Saunders Mac Lane

Categories for the Working Mathematician

Second Edition

Springer

Saunders Mac Lane
Professor Emeritus
Department of Mathematics
University of Chicago
Chicago, IL 60637-1514
USA

Mathematics Subject Classification (1991): 18-01

Library of Congress Cataloging-in-Publication Data
Mac Lane, Saunders, 1909–
Categories for the working mathematician/Saunders Mac Lane.—
2nd ed.
p. cm.—(Graduate texts in mathematics; 5)
Includes bibliographical references and index.
ISBN 0-387-98403-8 (hardcover: alk. paper)
1. Categories (Mathematics) I. Title. II. Series.
QA169.M33 1998
512′.55—dc21 97-45229

Printed on acid-free paper.

Production managed by Francine McNeill; manufacturing supervised by Thomas King.
Typeset by Asco Trade Typesetting Ltd., Hong Kong.
Printed and bound by Braun-Brumfield, Inc., Ann Arbor, MI.
Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387-98403-8 Springer-Verlag New York Berlin Heidelberg SPIN 10661688

Contents

Introduction

Category theory starts with the observation that many properties of mathematical systems can be unified and simplified by a presentation with diagrams of arrows. Each arrow $f: X \to Y$ represents a function; that is, a set X, a set Y, and a rule $x \mapsto fx$ which assigns to each element $x \in X$ an element $fx \in Y$; whenever possible we write fx and not $f(x)$, omitting unnecessary parentheses. A typical diagram of sets and functions is

$$\begin{tikzcd} & Y \arrow[dr, "g"] & \\ X \arrow[ur, "f"] \arrow[rr, "h"'] & & Z; \end{tikzcd}$$

it is commutative when h is $h = g \circ f$, where $g \circ f$ is the usual composite function $g \circ f : X \to Z$, defined by $x \mapsto g(fx)$. The same diagrams apply in other mathematical contexts; thus in the "category" of all topological spaces, the letters X, Y, and Z represent topological spaces while f, g, and h stand for continuous maps. Again, in the "category" of all groups, X, Y, and Z stand for groups, f, g, and h for homomorphisms.

Many properties of mathematical constructions may be represented by universal properties of diagrams. Consider the cartesian product $X \times Y$ of two sets, consisting as usual of all ordered pairs $\langle x, y \rangle$ of elements $x \in X$ and $y \in Y$. The projections $\langle x, y \rangle \mapsto x$, $\langle x, y \rangle \mapsto y$ of the product on its "axes" X and Y are functions $p: X \times Y \to X$, $q: X \times Y \to Y$. Any function $h: W \to X \times Y$ from a third set W is uniquely determined by its composites $p \circ h$ and $q \circ h$. Conversely, given W and two functions f and g as in the diagram below, there is a unique function h which makes the diagram commute; namely, $hw = \langle fw, gw \rangle$ for each w in W:

$$\begin{tikzcd} & W \arrow[dl, "f"'] \arrow[d, "h"] \arrow[dr, "g"] & \\ X & X \times Y \arrow[l, "p"] \arrow[r, "q"'] & Y. \end{tikzcd}$$

Thus, given X and Y, $\langle p, q\rangle$ is "universal" among pairs of functions from some set to X and Y, because any other such pair $\langle f, g\rangle$ factors uniquely (via h) through the pair $\langle p, q\rangle$. This property describes the cartesian product $X \times Y$ uniquely (up to a bijection); the same diagram, read in the category of topological spaces or of groups, describes uniquely the cartesian product of spaces or the direct product of groups.

Adjointness is another expression for these universal properties. If we write $\hom(W, X)$ for the set of all functions $f: W \to X$ and $\hom(\langle U, V\rangle, \langle X, Y\rangle)$ for the set of all pairs of functions $f: U \to X$, $g: V \to Y$, the correspondence $h \mapsto \langle ph, qh\rangle = \langle f, g\rangle$ indicated in the diagram above is a bijection

$$\hom(W, X \times Y) \cong \hom(\langle W, W\rangle, \langle X, Y\rangle).$$

This bijection is "natural" in the sense (to be made more precise later) that it is defined in "the same way" for all sets W and for all pairs of sets $\langle X, Y\rangle$ (and it is likewise "natural" when interpreted for topological spaces or for groups). This natural bijection involves two constructions on sets: The construction $W \mapsto W, W$ which sends each set to the diagonal pair $\Delta W = \langle W, W\rangle$, and the construction $\langle X, Y\rangle \mapsto X \times Y$ which sends each pair of sets to its cartesian product. Given the bijection above, we say that the construction $X \times Y$ is a *right adjoint* to the construction Δ, and that Δ is left adjoint to the product. Adjoints, as we shall see, occur throughout mathematics.

The construction "cartesian product" is called a "functor" because it applies suitably to sets *and* to the functions between them; two functions $k: X \to X'$ and $t: Y \to Y'$ have a function $k \times t$ as their cartesian product:

$$k \times t: X \times Y \to X' \times Y', \qquad \langle x, y\rangle \mapsto \langle kx, ty\rangle.$$

Observe also that the one-point set $1 = \{0\}$ serves as an identity under the operation "cartesian product", in view of the bijections

$$1 \times X \xrightarrow{\lambda} X \xleftarrow{\varrho} X \times 1 \tag{1}$$

given by $\lambda\langle 0, x\rangle = x$, $\varrho\langle x, 0\rangle = x$.

The notion of a monoid (a semigroup with identity) plays a central role in category theory. A monoid M may be described as a set M together with two functions

$$\mu: M \times M \to M, \qquad \eta: 1 \to M \tag{2}$$

such that the following two diagrams in μ and η commute:

$$\begin{tikzcd} M \times M \times M \arrow[r, "1 \times \mu"] \arrow[d, "\mu \times 1"'] & M \times M \arrow[d, "\mu"] \\ M \times M \arrow[r, "\mu"] & M, \end{tikzcd} \qquad \begin{tikzcd} 1 \times M \arrow[r, "\eta \times 1"] \arrow[d, "\lambda"] & M \times M \arrow[d, "\mu"] & M \times 1 \arrow[l, "1 \times \eta"'] \arrow[d, "\varrho"] \\ M \arrow[r, equal] & M \arrow[r, equal] & M; \end{tikzcd} \tag{3}$$

here 1 in $1 \times \mu$ is the identity function $M \to M$, and 1 in $1 \times M$ is the one-point set $1 = \{0\}$, while λ and ϱ are the bijections of (1) above. To say that these diagrams commute means that the following composites are equal:

$$\mu \circ (1 \times \mu) = \mu \circ (\mu \times 1)\,, \quad \mu \circ (\eta \times 1) = \lambda\,, \quad \mu \circ (1 \times \eta) = \varrho\,.$$

These diagrams may be rewritten with elements, writing the function μ (say) as a product $\mu(x, y) = xy$ for $x, y \in M$ and replacing the function η on the one-point set $1 = \{0\}$ by its (only) value, an element $\eta(0) = u \in M$. The diagrams above then become

$$\begin{tikzcd}
\langle x, y, z\rangle \arrow[r, mapsto] \arrow[d, mapsto] & \langle x, yz\rangle \arrow[d, mapsto] \\
\langle xy, z\rangle \arrow[r, mapsto] & (xy)z = x(yz)\,,
\end{tikzcd}
\quad
\begin{tikzcd}
\langle 0, x\rangle \arrow[r, mapsto] \arrow[d, mapsto] & \langle u, x\rangle \arrow[d, mapsto] \\
x \arrow[r, phantom, "="] & ux\,,
\end{tikzcd}
\quad
\begin{tikzcd}
\langle x, u\rangle \arrow[d, mapsto] & \langle x, 0\rangle \arrow[l, mapsto] \arrow[d, mapsto] \\
xu \arrow[r, phantom, "="] & x\,.
\end{tikzcd}$$

They are exactly the familiar axioms on a monoid, that the multiplication be associative and have an element u as left and right identity. This indicates, conversely, how algebraic identities may be expressed by commutative diagrams. The same process applies to other identities; for example, one may describe a group as a monoid M equipped with a function $\zeta : M \to M$ (of course, the function $x \mapsto x^{-1}$) such that the following diagram commutes:

$$\begin{tikzcd}
M \arrow[r, "\delta"] \arrow[d] & M \times M \arrow[r, "1 \times \zeta"] & M \times M \arrow[d, "\mu"] \\
1 \arrow[rr] & & M
\end{tikzcd}
\quad
\begin{tikzcd}
x \arrow[r, mapsto] \arrow[d, mapsto] & \langle x, x\rangle \arrow[r, mapsto] & \langle x, x^{-1}\rangle \arrow[d, mapsto] \\
0 \arrow[r, mapsto] & u \arrow[r, phantom, "="] & xx^{-1}\,,
\end{tikzcd}
\tag{4}$$

here $\delta : M \to M \times M$ is the diagonal function $x \mapsto \langle x, x\rangle$ for $x \in M$, while the unnamed vertical arrow $M \to 1 = \{0\}$ is the evident (and unique) function from M to the one-point set. As indicated just to the right, this diagram does state that ζ assigns to each element $x \in M$ an element x^{-1} which is a right inverse to x.

This definition of a group by arrows μ, η, and ζ in such commutative diagrams makes no explicit mention of group elements, so applies to other circumstances. If the letter M stands for a topological space (not just a set) and the arrows are continuous maps (not just functions), then the conditions (3) and (4) define a topological group – for they specify that M is a topological space with a binary operation μ of multiplication which is continuous (simultaneously in its arguments) and which has a continuous right inverse, all satisfying the usual group axioms. Again, if the letter M stands for a differentiable manifold (of class C^∞) while 1 is the one-point manifold and the arrows μ, η, and ζ are smooth mappings of manifolds, then the diagrams (3) and (4) become the definition of a Lie group. Thus groups, topological groups, and Lie groups can all be described as "diagrammatic" groups in the respective categories of sets, of topological spaces, and of differentiable manifolds.

This definition of a group in a category depended (for the inverse in (4)) on the diagonal map $\delta : M \to M \times M$ to the cartesian square $M \times M$. The definition of a monoid is more general, because the cartesian product $\times$ in $M \times M$ may be replaced by any other operation $\Box$ on two objects which is associative and which has a unit 1 in the sense prescribed by the isomorphisms (1). We can then speak of a monoid in the system $(C, \Box, 1)$, where C is the category, $\Box$ is such an operation, and 1 is its unit. Consider, for example, a monoid M in $(\mathbf{Ab}, \otimes, \mathbf{Z})$, where $\mathbf{Ab}$ is the category of abelian groups, $\times$ is replaced by the usual tensor product of abelian groups, and 1 is replaced by $\mathbf{Z}$, the usual additive group of integers; then (1) is replaced by the familiar isomorphism

$$\mathbf{Z} \otimes X \cong X \cong X \otimes \mathbf{Z}, \qquad X \text{ an abelian group}.$$

Then a monoid M in $(\mathbf{Ab}, \otimes, \mathbf{Z})$ is, we claim, simply a ring. For the given morphism $\mu : M \otimes M \to M$ is, by the definition of $\otimes$, just a function $M \times M \to M$, call it multiplication, which is bilinear; i.e., distributive over addition on the left and on the right, while the morphism $\eta : \mathbf{Z} \to M$ of abelian groups is completely determined by picking out one element of M; namely, the image u of the generator 1 of $\mathbf{Z}$. The commutative diagrams (3) now assert that the multiplication μ in the abelian group M is associative and has u as left and right unit – in other words, that M is indeed a ring (with identity $=$ unit).

The (homo)-morphisms of an algebraic system can also be described by diagrams. If $\langle M, \mu, \eta \rangle$ and $\langle M', \mu', \eta' \rangle$ are two monoids, each described by diagrams as above, then a morphism of the first to the second may be defined as a function $f : M \to M'$ such that the following diagrams commute:

$$\begin{tikzcd} M \arrow[d, "f"] \\ M', \end{tikzcd} \qquad \begin{tikzcd} M \times M \arrow[r, "\mu"] \arrow[d, "f \times f"] & M \arrow[d, "f"] \\ M' \times M' \arrow[r, "\mu'"] & M', \end{tikzcd} \qquad \begin{tikzcd} 1 \arrow[r, "\eta"] \arrow[d, equal] & M \arrow[d, "f"] \\ 1 \arrow[r, "\eta'"] & M'. \end{tikzcd} \tag{5}$$

In terms of elements, this asserts that $f(xy) = (fx)(fy)$ and $fu = u'$, with u and u' the unit elements; thus a homomorphism is, as usual, just a function preserving composite and units. If M and M' are monoids in $(\mathbf{Ab}, \otimes, \mathbf{Z})$, that is, rings, then a homomorphism f as here defined is just a morphism of rings (preserving the units).

Finally, an *action* of a monoid $\langle M, \mu, \eta\rangle$ on a set S is defined to be a function $\nu : M \times S \to S$ such that the following two diagrams commute:

$$\begin{tikzcd} M \times M \times S \arrow[r, "1 \times \nu"] \arrow[d, "\mu \times 1"'] & M \times S \arrow[d, "\nu"] \\ M \times S \arrow[r, "\nu"] & S, \end{tikzcd} \qquad \begin{tikzcd} 1 \times S \arrow[r, "\eta \times 1"] \arrow[dr, "\lambda"'] & M \times S \arrow[d, "\nu"] \\ & S. \end{tikzcd}$$

If we write $\nu(x, s) = x \cdot s$ to denote the result of the action of the monoid element x on the element $s \in S$, these diagrams state just that

$$x \cdot (y \cdot s) = (xy) \cdot s\,, \quad u \cdot s = s$$

for all $x, y \in M$ and all $s \in S$. These are the usual conditions for the action of a monoid on a set, familiar especially in the case of a group acting on a set as a group of transformations. If we shift from the category of sets to the category of topological spaces, we get the usual continuous action of a topological monoid M on a topological space S. If we take $\langle M, \mu, \eta\rangle$ to be a monoid in (**Ab**, $\otimes$, **Z**), then an action of M on an object S of **Ab** is just a left module S over the ring M.

I. Categories, Functors, and Natural Transformations

1. Axioms for Categories

First we describe categories directly by means of axioms, without using any set theory, and call them "metacategories". Actually, we begin with a simpler notion, a (meta)graph.

A *metagraph* consists of *objects* $a, b, c, \ldots$, *arrows* $f, g, h, \ldots$, and two operations, as follows:

Domain, which assigns to each arrow f an object $a = \operatorname{dom} f$;
Codomain, which assigns to each arrow f an object $b = \operatorname{cod} f$.

These operations on f are best indicated by displaying f as an actual arrow starting at its domain (or "source") and ending at its codomain (or "target"):

$$f : a \to b \quad \text{or} \quad a \xrightarrow{f} b .$$

A finite graph may be readily exhibited: Thus $\cdot \to \cdot \to \cdot$ or $\cdot \rightrightarrows \cdot$.

A *metacategory* is a metagraph with two additional operations:
Identity, which assigns to each object a an arrow $\mathrm{id}_a = 1_a : a \to a$;
Composition, which assigns to each pair $\langle g, f \rangle$ of arrows with $\operatorname{dom} g = \operatorname{cod} f$ an arrow $g \circ f$, called their *composite*, with $g \circ f : \operatorname{dom} f \to \operatorname{cod} g$. This operation may be pictured by the diagram

$$\begin{tikzcd} & b \arrow[dr, "g"] & \\ a \arrow[ur, "f"] \arrow[rr, "g \circ f"'] & & c \end{tikzcd}$$

which exhibits all domains and codomains involved. These operations in a metacategory are subject to the two following axioms:

Associativity. For given objects and arrows in the configuration

$$a \xrightarrow{f} b \xrightarrow{g} c \xrightarrow{k} d$$

one always has the equality

$$k \circ (g \circ f) = (k \circ g) \circ f . \tag{1}$$

This axiom asserts that the associative law holds for the operation of composition whenever it makes sense (i.e., whenever the composites on either side of (1) are defined). This equation is represented pictorially by the statement that the following diagram is commutative:

$$\begin{tikzcd}
a \arrow[rr, "k\circ(g\circ f)=(k\circ g)\circ f"] \arrow[d, "f"'] \arrow[drr, "g\circ f" near start] & & d \\
b \arrow[rr, "g"'] \arrow[urr, "k\circ g"' near end] & & c . \arrow[u, "k"']
\end{tikzcd}$$

Unit law. For all arrows $f: a \to b$ and $g: b \to c$ composition with the identity arrow 1_b gives

$$1_b \circ f = f \quad \text{and} \quad g \circ 1_b = g . \tag{2}$$

This axiom asserts that the identity arrow 1_b of each object b acts as an identity for the operation of composition, whenever this makes sense. The Eqs. (2) may be represented pictorially by the statement that the following diagram is commutative:

$$\begin{tikzcd}
a \arrow[r, "f"] \arrow[dr, "f"'] & b \arrow[d, "1_b"] \arrow[dr, "g"] & \\
& b \arrow[r, "g"'] & c .
\end{tikzcd}$$

We use many such diagrams consisting of vertices (labelled by objects of a category) and edges (labelled by arrows of the same category). Such a diagram is *commutative* when, for each pair of vertices c and c', any two paths formed from directed edges leading from c to c' yield, by composition of labels, equal arrows from c to c'. A considerable part of the effectiveness of categorical methods rests on the fact that such diagrams in each situation vividly represent the actions of the arrows at hand.

If b is any object of a metacategory C, the corresponding identity arrow 1_b is uniquely determined by the properties (2). For this reason, it is sometimes convenient to identify the identity arrow 1_b with the object b itself, writing $b: b \to b$. Thus $1_b = b = \mathrm{id}_b$, as may be convenient.

A metacategory is to be any interpretation which satisfies all these axioms. An example is the *metacategory of sets*, which has objects all sets and arrows all functions, with the usual identity functions and the usual composition of functions. Here "function" means a function with specified domain and specified codomain. Thus a function $f: X \to Y$ consists of a set X, its domain, a set Y, its codomain, and a rule $x \mapsto fx$ (i.e., a suitable set of ordered pairs $\langle x, fx \rangle$) which assigns, to each element $x \in X$, an element $fx \in Y$. These values will be written as fx, f_x, or $f(x)$, as may be convenient. For example, for any set S, the assignment $s \mapsto s$ for all $s \in S$ describes the *identity function* $1_S : S \to S$; if S is a subset of Y, the assignment $s \mapsto s$ also describes the *inclusion* or *insertion function* $S \to Y$; these functions are *different* unless $S = Y$. Given functions $f : X \to Y$ and $g : Y \to Z$, the *composite* function $g \circ f : X \to Z$ is defined by $(g \circ f)x = g(fx)$ for all $x \in X$. Observe that $g \circ f$ will mean first apply f, then g – in keeping with the practice of writing each function f to the left of its argument. Note, however, that many authors use the opposite convention.

To summarize, the metacategory of all sets has as objects, all sets, as arrows, all functions with the usual composition. The metacategory of all groups is described similarly: Objects are all groups G, H, K; arrows are all those functions f from the set G to the set H for which $f : G \to H$ is a homomorphism of groups. There are many other metacategories: All topological spaces with continuous functions as arrows; all compact Hausdorff spaces with the same arrows; all ringed spaces with their morphisms, etc. The arrows of any metacategory are often called its *morphisms*.

Since the objects of a metacategory correspond exactly to its identity arrows, it is technically possible to dispense altogether with the objects and deal only with arrows. The data for an *arrows-only metacategory* C consist of arrows, certain ordered pairs $\langle g, f \rangle$, called the composable pairs of arrows, and an operation assigning to each composable pair $\langle g, f \rangle$ an arrow $g \circ f$, called their composite. We say "$g \circ f$ is defined" for "$\langle g, f \rangle$ is a composable pair".

With these data one *defines* an identity of C to be an arrow u such that $f \circ u = f$ whenever the composite $f \circ u$ is defined *and* $u \circ g = g$ whenever $u \circ g$ is defined. The data are then required to satisfy the following three axioms:

(i) The composite $(k \circ g) \circ f$ is defined if and only if the composite $k \circ (g \circ f)$ is defined. When either is defined, they are equal (and this *triple composite* is written as kgf).

(ii) The triple composite kgf is defined whenever both composites kg and gf are defined.

(iii) For each arrow g of C there exist identity arrows u and u' of C such that $u' \circ g$ and $g \circ u$ are defined.

In view of the explicit definition given above for identity arrows, the last axiom is a quite powerful one; it implies that u' and u are unique in (iii), and it gives for each arrow g a codomain u' and a domain u. These axioms are equivalent to the preceding ones. More explicitly, given a metacategory of objects and arrows, its arrows, with the given composition, satisfy the "arrows-only" axioms; conversely, an arrows-only metacategory satisfies the objects-and-arrows axioms when the identity arrows, defined as above, are taken as the objects (Proof as exercise).

2. Categories

A category (as distinguished from a metacategory) will mean any interpretation of the category axioms within set theory. Here are the details. A *directed graph* (also called a "diagram scheme") is a set O of objects, a set A of arrows, and two functions

$$A \underset{\text{cod}}{\overset{\text{dom}}{\rightrightarrows}} O. \tag{1}$$

In this graph, the set of composable pairs of arrows is the set

$$A \times_O A = \{\langle g, f \rangle \mid g, f \in A \quad \text{and} \quad \operatorname{dom} g = \operatorname{cod} f\},$$

called the "product over O".

A *category* is a graph with two additional functions

$$\begin{array}{ll} O \xrightarrow{\text{id}} A, & A \times_O A \xrightarrow{\circ} A, \\ c \longmapsto \mathrm{id}_c, & \langle g, f \rangle \longmapsto g \circ f, \end{array} \tag{2}$$

called identity and composition also written as gf, such that

$$\operatorname{dom}(\operatorname{id} a) = a = \operatorname{cod}(\operatorname{id} a), \quad \operatorname{dom}(g \circ f) = \operatorname{dom} f, \quad \operatorname{cod}(g \circ f) = \operatorname{cod} g \tag{3}$$

for all objects $a \in O$ and all composable pairs of arrows $\langle g, f \rangle \in A \times_O A$, and such that the associativity and unit axioms (1.1) and (1.2) hold. In treating a category C, we usually drop the letters A and O, and write

$$c \in C \qquad f \text{ in } C \tag{4}$$

for "c is an object of C" and "f is an arrow of C", respectively. We also write

$$\hom(b, c) = \{f \mid f \text{ in } C, \ \operatorname{dom} f = b, \ \operatorname{cod} f = c\} \tag{5}$$

for the set of arrows from b to c. Categories can be defined directly in terms of composition acting on these "hom-sets" (§ 8 below); we do not follow this custom because we put the emphasis not on sets (a rather special category), but on axioms, arrows, and diagrams of arrows. We will later observe that our definition of a category amounts to saying that a category is a monoid for the product $\times_O$, in the general sense described in the introduction. For the moment, we consider examples.

0 is the empty category (no objects, no arrows);
1 is the category $\bullet$ with one object and one (identity) arrow;
2 is the category $\bullet \to \bullet$ with two objects a, b, and just one arrow $a \to b$ not the identity;

3 is the category with three objects whose non-identity arrows are arranged as in the triangle
$$\begin{tikzcd} & \cdot \arrow[dr] & \\ \cdot \arrow[ur] \arrow[rr] & & \cdot \end{tikzcd}$$
;

$\downdownarrows$ is the category with two objects a, b and just two arrows $a \rightrightarrows b$ not the identity arrows. We call two such arrows *parallel arrows*.

In each of the cases above there is only one possible definition of composition.

Discrete Categories. A category is *discrete* when every arrow is an identity. Every set X is the set of objects of a discrete category (just add one identity arrow $x \to x$ for each $x \in X$), and every discrete category is so determined by its set of objects. Thus, discrete categories are sets.

Monoids. A monoid is a category with one object. Each monoid is thus determined by the set of all its arrows, by the identity arrow, and by the rule for the composition of arrows. Since any two arrows have a composite, a monoid may then be described as a set M with a binary operation $M \times M \to M$ which is associative and has an identity ($=$ unit). Thus a monoid is exactly a semigroup with identity element. For any category C and any object $a \in C$, the set $\hom(a, a)$ of all arrows $a \to a$ is a monoid.

Groups. A group is a category with one object in which every arrow has a (two-sided) inverse under composition.

Matrices. For each commutative ring K, the set $\mathbf{Matr}_K$ of all rectangular matrices with entries in K is a category; the objects are all positive integers $m, n, \ldots$, and each $m \times n$ matrix A is regarded as an arrow $A : n \to m$, with composition the usual matrix product.

Sets. If V is any set of sets, we take $\mathbf{Ens}_V$ to be the category with objects all sets $X \in V$, arrows *all* functions $f : X \to Y$, with the usual composition of functions. By **Ens** we mean any one of these categories.

Preorders. By a preorder we mean a category P in which, given objects p and p', there is at most one arrow $p \to p'$. In any preorder P, define a binary relation $\leqq$ on the objects of P with $p \leqq p'$ if and only if there is an arrow $p \to p'$ in P. This binary relation is reflexive (because there is an identity arrow $p \to p$ for each p) and transitive (because arrows can be composed). Hence a preorder is a set (of objects) equipped with a reflexive and transitive binary relation. Conversely, any set P with such a relation determines a preorder, in which the arrows $p \to p'$ are exactly those ordered pairs $\langle p, p' \rangle$ for which $p \leqq p'$. Since the relation is transitive, there is a unique way of composing these arrows; since it is reflexive, there are the necessary identity arrows.

Preorders include *partial orders* (preorders with the added axiom that $p \leqq p'$ and $p' \leqq p$ imply $p = p'$) and *linear orders* (partial orders such that, given p and p', either $p \leqq p'$ or $p' \leqq p$).

Ordinal Numbers. We regard each ordinal number n as the linearly ordered set of all the preceding ordinals $n = \{0, 1, \ldots, n-1\}$; in particular, 0 is the empty set, while the first infinite ordinal is $\omega = \{0, 1, 2, \ldots\}$. Each ordinal n is linearly ordered, and hence is a category (a preorder). For example, the categories **1**, **2**, and **3** listed above are the preorders belonging to the (linearly ordered) ordinal numbers 1, 2, and 3. Another example is the linear order ω. As a category, it consists of the arrows

$$0 \to 1 \to 2 \to 3 \to \cdots,$$

all their composites, and the identity arrows for each object.

Δ is the category with objects all finite ordinals and arrows $f: m \to n$ all order-preserving functions ($i \leqq j$ in m implies $f_i \leqq f_j$ in n). This category Δ, sometimes called the *simplicial category*, plays a central role (Chapter VII).

Finord $=$ **Set**$_\omega$ is the category with objects all finite ordinals n and arrows $f: m \to n$ all functions from m to n. This is essentially the category of all finite sets, using just one finite set n for each finite cardinal number n.

Large Categories. In addition to the metacategory of all sets – which is not a set – we want an actual category **Set**, the category of all *small* sets. We shall assume that there is a big enough set U, the "universe", then describe a set x as "small" if it is a member of the universe, and take **Set** to be the category whose set U of objects is the set of all small sets, with arrows all functions from one small set to another. With this device (details in § 7 below) we construct other familiar large categories, as follows:

Set: Objects, all small sets; arrows, all functions between them.

Set$_*$: Pointed sets: Objects, small sets each with a selected base point; arrows, base-point-preserving functions.

Ens: Category of all sets and functions within a (variable) set V.

Cat: Objects, all small categories; arrows, all functors (§ 3).

Mon: Objects, all small monoids; arrows, all morphisms of monoids.

Grp: Objects, all small groups; arrows, all morphisms of groups.

Ab: Objects, all small (additive) abelian groups, with morphisms of such.

Rng: All small rings, with the ring morphisms (preserving units) between them.

CRng: All small commutative rings and their morphisms.

R-**Mod**: All small left modules over the ring R, with linear maps.

Mod-R: Small right R-modules.

K-**Mod**: Small modules over the commutative ring K.

Top: Small topological spaces and continuous maps.

Toph: Topological spaces, with arrows homotopy classes of maps.

Top$_*$: Spaces with selected base point, base point-preserving maps.

Particular categories (like these) will always appear in bold-face type. Script capitals are used by many authors to denote categories.

3. Functors

A *functor* is a morphism of categories. In detail, for categories C and B a functor $T: C \to B$ with domain C and codomain B consists of two suitably related functions: The *object function* T, which assigns to each object c of C an object Tc of B and the *arrow function* (also written T) which assigns to each arrow $f: c \to c'$ of C an arrow $Tf: Tc \to Tc'$ of B, in such a way that

$$T(1_c) = 1_{Tc}, \qquad T(g \circ f) = Tg \circ Tf, \tag{1}$$

the latter whenever the composite $g \circ f$ is defined in C. A functor, like a category, can be described in the "arrows-only" fashion: It is a function T from arrows f of C to arrows Tf of B, carrying each identity of C to an identity of B and each composable pair $\langle g, f \rangle$ in C to a composable pair $\langle Tg, Tf \rangle$ in B, with $Tg \circ Tf = T(g \circ f)$.

A simple example is the power set functor $\mathscr{P}: \mathbf{Set} \to \mathbf{Set}$. Its object function assigns to each set X the usual power set $\mathscr{P}X$, with elements all subsets $S \subset X$; its arrow function assigns to each $f: X \to Y$ that map $\mathscr{P}f: \mathscr{P}X \to \mathscr{P}Y$ which sends each $S \subset X$ to its image $fS \subset Y$. Since both $\mathscr{P}(1_X) = 1_{\mathscr{P}X}$ and $\mathscr{P}(g \;\; f) = \mathscr{P}g \;\; \mathscr{P}f$, this clearly defines a functor $\mathscr{P}: \mathbf{Set} \to \mathbf{Set}$.

Functors were first explicitly recognized in algebraic topology, where they arise naturally when geometric properties are described by means of algebraic invariants. For example, singular homology in a given dimension n (n a natural number) assigns to each topological space X an abelian group $H_n(X)$, the n-th homology group of X, and also to each continuous map $f: X \to Y$ of spaces a corresponding homomorphism $H_n(f): H_n(X) \to H_n(Y)$ of groups, and this in such a way that H_n becomes a functor $\mathbf{Top} \to \mathbf{Ab}$. For example, if $X = Y = S^1$ is the circle, $H_1(S^1) = \mathbf{Z}$, so the group homomorphism $H_1(f): \mathbf{Z} \to \mathbf{Z}$ is determined by an integer d (the image of 1); this integer is the usual "degree" of the continuous map $f: S^1 \to S^1$. In this case and in general, homotopic maps $f, g: X \to Y$ yield the same homomorphism $H_n(X) \to H_n(Y)$, so H_n can actually be regarded as a functor $\mathbf{Toph} \to \mathbf{Grp}$, defined on the homotopy category. The Eilenberg-Steenrod axioms for homology start with the axioms that H_n, for each natural number n, is a functor on $\mathbf{Toph}$, and continue with certain additional properties of these functors. The more recently developed extraordinary homology and cohomology theories are also functors on $\mathbf{Toph}$. The homotopy groups $\pi_n(X)$ of a space X can also be regarded as functors; since they depend on the choice of a base point in X, they are functors $\mathbf{Top}_* \to \mathbf{Grp}$. The leading idea in the use of functors in topology is that H_n or π_n gives an algebraic picture or image not just of the topological spaces, but also of all the continuous maps between them.

Functors arise naturally in algebra. To any commutative ring K the set of all non-singular $n \times n$ matrices with entries in K is the usual general linear group $\mathrm{GL}_n(K)$; moreover, each homomorphism $f : K \to K'$ of rings produces in the evident way a homomorphism $\mathrm{GL}_n f : \mathrm{GL}_n(K) \to \mathrm{GL}_n(K')$ of groups. These data define for each natural number n a functor $\mathrm{GL}_n : \mathbf{CRng} \to \mathbf{Grp}$. For any group G the set of all products of commutators $xyx^{-1}y^{-1}(x, y \in G)$ is a normal subgroup $[G, G]$ of G, called the *commutator* subgroup. Since any homomorphism $G \to H$ of groups carries commutators to commutators, the assignment $G \mapsto [G, G]$ defines an evident functor $\mathbf{Grp} \to \mathbf{Grp}$, while $G \mapsto G/[G, G]$ defines a functor $\mathbf{Grp} \to \mathbf{Ab}$, the factor-commutator functor. Observe, however, that the center $Z(G)$ of G (all $a \in G$ with $ax = xa$ for all x) does not naturally define a functor $\mathbf{Grp} \to \mathbf{Grp}$, because a homomorphism $G \to H$ may carry an element in the center of G to one not in the center of H.

A functor which simply "forgets" some or all of the structure of an algebraic object is commonly called a *forgetful* functor (or, an *underlying* functor). Thus the forgetful functor $U : \mathbf{Grp} \to \mathbf{Set}$ assigns to each group G the set UG of its elements ("forgetting" the multiplication and hence the group structure), and assigns to each morphism $f : G \to G'$ of groups the same function f, regarded just as a function between sets. The forgetful functor $U : \mathbf{Rng} \to \mathbf{Ab}$ assigns to each ring R the additive abelian group of R and to each morphism $f : R \to R'$ of rings the same function, regarded just as a morphism of addition.

Functors may be composed. Explicitly, given functors

$$C \xrightarrow{T} B \xrightarrow{S} A$$

between categories A, B, and C, the composite functions

$$c \mapsto S(Tc) \qquad f \mapsto S(Tf)$$

on objects c and arrows f of C define a functor $S \circ T : C \to A$, called the *composite* (in that order) of S with T. This composition is associative. For each category B there is an identity functor $I_B : B \to B$, which acts as an identity for this composition. Thus we may consider the metacategory of all categories: its objects are all categories, its arrows are all functors with the composition above. Similarly, we may form the category **Cat** of all small categories – but not the category of all categories.

An *isomorphism* $T : C \to B$ of categories is a functor T from C to B which is a bijection, both on objects and on arrows. Alternatively, but equivalently, a functor $T : C \to B$ is an isomorphism if and only if there is a functor $S : B \to C$ for which both composites $S \circ T$ and $T \circ S$ are identity functors; then S is the *two-sided inverse* $S = T^{-1}$.

Certain properties much weaker than isomorphism will be useful.

A functor $T : C \to B$ is *full* when to every pair c, c' of objects of C and to every arrow $g : Tc \to Tc'$ of B, there is an arrow $f : c \to c'$ of C with $g = Tf$. Clearly the composite of two full functors is a full functor.

A functor $T : C \to B$ is *faithful* (or an embedding) when to every pair c, c' of objects of C and to every pair $f_1, f_2 : c \to c'$ of parallel arrows of C the equality $Tf_1 = Tf_2 : Tc \to Tc'$ implies $f_1 = f_2$. Again, composites of faithful functors are faithful. For example, the forgetful functor **Grp**$\to$**Set** is faithful but not full and not a bijection on objects.

These two properties may be visualized in terms of hom-sets (see (2.5)). Given a pair of objects $c, c' \in C$, the arrow function of $T: C \to B$ assigns to each $f: c \to c'$ an arrow $Tf: Tc \to Tc'$ and so defines a function

$$T_{c,c'} : \hom(c, c') \to \hom(Tc, Tc'), \qquad f \mapsto Tf.$$

Then T is full when every such function is surjective, and faithful when every such function is injective. For a functor which is both full and faithful (i.e., "fully faithful"), every such function is a bijection, but this need not mean that the functor itself is an isomorphism of categories, for there may be objects of B not in the image of T.

A *subcategory* S of a category C is a collection of some of the objects and some of the arrows of C, which includes with each arrow f both the object $\operatorname{dom} f$ and the object $\operatorname{cod} f$, with each object s its identity arrow 1_s and with each pair of composable arrows $s \to s' \to s''$ their composite. These conditions ensure that these collections of objects and arrows themselves constitute a category S. Moreover, the injection (inclusion) map $S \to C$ which sends each object and each arrow of S to itself (in C) is a functor, the *inclusion functor*. This inclusion functor is automatically faithful. We say that S is a *full subcategory* of C when the inclusion functor $S \to C$ is full. A full subcategory, given C, is thus determined by giving just the set of its objects, since the arrows between any two of these objects s, s' are all morphisms $s \to s'$ in C. For example, the category $\mathbf{Set}_f$ of all finite sets is a full subcategory of the category **Set**.

Exercises

1. Show how each of the following constructions can be regarded as a functor: The field of quotients of an integral domain; the Lie algebra of a Lie group.
2. Show that functors $\mathbf{1} \to C$, $\mathbf{2} \to C$, and $\mathbf{3} \to C$ correspond respectively to objects, arrows, and composable pairs of arrows in C.
3. Interpret "functor" in the following special types of categories: (a) A functor between two preorders is a function T which is monotonic (i.e., $p \leqq p'$ implies $Tp \leqq Tp'$). (b) A functor between two groups (one-object categories) is a morphism of groups. (c) If G is a group, a functor $G \to \mathbf{Set}$ is a permutation representation of G, while $G \to \mathbf{Matr}_K$ is a matrix representation of G.
4. Prove that there is no functor **Grp**$\to$**Ab** sending each group G to its center (consider $S_2 \to S_3 \to S_2$, the symmetric groups).
5. Find two different functors T: **Grp**$\to$**Grp** with object function $T(G) = G$ the identity for every group G.

4. Natural Transformations

Given two functors $S, T: C \to B$, a *natural transformation* $\tau: S \dot{\to} T$ is a function which assigns to each object c of C an arrow $\tau_c = \tau c: Sc \to Tc$ of B in such a way that every arrow $f: c \to c'$ in C yields a diagram

$$\begin{tikzcd} c \arrow[d, "f"] & Sc \arrow[r, "\tau c"] \arrow[d, "Sf"'] & Tc \arrow[d, "Tf"] \\ c', & Sc' \arrow[r, "\tau c'"] & Tc' \end{tikzcd} \tag{1}$$

which is commutative. When this holds, we also say that $\tau_c: Sc \to Tc$ is *natural* in c. If we think of the functor S as giving a picture in B of (all the objects and arrows of) C, then a natural transformation τ is a set of arrows mapping (or, translating) the picture S to the picture T, with all squares (and parallelograms!) like that above commutative:

$$\begin{tikzcd} a \arrow[dr, "f"] \arrow[dd, "h"'] & \\ & b \arrow[dl, "g"] \\ c & \end{tikzcd} \qquad \begin{tikzcd} Sa \arrow[rrr, "\tau a"] \arrow[dr, "Sf"] \arrow[dd, "Sh"'] & & & Ta \arrow[dr, "Tf"] \arrow[dd] & \\ & Sb \arrow[rrr, "\tau b"] \arrow[dl, "Sg"] & & & Tb \arrow[dl, "Tg"] \\ Sc \arrow[rrr, "\tau c"'] & & & Tc & \end{tikzcd}$$

We call $\tau a, \tau b, \tau c, \ldots,$ the *components* of the natural transformation τ.

A natural transformation is often called a *morphism of functors*; a natural transformation τ with every component τc invertible in B is called a *natural equivalence* or better a *natural isomorphism*; in symbols, $\tau: S \cong T$. In this case, the inverses $(\tau c)^{-1}$ in B are the components of a natural isomorphism $\tau^{-1}: T \dot{\to} S$.

The determinant is a natural transformation. To be explicit, let $\det_K M$ be the determinant of the $n \times n$ matrix M with entries in the commutative ring K, while K^* denotes the group of units (invertible elements) of K. Thus M is non-singular when $\det_K M$ is a unit, and $\det_K$ is a morphism $\mathrm{GL}_n K \to K^*$ of groups (an arrow in **Grp**). Because the determinant is defined by the same formula for all rings K, each morphism $f: K \to K'$ of commutative rings leads to a commutative diagram

$$\begin{tikzcd} \mathrm{GL}_n K \arrow[r, "\det_K"] \arrow[d, "\mathrm{GL}_n f"'] & K^* \arrow[d, "f^*"] \\ \mathrm{GL}_n K' \arrow[r, "\det_{K'}"] & K'^*. \end{tikzcd} \tag{2}$$

This states that the transformation $\det: \mathrm{GL}_n \to (\ )^*$ is natural between two functors **CRng** $\to$ **Grp**.

For each group G the projection $p_G: G \to G/[G, G]$ to the factor-commutator group defines a transformation p from the identity functor on **Grp** to the factor-commutator functor $\mathbf{Grp} \to \mathbf{Ab} \to \mathbf{Grp}$. Moreover, p is natural, because each group homomorphism $f: G \to H$ defines the evident homomorphism f' for which the following diagram commutes:

$$\begin{array}{ccc} G & \xrightarrow{p_G} & G/[G,G] \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f'} \\ H & \xrightarrow{p_H} & H/[H,H]\,. \end{array} \tag{3}$$

The double character group yields a suggestive example in the category **Ab** of all abelian groups G. Let $D(G)$ denote the character group of G, so that $DG = \hom(G, \mathbf{R}/\mathbf{Z})$ is the set of all homomorphisms $t: G \to \mathbf{R}/\mathbf{Z}$ with the familiar group structure, where $\mathbf{R}/\mathbf{Z}$ is the additive group of real numbers modulo 1. Each arrow $f: G' \to G$ in **Ab** determines an arrow $Df: DG \to DG'$ (opposite direction!) in **Ab**, with $(Df)t = tf: G' \to \mathbf{R}/\mathbf{Z}$ for each t; for composable arrows, $D(g \circ f) = Df \circ Dg$. Because of this reversal, D is not a functor (it is a "contravariant" functor on **Ab** to **Ab**, see § II.2); however, the twice iterated character group $G \mapsto D(DG)$ and the identity $I(G) = G$ are both functors $\mathbf{Ab} \to \mathbf{Ab}$. For each group G there is a homomorphism

$$\tau_G: G \to D(DG)$$

obtained in a familiar way: To each $g \in G$ assign the function $\tau_G g: DG \to \mathbf{R}/\mathbf{Z}$ given for any character $t \in DG$ by $t \mapsto tg$; thus $(\tau_G g)t = t(g)$. One verifies at once that τ is a natural transformation $\tau: I \dot{\to} DD$; this statement is just a precise expression for the elementary observation that the definition of τ depends on no artificial choices of bases, generators, or the like. In case G is finite, τ_G is an isomorphism; thus, if we restrict all functors to the category $\mathbf{Ab}_f$ of finite abelian groups, τ is a natural isomorphism.

On the other hand, for each finite abelian group G there is an isomorphism $\sigma_G: G \cong DG$ of G to its character group, but this isomorphism depends on a representation of G as a direct product of cyclic groups and so cannot be natural. More explicitly, we can make D into a covariant functor $D': \mathbf{Ab}_{f,i} \to \mathbf{Ab}_{f,i}$ on the category $\mathbf{Ab}_{f,i}$ with objects all finite abelian groups and arrows all isomorphisms f between such groups, setting $D'G = DG$ and $D'f = Df^{-1}$. Then $\sigma_G: G \to D'G$ is a map $\sigma: I \to D'$ of functors $\mathbf{Ab}_{f,i} \to \mathbf{Ab}_{f,i}$, but it is not natural in the sense of our definition.

A parallel example is the familiar natural isomorphism of a finite-dimensional vector space to its double dual.

Another example of naturality arises when we compare the category **Finord** of all finite ordinal numbers n with the category $\mathbf{Set}_f$ of all finite sets (in some universe U). Every ordinal $n = \{0, 1, \dots, n-1\}$ is a finite set, so the inclusion S is a functor $S: \mathbf{Finord} \to \mathbf{Set}_f$. On the other hand, each finite set X determines an ordinal number $n = \# X$, the number of elements in X; we may choose for each X a bijection $\theta_X : X \to \# X$. For any function $f : X \to Y$ between finite sets we may then define a corresponding function $\# f : \# X \to \# Y$ between ordinals by $\# f = \theta_Y f \theta_X^{-1}$; this ensures that the diagram

$$\begin{tikzcd} X \arrow[r, "\theta_X"] \arrow[d, "f"] & \# X \arrow[d, "\# f"] \\ Y \arrow[r, "\theta_Y"] & \# Y \end{tikzcd}$$

will commute, and makes $\#$ a functor $\# : \mathbf{Set}_f \to \mathbf{Finord}$. If X is itself an ordinal number, we may take θ_X to be the identity. This ensures that the composite functor $\# \circ S$ is the identity functor I' of **Finord**. On the other hand, the composite $S \circ \#$ is not the identity functor $I : \mathbf{Set}_f \to \mathbf{Set}_f$, because it sends each finite set X to a special finite set – the ordinal number n with the same number of elements as X. However, the square diagram above does show that $\theta : I \dot{\to} S\#$ is a natural isomorphism. All told we have $I \cong S \circ \#$, $I' = \# \circ S$.

More generally, an *equivalence* between categories C and D is defined to be a pair of functors $S: C \to D$, $T: D \to C$ together with natural isomorphisms $I_C \cong T \circ S$, $I_D \cong S \circ T$. This example shows that this notion (to be examined in § IV.4) allows us to compare categories which are "alike" but of very different "sizes".

We shall use many other examples of naturality. As Eilenberg-Mac Lane first observed, "category" has been defined in order to be able to define "functor" and "functor" has been defined in order to be able to define "natural transformation".

Exercises

1. Let S be a fixed set, and X^S the set of all functions $h : S \to X$. Show that $X \mapsto X^S$ is the object function of a functor $\mathbf{Set} \to \mathbf{Set}$, and that evaluation $e_X : X^S \times S \dot{\to} X$, defined by $e(h, s) = h(s)$, the value of the function h at $s \in S$, is a natural transformation.
2. If H is a fixed group, show that $G \mapsto H \times G$ defines a functor $H \times - : \mathbf{Grp} \to \mathbf{Grp}$, and that each morphism $f : H \to K$ of groups defines a natural transformation $H \times - \dot{\to} K \times -$.
3. If B and C are groups (regarded as categories with one object each) and $S, T : B \to C$ are functors (homomorphisms of groups), show that there is a natural transformation $S \dot{\to} T$ if and only if S and T are conjugate; i.e., if and only if there is an element $h \in C$ with $Tg = h(Sg)h^{-1}$ for all $g \in B$.

4. For functors $S, T : C \to P$ where C is a category and P a preorder, show that there is a natural transformation $S \dot{\to} T$ (which is then unique) if and only if $Sc \leqq Tc$ for every object $c \in C$.
5. Show that every natural transformation $\tau : S \dot{\to} T$ defines a function (also called τ) which sends each arrow $f : c \to c'$ of C to an arrow $\tau f : Sc \to Tc'$ of B in such a way that $Tg \circ \tau f = \tau(gf) = \tau g \circ Sf$ for each composable pair $\langle g, f \rangle$. Conversely, show that every such function τ comes from a unique natural transformation with $\tau_c = \tau(1_c)$. (This gives an "arrows only" description of a natural transformation.)
6. Let F be a field. Show that the category of all finite-dimensional vector spaces over F (with morphisms all linear transformations) is equivalent to the category $\mathbf{Matr}_F$ described in § 2.

5. Monics, Epis, and Zeros

In categorical treatments many properties ordinarily formulated by means of elements (elements of a set or of a group) are instead formulated in terms of arrows. For example, instead of saying that a set X has just one element, one can say that for any other set Y there is exactly one function $Y \to X$. We now formulate a few more instances of such methods of "doing without elements".

An arrow $e : a \to b$ is *invertible* in C if there is an arrow $e' : b \to a$ in C with $e'e = 1_a$ and $ee' = 1_b$. If such an e' exists, it is unique, and is written as $e' = e^{-1}$. By the usual proof, $(e_1 e_2)^{-1} = e_2^{-1} e_1^{-1}$, provided the composite $e_1 e_2$ is defined and both e_1 and e_2 are invertible. Two objects a and b are *isomorphic* in the category C if there is an invertible arrow (an *isomorphism*) $e : a \to b$; we write $a \cong b$. The relation of isomorphism of objects is manifestly reflexive, symmetric, and transitive.

An arrow $m : a \to b$ is *monic* in C when for any two parallel arrows $f_1, f_2 : d \to a$ the equality $m \circ f_1 = m \circ f_2$ implies $f_1 = f_2$; in other words, m is monic if it can always be cancelled on the left (is *left cancellable*). In **Set** and in **Grp** the monic arrows are precisely the injections (monomorphisms) in the usual sense; i.e., the functions which are one-one into.

An arrow $h : a \to b$ is *epi* in C when for any two arrows $g_1, g_2 : b \to c$ the equality $g_1 \circ h = g_2 \circ h$ implies $g_1 = g_2$; in other words, h is epi when it is *right cancellable*. In **Set** the epi arrows are precisely the surjections (epimorphisms) in the usual sense; i.e., the functions onto.

For an arrow $h : a \to b$, a *right* inverse is an arrow $r : b \to a$ with $hr = 1_b$. A right inverse (which is usually not unique) is also called a *section* of h. If h has a right inverse, it is evidently epi; the converse holds in **Set**, but fails in **Grp**. Similarly, a *left inverse* for h is called a *retraction* for h, and any arrow with a left inverse is necessarily monic. If $gh = 1_a$, then g is a *split* epi, h a *split* monic, and the composite $f = hg$ is defined and is an idempotent. Generally, an arrow $f: b \to b$ is called *idempotent* when $f^2 = f$; an idempotent is said to *split* when there exist arrows g and h such that $f = hg$ and $gh = 1$.

An object t is *terminal* in C if to each object a in C there is exactly one arrow $a \to t$. If t is terminal, the only arrow $t \to t$ is the identity, and any two terminal objects of C are isomorphic in C. An object s is *initial* in C if to each object a there is exactly one arrow $s \to a$. For example, in the category **Set**, the empty set is an initial object and any one-point set is a terminal object. In **Grp**, the group with one element is both initial and terminal.

A *null object* z in C is an object which is both initial and terminal. If C has a null object, that object is unique up to isomorphism, while for any two objects a and b of C there is a unique arrow $a \to z \to b$ (the composite through z), called the *zero arrow* from a to b. Any composite with a zero arrow is itself a zero arrow. For example, the categories **Ab** and R-**Mod** have null objects (namely 0!), as does $\mathbf{Set}_*$ (namely the one-point set).

A *groupoid* is a category in which every arrow is invertible. A typical groupoid is the *fundamental groupoid* $\pi(X)$ of a topological space X. An object of $\pi(X)$ is a point x of X, and an arrow $x \to x'$ of $\pi(X)$ is a homotopy class of paths f from x to x'. (Such a path f is a continuous function $I \to X$, I the closed interval $I = [0, 1]$, with $f(0) = x$, $f(1) = x'$, while two paths f, g with the same end-points x and x' are homotopic when there is a continuous function $F: I \times I \to X$ with $F(t, 0) = f(t)$, $F(t, 1) = g(t)$, and $F(0, s) = x$, $F(1, s) = x'$ for all s and t in I.) The *composite* of paths $g: x' \to x''$ and $f: x \to x'$ is the path h which is "f followed by g", given explicitly by

$$\begin{aligned} h(t) &= f(2t), && 0 \leqq t \leqq 1/2, \\ &= g(2t-1), && 1/2 \leqq t \leqq 1. \end{aligned} \tag{1}$$

Composition applies also to homotopy classes, and makes $\pi(X)$ a category and a groupoid (the inverse of any path is the same path traced in the opposite direction).

Since each arrow in a groupoid G is invertible, each object x in G determines a group $\hom_G(x, x)$, consisting of all $g: x \to x$. If there is an arrow $f: x \to x'$, the groups $\hom_G(x, x)$ and $\hom_G(x', x')$ are isomorphic, under $g \mapsto fgf^{-1}$ (i.e., under conjugation). A groupoid is said to be *connected* if there is an arrow joining any two of its objects. One may readily show that a connected groupoid is determined up to isomorphism by a group (one of the groups $\hom_G(x, x)$) and by a set (the set of all objects). In this way, the fundamental groupoid $\pi(X)$ of a path-connected space X is determined by the set of points in the space and a group $\hom_{\pi(X)}(x, x)$ – the *fundamental group* of X.

Exercises

1. Find a category with an arrow which is both epi and monic, but not invertible (e.g., dense subset of a topological space).
2. Prove that the composite of monics is monic, and likewise for epis.
3. If a composite $g \circ f$ is monic, so is f. Is this true of g?
4. Show that the inclusion $\mathbf{Z} \to \mathbf{Q}$ is epi in the category **Rng**.
5. In **Grp** prove that every epi is surjective (Hint. If $\varphi : G \to H$ has image M not H, use the factor group H/M if M has index 2. Otherwise, let Perm H be the group of all permutations of the set H, choose three different cosets M, Mu and Mv of M, define $\sigma \in \operatorname{Perm} H$ by $\sigma(xu) = xv$, $\sigma(xv) = xu$ for $x \in M$, and σ otherwise the identity. Let $\psi : H \to \operatorname{Perm} H$ send each h to left multiplication ψ_h by h, while $\psi'_h = \sigma^{-1} \psi_h \sigma$. Then $\psi\varphi = \psi'\varphi$, but $\psi \neq \psi'$).
6. In **Set**, show that all idempotents split.
7. An arrow $f : a \to b$ in a category C is *regular* when there exists an arrow $g : b \to a$ such that $fgf = f$. Show that f is regular if it has either a left or a right inverse, and prove that every arrow in **Set** with $a \neq \emptyset$ is regular.
8. Consider the category with objects $\langle X, e, t\rangle$, where X is a set, $e \in X$, and $t : X \to X$, and with arrows $f : \langle X, e, t\rangle \to \langle X', e', t'\rangle$ the functions f on X to X' with $fe = e'$ and $ft = t'f$. Prove that this category has an initial object in which X is the set of natural numbers, $e = 0$, and t is the successor function.
9. If the functor $T : C \to B$ is faithful and Tf is monic, prove f monic.

6. Foundations

One of the main objectives of category theory is to discuss properties of totalities of Mathematical objects such as the "set" of all groups or the "set" of all homomorphisms between any two groups. Now it is the custom to regard a group as a set with certain added structure, so we are here proposing to consider a set of *all* sets with some given structure. This amounts to applying a comprehension principle: Given a property $\varphi(x)$ of sets x, form the set $\{x \mid \varphi(x)\}$ of *all* sets x with this property. However such a principle cannot be adopted in this generality, since it would lead to some of the famous paradoxical sets, such as the set of all sets not members of themselves.

For this reason, the standard practice in naive set theory, with the usual membership relation $\in$, is to restrict the application of the comprehension principle. One allows the formation from given sets u, v of the set $\{u, v\}$ (the set with exactly u and v as elements), of the ordered pair $\langle u, v\rangle$, of an infinite set (the set $\omega = \{0, 1, 2, \ldots\}$ of all finite ordinals), and of

The Cartesian Product	$u \times v = \{\langle x, y\rangle \mid x \in u \text{ and } y \in v\}$,
The Power Set	$\mathscr{P}u = \{v \mid v \subset u\}$,
The Union (of a set x of sets)	$\cup x = \{y \mid y \in z \text{ for some } z \in x\}$.