Category Theory

Davi Sales Barreira

March 12, 2022

Contents

List of Definitions

List of Theorems

List of Examples

Notes mostly based on Ribeiro [6], Bradley et al. [2], Borceux [1].

1 Preface

The study of Category Theory enables us to view Mathematics from a vantage point, and better understand how the different areas are connected. By looking at the subject from the distance (via Category Theory), we get a glimpse at the connections (and disconnections) between different fields.

Another interesting observation about Category Theory is that it's breaking the theoretical barrier and it's starting to be applied in real world applications. One prominent example is in programming, as highlighted by the book Milewski [5]. The book by Fong and Spivak [3] also focus on the application side of Category Theory.

For those interested in applying Category Theory, Julia has a very prominent package called Catlab.jl, which will be shown in these notes.

These notes are intended to serve as a quick introduction to people interested in applying Category Theory. Thus, instead of focusing on proving theorems or presenting a plethora of examples in mathematics, the goal is to rigorously (and quickly) introduce the field together with some intuition and useful examples.

The notes start by introducing the **big 5** of Category Theory, namely, categories, functors, natural transformations, limits/colimits and adjoints. After that, we focus on some applications of Category Theory, mainly it's application to data structures / data bases.

2 Categories

In this section, we formally define what a category is, and we provide some examples.

2.1 Universes, Sets and Classes

When working on Category Theory, it's common to find universal statements such as "for all topological spaces...". The issue with such statements is that, in a purely set-theoretical sense, we have to know whether such large collection ("all topological spaces") is indeed a set. We might be tempted to say that's true, but it's not so simple. The most known example of a possible failure of such which the answer is "no".

statements is Russel's Paradox of whether there is a set of all sets, for Therefore, in order to deal with such issue, we need a way to differentiate between a valid set and an arbitrary collection. Here is where the notion of a Universe comes in.

Definition 2.1 (Universe). We say that a set $\mathfrak{U}$ is a universe if[1]

(i) $x \in y$ and $y \in \mathfrak{U}$, then $x \in \mathfrak{U}$;

(ii) $I \in \mathfrak{U}$, and $\forall i \in I, x_i \in \mathfrak{U}$, then $\cup_{i \in I} x_i \in \mathfrak{U}$;

(iii) $x \in \mathfrak{U}$ then $\mathcal{P}(x) \in \mathfrak{U}$, where $\mathcal{P}(x)$ is the power set;

(iv) $x \in \mathfrak{U}$ and $f : x \to y$ is a surjective function, then $y \in \mathfrak{U}$;

(v) $\mathbb{N} \in \mathfrak{U}$.

From this definition, one can prove the following proposition.

Proposition 2.2. (i) $x \in \mathfrak{U}$ and $y \subset x$, then $y \in \mathfrak{U}$;

(ii) $x \in \mathfrak{U}$ and $y \subset x$, then $\{x, y\} \in \mathfrak{U}$;

(iii) $x \in \mathfrak{U}$ and $y \subset x$, then $x \times y \in \mathfrak{U}$;

(iv) $x \in \mathfrak{U}$ and $y \subset x$, then $y^x \in \mathfrak{U}$, where y^x is the set of functions $f : x \to y$.

With this definition, we state the axiom of existence universes.

Axiom 2.1. Every set S belongs to some universe $\mathfrak{U}$.

Definition 2.3. For a fixed universe $\mathfrak{U}$, if a set S is an element of $\mathfrak{U}$, then S is called a *small set*.

Talking about "small sets" and "big set" might become daunting, so instead, we use a different convention which is based on Gödel-Bernays theory of sets and classes. This theory states that:

Axiom 2.2 (Gödel-Bernays). Every set is a class, and a class is a set if and only if it belongs to some (other) class.

Note that using the notion of Universes, we can recover Gödel-Bernays theory. For that, use the following definition:

Definition 2.4. For a fixed universe $\mathfrak{U}$, we call S a *set* if it's an element of $\mathfrak{U}$, and call S a *class* if it's a subset of $\mathfrak{U}$. A class that is not a set is called a *proper class*.

[1] Definition from Borceux [1]

Since every set is a class, if $S \in \mathfrak{U}$, then S is a class, since U is a set and therefore a class, implying that S belongs to a class. On the other direction, if S is a class and $S \in V \subset \mathfrak{U}$, then since $V \subset \mathfrak{U}$, this means that $S \in \mathfrak{U}$, thus it's a set.

From now on, whenever we say *set* we are implying *small set* and whenever we say *class* we are implying either small or big sets, following Borceux [1] convention.

2.2 What is a Category?

Let's formally define a category and provide some examples.

Definition 2.5 (Category). A category $\mathcal{C} = \langle Ob_{\mathcal{C}}, Mor_{\mathcal{C}} \rangle$ consists of a class of objects $Ob_{\mathcal{C}}$ and a class of morphisms $Mor_{\mathcal{C}}$ satisfying the following conditions:

(i) Every morphism $f \in Mor_{\mathcal{C}}$ is associated to two objects $X, Y \in Ob_{\mathcal{C}}$ which is represented by $f : X \to Y$ or $X \xrightarrow{f} Y$, where $dom(f) = X$ is called the domain of f and $cod(f) = Y$ is the codomain. Moreover, we define $Mor_{\mathcal{C}}(X, Y)$ as

$$Mor_{\mathcal{C}}(X, Y) := \{f \in Mor_{\mathcal{C}} \ : \ X \in dom(f), \ Y \in cod(f)\};$$

(ii) For any three objects $X, Y, Z \in Ob_{\mathcal{C}}$, there exists a composition operator

$$\circ : Mor_{\mathcal{C}}(X, Y) \times Mor_{\mathcal{C}}(Y, Z) \to Mor_{\mathcal{C}}(X, Z),$$

(iii) For each object $X \in Ob_{\mathcal{C}}$ there exists a morfism $id_X \in Mor_{\mathcal{C}}(A, A)$ called the identity.

The composition operator must have the following properties:

(p.1) *Associative*: for every $f \in Mor_{\mathcal{C}}(A, B), g \in Mor_{\mathcal{C}}(B, C), h \in Mor_{\mathcal{C}}(C, D)$ then

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

(p.2) For any $f \in Mor_{\mathcal{C}}(X, Y)$, $g \in Mor_{\mathcal{C}}(Y, X)$,

$$f \circ id_X = f, \quad id_Y \circ g = g.$$

There are many ways to refers to the class of morphisms $Mor_{\mathcal{C}}(X, Y)$, such as $\mathcal{C}(X, Y)$ or $\hom_{\mathcal{C}}(X, Y)$. The reason for this is that this set is sometimes called hom-set. In this notes, we'll use either $Mor_{\mathcal{C}}(X, Y)$ or $\mathcal{C}(X, Y)$ when there is no ambiguity. Also, we'll use dom_f to mean $dom(f)$, and similarly for the codomain.

Another point about conventions. When talking about composition, it's convenient to use the operator $\fatsemi$, which is equivalent to the composition $\circ$, but with the inverted order, i.e.

$f \fatsemi g = g \circ f$. The convinience will become clearer once we introduce Hasse diagrams as a way to represent categories.

When the class of morphism $Mor_{\mathcal{C}}$ is a set, the category $\mathcal{C}$ is called a *locally small category*. If both $Ob_{\mathcal{C}}$ and $Mor_{\mathcal{C}}$ are sets, we then have a *small category*.

Finally, whenever it's not ambiguous, we might drop the subscript and use Ob to refer to the objects of $\mathcal{C}$ and Mor to refer to the morphisms of $\mathcal{C}$.

2.3 Examples of Categories

It's very common to represent categories via Hasse Diagrams. In these diagrams, the objects are represented as dots, and the morphisms as arrows. Let's show some examples.

Example 2.1 (Category 1). The category $\mathbf{1}$ consists of $Ob_{\mathbf{1}} := \{A\}$ and $Mor_{\mathbf{1}} = \{id_A\}$. The diagram for such category is shown below.

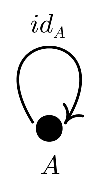

Figure 1: Hasse diagram of category **1**.

Example 2.2 (Category 2 and 3). The category $\mathbf{2}$ consists of $Ob_{\mathbf{2}} := \{A, B\}$ and $Mor_{\mathbf{1}} = \{id_A, id_B, f\}$, where $f : A \to B$. The diagram for such category is shown below.

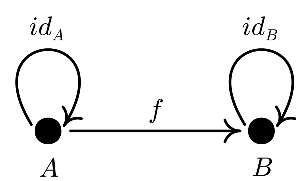

Figure 2: Hasse diagram of category **2**.

Since we know that identities are always present in categories, we'll omit them from future diagrams when there is no ambiguity. Thus, the figure below represents the same diagram as Figure 2.

$$
\begin{tikzcd}
A \arrow[r, "f"] & B
\end{tikzcd}
$$

Figure 3: Hasse diagram of category **2** omitting identity morphisms.

The category **3** has three morphisms besides the identities, given by f, g and their composition $f \fatsemi g$. The figure below illustrates the category with all it's morphisms.

$$
\begin{tikzcd}
A \arrow[loop above, "id_A"] \arrow[r, "f"] \arrow[rr, bend right, "f \fatsemi g"'] & B \arrow[loop above, "id_B"] \arrow[r, "g"] & C \arrow[loop above, "id_C"]
\end{tikzcd}
$$

Figure 4: Hasse diagram of category **3** showing all morphisms.

Drawing all the morphisms can make the diagram become too crowded, specially as the number of objects and morphisms grows. Hence, we simplify the diagram representation by ommiting not only the identity morphisms, but also the compositions. These can always be assumed to exist, since they are a necessery condition for every category. Thus, the figure below represents the same diagram as Figure 4.

$$
\begin{tikzcd}
A \arrow[r, "f"] & B \arrow[r, "g"] & C
\end{tikzcd}
$$

Figure 5: Hasse diagram of category **3** omitting identities and compositions.

Example 2.3 (Preorders). A preorder is defined by a tuple $(P, \leq)$, where P is a set of values, such that

(i) For $a, b \in P$, if $a \leq b$ and $b \leq c$, then $a \leq c$;

(ii) For every $a \in P$, $a \leq a$.

We can show that actually, this is a category, which we'll call $\mathfrak{P}$, where $Ob_{\mathfrak{P}} = P$ and each morphism f represents $a \leq b$, where $cod_f = a$ and $dom_f = b$. One example of preorder is the set of $\mathbb{N}$ equiped with the binary relation $\leq$ which is shown in the diagram below.

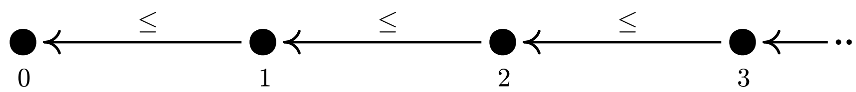

Figure 6: Hasse diagram of preorder category of natural numbers.

Note that in preorders, there is at most one morphism between each pair of objects. Thus, categories with such property are often referred as *thin categories* or *preorder category* (in Fong and Spivak [3], the authors call this a *preorder reflection*).

Example 2.4 (Monoids). A monoid is a triple $(M, \oplus, e_M)$ where $\oplus : M \times M \to M$ is a binary operation and e_M the neutral element, such that:

1. $a \oplus (b \oplus c) = (a \oplus b) \oplus c$
2. $a \oplus e_M = e_M \oplus a = a.$

Note that $(\mathbb{N} \cup \{0\}, +, 0)$ is a monoid.

Moreover, each single monoid can be defined as a category itself. For monoid $(M, \oplus, e_M)$, define a category $\mathcal{C}$ such that $Ob_{\mathcal{C}} := \{M\}$, and the set of morphisms are the elements of M, i.e. $Mor_{\mathcal{C}} := \{a \in M\}$. Finally, we define the composition operartion as $a \circ b := a \oplus b$ Thus, $(\mathbb{N} \cup \{0\}, +, 0)$ is the category illustrated in Figure 7.

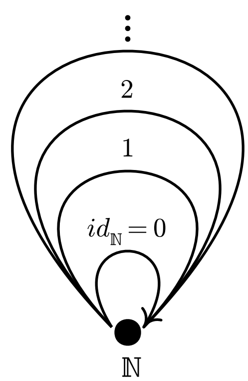

Figure 7: Hasse diagram of monoid category of natural numbers.

There are many other examples:

1. **Set** which is the category of sets, where the objects are sets and the morphisms are functions between sets.

2. **Top** is the category where topological spaces are the objects and continuous functions are the morphisms.

3. $\mathbf{Vec}_{\mathbb{F}}$ is the category where vector spaces over field $\mathbb{F}$ are the objects, and linear transformations are the morphisms.

4. **Gr** is the category of directed graphs, where $Ob_{\mathbf{Gr}} := \{\text{Vertex}, \text{Arrow}\}$, and the morphisms are
$$Mor_{\mathbf{Gr}} := \{src, tgt, id_{\text{Vertex}}, id_{\text{Arrow}}\}$$
where $src : \text{Arrow} \to \text{Vertex}$ returns the source vertex for each arrow and $tgt : \text{Arrow} \to \text{Vertex}$ returns the target vertex.

2.4 Programming with Category Theory

One might be surprised to find out that Category Theory, although very abstract in nature, has actual applications in the "real world". A very interesting example of this is in programming.

In programming languages such as Julia, we can think of 'Types' as objects and functions as morphisms.

2.5 Isomorphism, monomorphism and epimorphism

A very important definition in Category Theory is the notion of isomorphism. In Set Theory, we say that two sets are isomorphic if there is an invertible function between them. Yet, this concept is not restricted to Set Theory and can be generalized in Category Theory as follows:

Definition 2.6 (Categorical Isomorphism). Let $\mathcal{C}$ be a category with $X, Y \in Ob_{\mathcal{C}}$ and $f \in Mor_{\mathcal{C}}(X, Y)$.

(i) We say that f is *left invertible* if there exists $f_l \in Mor_{\mathcal{C}}(Y, X)$ such that $f_l \circ f = id_X$;

(ii) We say that f is *right invertible* if there exists $f_r \in Mor_{\mathcal{C}}(Y, X)$ such that $f \circ f_r = id_Y$;

(iii) We say that f is invertible if it's both left and right invertible.

When an invertible morphism exists between X and Y, we say that they are isomorphic.

Proposition 2.7. The following properties on inverses are true:

1. If f is an invertible morphism, then the left and right inverses are the same.

2. If f and g are invertible and composable, then $f \circ g$ is also invertible.

Proof. 1. Let f be invertible with left inverse f_l and right inverse f_r. Therefore,

$$f_l \circ id_Y = f_l \circ f \circ f_r = id_X \circ f_r = f_r.$$

2. Let $f : A \to B$ and $g : B \to C$ be invertible and composable, with $f \circ g : A \to C$. There exists the inverses $g^{-1} : C \to B$ and $f^{-1} : B \to A$. Note that $f^{-1} \circ g^{-1} : C \to A$, thus

$$(f^{-1} \circ g^{-1}) \circ (g \circ f) = f^{-1} \circ (g^{-1} \circ g) \circ f = (f^{-1} \circ id_B) \circ f = f^{-1} \circ f = id_A.$$

$$(g \circ f) \circ (f^{-1} \circ g^{-1}) = g \circ (f \circ f^{-1}) \circ g^{-1} = g \circ (id_B \circ g^{-1}) = g \circ g^{-1} = id_B.$$

We conclude that $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

□

Although similar to set isomorphism, categorical isomorphism is in a sense more general, and captures our intuition of isomorphism between categories better than the set theoretic case, even when we have finite objects.

Consider the following example. Let $\mathcal{P}$ be the category of Posets, where posets $(P, \leq_p)$ are the objects. Take two objects $P_1, P_2 \in Ob_{\mathcal{P}}$, where $P_1 := \{a, b\}$ with a and b **not** comparable, and $P_2 := \{0, 1\}$ where indeed $0 \leq 1$. The question is whether P_1 is "actually" isomorphic to P_2, and our intuition say that they should not be, since P_1 has two incomparable elements while P_2 has two comparable elements.

If we use the set theoretic definition, we would conclude that they **are** isomorphic, since there is a bijective function between P_1 and P_2. Take for example $f : P_1 \to P_2$ where $f(a) = 0$ and $f(b) = 1$. So the set theoretic isomorphism does not capture what we want. What about the categorical isomorphism? We can prove that this will not be an isomorphism using the categorical definition. Yet, in order to prove this, we need to specify what are the morphisms between the posets, and to do this, we need to define what are functors.

In the same way that set isomorphism is not the same as categorical isomorphism, the notions of injectivity and surjectivity are not equivalent to their categorical counterparts, which are called monomorphism and epimorphism.

Definition 2.8 (Monomorphism). Let $\mathcal{C}$ be a category and $f \in Mor_{\mathcal{C}}(A, B)$. We say that f is a monomorphism (or monic), if

$$f \circ g = f \circ h \implies g = h.$$

Definition 2.9 (Epimorphism). Let $\mathcal{C}$ be a category and $f \in Mor_{\mathcal{C}}(A, B)$. We say that f is an epimorphism (or epic), if

$$g \circ f = h \circ f \implies g = h.$$

Important! A morphism f can be both epic and monic, without being an isomorphism, which again highlights the difference between this concepts and their set-theoretic counterparts.

Proposition 2.10. The following properties on monomorphism and epimorphism are true:

1. f left-invertible $\implies$ f is monic. The converse are is not true.
2. f right-invertible $\implies$ f is epic. The converse are is not true.
3. f invertible $\implies$ f is monic and epic. The converse are is not true.
4. f monic and right-invertible $\implies$ f is isomorphism.
5. f epic and left-invertible $\implies$ f is isomorphism.

Proof. 1. Note $f : A \to B$ left-invertible implies that there exists a $f_l : B \to A$ such that $f_l \circ f = id$. Hence, for a $g : B \to C$ and $h : B \to C$, if

$$f \circ g = f \circ h,$$

then we have that

$$f_l \circ f \circ g = f_l \circ f \circ h \implies g = h.$$

To show that the converse is false, consider the category **2**(Figure 2). Note that $f : A \to B$ is monic, since the only morphism that composes with f is id_A and id_B. Yet, note that f is not left invertible, since there isn't even a morphism from B to A.

2. Use the same argument, but reversing the order of the compositions. For the converse, again consider the same category **2**. Note that $f : A \to B$ is epic, but it's not right invertible.

3. True since invertible means left and right invertible.

4. Since $f : A \to B$ right invertible, then there exists $f_r : B \to A$ such that $f \circ f_r = id_B$. Thus,

$$id_B \circ f = (f \circ f_r) \circ f = f \circ (f_r \circ f) = f \circ id_A \implies f_r \circ f = id_A.$$

5. Same argument. $\square$

2.6 Zero, Initial and Terminal Objects

Definition 2.11 (Zero, Initial and Terminal). Let $\mathcal{C}$ be a category.

1. An object $I \in Ob_{\mathcal{C}}$ is *initial* if for every $A \in Ob_{\mathcal{C}}$, there is exactly one morphism from I to A. Thus, from I to I there is only the identity.

2. An object $T \in Ob_{\mathcal{C}}$ is *terminal* if for every $A \in Ob_{\mathcal{C}}$, there is exactly one morphism from A to T. Thus, from I to I there is only the identity.

3. An object is *zero* if it's both terminal and initial.

Note that in the definitions above, we are defining these objects in terms of existence and uniqueness of morphisms, which is known in category theory as **universal constructions** (more on this later).

Theorem 2.12. Every *initial* object is unique up to an isomorphism, i.e. if in a category there are two *initial* objects, then they are isomorphic. Similarly, *terminal* objects are unique up to an isomorphism.

Proof. Let I_1, I_2 be initial. Then, there exists only $f : I_1 \to I_2$ and $g : I_2 \to I_1$. But since $g \circ f : I_1 \to I_1$ is a morphism from the initial object I_1, it must be equal to id_{I_1}. The same for I_2, which implies that f and g are inverses, and thus the objects are isomorphic. The same proof works for terminal objects. □

Example 2.5 (Terminal and Initial Objects in Set). Without thinking too much, one might assumet that in the category **Set** the empty set would be a zero object; but that's not true. In reality, the empty set is the initial object, since $f : \emptyset \to B$ is the only function from the empty set to any other set. Why is this valid?

Remember that in set theory, a function from two sets is defined as a binary relation such that for every $x \in dom_f$, there is a unique $y \in cod_f$, i.e. f is a triple (A, B, G), where $A = dom_f$, $B = cod_f$ and $G \subset A \times B$ such that $\forall x \in dom_f$, there exists a unique $y \in B$, such that $(x, y) \in G$.

Since $dom_f = \emptyset$, we have that $G \subset \emptyset \times B$, but this is actually empty. Why? If $\emptyset \times B$ is not empty, then there exists $(a, b) \in \times B$, which is false, since this would imply that $a \in \emptyset$, which contradicts the definition of the empty set that says that it can have no elements (note that $\emptyset \in \emptyset$ is actually false).

With this, we have that $G = \emptyset$, thus, the only possible function from $\emptyset$ to B is $f = (\emptyset, \emptyset, B)$. Which proves that the empty set is initial.

But what about terminal? The empty set actually does not have any morphisms that arrives on it, since there is no function $f : A \to \emptyset$. The terminal sets in **Set** are actually all the singletons (sets with only one element), since for any $\{a\}$, there will be only one function $g : A \to \{a\}$, which is $g(x) = a$.

Another definition we have is that of a *zero* morphism. The idea here is that this morphism must take the elements of an object A to the zero element in B, for example, a for two vector spaces $\mathbb{R}^n$ and $\mathbb{R}^m$, the zero linear transformation $z : \mathbb{R}^n \to \mathbb{R}^m$ should take every vector n-dimensional vector to the $\mathbf{0}$ m-dimensional vector. In Category Theory we do not talk about morphisms according to how they act on the elements, but only in the objects. So we cannot define z by saying to which element it maps. Yet, there is a way to do this in Category Theory, which gives rise to the *zero* morphism definition.

Definition 2.13 (Zero Morphism). Let $\mathcal{C}$ be a category, and 0 be a zero object. A morphism $z : A \to B$ is a zero morphism if there exists two morphisms $f : A \to 0$ and $g : 0 \to B$, such that

$$z = g \circ f.$$

See that this makes intuitive sense. In our example, since we wish to take $v \in \mathbb{R}^n$ to $0 \in \mathbb{R}^m$, we first take all v to the zero object, which in the category of vector spaces will be the zero vector space, i.e. $\{0\}$ the space where $0 \in \mathbb{R}$ is the only element. So now all v are 0. Note that every linear transformation from $\{0\}$ to $\mathbb{R}^m$ must take 0 to $\mathbf{0} \in \mathbb{R}^m$, otherwise, suppose that $g(0) = \mathbf{v} \neq \mathbf{0}$, hence for a scalar α,

$$g(0) = g(\alpha 0) = \mathbf{v} \neq \alpha \mathbf{v} = \alpha g(0).$$

This is a contradiction, since g is a linear transformation.

Theorem 2.14. Let $\mathcal{C}$ be a catgory with zero object 0. Then there exists a unique zero morphism between any two objects.

Proof. Let $A, B \in Ob_{\mathcal{C}}$. By the definition of the zero object, there exists a unique $f : A \to 0$ and $g : 0 \to B$, thus, $g \circ f$ is a zero morphism by definition and is unique, since there is no other f and g with these respective domain and codamain.

Moreover, note that if $z : A \to B$ is our zero morphism and $h : B \to C$, then

$$h \circ z = h \circ (g \circ f) = (h \circ g) \circ f.$$

But, $l = (h \circ g) : 0 \to C$, which means that $l \circ f$ is a zero morphism. The same argument works with a composition from the other direction. This means that compositions with zero morphisms return zero morphisms. $\square$

2.7 Understanding Duality

In several fields of Mathematics, one is faced with the informal notion of a dual. Mathematicians define a concept, and call them the dual in some sense, for example, the dual vector space, the dual of an optimization problem, and many more. I always found puzzling what exactly held these things together, i.e. what was the underlying principle that made something a dual of another.

Fortunately, Category Theory has a very elegant answer. For a given category $\mathcal{C}$, the dual category is denoted by $\mathcal{C}^{op}$ where are the objects are the same, but the morphisms are inverted. This means that $Ob_{\mathcal{C}} = Ob_{\mathcal{C}^{op}}$, and for every $f \in Mor_(A, B)$, we have $f^{op} \in Mor_{\mathcal{C}^{op}}(B, A)$.

This definition gives rise to a very interesting result (observation), called the *Duality Principle*.

Definition 2.15 (Dual Property and Dual Statement). We say p^{op} is the dual property for p if for all categories

$$\mathcal{C} \text{ has } p^{op} \iff \mathcal{C}^{op} \text{ has } p.$$

For a statement s about a category $\mathcal{C}$, the dual statement is the same statement, but with regards to $\mathcal{C}^{op}$.

For example, "a category has an initial object if and only if the dual category has a terminal object". In this example, the property of having an initial object is the dual property of having a terminal object, since the above statement is true for any category. What about the dual statement? The dual for the statement "the category $\mathcal{C}$ has an initial object" is "the category $\mathcal{C}^{op}$ has an initial object". Note that the dual statement is not always true. And here is where we get the duality principle.

Definition 2.16 (Duality Principle). If a statement s is true for every category, then the dual statement is also true for every category.

Let's digest a bit what this principle states. If we can prove that a certain statement is true for any arbitrary category, then it's dual will also be true without any effort what so ever. Roman et al. [7] gives a nice example of this. We already prove that for any category, if an initial object exists, this initial object is unique up to an isomorphism. Note that this is a statement that is true for any category, so the duality principle applies, i.e. the dual statement is true. And what is the dual statement? That for every $\mathcal{C}^{op}$ the initial object is unique up to an isomorphism. But an initial object in $\mathcal{C}^{op}$ is a terminal object in $\mathcal{C}$. So we have, without any effort, that every terminal object is unique up to an isomorphism.

2.8 Categorical Product and Coproduct

In set theory, we are used to the notion of a Cartesian product. Similarly to how we did for isomorphism, the idea of a product can be generalized via Category Theory. Here is how it's done.

Definition 2.17 (Span). Let A, B be objects in a category $\mathcal{C}$. A span on A and B is a triple (Z, f, g) where $f : Z \to A$ and $g : Z \to B$ are morphisms in $\mathcal{C}$. This is shown in figure 8.

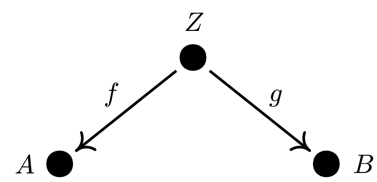

Figure 8: Diagrams showcasing a span between A and B.

Definition 2.18 (Categorical Product). Let A, B be objects in a category $\mathcal{C}$. A span $(A \times B, \pi_1, \pi_2)$ is called a product between A and B if for every span (Z, f, g) of A and B, there exists a unique morphism $h_{f,g} : Z \to A \times B$ such that $h_{f,g} \mathbin{⨾} \pi_1 = f$ and $h_{f,g} \mathbin{⨾} \pi_2 = g$. That is the same as saying that the diagram 9 commutes.

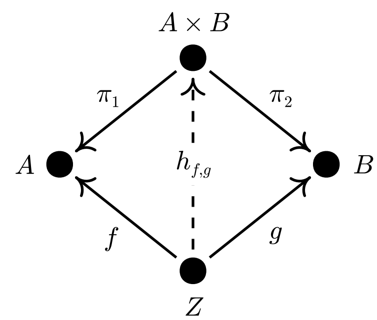

Figure 9: Diagrams showcasing the categorical product. Note that the dashed line is intended to highlight that the morphism $h_{f,g}$ is uniquely induced by f and g.

Note that this definition of a product is a **universal construction**, since it's done via existence and uniqueness. Another important aspect to note is that not every pair of objects in a category might have a product associated.

Theorem 2.19. For a category $\mathcal{C}$, a pair of objects A and B can have more than one product construction, but if this is the case, then both these constructions will be isomorphic.

Proposition 2.20 (Categorical Product vs Set Product). The categorical product generalizes the Cartesian product in set theory.

Proof. Consider the span $(X \times Y, \pi_1, \pi_2)$ where $\pi_1(x, y) = x$ and $\pi_2(x, y) = y$. Thus, for any span (Z, f, g) of A and B, make $h_{f,g}(z) = (f(z), g(z)) \in X \times Y$. This is how the Cartesian product works.

Let's drop the subscript in $h_{f,g}$. Now we have to show that h is a unique morphism, and that $h \mathbin{⨾} \pi_1 = f$ and $h \mathbin{⨾} \pi_2 = g$.

The second condition is trivially true, just note that

$$\forall z \in Z,\ \pi_1(h(z)) = \pi_1((f(z), g(z))) = f(z) \implies h \fatsemi \pi_1 = f,$$

and the same argument works for π_2 and g.

For uniqueness, consider $h' : Z \to X \times Y$ such that $h' \fatsemi \pi_1 = f$, and $h' \fatsemi \pi_2 = g$. If $h' \neq h$, then there is a $z \in Z$, such that $h(z) = (f(z), g(z)) \neq h'(z)$. Bu then, $\pi_1(h'(z)) \neq f(z)$ or $\pi_2(h'(z)) \neq g(z)$, which is a contradiction.

$\square$

From the definition of a product, it's easy to think of possible dual constructions by just inverting the arrows (morphisms).

Definition 2.21 (Cospan). Let A, B be objects in a category $\mathcal{C}$. A cospan on A and B is a triple (Z, f, g) where $f : A \to Z$ and $g : B \to Z$ are morphisms in $\mathcal{C}$. This is shown in Figure 10.

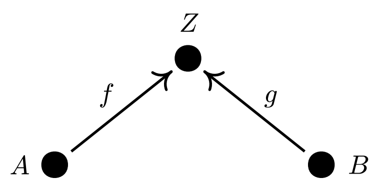

Figure 10: Diagrams showcasing a cospan between A and B.

Definition 2.22 (Categorical Coproduct). Let A, B be objects in a category $\mathcal{C}$. A cospan $(A + B, i_1, i_2)$ is called a product between A and B if for every span (Z, f, g) of A and B, there exists a unique morphism $h_{f,g} : Z \to A \times B$ such that $i_1 \fatsemi h_{f,g} = f$ and $i_2 \fatsemi h_{f,g} = g$. That is the same as saying that the diagram 11 commutes.

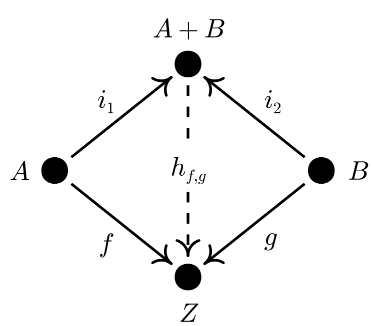

Figure 11: Diagrams showcasing the categorical coproduct. Note that the dashed line is intended to highlight that the morphism $h_{f,g}$ is uniquely induced by f and g.

While the categorical product was constructed to generalize the Cartesian set product, the coproduct was constructed in Category Theory, so the question is "to what does the coproduct corresponds in set theory?". The answer is the disjoint union! It's not a coincidence that we used "+" to symbolize it.

The idea of a product construction induced the notion of a product object $A \times B$. Yet, this construction also induces another definition, of the so called (co)product morphism.

Definition 2.23 (Product and Coproduct Morphism). Let $A, B, C, D \in Ob_{\mathcal{C}}$, $(A \times B, \pi_1^{A,B}, \pi_2^{A,B})$ and $(C \times D, \pi_1^{C,D}, \pi_2^{C,D})$ two products in the category $\mathcal{C}$, $f : A \to C$ and $g : B \to D$ two morphisms in $\mathcal{C}$. Hence, the morphism induced by $\pi_1^{A,B} \fatsemi f$ and $\pi_2^{A,B} \fatsemi g$ is called product morphism and denoted by $f \times g$, where

$$f \times g := (\pi_1^{A,B} \fatsemi f, \pi_2^{A,B} \fatsemi g) : A \times B \to C \times D.$$

Theorem 2.24. The product morphism $f \times g$ is the only morphism in $Mor_{\mathcal{C}}(A \times B, C \times D)$ such that

$$\pi_1^{C,D} \circ (f \times g) = f \circ \pi_1^{A,B}, \quad \pi_2^{C,D} \circ (f \times g) = g \circ \pi_2^{A,B}.$$

The coproduct morphism follows the same definition, but with coproducts.

Finally, one might be wondering what is an actual example of a product morphisms. For sets, it's the intuitive object, e.g. for two functions $f : A \to B$ and $g : C \to D$, the product morphism $f \times g : A \times B \to C \times D$ is just $(f \times g)(x, y) = (f(x), g(y))$.

2.9 Pullback, Pushout and Equalizers

3 Functors

Another central definition in Category Theory is that of Functors. While morphisms relate objects inside a category, a Functor establishes a relation between categories, thus, it's one level higher in terms of abstraction.

3.1 What is a Functor?

Let's formally define a Functor.

Definition 3.1 (Functor). Let $\mathcal{C}$ and $\mathcal{D}$ be two categories. A functor $F : \mathcal{C} \Rightarrow \mathcal{D}$ is a pair of mappings with the following properties:

(i) a mapping between objects

$$F : Ob_{\mathcal{C}} \to Ob_{\mathcal{D}},$$

where for each $A \in Ob_{\mathcal{C}}$, $F(A) \in Ob_{\mathcal{D}}$.

(ii) a mapping between morphisms

$$F : Mor_{\mathcal{C}} \to Mor_{\mathcal{D}},$$

where there are two possibilities:

(a) **Covariant Functor**, in which

$$F : Mor_{\mathcal{C}}(A, B) \to Mor_{\mathcal{D}}(F(A), F(B)),$$

hence for a morphism $f : A \to B$, then $F(f) : F(A) \to F(B)$.

(b) **Contravariant Functor**, in which

$$F : Mor_{\mathcal{C}}(A, B) \to Mor_{\mathcal{D}}(F(B), F(A)),$$

hence for a morphism $f : A \to B$, then $F(f) : F(B) \to F(A)$.

(iii) Identities morphisms are preserved, i.e. for $A \in Ob_{\mathcal{C}}$

$$F(id_A) = id_{F(A)}.$$

(iv) Compositions are preserved, i.e. for $f \in Mor_{\mathcal{C}}(A, B)$ and $g \in Mor_{\mathcal{C}}(B, C)$,

(a) For a **Covariant Functor**,

$$F(f \circ g) = F(f) \circ F(g).$$

(b) For a **Contravariant Functor**,

$$F(f \circ g) = F(g) \circ F(f).$$

It's common for authors to refer to covariant functors only as functors, i.e. whenever someone say that F is a functor, it might be implied that it's a covariant functor. We'll also use this convention whenever it's not ambiguous, and we'll always. Also, we'll sometimes use FA to mean $F(A)$.

Again, the use of diagrams may help understand what is going on. The figure below illustrates the identity and composition preservation of Functors.

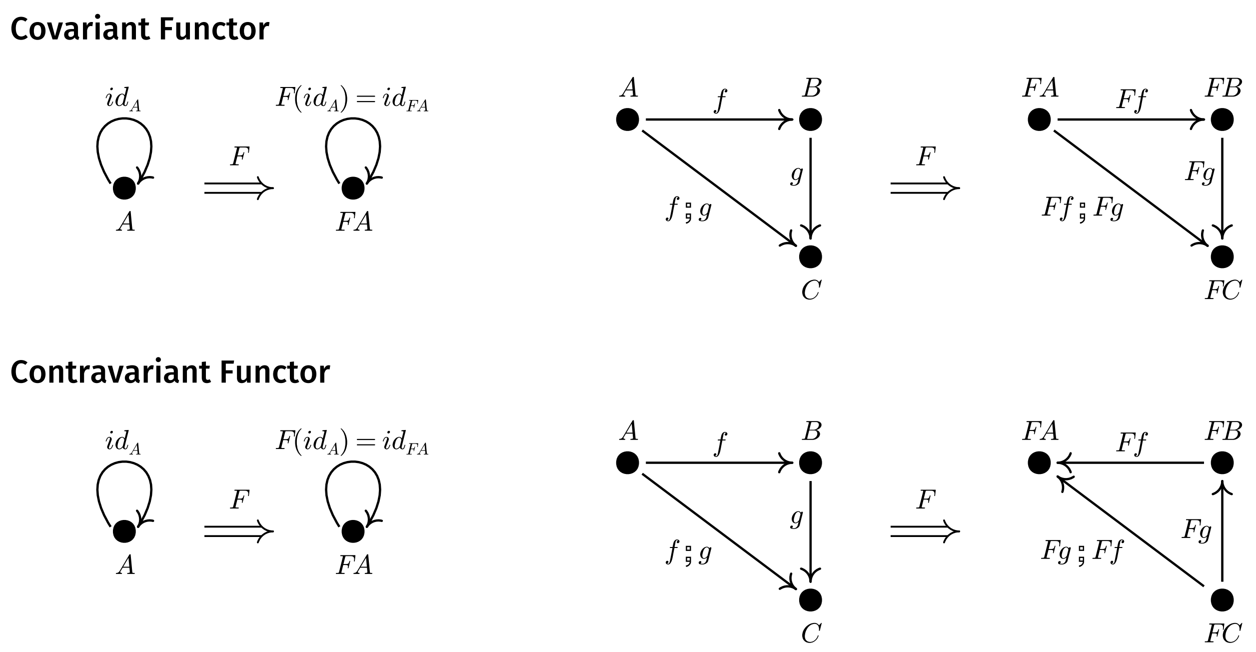

Figure 12: Diagrams showcasing the properties of Functors.

3.2 Category of Small Categories

One might realize that functors are acting on categories in a very similar way as morphisms do to objects. Indeed, we can define a functor composition to be such that for two functors $F : \mathcal{C} \Rightarrow \mathcal{D}$ and $G : \mathcal{D} \Rightarrow \mathcal{E}$, then $G \circ F$ is a functor from $\mathcal{C}$ to $\mathcal{E}$ where

1. For any $A \in Ob_{\mathcal{C}}$, $G \circ F(A) = G(F(A))$,
2. For any $f \in Mor_{\mathcal{C}}$, $G \circ F(f) = G(F(f))$.

We can also define an identity functor $I : \mathcal{C} \Rightarrow \mathcal{C}$, where $F(A) = A$ and $F(f) = f$.

Therefore, we might wonder whether there exists a category of all categories where objects are categories and morphisms are functors. The answer is "no". Similar to the set of all sets, it can be proven that this category does not exists. Yet, the category of all *small categories* does.

Remember, a small category is one where both morphisms and objects are sets. With this, let's prove our first theorem.

Theorem 3.2 (Category of Small Categories). Let $Ob_{\mathbf{SmCat}}$ be small categories and $Mor_{\mathbf{SmCat}}$ be functors. This constitutes a category.

Proof. To prove this, we'll use the fact that in Gödel-Bernays class set theory, the axiom 2.2 implies what is called a *comprehension scheme.*

Proposition 3.3 (Comprehension Scheme from Borceux [1]). If $\phi(x_1, ..., x_n)$ is a formula that the quantification occurs on set variables, then there exists a class A such that

$$(x_1, ..., x_n) \in A \iff \phi(x_1, ..., x_n).$$

Note that

$$\mathcal{C} := \langle Ob_{\mathcal{C}}, Mor_{\mathcal{C}} \rangle \in \mathbf{SmCat} \iff \mathcal{C} \text{ “is a category”}.$$

Since every $\mathcal{C}$ is small, then the formula to check whether $\mathcal{C}$ is a category iterates over set variables, i.e. $Ob_{\mathcal{C}}$ and $Mor_{\mathcal{C}}$. Thus, the Comprehension Scheme proposition guarantees that **SmCat** exists. □

3.3 Types of Functors

Before we go on to provide examples of functors, let's present a way to classify different functors.

Definition 3.4 (Faithful, Full, Fully Faithful and Embedding). Here we follow Roman et al. [7]. Let F be a functor between categories $\mathcal{C}$ and $\mathcal{D}$.

1. F is **faithful** if for every $A, B \in Ob_{\mathcal{C}}$, $F : Mor_{\mathcal{C}}(A, B) \to Mor_{\mathcal{C}}(FA, FB)$ is injective.
2. F is **full** if for every $A, B \in Ob_{\mathcal{C}}$, $F : Mor_{\mathcal{C}}(A, B) \to Mor_{\mathcal{C}}(FA, FB)$ is surjective.
3. F is **fully faithful** if for every $A, B \in Ob_{\mathcal{C}}$, $F : Mor_{\mathcal{C}}(A, B) \to Mor_{\mathcal{C}}(FA, FB)$ is bijective.
4. F is an **embedding** of $\mathcal{C}$ in $\mathcal{D}$ if F is fully faithful and $F : Ob_{\mathcal{C}} \to Ob_{\mathcal{D}}$ is injective.

3.4 Subcategories

Definition 3.5 (Subcategory). Let $\mathcal{D}$ be a category. We say that $\mathcal{C}$ is a subcategory of $\mathcal{D}$ if $Ob_{\mathcal{C}} \subset Ob_{\mathcal{D}}$ and $Mor_{\mathcal{C}} \subset Mor_{\mathcal{D}}$, such that $\mathcal{D}$ is a category.

If for every $A, B \in Ob_{\mathcal{D}}$, we have that $Mor_{\mathcal{D}}(A, B) = Mor_{\mathcal{C}}(A, B)$, then $\mathcal{D}$ is a *full subcategory.*

From the definition of a functor, one might wonder whether for any functor $F : \mathcal{C} \Rightarrow \mathcal{D}$, the image $F(\mathcal{C})$ is a subcategory of $\mathcal{D}$, i.e. if $\langle Ob_{F(\mathcal{C})}, Mor_{F(\mathcal{C})} \rangle$ is a category where

$$Ob_{F(\mathcal{C})} := \{F(A) : A \in Ob_{\mathcal{C}}\}, \quad Mor_{F(\mathcal{C})} := \{F(f) : f \in Mor_{\mathcal{C}}\}.$$

The answer is no.

3.5 Relevant Examples of Functors

Now that we know what a functor is, let's showcase some relevant examples that might be useful to someone applying Category Theory to another field.

Example 3.1 (Power Set Functor). The power set functor $\mathcal{P} : \mathbf{Set} \to \mathbf{Set}$ takes a set to it's power set and a function $f : A \to B$ to the image function $Imf : \mathcal{P}(A) \to \mathcal{P}(B)$, i.e. for a subset $S \subset A$ in the domain, returns $f(S) := \{f(x) \ : x \in S\}$.

Another even more relevant example is the *contravariante* power set functor $F : \mathbf{Set} \to \mathbf{Set}$ that takes A to $\mathcal{P}(A)$ and f to the inverse image f^{-1}.

Example 3.2 (Identity Functor). This does what one might expect from the name. The identity functor is $1_{\mathcal{C}} : \mathcal{C} \to \mathcal{C}$, such that $1_{\mathcal{C}}(A) = A \in Ob_{\mathcal{C}}$ and $1_{\mathcal{C}}(f) = f \in Mor_{\mathcal{C}}$.

Example 3.3 (Inclusion Functor). For a subcategory $\mathcal{S}$ of $\mathcal{C}$, the inclusion functor is $I_{\mathcal{C}} : \mathcal{S} \to \mathcal{C}$, such that $I_{\mathcal{C}}(A \in Ob_{\mathcal{S}}) = A \in Ob_{\mathcal{C}}$ and $I_{\mathcal{C}}(f \in Mor_{\mathcal{S}}) = f \in Mor_{\mathcal{C}}$.

3.6 Free Categories and $\mathcal{C}$-Sets

4 Natural Transformations

Similar to how morphisms define relations between objects in a category, natural transformations define relations between functors. The term "natural" in natural transformations was coined by Eilenberg and MacLane (the founders of Category Theory) due to the fact that these transformations were developed with the aim of explaining why some constructions in Mathematics were "natural" while others were not.

4.1 Defining Natural Transformations

Similar to functors, the formal definition of natural transformations are somewhat obscure at first sight, by as we dig a bit, we see that the is some intuition behind it that makes it easier to understand it.

Definition 4.1 (Natural Transformations). Let $\mathcal{C}$ and $\mathcal{D}$ be categories, and let $F, G : \mathcal{C} \to \mathcal{D}$ be functors. A natural transformation $\alpha : F \Rightarrow G$ is such that:

(i) For all $A \in \mathcal{C}$, there exists $\alpha_A : F(A) \to G(A)$ that is a morphism in $\mathcal{D}$, i.e $\alpha_A \in Mor_{\mathcal{D}}(F(A), G(A))$;

(ii) (Naturality) For all $f \in Mor_{\mathcal{C}}(A, B)$, then

$$F(f) \mathbin{⨟} \alpha_B = \alpha_A \mathbin{⨟} G(f).$$

Remember that for a morphism $f : A \to B$, we have that $F(f) : F(A) \to F(B)$ and $G(f) : G(A) \to G(B)$. Since $\alpha_A : F(A) \to G(A)$ and $\alpha_B : F(B) \to G(B)$, then $F(f)$ composes with α_B and $G(f)$ composes with α_A, and our definition above works.

Another way to represent property (ii) in the definition of natural transformations is to affirm that the diagram below.

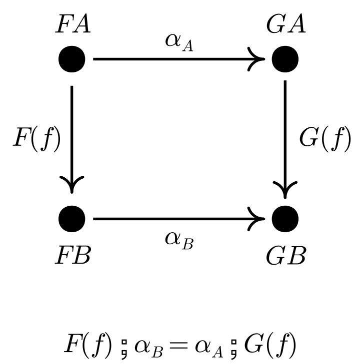

Figure 13: Commutative diagram of a natural transformation highlighting the commutative property of the definition.

From the definition, the natural transformation α is a transformation between functors that is in some sense associative, i.e. the order does not matter, we can apply α to FA and then apply Gf, or we can apply Ff to FA an then apply the natural transformation. Both return the same result.

4.2 The Category of Functors

The natural transformation is also ideal for defining a category of functors.

Definition 4.2. For categories $\mathcal{C}$ and $\mathcal{D}$, then $\mathcal{D}^{\mathcal{C}}$ is the category of functors from $\mathcal{C}$ to $\mathcal{D}$ where the morphisms are natural transformations, i.e.

$$Ob_{\mathcal{D}^{\mathcal{C}}} := \text{Functors from } \mathcal{C} \text{ to } \mathcal{D}$$

$$Mor_{\mathcal{D}^{\mathcal{C}}}(F, G) := \text{Natural Transformations from } F \text{ to } G.$$

Let's prove that the definition above is indeed a category. For that, we need to define a way to compose natural transformations that is associative, and we also need to define an identity morphism.

Proposition 4.3. In the category $\mathcal{D}^{\mathcal{C}}$, define the composition of two natural transformations $\alpha : F \to G$ and $\beta : G \to H$ as

$$\forall c \in \mathcal{C},\ \alpha_c \fatsemi \beta_c = (\alpha \fatsemi \beta)_c.$$

Then, using this definition, $\alpha \fatsemi \beta$ is indeed a normal transformation from $F \to H$, and this composition is associative.

Proof. First, since $\alpha_c \in Mor_{\mathcal{D}}(F(c), G(c))$, $\beta_c \in Mor_{\mathcal{D}}(G(c), H(c))$, $\gamma_c \in Mor_{\mathcal{D}}(H(c), J(c))$, then

$$\begin{aligned}
\alpha \fatsemi (\beta \fatsemi \gamma) \iff \forall c \in \mathcal{C}, \alpha_c \fatsemi (\beta \fatsemi \gamma)_c &= \alpha_c \fatsemi (\beta_c \fatsemi \gamma_c) \\
&= (\alpha_c \fatsemi \beta_c) \fatsemi \gamma_c \\
&= (\alpha_c \fatsemi \beta_c) \fatsemi \gamma_c \\
&= (\alpha \fatsemi \beta)_c \fatsemi \gamma_c \iff (\alpha \fatsemi \beta) \fatsemi \gamma.
\end{aligned}$$

We proved the associative part. Now, we need to show that such $\alpha \fatsemi \beta$ is a natural transformation from F to H. Take $f : c \to c' \in Mor_{\mathcal{C}}$. Consider $(\alpha \fatsemi \beta)_c : F(c) \to H(c)$, and $(\alpha \fatsemi \beta)_{c'} : F(c') \to H(c')$. Note that

$$F(f) \fatsemi \alpha_{c'} = \alpha_c \fatsemi G(f), \quad G(f) \fatsemi \beta_{c'} = \beta_c \fatsemi H(f).$$

Therefore, we have

$$\begin{aligned}
F(f) \fatsemi (\alpha \fatsemi \beta)_{c'} = (F(f) \fatsemi \alpha_{c'}) \fatsemi \beta_{c'} &= (\alpha_c \fatsemi G(f)) \fatsemi \beta_{c'} \\
&= \alpha_c \fatsemi (G(f) \fatsemi \beta_{c'}) \\
&= \alpha_c \fatsemi \beta_c \fatsemi H(f) \\
&= (\alpha \fatsemi \beta)_c \fatsemi H(f).
\end{aligned}$$

□

Example 4.1 (Natural Transformations in Sets). Consider the categories **1** and **Set**. We wish to construct the category $\mathbf{Set}^{\mathbf{1}}$, i.e. of functors from **1** to **Set**. Since **1** has only one object 1 and one morphisms id_1, then each functor $F : \mathbf{1} \to \mathbf{Set}$ is completely defined by $F(1) = S$, where S is a set, and, by the definition of a functor, $F(id_1) = id_S$. This means that every functor can be indexed by the set to which it takes the object of **1**, e.g. F_A is the functor that $F(1)$ is equal to set A.

Next, we want to define a natural transformation $\alpha : F_A \to F_B$. By definition, there are two criteria to satisfy. First, for every object $1 \in Ob_{\mathbf{1}}$ we have $\alpha_1 : F_A(1) \to F_B(1)$ such that $\alpha_1 \in Mor_{\mathbf{Set}}(F_A(1), F_B(1))$. Since the only object in **1** is 1, we only need to define α_1 to fully define α. Since α_1 must be a morphism of **Set**, then α_1 is a function from A to B.

Secondly, for every $f \in Mor_{\mathbf{1}}(1, 1)$, it's required that

$$F_A(f) \fatsemi \alpha_1 = \alpha_1 \fatsemi F_B(f).$$

Again, since the only morphism in $\mathbf{1}$ is id_1, we only need to prove that

$$F_A(id_1) \mathbin{⨾} \alpha_1 = \alpha_1 \mathbin{⨾} F_B(id_1).$$

Note that $F_A(id_1) = id_A$ and $F_B(id_1) = id_B$, and by the definition of the identity, we have that for $g \in Mor(A, B)$, $id_A \mathbin{⨾} g = g \mathbin{⨾} id_B$. Hence, the second condition is trivially satisfied.

With this, we conclude that the natural transformations from functors F_A to F_B are, in a sense, isomorphic to functions from sets A to B.

Considering the category **SmCat** of small categories, we can actually prove that **Set** and $\mathbf{Set}^{\mathbf{1}}$ are isomorphic.

Consider the functor $I : \mathbf{Set}^{\mathbf{1}} \to \mathbf{Set}$ where $I(F_A) = A \in \mathbf{Set}$ and for any natural transformation $\alpha^{(f)}$ where $\alpha_1^{(f)} = f$, we have $I(\alpha^{(f)}) = f$. Also, define $I^{-1}(A) = F_A$ and $I^{-1}(f) = \alpha_f$. Thus, I defines an isomorphism between the two categories.

This is only a sketch proof, since we would still need to prove that such functor I exists and is well-defined.

5 Limits, Colimits and Adjoints

6 What are Sets?

This section is based on Leinster [4].

When defining *small* and *locally small* categories, we need to differentiate between a *class* and a *set*. Anyone familiar with Russel's paradox on the set of all sets can appreciate why such distinction might be relevant.

One way to solve Russel's paradox was via Zermelo-Frankael and Choice (ZFC) axioms. Instead of strictly defining a set, the ZFC define what properties a set should have. Although this approach is the one assumed by most mathematicians, what ZFC calls a "set" does not actually match with how mathematicians use it. An example of the oddity in the definition of set's by ZFC is that elements of sets are also sets, so one could ask questions like "what are the elements of π?" [4].

Hence, instead of ZFC, we'll introduce here William Lawvere axioms as presented in Leinster [4]. Although less common, such system is more in sync with Category Theory, which is the subject at hand, and at the same time, it seems to more accurately describe what we mean by "sets".

6.0.1 Lawvere's Elementary Theory of the Category of Sets (ETCS)

As we said, to define a set we'll actually determine the properties that such object possesses. Thus, anything with such properties we'll be called a set. Of course, when stating such definition, we'll use terms that are again not tightly defined. But this is just part of life, since without such artifice, we would end up with circular definitions.

Let's now introduce the 10 axioms that make ETCS. This system of axioms is actually weaker (more general) than ZFC, and it can be shown to correspond to "Zermelo with bounded comprehension and choice" [4].

Although this axiomatization does not require Category Theory, we'll see that in some sense it has a categorical "flavor" to it.

Before stating the axioms, let's present some definitions that we'll be used in the axioms themselves. Note that these definitions only make sense once the axioms are established. But we present them now in order to make the exposition of ETCS cleaner.

Definition 6.1 (Terminal Set). A set T is called **terminal** in ETCS if for every set X there is only one function $f : X \to T$.

The terminal set is a way to define a single element set without relying on the definition of an element. In order to prove that this is indeed the case, we would need to clarify when two functions are the same, which will only be done after we present our axioms. It can be shown that every terminal set is unique up to an isomorphism, so one could use T to represent every terminal set.

Interestingly, if we are working in a context with a restricted collection of functions, then, a set T may behave as a single element set, while it may have multiple elements in another context. Consider for example, that $T = [0.5, 1]$, and we are in the context of functions that return natural numbers. Thus, for any set X, there exists only one function $f : X \to T$, which always returns 1.

As we've seen, the category of sets ($\boldsymbol{Set}$) will consist of $\langle Ob_{\boldsymbol{Set}}, Mor_{\boldsymbol{Set}} \rangle$, where $Ob_{\boldsymbol{Set}}$ is the collection of every set, and $Mor_{\boldsymbol{Set}}$ is the collection of every function. In the ETCS, the collection of every set will not be a set itself.

Definition 6.2 (Element of a Set). Given a set X, we write $x \in X$ to mean $x : T \to X$ where T is a terminal set.

Note that in this definition of an element, what we call an element of X is actually a function. Also, for $f : X \to Y$, then $f \circ x$ is a function from T to Y, i.e. it is an element of Y, which we write as $f(x) \in Y$.

Definition 6.3 (Cartesian Product). Given sets X and Y. The Cartesian product of X and Y is a set P, with functions $p_1 : P \to X$ and $p_2 : P \to Y$, such that for any set Z and functions $f_1 : Z \to P$ and $f_2 : Z \to P$, there exists a unique function $F = (f_1, f_2) : Z \to P$ where

$$p_1 \circ (f_1, f_2) = f_1, \quad p_2 \circ (f_1, f_2) = f_2.$$

Note that the Cartesian Product determines not only a product set, but also the projection functions. Similar to terminal sets, for any sets X and Y, the triple (P, p_1, p_2) are unique up to an isomorphism. Thus, we could fix (P, p_1, p_2) to be represented by $(X \times Y, \pi_1^{X \times Y}, \pi_2^{X \times Y})$.

Definition 6.4 (Function set). Let X and Y be two sets. A **function set** from X to Y is a tuple (F, ε), where F is a set and ε is a function $\varepsilon : F \times X \to Y$ such that for all sets Z and functions $q : Z \times X \to Y$, there exists a unique function $\overline{q} : Z \to F$ with $q(t, x) = \varepsilon(\overline{q}(t), x)$ for all $t \in Z$ and $x \in X$.

Definition 6.5 (Inverse Image). Let $f : X \to Y$ be a function and $y \in Y$. The **inverse image** of y under f is a tuple (A, j) where A is a set and $j : A \to X$ is a function such that $f \circ j(a) = y$ for every $a \in A$. Also, for every set Z and function $q : Z \to X$ such that $f(q(t)) = y$ for every $t \in Z$, there is a unique function $\overline{q} : Z \to A$ such that $q = j \circ \overline{q}$.

Again it can be shown that inverse images are unique up to an isomorphism.

Definition 6.6 (Injection). An injection $j : A \to X$ is a function with the property that $j(a) = j(a') \implies a = a'$ for every $a, a' \in A$.

Definition 6.7 (Surjection). A surjection $s : X \to Y$ is a function such that for every $y \in Y$ there exists an $x \in X$ such that $s(x) = y$.

Definition 6.8 (Right inverse). The right inverse of a function $s : X \to Y$ is a function $i : Y \to X$ such that $s \circ i = 1_Y$.

Definition 6.9 (Subset Classifier). The tuple $(\mathbf{2}, t)$ where $\mathbf{2}$ is a set and $t \in \mathbf{2}$ is called a subset classifier if for all sets A, X and injections $j : A \to X$, there is a unique function $\chi : X \to \mathbf{2}$, such that (A, j) is an inverse image of t under χ.

Note that in the definition above, the function χ can be seen as a characteristic function. Suppose that we wish to define χ_A. Hence, it's required that there exists a set $\mathbf{2}$ with $t \in \mathbf{2}$ such that $\chi_A(j(a)) = t$ for every $a \in A$.

Definition 6.10 (Natural Number System). A natural number system is a triple $(N, 0, s)$ where N is a set, $0 \in N$ and $s : N \to N$, such that for any set X, $a \in X$ and $r : X \to X$, there is a unique function $x : N \to X$ where $x(0) = a$ and $x(s(n)) = r(x(n))$ for every $n \in N$.

This comes from the idea that $s(n) \cong n+1$, that $x(0) \cong x_0$ and $x_n \cong x(s(n-1)) \cong r(x_n) \cong r(x(n))$. Once more, natural number systems are unique up to an isomorphism.

After all this definitions, we can finally state the axioms for Set Theory.

Definition 6.11 (ETCS). Lawvere's Elementary Theory of the Category of Sets consists on the following axioms:

(i) For all sets W, X, Y, Z, and functions $f : W \to X$, $g : X \to Y$, $h : Y \to Z$, we have

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

For every set X and Y and function $f : X \to Y$, there exist the identity functions 1_X and 1_Y, such that

$$f \circ 1_X = f = 1_Y \circ f.$$

(ii) There exists a terminal set T.

(iii) There exists a set with no elements, i.e. an empty set denoted by $\varnothing$.

(iv) For sets X, Y and functions $f : X \to Y$ and $g : X \to Y$, if $f(x) = g(x) \forall x \in X$, then $f = g$.

(v) Every pair of sets has a Cartesian product.

(vi) For all sets X and Y, there is a function set from X to Y.

(vii) For every $f : X \to Y$ and $y \in Y$, there is an inverse image of y with respect to f.

(viii) There exists a subset classifier. This can be thought as saying that for every set we can construct a characteristic function.

(ix) There exists a natural number system.

(x) Every surjection has a right inverse.

As we pointed out, these axioms are actually weaker than ZFC, but with one extra axiom, it can be shown to be as strong as ZFC. The last axiom is the one related to the Axiom of Choice. The first axiom states that sets form a category, and the following axioms distinguish this category from others.

With these axioms stated, we can now define the notion of a subset, and clearly differentiate objects that are and that aren't actually sets. One might think that "anything we can reasonably conceive" must be a set. But this is not the case.

Definition 6.12 (Subset). Given a set X, a subset of X is a function $f : X \to \mathbf{2}$. The subset $\chi_A : X \to \mathbf{2}$ is written as $A \subset X$, where χ_A is the characteristic function with $\chi_A^{-1}(t) = A$.

Corollary 6.13. A set T is terminal if and only if it has only a single element.

Proof. $\implies$) If T is terminal, then for any set X, we have a unique $f : X \to T$. For $t_1, t_2 \in T$, then $t_1 : T' \to T$ and $t_2 : T' \to T$ where T' is a terminal set. Note that $f : T' \to T$ is unique, hence, $t_1 = t_2$, meaning that T has only a single element.

$\impliedby$) If T has a single element $t \in T$, then for a set X, take $f_1 : X \to T$ and $f_2 : X \to T$. Since T has only one element, then $f_1(x) = t = f_2(x)$, which, by Axiom 3, implies that $f_1 = f_2$. $\square$

References

[1] Francis Borceux. *Handbook of categorical algebra: volume 1, Basic category theory*, volume 1. Cambridge University Press, 1994.

[2] Tai-Danae Bradley, Tyler Bryson, and John Terilla. *Topology: A Categorical Approach*. MIT Press, 2020.

[3] Brendan Fong and David I Spivak. *An invitation to applied category theory: seven sketches in compositionality*. Cambridge University Press, 2019.

[4] Tom Leinster. Rethinking set theory. *The American Mathematical Monthly*, 121(5): 403–415, 2014.

[5] Bartosz Milewski. *Category theory for programmers*. Blurb, 2018.

[6] Maico Ribeiro. *Teoria das Categorias para Matemáticos. Uma breve introdução*. 05 2020. ISBN 9786599039515.

[7] Steven Roman et al. *An Introduction to the Language of Category Theory*, volume 6. Springer, 2017.