a $K$-linear $G$-representation its

$f$-induced $H$-representation $\mathrm{Ind}_f(V)$. $\blacksquare$

# Functional Effects & Monads

We have various extensions of the notion of a function which we'll call functional effects.

## Partial Functions

A partial function from sets $X$ to $Y$ is a function from $X$ to $Y$ that may be undefined for some of its inputs.

We'll temporarily write $f : X \rightharpoonup Y$ for such functions.

**example** Define $f : \mathbb{Q} \rightharpoonup \mathbb{Q}$,

$$f(x) := \frac{1}{x}.$$

Then $f$ is defined everywhere except at $0 : \mathbb{Q}$.

Similarly, $g : \mathbb{Z} \rightharpoonup \mathbb{Z}$,

$$g(n) := \sqrt{n}$$

is defined only on $\{n : \mathbb{Z} \mid \exists x : \mathbb{Z},\ n = x^2\}$. ∎

We can view such a notion of a function as an extension of set functions with effect of partiality.

We can implement the effect of partiality $\perp$ as follows.

**define** Let $\forall X : \mathrm{Set}$,

$$\perp[X] = X + \{\perp\}. \ ∎$$

**define** Then a partial function $f$ from $X$ to $Y$ is a set function

$$f : X \longrightarrow \perp[Y]. \ ∎$$

Note that this precise definition exactly corresponds to the intuitive notion of a partial function:

$$f : X \rightharpoonup Y$$

* "undefined" on $\{x : X \mid f(x) = \perp\}$

Think of $\perp : \perp[Y]$ as an element signifying nothing, error, undefined result, or failure.

We can compose partial functions

**define** Let $X \xrightarrow{f} \perp[Y]$, $Y \xrightarrow{g} \perp[Z]$. We define

$$g \bullet f : X \longrightarrow \perp[Z]$$

$$(g \bullet f)(x) := \begin{cases} g(f(x)) & \text{if } f(x) : Y \\ \perp & \text{if } f(x) = \perp \end{cases} \ ∎$$

**example** Intuitively,

$$X \xrightarrow{f} Y \xrightarrow{g} Z$$

$$x_1 \rightsquigarrow y_1 \longrightarrow z_1$$
$$x_2 \rightsquigarrow y_2 \longrightarrow \perp$$
$$x_3 \rightsquigarrow \perp \longrightarrow \perp.$$

Try to see all the possible situations when we compose 3 or more partial functions ∎

Note that such composition is associative

$$h \bullet (g \bullet f) = (h \bullet g) \bullet f$$

To organize this into a category we need to make sense out of identity

$$id^{\perp}_X : X \rightharpoonup X$$

**define** $id^{\perp}_X : X \to \perp[X]$

$$id^{\perp}_X(x) := x. \ ∎$$

Then for $f : X \to \perp[X]$ we have

$$f \bullet id^{\perp}_X = f = id^{\perp}_Y \bullet f.$$

Often such an identity is called the unit of the effect, and is denoted by

$$\eta_X = id^{\perp}_X : X \longrightarrow \perp[X]$$

**define** Category $\mathbb{E}ff[\perp]$ of partial set functions consists of

object $X : \mathrm{Set}$

morphism given objects $X, Y$, a morphism is a set function

$$f : X \to \perp[Y]$$

& the previously defined composition & identity. ∎

Now for a partial function $f : X \rightharpoonup Y$ we can write

$$f : X \to \perp[Y] : \mathrm{Set}, \text{ or equivalently } f : X \to Y : \mathbb{E}ff[\perp]$$

We say that $\mathbb{E}ff[\perp]$ is the category of effects of functional effect $\perp$.

Each set function defines a partial function which happens to be defined everywhere.

**define** $\forall f : X \to Y : \mathrm{Set}$

$$\eta f : X \to \perp[Y] : \mathrm{Set}.$$

$$\eta f(x) := f(x) \ ∎$$

Note that the result of is the composition in $\mathrm{Set}$

$$X \xrightarrow{f} Y \xrightarrow{\eta^{\perp}_Y = \eta_Y} \perp[Y].$$

This partial composition extends the usual composition:

$$\forall X \xrightarrow{f} Y \xrightarrow{g} Z : \mathrm{Set}$$

$$\eta(g \bullet f) = \eta g \bullet \eta f.$$

Also $\eta\, id_X = \eta_X : X \to \perp[X]$.

So we get a functor

$$\eta : \mathrm{Set} \longrightarrow \perp[\mathrm{Set}]$$

$$X \mapsto X$$
$$f \mapsto \eta f$$
$$Y \mapsto Y$$

$\mathrm{Set}$ $\qquad$ $\mathbb{E}ff[\perp]$

Recall the process of function application that assigns to $f : X \to Y : \mathrm{Set}$ and $x : X$ the result $f(x) : Y$.

Its type is $(\cdot)(\cdot) : [X \to Y] \times X \longrightarrow Y$

We can extend function application to include partiality.

Given $f : X \to \perp[Y]$, it makes sense to combine it with any $p : \perp[X]$ — we'll denote the result by $f(p)$.

**define** $(\cdot)(\cdot) : [X \to \perp[Y]] \times \perp[X] \longrightarrow \perp[Y]$.

$\forall x : X$,

$$f(x) := f(x)$$
$$f(\perp) := \perp. \ ∎$$

Given $X \xrightarrow{f} Y \xrightarrow{g} Z : \mathrm{Set}$, we note the compatibility of application with composition and the identity

$$\forall p : \perp[X],$$
$$(g \bullet f)(p) = f(g(p))$$
$$id^{\perp}_X(p) = p.$$

Also $\forall x : X$,

$$(\eta \bullet f)(x) = g(f(x)).$$

Obviously we can extend $f : X \to \perp[Y]$, to

**define** $\bar{f} : \perp[X] \to \perp[Y]$

$$\bar{f}(p) := f(p) \ ∎$$

This allows us to think of partial application as the usual application.

Obviously $\overline{id^{\perp}_X} = id_{\perp[X]}$

It is not hard to see that we get a functor

$$\perp : \mathrm{Set} \longrightarrow \mathrm{Set}$$
$$X \rightsquigarrow \perp[X]$$
$$f \downarrow \rightsquigarrow \downarrow \bar{f}$$
$$Y \rightsquigarrow \perp[Y]$$

We started out with trying to formalize the notion of a partial function; this led us to extend the usual notions of composition, identity, function application.

By trying to organize these new extended constructs into categories and functors we were lead naturally to formulate basic laws of this new type of functions.

After one more example we will axiomatize the notion of a functional effect.

## Multivalued Functions

Now we'll formalize multivalued functions.

There are many notions of multivaluedness.

Let $X, Y : \mathrm{Set}$; in this example we'll focus on multivalued functions $f : X \rightharpoonup Y$ such that to each $x : X$, $f$ associates a sequence of outputs $y_1, \ldots, y_n$. For us this sequence is ordered, may be empty, and it can contain some finite number of elements.

We can formalize this with lists.

**define** $\forall X : \mathrm{Set}$,

$$List[X] := \{\text{finite sequence of elements in } X\}$$

We form lists with square brackets $[x_1, \ldots, x_n] : List[X]$.

Empty list is denoted by $[\,] : List[X]$.

Given two lists $xs := [x_1, \ldots, x_n]$, $xs' := [x'_1, \ldots, x'_m] : List[X]$, denote their concatenation by $xs \cdot ys := [x_1, \ldots, x_n, x'_1, \ldots, x'_m]$ ∎

**define** A List-multivalued function from $X$ to $Y$ is a map

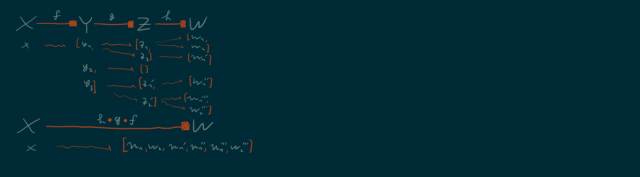
$$f : X \longrightarrow List[Y] : \mathrm{Set} \ ∎$$

Given $f : X \to List[Y]$, $g : Y \to List[Z]$, we can compose them $g \bullet f : X \longrightarrow List[Z]$,

suppose $x : X$, $f(x) := [y_1, y_2]$, $g(y_1) := [z_1, z_2]$, $g(y_2) := [z'_1, z'_2, z'_3]$.

Then $(g \bullet f)(x) = [z_1, z_2, z'_1, z'_2, z'_3]$.

**define** $(g \bullet f)(x) = g(y_1) \cdot g(y_2) \cdot \ldots \cdot g(y_n)$ where $[y_1, \ldots, y_n] := f(x)$ ∎



**define** A *Functional Effect* $\mathbb{T}$ consists of

a map $\mathbb{T} : \mathbf{Set} \longrightarrow \mathbf{Set}$ [we do not assume functoriality]

with (unit) $\eta : \forall X : \mathbf{Set}.$
$$X \longrightarrow \mathbb{T}[X]$$

(composition) $\bullet : \forall X, Y, Z : \mathbf{Set}.$
$$[Y \to \mathbb{T}[Z]] \times [X \to \mathbb{T}[Y]] \longrightarrow [X \to \mathbb{T}[Z]]$$

(application) $\cdot(\cdot) : \forall X, Y : \mathbf{Set}.$
$$[X \to \mathbb{T}[Y]] \to [\mathbb{T}[X] \to \mathbb{T}[Y]]$$

where (associativity) $\forall X \xrightarrow{f} \mathbb{T}[Y], Y \xrightarrow{g} \mathbb{T}[Z], Z \xrightarrow{h} \mathbb{T}[W].$
$$h \bullet (g \bullet f) = (h \bullet g) \bullet f$$

(neutrality) $\forall X \xrightarrow{f} \mathbb{T}[Y].$
$$f \bullet \eta_X = f = \eta_Y \bullet f$$

(composition compatibility) $\forall X \xrightarrow{f} \mathbb{T}[Y], Y \xrightarrow{g} \mathbb{T}[Z].$
$$g \bullet f = g(\!(f)\!) \blacksquare$$

We define category of $\mathbb{T}$-functional effects

**define** $\mathbf{Eff}[\mathbb{T}]$

obj $X : \mathbf{Set}$

mor $\dfrac{f : X \longrightarrow Y : \mathbf{Eff}[\mathbb{T}]}{f : X \longrightarrow \mathbb{T}[Y] : \mathbf{Set}}$

composition $\bullet : \forall X, Y, Z.$
$$(Y \to Z) : \mathbf{Eff}[\mathbb{T}] \times (X \to Y) : \mathbf{Eff}[\mathbb{T}] \longrightarrow (X \to Z) : \mathbf{Eff}[\mathbb{T}]$$

identity $\eta : \forall X.$
$$X \to X : \mathbf{Eff}[\mathbb{T}] \blacksquare$$

In literature this category is also known as the *Kleisli Category*.

Note that (composition compatibility) says that composition is determined by application, elementwise
$$\forall x : X.$$
$$(g \bullet f)(x) = g(\!(f)\!)(x)).$$

This implies an analogue of the usual function composition & application.

**prop** $\forall X \xrightarrow{f} \mathbb{T}[Y], Y \xrightarrow{g} \mathbb{T}[Z].$
$$\forall t : \mathbb{T}[X].$$
$$(g \bullet f)(t) = g(\!(f(\!(t)\!))\!)$$

**proof** Since $(g \bullet f)(\cdot) = g(\!(f(\!(\cdot)\!))\!)$,
so $(g \bullet f)(\!(\cdot)\!) = g(\!(f(\!(\cdot)\!))\!) \blacksquare$

Conversely application is determined by composition.

Given $f : X \to \mathbb{T}[Y] : \mathbf{Set}$ we interpret it as
$$f : X \to Y : \mathbf{Eff}[\mathbb{T}].$$

Also we interpret $\mathrm{id}_{\mathbb{T}(X)} : \mathbb{T}[X] \to \mathbb{T}[X] : \mathbf{Set}$ as
$$\eta_{\mathbb{T}[X]} : \mathbb{T}[X] \to X : \mathbf{Eff}[\mathbb{T}];$$

Then we may form
$$f \bullet \mathrm{id}_{\mathbb{T}[X]} : \mathbb{T}[X] \longrightarrow Y : \mathbf{Eff}[\mathbb{T}]$$
$$\text{or} \quad f \bullet \mathrm{id}_{\mathbb{T}[X]} : \mathbb{T}[X] \to \mathbb{T}[Y] : \mathbf{Set}.$$

**prop** $f(\!(\cdot)\!) = f \bullet \mathrm{id}_{\mathbb{T}[X]}$

**proof** $f \bullet \mathrm{id}_{\mathbb{T}[X]} = f(\!(\mathrm{id}_{\mathbb{T}(X)})\!) = f(\!(\cdot)\!) \blacksquare$

It is advantageous to write $\overline{f} := f(\!(\cdot)\!) : \mathbb{T}[X] \to \mathbb{T}[Y]$ as a kind of *lifted* version of $f$.

Usually functional effects are axiomatized by choosing only composition or only application, then devising suitable laws, and after that defining the other operation in terms of the first.

But then some of the axioms become too mysterious for the first exposure to these notions.

We've seen an axiomatization with both composition and application, but in practice it may be too cumbersome to check associativity directly.

For this reason I give an alternative axiomatization in terms of application.

We just need to develop analogues for neutrality and associativity.

**prop** [application neutrality]

Let $X : \mathbf{Set}$, $f : X \to \mathbb{T}[Y]$, then

(1) $\eta_X(\!(\cdot)\!) = \mathrm{id}_{\mathbb{T}[X]}$,

i.e. $\forall t : \mathbb{T}[X].$
$$\eta_X(\!(t)\!) = t$$

(2) $f(\!(\eta_X)\!) = f$

i.e. $\forall x : X.$
$$f(\!(\eta_X(x))\!) = f(x).$$

**proof** (1) $\eta_X(\!(\cdot)\!) = \eta_X \bullet \mathrm{id}_{\mathbb{T}[X]} = \mathrm{id}_{\mathbb{T}[X]}$

(2) $f(\!(\eta_X)\!) = f \bullet \eta_X = f \blacksquare$

Associativity is a bit more cumbersome

**prop** [application association]

Let $Y \xrightarrow{g} \mathbb{T}[Z]$, $Z \xrightarrow{h} \mathbb{T}[W]$, then
$$h(\!(g(\!(\cdot)\!))\!) = (h(\!(g)\!))(\!(\cdot)\!)$$

i.e. $\forall t : \mathbb{T}[Y].$
$$h(\!(g(\!(t)\!))\!) = (\lambda y.\, h(\!(g(\!(y)\!))\!))(\!(t)\!).$$

**proof** Let $t : \mathbb{T}[Y]$.

Consider the sequence of maps
$$Y \xrightarrow{\eta_Y} \mathbb{T}[Y], \quad Y \xrightarrow{g} \mathbb{T}[Z], \quad Z \xrightarrow{h} \mathbb{T}[W],$$

then $h \bullet (g \bullet \eta_Y) = (h \bullet g) \bullet \eta_Y$,

so $[h \bullet (g \bullet \eta_Y)](\!(t)\!) = [(h \bullet g) \bullet \eta_Y](\!(t)\!)$,

but then
$$\begin{aligned}[h \bullet (g \bullet \eta_Y)](\!(t)\!) &= h(\!((g \bullet \eta_Y)(\!(t)\!))\!) \\ &= h(\!(g(\!(\eta_Y(\!(t)\!))\!))\!) \\ &= h(\!(g(\!(t)\!))\!),\end{aligned}$$

and
$$\begin{aligned}[(h \bullet g) \bullet \eta_Y](\!(t)\!) &= (h \bullet g)(\!(\eta_Y(\!(t)\!))\!) \\ &= (h \bullet g)(\!(t)\!) \\ &= (\lambda y.\, h(\!(g(\!(y)\!))\!))(\!(t)\!) \blacksquare\end{aligned}$$

Reversing neutrality and associativity to *left* variation
$$\overline{\eta_X} = \mathrm{id}_{\mathbb{T}[X]}, \quad \overline{f} \circ \eta_X = f,$$
$$\overline{h} \circ \overline{g} = \overline{h \circ g}$$

**define** [2nd Axiomatization of *Functional Effect*]

A *Functional Effect* $\mathbb{T}$ consists of

a map $\mathbb{T} : \mathbf{Set} \longrightarrow \mathbf{Set}$ [we do not assume functoriality]

with (unit) $\eta : \forall X : \mathbf{Set}.$
$$X \longrightarrow \mathbb{T}[X]$$

(application) $\cdot(\cdot) : \forall X, Y : \mathbf{Set}.$
$$[X \to \mathbb{T}[Y]] \to [\mathbb{T}[X] \to \mathbb{T}[Y]]$$

where (application neutrality) (1) $\forall X : \mathbf{Set}$, $t : \mathbb{T}[X]$.
$$\eta_X(\!(t)\!) = t$$

(2) $\forall X \xrightarrow{f} \mathbb{T}[Y]$, $x : X$.
$$f(\!(\eta_X(x))\!) = f(x)$$

(application association) $\forall Y \xrightarrow{g} \mathbb{T}[Z]$, $Z \xrightarrow{h} \mathbb{T}[W]$, $t : \mathbb{T}[Y]$.
$$h(\!(g(\!(t)\!))\!) = (\lambda y.\, h(\!(g(\!(y)\!))\!))(\!(t)\!) \blacksquare$$

We've just seen that the first axiomatization satisfies the second.

Now we show the converse.

**prop** Let $\mathbb{T}$ with $\eta$ and $\cdot(\cdot)$ satisfy the second axiomatization.

Let $f : X \to \mathbb{T}[Y]$, $g : Y \to \mathbb{T}[Z]$, $h : Z \to \mathbb{T}[W]$, $x : X$, define the composition as
$$(g \bullet f)(x) := g(\!(f(x))\!).$$

Then (1) $f \bullet \eta_X = f = \eta_Y \bullet f$

and (2) $h \bullet (g \bullet f) = (h \bullet g) \bullet f$,

**proof** (1) $(f \bullet \eta_X)(x) = f(\!(\eta_X(x))\!) = f(x)$,

$(\eta_Y \bullet f)(x) = \eta_Y(\!(f(x))\!) = f(x)$ since $f(x) : \mathbb{T}[Y]$

(2)
$$\begin{aligned}[h \bullet (g \bullet f)](x) &= h(\!((g \bullet f)(x))\!) \\ &= h(\!(g(\!(f(x))\!))\!) \\ &= (\lambda y.\, h(\!(g(\!(y)\!))\!))(\!(f(x))\!) \quad \text{by (application associativity)}\end{aligned}$$

but
$$\begin{aligned}[(h \bullet g) \bullet f](x) &= (h \bullet g)(\!(f(x))\!) \\ &= (\lambda y.\, h(\!(g(\!(y)\!))\!))(\!(f(x))\!). \blacksquare\end{aligned}$$

Now we'll define the effect extension functor

**define** $\overline{\mathbb{T}} : \mathbf{Set} \longrightarrow \mathbf{Eff}[\mathbb{T}]$
$$X \mapsto X$$
$$X \xrightarrow{f} Y \mapsto X \xrightarrow{\overline{f}} \mathbb{T}[Y] := X \xrightarrow{f} Y \xrightarrow{\eta_Y} \mathbb{T}[Y] \blacksquare$$

This is indeed a functor.

**prop** (1) $\overline{\eta_X} = \eta_X$

(2) $\overline{(g \circ f)} = \overline{g} \circ \overline{f}$

**proof** (2) $\forall x : X.$
$$\begin{aligned}(\overline{g} \circ \overline{f})(x) &= \overline{g}(\overline{f}(x)) \\ &= \overline{g}(\eta_Y(f(x))) \\ &= \overline{g}(f(x)) \\ &= \overline{(g \circ f)}(x) \blacksquare\end{aligned}$$

Now $\mathbb{T}$ becomes a functor in a canonical way

**prop** $\mathbb{T} : \mathbf{Set} \longrightarrow \mathbf{Set}$ is a functor with
$$\forall X \xrightarrow{f} Y.$$
$$\mathbb{T}[f] : \mathbb{T}[X] \longrightarrow \mathbb{T}[Y]$$
$$\mathbb{T}[f](t) = (\eta_Y \circ f)(\!(t)\!)$$

**proof** Let $X \xrightarrow{f} Y \xrightarrow{g} Z$, $t : \mathbb{T}[X]$

(1) **goal** $\eta_Y \circ f(\!(t)\!) = \eta_Y(\!(\eta_Y \circ f(\!(t)\!))\!)$.
$$\begin{aligned}\eta_Y(\!(\eta_Y \circ f(\!(t)\!))\!) &= (\eta_Y \circ \eta_Y \circ f)(\!(t)\!) \\ &= \eta_Y(\eta_Y \circ f)(\!(t)\!).\end{aligned}$$

(2) **goal** $(\eta_Y \circ g)(\!(t)\!) = t$.

This is just $\eta_Y(\!(t)\!) = t \blacksquare$

Equivalently, we have $\mathbb{T}[g] = \overline{\eta_Y \circ f}$

**prop** $\eta : \forall X : \mathbf{Set}.$
$$X \longrightarrow \mathbb{T}[X]$$
is a natural transformation $\eta : \mathrm{id}_{\mathbf{Set}} \Longrightarrow \mathbb{T}$

**proof** Let $X \xrightarrow{f} Y : \mathbf{Set}$, $x : X$.

**goal** $\mathbb{T}[f](\eta_X(x)) = \eta_Y(f(x))$.
$$\begin{aligned}\mathbb{T}[f](\eta_X(x)) &= \eta_X(f(x)) \\ \eta_Y(f(\eta_X(x))) &= \eta_Y f(x). \blacksquare\end{aligned}$$

We have a functor in the other direction

**prop** $|\mathbb{T}| : \mathbf{Eff}[\mathbb{T}] \longrightarrow \mathbf{Set}$
$$X \longmapsto \mathbb{T}[X]$$
$$X \xrightarrow{f} \mathbb{T}[Y] \longmapsto \mathbb{T}[X] \xrightarrow{f(\cdot)} \mathbb{T}[Y]$$

is a functor

**proof** Let $X \xrightarrow{f} \mathbb{T}[Y]$, $Y \xrightarrow{g} \mathbb{T}[Y]$, $t : \mathbb{T}[X]$.

We need to show
$$(g \circ f)(\!(t)\!) = g(\!(f(\!(t)\!))\!)$$
$$\eta_X(\!(t)\!) = t$$
which we already know to be true $\blacksquare$

Note $\mathbb{T} := |\mathbb{T}| \circ \overline{\mathbb{T}}$, and almost by definition we get an adjunction,

because
$$\frac{X \longrightarrow \mathbb{T}[Y] : \mathbf{Set}}{X \longrightarrow Y : \mathbf{Eff}[\mathbb{T}]}$$

**prop**
$$\mathbf{Set} \underset{|\mathbb{T}|}{\overset{\overline{\mathbb{T}}}{\underset{\perp}{\rightleftarrows}}} \mathbf{Eff}[\mathbb{T}]$$

**proof** It is enough to show that
$$\forall X : \mathbf{Set}, \quad Y : \mathbf{Eff}[\mathbb{T}].$$
$$\frac{\dfrac{\dfrac{X \longrightarrow |\mathbb{T}||Y| : \mathbf{Set}}{X \longrightarrow \mathbb{T}[Y] : \mathbf{Set}}}{X \longrightarrow Y : \mathbf{Eff}[\mathbb{T}]}}{\overline{\mathbb{T}}[X] \longrightarrow Y : \mathbf{Eff}[\mathbb{T}]} \blacksquare$$

Note that so far we made no use of special properties of $\mathbf{Set}$.

So the definition and the above properties can be formulated over an arbitrary category $\mathbb{E}$ without change, use of elements can be trivially eliminated.

We'd have the following types

(unit) $\eta : \forall X : \mathbb{E}.$
$$X \longrightarrow \mathbb{T}[X] : \mathbb{E}$$

(composition) $\bullet : \forall X, Y, Z : \mathbf{Set}.$
$$[Y \to \mathbb{T}[Z]] : \mathbb{E} \times_{\mathbf{Set}} [X \to \mathbb{T}[Y]] : \mathbb{E} \xrightarrow{\bullet}_{\mathbf{Set}} [X \to \mathbb{T}[Z]] : \mathbb{E}$$

(application) $\cdot(\cdot) : \forall X, Y : \mathbf{Set}.$
$$[X \to \mathbb{T}[Y]] : \mathbb{E} \xrightarrow{\cdot(\cdot)}_{\mathbf{Set}} [\mathbb{T}[X] \to \mathbb{T}[Y]] : \mathbb{E}.$$

In this work we are interested only in effects on $\mathbf{Set}$ $\blacksquare$

**example** [$\perp$ partiality, List multiveladness]

We've already seen

$\perp$ the effect of partiality,

and List the effect of ordered, finitary multiveladness with repetitions.

We'll show that List is indeed a functional effect.

Let $X \xrightarrow{f} \mathrm{List}[Y]$, $Y \xrightarrow{g} \mathrm{List}[Z]$, $xs : \mathrm{List}[X]$, $x : X$.

Then (neutrality)

(1) $\eta_X(x_1, \dots, x_n) = [x_1, \dots, x_n]$
$= [x_1, \dots, x_n]$

(2) $f(\eta_X(x)) = f([x])$
$= f(x)$

(associativity)

We want to show that

$$g(f(xs)) = (\lambda x.\, g(f(x)))(xs).$$

Assume $xs = [x_1, \dots, x_n]$, and $f(x_i) = [y_1^i, \dots, y_{m_i}^i]$.

Then $f(xs) = [y_1^1, \dots, y_{m_1}^1, \dots\dots, y_1^n, \dots, y_{m_n}^n]$,

so $g(f(xs)) = g([y_1^1, \dots, y_{m_1}^1]) \cdot \dots \cdot g([y_1^n]) \cdot \dots \cdot g([y_{m_n}^n])$,

and $(\lambda x.\, g(f(x)))(xs) = g([f(x_1)]) \cdot \dots \cdot g([f(x_n)])$

$$= g([y_1^1, \dots, y_{m_1}^1]) \cdot \dots \cdot g([y_1^n, \dots, y_{m_n}^n])$$

$$= (g(y_1^1) \cdot \dots \cdot g(y_{m_1}^1)) \cdot \dots \cdot (g(y_1^n) \cdot \dots \cdot g(y_{m_n}^n)).$$

Since concatenation in $\mathrm{List}[Z]$ is associative, we know

$$g(f(xs)) = (\lambda x.\, g(f(x)))(xs). \blacksquare$$

**example** [Trivial Effect]

We may consider augmenting functions in a trivial way by

$\mathrm{Id} : \mathbf{Set} \longrightarrow \mathbf{Set}$

$\mathrm{Id}[X] := X$,

$\eta_X(x) := x$

$g \bullet f := g f$

$f(x) := f(x)$. $\blacksquare$

We return to special kind of multiveladness – non-determinism – by abandoning ordering. $\blacksquare$

**example** [$\{\!\{\cdot\}\!\}_{fin}$ finitary non-determinism]

A non-deterministic function $f : X \longrightarrow Y$ given $x : X$, returns an unordered finite sequence of outputs $y_1, \dots, y_n$.

To formalize this we need the notion of a multisubset.

A multisubset of $X$ is a subset of $X$ with multiplicities – an element can be included in a subset more than once.

We form them with double curly braces

$$\{\!\{a, b, b, c\}\!\} \subseteq X$$

Note $\{\!\{a, a\}\!\} \neq \{\!\{a\}\!\} \subseteq X$, as opposed to usual subsets $\{a, a\} = \{a\} \subseteq X$.

Since we do not care about the order, we have $\{\!\{a, b\}\!\} = \{\!\{b, a\}\!\} \subseteq X$.

To specify such a multisubset $S$ we give multiplicity to each element of $X$.

**define** (1) A multisubset of $X$ is a map

$$S : X \longrightarrow \mathbb{N}$$

Denote $\{\!\{X\}\!\} := X \longrightarrow \mathbb{N}$

Will write

multisubset $S \subseteq X$ for $S : \{\!\{X\}\!\}$.

(2) A multisubset $S \subseteq X$ is finite

$$\# \operatorname{supp} S = \{x : X \mid S(x) \neq 0\} \text{ is finite}$$

Denote by $\{\!\{X\}\!\}_{fin}$ the set of all finite multisubsets of $X$. $\blacksquare$

**example** Prime factorization of a positive integer is a multisubset of prime numbers; for example to $504 = 8 \cdot 9 \cdot 7 = 2^3 \cdot 3^2 \cdot 7$ we associate $\{\!\{2, 2, 2, 3, 3, 7\}\!\}$. $\blacksquare$

**example** Roots of a polynomial over a field $K$ form a multisubset of $K$ $\blacksquare$

**example** Eigenvalues of a linear operator form a multisubset in more than one way (algebraic vs. geometric multiplicities) $\blacksquare$

We can also write an element in $\{\!\{X\}\!\}_{fin}$ in additive notation. For example

$$\{\!\{a, b\}\!\}, \{\!\{a, a, a, b, b, c, d, d\}\!\}, \{\!\{\}\!\} : \{\!\{X\}\!\}_{fin}$$

becomes $a + b$, $3a + 2b + c + 2d$, $0 : \{\!\{X\}\!\}_{fin}$

Given $S := \{\!\{a, a, b\}\!\} = 2a + b$, $T := \{\!\{a, b, b, c, c\}\!\} = a + 2b + 2c \subseteq \{\!\{X\}\!\}_{fin}$,

they combine into $S + T := \{\!\{a, a, a, b, b, b, c, c\}\!\} = 3a + 3b + 2c$.

**define** For $S, T : \{\!\{X\}\!\}$, define

$$S + T : \{\!\{X\}\!\}$$

$$(S + T)(x) := S(x) + T(x). \blacksquare$$

**define** A function with $\{\!\{\cdot\}\!\}_{fin}$-non-deterministic effect is a map

$$f : X \longrightarrow \{\!\{Y\}\!\}_{fin} \quad \mathbf{Set} \blacksquare$$

We can extend it to work on $\{\!\{X\}\!\}_{fin}$, for example assuming $f(a) = \{\!\{c, d, d\}\!\}$, $f(b) = \{\!\{c, e\}\!\}$, then

$$f(\{\!\{a, a, b\}\!\}) = f(a) + f(a) + f(b)$$
$$= \{\!\{c, d, d\}\!\} + \{\!\{c, d, d\}\!\} + \{\!\{c, e\}\!\}$$
$$= \{\!\{c, c, c, d, d, d, d, d, e\}\!\}$$

i.e. we impose a kind of linearity on $f$

$$f(2 \cdot a + b) = 2 \cdot f(a) + f(b)$$
$$= 2 \cdot (c + 2 \cdot d) + (c + e)$$
$$= 3 \cdot c + 4 \cdot d + e$$

**define** Given $f : X \longrightarrow \{\!\{Y\}\!\}_{fin}$, $S : \{\!\{X\}\!\}_{fin}$,

define

$$f(S) : \{\!\{Y\}\!\}_{fin}$$

$$f(S) := \sum_{x : X} S(x) \cdot f(x) \blacksquare$$

**define** $\eta_X : X \longrightarrow \{\!\{X\}\!\}_{fin}$

$$\eta_X(x) := \{\!\{x\}\!\} \blacksquare$$

Thus we have the $\{\!\{\cdot\}\!\}$-effect.

Then for $f : X \to \{\!\{Y\}\!\}_{fin}$, $g : Y \to \{\!\{Z\}\!\}_{fin}$, we get composition

$$(g \bullet f)(x) := g(f(x)),$$

and applying the functor to $f : X \to Y$, we get

$$\{\!\{f\}\!\}_X : \{\!\{X\}\!\}_{fin} \longrightarrow \{\!\{Y\}\!\}_{fin}$$

$$\{\!\{f\}\!\}_X(S) := \sum_x f(x) S(x) \blacksquare$$

**example** [$\mathcal{P}_{fin}$ finitary non-determinism, no multiplicities]

When we discard both multiplicities and order, we have $\mathcal{P}_{fin}$-effect.

**define** $\mathcal{P}_{fin}[X] := \{\text{finite subsets of } X\}$ $\blacksquare$

For example $f : X \longrightarrow \mathcal{P}_{fin}[Y]$, $f(a, b, c\} : \mathcal{P}_{fin}[X]$,

$$f(\{a, b, c\}) = f(a) \cup f(b) \cup f(c),$$

$$\eta_X : X \longrightarrow \mathcal{P}_{fin}[X]$$

$$\eta_X(x) := \{x\}. \blacksquare$$

**example** [$\mathcal{P}$ unbounded non-determinism, no multiplicities]

We can generalize $\mathcal{P}_{fin}$-effect by abandoning finiteness.

By $\mathcal{P}[X]$ we denote the set of all subsets of $X$.

Then a function with $\mathcal{P}$-effect is just a map

$$R : X \longrightarrow \mathcal{P}[Y],$$

which amounts to the same thing as a relation from $X$ to $Y$:

$$x R y \quad :\Leftrightarrow \quad y \in R(x)$$

**define** For $S : \mathcal{P}[X]$,

$$R(S) := \bigcup_{x : S} R(s) \blacksquare$$

Given two such maps $R : X \to \mathcal{P}[Y]$, $S : Y \to \mathcal{P}[Z]$, their composition $S \bullet R$ corresponds to the composition of relations since

$$z : (S \bullet R)(x) \quad :\Leftrightarrow \quad z : S(R(x)) = \bigcup_{y : R(x)} S(y)$$

$$:\Leftrightarrow \quad \exists y.\, R(x).\ z : S(y)$$

$$:\Leftrightarrow \quad \exists y.\, x R y,\ y S z$$

So the category $\mathbf{Eff}[\mathcal{P}]$ is the category of sets with relations as morphisms.

Applying the functor map to $f : X \longrightarrow Y$ we get the direct image $f[\cdot] : \mathcal{P}[X] \longrightarrow \mathcal{P}[Y]$. $\blacksquare$

**example** [linear combinations in a field]

Let $K$ be a field. Then we can consider finite linear combinations over a set $X$ like

$$3 \cdot a + 5 \cdot b + (-2) \cdot c.$$

We formalize them as follows

**define** Let $X : \mathbf{Set}$, def.

$$\mathrm{vec}_K[X] := \left\{ a : X \to K \;\middle|\; a \text{ has finite support } \{x : X \mid a(x) \neq 0\} \right\} \blacksquare$$

Obviously this is a vector space over $K$.

We make this into an effect.

**define** $\eta_X : X \longrightarrow \mathrm{vec}_K[X]$

$$\eta_X(x) := x \blacksquare$$

**define** For $f : X \longrightarrow \mathrm{vec}_K[Y]$, $a_1 x_1 + \dots + a_n x_n : \mathrm{vec}_K[X]$,

set $f(a_1 x_1 + \dots + a_n x_n) := a_1 \cdot f(x_1) + \dots + a_n f(x_n)$. $\blacksquare$

Checking that this is a functional effect amounts to checking that a linear combination of linear combinations is a linear combination.

We can think of $f : X \longrightarrow \mathrm{vec}_K[Y]$ via its extension

$$\bar{f} : \mathrm{vec}_K[X] \longrightarrow \mathrm{vec}_K[Y]$$

as a linear map from vec. space with basis $X$ to a vec. space with basis $Y$.

And since any linear map $L : \mathrm{vec}_K[X] \longrightarrow \mathrm{vec}_K[Y]$ is determined on the basis, the category $\mathbf{Eff}[\mathrm{vec}_K]$ is canonically equivalent to the category of vector spaces with a basis and linear maps (that need not preserve the basis) $\blacksquare$

**example** Let $\mathrm{Bool} := \{0, 1\}$, and $\mathbb{F}_2 := \{0, 1\}$.

Then think of $\mathrm{Bool}$ as equipped with $\mathrm{or} : \mathrm{Bool}^2 \to \mathrm{Bool}$, false $0 : \mathrm{Bool}$ and $\mathbb{F}_2$ with $\oplus : \mathbb{F}_2^2 \to \mathbb{F}_2$ modulo 2 addition and $0 : \mathbb{F}_2$

Note $\mathcal{P}_{fin}[X] = \mathrm{vec}_{\mathbb{F}_2}[X]$.

Intuitively $\mathcal{P}_{fin}[X]$ are linear combinations valued in $\mathrm{Bool}$, while $\mathrm{vec}_{\mathbb{F}_2}[X]$ are linear combinations valued in $\mathbb{F}_2$.

Will think of $\mathcal{P}_{fin}$ as $\mathrm{vec}_{\mathrm{Bool}}$.

Note that $\mathbf{Eff}[\mathrm{vec}_{\mathbb{F}_2}]$ can be thought of as category of free vector spaces and matrices over the field $\mathbb{F}_2$.

In fact $\mathbf{Eff}[\mathcal{P}_{fin}]$ is similar, it is the category of free vector spaces and matrices, but multiplication of matrices is different:

$$\mathrm{vec}_{\mathrm{Bool}} \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 \vee 1 \vee 0 & 0 \vee 0 \vee 0 & 0 \vee 1 \vee 1 \\ 0 \vee 0 \vee 0 & 0 \vee 0 \vee 0 & 0 \vee 0 \vee 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\mathrm{vec}_{\mathbb{F}_2} \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 \oplus 1 \oplus 0 & 0 \oplus 0 \oplus 0 & 0 \oplus 1 \oplus 1 \\ 0 \oplus 0 \oplus 0 & 0 \oplus 0 \oplus 0 & 0 \oplus 0 \oplus 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Conversely recall that for $X, Y, Z$ finite,

$$f : X \longrightarrow Y : \mathbf{Eff}[\mathcal{P}_{fin}] \text{ are relations from } X \text{ to } Y.$$

So given $f : X \to \mathcal{P}[Y]$, $g : Y \to \mathcal{P}[Z]$, we have

$$z : (g \bullet f)\, x \quad :\Leftrightarrow \quad \exists y.\, Y.\ z \in g\, y,\ y \in f\, x$$

We can also think of $\mathbb{F}_2$-relations – these have different composition:

$$z : (g \bullet f)\, x \quad :\Leftrightarrow \quad \left| \{ y : Y \mid z \in g\, y,\ y \in f\, x \} \right| \text{ has odd number of elements.} \blacksquare$$

So for we considered overalloes with five operations associated to a functional effect $T$.

Composition $g \circ f$

application $f(t)$

left $\overline{f}$

unit $\eta_X(x)$

functor $T[f](t)$.

The last one is reduction $\mu(t\Sigma)$; it directly formalizes phrases such as "linear combination of linear combinations is a linear combination" and "list of lists is a list".

**example** [List] Consider lists of lists $\mathrm{List}[\mathrm{List}[X]]$; assuming $a,b,c,d : X$, then

$$\mu\left([[a,b,b],[c,d],[a]]\right) = [a,b,b,c,d,a] : \mathrm{List}[X].$$

So that $\mu([xs_1,\dots,xs_n]) = xs_1 \cdot \dots \cdot xs_n$ ■

**example** [$\mathbb{F}$-$\mathbb{N}$] Consider $\{\!\{ \{\!\{x\}\!\}_{x:X} \}\!\}_{x:X}$, assuming $a,b,b : X$, then

$$\mu\left(\{\!\{ \{\!\{a,b,b\}\!\}, \{\!\{b\}\!\}, \{\!\{a\}\!\} \}\!\}\right) = \{\!\{a,a,b,b,b\}\!\}$$ ■

**example** [$\mathrm{vec}_{\mathbb{R}}$] Consider $\mathrm{vec}_{\mathbb{R}}[\mathrm{vec}_{\mathbb{R}}[X]]$; assume $a,b,c : X$; we'll map linear combinations in brackets

$[5\cdot a + 2\cdot b]$, $[-7\cdot a + c] : \mathrm{vec}_{\mathbb{R}}[X]$,

so $[3\cdot[5\cdot a + 2\cdot b] + 7\cdot[-7\cdot a + c]] : \mathrm{vec}_{\mathbb{R}}[\mathrm{vec}_{\mathbb{R}}[X]]$

finally

$$\mu\left([3\cdot[5\cdot a + 2\cdot b] + 7\cdot[-7\cdot a + c]]\right) = [(3\cdot 5 - 7)\cdot a + (3\cdot 2)\cdot b + 7\cdot c] : \mathrm{vec}_{\mathbb{R}}[X],$$
$$= [8\cdot a + 6\cdot b + 7\cdot c]$$ ■

**example** [$\bot$] Let $X := \{a,b,c\}$, then we think that $\bot[X] = \{\eta(a), \eta(b), \eta(c), \bot\}$, and so $\bot\bot[X] = \{\eta\eta(a), \eta\eta(b), \eta\eta(c), \eta(\bot), \bot\}$, where of course $\eta(\bot) \neq \bot$.

The $\forall x : X$. $\mu(\eta\eta(x)) = \eta(x) : \bot[X]$,

and $\mu(\eta(\bot)) = \bot$ ■

We can define reduction for any functional effect $T$.

How can we reduce $tt : TT[X]$ to something in $T[X]$?

By lifting $id_{T[X]} : T[X] \longrightarrow T[X]$

to $\overline{id_{T[X]}} : TT[X] \longrightarrow T[X]$.

**define** Let $\mu : \forall X : \mathrm{Set}$.

$$TT[X] \longrightarrow T[X]$$
$$\mu_X(tt) = \overline{id_{T[X]}}(tt)$$

or equivalently,

$$\mu_X := \overline{id_{T[X]}}$$ ■

First of all reduction is a natural transformation $\mu : TT \longrightarrow T$ that is,

**prop** [$\mu$-naturality]

$\forall f : X \longrightarrow Y$.

$$\begin{tikzcd} TT[X] \arrow[r, "\mu_X"] \arrow[d, "TT[f]"'] \arrow[dr, phantom, "="] & T[X] \arrow[d, "T[f]"] \\ TT[Y] \arrow[r, "\mu_Y"'] & T[Y] \end{tikzcd}$$

**proof** Let $f : X \longrightarrow Y$.

It is enough to show that

$$\begin{tikzcd} T[X] \arrow[r, "id_{T[X]}"] \arrow[d, "\eta_{T[Y]} T[f]"'] \arrow[dr, phantom, "="] & T[X] \arrow[d, "T[f]"] \\ TT[Y] \arrow[r, "\mu_Y"'] & T[Y], \end{tikzcd}$$

and then apply $\overline{\ }$. The diagram above commutes because $\forall s : T[Y]$. $\mu_Y \eta_{T[Y]}(s) = s$,

since $\mu_Y \eta_{T[Y]}(s) = id_{T[Y]}(\eta_{T[Y]}(s)) = id_{T[Y]}(s) = s$. ■

Now we have equipped the set $T[X]$ with two maps

$$TT[X] \xrightarrow{\mu_X} T[X] \xleftarrow{\eta_X} X.$$

Will look at the analogous equipment on $TT[X]$.

**example** (2-unit $\eta$)

Given a list $[a,b,b,c] : \mathrm{List}[X]$, we can embed it into $\mathrm{List}[\mathrm{List}[X]]$ in two ways:

outer embedding by applying $\eta_{\mathrm{List}[X]}$

$$\eta_{\mathrm{List}[X]}([a,b,b,c]) = [[a,b,b,c]] : \mathrm{List}\,\mathrm{List}[X]$$

and inner embedding by applying the list functor to $\eta_X$

$$\mathrm{List}[\eta_X]([a,b,b,c]) = [[a],[b],[b],[c]] : \mathrm{List}\,\mathrm{List}[X].$$

These two maps are in general different, but are the same when restricted to inputs of the form $[a]$,

so $\eta_{\mathrm{List}[X]}(\eta(a)) = [[a]] = \mathrm{List}[\eta_X](\eta(a))$.

Also when we reduce their output, we get the same answer:

$$\mu([[a,b,b,c]]) = [a,b,b,c] = \mu([[a],[b],[b],[c]])$$

(reduction $\mu$)

Similarly with reduction:

Consider $xsss := \left[\,[[a,b,c]],\ [a],\ [c,b],\ \ [c,b],\ [c,b],\ [a]\,\right] : \mathrm{List}^3[X]$;

we can reduce its outer layer

$$\mu_{T[X]}(xsss) = [[a,b,c],\ [a],\ [c,b],\ [c,b],\ [a]] : \mathrm{List}^2[X],$$

or its inner layer

$$\mathrm{List}[\mu_X](xsss) = [[a,b,b,c,a,c,b],\ [c,b,a]] : \mathrm{List}^2[X].$$

When we reduce further, we get for both operations the list

$$[a,b,b,c,a,c,b,c,b,a] : \mathrm{List}[X],$$

that is inner and outer reductions to $\mathrm{List}[X]$ are coequalized by reduction to $\mathrm{List}[X]$.

When are those the same? We have three units into $\mathrm{List}^3[X]$:

outer $\eta_{\mathrm{List}[X]}$, middle $\mathrm{List}[\eta_{\mathrm{List}[X]}]$, inner $\mathrm{List}^3[\eta_X] : \mathrm{List}^2[X] \longrightarrow \mathrm{List}^3[X]$.

In the following example calculation we see that only the middle unit equalizes the inner and outer reductions

$$\begin{tikzcd} & \eta & & \mu \\ & {[[a,b,c],\ [d,e]]} \arrow[r, "\text{outer}"] \arrow[dr, "\text{middle}"] \arrow[dr, bend right, "\text{inner}"'] & {[[a,b,c],[d,e]]} \arrow[r, "\text{outer}"] & {[a,b,c],[d,e]} \\ {[[a,b,c],\ [d,e]]} & & {[[a,b,c]],\ [[d,e]]} \arrow[r, "\text{outer}"] \arrow[d, "\text{inner}"'] & {[a,b,c],[d,e]} \\ & & {[[a],[b],[c]],\ [[d],[e]]} \arrow[r, "\text{outer}"] \arrow[d, "\text{inner}"'] & {[[a],[b],[c],[d],[e]]} \\ & & {[a,b,c],\ [d,e]} \arrow[r, "\text{outer}"] & {[a,b,c],\ [d,e]} \end{tikzcd}$$

This is not a coincidence. ■

Now we have inner and outer version of the unit and reduction on $T^3[X]$

$$\begin{tikzcd} T^3[X] \arrow[r, shift left, "T[\mu_X]"] \arrow[r, shift right, "\mu_{T[X]}"'] & T^2[X] \arrow[l, "T[\eta_X]"'] \arrow[r, shift left, "T[\eta_X]"] \arrow[r, shift right, "\eta_{T[X]}"'] & T[X] \arrow[l, shift right, "\mu_{T[X]}"'] \end{tikzcd}$$

The map $\mu_X : T^3[X] \longrightarrow T[X]$ coequalizes both the units and reductions.

These coequalization properties deserve to be called neutrality and associativity for reasons to be explained later.

**prop** [$\mu$-$\eta$-neutrality]

Let $X : \mathrm{Set}$, then

$\forall t : T[X]$

$$\mu_X \eta_{T[X]}(t) = t = \mu_X T[\eta_X](t)$$

that is

$$\begin{tikzcd} T[X] \arrow[r, "\eta_{T[X]}"] \arrow[dr, "id_{T[X]}"'] & TT[X] \arrow[d, "\mu_X"] & T[X] \arrow[l, "T[\eta_X]"'] \arrow[dl, "id_{T[X]}"] \\ & T[X] & \end{tikzcd}$$

**proof**

$$\mu_X \eta_{T[X]}(t) = id_{T[X]}(\eta_{T[X]}(t))$$
$$= id_{T[X]}(t)$$
$$= t \text{ and}$$

$$\mu_X T[\eta_X](t) = id_{T[X]}(\eta_{T[X]}\eta_X(t))$$
$$= id_{T[X]}(\eta_{T[X]}(t))$$
$$= id_{T[X]}(t)$$
$$= t$$ ■

**prop** [$\mu$-associativity]

$$\begin{tikzcd} TTT[X] \arrow[r, "T[\mu_X]"] \arrow[d, "\mu_{T[X]}"] \arrow[dr, phantom, "="] & TT[X] \arrow[d, "\mu_X"] \\ TT[X] \arrow[r, "\mu_X"'] & T[X] \end{tikzcd}$$

**proof** It is enough to show that

$$\begin{tikzcd} TTT[X] \arrow[r, "T[\mu_X]"] \arrow[d, "id_{T[T[X]]}"] \arrow[dr, phantom, "="] & TT[X] \arrow[d, "\mu_X"] \\ TT[X] \arrow[r, "\mu_X"'] & T[X], \end{tikzcd}$$

then apply $\overline{\ }$.

But

$$\mu_X \overline{id_{T[X]}} \mu_X = \mu_X$$ ■

**prop** Equalization of inner and outer reduction by middle unit

(1) $\mu_{T[X]} T[\eta_{T[X]}] = id_{T[X]} = \mu_{T[X]} T[\eta_{T[X]}]$;

outer $\mu$ outer $\eta$ = id = inner $\mu$ inner $\eta$

(2) $\mu_{T[X]} \eta_{T[T[X]]} = id_{T[X]} = T[\mu_X]\, T^2(\eta_X)$

inner $\mu$ outer $\eta$ = outer $\eta$ $\mu$

(3) $T[\mu_X]\, \eta_{T[X]} = \eta_{T[X]}\, \mu_X$

outer $\mu$ inner $\eta$ = inner $\eta$ $\mu$

(4) $\mu_{T[X]}\, T[\eta_X] = T[\eta_X]\, \mu_X$

**proof** (3) and (4) follow from functoriality of $T$ and $\mu$-$\eta$-neutrality,

(1) from $\eta$-naturality

(2) from $\mu$-naturality ■

Statements of ($\mu$-$\eta$-neutrality) and ($\mu$-associativity) can be reformulated in terms of diagrams of natural transformations.

$$\begin{tikzcd} T \arrow[r, shift left, "\eta T"] \arrow[r, shift right, "T\eta"'] \arrow[dr, shift left, "id"] \arrow[dr, shift right, "id"'] & TT \arrow[d, shift left, "\mu"] \arrow[d, shift right, "id"'] & T \arrow[l, "\eta T"'] \arrow[l, "T\eta"] \arrow[dl, "id"] \\ & T & \end{tikzcd}$$

$$\begin{tikzcd} TTT \arrow[r, shift left, "T\mu"] \arrow[r, shift right, "\mu T"'] \arrow[d, shift left, "\mu T"] \arrow[d, shift right, "\mu T"'] \arrow[dr, phantom, "="] & TT \arrow[d, shift left, "\mu"] \arrow[d, shift right, "\mu"'] \\ TT \arrow[r, shift left, "\mu"] \arrow[r, shift right, "\mu"'] & T \end{tikzcd}$$

We package these laws together with functoriality into a new definition.

**define** [Monad] A monad on Set consists of

a functor $T : \mathrm{Set} \longrightarrow \mathrm{Set}$ with natural transformations

reduction $\mu : T^2 \Longrightarrow T$

unit $\eta : id \Longrightarrow T$

where ($\mu$-$\eta$-neutrality)

$$\begin{tikzcd} T \arrow[r, shift left, "\eta T"] \arrow[r, shift right, "T\eta"'] \arrow[dr, shift left, "id"] \arrow[dr, shift right, "id"'] & TT \arrow[d, shift left, "\mu"] \arrow[d, shift right, "\mu"'] & T \arrow[l, "\eta T"'] \arrow[l, "T\eta"] \arrow[dl, "id"] \\ & T & \end{tikzcd}$$

($\mu$-associativity)

$$\begin{tikzcd} TTT \arrow[r, shift left, "T\mu"] \arrow[r, shift right, "\mu T"'] \arrow[d, shift left, "\mu T"] \arrow[d, shift right, "\mu T"'] \arrow[dr, phantom, "="] & TT \arrow[d, shift left, "\mu"] \arrow[d, shift right, "\mu"'] \\ TT \arrow[r, shift left, "\mu"] \arrow[r, shift right, "\mu"'] & T \end{tikzcd}$$ ■

We have seen that a-$\eta$ functional effect gives rise to a monad, will see that this can be reversed such that monads are just another axiomatization of functional effects.

We illustrate how application can be recovered from a monad

**example** Let $f : X \longrightarrow \mathrm{List}[Y]$. Suppose $a,b,c : X$, $u,v,w : Y$,

$f(a) = [u,v]$, $f(b) = [w,v,v]$, $f(c) = []$.

Then $f([a,b,c]) = [u,v,w,v,v]$.

We can get at this result in terms of reduction and multiplication.

Note

$\mathrm{List}[f]([a,b,c]) = [[u,v],[w,v,v],[]]$,

then $\mu(\mathrm{List}[f]([a,b,c])) = [u,v,w,v,v]$,

more generally $f = \mu_Y \mathrm{List}[f]$ ■

**prop** Let $T$ be a functional effect, $f : X \longrightarrow T[Y]$,

then $\overline{f} = \mu_Y T[f]$.

**proof**

$$\mu_Y T[f] = \overline{id_Y}\,(\eta_{T[Y]} f)$$
$$= \overline{id_Y \eta_{T[Y]}} f$$
$$= \overline{\mu_Y \eta_{T[Y]}} f$$
$$= \overline{f}$$ ■

**prop** Let $T$ be a monad; define

$$\forall f : X \longrightarrow T[Y],$$
$$\bar{f} : T[X] \longrightarrow T[Y]$$
$$\bar{f} := \mu_Y T[f].$$

Then

(1) $T$ with the above defined lift is a functional effect

(2) $\mu_X = \overline{\eta_X}$

**proof** (1) Let $f : X \longrightarrow T[Y]$, $g : Y \longrightarrow T[Z]$.

We need to show (naturality) $\overline{\eta_Y} = id_{T[X]}$, $\bar{f}\,\eta_X = f$

and (associativity) $\overline{g}\,\bar{f} = \overline{gf}$

Now $\overline{\eta_X} = \mu_X T[\eta_X] = id_{T[X]}$ by ($\mu$-$\eta$-naturality)

and $\bar{f}\,\eta_X = \mu_Y T[f]\,\eta_X$ by ($\eta$-naturality)

$= \mu_Y \eta_{T[Y]} f$ by ($\mu$-$\eta$-naturality)

and $= f$.

Also $\overline{g}\,\bar{f} = \mu_Z T[\bar{g}\bar{f}]$ (functoriality)

$= \mu_Z T[\bar{g}]\,T[\bar{f}]$

$= \mu_Z T[\mu_Z T[g]]\,T[f]$ (functoriality)

$= \mu_Z T[\mu_Z]\,T[T[g]]\,T[f]$ ($\mu$-associativity)

$= \mu_Z \mu_{T[Z]}\,T[T[g]]\,T[f]$ ($\mu$-naturality)

$= \mu_Z T[T[g]]\,\mu_Y T[f]$

$= \overline{g}\,\bar{f}$.

(2) Note $\overline{\eta_X} = \mu_X T[\eta_X] = id_{T[X]}$ ($\mu$-$\eta$-naturality)

So $\overline{\overline{\eta_X}} = \overline{id_{T[X]}}$

$= \mu_X T[id_{T[X]}]$

$= \mu_X\, id_{T[X]}$ (functoriality)

$= \mu_X$ ■

So monads and functional effects are the same notion.

We can even define a morphism of functional effects and see that it corresponds to monad morphisms.

**define** A morphism $h$ of functional effects

consists of

$$h : \forall X : Set.$$
$$T[X] \longrightarrow U[X]$$

such (1) $h\,\eta^T_X = \eta^U_X$

(2) $\forall f : X \longrightarrow T[Y]$

$$h\,\bar{f} = \overline{h f}\, h$$

**example** We've already seen the $\{\!|\cdot|\!\}_m$-functional effect.

Is the finiteness of multisubsets important?

Consider arbitrary multisubsets $\{\!|X|\!\}$.

If this is a functional effect existing $\{\!|\cdot|\!\}_{fin}$, it has restriction

$$\mu : \{\!|\{\!|X|\!\}|\!\} \longrightarrow \{\!|X|\!\}.$$

Note that it still makes sense to add any two such multisubsets.

Consider a countable set $X := \{a_0, a_1, a_2, \dots\}$.

We consider a sequence of multisubsets

$$S_1 := \{\!|a_0|\!\},\quad S_2 := \{\!|a_0, a_1|\!\},\quad S_3 := \{\!|a_0, a_1, a_2|\!\}, \dots : \{\!|X|\!\}.$$

What should be the value of

$$\mu\left(\{\!|S_1, S_2, S_3, \dots|\!\}\right) = ?$$

Seems like the result would be a multiset where elements have infinite multiplicity, which is definitely not in $\{\!|X|\!\}$. ■

**example** Let $BinTree[X] := \{$ binary trees with leaves in $X\}$;

denote the empty tree by $() : BinTree[X]$;

given $x : X$, denote the singleton tree again by $x : BinTree[X]$;

given $t, t' : BinTree[X]$ denote the tree

with left branch $t$ and right branch $t'$ by $(t, t') : BinTree[X]$.

Then we visualize $((a, b), (c, d)), (e, f)) : BinTree[X]$ as

We get the obvious $\eta_X : X \longrightarrow BinTree[X]$; given a tree of trees we can reduce it in the obvious way. This defines a monad. ■

# Monoidal Categories

Now we introduce monoidal categories.

We do not define them, we only vaguely describe them.

They will only serve us as a device to motivate specific constructions with monads.

Recall:

**define** A monoid consists of $M : \mathbf{Set}$ with $\times : M \times M \to M$, $1 : M$

where $\forall a, b, c : M$:

$$1 \times a = a = a \times 1$$
$$a \times (b \times c) = (a \times b) \times c \quad \blacksquare$$

To generalize from monoids to monoidal categories, we replace sets by categories — we require an ability to multiply both objects and morphisms.

**define** [Product Category] Given categories $\mathbb{C}, \mathbb{D}$ their product category $\mathbb{C} \times \mathbb{D}$ consists of

objects pairs $\langle c \in \mathbb{C}, d : \mathbb{D} \rangle$

morphisms from $\langle c, d \rangle$ to $\langle c', d' \rangle$ are pairs $\langle f : c \to c', g : d \to d' \rangle$

identities of $\langle c, d \rangle$ is $\langle \mathrm{id}_c, \mathrm{id}_d \rangle$

composition $\langle f', g' \rangle \circ \langle f, g \rangle := \langle f' \circ f, g' \circ g \rangle \quad \blacksquare$

**define** [Bifunctor] Given categories $\mathbb{C}, \mathbb{D}, \mathbb{E}$, a bifunctor $B : \mathbb{C}, \mathbb{D} \to \mathbb{E}$ is just a functor $B : \mathbb{C} \times \mathbb{D} \to \mathbb{E} \quad \blacksquare$

Given such a bifunctor $B : \mathbb{C} \times \mathbb{D} \to \mathbb{E}$,

for each $c : \mathbb{C}$, we get a functor $B(c, \cdot) : \mathbb{D} \to \mathbb{E}$, and

for each $f : c \to c'$, we get a natural transformation $B(f, \cdot) : B(c, \cdot) \Rightarrow B(c', \cdot) : \mathbb{D} \to \mathbb{E}$.

Since we no longer can compare objects for equality in a general category as we could with elements of a set, we need to be more sophisticated in our formulation of monoid laws.

Given a category $\mathbb{E}$, it makes sense to compare two morphisms of the same type $f, g : a \to b : \mathbb{E}$, since $a \to b : \mathbb{E}$ is a set. But it is in general meaningless to state $a = b : \mathbb{E}$, since the objects of $\mathbb{E}$ need not form a set.

Still $\mathbb{E}$ possesses a canonical identity structure on objects, $a = b : \mathbb{E}$ as a true/false valued proposition is replaced by a set $a \cong b : \mathbb{E}$ of isomorphisms (equalities). Hence instead of $a$ and $b$ being equal we talk about having a specific isomorphism $\alpha : a \cong b : \mathbb{E}$.

Certain objects may become canonically equal, like $\mathrm{id}_a : a = a : \mathbb{E}$.

**description** A monoidal category consists of

a category $\mathbb{E}$ with a bifunctor $\otimes : \mathbb{E} \times \mathbb{E} \to \mathbb{E}$

object $\mathbb{1} : \mathbb{E}$

with canonical iso $\forall a, b, c : \mathbb{E}$,

$$l : \mathbb{1} \otimes a \cong a : \mathbb{E},$$
$$r : a \otimes \mathbb{1} \cong a : \mathbb{E},$$
$$\alpha : (a \otimes b) \otimes c \cong a \otimes (b \otimes c) \quad \blacksquare$$

What exactly is meant by canonical we will leave for the reader to find in some appropriate literature; we have no need to delve into details of monoidal categories.

If a monoidal category $\mathbb{E}$, since $\otimes$ is a bifunctor, we can multiply any morphisms $f : a \to b$, $g : a' \to b' : \mathbb{E}$ to get $f \otimes g : a \otimes a' \to b \otimes b' : \mathbb{E}$ in a functorial way

$$(f' f) \otimes (g' g) = (f' \otimes g')(f \otimes g),$$
$$\mathrm{id}_a \otimes \mathrm{id}_b = \mathrm{id}_{a \otimes b}.$$

**example** $\mathbf{Set}$ with cartesian product and 1-element set as unit forms a monoidal category $\blacksquare$

**example** Let $K$ be a field, then $\mathbf{Vec}_K$ with the usual tensor product of vector spaces and $K$ — the 1-dimensional vector space as a unit form a monoidal category $\blacksquare$

**example** $\mathbf{Set}$ with disjoint union and 0-element set as unit form a monoidal category $\blacksquare$

The additional data of $\otimes$ and $\mathbb{1}$ allows us to define new notions on $\mathbb{E}$. These new notions are analogues of what we already know from monoids, but they are formulated with diagrams.

First note how we can formulate the definitions relevant to monoids in terms of diagrams.

**define** A monoid consists of $M : \mathbf{Set}$ with $\mu : M \times M \to M$, $\eta : 1 \to M : \mathbf{Set}$

where

$$\begin{tikzcd} M \times 1 \arrow[dr, "\sim"'] \arrow[rr, "\mathrm{id}_M \times \eta"] & & M \times M \arrow[d, "\mu"] & 1 \times M \arrow[l, "\eta \times \mathrm{id}_M"'] \arrow[dl, "\sim"] \\ & & M & \end{tikzcd} \qquad \begin{tikzcd} M \times M \times M \arrow[r, "\mu \times \mathrm{id}_M"] \arrow[d, "\mathrm{id}_M \times \mu"'] \arrow[dr, phantom, "="] & M \times M \arrow[d, "\mu"] \\ M \times M \arrow[r, "\mu"'] & M \end{tikzcd} \quad \blacksquare$$

**define** A monoid homomorphism consists of $f : M \to N$, where

$$\begin{tikzcd} 1 \arrow[dr, "\eta"'] & M \times M \arrow[r, "f \times f"] \arrow[d, "\mu"] \arrow[dr, phantom, "="] & N \times N \arrow[d, "\mu"] \\ & M \arrow[r, "f"'] & N \end{tikzcd} \quad \blacksquare$$

**define** An $M$-set consists of $A : \mathbf{Set}$ with $\mu : M \times A \to A$

where

$$\begin{tikzcd} 1 \times A \arrow[r, "\eta \times \mathrm{id}_A"] \arrow[dr, "\sim"'] & M \times A \arrow[d, "\mu"] \\ & A \end{tikzcd} \qquad \begin{tikzcd} M \times M \times A \arrow[r, "\mathrm{id}_M \times \mu"] \arrow[d, "\mu \times \mathrm{id}_A"'] \arrow[dr, phantom, "="] & M \times A \arrow[d, "\mu"] \\ M \times A \arrow[r, "\mu"'] & A \end{tikzcd} \quad \blacksquare$$

**define** A $M$-set homomorphism of $M$-sets $A$ and $B$ consists of

$f : A \to B$

where

$$\begin{tikzcd} M \times A \arrow[r, "\mathrm{id}_M \times f"] \arrow[d, "\mu"] \arrow[dr, phantom, "="] & M \times B \arrow[d, "\mu"] \\ A \arrow[r, "f"] & B \end{tikzcd} \quad \blacksquare$$

Notice that the definitions were formulated using only the monoidal structure of $\mathbf{Set}, \times, 1$.

Fix a monoidal category $\mathbb{E}$.

Instead of writing diagrams like

$$\begin{tikzcd} m \otimes \mathbb{1} \arrow[r, "\mathrm{id} \otimes \eta"] \arrow[dr, "r"'] & m \otimes m \arrow[d, "\mu"] \arrow[dl, phantom, "="] \\ & m \end{tikzcd}$$

we'll hide the distinction between $m \otimes \mathbb{1}$ and $m$, so it becomes

$$\begin{tikzcd} m \arrow[r, "\mathrm{id} \otimes \eta"] \arrow[dr, "\mathrm{id}_m"'] & m \otimes m \arrow[d, "\mu"] \arrow[dl, phantom, "="] \\ & m \end{tikzcd}$$

Similarly for associativity.

**define** A monoid object in the monoidal category $\mathbb{E}$ consists of $m : \mathbb{E}$ with $\mu : m \otimes m \to m : \mathbb{E}$,

$\mu : \mathbb{1} \to m : \mathbb{E}$

where

$$\begin{tikzcd} m \arrow[r, "\mathrm{id} \otimes \eta"] \arrow[dr, "\mathrm{id}_m"'] & m \otimes m \arrow[d, "\mu"] \arrow[dl, phantom, "="] & m \otimes m \otimes m \arrow[r, "\mu \otimes \mathrm{id}_m"] \arrow[d, "\mathrm{id}_m \otimes \mu"'] \arrow[dr, phantom, "="] & m \otimes m \arrow[d, "\mu"] \\ & m & m \otimes m \arrow[r, "\mu"'] & m \end{tikzcd} \quad \blacksquare$$

**define** Let $m, n$ be monoid objects; then a homomorphism between them consists of $f : m \to n : \mathbb{E}$

where

$$\begin{tikzcd} \mathbb{1} \arrow[r, "\eta"] \arrow[dr, "\eta"'] & m \arrow[d, "f"] \arrow[dl, phantom, "="] & m \otimes m \arrow[r, "f \otimes f"] \arrow[d, "\mu"] \arrow[dr, phantom, "="] & n \otimes n \arrow[d, "\mu"] \\ & n & m \arrow[r, "f"'] & n \end{tikzcd} \quad \blacksquare$$

Denote the category of monoids in $\mathbb{E}$ by $\mathbf{Mon}(\mathbb{E})$.

**define** Given $m : \mathbf{Mon}(\mathbb{E})$, an $m$-object in $\mathbb{E}$ consists of $a : \mathbb{E}$ with $\mu : m \otimes a \to a : \mathbb{E}$

where

$$\begin{tikzcd} a \arrow[r, "\eta \otimes \mathrm{id}_a"] \arrow[dr, "\mathrm{id}_a"'] & m \otimes a \arrow[d, "\mu"] \arrow[dl, phantom, "="] & m \otimes m \otimes a \arrow[r, "\mu \otimes \mathrm{id}_a"] \arrow[d, "\mathrm{id}_m \otimes \mu"'] \arrow[dr, phantom, "="] & m \otimes a \arrow[d, "\mu"] \\ & a & m \otimes a \arrow[r, "\mu"'] & a \end{tikzcd} \quad \blacksquare$$

We call $\mu$ an $m$-action and say that $m$ acts on $a$ via $\mu$ $\blacksquare$

**define** Let $a, b$ be $m$-objects in $\mathbb{E}$; a homomorphism between them consists of $f : a \to b : \mathbb{E}$ where

$$\begin{tikzcd} m \otimes a \arrow[r, "\mathrm{id} \otimes f"] \arrow[d, "\mu"] \arrow[dr, phantom, "="] & m \otimes b \arrow[d, "\mu"] \\ a \arrow[r, "f"'] & b \end{tikzcd} \quad \blacksquare$$

We'll need a slight generalization of this; we will act by $m : \mathbb{E}$ on objects of some different category $\mathcal{X}$.

To make this possible we need a few additional notions.

There are many types of morphisms between monoidal categories, by a monoidal functor we mean what is usually known as a strong monoidal functor.

**description** Let $\mathbb{E}$ and $\mathbb{F}$ be monoidal categories.

A monoidal functor consists of a functor $F : \mathbb{E} \to \mathbb{F}$

with canonical iso

$$i : F[\mathbb{1}] \cong \mathbb{1}$$

$\forall a, b : \mathbb{E}$,

$$t : F[a] \otimes F[b] \cong F[a \otimes b] \quad \blacksquare$$

It is not hard to see that monoidal functor $F : \mathbb{E} \to \mathbb{F}$ extends to a functor $F : \mathbf{Mon}(\mathbb{E}) \to \mathbf{Mon}(\mathbb{F})$.

**description** Let $\mathbb{E}$ be a monoidal category, $\mathcal{X}$ a category.

Then an action of $\mathbb{E}$ on $\mathcal{X}$ consists of

a bifunctor $\otimes : \mathbb{E} \times \mathcal{X} \to \mathcal{X}$

with canonical iso

$\forall a, b : \mathbb{E}$, $X : \mathcal{X}$,

$$\mathbb{1} \otimes X \cong X : \mathcal{X}$$
$$(a \otimes b) \otimes X \cong a \otimes (b \otimes X) : \mathcal{X} \quad \blacksquare$$

**define** Let monoidal category $\mathbb{E}$ act on a category $\mathcal{X}$ via $\otimes$.

Let $m : \mathbf{Mon}(\mathbb{E})$, then $m$-object in $\mathcal{X}, \otimes$ consists of $A : \mathcal{X}$ with $\mu : m \otimes A \to A : \mathcal{X}$

where

$$\begin{tikzcd} A \arrow[r, "\eta \otimes \mathrm{id}_A"] \arrow[dr, "\mathrm{id}_A"'] & m \otimes A \arrow[d, "\mu_A"] \arrow[dl, phantom, "="] & m \otimes m \otimes A \arrow[r, "\mu \otimes \mathrm{id}_A"] \arrow[d, "\mu \otimes \mathrm{id}_A"'] \arrow[dr, phantom, "="] & m \otimes A \arrow[d, "\mu"] \\ & A & m \otimes A \arrow[r, "\mu"'] & A \end{tikzcd} \quad \blacksquare$$

**define** Given $A, B$ $m$-objects in $\mathcal{X}, \otimes$, a homomorphism between them consists of $f : A \to B : \mathcal{X}$

where

$$\begin{tikzcd} m \otimes A \arrow[r, "\mathrm{id}_m \otimes f"] \arrow[d, "\mu_A"'] \arrow[dr, phantom, "="] & m \otimes B \arrow[d, "\mu"] \\ A \arrow[r, "f"'] & B \end{tikzcd} \quad \blacksquare$$

This defines the category $m \otimes \mathcal{X}$ of $m$-objects in $\mathcal{X}$, it is also known as the Eilenberg-Moore Category.

**example** A monoid object $M$ in $\mathbf{Set}$ with cartesian product is just the usual monoid.

Monoidal category $\mathbf{Set}$ acts on itself via cartesian product. But the notation $M \times \mathbf{Set}$ for the category of $M$-sets is a bit unfortunate, since $M$ can be thought of as a 1-object category. For that reason we denote the action by $\triangleleft$, so it becomes $M \triangleleft \mathbf{Set}$ $\blacksquare$

**example** A monoid object in $\mathbf{Vec}_K, \otimes, K$ is exactly an associative algebra $M$ over $K$.

An $M$-object is just an $M$-module $\blacksquare$

**example** The functor $\mathrm{vec}_K : \mathbf{Set} \to \mathbf{Vec}_K$ that associates to $X : \mathbf{Set}$ the free vector space $\mathrm{vec}_K[X] : \mathbf{Vec}_K$ is monoidal; recall $\mathrm{vec}_K[X \times Y] = \mathrm{vec}_K[X] \otimes \mathrm{vec}_K[Y]$.

Given a monoid $M$ in $\mathbf{Set}$, the monoid $\mathrm{vec}_K[M]$ in $\mathbf{Vec}_K$ is the so called monoid algebra of $M : \mathbf{Set}$.

Its $\mathrm{vec}_K[M]$-objects are the linear representations of $M$ over $K$, that is they are vector spaces $V : \mathbf{Vec}_K$ with a monoid action $\times : M \times V \to V : \mathbf{Set}$ such that for each $m : M$ the map $V \to V$, $v \mapsto m \times v$ is $K$-linear $\blacksquare$

Now for our most relevant example.

**example** From each set $X$ we can construct a monoid $\mathrm{End}(X) := \{ f : X \to X \}$ of endomaps with composition as multiplication, and $\mathrm{id}$ as the neutral element.

Then given a monoid $M$, an $M$-action $\mu : M \times X \to X$ corresponds bijectively to monoid homomorphisms

$$\mu : M \to \mathrm{End}(X).$$

Since $\mathrm{id}_{\mathrm{End}(X)} : \mathrm{End}(X) \to \mathrm{End}(X)$ is a monoid homomorphism, we have the obvious action of $\mathrm{End}(X)$ on $X$

$$\mathrm{End}(X) \times X \to X$$
$$f, x \mapsto f(x).$$

This fully generalizes to monoidal categories. $\blacksquare$

**example** Let $\mathcal{X}$ be a category; we define a new (huge) category of endofunctors

**define** Category $\mathrm{End}(\mathcal{X})$ consists of

objects $F : \mathcal{X} \to \mathcal{X}$ functors

morphisms $\alpha : F \Rightarrow G$ natural transformations

composition for $F \xrightarrow{\alpha} G \xrightarrow{\beta} H : \mathcal{X} \to \mathcal{X}$ via

$$\beta \alpha : F \to G : \mathcal{X} \to \mathcal{X}$$
$$(\beta \alpha)_X = \beta_X \alpha_X : F[X] \to G[X] \quad \blacksquare$$

Now $\mathrm{End}(\mathcal{X})$ is strictly speaking too big to be a category, since for fixed functors $F, G$ the natural transformations $F \Rightarrow G : \mathcal{X} \to \mathcal{X}$ need not form a set; but this is not a big deal, since boolean valued equality $\alpha = \beta : F \Rightarrow G : \mathcal{X} \to \mathcal{X}$ still makes sense:

$$\alpha = \beta \iff \forall X : \mathcal{X},$$
$$\alpha_X = \beta_X : F[X] \to G[X].$$

The category of endofunctors has a canonical monoidal structure.

**define** Given $G, F : \mathrm{End}(\mathcal{X})$ we can multiply by just composition

$$G F : \mathrm{End}(\mathcal{X}).$$

The unit is $\mathrm{id}_{\mathcal{X}} : \mathrm{End}(\mathcal{X})$

Multiplication of $\beta : F \to F'$ with $\alpha : G \to G'$ is defined to be

$$\beta \otimes \alpha : F G \to F' G'$$
$$(\beta \otimes \alpha)_X = \beta_{G[X]} F[\alpha_X] = F[\alpha_X] \beta_{G[X]}$$

For convenience when we have

$$\mathrm{id}_F : F \Rightarrow F, \ \alpha : G \to G'$$

we write $F \alpha := \mathrm{id}_F \otimes \alpha : F G \to F G'$

since $(F \alpha)_X = F[\alpha_X]$

and for $\beta : F \Rightarrow F'$, $\mathrm{id}_G : G \to G$

we write $\beta G := \beta \otimes \mathrm{id}_G : F G \to F' G$

since $(\beta G)_X = \beta_{G[X]}$.

This makes $\mathrm{End}(\mathcal{X})$ into a monoidal category $\blacksquare$

Given a monoidal $\mathbb{E}$, then an $\mathbb{E}$-action $\otimes$ on $\mathcal{X}$ is the same thing as a monoidal functor $\mathcal{M} : \mathbb{E} \to \mathrm{End}(\mathcal{X})$

**define** We have an action of $\mathrm{End}(\mathcal{X})$ on $\mathcal{X}$:

given $F : \mathrm{End}(\mathcal{X})$, $X : \mathcal{X}$, we have $F[X] : \mathcal{X}$,

given $F : \mathrm{End}(\mathcal{X})$, $f : X \to Y : \mathcal{X}$, we have $F[f] : F[X] \to F[Y] : \mathcal{X}$

given $F \xrightarrow{\alpha} G : \mathrm{End}(\mathcal{X})$, $X : \mathcal{X}$, we have $\alpha_X : F[X] \to G[X] : \mathcal{X}$

This defines a bifunctor in $\mathrm{End}(\mathcal{X}) \times \mathcal{X} \to \mathcal{X}$.

Explicitly, given $F \xrightarrow{\alpha} G : \mathrm{End}(\mathcal{X})$, $f : X \to Y$,

we have

$$\alpha \otimes f : F[X] \to G[Y] : \mathcal{X},$$
$$\alpha \otimes f := G[f] \alpha_X = \alpha_Y F[f].$$

This is precisely the diagonal of the naturality square

$$\begin{tikzcd} F[X] \arrow[r, "\alpha_X"] \arrow[d, "F[f]"'] \arrow[dr, "\alpha \otimes f"] & G[X] \arrow[d, "G[f]"] \\ F[Y] \arrow[r, "\alpha_Y"'] & G[Y] \end{tikzcd} \quad \blacksquare$$

Note that a monoid object $\mathbb{T}$ in $\mathrm{End}(\mathcal{X})$ is an endofunctor $\mathbb{T} : \mathcal{X} \to \mathcal{X}$ with $\eta : \mathrm{id}_{\mathcal{X}} \to \mathbb{T}$, $\mu : \mathbb{T}^2 \Rightarrow \mathbb{T}$

where

$$\begin{tikzcd} \mathbb{T} \arrow[r, "\mathbb{T} \otimes \eta = \mathbb{T}\eta"] \arrow[dr, "\mathrm{id}_{\mathbb{T}}"'] & \mathbb{T}^2 \arrow[d, "\mu"] \arrow[dl, phantom, "="] & \mathbb{T} \arrow[l, "\eta \otimes \mathbb{T} = \eta\mathbb{T}"'] \arrow[dl, "\mathrm{id}_{\mathbb{T}}"] \\ & \mathbb{T} & \end{tikzcd}$$

$$\begin{tikzcd} \mathbb{T}^3 \arrow[r, "\mathrm{id}_{\mathbb{T}} \otimes \mu = \mathbb{T}\mu"] \arrow[d, "\mu \otimes \mathrm{id}_{\mathbb{T}} = \mu\mathbb{T}"'] \arrow[dr, phantom, "="] & \mathbb{T}^2 \arrow[d, "\mu"] \\ \mathbb{T}^2 \arrow[r, "\mu"'] & \mathbb{T} \end{tikzcd}$$

i.e. a monad on $\mathcal{X}$!

Given $\mathbb{T} : \mathbb{M}\mathrm{on}(\mathrm{End}(\mathcal{X}))$ i.e. a monad on $\mathcal{X}$,

a $\mathbb{T}$-object in $\mathcal{X}$ (via the canonical action of $\mathrm{End}(\mathcal{X})$ on $\mathcal{X}$)

consists of $A : \mathcal{X}$ with $\mu : \mathbb{T}[A] \longrightarrow A$

where
$$\begin{tikzcd} A \arrow[r, "\eta_A"] \arrow[dr, "=", "id_A"'] & \mathbb{T}[A] \arrow[d, "\mu"] \\ & A \end{tikzcd} \qquad \begin{tikzcd} \mathbb{T}[A] \arrow[r, "\mathbb{T}[\mu]"] \arrow[d, "\mu_2"'] \arrow[dr, phantom, "="] & \mathbb{T}[A] \arrow[d, "\mu"] \\ \mathbb{T}[A] \arrow[r, "\mu"'] & A \end{tikzcd}$$

And a homomorphism between $\mathbb{T}$-objects $A$ and $B$ in $\mathcal{X}$ is

a map $f : A \longrightarrow B : \mathcal{X}$ where
$$\begin{tikzcd} \mathbb{T}[A] \arrow[r, "\mathbb{T}[f]"] \arrow[d, "\mu"'] \arrow[dr, phantom, "="] & \mathbb{T}[B] \arrow[d, "\mu"] \\ A \arrow[r, "f"'] & B \end{tikzcd}$$

We denote the resulting category by $\mathbb{T}\mathcal{C}\mathcal{X}$ or even $\mathbb{T}[\mathcal{X}]$, we also call the objects of $\mathbb{T}\mathcal{C}\mathcal{X}$ the $\mathbb{T}$-models – the reason for this terminology will be explained later when we relate algebraic theories to monads. ■

To summarise we've started from monoids $M$ on Sets and their actions on sets $A$, then generalized them to monoidal categories $\mathbb{E}$ in $\mathcal{CAT}$ and actions on categories $\mathcal{X}$.

This allowed us to define monoid objects $m$ in $\mathbb{E}$ and their actions on objects in $\mathcal{X}$.

As a special case we've recovered monads $\mathbb{T}$ on $\mathcal{X}$ as monoids in $\mathrm{End}(\mathcal{X})$.

In the next chapter we'll concern ourselves with the category of $\mathbb{T}$-models in Set for a monad $\mathbb{T} : \mathrm{Set} \longrightarrow \mathrm{Set}$.

We've already seen that monads are the same thing as functional effects.

So now we can talk about the category of effects $\mathrm{Eff}[\mathbb{T}]$ of a monad $\mathbb{T}$.

The objects were sets $X, Y, \ldots$, but the maps where set functions enriched with a $\mathbb{T}$-effect $f : X \longrightarrow \mathbb{T}[Y]$.

Now we'll see that this effect notion can be defined in a greater generality.

Fix $\mathbb{E}$ monoidal category acting on $\mathcal{X}$ via $\otimes$.

Given $\alpha : a \longrightarrow b : \mathbb{E}$, $f : X \longrightarrow Y : \mathcal{X}$ we have,

$$\alpha \otimes f : a \otimes X \longrightarrow b \otimes Y : \mathcal{X}.$$

This is bifunctorial, meaning

for $a \xrightarrow{\alpha} b \xrightarrow{\beta} c : \mathbb{E}$, $X \xrightarrow{f} Y \xrightarrow{g} Z : \mathcal{X}$ we have,

$$(\beta \circ \alpha) \otimes (g \circ f) = (\beta \otimes g) \circ (\alpha \otimes f)$$
$$id_a \otimes id_X = id_{a \otimes X}.$$

For $a \xrightarrow{id_a} a : \mathbb{E}$, $X \xrightarrow{f} Y : \mathcal{X}$ and $a \xrightarrow{\alpha} b : \mathbb{E}$, $X \xrightarrow{id_X} X : \mathcal{X}$

we write

$$a \otimes f = id_a \otimes f : a \otimes X \longrightarrow a \otimes Y : \mathcal{X},$$
$$\alpha \otimes X = \alpha \otimes id_X : a \otimes X \longrightarrow b \otimes X : \mathcal{X}.$$

Then since $id_b \circ \alpha = \alpha = \alpha \circ id_a$ and $id_Y \circ f = f = f \circ id_X$,

we conclude

$$\alpha \otimes f = (\alpha \otimes Y) \circ (a \otimes f) = (b \otimes f) \circ (\alpha \otimes X).$$

Back to effects. Fix $m : \mathbb{M}on(\mathbb{E})$

**define** Category of $m$-effects on $\mathcal{X}$ [Kleisli Category] $\mathrm{Eff}[m \otimes \mathcal{X}]$

objects $X : \mathcal{X}$

morphisms $f : X \longrightarrow Y : \mathrm{Eff}[m \otimes \mathcal{X}]$

is just any $f : X \longrightarrow m \otimes Y$

identity $\eta_X : X \longrightarrow m \otimes X$

composition Given $f : X \longrightarrow m \otimes Y$, $g : Y \longrightarrow m \otimes Z$,

define

$$g \circ f = (\mu \otimes Z) \circ (m \otimes g) \circ f \quad ■$$

To check that this is indeed a category, define

**define** $\forall f : X \longrightarrow m \otimes Y : \mathcal{E}$.

$$\bar{f} = (\mu \otimes X) \circ (m \otimes f) : m \otimes X \longrightarrow m \otimes Y \quad ■$$

and using from $m$, a functor $m \otimes : \mathcal{X} \longrightarrow \mathcal{X}$,

and from

$$\eta : 1 \longrightarrow m : \mathbb{E},$$
$$\mu : m \circ m \longrightarrow m : \mathbb{E}$$

natural transformation

$$\eta_{\mathcal{X}} \otimes : id_{\mathcal{X}} \Longrightarrow m \otimes \mathcal{X} \longrightarrow \mathcal{X}$$
$$\mu \otimes : m \circ m \otimes \Longrightarrow m \otimes \mathcal{X} \longrightarrow \mathcal{X},$$

together with

$$\begin{tikzcd} m \otimes X \arrow[rr, "\eta \otimes m \otimes X"] \arrow[dr, "id_{m \otimes X}"'] & & m \otimes m \otimes X \arrow[d, "\mu \otimes X"] & & m \otimes \eta \otimes X \arrow[ll, ""'] \arrow[dl, "id_{m \otimes X}"] \\ & & m \otimes X & & \end{tikzcd},$$

$$\begin{tikzcd} m \otimes m \otimes m \otimes X \arrow[r, "\mu \otimes m \otimes X"] \arrow[d, "m \otimes \mu \otimes X"'] & m \otimes m \otimes X \arrow[d, "\mu \otimes X"] \\ m \otimes m \otimes X \arrow[r, "\mu \otimes X"'] & m \otimes X \end{tikzcd}$$

we can show the familiar

**prop** For $f : X \longrightarrow m \otimes Y$, $g : Y \longrightarrow m \otimes Z$, $h : Z \longrightarrow m \otimes W$

(1) $\overline{\eta \otimes X} = id_{m \otimes X}$

(2) $\overline{\overline{\eta \otimes X}} = \mu \otimes X$

(3) $\bar{f} \circ (\eta \otimes X) = f$

(4) $\overline{\bar{g} \circ f} = \bar{g} \circ \bar{f}$

(5) $g \circ f = \bar{g} \circ f$

(6) $f \circ (\eta \otimes X) = f = (\eta \otimes Y) \circ f$

(7) $h \circ (g \circ f) = (h \circ g) \circ f$ ■

**example** Consider Set with $\times, 1$. Let $M : \mathbb{M}on(\mathrm{Set})$.

We already know that $M\mathcal{C}\mathrm{Set}$ is the category of sets with $M$-actions.

What is $\mathrm{Eff}[M\mathcal{C}\mathrm{Set}]$?

Objects are sets, morphisms are maps $X \xrightarrow{f} M \times Y$.

Then we have

(1) the unit

$$\eta_X : X \longrightarrow M \times X$$
$$\eta_X(x) := \langle 1, x \rangle$$

(2) the reduction

$$\mu_X : M \times M \times X \longrightarrow M \times X$$
$$\mu_X(m, n, x) := \langle mn, x \rangle$$

(3) for $f : X \longrightarrow Y$, the functor application

$$Mf : MX \longrightarrow M \times Y$$
$$(Mf)(m, x) := \langle m, f(x) \rangle$$

(4) for $f : X \longrightarrow M \times Y$ the lift

$$\bar{f} : M \times X \longrightarrow M \times Y$$
$$\bar{f}(m, x) := \langle mn, y \rangle \text{ where } \langle n, y \rangle := f(x)$$

(5) for $X \xrightarrow{f} M \times Y$, $Y \xrightarrow{g} M \times Z$, the composition

$$g \circ f : X \longrightarrow M \times Z$$
$$(g \circ f)(x) = \langle mn, z \rangle \text{ where } \langle m, y \rangle = f(x),$$
$$\langle n, z \rangle = g(y).$$

This obviously defines a monad on Set which we denote $M \times : \mathrm{Set} \to \mathrm{Set}$.

So each monoid on Set leads to a monad.

Actually we can think of monoids as special monads.

We have the category of monoids $\mathrm{Mon}$ on Set,

and the category of monads $\mathrm{Monad}$ on Set.

**prop** $\mathrm{Mon}$ is embedded in $\mathrm{Monad}$

**proof** (1) Let $f : M \longrightarrow N$ be a monoid homomorphism.

We have a natural transformation

$$f \times : M \times \Longrightarrow N \times : \mathrm{Set} \to \mathrm{Set}$$

Obviously $f(\eta^M_X(x)) = f(1, x) = (f(1), x) = (1, x) = \eta^N_X(x)$,

and $f(\mu^M_X(m, n, x)) = f(mn, x) = (f(mn), f(n), x) = \mu^N_X(f(m), f(n), x)$,

so its a monad morphism.

(2) Let $f : M \times \Longrightarrow N \times : \mathrm{Set} \to \mathrm{Set}$ be a monad morphism.

Then from $f_1 : M \times 1 \longrightarrow M \times 1$ we can recover the underlying monoid homomorphism.

(3) These constructions are obviously mutually inverse. ■

Note that $M \times X$ has obvious $M$-action

$$\times : M \times (M \times X) \longrightarrow M \times X$$
$$m \times \langle n, x \rangle := \langle mn, x \rangle$$

We can think of each of $X : \mathrm{Eff}[M\mathcal{C}\mathrm{Set}]$ as an $M$-set $M \times X : M\mathcal{C}\mathrm{Set}$.

This extends to a functor since given $f : X \longrightarrow M \times Y$, that is $f : X \longrightarrow Y : \mathrm{Eff}[M\mathcal{C}\mathrm{Set}]$,

the lift $\bar{f} : M \times X \longrightarrow M \times Y$ is an $M$-set homomorphism

since

$$\bar{f}(mn', x) = \langle mn'n, y \rangle \qquad \text{where } \langle n, y \rangle = f(x)$$
$$= m \times \langle n'n, y \rangle$$
$$= m \times \bar{f}(n', y)$$

Actually this functor is an embedding:

Let $g : M \times X \longrightarrow M \times Y : M\mathcal{C}\mathrm{Set}$.

Define $f : X \longrightarrow M \times Y : \mathrm{Set}$

$$f(x) = g(1, x)$$

Then $g = \bar{f}$,

since

$$\bar{f}(m, x) = \bar{f}(m1, x)$$
$$= \langle mn, y \rangle \qquad \text{where } \langle n, y \rangle = g(1, x)$$
$$= m \times \langle n, y \rangle$$
$$= m \times g(1, x)$$
$$= g(m, x)$$

In fact $M \times X : M\mathcal{C}\mathrm{Set}$ with $\eta_X : X \longrightarrow M \times X$ is the free $M$-set over set $X$.

**prop** Fix $X : \mathrm{Set}$, then

$\forall Y : M\mathcal{C}\mathrm{Set}$.

$\forall f : X \longrightarrow Y : \mathrm{Set}$.

$\exists ! \bar{f} : M \times X \longrightarrow Y : M\mathcal{C}\mathrm{Set}$.

$\bar{f}(1, x) = f(x)$

**proof** Let $Y : M\mathcal{C}\mathrm{Set}$, $f : X \longrightarrow Y : \mathrm{Set}$.

We are forced to define

$$\bar{f}(m, x) := m \times f(1, x)$$

this is an $M$-set homomorphism, since

$$\bar{f}(mn, x) = mn \times f(1, x)$$
$$= m \times \bar{f}(n, x). \quad ■$$

So we can think of $\mathrm{Eff}[M\mathcal{C}\mathrm{Set}]$ as the category of free $M$-sets and $M$-set homomorphisms.

Since we have

$\forall X : \mathrm{Set}, Y : M\mathcal{C}\mathrm{Set}.$

$$\frac{M \times X \longrightarrow Y : M\mathcal{C}\mathrm{Set}}{X \longrightarrow Y : \mathrm{Set}} \quad ■$$

**example** The models of multiset $\{\!|\cdot|\!\}_{fin}$-monad are exactly the commutative monoids — the only difference between List and $\{\!|\cdot|\!\}_{fin}$ is that for $\{\!|\cdot|\!\}_{fin}$ order is ignored so $\{\!|a,b|\!\} = \{\!|b,a|\!\}$, but $[a,b] \neq [b,a]$, which translates into $ab = ba$ ■

**example** The models of the finite powerset $\mathcal{P}_{fin}$-monads are commutative idempotent monoids; that is we require $\{\!|a,a|\!\} = \{\!|a|\!\}$, which translates into

We have $a^2 = a$.

$$\mathcal{P}_{fin} \mathbb{C} Set \hookrightarrow \{\!|\cdot|\!\}_{fin} \mathbb{C} Set \hookrightarrow L_{fin} \mathbb{C} Set$$

Commutative idempotent monoids are also known as join (or meet) semilattices.

**define** A join semilattice consists of

$A : Set$ with $\vee : A \times A \to A$,
$\bot : A$

where $a \vee (b \vee c) = (a \vee b) \vee c$
$a \vee \bot = a = \bot \vee a$
$a \vee b = b \vee a$
$a \vee a = a$ ■

A join semilattice homomorphism preserves $\vee$ and $\bot$.

We can call these algebraic join semilattices.

We also have

**define** A poset join semilattice consists of a

$A : Set$ with $\leq$ a partial order
$\vee : A \times A \to A$
$\bot : A$

where $\bot \leq a$
$a \vee b$ is the least upper bound of $a$ and $b$ ■

It is easy to see that a poset join semilattice induces an algebraic join semilattice.

The surprise is that converse also holds.

**define** Let $A$ be algebraic join semilattice.

Define a relation $\leq$ on $A$

$a \leq b$ :iff $a \vee b = b$ ■

**prop** Let $A$ be algebraic join semilattice.

(1) $\bot \leq a$

(2) $a \vee b$ is the least upper bound of $a$ and $b$ ■

(3) Every $f : A \to B$ semi-lattice homomorphism is monotone that is

if $a \leq a' : A$ then $f(a) \leq f(a') : B$

**proof** (1) $\bot \vee a = a$

(2) Note $a, b \leq a \vee b$, since $a \vee (a \vee b) = a \vee a \vee b = a \vee b$

similarly for $b$.

Let $a, b \leq u$. We need to show that $a \vee b \leq u$.

Since $a \vee u = u = b \vee u$, we get

$$\begin{aligned}(a \vee b) \vee u &= a \vee (b \vee u) \\ &= a \vee u \\ &= u.\end{aligned}$$

(3) Let $a \leq a' : A$, then $f(a) \vee f(a') = f(a \vee a') = f(a')$.

So $f(a) \leq f(a')$ ■

The posets with top and bottom elements such that any two elements have both least upper bound and greatest upper bound are called lattices, and these too can be described purely in terms of operations and equations ■

**example** We've also seen the Powerset $\mathcal{P}$-monad.

What are its algebras?

Let $A : \mathcal{P}\mathbb{C}Set$. Obviously $A : \mathcal{P}_{fin}\mathbb{C}Set$

i.e. it is a semi-lattice.

But we have a reduction for all (not just finite) subsets,

$$\mu : \mathcal{P}[A] \longrightarrow A$$

Note $a \vee b = \mu(\{\!|a,b|\!\})$, $\bot = \mu(\{\}\!|)$.

For this reason we denote

$$\bigvee_{a \in S} a = \mu(S)$$

More generally for $a_i : I \longrightarrow A : Set$,

$$\bigvee_{i : I} a_i := \mu(\{\!|a_i \mid i : I|\!\})$$

Associativity of $\mu$ states that for $\mathcal{B} \in \mathcal{P}[A]$,

$$\bigvee_{a : \bigcup_{S \in \mathcal{B}} S} a = \bigvee_{S \in \mathcal{B}} \bigvee_{a \in S} a$$

For example

$$a_1 \vee a_2 \vee \cdots \vee b_1 \vee b_2 \vee \cdots \vee c_1 \vee c_2 \vee \cdots = (a_1 \vee a_2 \vee \cdots) \vee (b_1 \vee b_2 \vee \cdots) \vee (c_1 \vee c_2 \vee \cdots)$$

**prop** $\bigvee_{a \in S} a$ is the least upper bound of $S$.

**proof** (1) Let $a_0 : S$, then

$$a_0 \vee \bigvee_{a \in S} a = \bigvee_{a \in S \cup \{a_0\}} a = \bigvee_{a \in S} a \quad \text{since } S \cup \{a_0\} = S,$$

so $a_0 \leq \bigvee_{a \in S} a$.

(2) Let $u$ be such that $\forall a \in S \;\; a \leq u$.

Define $\mathcal{B} := \{\!|\{a\} \mid a : S|\!\} : \mathcal{P}^2[A]$. Then $\bigcup_{S \in \mathcal{B}} S = S \cup \{\}$

So

$$\begin{aligned}\bigvee_{a \in S} a \vee u &= \bigvee_{a \in S \cup \{u\}} a \\ &= \bigvee_{a \in S} (a \vee u) \\ &= \bigvee_{a \in S} u \\ &= u\end{aligned}$$ ■

So every $\mathcal{P}$-model is a complete join semilattice.

The converse holds also.

But if a poset has all joins, it automatically has all meets.

So the objects of $\mathcal{P}\mathbb{C}Set$ are complete lattices.

Yet it's all about morphisms!

Morphisms in $\mathcal{P}\mathbb{C}Set$ preserve all joins, but they do not have to preserve meets.

As we have seen already $Eff(\mathcal{P}\mathbb{C}Set)$ is the category of sets and relations.

For $X : Set$, the powerset $\mathcal{P}[X]$ with union is the free join-semilattice over $X$. ■

# Monads and Algebraic Theories

We finally relate algebraic theories with monads.

To do that we need another important class of monads coming from adjunctions.

## Adjunctions and Monads

Let $\mathbb{E}, \mathbb{S}$ be categories, $\mathbb{S} \underset{\mathcal{U}}{\overset{\mathcal{T}}{\rightleftarrows}} \mathbb{E}$ be an adjoint situation.

It is advantageous to think about $\mathbb{S}$ as the category of sets, $\mathbb{E}$ as a category of models of some monad, $\mathcal{U}$ as the forgetful functor, and $\mathcal{T}$ as the functor that to each set assigns its free model.

Informally this means that we have a canonical correspondence

$$\forall X : \mathbb{S},\ A : \mathbb{E}$$
$$f : X \longrightarrow \mathcal{U}(A) : \mathbb{S}$$
$$\bar{f} : \mathcal{T}(X) \longrightarrow A : \mathbb{E}.$$

From $X : \mathbb{S}$, $\mathcal{T}(X) : \mathbb{E}$, we have the unit

$$\frac{\eta_X : X \longrightarrow \mathcal{U}\mathcal{T}(X) : \mathbb{S}}{\mathrm{id}_{\mathcal{T}(X)} : \mathcal{T}(X) \longrightarrow \mathcal{T}(X) : \mathbb{E}},$$

from $\mathcal{U}(A) : \mathbb{S}$, $A : \mathbb{E}$, we have the counit

$$\frac{\varepsilon_A : \mathcal{U}(A) \longrightarrow \mathcal{U}(A) : \mathbb{S}}{\mu_A : \mathcal{T}\mathcal{U}(A) \longrightarrow A : \mathbb{E}} \quad [\text{often denoted } \varepsilon_A].$$

Note that this is just the analogue of the familiar

$$\overline{\eta_X} = \mathrm{id}_{\mathcal{T}(X)}, \quad \overline{\varepsilon_A} = \mu_A,$$

but instead we need to keep track of categories, so

$$\overline{\eta_X} = \mathrm{id}_{\mathcal{T}(X)}, \quad \overline{\varepsilon_A} = \mu_A.$$

Then $\mu_X$ is to be thought of as a homomorphism, whose underlying reduction is $\mathcal{U}(\mu_X)$; so we set $\mu_X := \mathcal{U}(\mu_{\mathcal{T}(X)}) : \mathcal{U}\mathcal{T}\mathcal{U}\mathcal{T}(X) \longrightarrow \mathcal{U}\mathcal{T}(X)$.

These satisfy analogous of neutrality

$$\mathcal{U}(\mu_A)\,\eta_{\mathcal{U}(A)} = \mathrm{id}_{\mathcal{U}(A)} \quad \text{and} \quad \mu_{\mathcal{T}(X)}\,\mathcal{T}[\eta_X] = \mathrm{id}_{\mathcal{T}(X)}.$$

In total we have natural transformations

$$\mu : \mathcal{T}\mathcal{U} \Longrightarrow \mathrm{id}_{\mathbb{E}}, \quad \eta : \mathrm{id}_{\mathbb{S}} \Longrightarrow \mathcal{U}\mathcal{T}$$

where

$$\begin{tikzcd} \mathcal{U} \arrow[r, Rightarrow, "\eta\mathcal{U}"] \arrow[dr, Rightarrow, "\mathcal{U}"'] & \mathcal{U}\mathcal{T}\mathcal{U} \arrow[d, Rightarrow, "\mu\mathcal{U}"] \\ & \mathcal{U} \end{tikzcd} \qquad \begin{tikzcd} \mathcal{T}\mathcal{U}\mathcal{T} \arrow[d, Rightarrow, "\mu\mathcal{T}"'] & \mathcal{T} \arrow[l, Rightarrow, "\mathcal{T}\eta"'] \arrow[dl, Rightarrow, "\mathcal{T}"] \\ \mathcal{T} & \end{tikzcd},$$

From these we can derive a functor associativity

$$\mu_X\,\mu_{\mathcal{T}\mathcal{U}(X)} = \mu_X\,\mathcal{T}\mathcal{U}[\mu_X];$$

which gives us

$$\mu_X\,\eta_{\mathcal{U}\mathcal{T}(X)} = \mathrm{id}_{\mathcal{U}\mathcal{T}(X)} = \mu_X\,\mathcal{U}\mathcal{T}[\eta_X]$$

and

$$\mu_X\,\mu_{\mathcal{U}\mathcal{T}(X)} = \mu_X\,\mathcal{U}\mathcal{T}[\mu_X]$$

So $\mathbb{T} := \mathcal{U}\mathcal{T} : \mathbb{S} \longrightarrow \mathbb{S}$ becomes a monad on $\mathbb{S}$.

**example** Let $\mathbb{T}$ be a finitary single sorted algebraic theory.

Then we have the adjunction

$$\mathbf{Set} \underset{\mathcal{U}}{\overset{\mathcal{T}}{\rightleftarrows}} \mathbb{T}\text{-}\mathbf{Set};$$

the monad of this adjunction is also denoted by $\mathbb{T}$.

It is special, since the underlying functor is finitary.

In fact finitary monads precisely correspond to finitary single sorted algebraic theories. We will see this in a little while. ■

From a monad $\mathbb{T}$ we define its finitary algebraic theory $\mathbb{T}_{fin}(\mathbb{T})$

## Monad Operations

As we have seen in all the examples, monads seem to give rise to operations.

**example** What's a List operation?

Consider the operation for monads

$$M \times M \longrightarrow M$$
$$\langle a, b \rangle \rightsquigarrow b\,a\,a$$

We can think of it as living in

$$[y, x, x] : \mathrm{List}[x, y].$$

Similarly

$$[x, y, z, z, y] : \mathrm{List}[x, y, z, w]$$

corresponds to

$$M^{\times 4} \longrightarrow M$$
$$\langle a, b, c, d \rangle \rightsquigarrow a\,b\,c\,c\,b.$$

Can we produce an interesting infinitary operation? Nope.

Consider operations of countable arity

$$M^{\widetilde{\times}} \longrightarrow M.$$

They live in $\mathrm{List}[x_0, x_1, x_2, \dots]$.

But any $xs : \mathrm{List}[x_0, x_1, x_2, \dots]$ is a finite sequence of elements in $\{x_0, x_1, \dots\}$,

so it lives in $\{x_0, x_1, \dots, x_n\}$ for some large enough $n : \mathbb{N}$,

that is only List operation of infinite arity is a finitary operation in disguise.

What are some trivial operations? For example the projection

$$M^{\times 3} \longrightarrow M$$
$$\langle a, b, c \rangle \rightsquigarrow b$$

comes from $[y] = \mathcal{U}(y) : \mathrm{List}[x, y, z]$. ■

**example** What's a $\mathcal{P}$-operation?

We have finitary operations such as

$$A^{\times 3} \longrightarrow A$$
$$\langle a, b, c \rangle \rightsquigarrow a \vee b \vee c$$

which comes from $\{x, y, z\} : \mathcal{P}[x, y, z]$.

But we also have

$$\{x_0, x_1, x_2, \dots\} : \mathcal{P}[x_0, x_1, x_2, \dots]$$

which gives us

$$A^{\widetilde{\times}} \longrightarrow A$$
$$(a_n)_{n \in \mathbb{N}} \rightsquigarrow \bigvee_{n \in \mathbb{N}} a_n.$$

Obviously $\bigvee_{\sim}$ does not come from any finitary operation.

The monad $\mathcal{P}$ has non-trivial operations of any cardinality. ■

Now it should be obvious that monads are very general notion of a theory that allows operations of any arity.

More abstractly

We can think of any $\omega : \mathbb{T}[x, y]$ as a binary operation.

Let $A : \mathbb{T}\text{-}\mathbf{Set}$, $a, b : A$.

Consider the map $\langle x : a, y : b \rangle : \{x, y\} \longrightarrow A$

Since $A$ is a $\mathbb{T}$-model, it lifts to

$$\overline{\langle x : a, y : b \rangle} : \mathbb{T}[x, y] \longrightarrow A$$

We define $[\![\omega]\!] : A \times A \longrightarrow A$

$$[\![\omega]\!](a, b) := \overline{\langle x : a, y : b \rangle}(\omega)$$

Given a $\mathbb{T}$-homomorphism $f : A \longrightarrow B : \mathbb{T}\text{-}\mathbf{Set}$, then $f$ preserves $\omega$ since

$$\begin{aligned} f([\![\omega]\!](a, a')) &= f\,\overline{\langle x : a, y : b \rangle}(\omega) \\ &= \overline{f \circ \langle x : a, y : b \rangle}(\omega) \\ &= \overline{\langle x : f(a), y : f(b) \rangle}(\omega) \\ &= [\![\omega]\!](f(a), f(b)). \end{aligned}$$

**define** Given $X : \mathbf{Set}$, a $\mathbb{T}$-operation of arity $X \to \cdot$ is any element of $\mathbb{T}[X]$ ■

**define** Given $X, Y : \mathbf{Set}$ a $\mathbb{T}$-operation of arity $X \longrightarrow Y$ is any element of $Y \longrightarrow \mathbb{T}[X]$ ■

**define** A trivial operation (a projection) of $\mathbb{T}$ of arity $X \to \cdot$ is any element of the form $\eta(x) : \mathbb{T}[X]$ for some $x : X$. ■

**define** A trivial operation of $\mathbb{T}$ determined by $\alpha : X \leftarrow Y$ is $\eta_X \alpha : X \longrightarrow Y$ ■

Let $\omega : X \longrightarrow Y : \mathbb{T}$, $\eta : Y \longrightarrow Z : \mathbb{T}$

that is $\omega : Y \longrightarrow \mathbb{T}[X]$, $\eta : Z \longrightarrow \mathbb{T}[Y]$.

We define the composition of operations to be

$$\eta\omega : X \longrightarrow Z : \mathbb{T}$$
$$\eta\omega = \omega \circ \eta$$

**prop** corresponds to $h : \mathbb{T} \longrightarrow \mathcal{U} : \mathrm{Monad}$

where $\forall X, Y, Z : \mathbf{Set}$

(1) $\forall \omega : X \longrightarrow Y$, $\eta : Y \longrightarrow \mathbb{T}$

$$h(\eta\omega) = h(\eta) \cdot h(\omega) : X \longrightarrow \mathcal{U}$$

(2) $\forall \varepsilon : X$

$$h(\eta(\varepsilon) : X \to \mathbb{T}) = \eta(\varepsilon) : X \longrightarrow \mathcal{U}$$

**define** Given $A : \mathbb{T}\text{-}\mathbf{Set}$, $\omega : X \longrightarrow Y : \mathbb{T}$,

$$[\![\omega]\!]_A : A^X \longrightarrow A^Y$$
$$[\![\omega]\!]_A(\alpha) = \bar{\alpha}\omega,$$

this makes sense since $\omega : Y \longrightarrow \mathbb{T}[X]$, $\alpha : X \longrightarrow A$ ■

**prop** Let $A, B : \mathbb{T}\text{-}\mathbf{Set}$, $f : A \longrightarrow B : \mathbf{Set}$.

Then $f$ is a $\mathbb{T}$-homomorphism iff $\forall \omega : X \longrightarrow Y : \mathbb{T}$, $f$ preserves $\omega$

**proof** (⇒) Let $f$ be a $\mathbb{T}$-homomorphism, $\omega : X \longrightarrow Y : \mathbb{T}$.

Then

$$\begin{aligned} f[\![\omega]\!]\alpha &= f\bar{\alpha}\omega \\ &= \overline{f\alpha}\,\omega \\ &= [\![\omega]\!](f\alpha) \end{aligned}$$

(⇐) Let $f$ preserve all operations,

then given $X : \mathbf{Set}$, we have the operation $\eta_{\mathbb{T}[X]} : X \longrightarrow \mathbb{T}[X] : \mathbb{T}$

then for $g : X \longrightarrow A$,

$$f[\![\eta_{\mathbb{T}[X]}]\!]g = [\![\eta_{\mathbb{T}[X]}]\!](fg),$$

that is $f\bar{g} = \overline{fg}$. ■

# Canonical Presentation

## Two ways to present a monoid

Let $\mathbf{Mon} := \mathbf{List}\,\mathbf{CSet}$, $M : \mathbf{Mon}$.

Then $\mu : \mathbf{List}[M] \to M : \mathbf{Mon}$ is surjective, since it has a section $\eta_M : M \to \mathbf{List}[M] : \mathbf{Mon}$.

Then $\mathbf{List}[M]/\mathrm{ker}\,\mu \cong M$, for example

$$[a, a', a'', a, a] = [aa', a'', aa] = [aa'a''aa] : \mathbf{List}[M]/\mathrm{ker}\,\mu,$$

and this element denotes $aa'a''aa : M$.

The the neutral element $1 : M$ for example corresponds to $[\,] = [1] = [1, 1] : \mathbf{List}[M]/\mathrm{ker}\,\mu$.

So we can think of $\mu : \mathbf{List}[M] \to M : \mathbf{Mon}$ as the canonical presentation of $M$ as a quotient of a free monoid.

There is also a different way of defining the kernel congruence $\mathrm{ker}_\mu$.

The monad $\mathbf{List}^2[M]$ is equipped with inner and outer reductions

$$\mathbf{List}^2[M] \underset{\mu}{\overset{\mathrm{inner},\,\mathrm{outer} : \mathbf{List}[\mu]}{\rightrightarrows}} \mathbf{List}[M] : \mathbf{Mon},$$

for example for $xss := [[a, a'], [b, b']]$,

$$\mathrm{inner}(xss) = [aa', bb'], \text{ and } \mathrm{outer}(xss) = [a, a', b, b'].$$

Note that the map $\mathbf{List}[\eta_M] : \mathbf{List}[M] \to \mathbf{List}^2[M] : \mathbf{Mon}$ makes inner and outer into a reflexive pair, for example let $xss := \mathbf{List}[\eta_M]([a, b, c]) = [[a], [b], [c]]$, so

$$\mathrm{inner}(xss) = [a, b, c] = \mathrm{outer}(xss).$$

The reflexive pair defines a reflexive monad relation $R$ on $\mathbf{List}[M]$:

$$ys\ R\ zs : \mathbf{List}[M] :\Leftrightarrow \exists\, xss : \mathbf{List}^2[M].\ ys = \mathrm{inner}(xss),\ \mathrm{outer}(xss) = zs.$$

For example from $[[a, a'], [b, b']] : \mathbf{List}^2[X]$ we have $[aa', bb']\ R\ [a, a', b, b']$.

As we've already proved in the chapter on coequalizers of models of finitary algebraic theories, the equivalence relation $\simeq := \{ys \mid xs \mid xs\ R\ ys\} : Q(M : \mathbf{Set})$ is a congruence.

According to $\simeq$ we have for example

$$[a, a', b, b'] = [aa', bb'] = [a, a'b'] = [aa'b'b'] : \mathbf{List}[X]/\simeq.$$

It is fairly obvious that $\simeq$ and $\mathrm{ker}_\mu$ are the same congruences.

## Presentation of models

This generalizes to all monads $\mathbb{T}$.

**prop** Let $M : \mathbb{T}\mathbf{CSet}$, then

$$M \text{ with } \mu : \mathbb{T}[M] \to M : \mathbb{T}\mathbf{CSet}$$

is the coequalizer of the reflexive pair

$$\mathbb{T}^2[M] \underset{\mathrm{outer} := \mu_M}{\overset{\mathrm{inner} := \mathbb{T}[\mu]}{\rightrightarrows}} \mathbb{T}[M] : \mathbb{T}\mathbf{CSet}$$

where the natural section is $\mathrm{refl} := \mathbb{T}[\eta_M] : \mathbb{T}[M] \to \mathbb{T}^2[M] : \mathbb{T}\mathbf{CSet}$.

**proof** (1) By definition of $M$ we have

$$\mu \circ \mathrm{inner} = \mu \circ \mathrm{outer}.$$

(2) Let $f : \mathbb{T}[M] \to N : \mathbb{T}\mathbf{CSet}$ such that $f \circ \mathrm{inner} = f \circ \mathrm{outer}$.

Then we have the universal of $f$, the map $f\eta_M : M \to N : \mathbb{T}\mathbf{CSet}$; now $f\eta_M = f$. We'll show that $f\eta_M$ is the unique map such that

$$\begin{tikzcd} \mathbb{T}[M] \arrow[r, "\mu"] \arrow[dr, "f"'] & M \arrow[d, "f\eta_M"] \\ & N \end{tikzcd}$$

To show that $f = f\eta_M \circ \mu$ just note that $f\eta_M = f\eta_M \circ \mu \circ \eta_M = f\eta_M$.

Let $g : M \to N : \mathbb{T}\mathbf{CSet}$ such that $f = g \circ \mu$.

Then $g = \mu \circ \mathbb{T}[g] = g \circ \mu = f$. ■

# Coproduct

## Coproducts of a pair of monoids

First we'll construct coproducts of monoids.

Let $M, N : \mathbf{Mon}$; then we have the disjoint union $M + N : \mathbf{Set}$.

We will construct the coproduct $M + N : \mathbf{Mon}$ with the help of $M + N : \mathbf{Set}$ and the free monoids.

This we will do by imposing a congruence on $\mathbf{List}[M + N : \mathbf{Set}]$.

If we had $M + N : \mathbf{Mon}$, then $M + N : \mathbf{Set} \subseteq M + N : \mathbf{Mon}$, and $\langle M + N : \mathbf{Set} \rangle = M + N : \mathbf{Mon}$.

So then we would have $\dfrac{\mathbf{List}[M + N : \mathbf{Set}]}{\mathrm{ker}_\mu[M + N : \mathbf{Set} \in M + N : \mathbf{Mon}]} \cong M + N : \mathbf{Mon}$.

Fortunately we can generate this kernel congruence in a way that doesn't make use of $M + N : \mathbf{Mon}$.

Looking at it differently, note the correspondence

$$\frac{\mathbf{List}[M + N : \mathbf{Set}] \longrightarrow S : \mathbf{Mon}}{M + N : \mathbf{Set} \longrightarrow S : \mathbf{Set}}$$

$$\frac{}{M \to S : \mathbf{Set},\ N \to S : \mathbf{Set}}$$

We want to restrict the set of homomorphisms in $\mathbf{List}[M + N : \mathbf{Set}] \to S$, so that the bottom part of the above correspondence reduces to pairs of homomorphisms $M \to S : \mathbf{Mon}$, $N \to S : \mathbf{Mon}$.

Denoting the elements of $M$ and $N$ by $a, a', a''$ and $b, b', b''$ respectively, we want to have a congruence $\approx$ such that for example

$$[a, a', b, a, b, b'] \approx [aa', b, a, bb'] : \mathbf{List}[M + N : \mathbf{Set}]/\approx,$$

or $[\,] \approx [1_M] \approx [1_M, 1_M, 1_M]$.

Recall that we have $\mu_M : \mathbf{List}[M] \to M : \mathbf{Mon}$ and $\mu_N : \mathbf{List}[N] \to N : \mathbf{Mon}$.

This gives us a set map $\mu_M + \mu_N : \mathbf{List}[M] + \mathbf{List}[N : \mathbf{Set}] \to (M + N : \mathbf{Set})$.

In this way we get an inner reduction

$$\mathrm{inner} := \mathbf{List}[\mu_M + \mu_N] : \mathbf{List}[\mathbf{List}[M] + \mathbf{List}[N : \mathbf{Set}]] \to \mathbf{List}[M + N : \mathbf{Set}] : \mathbf{Mon}.$$

The monad $\mathbf{List}[\mathbf{List}[M] + \mathbf{List}[N : \mathbf{Set}]]$ can be thought to contain "unmixed" lists of lists such as $xss := [[a, a'], [b, b, b''], [a]]$ but not $[[a, b, b''], [b, b'], [a', a]]$. Note

$$\mathrm{inner}(xss) = [a, a', b, b, b'', a].$$

We also have an outer reduction

$$\mathrm{outer}(xss) = [a, a', b, b, b'', a] \text{ defined}$$

$$\mathrm{outer} := \mathbf{List}[\mathbf{List}[M] + \mathbf{List}[N : \mathbf{Set}]] \xrightarrow{\mathbf{List}[(+) \circ (\mathbf{id}_M, \mathbf{id}_N) \circ \mathbf{List}[\eta_M]]} \mathbf{List}^2[M + N : \mathbf{Set}] \xrightarrow[\mathrm{inner}]{\mu_{M+N}} \mathbf{List}[M + N : \mathbf{Set}] : \mathbf{Mon}$$

where $M \xrightarrow{\mathrm{inj}_M} (M + N : \mathbf{Set}) \xleftarrow{\mathrm{inj}_N} N$ are set coproduct inclusions.

Then outer and inner form a reflexive pair via

$$\mathbf{List}[\eta_M + \eta_N] : \mathbf{List}[M + N : \mathbf{Set}] \longrightarrow \mathbf{List}[\mathbf{List}[M] + \mathbf{List}[N : \mathbf{Set}]] : \mathbf{Mon}.$$

Then the congruence generated by this reflexive pair is precisely the congruence we need.

Note for example for $xss := [[\eta_M, \eta_N], [\eta_M]]$ we have

$$\mathrm{inner}(xss) = [\eta_M, \eta_N, \eta_M], \quad \mathrm{outer}(xss) = [\,].$$

## Coproducts of models

**prop** Let $\mathbb{T}$ be a monad such that $\mathbb{T}\mathbf{CSet}$ has reflexive coequalizers — for example $\mathbb{T}$ is finitary.

Then $\mathbb{T}\mathbf{CSet}$ has all coproducts.

**proof** Let $M_x : X \longrightarrow \mathbb{T}\mathbf{CSet}$ be an $X : \mathbf{Set}$ indexed family.

Denote the reduction of $M_x$ by $\mu_x : \mathbb{T}[M_x] \to M_x : \mathbb{T}\mathbf{CSet}$, the inclusion by $\iota_x : M_x \longrightarrow \sum_{x:X} M_x : \mathbf{Set}$, and the unit by $\eta_x : M_x \longrightarrow \mathbb{T}[M_x] : \mathbb{T}\mathbf{CSet}$.

Then we have $\sum_{x:X} M_x : \mathbf{Set}$.

Let $\mathrm{inner} := \prod_{x:X} \mu_x : \prod_{x:X} \mathbb{T}[M_x : \mathbf{Set}] \longrightarrow \prod_{x:X} M_x : \mathbb{T}\mathbf{CSet}$,

$$\mathrm{outer} := \mu_{\sum_{x:X} M_x} \circ \prod_{x:X} \mathbb{T}[\mathrm{case}\, x.\, \mathbb{T}[\iota_x]]$$

where

$$\mu_{\sum_{x:X} M_x} : \mathbb{T}^2\Big[\sum_{x:X} M_x\Big] \longrightarrow \mathbb{T}\Big[\sum_{x:X} M_x\Big] : \mathbb{T}\mathbf{CSet}$$

$$\mathrm{case}\, x.\, \mathbb{T}[\iota_x] : \sum_{x:X} \mathbb{T}[M_x] \longrightarrow \mathbb{T}\Big[\sum_{x:X} M_x\Big] : \mathbf{Set},$$

and lastly

$$\mathrm{refl} := \mathbb{T}\Big[\sum_{x:X} \eta_x\Big] : \mathbb{T}\Big[\sum_{x:X} M_x\Big] \longrightarrow \mathbb{T}\Big[\sum_{x:X} \mathbb{T}[M_x : \mathbf{Set}]\Big] : \mathbb{T}\mathbf{CSet}.$$

We'll show

(1) inner, outer with refl is a reflexive pair

(2) Denote the reflexive coequalizer by

$$\mathbb{T}\Big[\sum_{x:X} \mathbb{T}[M_x : \mathbf{Set}]\Big] \underset{\mathrm{outer}}{\overset{\mathrm{inner}}{\rightrightarrows}} \mathbb{T}\Big[\sum_{x:X} M_x\Big] \xrightarrow{q} Q : \mathbb{T}\mathbf{CSet}$$

Then $Q$ can be made into a coproduct of $M_x : X \to \mathbb{T}\mathbf{CSet}$.

(1) Note

$$\begin{aligned} \mathrm{inner} \circ \mathrm{refl} &= \prod\Big[\sum_{x:X} \mu_x\Big] \prod\Big[\sum_{x:X} \eta_x\Big] \\ &= \prod\Big[\sum_{x:X} \mu_x \circ \sum_{x:X} \eta_x\Big] \\ &= \prod\Big[\sum_{x:X} \mu_x \circ \eta_x\Big] \\ &= \prod\Big[\sum_{x:X} \mathrm{id}_{M_x}\Big] \\ &= \mathrm{id}_{\sum_{x:X} M_x} \end{aligned}$$

and

$$\begin{aligned} \mathrm{outer} \circ \mathrm{refl} &= \mu_{\sum_{x:X} M_x} \circ \prod[\mathrm{case}\, x.\, \mathbb{T}[\iota_x]] \prod\Big[\sum_{x:X} \eta_x\Big] \\ &= \mu_{\sum_{x:X} M_x} \prod\Big[\mathrm{case}\, x.\, \mathbb{T}[\iota_x] \sum_{x:X} \eta_x\Big] \\ &= \mu_{\sum_{x:X} M_x} \prod\Big[\mathrm{case}\, x.\, \mathbb{T}[\iota_x]\, \eta_x\Big] \\ &= \mu_{\sum_{x:X} M_x} \prod\Big[\mathrm{case}\, x.\, \eta_{M_x} \circ \iota_x\Big] \\ &= \mu_{\sum_{x:X} M_x} \prod\Big[\mathrm{case}\, x.\, \eta_{M_x} \circ \mathrm{inj}_x\Big] \\ &= \mu_{\sum_{x:X} M_x} \prod\Big[\eta_{\sum_{x:X} M_x} \circ \mathrm{inj}_x\Big] \\ &= \mathrm{id}_{\sum_{x:X} M_x}. \end{aligned}$$

(2) We need to construct the homomorphism inclusion of the coproduct

$$\mathrm{inj} : \forall x : X.\ M_x \longrightarrow Q : \mathbb{T}\mathbf{CSet}.$$

We have $\iota_x : M_x \longrightarrow \sum_{x:X} M_x : \mathbf{Set}$; let $\mathrm{inj}_x := q \prod \iota_x \eta_x$.

Let $f_x : \forall x : X.\ M_x \longrightarrow S : \mathbb{T}\mathbf{CSet}$.

(existence) Define $g : \mathrm{case}\, x.\, f_x : \prod\Big[\sum_{x:X} M_x\Big] \longrightarrow S : \mathbb{T}\mathbf{CSet}$.

We will lift $g$ to $\hat{g} : Q \longrightarrow S : \mathbb{T}\mathbf{CSet}$ which will serve as the case map of $f_x$ in $\mathbb{T}\mathbf{CSet}$.

Note

$$\begin{aligned} g \circ \mathrm{inner} &= \mathrm{case}\, x.\, f_x \prod\Big[\sum_{x:X} \mu_x\Big] \\ &= \mathrm{case}\, x.\, f_x \sum_{x:X} \mu_x \\ &= \mathrm{case}\, x.\, f_x\, \mu_x \end{aligned}$$

and

$$\begin{aligned} g \circ \mathrm{outer} &= \mathrm{case}\, x.\, f_x \circ \mu_{\sum_{x:X} M_x} \prod[\mathrm{case}\, x.\, \mathbb{T}[\iota_x]] \\ &= \mathrm{case}\, x.\, f_x \prod[\mathrm{case}\, x.\, \mathbb{T}[\iota_x]] \\ &= \mathrm{case}\, x.\, f_x\ \mathrm{case}\, x.\, \mathbb{T}[\iota_x] \\ &= \mathrm{case}\, x.\, \mu_x \sum_{x:X} \mathbb{T}[f_x] \\ &= \mathrm{case}\, x.\, (\mu_x \circ \mathbb{T}[f_x]) \\ &= \mathrm{case}\, x.\, f_x\, \mu_x \end{aligned}$$

So we get the lift $\hat{g}$. Then

$$\begin{aligned} \hat{g}\, \mathrm{inj}_{x_0} &= \hat{g}\, q \prod \iota_x \eta_x \\ &= \mathrm{case}\, x.\, f_x \prod \iota_x \eta_x \\ &= \mathrm{case}\, x.\, f_x\ \iota_{x_0} \eta_{x_0} \\ &= \mathrm{case}\, x.\, f_x\ \iota_{x_0} \\ &= f_{x_0} \end{aligned}$$

(uniqueness) Let $h : Q \to S : \mathbb{T}\mathbf{CSet}$ such that

$$\forall x : X.\ h\, \mathrm{inj}_x = f_x.$$

To show $h = \hat{g}$, we'll show

$$h\, q = \hat{g}\, q = g.$$

To show that we'll show

$$h\, q\, \eta_{\sum_{x:X} M_x} = g\, \eta = \mathrm{case}\, x.\, f_x$$

Then it's enough to show

$$h\, q\, \eta_{\sum_{x:X} M_x}\, \iota_{x_0} = f_{x_0}$$

Note

$$\begin{aligned} h\, q\, \eta_{\sum_{x:X} M_x}\, \iota_{x_0} &= h\, q \prod[\iota_{x_0}]\, \eta_{x_0} \\ &= h\, \mathrm{inj}_{x_0} \\ &= f_{x_0}. \end{aligned}$$ ■

So finally

**prop** Let $\mathbb{T}$ be a single sorted finitary theory.

Then $\mathbb{T}\mathbf{CSet}$ is cocomplete.

**proof** We now showed that $\mathbb{T}\mathbf{CSet}$ has coproducts; in the chapter on colimits we showed that it has coequalizers. ■

# Split Coequalizers

Recall that for a monoid $M : \mathbf{Mon}$ we have its canonical presentation as a reflexive coequalizer of inner and outer reductions

$$\begin{tikzcd} \mathrm{List}^2[M] \arrow[r, shift left, "\text{inner} := \mathrm{List}[\mu]"] \arrow[r, shift right, "\text{outer} := \mu_M"'] & \mathrm{List}[M] \arrow[r, "q := \mu"] & M \end{tikzcd}$$

$$\text{where } \eta := \mathrm{List}[\mathrm{id}_M] : \mathrm{List}[M] \to \mathrm{List}^2[M]$$

In fact we have a different map of embedding $\mathrm{List}[M]$ into $\mathrm{List}^2[M]$. We can think of $\eta := \mathrm{List}[\mathrm{id}_M]$ as the inner embedding. $\eta_{in} := \mathrm{List}[\mathrm{id}_M]$, we also have the outer embedding

$$\eta_{out} := \eta_{\mathrm{List}[M]} : \mathrm{List}[M] \longrightarrow \mathrm{List}^2[M]$$

on which we now focus our attention.

Note that for $xs := [a, b, c]$,

$$\eta_{out}(xs) = [[a], [b], [c]], \text{ while } \eta_{in}(xs) = [[a], [b], [c]].$$

Even though $\eta_{in}$ is a common section of both reductions, the $\eta_{out}$ does not enjoy this property.

It is a section to the outer reduction.

for $xs := [a, b, c]$,

$$\text{outer } \eta_{out}(xs) = \text{outer}([[a], [b], [c]]) = [a, b, c],$$

meanwhile $\text{inner } \eta_{out}(xs) = \text{inner}([[a], [b], [c]]) = [a\,b\,c]$.

But $\mathrm{List}[M] \xrightarrow{q := \mu} M$ has a section $\eta_M : M \to \mathrm{List}[M]$ $[\mu([a]) = a]$,

and $\eta_M q([a, b, c]) = [a\,b\,c]$, so $\text{inner } \eta_{out} = \eta_M q$.

This is completely general.

**prop** Let $T$ be a monad, $M$ a $T$-model, and define

$$\begin{tikzcd} T^2[M] \arrow[r, shift left, "\text{inner} := T[\mu]"] \arrow[r, shift right, "\text{outer} := \mu_M"'] & T[M] \arrow[r, shift left, "q := \mu"] \arrow[l, bend left, "\eta_{out} := \eta_{T[M]}"] & M \arrow[l, shift left, "s := \eta_M"] \end{tikzcd}$$

Then (coequalizing) $q \text{ inner} = q \text{ outer}$ $[\mu T[\mu] = \mu \mu_M]$

($q$-section) $q s = \mathrm{id}_M$ $[\mu \eta_M = \mathrm{id}_M]$

(outer-section) $\text{outer } \eta_{out} = \mathrm{id}_{T[M]}$ $[\mu_M \eta_{T[M]} = \mathrm{id}_{T[M]}]$

(naturality) $q s \text{ inner } \eta_{out}$ $[\eta_M \mu = T[\mu] \eta_{T[M]}]$

**proof** We have (coequalizing) and ($q$-section) since $M$ is a $T$-algebra.

We have (outer-section) since $T[M]$ is a $T$-algebra with reduction $\mu_M$.

We have (naturality) since $\eta$ is a natural transformation. ∎

It is fruitful to abstract this situation out and call it a split coequalizer.

**define** In an arbitrary category a split coequalizer consists of

$$\text{maps } \begin{tikzcd} A \arrow[r, shift left, "\text{inner}"] \arrow[r, shift right, "\text{outer}"'] & B \arrow[r, shift left, "q"] \arrow[l, bend left, "\eta_{out}"] & Q \arrow[l, shift left, "s"] \end{tikzcd}$$

where (coequalizing) $q \text{ inner} = q \text{ outer}$

($q$-section) $q s = \mathrm{id}_Q$

(out-section) $\text{outer } \eta_{out} = \mathrm{id}_B$

(naturality) $\text{inner } \eta_{out} = s q$. ∎

Now amazingly we have the following

**prop** Let $A \underset{\text{outer}}{\overset{\text{inner}}{\rightrightarrows}} B \underset{s}{\overset{q}{\rightleftarrows}} Q$ be a split coequalizer.

Then $B \xrightarrow{q} Q$ is the coequalizer of the pair $A \underset{\text{outer}}{\overset{\text{inner}}{\rightrightarrows}} B$.

And since split coequalizers are formulated equationally, they are preserved by all functors.

**proof** Let $f : B \to S$ be such that $f \text{ in} = f \text{ out}$; define $\bar{f} := f s : Q \to S$.

We'll show

(1) $\bar{f} q = f$

(2) Let $g : Q \to S$ st $g q = f$.

Then $g = \bar{f}$

(1) $\bar{f} q = f s q$ (naturality)

$= f \text{ in } \eta q$ (coequalizing)

$= f \text{ out } \eta q$ (out-section)

$= f$

(2) $q$ is epimorphism since it has a section ($q$-section).

So its enough to show $g q = \bar{f} q$, but $g q = f = \bar{f} q$. ∎

# Monadicity

Now we return to the question: when is

a category $\mathbb{E}$ with an adjunction $\mathbf{Set} \underset{\mathcal{U}}{\overset{T}{\underset{\perp}{\rightleftarrows}}} \mathbb{E}$

a category of models of an algebraic theory $\mathbb{T}$

whose forgetful functor is basically $\mathcal{U}$?

We generalize the question: Given $S \underset{\mathcal{U}}{\overset{T}{\underset{\perp}{\rightleftarrows}}} \mathbb{E}$ an adjoint situation,

when do we have a nice equivalence of categories

$$\mathbb{E} \simeq \mathbb{T}\text{-}\mathbf{Cat}(S)$$

where $\mathbb{T} := \mathcal{U}T$?

Each object of $\mathbb{E}$ defines a $\mathbb{T}$-model in a functorial way.

**define** (1) For $A : \mathbb{E}$,

$$\mathcal{U}(A) : S \text{ with } \mathcal{U}(\mu_A) : \mathbb{T}[\mathcal{U}(A)] \longrightarrow \mathcal{U}(A) : S$$

is a $\mathbb{T}$-model.

(2) For $f : A \to B : \mathbb{E}$,

$$\mathcal{U}(f) : \mathcal{U}(A) \longrightarrow \mathcal{U}(B) : S \text{ is a } \mathbb{T}\text{-homomorphism.}$$

This lifts $\mathcal{U} : \mathbb{E} \to S$ to $\mathcal{U} : \mathbb{E} \to \mathbb{T}\text{-}\mathbf{Cat}(S)$. ∎

**define** The adjunction $S \underset{\mathcal{U}}{\overset{T}{\underset{\perp}{\rightleftarrows}}} \mathbb{E}$ is monadic

iff

the functor $\mathcal{U} : \mathbb{E} \to \mathbb{T}\text{-}\mathbf{Cat}(S)$ can be made

into an equivalence of categories. ∎

So our question becomes: when is an adjunction monadic?

Well try to figure out some additional assumptions on our adjoint situation so that we are able to construct some functor $\widehat{(\ )} : \mathbb{T}\text{-}\mathbf{Cat}(S) \longrightarrow \mathbb{E}$ going the other way.

First, what should happen when we restrict $\widehat{(\ )}$ to the free models?

Given $X : S$, the left $\mathcal{U} : \mathbb{E} \to \mathbb{T}\text{-}\mathbf{Cat}(S)$ sends $T(X) : \mathbb{E}$ to the free model $\mathbb{T}[X] : S$ with $\mu_X : \mathbb{T}^2[X] \to \mathbb{T}[X] : S$.

So we need to define $\widehat{\mathbb{T}[X]} : T[X] : \mathbb{E}$. Similarly given $f : X \to \mathbb{T}[Y] : S$, we have the left $\bar{f} : \mathbb{T}[X] \to \mathbb{T}[Y] : \mathbb{T}\text{-}\mathbf{Cat}(S)$

and it should be sent to $\widehat{\bar{f}} := \mu_{T(Y)} T[f] : \widehat{\mathbb{T}[X]} \longrightarrow \widehat{\mathbb{T}[Y]} : \mathbb{E}$.

In this way we get a functor $\widehat{(\ )} : \mathbb{E}_{\mathbb{T}}[\mathbb{T}\text{-}\mathbf{Cat}(S)] \longrightarrow \mathbb{E}$ which is a kind of left of $T$, it is easy to see that we get an embedding

**prop** $\widehat{(\ )} : \mathbb{E}_{\mathbb{T}}[\mathbb{T}\text{-}\mathbf{Cat}(S)] \longrightarrow \mathbb{E}$ is an embedding.

**proof** Let $f : X \to \mathcal{U}T(Y) : S$, then we apply $\widehat{(\ )}$ to get $\mu_{T(Y)} T[f] : T(X) \to T(Y) : \mathbb{E}$,

to prove this, given $g : T(X) \to T(Y) : \mathbb{E}$, we have $\mathcal{U}(g)\eta_X : X \to \mathcal{U}T(Y)$.

We need to show

(1) $\mu_{T(Y)} T[\mathcal{U}(g)\eta_X] = g$ and (2) $\mathcal{U}(\mu_{T(Y)} T[f])\eta_X = f$.

But (1) $\mu_{T(Y)} T[\mathcal{U}(g)\eta_X] = \mu_{T(Y)} T\mathcal{U}(g) T[\eta_X]$

$= g \mu_{T(X)} T[\eta_X]$

$= g$

and (2) $\mathcal{U}(\mu_{T(Y)} T[f])\eta_X = \mathcal{U}(\mu_{T(Y)}) \mathcal{U}T[f]\eta_X$

$= \mathcal{U}(\mu_{T(Y)})\eta_{\mathcal{U}T(Y)} f$

$= f$. ∎

How to generalize this from $\mathbb{E}_{\mathbb{T}}[\mathbb{T}\text{-}\mathbf{Cat}(S)]$ to $\mathbb{T}\text{-}\mathbf{Cat}(S)$?

Let $M : \mathbb{T}\text{-}\mathbf{Cat}(S)$

We have $T[M] : \mathbb{E}$, but this is not big to be $\widehat{M}$. We will attempt to construct it.

Note that we have

$$\begin{tikzcd} \mathbb{T}\mathbb{T}[M] \arrow[r, "T[\mu_M]"] \arrow[d, "\mu_{T[M]}"'] & \mathbb{T}[M] : \mathbb{E} \arrow[d, "\mathcal{U}"] \\ \mathbb{T}^2[M] \arrow[r, "\mu_M"'] & M : S \end{tikzcd}$$

The bottom diagram is crucially a split coequalizer.

So it would be nice if the top diagram had a coequalizer in $\mathbb{E}$.

**define** A $\mathcal{U}$-split pair of the adjunction $S \underset{\mathcal{U}}{\overset{T}{\underset{\perp}{\rightleftarrows}}} \mathbb{E}$ consists of

$$A \underset{g}{\overset{f}{\rightrightarrows}} B : \mathbb{E} \text{ with a split coequalizer } \begin{tikzcd} \mathcal{U}(A) \arrow[r, shift left, "\mathcal{U}(f)"] \arrow[r, shift right, "\mathcal{U}(g)"'] & \mathcal{U}(B) \arrow[r, shift left, "q"] \arrow[l, bend left, "\eta_{out}"] & Q : S \arrow[l, shift left, "s"] \end{tikzcd} ∎$$

**define** A $\mathcal{U}$-split adjunction consists of an adjunction $S \underset{\mathcal{U}}{\overset{T}{\underset{\perp}{\rightleftarrows}}} \mathbb{E}$ with

a way to turn $\mathcal{U}$-split pair $A \underset{g}{\overset{f}{\rightrightarrows}} B : \mathbb{E}$, $\mathcal{U}(A) \underset{\mathcal{U}(g)}{\overset{\mathcal{U}(f)}{\rightrightarrows}} \mathcal{U}(B) \underset{s}{\overset{q}{\rightleftarrows}} Q : S$,

into $\widehat{Q} : \mathbb{E}$ with $\bar{q} : B \longrightarrow \widehat{Q} : \mathbb{E}$

where (1) $A \underset{g}{\overset{f}{\rightrightarrows}} B \xrightarrow{\bar{q}} \widehat{Q} : \mathbb{E}$ a coequalizer diagram,

(2) the canonical map $\alpha$ in

$$\begin{tikzcd} \mathcal{U}(B) \arrow[r, "q"] \arrow[dr, "\mathcal{U}(\bar{q})"'] & Q : \mathbb{T}\text{-}\mathbf{Cat}(S) \arrow[d, "\alpha", "\simeq"'] \\ & \mathcal{U}(\widehat{Q}) \end{tikzcd}$$

has an inverse. ∎

**define** Now assume $S \underset{\mathcal{U}}{\overset{\mathbb{T}}{\rightleftarrows}} \mathbb{E}$ is a $\mathcal{U}$-split adjunction.

(1) Let $M : \mathbb{T}\circlearrowright\mathbf{Set}$. Define $\widehat{M} : \mathbb{E}$ to be in the coequalizer diagram

$$\mathbb{T}\mathbb{T}[M] \underset{\mu_{\mathbb{T}[M]}}{\overset{\mathbb{T}[\mu_M]}{\rightrightarrows}} \mathbb{T}[M] \xrightarrow{\xi} \widehat{M} : \mathbb{E}$$

(2) Let $M \xrightarrow{f} N : \mathbb{T}\circlearrowright\mathbf{Set}$. Then we have

$$\begin{tikzcd}
\mathbb{T}^2[M] \arrow[r, shift left, "\mathbb{T}[\mu_M]"] \arrow[r, shift right, "\mu_{\mathbb{T}[M]}"'] \arrow[d, "\mathbb{T}^2[f]"'] & \mathbb{T}[M] \arrow[r, "\mu_M"] \arrow[d, "\mathbb{T}[f]"'] \arrow[dr, phantom, "\circlearrowright"] & M : S \arrow[d, "f"] \\
\mathbb{T}^2[N] \arrow[r, shift left, "\mathbb{T}[\mu_N]"] \arrow[r, shift right, "\mu_{\mathbb{T}[N]}"'] & \mathbb{T}[N] \arrow[r, "\mu_N"] & N
\end{tikzcd}$$

define $\widehat{f} : \widehat{M} \to \widehat{N}$ using the universal property of coequalizers in

$$\begin{tikzcd}
\mathbb{T}\mathbb{T}[M] \arrow[r, shift left, "\mathbb{T}[\mu_M]"] \arrow[r, shift right, "\mu_{\mathbb{T}[M]}"'] \arrow[d, "\mathbb{T}\mathbb{T}[f]"'] & \mathbb{T}[M] \arrow[r, "\xi"] \arrow[d, "\mathbb{T}[f]"'] \arrow[dr, phantom, "\circlearrowright"] & \widehat{M} : \mathbb{E} \arrow[d, dashed, "\widehat{f}"] \\
\mathbb{T}\mathbb{T}[N] \arrow[r, shift left, "\mathbb{T}[\mu_N]"] \arrow[r, shift right, "\mu_{\mathbb{T}[N]}"'] & \mathbb{T}[N] \arrow[r, "\xi"] & \widehat{N} : \mathbb{E}
\end{tikzcd}$$

Again using universal property of coequalizers it is easy to see that we get a functor

$$\widehat{(\ )} : \mathbb{T}\circlearrowright S \longrightarrow \mathbb{E}. \blacksquare$$

Now is this the desired equivalence constructed from $\mathcal{U} : \mathbb{E} \longrightarrow \mathbb{T}\circlearrowright S$?

Consider

$$\begin{tikzcd}
\mathbb{E} \arrow[r, shift left, "\mathcal{U}"] & \mathbb{T}\circlearrowright S \arrow[l, shift left, "\perp", "\widehat{(\ )}"']
\end{tikzcd}$$

we wish to construct natural isos

(1) $\forall X : \mathbb{T}\circlearrowright S$.

$$\mathcal{U}(\widehat{X}) \cong X : \mathbb{T}\circlearrowright S$$

and (2) $\forall A : \mathbb{E}$.

$$\widehat{\mathcal{U}(A)} \cong A : \mathbb{E}$$

We can show (1) by only assuming $\mathcal{U}$-split adjunction.

**prop** Assume $S \underset{\mathcal{U}}{\overset{\mathbb{T}}{\rightleftarrows}} \mathbb{E}$ is a $\mathcal{U}$-split adjunction, $X : \mathbb{T}\circlearrowright S$.

Then $\mathcal{U}(\widehat{X}) \cong X : \mathbb{T}\circlearrowright S$.

**proof** We have the coequalizer diagram

$$\mathbb{T}\mathbb{T}[M] \underset{\mu_{\mathbb{T}[M]}}{\overset{\mathbb{T}[\mu_M]}{\rightrightarrows}} \mathbb{T}[M] \xrightarrow{\xi} \widehat{M} : \mathbb{E}.$$

coming from the $\mathcal{U}$-split data

$$\mathbb{T}^2[M] \underset{\mu_M}{\overset{\mathbb{T}[\mu_M]}{\rightrightarrows}} \mathbb{T}[M] \xrightarrow{\mu} M : S,$$

then by applying $\mathcal{U}$ to the first diagram we have

$$\begin{tikzcd}
\mathbb{T}[M] \arrow[r, shift left, "\mathbb{T}[\mu_M]"] \arrow[r, shift right, "\mu_M"'] & \mathbb{T}[M] \arrow[r, "\mu"] \arrow[dr, "\mathcal{U}(\widehat{\xi})"'] & M : \mathbb{T}\circlearrowright S \arrow[d, "\alpha"] \\
& & \mathcal{U}(\widehat{M})
\end{tikzcd}$$

where $\alpha$ is an iso. Naturality of $\alpha$ is obvious. $\blacksquare$

For (2) we need an additional hypothesis – $\mathcal{U}$ needs to reflect isos.

In fact if $\mathbb{E} = \mathbb{T}\circlearrowright\mathbf{Set}$, then it is easy to show that $\mathcal{U}$ reflects isos.

We have already seen this for the case of forgetful functors coming from finitary algebraic theories.

**prop** Let $M, N$ be models of some monad $\mathbb{T}$,

$f : M \to N$, $g : N \to M$

where $f \circ g$ homomorphism,

$gf = 1_M$, $fg = 1_N$

Then $g$ is a homomorphism.

**proof** $g\mu_N = g\mu_N \mathbb{T}[f]\mathbb{T}[g]$

$= gf\mu_M \mathbb{T}[g]$

$= \mu_M \mathbb{T}[g] \blacksquare$

Every monad model has a canonical presentation; this is not true in general for $\mathbb{E}$.

**prop** Assume $S \underset{\mathcal{U}}{\overset{\mathbb{T}}{\rightleftarrows}} \mathbb{E}$ is a $\mathcal{U}$-split adjunction, and $\mathcal{U}$ reflects isos. For $A : \mathbb{E}$, the diagram

$$\mathbb{T}\mathcal{U}\mathbb{T}\mathcal{U}(A) \underset{\mu_{\mathbb{T}\mathcal{U}(A)}}{\overset{\mathbb{T}\mathcal{U}(\mu_A)}{\rightrightarrows}} \mathbb{T}\mathcal{U}(A) \xrightarrow{\mu_A} A : \mathbb{E}$$

is a coequalizer.

In particular $\widehat{\mathcal{U}(A)} \cong A : \mathbb{E}$.

**proof** We have $\mathcal{U}(A) : \mathbb{T}\circlearrowright S$, and a coequalizer diagram

$$\mathbb{T}\mathbb{T}[\mathcal{U}(A)] \underset{\mu_{\mathbb{T}[\mathcal{U}(A)]}}{\overset{\mathbb{T}[\mathbb{T}[\mathcal{U}(A)]]}{\rightrightarrows}} \mathbb{T}[\mathcal{U}(A)] \xrightarrow{\xi} \widehat{\mathcal{U}(A)} : \mathbb{E}$$

coming from $\mathcal{U}$-split pair data

$$\mathbb{T}^2[\mathcal{U}(A)] \underset{\mu_{\mathcal{U}(A)}}{\overset{\mathbb{T}[\mathcal{U}(A)]}{\rightrightarrows}} \mathbb{T}[\mathcal{U}(A)] \xrightarrow{\mu_{\mathcal{U}(A)}} \mathcal{U}(A) : S$$

Hence we have

$$\begin{tikzcd}
\mathbb{T}\mathcal{U}\mathbb{T}\mathcal{U}(A) \arrow[r, shift left, "\mathbb{T}\mathcal{U}(\mu_A)"] \arrow[r, shift right, "\mu_{\mathbb{T}\mathcal{U}(A)}"'] \arrow[dr, "\mu_A"'] & \mathbb{T}\mathcal{U}(A) \arrow[r, "\xi"] \arrow[d, "\mu_A"] \arrow[dr, phantom, "\circlearrowright"] & \widehat{\mathcal{U}(A)} : \mathbb{E} \arrow[d, "\alpha"] \\
& A, & A : \mathbb{E}
\end{tikzcd}$$

Since $\mu_A$ is coequalized by the pair of maps above, applying $\mathcal{U}$ we have

$$\begin{tikzcd}
\mathbb{T}[\mathcal{U}(A)] \arrow[r, "\mathcal{U}(\xi)"] \arrow[dr, "\mathcal{U}(\mu_A)"'] & \mathcal{U}(\widehat{\mathcal{U}(A)}) : S \arrow[d, "\mathcal{U}(\alpha)"] \\
& \mathcal{U}(A)
\end{tikzcd}$$

where $\mathcal{U}(\alpha)$ is the canonical iso,

since both $\mathcal{U}(\mu_A)$ and $\mathcal{U}(\xi)$ are coequalizer maps.

But $\mathcal{U}$ reflects isos, so $\widehat{\mathcal{U}(A)} \xrightarrow{\alpha} A : \mathbb{E}$ is iso.

Naturality of $\alpha$ is obvious. $\blacksquare$

So we finally proved the monadicity theorem

**prop** [Monadicity]

Let $S \underset{\mathcal{U}}{\overset{\mathbb{T}}{\rightleftarrows}} \mathbb{E}$ be an adjunction,

then

the adjunction is monadic

iff (1) $\mathcal{U}$ reflects isomorphisms

(2) the adjunction is a $\mathcal{U}$-split adjunction $\blacksquare$

Back to finitary theories.

**prop** Let $\mathbf{Set} \underset{\mathcal{U}}{\overset{\mathbb{T}}{\rightleftarrows}} \mathbb{E}$ be an adjunction;

then

this is an adjunction coming from some finitary single-sorted algebraic theory

iff

(1) $\mathcal{U}$ reflects isomorphisms,

(2) the adjunction is a $\mathcal{U}$-split adjunction,

(3) the functor $\mathcal{U}\mathbb{T} : \mathbf{Set} \to \mathbf{Set}$ is finitary. $\blacksquare$

# Possibility Spaces

Consider a toss of a coin. If the coin is fair, we say there is 50% probability that it will be Heads, and equally 50% probability that it will be Tails.

Denoting the outcomes by H and T, we might denote the answer to the question 'what is the final outcome' by the expression

$$50\%\,H + 50\%\,T.$$

It is neither H or T, but a precisely quantified superposition or possibility of H or T.

If the coin is unfair the final possibility might be

$$30\%\,H + 70\%\,T$$

or even $1\%\,H + 99\%\,T$ [magnets].

The coefficients of such possibilities might signify chance, or lack of certainty – that is they are probabilities and must sum to 1.

In general, given a set $X$ of outcomes, the possibilities of $X$ are expressions of the form

$$p_1 x_1 + p_2 x_2 + \ldots + p_n x_n \quad \text{where } x_1, \ldots, x_n : X, \quad p_1, \ldots, p_n : [0;1], \quad p_1 + \ldots + p_n = 1.$$

More formally

**define** The possibility space over $X : \mathbf{Set}$ is defined to be

$$\mathbf{Pos}[X] := \left\{ \omega : X \longrightarrow [0;1] \;\middle|\; \begin{array}{l} \omega \text{ has finite support} \\ \sum_{x:X} \omega(x) = 1 \end{array} \right\}$$

**example** $\mathbf{Pos}[H,T] = \{ p\cdot H + q\cdot T \mid p, q : [0;1],\ p+q=1 \}$.

Possibilities on $X$ are also known as probability distributions or convex combinations.

Another way to think of possibilities is as random elements of $X$.

Suppose we bet on the outcome of a six-sided dice, if the outcome is ⚅ we get 60\$, otherwise we pay 20\$. How much do we get after the roll? Let $\mathrm{Face} := \{ ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ \}$.

Suppose the dice is fair, so it is governed by the possibility

$$P := \tfrac{1}{6}\cdot⚀ + \tfrac{1}{6}\cdot⚁ + \tfrac{1}{6}\cdot⚂ + \tfrac{1}{6}\cdot⚃ + \tfrac{1}{6}\cdot⚄ + \tfrac{1}{6}\cdot⚅ : \mathbf{Pos}[\mathrm{Faces}]$$

The value of the outcome is given by the function

$$\begin{aligned} v : \mathrm{Face} &\longrightarrow \mathrm{Money} \\ v(⚅) &:= 60\$ \\ v(_) &:= -20\$ \end{aligned}$$

We apply function $v$ to possibility $P$ to get the possibility

$$\begin{aligned} v(P) &= \tfrac{1}{6}\cdot v(⚀) + \tfrac{1}{6}\cdot v(⚁) + \tfrac{1}{6}\cdot v(⚂) + \tfrac{1}{6}\cdot v(⚃) + \tfrac{1}{6}\cdot v(⚄) + \tfrac{1}{6}\cdot v(⚅) \\ &= \tfrac{1}{6}\cdot -20\$ + \tfrac{1}{6}\cdot -20\$ + \tfrac{1}{6}\cdot -20\$ + \tfrac{1}{6}\cdot -20\$ + \tfrac{1}{6}\cdot -20\$ + \tfrac{1}{6}\cdot 60\$ \\ &= \tfrac{5}{6}\cdot -20\$ + \tfrac{1}{6}\cdot 60\$. \end{aligned}$$

This is the exact answer; we can ask for the average (expected) outcome by equipping Money with an expectation operator

$$E : \mathbf{Pos}[\mathrm{Money}] \longrightarrow \mathrm{Money}$$

so that

$$\begin{aligned} E(v(p)) &= E\left[ \tfrac{5}{6}\cdot -20\$ + \tfrac{1}{6}\cdot 60\$ \right] \\ &= \left( \tfrac{5}{6}\cdot -20 + \tfrac{1}{6}\cdot 60 \right)\$ \\ &= -6.66\$. \end{aligned}$$

Abstracting out

**define** We can extend $f : X \longrightarrow Y : \mathbf{Set}$ to

$$\begin{aligned} f : \mathbf{Pos}[X] &\longrightarrow \mathbf{Pos}[Y] \\ f(p_1 x_1 + \ldots + p_n x_n) &:= p_1 f(x_1) + \ldots + p_n f(x_n) \end{aligned}$$

**example** Suppose $X := \{a_1, a_2, a_3\}$, $Y := \{b_1, b_2, b_3, b_4\}$,

$$\begin{aligned} f : X &\longrightarrow Y \\ a_1 &\longrightarrow b_1 \\ a_2 &\longrightarrow b_2 \\ a_3 &\longrightarrow b_3 \\ & \quad\ \ b_4 \end{aligned}$$

Then $f(30\%\cdot a_1 + 20\%\cdot a_2 + 50\%\cdot a_3) = 30\%\cdot b_1 + 70\%\cdot b_2$.

Each element $x : X$ defines a certain possibility $x : \mathbf{Pos}[X]$, for example $H := 1H + 0T : \mathbf{Pos}[X]$.

So we get a map

**define** $\eta_X : X \longrightarrow \mathbf{Pos}[X]$.

Might this be a monad?

We can think of functions

$$f : X \longrightarrow \mathbf{Pos}[Y]$$

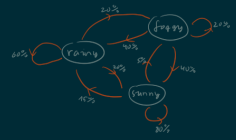
as maps from $X$ to $Y$ that for each input $x : X$, return a random output $y : Y$, that is $f$ is a function with randomness effect.

**example** Suppose that weather each day can be only in three mutually exclusive states: sunny, foggy, raining. If it raining today, there is a 60% chance of raining tomorrow and 20% chance of fog; if its foggy, tomorrow there is a 40% chance of rain and 40% chance of sun; if its sunny, there is 80% chance of sun, and 15% of rain.

Let $\mathrm{Weather} := \{\mathrm{raining}, \mathrm{foggy}, \mathrm{sunny}\}$. This can be modelled as a function

$$\begin{aligned} f : \mathrm{Weather} &\longrightarrow \mathbf{Pos}[\mathrm{Weather}] \\ \mathrm{raining} &\longrightarrow 60\%\cdot\mathrm{raining} + 20\%\cdot\mathrm{foggy} + 20\%\cdot\mathrm{sunny} \\ \mathrm{foggy} &\longrightarrow 40\%\cdot\mathrm{raining} + 20\%\cdot\mathrm{foggy} + 40\%\cdot\mathrm{sunny} \\ \mathrm{sunny} &\longrightarrow 15\%\cdot\mathrm{raining} + 5\%\cdot\mathrm{foggy} + 80\%\cdot\mathrm{sunny} \end{aligned}$$

We can visualize it as a probabilistic state machine




or think of it as a matrix whose columns sum up to 1

$$\begin{array}{c|ccc} & \mathrm{rain} & \mathrm{fogg} & \mathrm{sunny} \\ \hline \mathrm{rain} & 60\% & 40\% & 15\% \\ \mathrm{fogg} & 20\% & 20\% & 5\% \\ \mathrm{sunny} & 20\% & 40\% & 80\% \end{array}$$

Suppose today is sunny, what kind of a weather will the day after tomorrow bring?

The weather of tomorrow is

$$f(\mathrm{raining}) = 60\%\cdot\mathrm{raining} + 20\%\cdot\mathrm{foggy} + 20\%\cdot\mathrm{sunny}$$

The day after that

$$\begin{aligned} f(60\%\cdot\mathrm{raining} + 20\%\cdot\mathrm{foggy} + 20\%\cdot\mathrm{sunny}) &= 60\%\cdot f(\mathrm{raining}) + 20\%\cdot f(\mathrm{foggy}) + 20\%\cdot f(\mathrm{sunny}) \\ &= (36\%\cdot\mathrm{raining} + 12\%\cdot\mathrm{foggy} + 12\%\cdot\mathrm{sunny}) \\ &\quad + (8\%\cdot\mathrm{raining} + 4\%\cdot\mathrm{foggy} + 8\%\cdot\mathrm{sunny}) \\ &\quad + (3\%\cdot\mathrm{raining} + 1\%\cdot\mathrm{foggy} + 16\%\cdot\mathrm{sunny}) \\ &= 47\%\cdot\mathrm{raining} + 17\%\cdot\mathrm{foggy} + 36\%\cdot\mathrm{sunny} \end{aligned}$$

Note $47\% + 17\% + 36\% = 100\%$

We can get at this by multiplying

$$\begin{bmatrix} 47\% \\ 17\% \\ 36\% \end{bmatrix} = \begin{bmatrix} 60\% & 40\% & 15\% \\ 20\% & 20\% & 5\% \\ 20\% & 40\% & 80\% \end{bmatrix} \begin{bmatrix} 60\% \\ 20\% \\ 20\% \end{bmatrix}$$

Abstracting

**define** Let $f : X \longrightarrow \mathbf{Pos}[Y]$, define its lift

$$\begin{aligned} \bar{f} : \mathbf{Pos}[X] &\longrightarrow \mathbf{Pos}[Y] \\ \bar{f}(p_1 x_1 + \ldots + p_n x_n) &:= p_1 q_1^1 y_1^1 + \ldots + p_1 q_{m_1}^1 y_{m_1}^1 \\ &\quad + \ldots + \\ &\quad p_n q_1^n y_1^n + \ldots + p_n q_{m_n}^n y_{m_n}^n \end{aligned}$$

where $q_1^i y_1^i + \ldots + q_{m_i}^i y_{m_i}^i = f(x_i)$.

This allows us to compose random valued maps

**define** For $f : X \longrightarrow \mathbf{Pos}[Y]$, $g : Y \longrightarrow \mathbf{Pos}[Z]$,

$$\begin{aligned} g \cdot f : X &\longrightarrow \mathbf{Pos}[Z] \\ (g \cdot f)(x) &:= \bar{g}(f(x)) \end{aligned}$$

And reduction

**define** $\mu : \mathbf{Pos}[\mathbf{Pos}[X]] \longrightarrow \mathbf{Pos}[X]$

$$\mu := \overline{id_{\mathbf{Pos}[X]}}$$

It is much more intuitive to illustrate $\mu$ on small possibilities

**example** For clarity add parentheses around a possibility, e.t.

for $a, a', b, b', c, c' : [0;1]$, $a + a' = 1$, $b + b' = 1$, $c + c' = 1$,

$x, y, z, w : X$ be fixed

$$(b\cdot x + c\cdot y) : \mathbf{Pos}[X],$$

and

$$\left( a\cdot(b\cdot x + b'\cdot y) + a'\cdot(c\cdot z + c'\cdot w) \right) : \mathbf{Pos}\mathbf{Pos}[X],$$

then

$$\mu\left[ \left( a\cdot(b\cdot x + b'\cdot y) + a'\cdot(c\cdot z + c'\cdot w) \right) \right] = (ab\cdot x + ab'\cdot y + a'c\cdot z + a'c'\cdot w).$$

So $\mathbf{Pos}$ is a monad.

Functions $f : X \longrightarrow \mathbf{Pos}[Y]$ are also known as discrete Markov Chains and the category of effects $\mathbf{Eff}[\mathbf{Pos}(\mathbf{Set})]$ can be thought of as the category of discrete Markov Chains.

What about the $\mathbf{Pos}$-models? We'll call them possibility spaces. We've already met one –

Moneys with expectation operator $E : \mathbf{Pos}[\mathrm{Money}] \longrightarrow \mathrm{Money}$.

**define** A possibility space consists of

$$A : \mathbf{Set} \text{ with } E : \mathbf{Pos}[A] \longrightarrow A$$

where (normalization) $\forall a : A$

$$E[a] = a$$

(convexity)

$$E\,E = E\,\mu_A.$$

**example** Note $\mathrm{Moneys} \cong \mathbb{R}$ [non-canonically].

Let $E : \mathbf{Pos}[\mathbb{R}] \longrightarrow \mathbb{R}$

$$E[p_1 r_1 + \ldots + p_n r_n] := p_1 r_1 + \ldots + p_n r_n$$

We illustrate associativity on an example:

$$\begin{aligned} E\left[ \mu\left[ (a\cdot(b\cdot r + c\cdot s) + a'\cdot(b\cdot r + c\cdot s)) \right] \right] &= E\left[ ab\cdot r + ac\cdot s + a'b'\cdot r + a'c'\cdot s \right] \\ &= (ab)\cdot r + (ac)\cdot s + (a'b')\cdot r + (a'c')\cdot s, \end{aligned}$$

$$\begin{aligned} E\left[ E\left[ (a\cdot(b\cdot r + c\cdot s) + a'\cdot(b\cdot r + c\cdot s)) \right] \right] &= E\left[ (a\cdot E(b\cdot r + c\cdot s) + a'\cdot E(b\cdot r + c\cdot s)) \right] \\ &= E\left[ (a\cdot(b\cdot r + c\cdot s) + a'\cdot(b'\cdot r + c'\cdot s)) \right] \\ &= a(br + cs) + a'(b'r + c's). \end{aligned}$$

For $X : \mathbf{Set}$, in probability theory functions

$$f : X \longrightarrow \mathbb{R}$$

are called random variables and have their own expectation. Here it is obvious that we can't speak of expectation as a real number of $f$ without first choosing a possibility on $X$.

So the expectation is possibility dependent function

$$\begin{aligned} E[f] : \mathbf{Pos}[X] &\longrightarrow \mathbb{R} \\ E[f](P) &:= E[f(P)] \end{aligned}$$

This is exactly the monadic lift of $f$ with respect to $\mathbb{R} : \mathbf{Pos} \circlearrowleft \mathbf{Set}$.

Obviously the monad $\mathbf{Pos}$ is finitary. Does it have any special operations?

Note that $\mathbf{Pos}[\varnothing]$ is empty since the only map $\omega : \varnothing \to [0;1]$ sums to 1, so there are no special constants (as opposed to vector spaces – they have 0).

Since $\mathbf{Pos}[X] := \{x\}$, the only nullary operation is identity $id : \bullet \longrightarrow \mathbf{Pos}$.

Note

$$\mathbf{Pos}[X, Y] := \left\{ p\cdot X + q\cdot Y \;\middle|\; \begin{array}{l} p, q : [0;1], \\ p + q = 1 \end{array} \right\}$$

So assuming $\langle X, Y \rangle : \mathbf{Set} \cdot \mathbf{Set} : \mathbf{Set}^2 \longrightarrow \mathbf{Set}$ the only binary operation is a convex combination $p\cdot X + q\cdot Y : \mathbf{Set}^2 \longrightarrow \mathbf{Set}$ where $p, q : [0;1]$, $p + q = 1$; these are parametrized by a single number in $p : [0;1]$ – they have the form

$$p\cdot X + (1-p)\cdot Y : \mathbf{Set}^2 \longrightarrow \mathbf{Set}.$$

All of the other operations can be generated from the binary one.

For example any ternary has the form

$$a\cdot X + b\cdot Y + c\cdot Z : \mathbf{Set}^3 \longrightarrow \bullet, \quad a, b, c : [0;1],\ a + b + c = 1.$$

Then

case $a = 1$: Then the operation is the projection $X : \mathbf{Set}^3 \longrightarrow \bullet$.

case $a \neq 1$: Then

$$a\cdot X + b\cdot Y + c\cdot Z = a\cdot X + (1-a)\left[ \frac{b}{1-a}\cdot Y + \frac{c}{1-a}\cdot Z \right] : \mathbf{Set}^3 \longrightarrow \bullet$$

where $\frac{b}{1-a} + \frac{c}{1-a} = \frac{b+c}{1-a} = \frac{1-a}{1-a} = 1$.

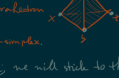
Similarly for operations of higher arity.

So we can think of possibility spaces as sets where the notion of a convex combination makes sense; for this reason they are called convex spaces.

We can visualize $\mathbf{Pos}[x, y]$ as a line segment $x \longrightarrow y$;

$\mathbf{Pos}[x, y, z]$ is a triangle

$\mathbf{Pos}[x, y, z, w]$ is a tetrahedron




So $\mathbf{Pos}[x_0, x_1, \ldots, x_n]$ is an $n$-simplex.

But it's not that helpful, we will stick to the probabilistic intuition.

**example** Intervals $[0;1]$, $[0;\infty)$ are sub-possibility spaces of $\mathbb{R}$.

Even any interval $(a;b)$ or $[a;b]$ is a sub-possibility space of $\mathbb{R}$.

**example** Let $V$ be an $\mathbb{R}$-vector space; then $V$ is canonically a possibility space. Any convex subset of $V$ is a sub-possibility space of $V$.

It is possible to give a fairly small axiomatization of the operation $p \cdot X + (1-p) \cdot Y : 2 \to -$, so that we could present the theory of possibility spaces in terms of a family of binary operations satisfying universally quantified equations. Such presentation of possibility spaces is not as natural as the standard presentations of vector and affine spaces; this is probably the only reason why possibility spaces are not as popular as vector and affine spaces.

Through the lens of monads and algebraic theories, where we consider all operations at once, the three concepts do not seem so different from each other:

vector spaces are sets where we can take linear combinations,
affine spaces are sets where we can take affine combinations,
convex spaces are sets where we can take convex combinations.

One might wonder what sorts of combinations are there?

We give a few further examples.

**example** Conical (or positive) combinations are for example of the form

$$a \cdot x + b \cdot y + c \cdot z \quad \text{where } a, b, c : [0, \infty) \blacksquare$$

**example** In Quantum Information Theory it seems that the basic mathematical objects are convex sets where the real unit interval $[0,1]$ is replaced by the complex unit disk $D^2 := \{ z : \mathbb{C} \mid |z| \leq 1 \}$.

More precisely define

$$D^2[X] := \left\{ \omega : X \to D^2 \;\middle|\; \begin{array}{l} \omega \text{ has finite support,} \\ \sum_{x : X} |\omega(x)|^2 = 1 \end{array} \right\}.$$

Note that $D^2[X]$ is closed under complex multiplication.

We call elements of $[0,1]$ probabilities; the elements of $D^2$ are called complex amplitudes.

Instead of the possibility of a coin flip $p \cdot H + q \cdot T$ one considers the state of an electron in a hydrogen atom, the analogues of the $H$ & $T$ are $|0\rangle$ the ground state and $|1\rangle$ the excited state, the analogue of the possibility $p \cdot H + q \cdot T$ is the hydrogen electron state

$$\alpha \cdot |0\rangle + \beta \cdot |1\rangle : D^2[\{|0\rangle, |1\rangle\}] \quad \text{where } |\alpha|^2 + |\beta|^2 = 1$$

Each such state defines a possibility $|\alpha|^2 \cdot |0\rangle + |\beta|^2 \cdot |1\rangle : Pos[\{|0\rangle, |1\rangle\}]$

For example the state $\left(\frac{1}{2} + \frac{1}{2} i\right) |0\rangle - \frac{1}{12} |1\rangle$ leads to the possibility $\frac{1}{2} |0\rangle + \frac{1}{2} |1\rangle$.

For more information see Quantum Mechanics and Quantum Computation course by Umesh Vazirani on edx.org. ■

**example** We can consider an countable-infinitary extension of $Pos$.

Instead of finite possibilities

$$p_1 \cdot x_1 + \ldots + p_n \cdot x_n, \quad p_1, \ldots, p_n : [0,1], \; p_1 + \ldots + p_n = 1$$

we consider

$$\sum_{n=0}^{\infty} p_n \cdot x_n \quad \text{where } p_i : [0,1], \text{ the series } \sum_{n=0}^{\infty} p_n \text{ converges to } 1.$$

For example $\frac{1}{2} \cdot x_1 + \frac{1}{4} \cdot x_2 + \frac{1}{8} \cdot x_3 + \ldots$ is a valid possibility.

Then

$$Pos_\omega[X] := \left\{ \omega : X \to [0,1] \;\middle|\; \begin{array}{l} \omega \text{ has at most countable support,} \\ \sum_{x : X} \omega(x) \text{ converges to } 1 \end{array} \right\}$$

also has a monad structure.

This monad is not finitary, but it is bounded - obviously

$$T[X] \text{ can be reconstructed from } T[S] \text{ where } S \subseteq X \text{ at most countable} \blacksquare$$

What are the homomorphisms of possibility spaces?

Given $A, B$ possibility spaces, a homomorphism is a map $f : A \to B$ commuting with the expectation operator

$$f E = E f;$$

equivalently they are the maps that preserve the convex combination

$$f(p \cdot a + q \cdot a') = p \cdot f(a) + q \cdot f(a') \quad \text{where } p + q = 1 : [0,1].$$

These are known as convex-linear maps.

**example** Given $X : Set$, then $Pos[X]$ is at $\mathcal{M} : Pos[Pos[X]] \to Pos[X]$

if the free possibility space, so for example
convex-linear maps $f : Pos[X] \to \mathbb{R}$
correspond to random variables $f : X \to \mathbb{R}$,
and convex-linear maps $f : Pos[X] \to Pos[Y]$
correspond to discrete markov chains $f : X \to Pos[Y]$. ■

## Tensor Product

Suppose we flip a coin twice.

The final result is one of the pairs

$$\langle H, H \rangle, \langle H, T \rangle, \langle T, H \rangle, \langle T, T \rangle$$

Can we model this situation with possibility spaces?

Let $Side := \{H, T\}$.

Supposing that the coin flip is governed by $50\% \cdot H + 50\% \cdot T : Pos[Side]$, then flipping it twice is governed by

$$25\% \cdot \langle H, H \rangle + 25\% \cdot \langle H, T \rangle + 25\% \cdot \langle T, H \rangle + 25\% \cdot \langle T, T \rangle : Pos[Side \times Side].$$

For clarity denote the pair $\langle x, y \rangle$ by $x \otimes y$ (read $x$ tensor $y$).

Then we write

$$P := 25\% \cdot H \otimes H + 25\% \cdot H \otimes T + 25\% \cdot T \otimes H + 25\% \cdot T \otimes T$$

for the above possibility; it can be suggestively "factored" into a product of possibilities

$$(50\% \cdot H + 50\% \cdot T) \otimes (50\% \cdot H + 50\% \cdot T),$$

the possibility $P$ can be recovered by multiplying the above expression out.

Now we make sense of this.

In mathematics given elements $a : X$, $b : Y$ it took a long time to recognize that we can think of $a$ and $b$ as being coupled together into one mathematical object called the pair of $a$ and $b$ denoted $\langle a, b \rangle$ living in a new set $X \times Y$.

Such a step was probably not taken because of its utter triviality, why would you even consider such a thing?

But after mathematicians were forced to grapple with concepts such as vectors, complex numbers, or multivariable functions, the pair was born.

We can perform a similar abstraction when replacing elements with possibilities.

If we think of $Pos[X]$ and $Pos[Y]$ as the random elements of $X$ and $Y$, what should be a random pair? Perhaps a pair of random elements?

The set $Pos[X] \times Pos[Y]$ is a possibility space, but that's not the right candidate.

An outcome of a possibility of pairs should be a pair - so a random pair is just a random element of $X \times Y$.

**define** We call the space of all random pairs the tensor product of $Pos[X]$ with $Pos[Y]$

$$Pos[X] \otimes Pos[Y] := Pos[X \times Y] \blacksquare$$

Given possibilities $P := \sum_i p_i \cdot x_i : Pos[X]$, $Q := \sum_j q_j \cdot y_j : Pos[Y]$, then we can construct their product which we denote

$$P \otimes Q := \left( \sum_i p_i \cdot x_i \right) \otimes \left( \sum_j q_j \cdot y_j \right) : Pos[X] \otimes Pos[Y]$$

$$:= \sum_i \sum_j p_i q_j \cdot x_i \otimes y_j$$

This is a possibility on $X \times Y$ since $\sum_i \sum_j p_i q_j = \sum_i p_i \left( \sum_j q_j \right) = \sum_i p_i = 1$.

So we get a map

$$\otimes : Pos[X] \times Pos[Y] \longrightarrow Pos[X] \otimes Pos[Y] : Set.$$

The possibilities that can be constructed in this way are special - they might be called independent.

For example the possibility

$$P := 21\% \cdot HH + 18\% \cdot HT + 28\% \cdot TH + 42\% \cdot TT$$
$$= (30\% \cdot H + 70\% \cdot T) \otimes (40\% \cdot H + 60\% \cdot T),$$

can be factored, while the possibility

$$Q := 15\% \cdot HH + 35\% \cdot HT + 20\% \cdot TH + 30\% \cdot TT$$

can't be factored; if it could be factored, then

$$(aH + bT) \otimes (cH + dT) = ac \cdot HH + ad \cdot HT + bc \cdot TH + bd \cdot TT$$

we would have

$$ac = 0.15, \quad ad = 0.35, \quad bc = 0.2, \quad bd = 0.3.$$

Then $a = 0$ or $a \neq 0$

case $a = 0$: Then $0 = ac = 0.15$, a contradiction.

case $a \neq 0$: Obviously we can assume $b \neq 0$.

Then $c = \frac{0.15}{a} = \frac{0.2}{b}$, $d = \frac{0.35}{a} = \frac{0.3}{b}$.

So $0.2a - 0.15b = 0$,
$0.3a - 0.35b = 0$.

This system has unique solution $a = 0$, $b = 0$, a contradiction.

We can easily give meaning to both $P$ and $Q$.

Consider the situation where we flip two unfair coins, the first is governed by $30\% \cdot H + 70\% \cdot T$ and the second by $40\% \cdot H + 60\% \cdot T$.

Then the final outcome of both is governed by

$$(30\% \cdot H + 70\% \cdot T) \otimes (40\% \cdot H + 60\% \cdot T) = P$$

Another situation is when we have three coins $C$, $D$, $D'$ where $C$ is fair, so it is governed by $50\% \cdot H + 50\% \cdot T$, and $D$, $D'$ are unfair, and are governed by $30\% \cdot H + 70\% \cdot T$ and $40\% \cdot H + 60\% \cdot T$ respectively.

Suppose we flip $C$, and then if $C$ is Heads we flip $D$, otherwise we flip $D'$.

The result is governed by

$$50\% \cdot H \otimes (30\% \cdot H + 70\% \cdot T) + 50\% \cdot T \otimes (40\% \cdot H + 60\% \cdot T)$$
$$= 15\% \cdot HH + 35\% \cdot HT + 20\% \cdot TH + 30\% \cdot TT$$
$$= Q.$$

So the probability that we get both Tails is $30\%$.

From $Q$ we can easily answer:

what is the probability of the first flip being Heads, given that the second flip was Tails

We rearrange

$$0.15 \cdot HH + 0.35 \cdot HT + 0.2 \cdot TH + 0.3 \cdot TT$$
$$= 0.15 \cdot HH + 0.2 \cdot TH + 0.35 \cdot HT + 0.3 \cdot TT$$
$$= (\cdots) \otimes (1-c) \cdot H + (aH + bT) \otimes c \cdot T$$
$$= (\cdots) \otimes (1-c) \cdot H + (aH + bT) \otimes c \cdot T, \quad \begin{array}{l} ac = 0.15, \; bc = 0.3 \\ a + b = 1 \\ \text{so } c = 0.35 + 0.3 = 0.65 \end{array}$$
$$= (\cdots) \otimes 0.35 \cdot H + \left( \frac{0.15}{0.65} \cdot H + \frac{0.3}{0.65} \cdot T \right) \otimes 0.65 \cdot T$$

So the probability is $\frac{0.15}{0.65} \approx 0.5385$

So we see that tensor product $Pos[X] \otimes Pos[Y]$ seamlessly encodes information about independence and conditional probability.

**define** Let $X : Set$, by an event we mean any subset $E \subseteq X$ (or equivalently a property of $X$)

Given a possibility $P := p_1 \cdot a_1 + \ldots + p_n \cdot a_n : Pos[X]$, define the probability of event $E$ with respect to possibility $P$

$$P[E] := \sum_{\substack{i : n \\ a_i \in E}} p_i.$$

Given $a : X$, denote the Event that $a$ occurred by

$$E_a := \{a\}$$

Then $P[E_{a_k}] = p_k$, that is

$$P = P[E_{a_1}] \cdot a_1 + \ldots + P[E_{a_n}] \cdot a_n \blacksquare$$

**prop** Then $P[-] : \mathcal{P}[X] \longrightarrow [0,1]$ is a finitary probability measure:

so $P[\emptyset] = 0$, $P[X] = 1$,

for mutually disjoint events $E_1, \ldots, E_n$

$$P[E_1 \cup \ldots \cup E_n] = P[E_1] + \ldots + P[E_n],$$

but especially

$$\exists \text{finite } E \subseteq X.$$
$$P[E] = 0. \blacksquare$$

**define** Let $P := \sum_{i:n} p_i \cdot a_i : Pos[X]$, $T$ event on $X$ w/ $P[T] \neq 0$.

Then we can define the restricted (posterior) possibility

$$P|_T := \frac{1}{P[T]} \sum_{\substack{i : n \\ a_i \in T}} p_i \cdot a_i : Pos[T]$$

Given event $E$ on $X$, we define the conditional probability of $E$ with respect to $P$ over $T$ as

$$P[E \mid T] := P|_T[E \cap T] = \frac{1}{P[T]} \sum_{\substack{i : n \\ a_i \in E \cap T}} p_i \blacksquare$$

**example** Let $X := \{t, \bar{t}\}$, $Y := \{e_1, e_2, e_3\}$

Consider the possibility

$$P := p \cdot t \otimes (a \cdot e_1 + b \cdot e_2 + c \cdot e_3) + q \cdot \bar{t} \otimes (a' \cdot e_1 + b' \cdot e_2 + c' \cdot e_3) : Pos[X] \otimes Pos[Y]$$

where $p, q, a, a', b, b', c, c' \geq 0$.

We implicitly assumed $p + q = 1$, $a + b + c = 1$, $a' + b' + c' = 1$

Then

$$P = (pa \cdot t \otimes e_1 + pb \cdot t \otimes e_2 + pc \cdot t \otimes e_3) + (qa' \cdot \bar{t} \otimes e_1 + qb' \cdot \bar{t} \otimes e_2 + qc' \cdot \bar{t} \otimes e_3)$$
$$= (pa \cdot t \otimes e_1 + qa' \cdot \bar{t} \otimes e_1) + (pb \cdot t \otimes e_2 + qb' \cdot \bar{t} \otimes e_2) + (pc \cdot t \otimes e_3 + qc' \cdot \bar{t} \otimes e_3)$$
$$= \left( \frac{pa}{pa + qa'} \cdot t + \frac{qa'}{pa + qa'} \cdot \bar{t} \right) \otimes (pa + qa') \cdot e_1$$
$$+ \left( \frac{pb}{pb + qb'} \cdot t + \frac{qb'}{pb + qb'} \cdot \bar{t} \right) \otimes (pb + qb') \cdot e_2$$
$$+ \left( \frac{pc}{pc + qc'} \cdot t + \frac{qc'}{pc + qc'} \cdot \bar{t} \right) \otimes (pc + qc') \cdot e_3.$$

Let

$$T := \{ \langle t, e_1 \rangle, \langle t, e_2 \rangle, \langle t, e_3 \rangle \} \subseteq X \times Y$$
$$E_i := \{ \langle t, e_i \rangle, \langle \bar{t}, e_i \rangle \}$$

Think of $T$ and $E_i$ as the events where $t$ and $e_i$ happened respectively.

Then

$$P = P[T] \cdot t \otimes (P[E_1 \mid T] \cdot e_1 + P[E_2 \mid T] \cdot e_2 + P[E_3 \mid T] \cdot e_3)$$
$$+ P[\bar{T}] \cdot \bar{t} \otimes (P[E_1 \mid \bar{T}] \cdot e_1 + P[E_2 \mid \bar{T}] \cdot e_2 + P[E_3 \mid \bar{T}] \cdot e_3)$$
$$= P[T \cap E_1] \cdot t \otimes e_1 + P[T \cap E_2] \cdot t \otimes e_2 + P[T \cap E_3] \cdot t \otimes e_3$$
$$+ P[\bar{T} \cap E_1] \cdot \bar{t} \otimes e_1 + P[\bar{T} \cap E_2] \cdot \bar{t} \otimes e_2 + P[\bar{T} \cap E_3] \cdot \bar{t} \otimes e_3$$
$$= (P[T \mid E_1] \cdot t + P[\bar{T} \mid E_1] \cdot \bar{t}) \otimes P[E_1] \cdot e_1$$
$$+ (P[T \mid E_2] \cdot t + P[\bar{T} \mid E_2] \cdot \bar{t}) \otimes P[E_2] \cdot e_2$$
$$+ (P[T \mid E_3] \cdot t + P[\bar{T} \mid E_3] \cdot \bar{t}) \otimes P[E_3] \cdot e_3$$

Consider $\frac{pa}{pa + qa'} = P[T \mid E_1]$.

Since $p = P[T]$, $a = P[E_1 \mid T]$, $pa + qa' = P[T \cap E_1] + P[\bar{T} \cap E_1] = P[E_1]$, from the above expressions we get the Bayes Theorem

$$P[T \mid E_1] = \frac{P[T] \, P[E_1 \mid T]}{P[E_1]} \blacksquare$$

## Hom

Consider the set of all convex-linear maps

$$A \to B : \mathrm{Pro}_{\mathbb{R}}.$$

Given $f_1, \dots, f_n : A \to B : \mathrm{Pro}_{\mathbb{R}}$, the convex combination

$$p_1 \cdot f_1 + \dots + p_n \cdot f_n : A \to B$$

$$(p_1 \cdot f_1 + \dots + p_n \cdot f_n)(a) = p_1 \cdot f_1(a) + \dots + p_n \cdot f_n(a)$$

makes sense, and the result is again a convex-linear map

$$\begin{aligned}(p_1 \cdot f_1 + \dots + p_n \cdot f_n)(p \cdot a + q \cdot a') &= p_1 \cdot f_1(p \cdot a + q \cdot a') + \dots + p_n \cdot f_n(p \cdot a + q \cdot a') \\ &= p_1 \cdot p \cdot f_1(a) + \dots + p p_n f_n(a) + q p_1 f_1(a') + \dots + q p_n f_n(a') \\ &= p \cdot (p_1 \cdot f_1 + \dots + p_n \cdot f_n)(a) + q \cdot (p_1 \cdot f_1 + \dots + p_n \cdot f_n)(a'),\end{aligned}$$

note that we made an essential use of commutativity.

We denote this possibility space by

$$[A \to B] := A \to B : \mathrm{Pro}_{\mathbb{R}} \quad \text{so} \quad [A \to B] : \mathrm{Pos}.$$

Recall that for $X, Y, Z : \mathrm{Set}$, we have a canonical bijection

$$X \to [Y \to Z] \cong Y \to [X \to Z] : \mathrm{Set}$$

$$f \quad \mapsto \quad \bar{f} := \lambda y.\ \lambda x.\ f(x, y)$$

This generalizes to possibility spaces; let $A, B, C : \mathrm{Pro}_{\mathbb{R}}\mathrm{Set}$.

Then the above bijection restricts to an isomorphism of possibility spaces

$$\overline{[A \to [B \to C]]} \cong [B \to [A \to C]] : \mathrm{Pro}_{\mathbb{R}}\mathrm{Set}$$

since $\overline{p \cdot f + q \cdot g}(b)(a) = (p \cdot f + q \cdot g)(a)(b)$

$$= p \cdot f(a)(b) + q \cdot g(a)(b)$$

$$= p \cdot \bar{f}(b)(a) + q \cdot \bar{g}(b)(a).$$

This allows us to talk about biconvex-linear maps — maps of that take two arguments and are convex-linear in each of them.

**define** Say $f : A \times B \to C : \mathrm{Set}$ is biconvex-linear

iff

(1) $\forall a : A$

$\bar{f}(a, \cdot) : B \to C$ is convex-linear

(2) $\forall b : B$

$\bar{f}(\cdot, b) : A \to C$ is convex-linear.

or equivalently

$$f\Big(\sum_i p_i \cdot a_i, \sum_j q_j \cdot b_j\Big) = \sum_i \sum_j p_i q_j \cdot f(a_i, b_j)$$

Let $[A, B \to C] := \{ f : A \times B \to C \text{ biconvex-linear} \}$ ■

The set of biconvex-linear maps is a possibility space; obviously

$$[A, B \to C] = [A \to [B \to C]] = [B \to [A \to C]] = [B, A \to C].$$

Note that since $\mathrm{Pos}_{\mathbb{R}}[X] \otimes \mathrm{Pos}_{\mathbb{R}}[Y] = \mathrm{Pos}_{\mathbb{R}}[X \times Y]$, we have

$$\big[\mathrm{Pos}_{\mathbb{R}}[X], \mathrm{Pos}_{\mathbb{R}}[Y] \to C\big] = [X \times Y \to C : \mathrm{Set}] = \big[\mathrm{Pos}_{\mathbb{R}}[X] \otimes \mathrm{Pos}_{\mathbb{R}}[Y] \to C\big] : \mathrm{Pos}_{\mathbb{R}}\mathrm{Set},$$

so biconvex-linear maps $f : \mathrm{Pos}_{\mathbb{R}}[X] \times \mathrm{Pos}_{\mathbb{R}}[Y] \to C$ correspond to convex-linear maps $f : \mathrm{Pos}_{\mathbb{R}}[X] \otimes \mathrm{Pos}_{\mathbb{R}}[Y] \to C$.

This can be generalized from biconvex-linear maps on free-possibility spaces $\mathrm{Pos}_{\mathbb{R}}[X], \mathrm{Pos}_{\mathbb{R}}[Y]$, to biconvex-linear maps on arbitrary pair of possibility spaces $A, B$; we'll be able to construct their tensor product $A \otimes B : \mathrm{Pos}_{\mathbb{R}}\mathrm{Set}$ such that biconvex-linear maps $f : A, B \to C$ correspond to convex-linear $f : A \otimes B \to C$.

More concretely, given $a : A$, $b : B$, we'll be able to construct their independent product $a \otimes b : A \otimes B$; but since $A \otimes B$ is a possibility space, expressions such as $30\% \cdot a \otimes b + 60\% \cdot a' \otimes b' + 10\% \cdot a' \otimes b' : A \otimes B$ also make sense, and they need not be factorable into $a \otimes b$, $A \otimes B$ — that is, elements of $A \otimes B$ can be entangled.

We'll be able to "distribute"

$$(p \cdot a + q \cdot a') \otimes (r \cdot b + s \cdot b') = pr \cdot a \otimes b + ps \cdot a \otimes b' + qr \cdot a' \otimes b + qs \cdot a' \otimes b' : A \otimes B.$$

More generally

$$\sum_i p_i \cdot a_i \otimes \sum_j r_j \cdot b_j = \sum_i \sum_j p_i r_j \cdot a_i \otimes b_j : A \otimes B$$

Now we formulate the universal property of $A \otimes B$

**define** Let $A, B$ be possibility spaces, then their tensor product consists of

a possibility space $A \otimes B$ with a biconvex-linear $\otimes : A \times B \to A \otimes B$

where $\forall$ possibility spaces $C$,

$\forall$ biconvex-linear $f : A \times B \to C$,

$\exists!$ convex-linear $\bar{f} : A \otimes B \to C$,

$\forall a : A, b : B$,

$\bar{f}(a \otimes b) = f(a, b)$ ■

**implementation** First instinct might be to try to define $A \otimes B := \mathrm{Pos}_{\mathbb{R}}[A \times B]$, but this space is too big. Note

$f : \mathrm{Pos}_{\mathbb{R}}[A \times B] \to C$ correspond to all set maps $f : A \times B \to C$,

but we need them to correspond to those $f : A \times B \to C$ that are biconvex-linear.

To restrict to biconvex-linear maps, we identify some of the pairs of elements in $\mathrm{Pos}_{\mathbb{R}}[A \times B]$. Denote $a \otimes b := \langle a, b \rangle : \mathrm{Pos}_{\mathbb{R}}[A \times B]$.

Recall that the equivalence relation $\equiv$ on a possibility space $T$ that make $T \equiv$ into a possibility spaces are called congruences; they must preserve the $\mathrm{Pos}_{\mathbb{R}}$-operations. Actually it's enough to check that if $a \equiv a'$, $b \equiv b' : T$, then $p \cdot a + q \cdot b \equiv p \cdot a' + q \cdot b' : T$.

Let us generate the equivalence as follows

$$A \otimes B := \frac{\mathrm{Pos}_{\mathbb{R}}[A \times B]}{\Big\langle \Big(\sum_i p_i \cdot a_i\Big) \otimes \Big(\sum_j r_j \cdot b_j\Big) = \sum_i \sum_j p_i r_j \cdot a_i \otimes b_j \ \Big|\ a_i : A,\ b_j : B \Big\rangle}$$

Now this really implements the tensor product of $A$ and $B$, and the details are spelled out in the Chapter on Commutative Theories in much greater generality ■

Usually students encounter the tensor product for the first time in a more advanced course on Linear Algebra.

One of the biggest problems in Linear Algebra education is that from the beginning the emphasis is on interpreting vector spaces geometrically as $\mathbb{R}^n$. Sure, we hear about function spaces, but no one dwells on these too much. Even the example of the free vector space $\mathrm{Vec}[\text{apple}, \text{orange}]$ would help tremendously, but rarely anyone bothers.

If you interpret everything geometrically, it is very hard to make any sense out of quotients $V/S$, duals $V^*$, and especially tensor products $V \otimes W$, causing the emergence of (very nice) tutorials with names such as

How to lose your fear of tensor products.

Geometry helps out only with direct sums $V \oplus W$.

So tensor product inherit all the misunderstandings caused mostly by geometric interpretation. Tensor products are usually introduced in the following sequence of steps:

0. (optimal) extension by scalars
1. tensor product of free vector spaces
2. (optimal) tensor product of duals $V^* \otimes W^*$
3. universal property of $V \otimes W$
4. implementation $K[V \times W] / S$.

Beside the inadequacy of geometric interpretation the problem with this sequence are that for all quotients $V/S$ are misunderstood because for vector spaces, convexness and congruences are naturally in a bijection, so teachers ignore the much more appropriate concept of congruence.

In 3. the concept of universal property is one of the most important notions in modern mathematics; the pity is that most of the time it is initially encountered in the context of tensor products thereby concentrating the complexity of both at the same time. Why is it that not once is direct sum introduced in terms of its universal property (at least directly).

Is it perhaps because geometric intuition works well here?

Yet there is no recourse to natural examples of tensor products?

But most importantly 1. The first non-trivial tensor product you can make is $\mathbb{R}^2 \otimes \mathbb{R}^2$, yet what is its geometric meaning?

Here is where the obviousness of $\mathrm{Pos}_{\mathbb{R}}[X] \otimes \mathrm{Pos}_{\mathbb{R}}[Y]$ kicks shines, it has a solid probabilistic interpretation, and it feels very natural, there is definitely the feeling of inevitability of the concept of tensor product.

The basic examples are high school level material.

Immediately one can do useful calculations.

On the other hand the most likely first application of tensor product of vector spaces that one encounters is in Abstract Differential Geometry, which is way more inaccessible than Finitary Probability Theory.

## Giry Monad

We've seen that possibilities are just finitary probability measures.

The idea of possibilities can be pushed much further to general probability measures.

Recall that a measurable space is a set $X$ together with a $\sigma$-algebra structure on it $\sigma(X)$ [collection of subsets closed under $\emptyset, X$, complementation, countable unions and intersections].

Will call the elements of $\sigma(X)$ the events on $X$.

Given measurable spaces $X, Y$ a map $f : X \to Y$ : Set is measurable

iff $\forall T$ event of $Y$,

$f^{\leftarrow}(T)$ is an event of $X$.

This gives us the category Meas of measurable spaces.

A probability measure is a map

$$P : \sigma(X) \to [0, 1]$$

where $P[\emptyset] = 0$, $P[X] = 1$,

for mutually disjoint $E_0, E_1, \dots : \sigma(X)$

$$P\Big[\bigcup_{n \geq 0} E_n\Big] = \sum_{n \geq 0}^{\infty} P[E_n].$$

Define $\mathrm{Pro}_{\mathbb{R}}[X] := \{ P : \sigma(X) \to [0, 1] \text{ probability measure} \}$

It is easy to see that countable convex combinations make sense in $\mathrm{Pro}_{\mathbb{R}}[X]$, but $\mathrm{Pro}_{\mathbb{R}}[X]$ has much more structure, its atleast a measurable space.

But first recall that we can combine events and probability measures via integration; given $P : \mathrm{Pro}_{\mathbb{R}}[X]$, $E : \sigma(X)$ we have

$$P[E] := \int_E 1 \cdot dP : [0, 1]$$

The interval $[0; 1]$ is a measurable space (Borel $\sigma$-algebra is generated with open intervals in the Euclidean Topology).

We can combine $P : \mathrm{Pro}_{\mathbb{R}}[X]$ with any measurable $f : X \to [0, 1]$ to get $E[f](P) := P[f] := \int_X f \, dP$

This extends $P[E] := \int_X \chi_E \, dP$ where $\chi_E : X \to [0, 1]$ is the characteristic function of $E$.

We get a map $\int : \mathrm{Pro}_{\mathbb{R}}[X] \times [X \to [0, 1] : \mathrm{Meas}] \to [0, 1]$

The $\sigma$-algebra on $\mathrm{Pro}_{\mathbb{R}} X$ is the least one such that $\forall f : X \to [0, 1] : \mathrm{Meas}$,

$$\int_X f \, d\cdot : \mathrm{Pro}_{\mathbb{R}}[X] \to [0, 1] \text{ is measurable.}$$

So $\mathrm{Pos}_{\mathbb{R}}[X] : \mathrm{Meas}$.

**define** Given $f : X \to Y : \mathrm{Meas}$, $P : \mathrm{Pro}_{\mathbb{R}}[X]$, we can pushforward $P$ along $f$ to

$$f(P) : \mathrm{Pro}_{\mathbb{R}}[Y]$$

$$f(P)[T] := P[f^{\leftarrow}[T]] \ ■$$

This actually makes $\mathrm{Pro}_{\mathbb{R}} : \mathrm{Meas} \to \mathrm{Meas}$ into a functor.

Actually it's a monad on Meas:

**define** $\mu : \mathrm{Pro}_{\mathbb{R}}\mathrm{Pro}_{\mathbb{R}}[X] \to \mathrm{Pro}_{\mathbb{R}}[X]$

$$\mu[\mathbb{P}](E] := \int_{\mathrm{Pro}_{\mathbb{R}}[X]} [E] \, d\mathbb{P}$$

$$= \int_{\mathrm{Pro}_{\mathbb{R}}[X]} \int_E 1 \, dP \, d\mathbb{P} \ ■$$

The unit $\eta : \mathrm{id}_{\mathrm{Meas}} \to \mathrm{Meas}$ is just

**define** $\eta_X : X \to \mathrm{Pro}_{\mathbb{R}}[X]$

$\eta_X[E] := 1$ if $x \in E$ otherwise $0$. ■

This is known as the Giry Monad; see for example Codensity and the Giry Monad by Tom Avery ■

# Commutative Theories

In this chapter we introduce a special class of theories known as commutative theories; they generalise theories of vector and convex spaces so that we can their models can be tensored and homsets become models of the theory.

## Commuting Operations

**example** The theory of monoids is not commutative.

Let $M, N$ be monoids. Is the set of all monoid homomorphisms $A \to B$ Mon canonically a monoid?

First, given $X : Set$ the set of all set maps $X \to N : Set$ carries a pointwise monoid structure: that is given $f, g : X \to N$, $(f \star g)(x) := f(x) \star g(x)$.

The neutral element in $X \to N$ is the constant 1 function.

So $M \to N : Set$ is a monoid; is $M \to N$ Mon a submonoid?

The constant 1 function is a homomorphism.

Let $f, g : M \to N$ Mon; is $f \star g$ a homomorphism?

Note
$$(f \star g)(a \star b) = f(a \star b) \star g(a \star b) = f(a) \star f(b) \star g(a) \star g(b)$$
but
$$(f \star g)(a) \star (f \star g)(b) = f(a) \star g(a) \star f(b) \star g(b),$$
and here we are stuck, since $f(b)$ need not commute with $g(a)$. ■

More generally, fix a finitary finite sorted algebraic theory $\mathbb{T}$.

**example** Then given $X : Set$, $N \mathbb{T} \odot Set$, the set of all set maps $X \to N : Set$ is a $\mathbb{T}$-model: for example for $\omega : {}^{n}\!\bullet \to \mathbb{T}$, we have $\omega(f, g, h)(x) := \omega(f(x), g(x), h(x)) : N$.

Given $M, N : \mathbb{T} \odot Set$, is $M \to N \mathbb{T} \odot Set$ a $\mathbb{T}$-submodel of $M \to N Set$?

Fix $\omega : {}^{2}\!\bullet \to \mathbb{T}$, $f, g, h : M \to N \mathbb{T} \odot Set$. Is $\omega(f, g, h) : M \to N$ a $\mathbb{T}$-homomorphism?

For example does it preserve $\eta : {}^{2}\!\bullet \to \mathbb{T}$? Note
$$\omega(f, g, h)(\eta(a, b)) = \omega(f(\eta(a, b)), g(\eta(a, b)), h(\eta(a, b)))$$
$$= \omega(\eta(f(a), f(b)), \eta(g(a), g(b)), \eta(h(a), h(b)))$$
but $\eta(\omega(f, g, h)(a), \omega(f, g, h)(b)) = \eta(\omega(f(a), g(a), h(a)), \omega(f(b), g(b), h(b)))$.

So we need
$$\omega(\eta(f(a), f(b)), \eta(g(a), g(b)), \eta(h(a), h(b))) = \eta(\omega(f(a), g(a), h(a)), \omega(f(b), g(b), h(b)))$$

When are these expressions equal for all possible models $A, B$ and homomorphisms $f, g, h$?

Fixing $\langle a, b, c, b', a'', b'' \rangle = \mathrm{id} : {}^{6}\!\bullet \to {}^{6}\!\bullet$, it is as if we needed the following operation to be the one
$$\omega(\eta(a, b), \eta(a', b'), \eta(a'', b'')) : {}^{6}\!\bullet \to \mathbb{T},$$
If that is the case we say that $\omega$ commutes with $\eta$.

It might be tempting to write $\omega(\eta, \eta, \eta) = \eta(\omega, \omega)$, but that's wrong, since the arguments in the right expressions are permuted differently.

**define** Let $\omega : {}^{n}\!\bullet \to \mathbb{T}$, $\eta : {}^{m}\!\bullet \to \mathbb{T}$.

Say $\omega$ commutes with $\eta$ if
$$\omega(\eta(a_1^1, \dots, a_m^1), \dots, \eta(a_1^n, \dots, a_m^n)) = \eta(\omega(a_1^1, \dots, a_1^n), \dots, \omega(a_m^1, \dots, a_m^n)) : {}^{nm}\!\bullet \to \mathbb{T},$$
written more compactly
$$\omega(\lambda_i \eta(\lambda_j a_i^j)) = \eta(\lambda_j \omega(\lambda_i a_i^j)) : {}^{mn}\!\bullet \to \mathbb{T}.$$ ■

Note that if $\omega$ commutes with $\eta$, then so does $\eta$ commute with $\omega$.

**example** Two constants $c, c' : 1 \to \mathbb{T}$ commute if $c = c'$. So each constant commutes with itself.

Given a nullary operation $u : \bullet \to \mathbb{T}$, it commutes with a constant $c$ if $u(c) = c : 1 \to \mathbb{T}$.

A $\omega : {}^{2}\!\bullet \to \bullet$ commutes with $c : 1 \to \bullet$ if $\omega(c, c) = c : 1 \to \bullet$ and so on. ■

**example** Two nullary operations $u, t : \bullet \to \bullet$ commute if $u.t = t.u : \bullet \to \bullet$.

So any nullary operation commutes with itself. ■

**example** A nullary $u : \bullet \to \bullet$ commutes with binary $\omega : {}^{2}\!\bullet \to \bullet$ if $\omega(u(a), u(b)) = u(\omega(a, b)) : {}^{2}\!\bullet \to \bullet$. ■

**example** The binary operations $\omega, \eta : {}^{2}\!\bullet \to \bullet$ commute if $\omega(\eta(a, b), \eta(a', b')) = \eta(\omega(a, a'), \omega(b, b'))$.

Writing them in an infix prefixed notation $\cdot, \otimes : {}^{2}\!\bullet \to \bullet$, this says $(a \cdot b) \otimes (a' \cdot b') = (a \otimes a') \cdot (b \otimes b')$.

We have encountered something similar when we considered bifunctor $\otimes : \mathbb{A} \times \mathbb{B} \to \mathbb{C}$, then bifunctoriality states for $x \xrightarrow{f} x' \xrightarrow{g} x'' : \mathbb{A}$, $y \xrightarrow{f'} y' \xrightarrow{g'} y'' : \mathbb{B}$ we have
$$(g.f) \otimes (g'.f') = (g \otimes g').(f \otimes f') : x \otimes y \to x'' \otimes y'' : \mathbb{C}.$$
Changing notation once again, let $1, -: A \times A \to A : Set$, we'll write $a \mid b$ and $\frac{a}{b}$.

Then the above condition becomes a kind of 2-dimensional associativity
$$\frac{a \mid b}{c \mid d} = \frac{a}{c} \Big| \frac{b}{d}$$
Assuming the operations share a neutral element $0 : A$, and are both associative,
$$a \mid b = \frac{a}{0} \Big| \frac{0}{b} = \frac{a \mid 0}{0 \mid b} = \frac{a}{b}$$
that is both operations are equal.

When does a binary operation $\cdot : {}^{2}\!\bullet \to \bullet$ commute with itself? If $(x \star (a)) \star (b \star y) = (x \star b) \star (a \star y) : {}^{4}\!\bullet \to \bullet$.

This is not quite the commutativity $a \star b = b \star a$.

If $\star$ has a neutral constant $e : 1 \to \bullet$ so that $e \star x = x = x \star e : \bullet \to \bullet$, then self commutation implies commutativity.

It even must be associative since
$$a \star (b \star c) = (a \star 1) \star (b \star c) = (a \star b) \star (1 \star c) = (a \star b) \star c$$
This is known as the Eckmann–Hilton argument.

**prop** Any two commuting operations such that they have an element 1 that's left neutral for one operation and right neutral for the other must be the same commutative, associative operation. ■

As an application, let $\mathbb{E}$ be a category, then we have the monoidal category $\mathrm{End}(\mathbb{E})$.

Consider endomorphisms of the identity functor $\mathrm{id}_{\mathbb{E}} : \mathbb{E} \to \mathbb{E}$, that is natural transformations $\alpha : \mathrm{id}_{\mathbb{E}} \Rightarrow \mathrm{id}_{\mathbb{E}}$.

They can be combined via the composition
$$\mathrm{id}_{\mathbb{E}} \xRightarrow{\alpha} \mathrm{id}_{\mathbb{E}} \xRightarrow{\beta} \mathrm{id}_{\mathbb{E}} \rightsquigarrow \mathrm{id}_{\mathbb{E}} \xRightarrow{\beta.\alpha} \mathrm{id}_{\mathbb{E}},$$
and they can be tensored
$$\mathrm{id}_{\mathbb{E}} \xRightarrow{\alpha} \mathrm{id}_{\mathbb{E}} \rightsquigarrow \mathrm{id}_{\mathbb{E}} \xRightarrow{\alpha \otimes \beta} \mathrm{id}_{\mathbb{E}}$$
$$\mathrm{id}_{\mathbb{E}} \xRightarrow{\beta} \mathrm{id}_{\mathbb{E}}$$

So $\mathrm{End}(\mathrm{id}_{\mathbb{E}})$ is equipped with two operations. Of course they both commute.

But $\mathrm{id} : \mathrm{id}_{\mathbb{E}} \Rightarrow \mathrm{id}_{\mathbb{E}}$ is a neutral element for both
$$\mathrm{id}.\alpha = \alpha = \alpha.\mathrm{id},$$
$$\mathrm{id} \otimes \alpha = \alpha = \alpha \otimes \mathrm{id},$$
hence by Eckmann–Hilton $\alpha.\beta = \beta.\alpha = \beta \otimes \alpha = \alpha \otimes \beta$.

The commutative monoid $\mathrm{End}(\mathrm{id}_{\mathbb{E}})$ is called the center of category $\mathbb{E}$; when $\mathbb{E} = M$ is a monoid (one object category) then the categorical center coincides with the monoid center $\{c : M \mid \forall m : M.\ cm = mc\}$. ■

Let $\omega : {}^{5}\!\bullet \to \bullet$, $\eta : {}^{4}\!\bullet \to \bullet$. Assuming $\langle a, b, c, d, e, f \rangle = \mathrm{id} : {}^{6}\!\bullet \to {}^{6}\!\bullet$, let us denote the commutation of $\omega$ with $\eta$ by
$$\omega \otimes \eta(a, b, c, d, e, f) := \omega(\eta(a, b), \eta(c, d), \eta(e, f)).$$
Generally,

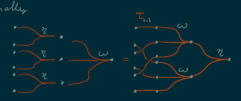
**define** For $\omega : {}^{n}\!\bullet \to \bullet$, $\eta : {}^{m}\!\bullet \to \bullet$, $\langle x_1, x_2, \dots, x_{nm} \rangle = \mathrm{id} : {}^{nm}\!\bullet \to {}^{nm}\!\bullet$ set
$$\omega \otimes \eta(x_1, \dots, x_{nm}) := \omega(\eta(x_1, \dots, x_n), \dots, \eta(x_{(n-1)m+1}, \dots, x_{nm})).$$ ■

Assuming $\omega : {}^{n}\!\bullet \to \bullet$ commutes with $\eta : {}^{m}\!\bullet \to \bullet$ doesn't mean
$$\omega \otimes \eta = \eta \otimes \omega.$$
Assume $\langle a, b, a', b', a'', b'' \rangle = \mathrm{id} : {}^{6}\!\bullet \to {}^{6}\!\bullet$, then
$$\omega \otimes \eta = \omega(\eta(a, b), \eta(a', b'), \eta(a'', b''))$$
but $\eta \otimes \omega = \eta(\omega(a, b, a''), \omega(b', a', b''))$.

and not $\eta(\omega(a, a', a''), \omega(b, b', b''))$.

We denote the permutation by $\tau_{3,2} : {}^{6}\!\bullet \to {}^{6}\!\bullet$ so that
$$\tau_{3,2}(a, b, c, d, e, f) = \langle a, c, e, b, d, f \rangle$$
so that $\eta \otimes \omega = \omega \otimes \eta . \tau_{3,2}$.

Visually




**define** An operation $\omega : {}^{n}\!\bullet \to \mathbb{T}$ is central if $\forall \eta : {}^{m}\!\bullet \to \mathbb{T}$ $\omega$ and $\eta$ commute. ■

Central operation are special because their interpretations are $\mathbb{T}$-homomorphisms. For example consider a central $\omega : {}^{2}\!\bullet \to \bullet$; then given $M, \mathbb{T} \odot Set$, we have
$$\omega : M \times M \times M \to M,$$
and for $\eta : {}^{3}\!\bullet \to \bullet$,
$$\omega(\eta((a, b, c), (a', b', c'))) = \omega(\eta(a, a'), \eta(b, b'), \eta(c, c'))$$
$$= \eta(\omega(a, b, c), \omega(a', b', c'))$$
that is $\omega : M^3 \to M$ preserves $\eta : {}^{3}\!\bullet \to \bullet$.

More generally,

**prop** Let $\omega : {}^{n}\!\bullet \to \mathbb{T}$, $\eta : {}^{m}\!\bullet \to \mathbb{T}$. Then

(1) $\omega$ commutes with $\eta$

$\Leftrightarrow$ (2) $\forall M \mathbb{T} \odot Set$
$$\omega : M^n \to M \text{ preserves } \eta$$

$\Leftrightarrow$ (3) $\forall f_1, \dots, f_n : M \to N : \mathbb{T} \odot Set$, the map
$$\omega(f_1, \dots, f_n) : M \to N \text{ preserves } \eta$$

**proof** We've already seen that (1) $\Rightarrow$ (2), (3).

It is easy to see that (2) $\Leftrightarrow$ (3) since $\omega = \omega(\pi_1^n, \dots, \pi_n^n)$, and (3) $\Rightarrow$ (2).

We'll show (2) $\Rightarrow$ (1). Assume (2).

We'll illustrate the proof for $n = 3$, $m = 2$.

Consider $\mathbb{T}[a, b, a', b', a'', b'']$. We can think that the operation $\omega \otimes \eta$ and $\eta \otimes \omega$ $\tau_{3,2}$ live in $\mathbb{T}[a, b, a', b', a'', b'']$.

We can interpret $\omega$, in one set as $\mathbb{T}$-homomorphism
$$\omega : \mathbb{T}[a, b, a', b', a'', b'']^3 \to \mathbb{T}[a, b, a', b', a'', b''].$$
Then it preserves $\eta$ at inputs $\langle a, b \rangle$, $\langle a', b' \rangle$, $\langle a'', b'' \rangle$, so
$$\omega(\eta(a, b), \eta(a', b'), \eta(a'', b'')) = \eta(\omega(a, a', a''), \omega(b, b', b'')) \ ■$$

**define** $M : \mathbb{T} \odot Set$ is commutative if $\forall \omega : {}^{n}\!\bullet \to \mathbb{T}$
$$[\omega] : M^n \to M \text{ is a } \mathbb{T}\text{-homomorphism}. \ ■$$

Obviously

**prop** Let $M, N : \mathbb{T} \odot Set$ where $N$ is commutative. Then $M \to N \mathbb{T} \odot Set$ is a $\mathbb{T}$-model. ■

## Commutative Theories

**define** A single sorted algebraic theory $\mathbb{A}$ is commutative if any of the equivalent conditions hold

(1) $\forall \omega : {}^{n}\!\bullet \to \bullet$, $\eta : {}^{m}\!\bullet \to \bullet : \mathbb{A}$, $\omega$ and $\eta$ commute.

(2) Every model of $\mathbb{A}$ is commutative.

(3) $\forall M, N : \mathbb{A} \odot Set$,
$$M \to N \mathbb{A} \odot Set \text{ is an } \mathbb{A}\text{-model}. \ ■$$

For a commutative theory $\mathbb{A}$.

Obviously

**prop** Commutative Theories $\mathbb{A}$ has at most one constant, many operations $\bullet \to \bullet : \mathbb{A}$ form a commutative monoid.

**example** The following theories are commutative:

theories of commutative monoids, abelian groups, sets, pointed sets, vector spaces over a field, modules over a commutative ring, $M$-sets for commutative monoid $M$, convex spaces, affine spaces, semi lattices.

The following theories are not commutative:

theories of monoids, groups, rings, $M$-sets for non-commutative monoid $M$, but also commutative rings, boolean algebras or any theory with more than one constant. ■

## Hom and Bihomomorphisms

**define** Let $M, N : \mathbb{A}$, set
$$[M \to N] := \{f : M \to N : \mathbb{A} \odot Set\}$$
Obviously $[M \to N]$ is an $\mathbb{A}$-model. ■

**define** (1) Let $M, N, S : \mathbb{A} \odot Set$. Then a $\mathbb{A}$-bihomomorphism $f : M, N \to S : \mathbb{A} \odot Set$ is a map $f : M \times N \to S : Set$ where

(1) $\forall a : M$,
$$f(a, -) : N \to S \text{ is an } \mathbb{A}\text{-homomorphism}$$

(2) $\forall b : M$,
$$f(-, b) : M \to S \text{ is an } \mathbb{A}\text{-homomorphism}.$$

Equivalently $\forall \omega : {}^{n}\!\bullet \to \bullet$, $\eta : {}^{m}\!\bullet \to \bullet : \mathbb{A}$, $\lambda_i a_i : M^n$, $\lambda_j b_j : N^m$
$$\omega(\lambda_i \eta(\lambda_j f(a_i, b_j))) = f(\omega(\lambda_i a_i), \eta(\lambda_j b_j)) = \eta(\lambda_j \omega(\lambda_i f(a_i, b_j)))$$
(of course since $\mathbb{A}$ is commutative, only one of the equalities is needed)

(11) Let $M_1, \dots, M_k, S : \mathbb{A} \odot Set$, then an $\mathbb{A}$-$k$-homomorphism $f : M_1, \dots, M_k \to S : \mathbb{A} \odot Set$ is a map $f : M_1 \times \dots \times M_k \to S : Set$ where $\forall \omega_i : {}^{n_i}\!\bullet \to \bullet, \dots, \omega_k : {}^{n_k}\!\bullet \to \bullet : \mathbb{A}$,
$$f(\omega_1(\lambda_{i_1} a_{i_1}), \dots, \omega_k(\lambda_{i_k} a_{i_k})) = \omega_1(\lambda_{i_1} \cdots \omega_k(\lambda_{i_k} f(a_{i_1}, \dots, a_{i_k})) \cdots)$$
Denote the set of $n$-homomorphisms by $[M_1, \dots, M_n \to S]$. ■

**prop** Let $M, N, S : \mathbb{A} \odot Set$, then
$$[M \to [N \to S]] \overset{(1)}{\cong} [M, N \to S] : Set$$
$$[M \to [N \to S]] \overset{(11)}{\cong} [N \to [M \to S]]$$

**proof** (1) Let $f : M \to [N \to S] : Set$; define $\bar{f} : M \times N \to S$, $\bar{f}(a, b) := f(a)(b)$.

Then our task is to show
$$f : [M \to [N \to S]] \Leftrightarrow \bar{f} : [M, N \to S].$$

$\Rightarrow$: $\bar{f}(\omega(\lambda_i a_i), \eta(\lambda_j b_j)) = f(\omega(\lambda_i a_i))(\eta(\lambda_j b_j))$
$$= \eta(\lambda_j f(\omega(\lambda_i a_i))(b_j))$$
$$= \eta(\lambda_j \omega(\lambda_i f(a_i)(b_j)))$$
$$= \eta(\lambda_j \omega(\lambda_i \bar{f}(a_i, b_j)))$$

$\Leftarrow$: We'll show $f(\omega(\lambda_i a_i)) = \omega(\lambda_i f(a_i))$:
$$f(\omega(\lambda_i a_i))(b) = \bar{f}(\omega(\lambda_i a_i), b)$$
$$= \omega(\lambda_i \bar{f}(a_i, b))$$
$$= \omega(\lambda_i f(a_i)(b))$$
$$= \omega(\lambda_i f(a_i))(b).$$
This allows us to transport $\mathbb{A}$-module structure from $[M \to [N \to S]]$ to $[M, N \to S]$.

(2) Let $f : M \to [N \to S] : Set$; define $\bar{f} : N \to M \to S$, $\bar{f}(b)(a) := f(a)(b)$.

Then $f : M \to [N \to S] \Leftrightarrow \bar{f} : N \to [M \to S]$ is obvious.

We need to show that $\bar{(-)}$ is a homomorphism:
$$\overline{\omega(\lambda_i f_i)}(b)(a) = \omega(\lambda_i f_i)(a)(b)$$
$$= \omega(\lambda_i f_i(a))(b)$$
$$= \omega(\lambda_i f_i(a)(b))$$
$$= \omega(\lambda_i \bar{f}_i(b)(a))$$
$$= \omega(\lambda_i \bar{f}_i(b))(a). \ ■$$

# Symmetric Closed Monoidal Structure

Now we formulate the universal property of tensor products of modules.

**define** Let $M, N : A\text{-}\mathbf{Set}$. Then a tensor product of $M$ and $N$ consists of

$$M \otimes N : A\text{-}\mathbf{Set} \text{ with } \otimes : [M, N \multimap M \otimes N]$$

where $\forall f : [M, N \multimap S]$,

$$\exists ! \bar{f} : [M \otimes N \multimap S],$$
$$\bar{f}(a \otimes b) = f(a, b). \blacksquare$$

We implement it for the special case of free models.

Recall that for $X : \mathbf{Set}$, $M : A\text{-}\mathbf{Set}$, to define a homomorphism $f : \mathbb{A}[X] \longrightarrow M\ A\text{-}\mathbf{Set}$, it's enough to specify its restriction $f : X \longrightarrow M : \mathbf{Set}$.

Also at one think of $\mathbb{A}[X]$ as the set of all operations $\omega : \overset{X}{\bullet\!\cdots\!\bullet} \to \bullet : A$, and a map $f : X \to M : \mathbf{Set}$ as an $X$-tuple of $M$-elements $\underset{i}{\lambda} a_i : M^X$, then its lift on $\omega : \overset{X}{\bullet\!\cdots\!\bullet} \to A$ is $\underset{i}{\lambda} \overline{a_i}(\omega) = \omega(\underset{i}{\lambda} a_i) : M$.

Back to tensor products.

Let $X, Y : \mathbf{Set}$. A candidate for $\mathbb{A}[X] \otimes \mathbb{A}[Y]$ we take the free model $\mathbb{A}[X \times Y]$.

We need a bihomomorphism

$$\otimes : \mathbb{A}[X] \times \mathbb{A}[Y] \longrightarrow \mathbb{A}[X \times Y];$$

there are two candidates corresponding to

$$\mathbb{A}[X] \longrightarrow [\mathbb{A}[Y] \multimap \mathbb{A}[X \times Y]] : A\text{-}\mathbf{Set}$$
$$x \rightsquigarrow y \rightsquigarrow \langle x, y \rangle$$

and using the canonical iso $[m \multimap [n \multimap s]] \cong [n \multimap [m \multimap s]]$,

$$\mathbb{A}[Y] \longrightarrow [\mathbb{A}[X] \multimap \mathbb{A}[X \times Y]] : A\text{-}\mathbf{Set}.$$
$$y \rightsquigarrow x \rightsquigarrow \langle x, y \rangle$$

To clarify consider the following calculation with $X = \{a, b, c\}$, $Y = \{\eta, z\}$, $X \times Y = \{a_\eta, a_z, b_\eta, b_z, c_\eta, c_z\}$.

$$\mathbb{A}[a, b, c] \longrightarrow \mathbb{A}[\eta, z] \longrightarrow \mathbb{A}[a_\eta, a_z, b_\eta, b_z, c_\eta, c_z]$$

| | | |
|---|---|---|
| $x$ | $y$ | $\langle x, y \rangle$ |
| $a$ | $\eta$ | $a_\eta$ |
| | $z$ | $a_z$ |
| | $\eta$ | $\eta(a_\eta, a_\eta)$ |
| $b$ | $z$ | $\eta(b_\eta, b_z)$ |
| $c$ | $\eta$ | $\eta(c_\eta, c_\eta)$ |
| $\omega$ | $\eta$ | $\omega(\eta(a_\eta, a_\eta), \eta(b_\eta, b_z), \eta(c_\eta, c_\eta))$ |

and

$$\mathbb{A}[\eta, z] \longrightarrow \mathbb{A}[a, b, c] \longrightarrow \mathbb{A}[a_\eta, a_z, b_\eta, b_z, c_\eta, c_z]$$

| | | |
|---|---|---|
| $y$ | $x$ | $\langle x, y \rangle$ |
| $\eta$ | $a$ | $a_\eta$ |
| | $b$ | $b_\eta$ |
| | $c$ | $c_\eta$ |
| | $\omega$ | $\omega(a_\eta, b_\eta, c_\eta)$ |
| $z$ | $\omega$ | $\omega(a_z, b_z, c_z)$ |
| $\eta$ | $\omega$ | $\eta(\omega(a_\eta, b_\eta, c_\eta), \omega(a_z, b_z, c_z))$. |

These are the two ways to commutate operations $\omega$ and $\eta$!

And since $A$ is commutative, these two ways are the same, so we have a canonical bihomomorphism $\otimes : \mathbb{A}[X] \times \mathbb{A}[Y] \longrightarrow \mathbb{A}[X \times Y]$, and we see that commutativity of $A$ is essential.

The uniqueness of this map can also be used as the definition of a commutative monad.

**prop** Let $X, Y : \mathbf{Set}$, then we can implement the tensor product of $\mathbb{A}[X]$, $\mathbb{A}[Y]$ as

$$\mathbb{A}[X] \otimes \mathbb{A}[Y] := \mathbb{A}[X \times Y] \text{ with } \otimes : [X, Y \multimap \mathbb{A}[X \times Y]].$$

**proof** Let $f : \mathbb{A}[X] \times \mathbb{A}[Y] \longrightarrow S$ be a bihomomorphism.

Then we can restrict it to $f|_{X \times Y} : X \times Y \longrightarrow S : \mathbf{Set}$.

This lifts to $\overline{f|_{X \times Y}} : \mathbb{A}[X \times Y] \longrightarrow S\ A\text{-}\mathbf{Set}$, and

$$\overline{f|_{X \times Y}}(x, y) = f|_{X \times Y}(x, y) = f(x, y).$$

Given another $g : \mathbb{A}[X \times Y] \longrightarrow S : A\text{-}\mathbf{Set}$ s.t. $g(x, y) = f(x, y)$, we must have $g = \overline{f|_{X \times Y}}$. $\blacksquare$

Now for the general tensor product.

Let $M, N : A\text{-}\mathbf{Set}$. Denote $a \otimes b := \eta(a, b) : \mathbb{A}[M \times N]$.

On $\mathbb{A}[M \times N]$ define an $A$-congruence $\simeq$ generated by

$$\omega(\underset{i}{\lambda} a_i) \otimes \eta(\underset{j}{\lambda} b_j) \simeq \omega\left(\underset{i}{\lambda} \eta(\underset{j}{\lambda} a_i \otimes b_j)\right) \quad \Big| \quad \omega : \overset{n}{\bullet\!\cdots\!\bullet} \to \bullet : A,$$
$$\simeq \eta\left(\underset{j}{\lambda} \omega(\underset{i}{\lambda} a_i \otimes b_j)\right) \quad \Big| \quad \underset{i}{\lambda} a_i : M^n,\ \underset{j}{\lambda} b_j : N^n.$$

Obviously the $\otimes$ map corestricts to a bihomomorphism

$$\otimes : M \times N \longrightarrow \frac{\mathbb{A}[M \times N]}{\simeq}$$

**prop** $M \otimes N := \dfrac{\mathbb{A}[M \times N]}{\simeq}$ with $\otimes : [M, N \multimap \dfrac{\mathbb{A}[M \times N]}{\simeq}]$ implements the tensor product of $M, N$.

**proof** Let $f : [M, N \multimap S]$.

(Left)

Since $f : M \times N \multimap S : \mathbf{Set}$, it lifts to $\bar{f} : \mathbb{A}[M \times N] \longrightarrow S : A\text{-}\mathbf{Set}$.

It respects the congruence $\simeq$, since

$$\bar{f}\left(\omega(\underset{i}{\lambda} a_i) \otimes \eta(\underset{j}{\lambda} b_j)\right) = f\left(\omega(\underset{i}{\lambda} a_i), \eta(\underset{j}{\lambda} b_j)\right) : S$$
$$= \omega\left(\underset{i}{\lambda} \eta(\underset{j}{\lambda} f(a_i, b_j))\right)$$

and $\bar{f}\left(\omega\left(\underset{i}{\lambda} \eta(\underset{j}{\lambda} a_i \otimes b_j)\right)\right) = \omega\left(\underset{i}{\lambda} \bar{f}(\eta(\underset{j}{\lambda} a_i \otimes b_j))\right) : S$

$$= \omega\left(\underset{i}{\lambda} \eta(\underset{j}{\lambda} f(a_i \otimes b_j))\right)$$
$$= \omega\left(\underset{i}{\lambda} \eta(\underset{j}{\lambda} f(a_i \otimes b_j))\right).$$

Similarly $\bar{f}\left(\omega(\underset{i}{\lambda} a_i) \otimes \eta(\underset{j}{\lambda} b_j)\right) = \bar{f}\left(\eta\left(\underset{j}{\lambda} \omega(\underset{i}{\lambda} a_i \otimes b_j)\right)\right)$.

So $\bar{f}$ restricts to $\bar{f} : M \otimes N \longrightarrow S : A\text{-}\mathbf{Set}$.

(Uniqueness)

Let $g : M \otimes N \longrightarrow S : A\text{-}\mathbf{Set}$ where $g(a \otimes b) = f(a, b)$.

Note that this gives us a map $g : \mathbb{A}(M \times N) \longrightarrow S\ A\text{-}\mathbf{Set}$ that respects $\simeq$.

Then $g \otimes : M \times N \longrightarrow S$ is a bihomomorphism and $g \otimes = f$.

So $\overline{g \otimes} = \bar{f} : \mathbb{A}[M \times N] \longrightarrow S$. But $\overline{g \otimes} = g : \mathbb{A}(M \times N) \to S$ (they agree on $M \times N$), so $g = \bar{f} : M \otimes N \longrightarrow S$.

Obviously the restriction

$$[M \otimes N \multimap S] \longrightarrow [M, N \multimap S]$$
$$f \longrightarrow f|_{M \times N}$$

is an $A$-homomorphism, so $[M \otimes N \multimap S] \cong [M, N \multimap S]\ A\text{-}\mathbf{Set}$. $\blacksquare$

With the tensor product model $M \otimes N$ and the Hom model $[M \multimap N]$ the category of models $A\text{-}\mathbf{Set}$ becomes a symmetric monoidal closed category.

**define** Let $f : M \to M'$, $g : N \to N' : A\text{-}\mathbf{Set}$, define $f \otimes g : M \otimes N \longrightarrow M' \otimes N' : A\text{-}\mathbf{Set}$,

$$(f \otimes g)(a \otimes b) := f(a) \otimes g(b). \blacksquare$$

This makes tensor product into a bifunctor

$$\otimes : A\text{-}\mathbf{Set} \times A\text{-}\mathbf{Set} \longrightarrow A\text{-}\mathbf{Set}$$

since

**prop** $\mathrm{id}_M \otimes \mathrm{id}_N = \mathrm{id}_{M \otimes N}$,

for $M \xrightarrow{f} M' \xrightarrow{f'} M''$, $N \xrightarrow{g} N' \xrightarrow{g'} N'' : A\text{-}\mathbf{Set}$,

$$(f' f) \otimes (g' g) = (f' \otimes g') \circ (f \otimes g) : M \otimes N \longrightarrow M'' \otimes N'' : A\text{-}\mathbf{Set}. \blacksquare$$

The hom functor $[- \multimap -] : A\text{-}\mathbf{Set}^{op} \times A\text{-}\mathbf{Set} \longrightarrow \mathbf{Set}$ lifts to $[- \multimap -] : A\text{-}\mathbf{Set}^{op} \times A\text{-}\mathbf{Set} \longrightarrow A\text{-}\mathbf{Set}$.

The composition is a bihomomorphism

**prop** For $\omega : \overset{k}{\bullet\!\cdots\!\bullet} \to \bullet$, $\eta : \overset{l}{\bullet\!\cdots\!\bullet} \to \bullet : A$, $\underset{i}{\lambda} f_i : [M \multimap N]^k$, $\underset{j}{\lambda} g_j : [N \multimap S]^l$, we have

$$\eta(\underset{j}{\lambda} g_j) \circ \omega(\underset{i}{\lambda} f_i) = \eta(\underset{j}{\lambda} \omega(\underset{i}{\lambda} g_j f_i))$$

is equal to the lift of composition in

$$[N \multimap S] \otimes [M \multimap N] \longrightarrow [M \multimap S] : A\text{-}\mathbf{Set}$$

**proof**

$$(\eta(\underset{j}{\lambda} g_j) \circ \omega(\underset{i}{\lambda} f_i))(m) = \eta(\underset{j}{\lambda} g_j)\left(\omega(\underset{i}{\lambda} f_i)(m)\right)$$
$$= \eta(\underset{j}{\lambda} g_j)\left(\omega(\underset{i}{\lambda} f_i(m))\right)$$
$$= \eta(\underset{j}{\lambda} g_j(\omega(\underset{i}{\lambda} f_i(m)))$$

$$\eta(\underset{j}{\lambda} \omega(\underset{i}{\lambda} g_j f_i))(m) = \eta(\underset{j}{\lambda}, \omega(\underset{i}{\lambda} g_j f_i)(m))$$
$$= \eta(\underset{j}{\lambda}, \omega(\underset{i}{\lambda} g_j(f_i(m)))$$
$$= \eta(\underset{j}{\lambda} g_j(\omega(\underset{i}{\lambda} f_i(m))). \blacksquare$$

**prop** We have canonically

(a) $[M \otimes N \multimap S] \cong [M \multimap [N \multimap S]]$,

(b) $M \otimes N \cong N \otimes M$

(c) $A[1] \otimes M \cong M \cong M \otimes A[1]$

(d) $M \otimes (N \otimes S) \cong (M \otimes N) \otimes S$

**proof** (a) That we have such a bijection is just a syntactic reformulation of the universal property of tensor products since

$$[M \multimap [N \multimap S]] \cong [M, N \multimap S].$$

(c)

$$\frac{M \otimes N \longrightarrow T : A\text{-}\mathbf{Set}}{\dfrac{\dfrac{M \longrightarrow [N \multimap T] : A\text{-}\mathbf{Set}}{N \longrightarrow [M \multimap T] : A\text{-}\mathbf{Set}}}{N \otimes M \longrightarrow T : A\text{-}\mathbf{Set}}}$$

(d)

$$\frac{A[1] \otimes M \longrightarrow T : A\text{-}\mathbf{Set}}{\dfrac{\dfrac{A[1] \longrightarrow [M \multimap T] : A\text{-}\mathbf{Set}}{1 \longrightarrow [M \multimap T] : \mathbf{Set}}}{M \longrightarrow T : A\text{-}\mathbf{Set}}}$$

(e) To prove associativity we would have to develop the notion of a triple tensor product $M \otimes N \otimes S$: its universal property is

$$[M \otimes N \otimes S \multimap T] \cong [M, N, S \multimap T].$$

Also we would need a few obvious properties of trihomomorphisms.

$$\frac{M \otimes (N \otimes S) \longrightarrow T : A\text{-}\mathbf{Set}}{\dfrac{\dfrac{M \longrightarrow [N \otimes S \multimap T] : A\text{-}\mathbf{Set}}{M \longrightarrow [N, S \multimap T] : A\text{-}\mathbf{Set}}}{[M, N, S \longrightarrow T]}}$$

$$\frac{M \otimes N \otimes S \longrightarrow T : A\text{-}\mathbf{Set}}{[M, N \multimap [S \multimap T]] \cong [M, N, S \multimap T]}. \blacksquare$$

Similarly for $(M \otimes N) \otimes S$ where we would have to use

**example** For the theory $\mathbf{Vec}_K$ of vector spaces over a field $K$ the symmetric monoidal closed structure on $\mathbf{Vec}_K := \mathbf{Vec}_K\ \mathbf{Set}$ is the usual Hom and tensor product.

Similarly for the theory $\mathbf{ab}$ of abelian groups, mod$_R$ of modules over a commutative ring $R$ or commutative algebra $A$ over field $K$. $\blacksquare$

**example** The initial theory of sets determines the cartesian closed structure on $\mathbf{Set}$. $\blacksquare$

**example** Given pointed sets $X, Y, Z$ denote the special element by $0$; then a bihomomorphism $f : X \times Y \to Z$ is just a map $f : X \times Y \to Z : \mathbf{Set}$ where $f(0, b) = 0 = f(a, 0)$.

Hence tensor product $X \otimes Y$ is the quotient

$$\frac{X \times Y}{\langle 0, b \rangle = \langle a, 0 \rangle = \langle a, 0 \rangle \mid a : X,\ b : Y}$$

also known as smash product & denoted $X \wedge Y$ for some reason. $\blacksquare$

# Models with Monoid Structure

What about the commutative rings?

The theory of commutative rings is not commutative,

yet the tensor product of commutative rings still makes sense.

First note that the theory of abelian groups $\mathsf{ab}$ is commutative,

so $\mathsf{Ab} := \mathsf{ab}\circ\mathsf{Set}$ is symmetric monoidal closed category.

So given $A, B : \mathsf{Ab}$, it makes sense to talk about $A \otimes B$, $A \to B$ : $\mathsf{Ab}$;

bilinear maps $f : A \times B \longrightarrow T : \mathsf{Ab}$ are just bilinear maps

$$f(n\,a + m\,a', k\,b + l\,b') = nk\,f(a,b) + nl\,f(a,b') + mk\,f(a',b) + ml\,f(a',b')$$

Since $\mathsf{Ab}$ is atleast monoidal, it makes sense to talk about monoid objects in $\mathsf{Ab}$

Its just an $A : \mathsf{Ab}$ with $\times : A \otimes A \longrightarrow A$, $1 : A$

where (a) $a \times 1 = a = 1 \times a$

(a)
$$\begin{tikzcd} A\otimes A\otimes A \arrow[r, "\mathrm{id}\otimes\times"] \arrow[d, "\times\otimes\mathrm{id}"'] & A\otimes A \arrow[d, "\times"] \\ A\otimes A \arrow[r, "\times"'] & A \end{tikzcd}$$

This amounts to the same thing as a ring.

Since $\mathsf{Ab}$ is symmetric monoidal, we can define commutative monoid objects – they satisfy

$$\begin{tikzcd} A\otimes A \arrow[rr, "\sigma"] \arrow[dr, "\times"'] & & A\otimes A \arrow[dl, "\times"] \\ & A & \end{tikzcd}$$

where $\sigma : \forall A, B : \mathsf{Ab}$,
$$A\otimes B \xrightarrow{\sim} B\otimes A$$
is the canonical symmetric map.

Commutative monoids in $\mathsf{Ab}$ are just commutative rings.

In fact $\mathsf{Mon}(\mathsf{Ab}) = \mathsf{Ring}$, $\mathsf{CMon}(\mathsf{Ab}) = \mathsf{CRing}$.

We'll look at such a situation more generally.

Fix $\mathcal{A}$ a commutative theory, then a monoid object in $\mathcal{A}\circ\mathsf{Set}$ amounts to

**define** A monoid in $\mathcal{A}\circ\mathsf{Set}$ consists of

$A : \mathcal{A}\circ\mathsf{Set}$ with $\times : [A, A \longrightarrow A]$, $1 : A$

where (monoidal) $1 \times a = a = a \times 1$

(associativity) $a \times (b \times c) = (a \times b) \times c$.

If it is commutative

"$\Psi$" (commutativity) $a \times b = b \times a$. ■

Is the tensor product of monoids canonically a monoid?

Not necessarily.

**define** Let $A, B$ be monoids. We make $A \otimes B$ into a monoid via:

(neutral) $1_{A\otimes B} := 1_A \otimes 1_B : A \otimes B$.

(mult) Define a binary operation on $A \otimes B$ on the generators of $A \otimes B$

$$(a \otimes b) \times (a' \otimes b') := (a \times a') \otimes (b \times b') \;■$$

Does this make sense? Actually this only defines a map

$$(A \times B) \times (A \times B) \longrightarrow A \otimes B : \mathsf{Set}$$
$$(a, b), (a', b') \longmapsto (a \times a') \otimes (b \times b').$$

But this map is a 4-homomorphism, so it lifts to a bihomomorphism

$$[A \otimes B, A \otimes B \longrightarrow A \otimes B]$$

(neutrality) Neutrality follows from

$$(1_A \otimes 1_B) \times (a \otimes b) = a \otimes b = (a \otimes b) \times (1_A \otimes 1_B).$$

(associativity) Associativity follows from

$$((a \otimes b) \times (a' \otimes b')) \times (a'' \otimes b'') = (a \times a' \times a'') \otimes (b \times b' \times b'') = (a \otimes b) \times ((a' \otimes b') \times (a'' \otimes b'')). \;■$$

We have the map $i_A : A \longrightarrow A \otimes B$ and $i_B : B \longrightarrow A \otimes B$,
$$a \longmapsto a \otimes 1 \qquad b \longmapsto 1 \otimes b$$

they are monoid homomorphisms, for example

$$i_A(1) = 1 \otimes 1 \quad \text{and} \quad i_A(a \times a') = (a \times a') \otimes 1 = (a \times a') \otimes (1 \times 1) = (a \otimes 1) \times (a' \otimes 1) = i_A(a) \times i_A(a').$$

**define** Let $f : A \longrightarrow C : \mathsf{Mon}[\mathcal{A}\circ\mathsf{Set}]$, $g : B \longrightarrow C : \mathsf{Mon}[\mathcal{A}\circ\mathsf{Set}]$

Define $f \times g : A \otimes B \longrightarrow C : \mathcal{A}\circ\mathsf{Set}$

$$(f \times g)(a \otimes b) := f(a) \times f(b). \;■$$

Is $f \times g$ a monoid homomorphism? Note

$$(f \times g)(1 \otimes 1) = f(1) \times f(1) = 1 \times 1 = 1,$$

but $(f \times g)(a \times a') \otimes (b \times b') = f(a) \times f(a') \times g(a) \times g(b)$,

and $(f \times g)(a \otimes b) \times (f \times g)(a' \otimes b') = f(a) \times g(b) \times f(a') \times g(b')$,

so we see that $C$ be a commutative monoid.

When $A$ and $B$ are commutative, $A \otimes B$ is also commutative,

and it has a very nice universal property in $\mathsf{CMon}[\mathcal{A}\circ\mathsf{Set}]$.

**prop** $A \otimes B : \mathsf{CMon}[\mathcal{A}\circ\mathsf{Set}]$ with $i_A : A \to A \otimes B : \mathsf{CMon}[\mathcal{A}\circ\mathsf{Set}]$
$i_B : B \to A \otimes B : \mathsf{CMon}[\mathcal{A}\circ\mathsf{Set}]$

implements the coproduct universal property in $\mathsf{CMon}[\mathcal{A}\circ\mathsf{Set}]$.

**proof** Let $f : A \longrightarrow C : \mathsf{CMon}[\mathcal{A}\circ\mathsf{Set}]$, $g : B \longrightarrow C : \mathsf{CMon}[\mathcal{A}\circ\mathsf{Set}]$.

We need a map $h : A \otimes B \longrightarrow C : \mathsf{CMon}[\mathcal{A}\circ\mathsf{Set}]$ such that

$$h(a \otimes 1) = f(a), \quad h(1 \otimes b) = g(b),$$

but then
$$\begin{aligned} h(a \otimes b) &= h((a \times 1) \otimes (1 \times b)) \\ &= h((a \otimes 1) \times (1 \otimes b)) \\ &= f(a) \times g(b), \end{aligned}$$

so we are forced to take $h := f \times g$, which we already

know to be in $\mathsf{CMon}[\mathcal{A}\circ\mathsf{Set}]$. ■