Category Theory

Davi Sales Barreira

February 22, 2022

Contents

List of Definitions

List of Theorems

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

1 Why study Category Theory

The study of Category Theory enables us to view Mathematics from a vantage point, and better understand how the different areas are connected. For example, it might not always be clear which properties are *topological*, and which aren't. By looking at the subject from the distance (via Category Theory), we get a glimpse at the connections (and disconnections) between different fields.

Another very interesting observation about Category Theory is that it's becoming very popular in programming. This is highlighted for example in the book Milewski [5]. In order to help the understanding of the subject, I'll be using "applications" of Category Theory, mainly inspired by Fong and Spivak [3]. We'll also do coding examples using Catlab.jl, a Julia package for applied Category Theory.

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 statements is Russel's Paradox of whether there is a set of all sets, for which the answer is "no".

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 $\cup_{i \in I} x_i \in \mathfrak{U}$;

[1] Definition from Borceux [1]

(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*.

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 $f \in Mor_{\mathcal{C}}(X, Y)$, $g \in Mor_{\mathcal{C}}(Y, X)$,

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

There are many ways to refers to the set 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 **1** consists of $Ob_{\mathbf{1}} := \{a\}$ and $Mor_{\mathbf{1}} = \{id_a\}$. The diagram for such Category is shown below.

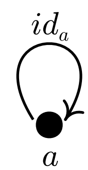

Figure 1: Hasse diagram of Category **1**.

Example 2.2 (Category 2 and 3). The Category **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.

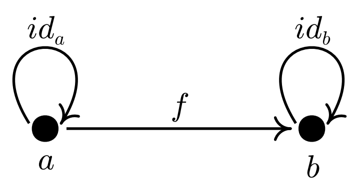

Figure 2: Hasse diagram of Category **2**.

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.

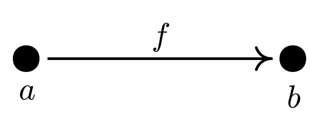

Figure 3: Hasse diagram of Category **2** omitting identity morphisms.

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.

$$
\begin{tikzcd}
a \arrow[loop above, "id"] \arrow[r, "f"] \arrow[rr, bend right, "f \fatsemi g"'] & b \arrow[loop above, "id"] \arrow[r, "g"] & c \arrow[loop above, "id"]
\end{tikzcd}
$$

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.

$$0 \xleftarrow{\leq} 1 \xleftarrow{\leq} 2 \xleftarrow{\leq} 3 \longleftarrow \cdots$$

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*).

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 Brief words on isomorphism

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 $g \in Mor_{\mathcal{C}}(Y, X)$ such that $g \circ f = id_X$;

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

(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.

Note that when f is invertible, the morphism that inverts f is unique with the left and right inverses coinciding, since $g \circ id_Y = g \circ f \circ h = id_X \circ h = h$.

3 Functors

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.

4 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".

4.0.1 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 4.1 (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 (***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 4.2 (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 4.3 (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 4.4 (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 4.5 (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 4.6 (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 4.7 (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 4.8 (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 4.9 (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 4.10 (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 4.11 (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 4.12 (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 4.13. 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.