Category Theory

Davi Sales Barreira

February 8, 2022

Contents

List of Definitions

List of Theorems

Notes mostly based on Ribeiro [4], Bradley et al. [1] and Milewski [3].

1 What are Categories?

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 a glimpse at the connections (and disconnections) between different fields.

Another very intresting observation about Category Theory is that it's becoming very popular in programming. This is highlighted for example in the book Milewski [3]. Thus, in order to help the understanding of the subject, I'll be using programming as a real world example of application of Category Theory. Even more specifically, the example shall be done using Julia.

Definition 1.1 (Category). A category $\mathcal{C} = \langle Ob_{\mathcal{C}}, Mor_{\mathcal{C}} \rangle$ is a collection of objects $Ob_{\mathcal{C}}$ and 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

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

(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.

Definition 1.2 (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 $g \circ id_Y = g \circ f \circ h = id_X \circ h = h$.

1.1 Programming as Category

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

1.2 What are Sets?

This section is based on Leinster [2].

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?" [2].

Hence, instead of ZFC, we'll introduce here William Lawvere axioms as presented in Leinster [2]. 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".

1.2.1 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" [2].

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 1.3 (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 1.4 (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 1.5 (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 1.6 (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 1.7 (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 1.8 (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 1.9 (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 1.10 (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 1.11 (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 1.12 (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 1.13 (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 1.14 (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 1.15. 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] Tai-Danae Bradley, Tyler Bryson, and John Terilla. *Topology: A Categorical Approach.* MIT Press, 2020.

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

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

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