

Notes on Initial Algebras

Gabriele Rastello

April 11, 2022

Contents

Contents	1
1 Preliminaries	2
1.1 Limits and colimits	3
1.2 Other notations	6
2 Algebras for endofunctors	7
2.1 F -algebras	8
2.2 Initial Algebras	11
2.3 Recursion and Induction	16
3 Initial algebras from finitary iteration	19
3.1 Adámek's Theorem	20
4 Initial algebras from transfinite iteration	26
4.1 Colimits of chains	27
4.2 Transfinite iteration	28
Bibliography	31

Chapter 1

Preliminaries

1.1 Limits and colimits

Through this section let \mathcal{A}, \mathcal{D} be categories with \mathcal{D} small. We recall what limits and colimits are, some significant examples (particularly products and coproducts) and some of their basic properties. We also introduce some notation. For detailed proofs see [Bor08, Chapter 2].

Definition 1.1.1. Given a functor $F: \mathcal{D} \rightarrow \mathcal{A}$ a **cone** on F is an object $C \in \mathcal{A}$ and a family of arrows $(\pi_D: C \rightarrow FD)_{D \in \mathcal{D}}$ such that for every arrow $d: D_1 \rightarrow D_2$ of \mathcal{D} we have

$$Fd \circ \pi_{D_1} = \pi_{D_2}. \quad (1.1.1)$$

The arrows π_D are called **projections** of the cone, the object C the **vertex**.

Definition 1.1.2. Given a functor $F: \mathcal{D} \rightarrow \mathcal{A}$ a **limit** of F is a cone $(L, (\pi_D)_{D \in \mathcal{D}})$ such that for every other cone $(C, (\rho_D)_{D \in \mathcal{D}})$ there is a unique arrow $m: C \rightarrow L$ such that

$$\rho_D = \pi_D \circ m \quad \text{for every } D \in \mathcal{D}.$$

Proposition 1.1.3. When a functor F has a limit that limit is unique (up to isomorphism).

The following proposition is a way of proving equality of two arrows into a limit. It is most useful when working in abstract categories where arrows are not (generally) functions.

Proposition 1.1.4. Let $(L, (\pi_D)_{D \in \mathcal{D}})$ be the limit of F and consider two arrows $f, g: C \rightarrow L$. If for every $D \in \mathcal{D}$, $\pi_D \circ f = \pi_D \circ g$ then $f = g$.

The definitions of cone and of limit are dualized to yield those of cocone and colimit.

Definition 1.1.5. Given a functor $F: \mathcal{D} \rightarrow \mathcal{A}$ a **cocone** on F is an object $C \in \mathcal{A}$ and a family of arrows $(\sigma_D: FD \rightarrow C)_{D \in \mathcal{D}}$ such that for every arrow $d: D_1 \rightarrow D_2$ of \mathcal{D} we have

$$\sigma_{D_2} \circ Fd = \sigma_{D_1}. \quad (1.1.2)$$

The arrows σ_D are called **coprojections** of the cocone.

Definition 1.1.6. Given a functor $F: \mathcal{D} \rightarrow \mathcal{A}$ a **colimit** of F is a cocone $(L, (\sigma_D)_{D \in \mathcal{D}})$ such that for every other cocone $(C, (\tau_D)_{D \in \mathcal{D}})$ there is a unique arrow $m: L \rightarrow C$ such that

$$\tau_D = m \circ \sigma_D \quad \text{for every } D \in \mathcal{D}.$$

Propositions 1.1.3 and 1.1.4 are dualized as follows.

Proposition 1.1.7. When a functor F has a colimit that colimit is unique (up to isomorphism).

Proposition 1.1.8. Let $(L, (\sigma_D)_{D \in \mathcal{D}})$ be the colimit of F and consider two arrows $f, g: L \rightarrow C$. If for every $D \in \mathcal{D}$, $f \circ \sigma_D = g \circ \sigma_D$ then $f = g$.

Products and coproducts

We now turn to two particular classes of (co)limits: products and coproducts. Recall that a discrete category is a category that has no non-identity arrow.

Definition 1.1.9. Given a functor $F: \mathcal{D} \rightarrow \mathcal{A}$ where \mathcal{D} is some discrete category a limit of F is called a **product** while a colimit a **coproduct**.

product, coproduct

Notation 1.1.10. Notice that to give a functor from a discrete category to \mathcal{A} is equivalent to picking an element of \mathcal{A} for every element of \mathcal{D} . We thus speak of the “product of A and B ” for $A, B \in \mathcal{A}$ without explicit reference to any functor. Moreover we denote the product of A and B in \mathcal{A} , when it exists, by $A \times B$. Similarly we speak of the coproduct of two elements of \mathcal{A} and denote it by $A + B$ when it exists.

Proposition 1.1.11. In a category, when the interested (co)products exists, we have that

1. $A \times B$ is isomorphic to $B \times A$;
2. $A + B$ is isomorphic to $B + A$;
3. $(A \times B) \times C$ is isomorphic to $A \times (B \times C)$;
4. $(A + B) + C$ is isomorphic to $A + (B + C)$.

In light of this proposition we write $A \times B \times C$ and $A + B + C$ with no parenthesis and similarly for any finite number of factors or addenda.

Notation 1.1.12. When we take the (co)product of an infinite number of objects of \mathcal{A} we use the notation $\prod_{i \in I} A_i$ (for products) and $\coprod_{i \in I} A_i$ (for coproducts). Notice that the order of the factors/addenda does not matter.

Notation 1.1.13. Let $(L = \prod_{i \in I} A_i, (\pi_i)_{i \in I})$ be a product in \mathcal{A} . Then by Proposition 1.1.4 there is a one-to-one correspondance between arrows $f: C \rightarrow L$ and cones $(C, (\rho_i: C \rightarrow A_i)_{i \in I})$. Moreover any collection of arrows $(\rho_i: C \rightarrow A_i)_{i \in I}$ satisfies condition (1.1.1) so it is a cone. This means that to give an arrow $f: C \rightarrow L$ is equivalent to giving a family of functions $(\rho_i: C \rightarrow A_i)_{i \in I}$ so we write

$$f = \langle \rho_1, \rho_2, \rho_3, \dots \rangle; \quad (1.1.3)$$

given a convenient ordering on I .

By a dual argument we obtain that arrows $f: L \rightarrow C$ out of a coproduct $L = \coprod_{i \in I} A_i$ are in one-to-one correspondence with families of arrows $(\tau_i: A_i \rightarrow C)$ and write

$$f = [\tau_1, \tau_2, \tau_3, \dots] \quad (1.1.4)$$

for a convenient ordering on I .

Notation 1.1.14. Let A, B, C, D be objects of \mathcal{A} such that $A \times B, C \times D$ exist and let $f: A \rightarrow C, g: B \rightarrow D$ be arrows. With reference to the following diagram we denote $\langle f \circ \pi_1, g \circ \pi_2 \rangle$ by $f \times g$. Dually if $A + B, C + D$ exist we denote $[\sigma_1 \circ f, \sigma_2 \circ g]$ by $f + g$.

$f \times g, f + g$

$$\begin{array}{ccc}
A & \xrightarrow{f} & C \\
\uparrow \pi_1 & & \uparrow \sigma_1 \\
A \times B & \xrightarrow{f \times g} & C \times D \\
\downarrow \pi_2 & & \downarrow \sigma_2 \\
B & \xrightarrow{g} & D
\end{array}
\qquad
\begin{array}{ccc}
A & \xrightarrow{f} & C \\
\downarrow & & \downarrow \\
A + B & \xrightarrow{f+g} & C + D \\
\uparrow & & \uparrow \\
B & \xrightarrow{g} & D
\end{array}$$

The next proposition shows how the notation introduced above for products interacts with the composition. It is most useful for performing calculations; we will use it without explicit reference particularly in the proof of the Primitive Recursion Theorem (see 2.3.3).

Proposition 1.1.15. Consider the following two diagrams.

$$\begin{array}{ccc}
& A & \xrightarrow{g_1} D \\
f_1 \nearrow & \uparrow & \uparrow \\
C & \xrightarrow{\langle f_1, f_2 \rangle} A \times B & \xrightarrow{g_1 \times g_2} D \times E \\
f_2 \searrow & \downarrow & \downarrow \\
& B & \xrightarrow{g_2} E
\end{array}
\qquad
\begin{array}{ccc}
& A & \\
f_1 \nearrow & \uparrow & \\
D \xrightarrow{g} C & \xrightarrow{\langle f_1, f_2 \rangle} A \times B & \\
f_2 \searrow & \downarrow & \\
& B &
\end{array}$$

We have that

1. $(g_1 \times g_2) \circ \langle f_1, f_2 \rangle = (g_1 \circ f_1, g_2 \circ f_2)$;
2. $\langle f_1, f_2 \rangle \circ g = \langle f_1 \circ g, f_2 \circ g \rangle$.

Initial and terminal objects

Definition 1.1.16. An **initial object** for a category \mathcal{A} is an object $0 \in \mathcal{A}$ such that for every other object $A \in \mathcal{A}$ there is exactly one arrow from 0 to A . Dually a **terminal object** for \mathcal{A} is an object $1 \in \mathcal{A}$ such that for every other object $A \in \mathcal{A}$ there is exactly one arrow from A to 1 .

initial object

terminal object

Remark 1.1.17. Initial objects, when they exist, are colimits of the functor from the empty category into \mathcal{A} ; as such they are all isomorphic and we speak of *the* initial object. Dually terminal objects are limits of the functor from the empty category and are unique as well.

(Co)limit-preserving functors

Definition 1.1.18. A functor $G: \mathcal{A} \rightarrow \mathcal{B}$ **preserve limits** if, for every small category \mathcal{D} and functor $F: \mathcal{D} \rightarrow \mathcal{A}$, when the limit of F exists the limit of $G \circ F$ is obtained by applying G to the limit cone. That is: if $(L, (\pi_D)_{D \in \mathcal{D}})$ is the limit of F then $(GL, (G\pi_D)_{D \in \mathcal{D}})$ is the limit of $G \circ F$. As expected G is said to preserve colimits when, for every small category \mathcal{D} and functor $F: \mathcal{D} \rightarrow \mathcal{A}$, when the colimit of F exists then the colimit of $G \circ F$ is the image through G of the colimit cocone.

limit preservation

colimit preservation

By restricting the definition above to a specific small category \mathcal{D} we obtain the notion of a functor that preserves \mathcal{D} -(co)limits. We will be interested in functors that preserve ω -(co)limits where ω is the set of natural numbers with the usual ordering regarded as a category.

1.2 Other notations

Notation 1.2.1. When we are working with a category \mathcal{A} that has enough structure¹ we will often define endofunctors $F: \mathcal{A} \rightarrow \mathcal{A}$ by their action on objects only such as

$$FX = X \times X + 1.$$

This means that the functor F sends objects A to $A \times A + 1$ (where 1 is the terminal object of \mathcal{A}) and arrows f to $f \times f + \text{id}_1$.

¹i.e. that has all the (co)limits needed for our definition to make sense.

Chapter 2

Algebras for endofunctors

2.1 F -algebras

Though this section let \mathcal{A} be an arbitrary category that we call the **base category** and $F: \mathcal{A} \rightarrow \mathcal{A}$ an endofunctor.

Definition 2.1.1. An **algebra** for F (or an **F -algebra**) is a pair (A, α) where A is an object of \mathcal{A} (that we call the **carrier** or **base object**) and α is an arrow of type $FA \rightarrow A$ (that we call the **structure** of the algebra).

F -algebra

As one should expect F -algebras form a category of their own.

Definition 2.1.2. Given F -algebras (A, α) and (B, β) a **morphism** between them is an arrow f of \mathcal{A} such that

$$\begin{array}{ccc} FA & \xrightarrow{\alpha} & A \\ \downarrow Ff & & \downarrow f \\ FB & \xrightarrow{\beta} & B \end{array}$$

commutes.

Remark 2.1.3. Given morphisms of algebras $f: (A, \alpha) \rightarrow (B, \beta)$ and $g: (B, \beta) \rightarrow (C, \gamma)$ their composition is exactly $g \circ f$; one can easily check that $g \circ f$ makes the diagram above commute using that f and g are morphisms of F -algebras and that F is a functor. The identity on an algebra (A, α) is exactly id_A .

We call the category of F -algebras and morphisms of F -algebras $\text{Alg}F$.

$\text{Alg}F$

Example 2.1.4. In Set let 1 be the terminal object $\{*\}$ and consider the functor

$$FX = X + 1.$$

To give an arrow $\alpha: A + 1 \rightarrow A$ is equivalent to giving two functions $\alpha_0: 1 \rightarrow A$, $\alpha_1: A \rightarrow A$; that is $\alpha = [\alpha_1, \alpha_0]$. So an algebra for F is set A with a constant α_0 and a unary operation α_1 . A morphism of F -algebras is a function f such that

$$\begin{array}{ccc} A + 1 & \xrightarrow{[\alpha_1, \alpha_0]} & A \\ \downarrow f + \text{id}_1 & & \downarrow f \\ B + 1 & \xrightarrow{[\beta_1, \beta_0]} & B \end{array}$$

commutes. This translates, using Proposition 1.1.8, into the following two conditions

$$f \circ \alpha_1 = \beta_1 \circ f, \quad f \circ \alpha_0 = \beta_0.$$

Which together give that f preserves the constant and commutes with the unary operation of the algebras.

Example 2.1.5. Similarly algebras for the endofactor $FX = X \times X + 1$ are sets equipped with a constant and a binary operation. Morphisms of such algebras preserve the constant and commute with the operation. Indeed if (A, a, \otimes_A) and (B, b, \otimes_B) are two F -algebras a morphism $f: A \rightarrow B$ is a function such that

1. $f(a) = b$;
2. $f(x \otimes_A y) = f(x) \otimes_B f(y)$ for all $x, y \in A$.

This is reminiscent of structures such as groups and monoids.

Example 2.1.6. Let Σ_1 be a set; then algebras for $FX = \Sigma_1 \times X$ are sets with a unary operation for every element $\sigma \in \Sigma_1$. Indeed let A be an F -algebra and let $\alpha: \Sigma_1 \times A \rightarrow A$ be its structure; for a fixed $\sigma \in \Sigma_1$ we have that $\alpha(\sigma, -)$ is a unary operation on A . Morphisms of F -algebras in this case commute with all such unary operations.

Examples 2.1.4, 2.1.5, 2.1.6 can be generalized with the introduction of polynomial functors over \mathbf{Set} . This captures the idea of a Σ -algebra from universal algebra and provides a generalization to other base categories with sufficient structure.

Definition 2.1.7. A **signature** Σ is a collection of sets $(\Sigma_n)_{n < \omega}$ where Σ_n is called the set of n -ary symbols. The symbols in Σ_0 are also called constant symbols (or symbols for constants).

signature

Definition 2.1.8. Given a signature Σ a **Σ -algebra** is a set A together with interpretations for every symbol of Σ . That is: to every $\sigma \in \Sigma_0$ we associate a fixed element of A and to every $\sigma \in \Sigma_n$ an n -ary function on A . We denote the interpretation of σ in A by σ^A .

 Σ -algebra

Given two Σ -algebras A and B a **morphism** of Σ -algebras is a function $f: A \rightarrow B$ such that

$$f\sigma^A = \sigma^B \quad (2.1.1)$$

for every $\sigma \in \Sigma_0$ and

$$f(\sigma^A(x_1, \dots, x_n)) = \sigma^B(f(x_1), \dots, f(x_n)) \quad (2.1.2)$$

for every $\sigma \in \Sigma_n$ and $x_1, \dots, x_n \in A$.

We will now observe how Σ -algebras can be realized as algebras for specific endofunctors. To every signature we associate a **polynomial functor** $H_\Sigma: \mathbf{Set} \rightarrow \mathbf{Set}$ that operates as follows on objects X and arrows f .

polynomial functor

$$\begin{aligned} H_\Sigma X &= \coprod_{n < \omega} \Sigma_n \times X^n \\ H_\Sigma f &= \coprod_{n < \omega} \text{id}_{\Sigma_n} \times \underbrace{f \times \dots \times f}_{n \text{ times}} \end{aligned}$$

Then if A is a Σ -algebra we can define functions $\alpha_n: \Sigma_n \times A^n \rightarrow A$ naturally as

$$\alpha_n(\sigma, x_1, \dots, x_n) := \sigma^A(x_1, \dots, x_n) \quad (2.1.3)$$

and then $\alpha := [\alpha_0, \alpha_1, \dots]$ makes A into a H_Σ -algebra. Conversely if we have a H_Σ -algebra (A, α) we can define an interpretation of every symbol by reading (2.1.3) the other way around.

For morphisms consider the following square that we know commutes if f is a morphism of H_Σ -algebras.

$$\begin{array}{ccc} \coprod_{n < \omega} \Sigma_n \times A^n & \xrightarrow{\alpha} & A \\ \downarrow H_\Sigma f & & \downarrow f \\ \coprod_{n < \omega} \Sigma_n \times B^n & \xrightarrow{\beta} & B \end{array}$$

Since the diagonal of the square is an arrow out of a coproduct we know from Proposition 1.1.8 that $f \circ \alpha$ and $\beta \circ H_\Sigma f$ are equal when composed with the coprojections of the cocone and this can happen if and only if conditions (2.1.1) and (2.1.2) hold.

We now turn to continuous algebras.

Definition 2.1.9. A **complete partial order** (or a **CPO**) is a poset (A, \sqsubseteq) in which all ω -chains

$$a_1 \sqsubseteq a_2 \sqsubseteq a_3 \sqsubseteq \dots$$

have a join i.e. a least upper bound; we denote it by $\bigvee_{n < \omega} a_n$. A function $f: (A, \sqsubseteq) \rightarrow (B, \sqsubseteq)$ is **continuous** if

1. it is monotone i.e. $a \sqsubseteq b$ in A implies $f(a) \sqsubseteq f(b)$ in B ;
2. given an ω -chain $a_1 \sqsubseteq a_2 \sqsubseteq a_3 \sqsubseteq \dots$ we have $f(\bigvee_{n < \omega} a_n) = \bigvee_{n < \omega} f(a_n)$.

Note that by 1 any chain in A is preserved by f and since B is a CPO it must have a join; condition 2 forces that join to be the image of the join of the original chain under f .

Remark 2.1.10. CPOs and continuous functions form a category CPO. In CPO (co)products are formed as in Set and have, for products, the pointwise order and, for coproducts, each component keeps its own ordering and elements in different components are never comparable. Moreover CPO has a terminal object 1 that is the singleton $\{*\}$ with the only possible partial order on it.

Example 2.1.11. Consider the functor $FX = X + 1$ on CPO. Algebras for F are CPOs with a constant and a continuous unary operation. Similarly algebras for the functor $FX = X \times X + 1$ are CPOs with a constant and a continuous binary operation. The continuity of the operation (which we denote here as \otimes) in this case gives us that if

$$a_1 \sqsubseteq a_2 \sqsubseteq a_3 \sqsubseteq \dots \quad \text{and} \quad b_1 \sqsubseteq b_2 \sqsubseteq b_3 \sqsubseteq \dots$$

are chains then we have a chain $(a_1, b_1) \sqsubseteq (a_2, b_2) \sqsubseteq \dots$ is a chain in the product and

$$\bigvee_{n < \omega} a_n \otimes \bigvee_{n < \omega} b_n = \bigvee_{n < \omega} a_n \otimes b_n.$$

complete partial
order

continuous function

Example 2.1.12. Now let $FX = X_\perp$ be the functor that adds to a CPO a bottom element (i.e. an element \perp such that $\perp \sqsubseteq x$ for every x in the poset) and that sends a continuous function $f: X \rightarrow Y$ to $f_\perp: X_\perp \rightarrow Y_\perp$ defined as

$$f_\perp(x) = \begin{cases} f(x) & \text{if } x \neq \perp \\ \perp & \text{if } x = \perp \end{cases}.$$

An algebra (A, α) for F has a unary continuous function $\alpha_1: A \rightarrow A$ and a constant α_\perp such that $\alpha_\perp \sqsubseteq \alpha(a)$ for every $a \in A$ by monotony of α . Comparing this to Example 2.1.11 by changing the functor we now have a condition on our constant.

Finally let CPO_\perp be the category of CPOs with a bottom element and continuous functions that preserve the bottom i.e. such that $f(\perp) = \perp$. We call such functions **strict continuous functions**.

strict continuous
function

Remark 2.1.13. In CPO_\perp products work as in CPO (the bottom element is naturally the one whose components are all \perp) but the construction of coproducts must be changed. Indeed coproducts in CPO_\perp are formed as in CPO but the bottoms of all the components are unified into a single element. The initial object is still the singleton as it clearly has a bottom.

Remark 2.1.14. Notice that the functor $FX = X + 1$ on CPO_\perp is the identity functor so its algebras are CPOs with a bottom and a strict continuous unary operation. If we want a constant on top of the unary operation we can consider the functor $FX = X_\perp + 1_\perp$, $Ff = f_\perp + \text{id}_{1_\perp}$. Indeed if (A, α) is an F -algebra then $\alpha = [\alpha_1, \alpha_0]$ with $\alpha_1: A_\perp \rightarrow A$ and $\alpha_0: 1_\perp \rightarrow A$ are strict continuous functions. Indeed α_0 picks an element of A : we need the domain to be 1_\perp because otherwise we would be forced to pick the bottom of A as the function is strict. Similarly α_1 gives a continuous unary operation on A (not necessarily a strict one!) as adding a new bottom to A “frees” the first one from being preserved.

2.2 Initial Algebras

Definition 2.2.1. An **initial algebra** for an endofunctor F is an initial object in $\text{Alg}F$. When an initial algebra exists we denote it by $(\mu F, \iota)$.

initial algebra

Remark 2.2.2. If 0 is the initial object of \mathcal{A} and F preserves initial objects then $(0, \text{id}_0)$ is the initial algebra for F .

We will now look at some non-trivial examples of initial algebras.

Example 2.2.3. Consider the functor $FX = X + 1$ on Set . The initial algebra for F is (\mathbb{N}, ι) where $\iota = [\iota_1, \iota_0]$ with $\iota_0 = 0$ and ι_1 the successor function on the naturals. Indeed let (B, β) be another algebra with $\beta = [\beta_1, \beta_0]$ and consider the function $f: \mathbb{N} \rightarrow B$ defined as

$$\begin{cases} f(0) & = \beta_0 \\ f(n+1) & = \beta_1(f(n)) \end{cases} \quad (2.2.1)$$

This is a morphism of algebras:

1. $f(0) = \beta_0$;
2. $(f \circ \iota_1)(n) = f(n+1) = (\beta_1 \circ f)(n)$ for all $n \in \mathbb{N}$ so $f \circ i_1 = \beta_1 \circ f$;

and clearly unique because any other morphism must satisfy 1 and 2, thus a simple induction gives us that it must be f .

Example 2.2.4. Let $\mathcal{P}_f: \text{Set} \rightarrow \text{Set}$ be the finite power-set functor. This functor sends a set X to $\mathcal{P}_f X = \{Z \subseteq X : Z \text{ is finite}\}$ and a function $f: X \rightarrow Y$ to the function $\mathcal{P}_f f$ that sends a finite subset of X to its (obviously finite) f -image in Y . Let V_ω be the set of hereditarily finite sets. That is: the set containing the empty set and all finite sets whose elements are hereditarily finite sets as well. We will see in Section ?? that (V_ω, id) is the initial algebra of \mathcal{P}_f .

finite power-set
functor

Example 2.2.5. Let \mathcal{A} be a category with countable coproducts. If $A \in \mathcal{A}$ we write $\mathbb{N} \bullet A$ for the coproduct of A with itself countably many times. We shall now consider the functor $FX = X + A$ for which we prove the initial algebra to be $\mathbb{N} \bullet A$.

First we need an algebra structure ι on $\mathbb{N} \bullet A$. If $\text{in}_k: A \rightarrow \mathbb{N} \bullet A$ is the k -th coprojection then we set

$$\iota = [\alpha_1, \text{in}_0]: (\mathbb{N} \bullet A) + A \rightarrow \mathbb{N} \bullet A \quad (2.2.2)$$

where α_1 is obtained from the universal property of the coproduct applied to the cone $(\text{in}_k)_{1 \leq k < \omega}$. We thus have that the following triangles commute for every $k < \omega$.

$$\begin{array}{ccc} \mathbb{N} \bullet A & \xrightarrow{\alpha_1} & \mathbb{N} \bullet A \\ & \nwarrow \text{in}_k \quad \nearrow \text{in}_{k+1} & \\ & A & \end{array}$$

Now let (B, β) with $\beta = [\beta_1, \beta_0]$ be an algebra and f a morphism from $(\mathbb{N} \bullet A, \iota)$. We have that the following square commutes.

$$\begin{array}{ccc} \mathbb{N} \bullet A + A & \xrightarrow{[\alpha_1, \text{in}_0]} & \mathbb{N} \bullet A \\ \downarrow f + \text{id} & & \downarrow f \\ B + A & \xrightarrow{[\beta_1, \beta_0]} & B \end{array}$$

So any such f must be such that

1. $f \circ \text{in}_0 = \beta_0$;
2. $f \circ \alpha_1 \circ \text{in}_k = f \circ \text{in}_{k+1} = \beta_1 \circ f \circ \text{in}_k$ for all $k \in \mathbb{N}$.

This gives us that $[\beta_0, \beta_1 \circ \beta_0, \beta_1 \circ \beta_1 \circ \beta_0, \dots]$ is the unique morphism from $(\mathbb{N} \bullet A, \iota)$ to (B, β) which proves our claim.

Finally we describe initial algebras for polynomial functors in two equivalent ways: using closed terms and using trees.

Definition 2.2.6. Let Σ be a signature. A **closed term** for Σ is a string t of symbols of Σ such that:

closed term

1. t is a constant symbol;
2. t is of the form $\sigma(t_1, \dots, t_n)$ where $\sigma \in \Sigma_n$ and t_1, \dots, t_n are closed terms.

Remark 2.2.7. Let μH_Σ be the set of all closed terms; this is naturally a Σ -algebra:

1. if σ is a constant symbol $\sigma^{\mu H_\Sigma}$ is σ itself but regarded as a term;
2. if σ is an n -ary symbol then it defines an n -ary operation on μH_Σ that takes terms t_1, \dots, t_n to the term $\sigma(t_1, \dots, t_n)$.

Remark 2.2.8. Let A be a Σ -algebra and t a closed term. We can define an evaluation function $\text{ev}: \mu H_\Sigma \rightarrow A$ as follows:

1. if t is a constant symbol σ then $\text{ev}(t) = \sigma^A$;
2. if t is of the form $\sigma(t_1, \dots, t_n)$ then $\text{ev}(t) = \sigma^A(\text{ev}(t_1), \dots, \text{ev}(t_n))$.

Clearly ev is a morphism and moreover it is unique so μH_Σ is the initial algebra for H_Σ .

Definition 2.2.9. Given a signature Σ a Σ -**tree** is an ordered tree where every node of k children is labelled by a k -ary symbol of Σ .

Remark 2.2.10. Every n -ary symbol of Σ defines an n -ary operation on Σ -trees which takes n trees to the tree obtained by connecting each root to a new node labelled by σ , which becomes the root of a new tree. We call this operation **tree-tupling**. Note that the order of the trees matters.

tree-tupling

Proposition 2.2.11. The initial algebra μH_Σ is the algebra of finite Σ -trees with tree-tupling.

Proof. Let T be the algebra of finite Σ -trees with tree-tupling; we shall find an isomorphism $f: \mu H_\Sigma \rightarrow T$ of algebras. Define f by structural recursion as follows

1. if t is a constant term σ let $f(t)$ be the tree formed by a single node labelled by σ ;
2. if t is a term of the form $\sigma(t_1, \dots, t_n)$ let $f(t)$ be the tree obtained by tree-tupling $f(t_1), \dots, f(t_n)$ with the new root labelled by σ .

This function is a morphism by definition and has an inverse, defined similarly by structural recursion, which is again a morphism. \square

Example 2.2.12. By the discussion above the initial algebra for the functor $FX = X \times X + 1$ on Set is the algebra of all finite binary trees.

Example 2.2.13. Consider the functor $FX = B \times X + 1$ on \mathbf{Set} with $B \in \mathbf{Set}$. Algebras for F are sets with a constant and $|B|$ unary operations. By the discussion above we know that μF is the algebra of finite trees for the signature $\Sigma = (\Sigma_0 = \{*\}, \Sigma_1 = B)$ so its elements are “linear” trees such as



with $a, b, c \in B$. We immediately deduce that μF can also be realized as the set of words over B .

Remark 2.2.14. Notice the importance of constant symbols. Indeed if a polynomial functor on \mathbf{Set} has no constant then it preserves the initial object 0 so by Remark 2.2.2 the initial algebra is trivial. Equivalently one can observe that, without constant symbols, the set of closed terms is empty.

Example 2.2.15. Consider $FX = X_\perp$ on \mathbf{CPO}_\perp . Let \mathbb{N}^\top be the set of the natural numbers with an added topmost element ∞ ordered naturally. Notice that $(\mathbb{N}^\top)_\perp$ is isomorphic (as an order) to \mathbb{N}^\top and consider the successor function $s: (\mathbb{N}^\top)_\perp \rightarrow \mathbb{N}^\top$ with $s(\infty) = \infty$. We claim this is the initial algebra for F .

Let (A, α) be another F -algebra and $f: (\mathbb{N}^\top, s) \rightarrow (A, \alpha)$ a morphism. Because f is an arrow of \mathbf{CPO}_\perp it must be strict so $f(0) = \perp$ and because it is a morphism we must have $f(s(n)) = \alpha(f(n))$; this defines f inductively on \mathbb{N} . Finally we recall that f must be continuous (as an arrow of \mathbf{CPO}_\perp) so

$$f(\infty) = f\left(\bigvee_{n < \omega} n\right) = \bigvee_{n < \omega} (f(n)).$$

This implies that there is really a unique morphism from (\mathbb{N}^\top, s) to (A, α) so (\mathbb{N}^\top, s) is initial.

We conclude this section with the a classic lemma of Lambek’s which gives necessary conditions for functors to have an initial algebra.

Definition 2.2.16. A **fixed point** of an endofunctor F is an element $A \in \mathcal{A}$ that is isomorphic to FA .

Lemma 2.2.17 (Lambek’s Lemma). An initial algebra for F is always a fixed point.

Proof. Let $(\mu F, \iota)$ be the initial algebra of F . Notice that $(F(\mu F), F\iota)$ is also an algebra and so there is a unique morphism $f: (\mu F, \iota) \rightarrow (F(\mu F), F\iota)$. We have that the following

diagram commutes

$$\begin{array}{ccc}
 F(\mu F) & \xrightarrow{\iota} & \mu F \\
 \downarrow Ff & & \downarrow f \\
 F(F(\mu F)) & \xrightarrow{F\iota} & F(\mu F) \\
 \downarrow F\iota & & \downarrow \iota \\
 F(\mu F) & \xrightarrow{\iota} & \mu F
 \end{array}$$

so $\iota \circ f$ is an endomorphism of algebras on the initial algebra; hence it is $\text{id}_{\mu F}$. Now

$$f \circ \iota = F\iota \circ Ff = F(\iota \circ f) = F(\text{id}_{\mu F}) = \text{id}_{F(\mu F)}$$

so ι is an isomorphism. This shows that μF is a fixed point of F . \square

Remark 2.2.18. We know from Cantor's Theorem that there is no surjection from a given set X into its power set $\mathcal{P}X$ so the functor \mathcal{P} on Set has no fixed point hence no initial algebra.

Lambeck's lemma can be seen as a generalization of the following order-theoretic lemma.

Definition 2.2.19. Let (P, \sqsubseteq) be a poset and $f: P \rightarrow P$ a monotone function. An element $x \in P$ is a **pre-fixed point** of f if $f(x) \sqsubseteq x$.

Lemma 2.2.20. Given a poset (P, \sqsubseteq) and a monotone function $f: P \rightarrow P$ let $A = \{x \in P: f(x) \sqsubseteq x\}$ be the set of pre-fixed points of f . If \bar{a} is the meet of A then \bar{a} is the least fixed point of f .

Proof. If $x \in A$ then $\bar{a} \sqsubseteq x$ and $f(\bar{a}) \sqsubseteq f(x) \sqsubseteq x$ follows from monotony of f and definition of A . This shows that $f(\bar{a})$ is a lower bound for A so $f(\bar{a}) \sqsubseteq \bar{a}$. Now by applying f again we obtain $f(f(\bar{a})) \sqsubseteq f(\bar{a})$ which gives us $f(\bar{a}) \in A$ and thus $\bar{a} \sqsubseteq f(\bar{a})$. Moreover \bar{a} must be the least fixed point of f since all fixed points are pre-fixed points. \square

This lemma then gives us the following well-known theorem for free.

Theorem 2.2.21 (Knaster-Tarski Theorem). Let (L, \sqsubseteq) be a complete lattice. Then every monotone function of L has a fixed point.

Proof. Let f be a monotone function on L and A be the set of pre-fixed points of f . It must have a meet \bar{a} because the lattice is complete so by the Lemma we have a fixed point. \square

The generalization of order theoretic results such as Lemma 2.2.20 will be a recurrent theme. Indeed we can see orders as categories in the usual way, order-preserving functions as functors and pre-fixed points as algebras for these functors.

2.3 Recursion and Induction

Recursion is a way of defining a function f on the natural numbers by specifying a value for $f(0)$ and a way of deriving the value of $f(n + 1)$ from the value of $f(n)$. Recursion, however, can also be used to define functions on (among other things) trees; in this case it is usually called *structural* induction (we have an example of this in Remark 2.2.8). Here we show how one can use the concept of an initial algebra to provide a general definition of recursion of which the cited examples as particular cases.

Definition 2.3.1. Let F be an endofunctor and μF its initial algebra. A morphism $f: \mu F \rightarrow A$ is **recursively specified** if there exists an algebra structure $\alpha: FA \rightarrow A$ on A that makes f into an algebra homomorphism (actually the unique algebra homomorphism).

recursively
specified

Example 2.3.2. Consider the functor $FX = X \times X + 1$ on \mathbf{Set} and let T be its initial algebra that we know from Example 2.2.12 is the algebra of finite binary trees with tree-tupling. Let $h: T \rightarrow \mathbb{N}$ be the function that assigns 0 to the root-only tree and, given a tree t obtained by tree-tupling of t_1 and t_2 , $h(t) = 1 + \max\{h(t_1), h(t_2)\}$. We now show that h is recursively specified according to our definition.

We need an algebra structure on \mathbb{N} i.e. an arrow $\alpha: \mathbb{N} \times \mathbb{N} + 1 \rightarrow \mathbb{N}$. Since α is an arrow out of a coproduct we have $\alpha = [\alpha_1, \alpha_0]$ so we can let

$$\begin{aligned}\alpha_0 &= 0; \\ \alpha_1(n, m) &= 1 + \max\{n, m\}.\end{aligned}$$

Indeed if ι_1 is tree-tupling and ι_0 is the root-only tree the following square trivially commutes thus h is recursively specified.

$$\begin{array}{ccc} T \times T + 1 & \xrightarrow{[\iota_1, \iota_0]} & T \\ \downarrow h \times h + \text{id}_1 & & \downarrow h \\ \mathbb{N} \times \mathbb{N} + 1 & \xrightarrow{[\alpha_1, \alpha_0]} & \mathbb{N} \end{array}$$

The following theorem is a generalization of the classic recursion with parameters theorem.

Theorem 2.3.3 (Primitive Recursion). Assume the base category \mathcal{A} to have finite products and let F be an endofunctor with an initial algebra μF . Then for every $\alpha: F(A \times \mu F) \rightarrow A$ there is a unique $h: \mu F \rightarrow A$ such that the following square commutes.

$$\begin{array}{ccc} F(\mu F) & \xrightarrow{\iota} & \mu F \\ \downarrow F\langle h, \text{id}_{\mu F} \rangle & & \downarrow h \\ F(A \times \mu F) & \xrightarrow{\alpha} & A \end{array} \quad (2.3.1)$$

Proof. Let π_1, π_2 be the projections of the product $A \times \mu F$ and consider the arrow

$$\bar{\alpha}: F(A \times \mu F) \xrightarrow{\langle \text{id}_{F(A \times \mu F)}, F\pi_2 \rangle} F(A \times \mu F) \times F(\mu F) \xrightarrow{\alpha \times \iota} A \times \mu F.$$

This gives an algebra structure on $A \times \mu F$ so let $\bar{h}: \mu F \rightarrow A \times \mu F$ be the unique algebra homomorphism given by the initiality of μF .

$$\begin{array}{ccccc}
 F(\mu F) & \xrightarrow{\quad \iota \quad} & & \mu F & \\
 \downarrow F\bar{h} & & (1) & \downarrow \bar{h} & \\
 F(A \times \mu F) & \xrightarrow{\langle \text{id}_{F(A \times \mu F)}, F\pi_2 \rangle} & F(A \times \mu F) \times F(\mu F) & \xrightarrow{\alpha \times \iota} & A \times \mu F \\
 \downarrow F\pi_2 & \nearrow \pi_2 & & \downarrow \pi_2 & \\
 F(\mu F) & \xrightarrow{\quad \iota \quad} & & \mu F &
 \end{array} \quad (2.3.2)$$

The outer square in (2.3.2) commutes because (1), (2) and (3) do. Indeed (1) commutes because \bar{h} is an algebra homomorphism, (2) by Notation 1.1.14 and (3) because of Notation 1.1.13. By functoriality of F we then have that $\pi_2 \circ \bar{h}$ is an endomorphism on the initial algebra thus $\pi_2 \circ \bar{h} = \text{id}_{\mu F}$.

Now set $h := \pi_1 \circ \bar{h}$ so $\bar{h} = \langle h, \text{id}_{\mu F} \rangle$. Extending (1) by π_1 we obtain the following diagram that we know commutes.

$$\begin{array}{ccccccc}
 F(\mu F) & \xrightarrow{\quad \iota \quad} & & \mu F & & & \\
 \downarrow F\bar{h} = F\langle h, \text{id}_{\mu F} \rangle & & & \downarrow \bar{h} & \searrow h & & \\
 F(A \times \mu F) & \xrightarrow{\langle \text{id}_{F(A \times \mu F)}, F\pi_2 \rangle} & F(A \times \mu F) \times F(\mu F) & \xrightarrow{\alpha \times \iota} & A \times \mu F & \xrightarrow{\pi_1} & A
 \end{array}$$

But notice that

$$\pi_1 \circ (\alpha \times \iota) \circ \langle \text{id}_{F(A \times \mu F)}, F\pi_2 \rangle = \pi_1 \circ \langle \alpha, \iota \circ F\pi_2 \rangle = \alpha.$$

so we have (2.3.1).

To prove uniqueness consider $h: \mu F \rightarrow A$ homomorphism of algebras such that (2.3.1) commutes. We claim that $\bar{h} = \langle h, \text{id}_{\mu F} \rangle$ so that $h = \pi_1 \circ \bar{h}$; thus uniqueness. In order to do it we show that $\langle h, \text{id}_{\mu F} \rangle$ is an algebra homomorphism and conclude it must be \bar{h} by initiality of μF . Indeed we have

$$\begin{aligned}
 \pi_1 \circ (\alpha \times \iota) \circ \langle \text{id}_{F(A \times \mu F)}, F\pi_2 \rangle \circ F\langle h, \text{id}_{\mu F} \rangle &= \alpha \circ \text{id}_{F(A \times \mu F)} \circ F\langle h, \text{id}_{\mu F} \rangle \\
 &= h \circ \iota \\
 &= \pi_1 \circ \langle h, \text{id}_{\mu F} \rangle \circ \iota;
 \end{aligned}$$

$$\begin{aligned}
 \pi_2 \circ (\alpha \times \iota) \circ \langle \text{id}_{F(A \times \mu F)}, F\pi_2 \rangle \circ F\langle h, \text{id}_{\mu F} \rangle &= \iota \circ F\pi_2 \circ F\langle h, \text{id}_{\mu F} \rangle \\
 &= \iota \circ F(\pi_2 \circ \langle h, \text{id}_{\mu F} \rangle) \\
 &= \iota \circ F\text{id}_{\mu F} \\
 &= \iota \circ \text{id}_{F(\mu F)} \\
 &= \iota \\
 &= \pi_2 \circ \langle h, \text{id}_{\mu F} \rangle \circ \iota;
 \end{aligned}$$

so we conclude by Proposition 1.1.4. \square

Recall the classic recursion with parameters theorem. We shall prove it using Theorem 2.3.3; for a classical proof see [And19, Theorem 18.1].

Theorem 2.3.4 (Recursion with parameters). Let A be a set, $a \in A$ and $f: A \times \mathbb{N} \rightarrow A$ be a function. Then there is a function $g: \mathbb{N} \rightarrow A$ such that

$$\begin{cases} g(0) = a \\ g(n+1) = f(g(n), n) \end{cases}.$$

Proof. Consider the functor $FX = X + 1$ on \mathbf{Set} and form the arrow

$$\alpha: F(A \times \mu F) = A \times \mathbb{N} + 1 \xrightarrow{[f, a]} A.$$

Now by Theorem 2.3.3 there is a function $g: \mathbb{N} \rightarrow A$ such that (2.3.1) commutes. This gives us the diagram

$$\begin{array}{ccc} \mathbb{N} \times 1 & \xrightarrow{[\iota_1, \iota_0]} & \mathbb{N} \\ \downarrow \langle g, \text{id}_{\mathbb{N}} \rangle + \text{id}_1 & & \downarrow g \\ A \times \mathbb{N} + 1 & \xrightarrow{[f, a]} & A \end{array} \quad (2.3.3)$$

that yields the conditions

1. $g \circ \iota_0 = a \circ \text{id}_1$;
2. $g \circ \iota_1 = f \circ \langle g, \text{id}_{\mathbb{N}} \rangle$.

Now 1 gives us $g(0) = a$ and 2 that

$$g(n+1) = (g \circ \iota_1)(n) = (f \circ \langle g, \text{id}_{\mathbb{N}} \rangle)(n) = f(g(n), n) \quad \text{for all } n < \omega.$$

So g has the required properties. \square

The induction principle, that if a subset of \mathbb{N} contains 0 and is closed under the successor operation then it is \mathbb{N} , can also be generalized to the more general setting of F -algebras. Recall that a subobject in a category is represented by a monomorphism. We call subalgebras subobjects in $\mathbf{Alg} F$.

Theorem 2.3.5 (Induction Principle). Let F be an endofunctor and μF its initial algebra. Then every subalgebra of F is an isomorphism.

Proof. Let $m: (B, \beta) \rightarrow (\mu F, \iota)$ be a subalgebra. By initiality there is a unique homomorphism of F -algebras $h: (\mu F, \iota) \rightarrow (B, \beta)$. Now $m \circ h$ is an endomorphism of the initial object of $\mathbf{Alg} F$ so $m \circ h = \text{id}_{\mu F}$, but notice that we have m monic and split-epi thus an isomorphism. \square

Chapter 3

Initial algebras from finitary iteration

3.1 Adámek's Theorem

We begin this section by recalling Kleene's Fixed Point Theorem and its proof with the goal of generalizing it to the categorical setting.

Theorem 3.1.1 (Kleene's Fixed Point Theorem). Let A be a CPO with bottom \perp ; then every continuous function $F: A \rightarrow A$ has a least fixed point $\mu F = \sup_{n < \omega} F^n(\perp)$ ¹.

Proof. Consider the ω -chain $\perp \leq F(\perp) \leq F^2(\perp) \leq \dots$ and let μF be its join. By continuity $F(\mu F) = \bigvee_{n < \omega} F(F^n(\perp))$ but we know that $\bigvee_{n < \omega} F^n(\perp) = \bigvee_{n < \omega} F^{n+1}(\perp)$ so $\mu F = F(\mu F)$, a fixed point.

Now let $F(x) \leq x$ be a pre-fixed point. As $\perp \leq x$ we have $F(\perp) \leq F(x) \leq x$ and by induction we obtain that $F^n(\perp) \leq x$ for all $n < \omega$ so x . But this shows that x is an upper bound so $\mu F \leq x$ by definition and, finally, μF must be the least fixed point because it is less than any pre-fixed point and fixed points are trivially pre-fixed points. \square

Unsurprisingly we shall replace CPOs with bottom by categories with an initial object and endofunctions by endofunctors. The ω -chain used in the proof of Theorem 3.1.1 becomes a diagram and the continuity condition become a colimit-preservation one.

Definition 3.1.2. The **initial-algebra ω -chain** of an endofunctor F is the diagram

initial-algebra
 ω -chain

$$0 \xrightarrow{!} F0 \xrightarrow{F!} F^2 0 \xrightarrow{F^2!} F^3 0 \xrightarrow{F^3!} \dots \quad (3.1.1)$$

where we denote by $!$ the unique arrow out of the initial object. This diagram can be realized as a functor from ω regarded as a category.

Remark 3.1.3. Let (A, α) be an F -algebra. Then it induces a canonical cocone $(\alpha_n: F^n 0 \rightarrow A)_{n < \omega}$ on (3.1.1) by

$$\begin{aligned} \alpha_0 &= ! \\ \alpha_{n+1} &= F F^n 0 \xrightarrow{F \alpha_n} F A \xrightarrow{\alpha} A. \end{aligned}$$

To check that the α_n 's are a cocone we prove $\alpha_n = \alpha_{n+1} \circ F^n!$ by induction on n . When $n = 0$ the condition becomes $! = \alpha_1 \circ !$ which is true because both sides are arrows out of the initial object. Now suppose the condition holds for $n - 1$:

$$\begin{aligned} \alpha_{n+1} \circ F^n! &= \alpha \circ F \alpha_n \circ F^k! && \text{definition of } \alpha_{n+1} \\ &= \alpha \circ F(\alpha_n \circ F^{n-1}!) && \text{functoriality of } F \\ &= \alpha \circ F \alpha_{n-1} && \text{inductive hypothesis} \\ &= \alpha_n && \text{definition of } \alpha_n. \end{aligned}$$

Remark 3.1.4. If we apply F to the initial-algebra ω -chain we obtain

$$F0 \xrightarrow{F!} F^2 0 \xrightarrow{F^2!} F^3 0 \xrightarrow{F^3!} \dots \quad (3.1.2)$$

¹We write F^n for $\underbrace{F \dots F}_{n \text{ times}}$ and F^0 is the identity function. We use a similar notation for functors.

This ω -chain has the same colimit as the original one. Indeed since the first element of (3.1.1) is the initial object there is an obvious one-to-one correspondence between cocones on (3.1.1) and cocones on the new chain above; and the same factorization morphisms work.

We are ready to state and prove the main theorem of this chapter.

Theorem 3.1.5 (Adámek). let \mathcal{A} be a category with an initial object 0 and colimits of ω -chains and F an endofunctor that preserves ω -colimits. Then F has an initial algebra $\mu F = \text{colim}_{n < \omega} F^n 0$.

Proof. Let $(\mu F, (c_n: F^n 0 \rightarrow \mu F)_{n < \omega})$ be the colimit of the initial-algebra ω -chain. Since F preserves ω -colimits we have that $(F(\mu F), (Fc_n: F^{n+1} 0 \rightarrow F(\mu F))_{n < \omega})$ is the colimit of (3.1.2) but $(c_{n+1}: F^{n+1} 0 \rightarrow \mu F)_{n < \omega}$ is a cocone on (3.1.2) so there is a unique arrow ι from $F(\mu F)$ to μF such that

$$\iota \circ Fc_n = c_{n+1} \quad \text{for all } n < \omega. \quad (3.1.3)$$

This gives an F -algebra structure to μF that we now check is initial.

Let (A, α) be a generic F -algebra. We have the induced cocone $(\alpha_n: F^n 0 \rightarrow \mu F)_{n < \omega}$ as in Remark 3.1.3 and this gives a unique arrow $h: \mu F \rightarrow A$ such that $h \circ c_n = \alpha_n$ for $n < \omega$. We shall now prove that h is a morphism of F -algebras and that it is unique.

To prove that $h \circ \iota = \alpha \circ Fh$ we check that both arrows are factorization morphisms for the cocone $(\alpha_{n+1}: F^{n+1} 0 \rightarrow A)_{n < \omega}$. In practice we shall check that

$$h \circ \iota \circ Fc_n = \alpha_{n+1} = \alpha \circ Fh \circ Fc_n$$

for all $n < \omega$. For the first equality notice that $h \circ \iota \circ Fc_n = h \circ c_{n+1} = \alpha_{n+1}$. For the second equality consider the following diagram

$$\begin{array}{ccc} F^{n+1}0 & \xrightarrow{\alpha_{n+1}} & A \\ Fc_n \downarrow & \searrow F\alpha_n & \uparrow \alpha \\ F(\mu F) & \xrightarrow{Fh} & FA \end{array} .$$

The upper-right triangle commutes by definition of α_{n+1} (see Remark 3.1.3). The lower-left triangle commute because h is the factorization arrow and by functoriality of F . Then, as the whole square commutes, $\alpha \circ Fh \circ Fc_n = \alpha_{n+1}$. This proves that $h: \mu F \rightarrow A$ is a homomorphism of F -algebras.

To prove uniqueness suppose there is an arrow $k: \mu F \rightarrow A$ such that it is also an F -algebra homomorphism i.e. $k \circ \iota = \alpha \circ Fk$. We will show that $k \circ c_n = \alpha_n$ for all $n < \omega$ and conclude by uniqueness of h . Working by induction, if $n = 0$ then $k \circ c_0 = \alpha_0$

is trivial because both arrows are out of the initial object. Now suppose that $k \circ c_n = \alpha_n$:

$$\begin{aligned}
 k \circ c_{n+1} &= k \circ \iota \circ Fc_n && (3.1.3) \\
 &= \alpha \circ Fk \circ Fc_n && k \text{ homomorphism} \\
 &= \alpha \circ F(k \circ c_n) && F \text{ functor} \\
 &= \alpha \circ F\alpha_n && \text{inductive hypothesis} \\
 &= \alpha_{n+1} && \text{definition of } \alpha_{n+1}.
 \end{aligned}$$

This concludes the proof. \square

Remark 3.1.6. Looking at the proof of Theorem 3.1.5 we notice that it isn't really necessary for \mathcal{A} to have colimits of *all* ω -chains or for F to preserve *all* of them as we care for these properties only when applied to the initial-algebra ω -chain. However in practice it may be more convenient to check the general properties rather than the particular case.

We will now look at some examples of functors to which one can apply Adámek's Theorem. We urge the reader to notice that the Theorem does not only guarantee the existence of an initial algebra, but also that such algebra is a colimit. By looking at the diagram in question (the initial-algebra ω -chain) we can often build an intuition of how the initial algebra works.

Remark 3.1.7. We shall prove that in a category \mathcal{A} with binary coproducts, a terminal object and colimits of ω -chains, the functor $FX = X + 1$ preserves colimits of ω -chains.

Consider an ω -chain

$$A_0 \xrightarrow{a_0} A_1 \xrightarrow{a_1} A_2 \xrightarrow{a_2} A_3 \xrightarrow{a_3} \dots \quad (3.1.4)$$

and let $(A, \alpha_i : A_i \rightarrow A)_{i < \omega}$ be its colimit. We want to prove that $(A + 1, \alpha_i + \text{id}_1)_{i < \omega}$ is the colimit of

$$A_0 + 1 \xrightarrow{a_0 + \text{id}_1} A_1 + 1 \xrightarrow{a_1 + \text{id}_1} A_2 + 1 \xrightarrow{a_2 + \text{id}_1} A_3 + 1 \xrightarrow{a_3 + \text{id}_1} \dots \quad (3.1.5)$$

By functoriality of F we have immediately that it is a cocone on (3.1.5) so we only need to show it is universal. Let $(B, \beta_i : A_i + 1 \rightarrow B)_{i < \omega}$ be a cocone on (3.1.5) and note that for all $i < \omega$, $\beta_i = [\beta_i^1, \beta_i^2]$ with $\beta_i^1 : A_i \rightarrow B$ and $\beta_i^2 : 1 \rightarrow B$. By composition with the coprojections of $A_i + 1$ we obtain that

1. $(B, \beta_i^1)_{i < \omega}$ is a cocone on (3.1.4);
2. $(B, \beta_i^2)_{i < \omega}$ is a cocone on

$$1 \xrightarrow{\text{id}_1} 1 \xrightarrow{\text{id}_1} 1 \xrightarrow{\text{id}_1} 1 \xrightarrow{\text{id}_1} \dots \quad (3.1.6)$$

Now (3.1.4) has colimit $(A, \alpha_i)_{i < \omega}$ and (3.1.6) has colimit $(1, \gamma_i = \text{id}_1)_{i < \omega}$ so we obtain unique arrows $m^1 : A \rightarrow B$ and $m^2 : 1 \rightarrow B$ such that

$$m^1 \circ \alpha_i = \beta_i^1 \quad \text{and} \quad m^2 \circ \text{id}_1 = \beta_i^2 \quad \text{for all } i < \omega.$$

Finally we set $m = [m^1, m^2]$ and this is a factorization arrow for $(B, \beta_i)_{i < \omega}$ as the following calculation shows

$$\begin{aligned} m \circ (\alpha_i + \text{id}_1) &= [m^1, m^2] \circ (\alpha_i + \text{id}_1) \\ &= [m^1 \circ \alpha_i, m^2 \circ \text{id}_1] \\ &= [\beta_i^1, \beta_i^2] \\ &= \beta_i. \end{aligned}$$

We only need to show that such $m : A + 1 \rightarrow B$ is unique. Let $n : A + 1 \rightarrow B$ be such that $n \circ (\alpha_i + \text{id}_1) = \beta_i$ for all $i < \omega$. Since n is an arrow out of a coproduct we have $n = [n^1, n^2]$ with $n^1 : A \rightarrow B, n^2 : 1 \rightarrow B$. Then from the factorization condition we obtain $[n^1 \circ \alpha_i, n^2 \circ \text{id}_1] = [\beta_i^1, \beta_i^2]$ and, by composing with the coprojects, $n^1 \circ \alpha_i = \beta_i^1$ and $n^2 \circ \text{id}_1 = \beta_i^2$ for all $i < \omega$. But now we must have $n^1 = m^1$ and $n^2 = m^2$ because m^1, m^2 are unique. Finally $n = m$ and F preserves colimits of ω -chains.

By a similar argument one can show that all polynomial functors preserve colimits of ω -chains.

Example 3.1.8. In a category with binary coproducts and an initial object 1 the initial algebra for $FX = X + 1$, if it exists, is called the **natural numbers object** (or **NNO**) of the category. Notice that a NNO, being an initial algebra, is unique so the use of “the” is justified.

natural numbers
object

Explicitly a NNO is a tuple $(N, \iota : N + 1 \rightarrow N)$ such that for every other $(A, \alpha = [\alpha_0, \alpha_1] : A + 1 \rightarrow A)$ there is a unique $h : N \rightarrow A$ such that

$$\begin{array}{ccc} N + 1 & \xrightarrow{\iota} & N \\ \downarrow h + \text{id}_1 & & \downarrow h \\ A + 1 & \xrightarrow{\alpha} & A \end{array}$$

commutes. By Adámek's Theorem the NNO is the colimit of the ω -chain

$$0 \longrightarrow 1 \longrightarrow 1 + 1 \longrightarrow 1 + 1 + 1 \longrightarrow \cdots$$

So if $\mathcal{A} = \text{Set}$ the chain above becomes the sequence of natural numbers, where each $n \in \mathbb{N}$ is regarded as the set $\{0, \dots, n - 1\}$, and N will thus be exactly \mathbb{N} hence the name “natural numbers object”.

Example 3.1.9. On Pos , the category of partially ordered sets and order-preserving functions, consider the functor $FX = X_{\perp}$. Its initial-algebra ω -chain is

DIAGRAM HERE

So the n -th element of the ω -chain can be thought of as the set $n = \{0, \dots, n - 1\}$ with the natural order upon it while the arrows are inclusions. The colimit is \mathbb{N} with the standard ordering.

Definition 3.1.10. A function $f : M \rightarrow M'$ between metric spaces $(M, d_M), (M', d_{M'})$ is **non-expanding** if $d_{M'}(f(x), f(y)) \leq d_M(x, y)$ for all $x, y \in M$. A metric space (M, d_M) is 1-bounded if $d_M(x, y) \leq 1$ for all $x, y \in M$. We denote by **MS** the category of 1-bounded metric spaces and non-expanding maps.

non-expanding
map

MS

Remark 3.1.11. Coproducts in **MS** are disjoint unions where each summand retains its metric and points in different summands have distance 1. The product of $M, M' \in \mathbf{MS}$ is the cartesian product $M \times M'$ endowed with the maximum metric

$$d_{M \times M'}((x, x'), (y, y')) = \max\{d_M(x, y), d_{M'}(x', y')\}.$$

Moreover **MS** has a terminal object $1 = \{*\}$ i.e. the singleton with its unique metric.

Example 3.1.12. We can now consider $FX = X + 1$ on **MS**. The objects of the initial-algebra ω -chain are again the natural numbers $n = \{0, \dots, n-1\}$, but this time regarded as discrete spaces (that is: spaces where all points have distance 1) and the connecting arrows are the inclusions. Thus the colimit, and initial algebra of F , is \mathbb{N} endowed with the discrete metric.

Now consider the functor $FX = \frac{1}{2}X + 1$ where we denote by $\frac{1}{2}X$ the space obtained by scaling the metric of X by a factor of $\frac{1}{2}$. This functor preserves colimits of ω -chains by Remark 3.1.7, mutatis mutandis. Its initial-algebra ω -chain in the following sequence

diagram

where 0 always denotes the new point added by the functor. The metric on $F^n 0$ is

$$d(i, j) = \begin{cases} 0 & \text{if } i = j \\ 2^{-\min\{i, j\}} & \text{otherwise} \end{cases}$$

and the connecting morphisms are (non-expanding) inclusions. We then obtain that μG is again \mathbb{N} , but this time endowed with the metric d above.

In the previous examples we have seen how Adámek's Theorem allows us to look at “finite pieces” of the initial algebra. The next one shows that we can use it also to understand how the structure on the initial algebra works.

Example 3.1.13. Consider $FX = X \times X + 1$ on **Set**. This functor preserves colimits of ω -chains and we know from Example 2.2.12 that μF is the algebra of all finite binary trees. We will now apply Adámek's Theorem to obtain a “constructive” reason for why this is so.

If we denote the pair $(x, y) \in X \times X$ by the tree $[\text{TREE}]$ we have that

DIAGRAM

so $F^n 0$ is the set of binary trees of height less than n . The connecting arrows $F^n 0 \rightarrow F^{n+1} 0$ are inclusions and so the colimit of the chain, and thus the initial algebra, is the union $\bigcup_{n < \omega} F^n 0$ i.e. the set of all finite binary trees.

Recall (again from Example 2.2.12) that the structure $\iota : \mu F \times \mu F + 1 \rightarrow \mu F$ is given by tree-tupling. This information too can be obtained by Adámek's Theorem. Recall from

the proof that, if $(\mu F, c_n : F^n 0 \rightarrow \mu F)_{i < \omega}$ is the colimit of the initial-algebra ω -chain, then ι is the unique arrow such that

$$c_n = \iota \circ Fc_{n-1} \quad \text{for all } 0 < n < \omega \quad (3.1.7)$$

and, moreover, the c_n 's are inclusions. Now fix $0 < k < \omega$ and consider a tree $[TREE\ t_1, t_2]$ in $F^k 0$ that we also regard as the pair (t_1, t_2) for $t_1, t_2 \in F^{k-1} 0$. We have

$$\begin{aligned} c_k([TREE\ t_1, t_2]) &= [TREE\ t_1, t_2] \in \mu F \\ \iota \circ Fc_{k-1}((t_1, t_2)) &= \iota(\underbrace{(t_1, t_2)}_{\in \mu F \times \mu F + 1}) \end{aligned}$$

so by (3.1.7) ι is indeed tree-tupling.

Chapter 4

Initial algebras from transfinite iteration

TODO: maybe intro on cardinals? TODO: find better name for this section.

4.1 Colimits of chains

Through this chapter we regard ordinals $j \in \text{Ord}$ as linear orders and thus as categories. A j -chain in some category \mathcal{A} is a functor $C : j \rightarrow \mathcal{A}$; we denote the elements of the chain by C_i , for $i \in j$, and its arrows by $c_{i,i'} : C_i \rightarrow C_{i'}$ for $i \leq i' < j$. We say that \mathcal{A} has colimits of chains if for all $j \in \text{Ord}$ every j -chain has a colimit; this includes the existence of an initial object (as the colimit of the empty chain) but does not include colimits of Ord-chain.

j -chain

Lemma 4.1.1. Let j be an ordinal and $C : j \rightarrow \mathcal{A}$ a j -chain into a category \mathcal{A} with colimits of chains. Now let $j' \subseteq j$ be cofinal i.e. such that for every $i \in j$ there is some $i' \in j'$ such that $i \leq i'$ and denote by C' the restriction of C to j' . Then C and C' have the same colimit.

Proof. The first thing we notice is that there is a canonical bijection between cocones on C and cocones on C' . Given a cocone $(A, \alpha_i : C_i \rightarrow A)_{i \in j}$ on C throwing away all the indexes not in j' yields $(A, \alpha_i : C_i \rightarrow A)_{i \in j'}$: a cocone on C' . Conversely given $(A, \alpha_i : C_i \rightarrow A)_{i \in j'}$ we need to define $\alpha_i : C_i \rightarrow A$ for all $i \in j \setminus j'$; but there's only one way to do so. Indeed by cofinality let $i' \in j'$ be such that $i \leq i'$; then by the cocone condition we must set $\alpha_i = \alpha_{i'} \circ c_{i,i'}$ so $(A, \alpha_i : C_i \rightarrow A)_{i \in j'}$ can be extended to the whole chain in only one way.

Let $(L, \sigma_i : C_i \rightarrow L)_{i \in j}$ be the colimit of C ; then $(L, \sigma_i : C_i \rightarrow L)_{i \in j'}$ is the colimit of C' . Let $(A, \alpha_i : C_i \rightarrow A)_{i \in j'}$ be a cocone on C' that we extend to a cocone on C as explained above. Now, since L is the colimit, let $m : L \rightarrow A$ be the unique factorization arrow such that

$$m \circ \sigma_i = \alpha_i \quad \text{for all } i \in j.$$

Particularly for all $i \in j'$ so m is a factorization of $(A, \alpha_i : C_i \rightarrow A)_{i \in j'}$ through $(L, \sigma_i : C_i \rightarrow L)_{i \in j'}$. To prove it is unique let $n : L \rightarrow A$ be such that $n \circ \sigma_i = \alpha_i$ for all $i \in j'$; we will show that n actually works for all $i \in j$ and thus must be m by uniqueness. Let $i \in j \setminus j'$ and $i' \in j'$ such that $i \leq i'$ by cofinality, then we have

$$n \circ \sigma_i = n \circ \sigma_{i'} \circ c_{i,i'} = \alpha_{i'} \circ c_{i,i'} = \alpha_i.$$

For the converse let $(L, \sigma_i : C_i \rightarrow L)_{i \in j'}$ be the colimit of C' . One can prove that the unique extension $(L, \sigma_i : C_i \rightarrow L)_{i \in j}$ is the colimit of C by a similar argument. Finally we have that C and C' have the same colimit (modulo restriction/extension) and that the factorization arrows are the same. \square

Thanks to this lemma we can safely calculate colimits of chains, or apply Proposition 1.1.8, on convenient cofinal subchains instead of on the full chain. This allows us to “skip” limit ordinals because given j limit the subset $\{i \in j : i \text{ successor}\}$ is cofinal in j .

4.2 Transfinite iteration

The main result of this section is again a generalization to the categorical setting of an order-theoretic result; this time from Zermelo. Let P be a chain-complete poset and $f : P \rightarrow P$ an order-preserving map. We can define an Ord-indexed sequence by transfinite recursion as follows

$$\begin{aligned} f^0(\perp) &= \perp && \text{base case;} \\ f^{j+1}(\perp) &= f(f^j(\perp)) && \text{for all ordinals } j; \\ f^j(\perp) &= \bigvee_{i < j} f^i(\perp) && \text{for limit ordinals } j. \end{aligned}$$

This sequence is clearly a chain.

Theorem 4.2.1 (Zermelo). Let P be a chain-complete poset. Then every order-preserving $f : P \rightarrow P$ has a lest fixed point μf and $\mu f = f^j(\perp)$ for some $j \in \text{Ord}$.

Proof. By Hartog's Lemma (insert here a reference) there is an ordinal i such that there is no injection $i \rightarrow P$. Thus there are $j < k < i$ such that $f^j(\perp) = f^k(\perp)$ and we have $f^j(\perp) \leq f^{j+1}(\perp) \leq f^k(\perp)$ because $(f^j(\perp))_{j \in \text{Ord}}$ is a chain. So $f(f^j(\perp)) = f^{j+1}(\perp) = f^j(\perp)$ hence $f^j(\perp)$ is a fixed point.

Now let $x = f(x)$ be another fixed point. By transfinite induction we can prove $x \geq f^i(\perp)$ for all $i \in \text{Ord}$ (similarly to what we did in Theorem 3.1.1) thus when $i = j$ we obtain $x \geq f^j(\perp)$. This proves that $f^j(\perp)$ is indeed the least fixed point of f . \square

In order to generalize this theorem and its proof we begin by defining a transfinite version of the initial-algebra ω -chain. This will be a functor $W : \text{Ord} \rightarrow \mathcal{A}$ into a category \mathcal{A} with colimits of chains. It is defined by transfinite recursion as follows.

$$\begin{aligned} W_0 &= 0 && \text{base case;} \\ W_{j+1} &= FW_j && \text{for all ordinals } j; \\ W_j &= \text{colim}_{i < j} W_i && \text{for all limit ordinals } j. \end{aligned}$$

As for arrows $w_{0,1}$ is the unique arrow $0 \rightarrow W_1 = F0$ and $w_{i+1,j+1} = Fw_{i,j}$ for all $i, j \in \text{Ord}$. When j is limit we have $W_j = \text{colim}_{i < j} W_i$ so we let $\{w_{i,j}\}_{i < j}$ be the arrows forming the colimit cocone. We are only missing $w_{j,j+1}$ for j limit. Such arrows are obtained using the universal property of W_j . First notice that we can leave out the initial object 0 and all the limit ordinals without changing the colimit i.e.

$$W_j = \text{colim}_{i < j} W_i = \text{colim}_{i < j \text{ successor}} W_i$$

by Lemma 4.1.1. Now $(Fw_{i,j} : FW_i = W_{i+1} \rightarrow FW_j = W_{j+1})_{i < j}$ is a cocone on this reduced diagram. Thus we obtain a unique factorization arrow $W_j \rightarrow W_{j+1}$ that we declare to be our $w_{j,j+1}$ so the definition of our transfinite chain can continue. Notice also that, being the factorization arrow, $w_{j,j+1}$ is such that

$$w_{j,j+1} \circ w_{i+1,j} = Fw_{i,j} \quad \text{for all } i < j. \quad (4.2.1)$$

We now generalize Remark 3.1.3 to the transfinite case.

Remark 4.2.2. Let (A, α) be an F -algebra; then it induces a cocone $(\alpha_i : W_i \rightarrow A)_{i \in \text{Ord}}$ on W as follows.

$$\begin{aligned} \alpha_0 &= ! && \text{base case;} \\ \alpha_{i+1} &= \alpha \circ F\alpha_i && \text{for all ordinals } i; \end{aligned}$$

for limit ordinals j we define $\alpha_j : W_j \rightarrow A$ to be the unique factorization arrow such that $\alpha_j \circ w_{i,j} = \alpha_i$ for all $i < j$. In order for this construction to work we need to prove that, at stage j , $(\alpha_i : W_i \rightarrow A)_{i < j}$ is a cocone i.e. that

$$\alpha_i \circ w_{k,i} = \alpha_k \quad (4.2.2)$$

for all $k < i < j$. Assuming that all α_i 's for $i < j$ have been constructed we have that:

1. for $w_{0,1}$ condition (4.2.2) is trivial;
2. for $w_{k,i}$ with i limit it is trivial by construction of α_i ;
3. for $w_{i,i+1}$ with i limit we can prove (4.2.2) holds by showing

$$\alpha_{i+1} \circ w_{i,i+1} \circ w_{k,i} = \alpha_i \circ w_{k,i} = \alpha_k$$

for all $k < i$ and conclude by uniqueness of α_i . It is enough to prove the relation for all $k < i$ successors by Lemma 4.1.1 and indeed we have

$$\begin{aligned} \alpha_{i+1} \circ w_{i,i+1} \circ w_{k,i} &= \alpha \circ F\alpha_i \circ w_{i,i+1} \circ w_{k,i} && \text{definition of } \alpha_{i+1} \\ &= \alpha \circ F\alpha_i \circ Fw_{k-1,i} && (4.2.1) \\ &= \alpha \circ F(\alpha_i \circ w_{k-1,i}) && \text{functoriality of } F \\ &= \alpha \circ F\alpha_{k-1} && \text{definition of } \alpha_i \\ &= \alpha_k && \text{definition of } \alpha_k. \end{aligned}$$

4. for $w_{i,i+1}$ with i non-limit we proceed inductively like in Remark 3.1.3; the base case is provided by item 1 if i is a natural number and by item 3 otherwise.

Finally we conclude that $(\alpha_i : W_i \rightarrow A)_{i \in \text{Ord}}$ is well defined and a cocone on W since, as we just proved, for every j limit $(\alpha_i : W_i \rightarrow A)_{i < j}$ is a cocone.

Definition 4.2.3. We say that the initial-algebra chain **converges** in λ steps if $w_{\lambda, \lambda+1}$ is an isomorphism and that it converges in **exactly** λ steps when λ is the least such ordinal.

converge of the
initial-algebra
chain

We are now ready to state and prove the transfinite version of Adámek's Theorem.

Theorem 4.2.4. Let \mathcal{A} be a category with colimits of chains and F an endofunctor. If the initial-algebra chain of F converges in j steps then W_j is the initial algebra μF with structure given by

$$(w_{j,j+1})^{-1} : W_{j+1} = FW_j \rightarrow W_j.$$

Proof. We only need to prove that W_j with the algebra structure given above is initial. To this end let (A, α) be an F -algebra and let h be the arrow $\alpha_j : W_j \rightarrow A$ of the cocone induced by (A, α) in the sense of Remark 4.2.2. We shall prove that h is the unique F -algebra homomorphism from W_j to A .

$$\begin{array}{ccc}
 W_{j+1} = FW_j & \xrightarrow{(w_{j,j+1})^{-1}} & W_j \\
 \downarrow Fh = F\alpha_j & \searrow \alpha_{j+1} & \downarrow h = \alpha_j \\
 FA & \xrightarrow{\alpha} & A
 \end{array}$$

In the above diagram the lower left triangle commute by construction of α_{j+1} and the upper right one because the α_i 's form a cocone. Thus the outer square commutes i.e. h is an homomorphism.

To prove that h is unique assume that $k : W_j \rightarrow A$ is another homomorphism. We will show by transfinite induction that $h \circ w_{i,j} = k \circ w_{i,j}$ for all $i \leq j$ and then conclude by choosing $i = j$. Notice that $h \circ w_{i,j} = \alpha_j \circ w_{i,j} = \alpha_i$ so we are really interested in proving

$$\alpha_i = k \circ w_{i,j} \quad (4.2.3)$$

for all $i \leq j$. If $i = 0$ then (4.2.3) is obvious because both sides of the equality are arrows out of the initial object. If we have $\alpha_i = k \circ w_{i,j}$ then

$$\begin{aligned}
 \alpha_{i+1} &= \alpha \circ F\alpha_i && \text{definition of } \alpha_i \\
 &= \alpha \circ F(k \circ w_{i,j}) && \text{inductive hypothesis} \\
 &= \alpha \circ Fk \circ w_{i+1,j+1} && \text{factoriality of } F \\
 &= k \circ (w_{j,j+1})^{-1} \circ w_{i+1,j+1} && k \text{ homomorphism} \\
 &= k \circ w_{i+1,j}.
 \end{aligned}$$

Finally if i is limit we know α_i is the unique arrow such that $\alpha_i \circ w_{s,i} = \alpha_s$ for all $s < i$; so if we show that $k \circ w_{i,j}$ does the same then we have (4.2.3). Indeed for $s < i$ we have

$$\begin{aligned}
 k \circ w_{i,j} \circ w_{s,i} &= k \circ w_{s,j} \\
 &= h \circ w_{s,j} && \text{inductive hypothesis since } s < i \\
 &= \alpha_j \circ w_{s,j} \\
 &= \alpha_s && \text{the } \alpha_i \text{'s are a cocone.}
 \end{aligned}$$

This concludes the proof. □

Bibliography

- [Bor08] Francis Borceux. *Handbook of Categorical Algebra: Volume 1, Basic Category Theory*. 1st ed. Cambridge University Press, 2008.
- [And19] Alessandro Andretta. *Elements of Mathematical Logic*. July 16, 2019. URL: http://torino.logicgroup.altervista.org/torino/materiale/Andretta_Elements_20190716.pdf.
- [Jř21] Lawrence S. Moss Jiří Adámekm Stefan Milius. *Initial Algebras, Terminal Coalgebras, and the Theory of Fixed Points of Functors*. Aug. 27, 2021. URL: <https://www8.cs.fau.de/ext/milius/publications/files/CoalgebraBook.pdf>.