

Category theory and diagrammatic reasoning

30th January 2019

Last updated: 30th January 2019

1 Categories, functors and diagrams

It is a common opinion that sets are the most basic mathematical objects. The intuition of a set is a collection of elements with no additional structure.

When the set is finite, especially when it has a small number of elements, then we can describe it simply by listing its elements one by one. With infinite sets, we usually give some algorithmic procedure which generates all the elements, but this rarely completes the description: that only happens when the set is “freely generated” in some sense.

Example 1. The set \mathbb{N} of natural numbers is freely generated by the procedure:

- add one element, 0, to the empty set;
- at every step, if x is the last element you added, add an element Sx .

It is more common that a set X is *presented* by first giving a set X_0 of *terms* that denote elements of X , and which is freely generated; and then by giving an *equivalence relation* on X_0 which tells us when two terms of X_0 should be considered equal as elements of X . This is sometimes called a *setoid*.

We can present the equivalence relation on X_0 as a set X_1 , together with two functions

$$X_1 \begin{array}{c} \xrightarrow{s} \\ \xrightarrow{t} \end{array} X_0. \quad (1)$$

You should think of an element $f \in X_1$ as a “witness of the fact that $s(f)$ is equal to $t(f)$ in X ”. To transform this into an equivalence relation on X_0 , we take the relation

$$x \sim y \text{ if and only if there exists } f \in X_1 \text{ such that } s(f) = x \text{ and } t(f) = y$$

and then close it under the reflexive, symmetric, and transitive properties.

As a mathematical structure, (1) is a directed graph with labelled vertices and edges (from now on, just a *graph*). This graph has X_0 as its set of vertices, and it has an edge from $s(f)$ to $t(f)$, labelled f , for all $f \in X_1$:

$$\begin{array}{ccc} s(f) & & t(f) \\ \bullet & \xrightarrow{f} & \bullet \end{array}$$

We will write $f : x \rightarrow y$, in short, for an edge that has $s(f) = x$ and $t(f) = y$.

The fact is, some equivalence relations can be quite useless. For example, say X_0 has two elements $\{x, y\}$, and $x \sim y$ if and only if the Riemann hypothesis is true: neither of us can tell whether $x = y$ or $x \neq y$ in X .

In practice, many useful equivalence relations come with a procedure to *decide* when two elements are equal, performing certain computations with terms; and when we compute something, there should be a *direction* towards which we are going, or we may end up stuck in a loop. In these situations, the “proofs of equality” $f \in X_1$ may be seen as *processes* going from $s(f)$ to $t(f)$.

Example 2. The set $\mathbb{Q}_{>0}$ of positive rational numbers is generated by pairs (n, m) of non-zero natural numbers, together with the equivalence relation $(n_1, m_1) \sim (n_2, m_2)$ if and only if $n_1 m_2 = n_2 m_1$.

Any positive rational number has a “minimal” representative given by (p, q) such that $\gcd(p, q) = 1$. So starting from (n, m) , we may want to reduce it to this representative, dividing n and m by their common divisors. We can describe each reduction step by a triple (n, m, k) , where $k > 1$ is a common divisor of n and m : the triple is telling us that we can reduce (n, m) by dividing both numbers by k . This corresponds to the graph

$$X_1 := \{(n, m, k) \mid k \text{ divides both } n \text{ and } m\},$$

$$s(n, m, k) := (n, m), \quad t(n, m, k) := \left(\frac{n}{k}, \frac{m}{k}\right).$$

Exercise 3. Prove that this graph presents $\mathbb{Q}_{>0}$.

Even in very simple situations, if we keep track of all the computations we do with elements of sets, we are forced to admit we are actually working at least with directed graphs. When we write something as simple as $2 + 2 = 4$, it is not actually the case that the two sides have the same computational content: for example, by computation, we can go uniquely from $2 + 2$ to a single natural number, 4, but we cannot uniquely decompose 4 as a sum of two numbers. In other words, the “equality” $2 + 2 = 4$ has a preferred computational direction $2 + 2 \rightarrow 4$.

Of course, the computational interpretation of graphs goes beyond the presentation of set(oid)s. The structure (1) is the same as a *labelled transition system*: the elements of X_0 are *states* of a system, and the edges $f \in X_1$ are possible *transitions* from the state $s(f)$ to the state $t(f)$. It is also the same as a labelled *abstract rewriting system*. In this context, we may not want to consider $s(f)$ and $t(f)$ as “equivalent”.

There are two natural things that we can do with computations:

1. we are in state x and we do nothing, just stay at x ;
2. we proceed from x to y by a process $f : x \rightarrow y$, then we proceed from y to z by a process $g : y \rightarrow z$, and so on.

Interpreting the first one requires the structure of a *reflexive graph*.

Definition 4. A *reflexive graph* is a graph X together with a function $\text{id}_{(-)} : X_0 \rightarrow X_1$ satisfying $s(\text{id}_x) = t(\text{id}_x) = x$ for all $x \in X_0$.

We call id_x the *identity* on x .

For the second one, we need to be able to interpret any *finite path* in the graph X as an edge of X . Let

$$X_1^+ := \{f_1, \dots, f_n \mid f_i \in X_1, t(f_i) = s(f_{i+1})\}$$

be the set of finite paths in a reflexive graph X . The structure that we need is a *composition* function $m : X_1^+ \rightarrow X_1$ such that $s(m(f_1, \dots, f_n)) = s(f_1)$ and $t(m(f_1, \dots, f_n)) = t(f_n)$. Furthermore, we expect it to satisfy some properties:

1. the composite of a path f of length 1 is f itself;
2. composing parts of a path, then composing the composites, is the same as composing the entire path;
3. composing with an identity “does nothing”.

Definition 5. A *category* is a reflexive graph X together with a *composition* operation $m : X_1^+ \rightarrow X_1$, satisfying the following equations whenever the left-hand side is defined:

1. $s(m(f_1, \dots, f_n)) = s(f_1)$ and $t(m(f_1, \dots, f_n)) = t(f_n)$;
2. $m(f) = f$;
3. $m(f_1, \dots, f_n) = m(m(f_1, \dots, f_j), \dots, m(f_k, \dots, f_n))$;
4. $m(f, \text{id}_x) = m(\text{id}_x, f) = f$.

A category has an underlying graph. The vertices of this graph are usually called the *objects* of the category. The edges of this graph are called the *morphisms*, or *arrows* of the category. In the prospect of higher-dimensional generalisations, objects and morphisms may also be called *0-cells* and *1-cells*, respectively.

If f is a morphism, then $s(f)$ is called the *source*, and $t(f)$ the *target* of f . Other common terms are *input* and *output*, or *domain* and *codomain*, respectively.

We will write $f;g$ for $m(f,g)$, and similarly $f_1; \dots; f_n$ for $m(f_1, \dots, f_n)$, whenever defined. Other common notations for $f;g$ are $g \circ f$ and gf ; notice that these reverse the order in which the morphisms are written.

Remark 6. Strictly speaking, this is the definition what is called a *small* category; in general, objects and morphisms may not be required to form a set. For now you should not be too worried about it.

Remark 7. The more common definition requires a composition \tilde{m} only of paths of length two, satisfying

$$\tilde{m}(\tilde{m}(f, g), h) = \tilde{m}(f, \tilde{m}(g, h)), \quad \tilde{m}(f, \text{id}_x) = \tilde{m}(\text{id}_x, f) = f,$$

whenever either side is defined. This is simpler to check, and what you will want to use in practice.

Exercise 8. Prove that the two definitions are equivalent.

Let us look at a few examples.

Example 9. Let X be a graph. We can form a reflexive graph X^* with

$$X_1^* := X_1^+ + \{\text{id}_x : x \rightarrow x \mid x \in X_0\},$$

that is, the edges of X^* are finite paths, possibly of “length 0”, in X . This becomes a category by letting the composite of two paths be their concatenation. This is called the *free category* on a graph.

Example 10. A *preorder* on a set P is a binary relation \leq which is *reflexive* and *transitive*, that is, for all $x, y, z \in P$,

- $x \leq x$, and
- if $x \leq y$ and $y \leq z$, then $x \leq z$.

A preorder determines a category as follows:

- the objects are the elements of P ;
- there is a unique morphism $x \rightarrow y$ if and only if $x \leq y$;
- the identity on x is the unique morphism $x \rightarrow x$, which exists because \leq is reflexive;
- the composite of two morphisms $x \rightarrow y$ and $y \rightarrow z$ is the unique morphism $x \rightarrow z$, which exists because \leq is transitive.

Conversely, any category with the property that between any two objects x, y there is *at most* one morphism determines a preorder on its set of objects.

Definition 11. An *isomorphism* $f : x \rightarrow y$ in a category X is a morphism with the following property: there is an *inverse* morphism $f^{-1} : y \rightarrow x$ such that $f; f^{-1} = \text{id}_x$ and $f^{-1}; f = \text{id}_y$. We say that two objects x, y are *isomorphic* if there is an isomorphism $f : x \rightarrow y$.

A *groupoid* is a category with the property that all its morphisms are isomorphisms.

Exercise 12. Prove that “being isomorphic” is an equivalence relation on the objects of a category.

Example 13. In a preorder, two objects x, y are isomorphic if and only if $x \leq y$ and $y \leq x$. A *partial order* is exactly a preorder with the property that any two isomorphic objects are equal. An *equivalence relation* is a preorder which is also a groupoid.

Example 14. A *monoid* is a set M together with a multiplication operation $(x, y) \mapsto x \cdot y$ and a chosen element e , the *unit*, satisfying

$$(x \cdot y) \cdot z = x \cdot (y \cdot z), \quad x \cdot e = e \cdot x = x$$

for all $x, y, z \in M$. Looking at Remark 7, you should see that a monoid is the same as a category with a single object $*$: the elements of M are morphisms $x : * \rightarrow *$, the unit is id_* , and multiplication is composition of morphisms.

A monoid is a *group* if all its elements have two-sided inverses: this is the same as a *groupoid* with a single object.

Example 15. Most mathematical structures with an underlying set form categories together with their structure-preserving functions; these are called *concrete* categories. In all of these, composition is composition of functions, and identities are identity functions. These categories are often “large” in the sense that their objects do not form a set, but we may reduce them to “small” categories by restricting to objects below a certain size (e.g. sets below a certain cardinality).

Examples are:

- **Set** whose objects are sets, morphisms are functions;
- **Grp** whose objects are groups, morphisms are group homomorphisms;
- **Vec_k** whose objects are vector spaces over a field k , and morphisms are linear maps;
- **Top** whose objects are topological spaces, and morphisms are continuous maps;

and many, many more.

Definition 16. Let x, y be two objects of a category X . The *hom-set* of morphisms from x to y is the subset

$$\text{Hom}_X(x, y) := \{f : x \rightarrow y\}$$

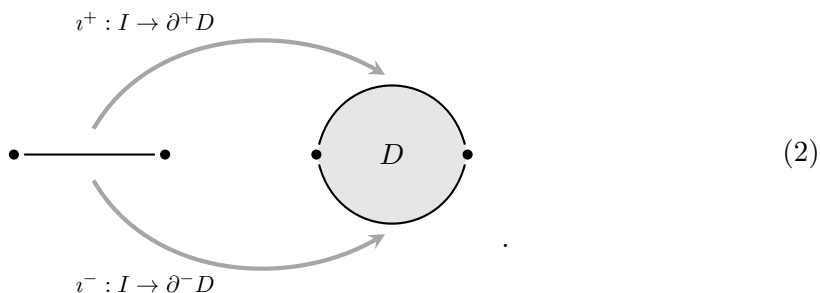
of the set of morphisms X_1 .

A morphism $f : x \rightarrow x$ is called an *endomorphism* of x . For each x , the hom-set $\text{Hom}_X(x, x)$ admits the structure of a monoid, whose multiplication is composition of endomorphisms.

Example 17. If you know some algebraic topology, you may have encountered the *fundamental groupoid* of a space X . This is a category defined as follows.

A *path* in X is a continuous map $f : I \rightarrow X$, where I is the closed interval $[0, 1]$. Let D be the topological disk; we can subdivide its boundary into two copies $\partial^- D$ and $\partial^+ D$

of I , connected by their extremities:



A *homotopy relative to ∂I* between two paths $f, g : I \rightarrow X$ is a map $h : D \rightarrow X$ such that $i^-; h = f$ and $i^+; h = g$. This implies that $f(0) = g(0)$ and that $f(1) = g(1)$. We say that two paths are homotopic relative to ∂I if there exists such a homotopy: it can be shown that this is an equivalence relation on paths in X .

The *fundamental groupoid* $\pi(X)$ of X is the category so defined:

- the objects are the points of the space X ;
- the morphisms $f_\sim : x \rightarrow y$ are equivalence classes of paths $f : I \rightarrow X$ with $f(0) = x$ and $f(1) = y$, modulo homotopy relative to ∂I ;
- the identity on x is the (equivalence class of) the constant path on x ;
- the composition of $f_\sim : x \rightarrow y$ and $g_\sim : y \rightarrow z$ is the equivalence class of the *concatenation* $f * g : [0, 2] \rightarrow X$ of the two paths, pre-composed with a fixed homeomorphism $[0, 1] \rightarrow [0, 2]$ (for example $t \mapsto 2t$).

This is well-defined, and is in fact a groupoid: the inverse of $f_\sim : x \rightarrow y$ is the equivalence class of the *reversed* path $f^* : I \rightarrow X$, which is f precomposed with $t \mapsto (1 - t)$.

For each point $x \in X$, the *fundamental group* $\pi(X, x)$ of X at x is the monoid of endomorphisms $\text{Hom}_{\pi(X)}(x, x)$.

Definition 18. Let X and Y be two categories. A *functor* $F : X \rightarrow Y$ is a pair of functions $F : X_0 \rightarrow Y_0$ and $F : X_1 \rightarrow Y_1$ satisfying the following equations, whenever the left-hand side makes sense:

1. $s(F(f)) = F(s(f))$ and $t(F(f)) = F(t(f))$,
2. $F(f; g) = F(f); F(g)$, and
3. $F(\text{id}_x) = \text{id}_{F(x)}$.

The functor F is an *isomorphism* if there exists a functor $F^{-1} : Y \rightarrow X$ such that $F; F^{-1} : X_i \rightarrow X_i$ and $F^{-1}; F : Y_i \rightarrow Y_i$ are the identity function, for $i = 0, 1$. Two categories are *isomorphic* if there exists an isomorphism between them.

A functor is an *inclusion* if it is injective on both objects and morphisms.

In words, a functor maps a morphism $f : x \rightarrow y$ to a morphism $F(f) : F(x) \rightarrow F(y)$, in such a way that the image of a composite is the composite of the images, and the image of an identity is an identity.

Definition 19. The *identity functor* on a category X is the functor $\text{id}_X : X \rightarrow X$ whose components are identity functions on X_i .

Given two functors $F : X \rightarrow Y$ and $G : Y \rightarrow Z$, their *composite* $F;G : X \rightarrow Z$ is the functor with component functions $F;G : X_i \rightarrow Z_i$, for $i = 0, 1$.

Restricting to categories of small size (so there is a set of them), this defines a category **Cat**.

Example 20. The concrete categories of Example 15 have a *forgetful* functor to **Set**, which sends their objects to their underlying sets, and the morphisms to their underlying functions of sets.

A forgetful functor may factor as a composite of functors progressively forgetting more and more structure: for example, the forgetful functor $U : \mathbf{Vec}_k \rightarrow \mathbf{Set}$ factors as $U : \mathbf{Vec}_k \rightarrow \mathbf{Grp}$, sending a vector space to its underlying abelian group with addition, followed by $U : \mathbf{Grp} \rightarrow \mathbf{Set}$.

Exercise 21. What are functors between preorders? And functors between monoids?

Lemma 22. Let $f : x \rightarrow y$ be an isomorphism in X and $F : X \rightarrow Y$ a functor. Then $F(f)$ is an isomorphism in Y .

Proof. Let $f^{-1} : y \rightarrow x$ be the inverse of f in X . Then

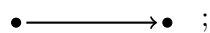
$$F(f);F(f^{-1}) = F(f;f^{-1}) = F(\text{id}_x) = \text{id}_{F(x)},$$

$$F(f^{-1});F(f) = F(f^{-1};f) = F(\text{id}_y) = \text{id}_{F(y)},$$

that is, $F(f^{-1})$ is an inverse to $F(f)$ in Y . □

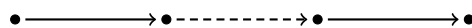
Example 23. Let 1 be the free category on the graph with one vertex and no edges; this has a single object $*$, and only the identity morphism $\text{id}_* : * \rightarrow *$. A functor $F : 1 \rightarrow X$ maps $*$ to an object $F(*)$ of X , and the identity on $*$ to the identity on $F(*)$: equivalently, it is the choice of an object of X .

Example 24. Let \vec{I} be the free category on the graph



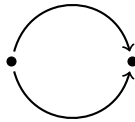
then a functor $F : \vec{I} \rightarrow X$ is, equivalently, the choice of a morphism of X .

Similarly, letting $\#^n \vec{I}$ be the free category on the linear graph

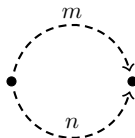


with n consecutive edges (the dotted line stands for a sequence of edges), a functor $F : \#^n \vec{I} \rightarrow X$ corresponds to a composable sequence of n morphisms in X .

Then, functors from the free category ∂O^2 on the graph



are the same as *parallel* pairs of morphisms in X , that is, pairs of morphisms with the same source and target; and functors from the free category $\partial O^2_{n,m}$



where the lower semicircle is a sequence of n edges, and the upper one a sequence of m edges, are the same as parallel *paths* of morphisms in X .

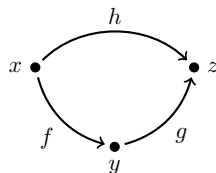
Exercise 25. What is a functor $F : \partial O^2 \rightarrow \mathbf{Set}$?

In general, we can depict functors from free categories on a graph by labelling the vertices and edges of these graphs with the names of their images in X . For example, the functor $\partial O^2 \rightarrow X$ corresponding to a pair of morphisms $f, g : x \rightarrow y$ is depicted as the diagram



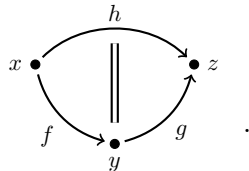
Definition 26. Let X be a graph, and $F : X^* \rightarrow Y$ a functor from the free category on X to a category Y . We say that F is a *commutative diagram* in Y if the images of all parallel paths in X have equal composites in Y .

Example 27. The functor $F : \partial O^2 \rightarrow X$ depicted in (3) is a commutative diagram if and only if $f = g$. A functor $F : \partial O^2_{2,1} \rightarrow X$, that is, a diagram

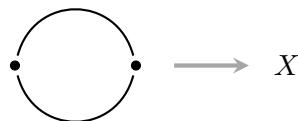


is commutative if and only if $h = f;g$.

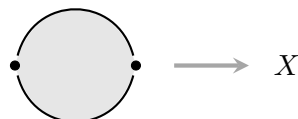
We will often express the fact that a diagram is commutative by placing an equality sign between two parallel paths, e.g.



Remark 28. In the fundamental groupoid of a space X , two paths f, g with the same endpoints are declared equal whenever the map

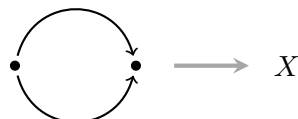


from the circle, which is equal to f on the upper semicircle and to g on the lower semicircle, can be “completed” to (factors through) a map

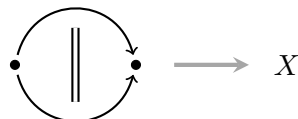


from the disk.

The intuition of a commutative diagram in a category X is the same: any parallel pair of paths determines a functor



which is a commutative diagram if and only if it factors through a functor



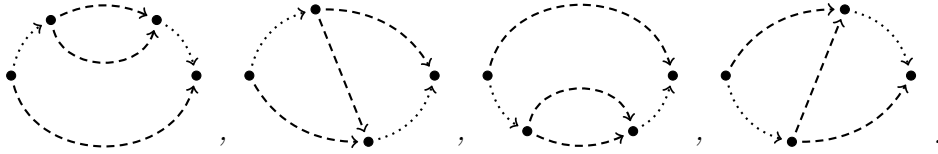
from the category where the two parallel edges are declared equal (isomorphic to \vec{I}).

Lemma 29. *Let $F : X^* \rightarrow Y$ be a commutative diagram in Y , and $G : Y \rightarrow Z$ a functor. Then $F;G : X^* \rightarrow Z$ is a commutative diagram in Z .*

Proof. Let p and q be any pair of parallel paths in X . Then $F(p) = F(q)$, and consequently $F;G(p) = G(F(p)) = G(F(q)) = F;G(q)$. \square

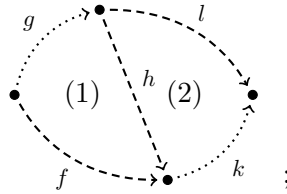
The most important fact about commutative diagrams is that they can be composed together. In the following pictures, a “dotted arrow” means a path of 0 or more edges, and a “dashed arrow” means a path of 1 or more edges.

Proposition 30. *Let $F : X^* \rightarrow Y$ be a functor from the free category on a graph X which has one of the following shapes:*



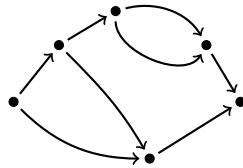
If the restrictions of F to the two smaller circles composing the larger circle are both commutative, then F is commutative.

Proof. We consider the case of a diagram



the others are analogous. We have $F(f;k) = F(f);F(k) = F(g;h);F(k)$ by commutativity of the restriction to (1). This is equal to $F(g);F(h);F(k) = F(g);F(h;k) = F(g);F(l)$ by commutativity of the restriction to (2). Hence $F(f;k) = F(g;l)$. \square

This “composition of commutative diagrams” can be iterated. For example, a diagram of shape



is commutative if and only if its restrictions to the three components are: it suffices to focus on a pair, first, to obtain commutativity of its outer boundary; and then repeat for the outer boundary of the pair and the third component.

We conclude with a couple of constructions of categories from other categories.

Construction 31. Let X be a category. There is a category X^{op} , the *opposite* category of X , whose objects are the same as those of X , and there is a morphism $f^{\text{op}} : y \rightarrow x$ for each morphism $f : x \rightarrow y$ of X : that is, X^{op} is X with “all arrows reversed”. The identity on x in X^{op} is id_x^{op} , and the composite $g^{\text{op}}; f^{\text{op}}$ is defined to be $(f; g)^{\text{op}}$.

Given a functor $F : X \rightarrow Y$, there is a functor $F : X^{\text{op}} \rightarrow Y^{\text{op}}$, equal to F on objects, and defined by $F(f^{\text{op}}) := (F(f))^{\text{op}}$ on morphisms.

Definition 32. A functor $F : X^{\text{op}} \rightarrow Y$ is called a *contravariant* functor from X to Y .

It is common to treat contravariant functor as if they were defined on X , rather than X^{op} , but reversed the source and target and the direction of composition: for example, given $f : x \rightarrow y$ in X , write $F(f) : F(y) \rightarrow F(x)$ for $F(f^{\text{op}})$.

Example 33. The *powerset* functor $\mathcal{P} : \mathbf{Set}^{\text{op}} \rightarrow \mathbf{Set}$ is the contravariant functor mapping a set A to its powerset \mathcal{P} , and a function $f : A \rightarrow B$ to the *inverse image* function $\mathcal{P}f : \mathcal{P}B \rightarrow \mathcal{P}A$, sending $U \subseteq B$ to $f^{-1}(U) \subseteq A$.

Construction 34. Let X and Y be two categories. There is a category $X \times Y$, the *product* of the categories X and Y , defined as follows:

- objects are ordered pairs (x, y) of an object of X and an object of Y ;
- morphisms $(x, y) \rightarrow (x', y')$ are pairs (f, g) of a morphism $f : x \rightarrow x'$ in X and a morphism $g : y \rightarrow y'$ in Y ;
- the unit on (x, y) is $(\text{id}_x, \text{id}_y)$;
- the composite of (f, g) and (f', g') is $(f; f', g; g')$.

Remark 35. The products $X \times 1$ and $1 \times X$ are both isomorphic to X .

For all categories X, Y , there are *projection* functors $p_1 : X \times Y \rightarrow X$ and $p_2 : X \times Y \rightarrow Y$, sending (x, y) to x and y , respectively, and (f, g) to f and g , respectively.

For all objects y of Y , there is an inclusion $(-, y) : X \rightarrow X \times Y$ sending x to (x, y) and f to (f, id_y) . Similarly, for all objects x of X , there is an inclusion $(x, -) : Y \rightarrow X \times Y$ sending y to (x, y) and g to (id_x, g) .

Example 36. Let X be a category. There is a functor $\text{Hom}_X : X^{\text{op}} \times X \rightarrow \mathbf{Set}$, sending the object (x, y) to the homset $\text{Hom}_X(x, y)$, and the morphism $(f^{\text{op}}, h) : (x', y) \rightarrow (x, y')$, corresponding to a pair $f : x \rightarrow x'$ and $h : y \rightarrow y'$ of morphisms of X , to the function

$$\begin{aligned} \text{Hom}_X(x', y) &\rightarrow \text{Hom}_X(x, y'), \\ g &\mapsto f; g; h. \end{aligned}$$

Precomposing with an inclusion $(-, x) : X^{\text{op}} \rightarrow X^{\text{op}} \times X$ gives a contravariant functor $\text{Hom}_X(-, x) : X^{\text{op}} \rightarrow \mathbf{Set}$, and precomposing with $(x, -)$ a functor $\text{Hom}_X(x, -) : X \rightarrow \mathbf{Set}$. These are called *contravariant* and *covariant* homset functors, respectively.

Exercise 37. Given two categories X and Y , consider the graph whose vertices are of the form $x \boxtimes y$, where x is an object of X and y an object of Y , and edges are of the form $f \boxtimes y : x \boxtimes y \rightarrow x' \boxtimes y$ or $x \boxtimes g : x \boxtimes y \rightarrow x \boxtimes y'$ for all morphisms $f : x \rightarrow x'$ in X and $g : y \rightarrow y'$ in Y .

Then, let $X \boxtimes Y$ be the category whose objects are the vertices of this graph, and morphisms are finite paths in this graph, quotiented by the equations

$$\begin{aligned} (f \boxtimes y); (f' \boxtimes y) &= (f; f') \boxtimes y, \\ (x \boxtimes g); (x \boxtimes g') &= x \boxtimes (g; g'), \\ \text{id}_x \boxtimes y &= x \boxtimes \text{id}_y = \text{id}_{x \boxtimes y}, \\ (f \boxtimes y); (x' \boxtimes g) &= (x \boxtimes g); (f \boxtimes y'), \end{aligned}$$

whenever either side is defined. Prove that $X \boxtimes Y$ is isomorphic to the product $X \times Y$.