A category theory primer

Ralf Hinze

Computing Laboratory, University of Oxford
Wolfson Building, Parks Road, Oxford, OX1 3QD, England
ralf.hinze@comlab.ox.ac.uk
http://www.comlab.ox.ac.uk/ralf.hinze/

1 Categories, functors and natural transformations

Category theory is a game with objects and arrows between objects. We let $\mathbb{C}$, $\mathbb{D}$ etc range over categories.

A category is often identified with its class of objects. For instance, we say that **Set** is the category of sets. In the same spirit, we write $A \in \mathbb{C}$ to express that A is an object of $\mathbb{C}$. We let A, B etc range over objects.

However, equally, if not more important are the arrows of a category. So, **Set** is really the category of sets and total functions. (There is also **Rel**, the category of sets and relations.) If the objects have additional structure (monoids, groups etc.) then the arrows are typically structure-preserving maps. For every pair of objects $A, B \in \mathbb{C}$ there is a class of arrows from A to B, denoted $\mathbb{C}(A, B)$. If $\mathbb{C}$ is obvious from the context, we abbreviate $f \in \mathbb{C}(A, B)$ by $f : A \to B$. We will also loosely speak of $A \to B$ as the type of f. We let f, g etc range over arrows.

For every object $A \in \mathbb{C}$ there is an arrow $id_A \in \mathbb{C}(A, A)$, called the identity. Two arrows can be composed if their types match: if $f \in \mathbb{C}(A, B)$ and $g \in \mathbb{C}(B, C)$, then $g \cdot f \in \mathbb{C}(A, C)$. We require composition to be associative with identity as its neutral element.

Every structure comes equipped with structure-preserving maps, so do categories, where these maps are called *functors*. Since a category consists of two parts, objects and arrows, a functor $F : \mathbb{C} \to \mathbb{D}$ consists of a mapping on objects and a mapping on arrows. It is common practise to denote both mappings by the same symbol. We will also loosely speak of F's arrow part as a 'map'. The action on arrows has to respect the types: if $f \in \mathbb{C}(A, B)$, then $F\, f \in \mathbb{D}(F\, A, F\, B)$. Furthermore, F has to preserve identity, $F\, id_A = id_{F\, A}$, and composition $F\,(g \cdot f) = F\, g \cdot F\, f$. The force of functoriality lies in the action on arrows and in the preservation of composition. There is an identity functor, $\mathsf{Id}_{\mathbb{C}} : \mathbb{C} \to \mathbb{C}$, and functors can be composed: $(\mathsf{G} \circ F)\, A = \mathsf{G}\,(F\, A)$ and $(\mathsf{G} \circ F)\, f = \mathsf{G}\,(F\, f)$. This data turns small categories and functors into a category, called **Cat**.[1] We let F, G etc range over functors.

Let $F, \mathsf{G} : \mathbb{C} \to \mathbb{D}$ be two parallel functors. *A natural transformation* $\alpha : F \dot{\to} \mathsf{G}$ is a collection of arrows, so that for each object $A \in \mathbb{C}$ there is an arrow

[1] To avoid paradoxes, we have to require that the objects of **Cat** are small, where a category is called small if the class of objects and the class of all arrows are sets.

$\alpha\, A \in \mathbb{D}(F\, A, \mathsf{G}\, A)$ such that

$$\mathsf{G}\, h \cdot \alpha\, A' = \alpha\, A'' \cdot F\, h, \tag{1}$$

for all arrows $h \in \mathbb{C}(A', A'')$. Given α and h, there are essentially two ways of turning $F\, A'$ things into $\mathsf{G}\, A''$ things. The coherence condition (1) demands that they are equivalent. We also write $\alpha : \forall A \,.\, F\, A \to \mathsf{G}\, A$ and furthermore $\alpha : \forall A \,.\, F\, A \cong \mathsf{G}\, A$, if α is a natural isomorphism. There is an identity transformation $id_F : F \dot{\to} F$ defined $id_F\, A = id_{F\, A}$. Natural transformations can be composed: if $\alpha : F \dot{\to} \mathsf{G}$ and $\beta : \mathsf{G} \dot{\to} \mathsf{H}$, then $\beta \cdot \alpha : F \dot{\to} \mathsf{H}$ is defined $(\beta \cdot \alpha)\, A = \beta\, A \cdot \alpha\, A$. Thus, functors of type $\mathbb{C} \to \mathbb{D}$ and natural transformations between them form a category, the functor category $\mathbb{D}^{\mathbb{C}}$. (Functor categories are exponentials in **Cat**, hence the notation.) We let α, β etc range over natural transformations.

2 Constructions on categories

New categories from old.

Let $\mathbb{C}$ be a category. The *opposite category* $\mathbb{C}^{\mathsf{op}}$ has the same objects as $\mathbb{C}$, arrows and composition, however, are flipped: $f \in \mathbb{C}^{\mathsf{op}}(A, B)$ if $f \in \mathbb{C}(B, A)$, and $g \cdot f \in \mathbb{C}^{\mathsf{op}}(A, C)$ if $f \cdot g \in \mathbb{C}(C, A)$. A functor of type $\mathbb{C}^{\mathsf{op}} \to \mathbb{D}$ or $\mathbb{C} \to \mathbb{D}^{\mathsf{op}}$ is sometimes called a *contravariant* functor from $\mathbb{C}$ to $\mathbb{D}$, the usual kind being styled *covariant*. An incestuous example of a contravariant functor is $\mathbb{C}(-, B) : \mathbb{C}^{\mathsf{op}} \to \mathbf{Set}$, whose action on arrows is given by $\mathbb{C}(h, B)\, f = f \cdot h$.[2] The functor $\mathbb{C}(-, B)$ maps an object A to the *set* of arrows from A to a fixed B, and it takes an arrow $h \in \mathbb{C}(A'', A')$ to a function $\mathbb{C}(h, B) : \mathbb{C}(A', B) \to \mathbb{C}(A'', B)$. Conversely, $\mathbb{C}(A, -) : \mathbb{C} \to \mathbf{Set}$ is a covariant functor defined $\mathbb{C}(A, k)\, f = k \cdot f$.

Let $\mathbb{C}_1$ and $\mathbb{C}_2$ be a categories. An object of the *product category* $\mathbb{C}_1 \times \mathbb{C}_2$ is a pair $\langle A_1, A_2 \rangle$ of objects $A_1 \in \mathbb{C}_1$ and $A_2 \in \mathbb{C}_2$; an arrow of $(\mathbb{C}_1 \times \mathbb{C}_2)(\langle A_1, A_2 \rangle, \langle B_1, B_2 \rangle)$ is a pair $\langle f_1, f_2 \rangle$ of arrows $f_1 \in \mathbb{C}_1(A_1, B_1)$ and $f_2 \in \mathbb{C}_2(A_2, B_2)$. Identity and composition are defined component-wise:

$$id = \langle id, id \rangle \qquad \text{and} \qquad \langle g_1, g_2 \rangle \cdot \langle f_1, f_2 \rangle = \langle g_1 \cdot f_1, g_2 \cdot f_2 \rangle.$$

The projection functors $\mathit{Outl} : \mathbb{C}_1 \times \mathbb{C}_2 \to \mathbb{C}_1$ and $\mathit{Outr} : \mathbb{C}_1 \times \mathbb{C}_2 \to \mathbb{C}_2$ are given by $\mathit{Outl}\, \langle A_1, A_2 \rangle = A_1$, $\mathit{Outl}\, \langle f_1, f_2 \rangle = f_1$ and $\mathit{Outr}\, \langle A_1, A_2 \rangle = A_2$, $\mathit{Outr}\, \langle f_1, f_2 \rangle = f_2$. Product categories avoid the need for functors of several arguments. Functors from a product category are sometimes called *bifunctors*. An example is the *hom-functor* $\mathbb{C}(-, =) : \mathbb{C}^{\mathsf{op}} \times \mathbb{C} \to \mathbf{Set}$, which maps a pair of objects to the set of arrows between them. Its action on arrows is given by $\mathbb{C}(f, g)\, h = g \cdot h \cdot f$. The diagonal functor $\Delta : \mathbb{C} \to \mathbb{C} \times \mathbb{C}$ is an example of a functor into a product category: it duplicates its argument $\Delta A = \langle A, A \rangle$ and $\Delta f = \langle f, f \rangle$.

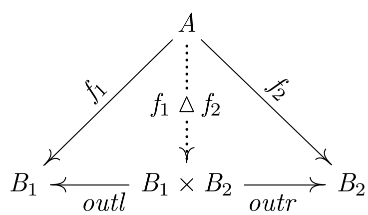
[2] Partial applications of mappings and operators are written using 'categorical dummies', where $-$ marks the first and $=$ the optional second argument.

3 Products and coproducts

Constructions in category theory are typically given using so-called universal properties. The paradigmatic example of this approach is the definition of products — in fact, this is also historically the first example. The *product* of two objects B_1 and B_2 consists of an object written $B_1 \times B_2$ and a pair of arrows $outl : B_1 \times B_2 \to B_1$ and $outr : B_1 \times B_2 \to B_2$. These three things have to satisfy the following *universal property*: for each object A and for each pair of arrows $f_1 : A \to B_1$ and $f_2 : A \to B_2$, there exists an arrow $f_1 \vartriangle f_2 : A \to B_1 \times B_2$ (pronounced "split") such that

$$f_1 = outl \cdot g \quad \wedge \quad f_2 = outr \cdot g \quad \Longleftrightarrow \quad f_1 \vartriangle f_2 = g, \tag{2}$$

for all $g : A \to B_1 \times B_2$. The property states the existence of the arrow $f_1 \vartriangle f_2$ and furthermore that it is the *unique* arrow satisfying the property on the left. (It is also called the *mediating arrow*). The following diagram summarises the type information.



This is an example of a commuting diagram: all paths from the same source to the same target lead to the same result by composition. The dotted arrow indicates that $f_1 \vartriangle f_2$ is the unique arrow from A to $B_1 \times B_2$ that makes the diagram commute.

A universal property such as (2) has two immediate consequences that are worth singling out. If we substitute the right-hand side into the left-hand side, we obtain the *computation laws* (also known as β-rules):

$$f_1 = outl \cdot (f_1 \vartriangle f_2); \tag{3}$$

$$f_2 = outr \cdot (f_1 \vartriangle f_2). \tag{4}$$

They can be seen as defining equations for the arrow $f \vartriangle g$.

Instantiating g in (2) to the identity $id_{A \times B}$ and substituting into the right-hand side, we obtain the *reflection law* (also known as the simple η-rule):

$$outl \vartriangle outr = id_{A \times B}. \tag{5}$$

The universal property enjoys two further consequences, which we shall later identify as naturality properties. The first consequence is the *fusion law* that allows us to fuse a split with an arrow to form another split:

$$(f_1 \vartriangle f_2) \cdot h = f_1 \cdot h \vartriangle f_2 \cdot h, \tag{6}$$

for all $h : A' \to A''$. The law states that $\vartriangle$ is natural in A. For the proof we reason

$$
\begin{aligned}
& f_1 \cdot h \vartriangle f_2 \cdot h = (f_1 \vartriangle f_2) \cdot h \\
\iff & \quad \{ \text{ universal property (2) } \} \\
& f_1 \cdot h = \mathit{outl} \cdot (f_1 \vartriangle f_2) \cdot h \quad \wedge \quad f_2 \cdot h = \mathit{outr} \cdot (f_1 \vartriangle f_2) \cdot h \\
\iff & \quad \{ \text{ computation (3) and (4) } \} \\
& f_1 \cdot h = f_1 \cdot h \quad \wedge \quad f_2 \cdot h = f_2 \cdot h
\end{aligned}
$$

The definition of products is also parametric in B_1 and B_2 — note that both objects are totally passive in the description above. We capture this property by turning $\times$ into a functor of type $\mathbb{C} \times \mathbb{C} \to \mathbb{C}$ (to define products in $\mathbb{C}$ we need products in **Cat**, to define products in **Cat** we need products in ...). The action of $\times$ on arrows is defined

$$
f_1 \times f_2 = f_1 \cdot \mathit{outl} \vartriangle f_2 \cdot \mathit{outr}. \tag{7}
$$

We postpone the proof that $\times$ preserves identity and composition.

The *functor fusion law* states that we can fuse a map after a split to form another split:

$$
(k_1 \times k_2) \cdot (f_1 \vartriangle f_2) = k_1 \cdot f_1 \vartriangle k_2 \cdot f_2, \tag{8}
$$

for all $k_1 : B_1' \to B_1''$ and $k_2 : B_2' \to B_2''$. It formalises that $\vartriangle$ is natural in B_1 and B_2. The proof of (8) builds on fusion and computation.

$$
\begin{aligned}
& (k_1 \times k_2) \cdot (f_1 \vartriangle f_2) \\
= & \quad \{ \text{ definition of } \times \text{ (7) } \} \\
& (k_1 \cdot \mathit{outl} \vartriangle k_2 \cdot \mathit{outr}) \cdot (f_1 \vartriangle f_2) \\
= & \quad \{ \text{ fusion (6) } \} \\
& k_1 \cdot \mathit{outl} \cdot (f_1 \vartriangle f_2) \vartriangle k_2 \cdot \mathit{outr} \cdot (f_1 \vartriangle f_2) \\
= & \quad \{ \text{ computation (3) and (4) } \} \\
& k_1 \cdot f_1 \vartriangle k_2 \cdot f_2
\end{aligned}
$$

Given these prerequisites, it is straightforward to show that $\times$ preserves identity

$$
\begin{aligned}
& \mathit{id}_A \times \mathit{id}_B \\
= & \quad \{ \text{ definition of } \times \text{ (7) } \} \\
& \mathit{id}_A \cdot \mathit{outl} \vartriangle \mathit{id}_B \cdot \mathit{outr} \\
= & \quad \{ \text{ identity and reflection (5) } \} \\
& \mathit{id}_{A \times B}
\end{aligned}
$$

and composition

$$
(g_1 \times g_2) \cdot (f_1 \times f_2)
$$

$$
\begin{aligned}
&= \quad \{ \text{definition of } \times \text{ (7)} \} \\
&\quad (g_1 \times g_2) \cdot (f_1 \cdot \mathit{outl} \vartriangle f_2 \cdot \mathit{outr}) \\
&= \quad \{ \text{functor fusion (8)} \} \\
&\quad g_1 \cdot f_1 \cdot \mathit{outl} \vartriangle g_2 \cdot f_2 \cdot \mathit{outr} \\
&= \quad \{ \text{definition of } \times \text{ (7)} \} \\
&\quad g_1 \cdot f_1 \times g_2 \cdot f_2 .
\end{aligned}
$$

The projection arrows, *outl* and *outr* are natural transformations, as well.

$$k_1 \cdot \mathit{outl} = \mathit{outl} \cdot (k_1 \times k_2); \tag{9}$$
$$k_2 \cdot \mathit{outr} = \mathit{outr} \cdot (k_1 \times k_2). \tag{10}$$

This is a direct consequence of the computation laws.

$$
\begin{aligned}
&\quad \mathit{outl} \cdot (k_1 \times k_2) \\
&= \quad \{ \text{definition of } \times \text{ (7)} \} \\
&\quad \mathit{outl} \cdot (k_1 \cdot \mathit{outl} \vartriangle k_2 \cdot \mathit{outr}) \\
&= \quad \{ \text{computation (3)} \} \\
&\quad k_1 \cdot \mathit{outl}
\end{aligned}
$$

The naturality of $\vartriangle$ can be captured precisely using product categories and hom-functors.

$$(\vartriangle) : \forall A\ B\ .\ (\mathbb{C} \times \mathbb{C})(\varDelta A, B) \to \mathbb{C}(A, \times B)$$

Split takes a pair of arrows as an argument and delivers an arrow to a product (B is an object in a product category). The fusion law captures naturality in A,

$$\mathbb{C}(h, \times B) \cdot (\vartriangle) = (\vartriangle) \cdot (\mathbb{C} \times \mathbb{C})(\varDelta h, B),$$

and the functor fusion law naturality in B,

$$\mathbb{C}(A, \times k) \cdot (\vartriangle) = (\vartriangle) \cdot (\mathbb{C} \times \mathbb{C})(\varDelta A, k).$$

(A transformation between bifunctors is natural if and only if it is natural in both arguments separately.)

The naturality of *outl* and *outr* can be captured using the projection functors *Outl* and *Outr*.

$$
\begin{aligned}
\mathit{outl} &: \forall B\ .\ \mathbb{C}(\times B, \mathit{Outl}\ B); \\
\mathit{outr} &: \forall B\ .\ \mathbb{C}(\times B, \mathit{Outr}\ B).
\end{aligned}
$$

The naturality conditions are

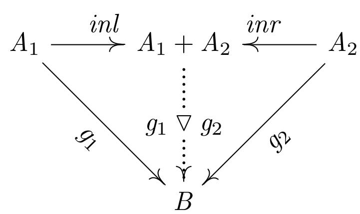
$$
\begin{aligned}
\mathit{Outl}\ k \cdot \mathit{outl} &= \mathit{outl} \cdot \times k, \\
\mathit{Outr}\ k \cdot \mathit{outr} &= \mathit{outr} \cdot \times k.
\end{aligned}
$$

The import of all this is that $\times$ is right adjoint to the diagonal functor $\varDelta$, see Sec. 6.

The construction of products nicely dualises to coproducts, which are products in the opposite category. The *coproduct* of two objects A_1 and A_2 consists of an object written $A_1 + A_2$ and a pair of arrows $inl : A_1 \to A_1 + A_2$ and $inr : A_2 \to A_1 + A_2$. These three things have to satisfy the following *universal property*: for each object B and for each pair of arrows $g_1 : A_1 \to B$ and $g_2 : A_2 \to B$, there exists an arrow $g_1 \triangledown g_2 : A_1 + A_2 \to B$ (pronounced "join") such that

$$f = g_1 \triangledown g_2 \iff f \cdot inl = g_1 \quad \wedge \quad f \cdot inr = g_2, \tag{11}$$

for all $f : A_1 + A_2 \to B$.



Like for products, the universal property implies computation, reflection, fusion and functor fusion laws. *Computation law*:

$$(g_1 \triangledown g_2) \cdot inl = g_1; \tag{12}$$
$$(g_1 \triangledown g_2) \cdot inr = g_2. \tag{13}$$

Reflection law:

$$id_{A+B} = inl \triangledown inr. \tag{14}$$

Fusion law:

$$k \cdot (g_1 \triangledown g_2) = k \cdot g_1 \triangledown k \cdot g_2. \tag{15}$$

The arrow part of the coproduct functor is defined

$$g_1 + g_2 = inl \cdot g_1 \triangledown inr \cdot g_2. \tag{16}$$

Functor fusion law:

$$(g_1 \triangledown g_2) \cdot (h_1 + h_2) = g_1 \cdot h_1 \triangledown g_2 \cdot h_2. \tag{17}$$

The two fusion laws identify $\triangledown$ as a natural transformation:

$$(\triangledown) : \forall A\ B\ .\ \mathbb{C}(+A, B) \to (\mathbb{C} \times \mathbb{C})(A, \Delta B).$$

Finally, the injection arrows are natural transformations, as well.

$$(h_1 + h_2) \cdot inl = inl \cdot h_1 \tag{18}$$
$$(h_1 + h_2) \cdot inr = inr \cdot h_2 \tag{19}$$

The import of all this is that $+$ is left adjoint to the diagonal functor Δ.

4 Initial and final objects

An object is called initial if for each object $B \in \mathbb{C}$ there is exactly one arrow from the initial object to B. Any two initial objects are isomorphic, which is why we usually speak of *the* initial object. It is denoted 0, and the unique arrow from 0 to B is written $¡_B$ (pronounce "gnab").

$$0 \overset{¡_B}{\dashrightarrow} B$$

The uniqueness can also be expressed as a *universal property*:

$$f = ¡_B \quad \iff \quad \mathit{true},$$

for all $f : 0 \to B$. Instantiating f to the identity id_0, we obtain the *reflection law*: $id_0 = ¡_0$. An arrow after a gnab can be fused into a single gnab.

$$k \cdot ¡_{B'} = ¡_{B''},$$

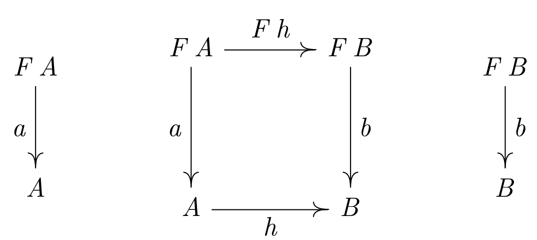
for all $k : B' \to B''$. The *fusion law* expresses that $¡_B$ is natural in B.

Dually, 1 is a final object if for each object $A \in \mathbb{C}$ there is a unique arrow from A to 1, denoted $!_A$ (pronounce "bang").

$$A \overset{!_A}{\dashrightarrow} 1$$

5 Initial algebras and final coalgebras

Let $F : \mathbb{C} \to \mathbb{C}$ be an endofunctor. An F-*algebra* is a pair $\langle A, a\rangle$ consisting of an object $A \in \mathbb{C}$ and an arrow $a \in \mathbb{C}(F\,A, A)$. An F-*homomorphism* between algebras $\langle A, a\rangle$ and $\langle B, b\rangle$ is an arrow $h \in \mathbb{C}(A, B)$ such that $h \cdot a = b \cdot F\,h$.



Identity is an F-homomorphism and F-homomorphisms compose. Consequently, the data defines a category, called $\mathbf{Alg}(F)$. The initial object in this category — if it exists — is the so-called *initial* F-*algebra* $\langle \mu F, in\rangle$. The import of initiality is that there is a unique arrow from $\langle \mu F, in\rangle$ to any other F-algebra $\langle B, b\rangle$. This unique arrow is written $(\!|b|\!)$ and is called *fold* or *catamorphism*. Expressed in terms of the base category, it satisfies the following *universal property*.

$$f = (\!|b|\!) \quad \iff \quad f \cdot in = b \cdot F\,f \tag{20}$$

Like for products, the universal property has two immediate consequences. Substituting the left-hand side into the right-hand side gives the *computation law*:

$$(\!|b|\!) \cdot in = b \cdot F\,(\!|b|\!). \tag{21}$$

Setting $f := id$ and $b := in$, we obtain the *reflection law*:

$$id = (\!|in|\!). \tag{22}$$

Since the initial algebra is an initial object, we also have a *fusion law* for fusing an arrow with a fold to form another fold.

$$k \cdot (\!|b'|\!) = (\!|b''|\!) \quad \Longleftarrow \quad k \cdot b' = b'' \cdot F\,k \tag{23}$$

The proof is trivial if phrased in terms of the category $\mathbf{Alg}(F)$. However, we can also execute the proof in the underlying category $\mathbb{C}$.

$$\begin{aligned}
& k \cdot (\!|b'|\!) = (\!|b''|\!) \\
\Longleftrightarrow\;& \{\ \text{universal property (20)}\ \} \\
& k \cdot (\!|b'|\!) \cdot in = b'' \cdot F\,(k \cdot (\!|b'|\!)) \\
\Longleftrightarrow\;& \{\ \text{computation (21)}\ \} \\
& k \cdot b' \cdot F\,(\!|b'|\!) = b'' \cdot F\,(k \cdot (\!|b'|\!)) \\
\Longleftrightarrow\;& \{\ F \text{ functor}\ \} \\
& k \cdot b' \cdot F\,(\!|b'|\!) = b'' \cdot F\,k \cdot F\,(\!|b'|\!) \\
\Longleftarrow\;& \{\ \text{cancel } - \cdot F\,(\!|b'|\!) \text{ on both sides}\ \} \\
& k \cdot b' = b'' \cdot F\,k.
\end{aligned}$$

The fusion law states that $(\!|-|\!)$ is natural in $\langle B, b\rangle$, that is, as an arrow in $\mathbf{Alg}(F)$. This does *not* imply naturality in the underlying category $\mathbb{C}$ (as an arrow in $\mathbb{C}$ it is a strong dinatural transformation).

Using these laws we can show that μF is indeed a *fixed point* of the functor: $F\,(\mu F) \cong \mu F$. The isomorphism is witnessed by the arrows $in \in \mathbb{C}(F\,(\mu F), \mu F)$ and $(\!|F\,in|\!) \in \mathbb{C}(\mu F, F\,(\mu F))$. We calculate

$$\begin{aligned}
& in \cdot (\!|F\,in|\!) = id \\
\Longleftrightarrow\;& \{\ \text{reflection}\ \} \\
& in \cdot (\!|F\,in|\!) = (\!|in|\!) \\
\Longleftarrow\;& \{\ \text{fusion (23)}\ \} \\
& in \cdot F\,in = in \cdot F\,in
\end{aligned}$$

For the reverse direction, we reason

$$\begin{aligned}
& (\!|F\,in|\!) \cdot in \\
=\;& \{\ \text{computation}\ \} \\
& F\,in \cdot F\,(\!|F\,in|\!)
\end{aligned}$$

$$
\begin{aligned}
&= \quad \{\ F \text{ functor}\ \} \\
&\quad F\,(in \cdot (\!|F\ in|\!)) \\
&= \quad \{\ \text{see proof above}\ \} \\
&\quad F\ id \\
&= \quad \{\ F \text{ functor}\ \} \\
&\quad id.
\end{aligned}
$$

Perhaps surprisingly, folds also enjoy a functor fusion law. To be able to formulate the law, we have to turn μ into a higher-order functor of type $\mathbb{C}^{\mathbb{C}} \to \mathbb{C}$. The object part of this functor maps a functor to its initial algebra. The arrow part, which maps a natural transformation $\alpha : F \dot{\to} \mathsf{G}$ to an arrow $\mu\alpha \in \mathbb{C}(\mu F, \mu \mathsf{G})$, is given by

$$\mu\alpha = (\!|in \cdot \alpha|\!). \tag{24}$$

(To reduce clutter we have omitted the type argument of α on the right-hand side, which should read $(\!|in \cdot \alpha\,(\mu\mathsf{G})|\!)$). Like for products, we postpone the proof that μ preserves identity and composition.

The *functor fusion law* states that we can fuse a fold after a map to form another fold:

$$(\!|b \cdot \alpha|\!) = (\!|b|\!) \cdot \mu\alpha, \tag{25}$$

for all $\alpha : F' \dot{\to} F''$. To establish functor fusion we reason

$$
\begin{aligned}
&\quad (\!|b|\!) \cdot \mu\alpha = (\!|b \cdot \alpha|\!) \\
\Longleftrightarrow &\quad \{\ \text{definition of } \mu \text{ (24)}\ \} \\
&\quad (\!|b|\!) \cdot (\!|in \cdot \alpha|\!) = (\!|b \cdot \alpha|\!) \\
\Longleftarrow &\quad \{\ \text{fusion (23)}\ \} \\
&\quad (\!|b|\!) \cdot in \cdot \alpha = b \cdot \alpha \cdot F'\,(\!|b|\!) \\
\Longleftrightarrow &\quad \{\ \text{computation (21)}\ \} \\
&\quad b \cdot F''\,(\!|b|\!) \cdot \alpha = b \cdot \alpha \cdot F'\,(\!|b|\!) \\
\Longleftrightarrow &\quad \{\ \text{naturality of } \alpha\ \} \\
&\quad b \cdot \alpha \cdot F'\,(\!|b|\!) = b \cdot \alpha \cdot F'\,(\!|b|\!).
\end{aligned}
$$

Given these prerequisites, it is straightforward to show that μ preserves identity

$$
\begin{aligned}
&\quad \mu id \\
&= \quad \{\ \text{definition of } \mu \text{ (24)}\ \} \\
&\quad (\!|in \cdot id|\!) \\
&= \quad \{\ \text{identity and reflection (22)}\ \} \\
&\quad id
\end{aligned}
$$

and composition

$$
\begin{aligned}
& \mu\beta \cdot \mu\alpha \\
= \quad & \{ \text{ definition of } \mu \text{ (24) } \} \\
& (\!|in \cdot \beta|\!) \cdot \mu\alpha \\
= \quad & \{ \text{ functor fusion (25) } \} \\
& (\!|in \cdot \beta \cdot \alpha|\!) \\
= \quad & \{ \text{ definition of } \mu \text{ (24) } \} \\
& \mu(\beta \cdot \alpha).
\end{aligned}
$$

To summarise, functor fusion expresses that $(\!|-|\!)$ is natural in F:

$$(\!|-|\!) : \forall F \,.\, \mathbb{C}(F\,B, B) \to \mathbb{C}(\mu F, B).$$

The arrow $in : F\,(\mu F) \to \mu F$ is natural in F, as well. The arrow part of the higher-order functor $\Lambda F \,.\, F\,(\mu F)$ is $\lambda\,\alpha \,.\, F''\,(\mu\alpha) \cdot \alpha = \lambda\,\alpha \,.\, \alpha \cdot F'\,(\mu\alpha)$.

$$\mu\alpha \cdot in = in \cdot \alpha \cdot F\,(\mu\alpha) \tag{26}$$

We reason

$$
\begin{aligned}
& \mu\alpha \cdot in \\
= \quad & \{ \text{ definition of } \mu \text{ (24) } \} \\
& (\!|in \cdot \alpha|\!) \cdot in \\
= \quad & \{ \text{ computation (21) } \} \\
& in \cdot \alpha \cdot F\,(\!|in \cdot \alpha|\!) \\
= \quad & \{ \text{ definition of } \mu \text{ (24) } \} \\
& in \cdot \alpha \cdot F\,(\mu\alpha).
\end{aligned}
$$

The development nicely dualises to *F-coalgebras* and *unfolds*. An *F-coalgebra* is a pair $\langle A, a\rangle$ consisting of an object $A \in \mathbb{C}$ and an arrow $a \in \mathbb{C}(A, F\,A)$. An *$F$-homomorphism* between coalgebras $\langle A, a\rangle$ and $\langle B, b\rangle$ is an arrow $h \in \mathbb{C}(A, B)$ such that $F\,h \cdot a = b \cdot h$. Identity is an F-homomorphism and F-homomorphisms compose. Consequently, the data defines a category, called **Coalg**(F). The final object in this category — if it exists — is the so-called *final F-coalgebra* $\langle \nu F, out\rangle$. The import of finality is that there is a unique arrow to $\langle \nu F, out\rangle$ from any other F-coalgebra $\langle A, a\rangle$. This unique arrow is written $[\!(a)\!]$ and is called *unfold* or *anamorphism*. Expressed in terms of the base category, it satisfies the following *universal property*.

$$F\,g \cdot a = out \cdot g \quad\iff\quad [\!(a)\!] = g \tag{27}$$

Like for initial algebras, the universal property implies computation, reflection, fusion and functor fusion laws. *Computation law*:

$$F\,[\!(a)\!] \cdot a = out \cdot [\!(a)\!]. \tag{28}$$

Reflection law:

$$\llparenthesis out \rrparenthesis = id. \tag{29}$$

Fusion law:

$$\llparenthesis a' \rrparenthesis = \llparenthesis a'' \rrparenthesis \cdot h \quad \Longleftarrow \quad F\,h \cdot a = a'' \cdot h. \tag{30}$$

The object part of the functor ν is defined

$$\nu\mathbb{a} = \llparenthesis \mathbb{a} \cdot out \rrparenthesis. \tag{31}$$

Functor fusion law:

$$\nu\mathbb{a} \cdot \llparenthesis a \rrparenthesis = \llparenthesis \mathbb{a} \cdot a \rrparenthesis \tag{32}$$

Finally, *out* is a natural transformation.

$$\mathbb{a} \cdot F\,(\nu\mathbb{a}) \cdot out = out \cdot \nu\mathbb{a}$$

6 Adjunctions

We have noted in Sec. 3 that products and coproducts are part of an adjunction. In this section, we explore the notion of an adjunction in greater depth.

Let $\mathbb{C}$ and $\mathbb{D}$ be categories. The functors L and R are *adjoint*, denoted $\mathsf{L} \dashv \mathsf{R}$,

$$\mathbb{C} \underset{\mathsf{R}}{\overset{\mathsf{L}}{\underset{\longrightarrow}{\overset{\longleftarrow}{\bot}}}} \mathbb{D}$$

if and only if there is a bijection

$$\phi : \forall A\,B\;.\;\mathbb{C}(\mathsf{L}\,A, B) \cong \mathbb{D}(A, \mathsf{R}\,B),$$

that is natural both in A and B. The isomorphism ϕ is called the *adjoint transposition*. It is also called the *left adjunct* with ϕ° being the *right adjunct*. That ϕ and ϕ° are mutually inverse, can be captured using an equivalence.

$$f = \phi^{\circ}\,g \quad \Longleftrightarrow \quad \phi\,f = g \tag{33}$$

(The left-hand side lives in $\mathbb{C}$, and the right-hand side in $\mathbb{D}$.) The formula is reminiscent of the universal property of products. That the latter indeed defines an adjunction can be seen more clearly if we re-formulate (2) in terms of the categories involved.

$$f = \langle outl, outr \rangle \cdot \mathit{\Delta}\,g \quad \Longleftrightarrow \quad \triangle f = g$$

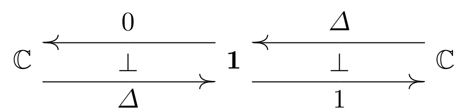
The right part of the diagram below explicates the categories involved.

$$\mathbb{C} \underset{\mathit{\Delta}}{\overset{+}{\underset{\longrightarrow}{\overset{\longleftarrow}{\bot}}}} \mathbb{C} \times \mathbb{C} \underset{\times}{\overset{\mathit{\Delta}}{\underset{\longrightarrow}{\overset{\longleftarrow}{\bot}}}} \mathbb{C}$$

We actually have a double adjunction with $+$ being left adjoint to Δ. Rewritten in terms of product categories, the universal property of coproducts (11) becomes

$$f = \nabla g \iff \Delta f \cdot \langle inl, inr \rangle = g.$$

Initial objects and final objects also define (a rather trivial adjunction) between the category $\mathbf{1}$ and $\mathbb{C}$.



The category $\mathbf{1}$ consists of a single object $*$ and a single arrow id_*. The diagonal functor is now defined $\Delta A = *$ and $\Delta f = id_*$. The objects 0 and 1 are seen as constant functors from $\mathbf{1}$. (An object $A \in \mathbb{C}$ seen as a functor $A : \mathbf{1} \to \mathbb{C}$ maps $*$ to A and id_* to id_A.)

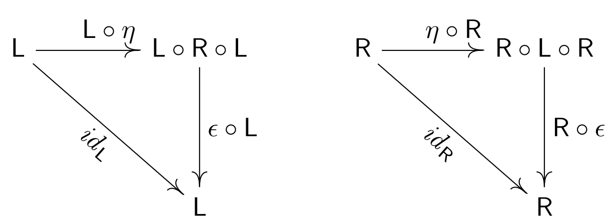
$$f = ¡_B \cdot 0\, g \iff \Delta f \cdot id_* = g \tag{34}$$
$$f = id_* \cdot \Delta g \iff 1 f \cdot !_A = g \tag{35}$$

The universal properties are a bit degenerated as the right-hand side of (34) and the left-hand side of (35) is vacuously true.

An adjunction can be defined in a variety of ways. An alternative approach makes use of two natural transformations: the *counit* $\epsilon : \mathsf{L} \circ \mathsf{R} \mathrel{\dot{\to}} Id$ and the *unit* $\eta : Id \mathrel{\dot{\to}} \mathsf{R} \circ \mathsf{L}$. For products, the counit is the pair of arrows $\langle outl, outr \rangle$ and the unit is the diagonal arrow $\delta = id \vartriangle id$. The units must satisfy

$$(\epsilon \circ \mathsf{L}) \cdot (\mathsf{L} \circ \eta) = id_{\mathsf{L}} \qquad \text{and} \qquad (\mathsf{R} \circ \epsilon) \cdot (\eta \circ \mathsf{R}) = id_{\mathsf{R}},$$

where $\circ$ denotes (horizontal) composition of a natural transformation with a functor: $(F \circ \alpha)\, A = F\,(\alpha\, A)$ and $(\alpha \circ F)\, A = \alpha\,(F\, A)$. It is useful to explicate the typing information.



You may want to think of L and R as closure operations. The unit laws express that going left-right-left is the same as going left once and likewise for going right.

All in all, an adjunction consists of six entities: two functors, two adjuncts, and two units. Every single of those can be defined in terms of the others:

$$\begin{array}{lll} \phi^\circ\, g = \epsilon \cdot \mathsf{L}\, g & \epsilon = \phi^\circ\, id & \mathsf{L}\, g = \phi^\circ\,(\eta \cdot g) \\ \phi\, f = \mathsf{R}\, f \cdot \eta & \eta = \phi\, id & \mathsf{R}\, f = \phi\,(f \cdot \epsilon). \end{array}$$

In terms of programming language concepts, adjuncts correspond to introduction and elimination rules ($\vartriangle$ introduces a pair, $\triangledown$ eliminates a sum). The units can be seen as simple variants of these rules ($\langle \mathit{outl}, \mathit{outr} \rangle$ eliminates a pair and $\langle \mathit{inl}, \mathit{inr} \rangle$ introduces a sum). When we discussed products, we derived a variety of laws from the universal property. Table 1 re-formulates these laws using the new vocabulary. For instance, from the perspective of the right adjoint $f = \phi^{\circ}\,(\phi\, f)$ corresponds to a computation law or β-rule, viewed from the left it is an η-rule.[3] The table merits careful study. Table 2 lists some examples of adjunctions.

Table 1. Adjunctions and laws (view from the left / right).

ϕ° *introduction / elimination* $\phi^{\circ} : \mathbb{D}(A, \mathsf{R}\, B) \to \mathbb{C}(\mathsf{L}\, A, B)$	ϕ *elimination / introduction* $\phi : \mathbb{C}(\mathsf{L}\, A, B) \to \mathbb{D}(A, \mathsf{R}\, B)$
Universal property $f = \phi^{\circ}\, g \iff \phi\, f = g$	
$\epsilon \in \mathbb{C}(\mathsf{L}\,(\mathsf{R}\, B), B)$ $\epsilon = \phi^{\circ}\, \mathit{id}$	$\eta \in \mathbb{D}(A, \mathsf{R}\,(\mathsf{L}\, A))$ $\phi\, \mathit{id} = \eta$
— / *computation law* η-rule / β-rule $f = \phi^{\circ}\,(\phi\, f)$	*computation law* / — β-rule / η-rule $\phi\,(\phi^{\circ}\, g) = g$
reflection law / — simple η-rule / simple β-rule $\mathit{id} = \phi^{\circ}\, \eta$	— / *reflection law* simple β-rule / simple η-rule $\phi\, \epsilon = \mathit{id}$
functor fusion law / — ϕ° is natural in A $\phi^{\circ}\, g \cdot \mathsf{L}\, h = \phi^{\circ}\,(g \cdot h)$	— / *fusion law* ϕ is natural in A $\phi\, f \cdot h = \phi\,(f \cdot \mathsf{L}\, h)$
fusion law / — ϕ° is natural in B $k \cdot \phi^{\circ}\, g = \phi^{\circ}\,(\mathsf{R}\, k \cdot g)$	— / *functor fusion law* ϕ is natural in B $\mathsf{R}\, k \cdot \phi\, f = \phi\,(k \cdot f)$
ϵ is natural in B $k \cdot \epsilon = \epsilon \cdot \mathsf{L}\,(\mathsf{R}\, k)$	η is natural in A $\mathsf{R}\,(\mathsf{L}\, h) \cdot \eta = \eta \cdot h$

Since the components of an adjunction are inter-definable, an adjunction can be specified by providing only part of the data. Surprisingly little is needed: for products only the functor L and the counit ϵ were given, the other ingredients were derived from those. In the rest of this section, we replay the derivation in terms of adjunctions. Let $\mathsf{L} : \mathbb{D} \to \mathbb{C}$ be a functor, and let $\epsilon \in \mathbb{C}(\mathsf{L}\,(\mathsf{R}\, B), B)$ be a *universal arrow*. Universality means that for each $f \in \mathbb{C}(\mathsf{L}\, A, B)$ there exists an

[3] It is a coincidence that the same Greek letter is used both for extensionality (η-rule) and for the unit of an adjunction.

Table 2. Examples of adjunctions.

adjunction	initial object	final object	coproduct	product	exponential
L	0	Δ	$+$	Δ	$- \times X$
R	Δ	1	Δ	$\times$	$(-)^X$
ϕ°			∇		$\mathit{uncurry}$
ϕ				$\vartriangle$	λ
ϵ	¡			$\langle \mathit{outl}, \mathit{outr} \rangle$	apply
η		!	$\langle \mathit{inl}, \mathit{inr} \rangle$		

arrow $\phi\, f \in \mathbb{D}(A, \mathsf{R}\, B)$ such that

$$f = \epsilon \cdot \mathsf{L}\, g \quad \iff \quad \phi\, f = g, \tag{36}$$

for all $g \in \mathbb{D}(A, \mathsf{R}\, B)$. The formula suggests that $\epsilon \cdot \mathsf{L}\, g = \phi^\circ\, g$. *Computation law:* substituting the right-hand side into the left-hand side, we obtain

$$f = \epsilon \cdot \mathsf{L}\, (\phi\, f). \tag{37}$$

Reflection law: setting $f := \epsilon$ and $g := \mathit{id}$, yields

$$\phi\, \epsilon = \mathit{id}. \tag{38}$$

Fusion law: to establish

$$\phi\, (f \cdot \mathsf{L}\, h) = \phi\, f \cdot h, \tag{39}$$

we appeal to the universal property:

$$f \cdot \mathsf{L}\, h = \epsilon \cdot \mathsf{L}\, (\phi\, f \cdot h) \quad \iff \quad \phi\, (f \cdot \mathsf{L}\, h) = \phi\, f \cdot h.$$

To show the left-hand side, we calculate

$$\begin{aligned} & \epsilon \cdot \mathsf{L}\, (\phi\, f \cdot h) \\ = \quad & \{\ \mathsf{L} \text{ functor}\ \} \\ & \epsilon \cdot \mathsf{L}\, (\phi\, f) \cdot \mathsf{L}\, h \\ = \quad & \{\ \text{computation (37)}\ \} \\ & f \cdot \mathsf{L}\, h. \end{aligned}$$

The type constructor R can be turned into a functor whose action on arrows is defined $\mathsf{R} f = \phi\, (f \cdot \epsilon)$. (The definition is suggested by combining reflection and functor fusion: $\mathsf{R} f = \mathsf{R} f \cdot \phi\, \epsilon = \phi\, (f \cdot \epsilon)$.) *Functor fusion law:*

$$\mathsf{R}\, k \cdot \phi\, f = \phi\, (k \cdot f). \tag{40}$$

For the proof, we reason

$$\begin{aligned}
& \mathsf{R}\,k \cdot \phi\,f \\
= & \quad \{ \text{ definition of } \mathsf{R} \ \} \\
& \phi\,(k \cdot \epsilon) \cdot \phi\,f \\
= & \quad \{ \text{ fusion (39) } \} \\
& \phi\,(k \cdot \epsilon \cdot \mathsf{L}\,(\phi\,f)) \\
= & \quad \{ \text{ computation (37) } \} \\
& \phi\,(k \cdot f).
\end{aligned}$$

Functoriality: R preserves identity

$$\begin{aligned}
& \mathsf{R}\,id \\
= & \quad \{ \text{ definition of } \mathsf{R} \ \} \\
& \phi\,(id \cdot \epsilon) \\
= & \quad \{ \text{ identity and reflection (38) } \} \\
& id
\end{aligned}$$

and composition

$$\begin{aligned}
& \mathsf{R}\,g \cdot \mathsf{R}\,f \\
= & \quad \{ \text{ definition of } \mathsf{R} \ \} \\
& \mathsf{R}\,g \cdot \phi\,(f \cdot \epsilon) \\
= & \quad \{ \text{ functor fusion (40) } \} \\
& \phi\,(g \cdot f \cdot \epsilon) \\
= & \quad \{ \text{ definition of } \mathsf{R} \ \} \\
& \mathsf{R}\,(g \cdot f).
\end{aligned}$$

Fusion and functor fusion show that ϕ is natural both in A and in B. Finally, the counit ϵ is natural in B.

$$\begin{aligned}
& \epsilon \cdot \mathsf{L}\,(\mathsf{R}\,k) \\
= & \quad \{ \text{ definition of } \mathsf{R} \ \} \\
& \epsilon \cdot \mathsf{L}\,(\phi\,(k \cdot \epsilon)) \\
= & \quad \{ \text{ computation (37) } \} \\
& k \cdot \epsilon
\end{aligned}$$

Dually, a functor R and a universal arrow $\eta \in \mathbb{C}(A, \mathsf{R}\,(\mathsf{L}\,A))$ are sufficient.

$$f = \phi^{\circ}\,g \quad \Longleftrightarrow \quad \mathsf{R}\,f \cdot \eta = g.$$

Define $\phi\,f = \mathsf{R}\,f \cdot \eta$ and $\mathsf{L}\,g = \phi^{\circ}\,(\eta \cdot g)$.