OXFORD LOGIC GUIDES

Series Editors

D.M. GABBAY
A.J. MACINTYRE
D.S. SCOTT

OXFORD LOGIC GUIDES

Available books in the series:

10. Michael Hallett: *Cantorian set theory and limitation of size*
17. Stewart Shapiro: *Foundations without foundationalism*
18. John P. Cleave: *A study of logics*
21. C. McLarty: *Elementary categories, elementary toposes*
22. R.M. Smullyan: *Recursion theory for metamathematics*
23. Peter Clote and Jan Krajíček: *Arithmetic, proof theory, and computational complexity*
24. A. Tarski: *Introduction to logic and to the methodology of deductive sciences*
25. G. Malinowski: *Many valued logics*
26. Alexandre Borovik and Ali Nesin: *Groups of finite Morley rank*
27. R.M. Smullyan: *Diagonalization and self-reference*
28. Dov M. Gabbay, Ian Hodkinson, and Mark Reynolds: *Temporal logic: Mathematical foundations and computational aspects: volume 1*
29. Saharon Shelah: *Cardinal arithmetic*
30. Erik Sandewall: *Features and fluents: Volume I: A Systematic approach to the representation of knowledge about dynamical systems*
31. T.E. Forster: *Set theory with a universal set: Exploring an untyped universe, second edition*
32. Anand Pillay: *Geometric stability theory*
33. Dov M. Gabbay: *Labelled deductive systems*
35. Alexander Chagrov and Michael Zakharyaschev: *Modal logic*
36. G. Sambin and J. Smith: *Twenty-five years of Martin-Löf constructive type theory*
37. María Manzano: *Model theory*
38. Dov M. Gabbay: *Fibring logics*
39. Michael Dummet: *Elements of Intuitionism, second edition*
40. D.M. Gabbay, M.A. Reynolds, and Marcelo Finger: *Temporal logic: Mathematical foundations and computational aspects volume 2*
41. J.M. Dunn and G. Hardegree: *Algebraic Methods in Philosophical Logic*
42. H. Rott: *Change, Choice and Inference: A study of belief revision and non-monotonic reasoning*
43. P.T. Johnstone: *Sketches of an Elephant: A topos theory compendium: Volume 1*
44. P.T. Johnstone: *Sketches of an Elephant: A topos theory compendium: Volume 2*
45. David J. Pym and Eike Ritter: *Reductive Logic and Proof Search: Proof theory, semantics and control*
46. D.M. Gabbay and L. Maksimova: *Interpolation and Definability: Modal and Intuitionistic Logics*
47. John L. Bell: *Set Theory: Boolean-valued models and independence proofs, third edition*
48. Laura Crosilla and Peter Schuster: *From Sets and Types to Topology and Analysis: Towards practicable foundations for constructive mathematics*
49. Steve Awodey: *Category Theory*

Category Theory

STEVE AWODEY
Carnegie Mellon University

CLARENDON PRESS • OXFORD
2006

OXFORD
UNIVERSITY PRESS
Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

© Steve Awodey, 2006

The moral rights of the author have been asserted
Database right Oxford University Press (maker)

First published 2006

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover
and you must impose the same condition on any acquirer

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging in Publication Data
Data available

Typeset by Newgen Imaging Systems (P) Ltd., Chennai, India
Printed in Great Britain
on acid-free paper by
Biddles Ltd., King's Lynn, Norfolk

ISBN 0–19–856861–4 978–0–19–856861–2

1 3 5 7 9 10 8 6 4 2

in memoriam
Saunders Mac Lane

PREFACE

Why write a new textbook on Category Theory, when we already have Mac Lane's *Categories for the Working Mathematician*? Simply put, because Mac Lane's book is for the working (and aspiring) mathematician. What is needed now, after 30 years of spreading into various other disciplines and places in the curriculum, is a book for everyone else.

This book has grown from my courses on Category Theory at Carnegie Mellon University over the last 10 years. In that time, I have given numerous lecture courses and advanced seminars to undergraduate and graduate students in Computer Science, Mathematics, and Logic. The lecture course based on the material in this book consists of two, 90-minute lectures a week for 15 weeks. The germ of these lectures was my own graduate student notes from a course on Category Theory given by Mac Lane at the University of Chicago. In teaching my own course, I soon discovered that the mixed group of students at Carnegie Mellon had very different needs than the Mathematics graduate students at Chicago and my search for a suitable textbook to meet these needs revealed a serious gap in the literature. My lecture notes evolved over a time to fill this gap, supplementing and eventually replacing the various texts I tried using.

The students in my courses often have little background in Mathematics beyond a course in Discrete Math and some Calculus or Linear Algebra or a course or two in Logic. Nonetheless, eventually, as researchers in Computer Science or Logic, many will need to be familiar with the basic notions of Category Theory, without the benefit of much further mathematical training. The Mathematics undergraduates are in a similar boat: mathematically talented, motivated to learn the subject by its evident relevance to their further studies, yet unable to follow Mac Lane because they still lack the mathematical prerequisites. Most of my students do not know what a free group is (yet), and so they are not illuminated to learn that it is an example of an adjoint.

This, then, is intended as a text and reference book on Category Theory, not only for students of Mathematics, but also for researchers and students in Computer Science, Logic, Linguistics, Cognitive Science, Philosophy, and any of the other fields that now make use of it. The challenge for me was to make the basic definitions, theorems, and proof techniques understandable to this readership, and thus without presuming familiarity with the main (or at least original) applications in algebra and topology. It will not do, however, to develop the subject in a vacuum, simply skipping the examples and applications. Material at this level of abstraction is simply incomprehensible without the applications and examples that bring it to life.

Faced with this dilemma, I have adopted the strategy of developing a few basic examples from scratch and in detail—namely posets and monoids—and

then carrying them along and using them throughout the book. This has several didactic advantages worth mentioning: both posets and monoids are themselves special kinds of categories, which in a certain sense represent the two "dimensions" (objects and arrows) that a general category has. Many phenomena occurring in categories can best be understood as generalizations from posets or monoids. On the other hand, the categories of posets (and monotone maps) and monoids (and homomorphisms) provide two further, quite different examples of categories in which to consider various concepts. The notion of a limit, for instance, can be considered both in a given poset and in the category of posets.

Of course, many other examples besides posets and monoids are treated as well. For example, the chapter on groups and categories develops the first steps of Group Theory up to kernels, quotient groups, and the homomorphism theorem, as an example of equalizers and coequalizers. Here, and occasionally elsewhere (e.g. in connection with Stone duality), I have included a bit more Mathematics than is strictly necessary to illustrate the concepts at hand. My thinking is that this may be the closest some students will ever get to a higher Mathematics course, so they should benefit from the labor of learning Category Theory by reaping some of the nearby fruits.

Although the mathematical prerequisites are substantially lighter than for Mac Lane, the standard of rigor has (I hope) not been compromised. Full proofs of all important propositions and theorems are given, and only occasional routine lemmas are left as exercises (and these are then usually listed as such at the end of the chapter). The selection of material was easy. There is a standard core that must be included: categories; functors; natural transformations; equivalence; limits and colimits; functor categories; representables; Yoneda's Lemma; adjoints; and monads. That nearly fills a course. The only "optional" topic included here is cartesian closed categories and the lambda-calculus, which is a must for computer scientists, logicians, and linguists. Several other obvious further topics were purposely not included: 2-categories, toposes (in any depth), and monoidal categories. These topics are treated in Mac Lane, which the student should be able to read after having completed the course.

Finally, I take this opportunity to thank Wilfried Sieg for his exceptional support of this project; Peter Johnstone and Dana Scott for helpful suggestions and support; André Carus for advice and encouragement; Bill Lawvere for many very useful comments on the text; and the many students in my courses who have suggested improvements to the text, clarified the content with their questions, tested all of the exercises, and caught countless errors and typos. For the latter, I also thank the many readers who took the trouble to collect and send helpful corrections, particularly Brighten Godfrey, Peter Gumm, Bob Lubarsky and Dave Perkinson. Andrej Bauer and Kohei Kishida are to be thanked for providing Figures 9.1 and 8.1, respectively. Of course, Paul Taylor's macros for commutative diagrams must also be acknowledged. And my dear Karin deserves thanks for too many things to mention. Finally, I wish to record here my debt of

gratitude to my mentor Saunders Mac Lane, not only for teaching me category theory, and trying to teach me how to write, but also for helping me to find my place in Mathematics. I dedicate this book to his memory.

Steve Awodey
Pittsburgh
September 2005

CONTENTS

Preface vi

1 Categories 1
- 1.1 Introduction 1
- 1.2 Functions of sets 3
- 1.3 Definition of a category 4
- 1.4 Examples of categories 5
- 1.5 Isomorphisms 11
- 1.6 Constructions on categories 13
- 1.7 Free categories 16
- 1.8 Foundations: large, small, and locally small 21
- 1.9 Exercises 23

2 Abstract structures 25
- 2.1 Epis and monos 25
- 2.2 Initial and terminal objects 28
- 2.3 Generalized elements 29
- 2.4 Sections and retractions 33
- 2.5 Products 34
- 2.6 Examples of products 36
- 2.7 Categories with products 41
- 2.8 Hom-sets 42
- 2.9 Exercises 45

3 Duality 47
- 3.1 The duality principle 47
- 3.2 Coproducts 49
- 3.3 Equalizers 54
- 3.4 Coequalizers 57
- 3.5 Exercises 63

4 Groups and categories 65
- 4.1 Groups in a category 65
- 4.2 The category of groups 68
- 4.3 Groups as categories 70
- 4.4 Finitely presented categories 73
- 4.5 Exercises 74

5 Limits and colimits 77
- 5.1 Subobjects 77
- 5.2 Pullbacks 80

5.3 Properties of pullbacks 84
5.4 Limits 89
5.5 Preservation of limits 94
5.6 Colimits 95
5.7 Exercises 102

6 Exponentials 105
6.1 Exponential in a category 105
6.2 Cartesian closed categories 108
6.3 Heyting algebras 113
6.4 Equational definition 118
6.5 λ-calculus 119
6.6 Exercises 123

7 Functors and naturality 125
7.1 Category of categories 125
7.2 Representable structure 127
7.3 Stone duality 131
7.4 Naturality 133
7.5 Examples of natural transformations 135
7.6 Exponentials of categories 139
7.7 Functor categories 142
7.8 Equivalence of categories 146
7.9 Examples of equivalence 150
7.10 Exercises 155

8 Categories of diagrams 159
8.1 Set-valued functor categories 159
8.2 The Yoneda embedding 160
8.3 The Yoneda Lemma 162
8.4 Applications of the Yoneda Lemma 166
8.5 Limits in categories of diagrams 167
8.6 Colimits in categories of diagrams 168
8.7 Exponentials in categories of diagrams 172
8.8 Topoi 174
8.9 Exercises 176

9 Adjoints 179
9.1 Preliminary definition 179
9.2 Hom-set definition 183
9.3 Examples of adjoints 187
9.4 Order adjoints 191
9.5 Quantifiers as adjoints 193
9.6 RAPL 197
9.7 Locally cartesian closed categories 202

9.8 Adjoint functor theorem 210
9.9 Exercises 219

10 Monads and algebras 223
10.1 The triangle identities 223
10.2 Monads and adjoints 225
10.3 Algebras for a monad 229
10.4 Comonads and coalgebras 234
10.5 Algebras for endofunctors 236
10.6 Exercises 244

References 249

Index 251

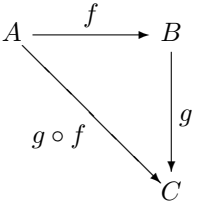
This page intentionally left blank

1

CATEGORIES

1.1 Introduction

What is category theory? As a first approximation, one could say that category theory is the mathematical study of (abstract) *algebras of functions.* Just as group theory is the abstraction of the idea of a system of permutations of a set or symmetries of a geometric object, category theory arises from the idea of a system of functions among some objects.



We think of the composition $g \circ f$ as a sort of "product" of the functions f and g, and consider abstract "algebras" of the sort arising from collections of functions. A category is just such an "algebra," consisting of objects $A, B, C, \ldots$ and arrows $f : A \to B$, $g : B \to C, \ldots$, that are closed under composition and satisfy certain conditions typical of the composition of functions. A precise definition is given later in this chapter.

A branch of abstract algebra, category theory was invented in the tradition of Felix Klein's *Erlanger Programm*, as a way of studying and characterizing different kinds of mathematical structures in terms of their "admissible transformations." The general notion of a category provides a characterization of the notion of a "structure-preserving transformation," and thereby of a species of structures admitting such transformations.

The historical development of the subject has been, very roughly, as follows:

1945 Eilenberg and Mac Lane's "General theory of natural equivalences" was the original paper, in which the theory was first formulated.

late 1940s The main applications were originally in the fields of algebraic topology, particularly homology theory, and abstract algebra.

1950s A. Grothendieck et al. began using category theory with great success in algebraic geometry.

1960s F.W. Lawvere and others began applying categories to logic, revealing some deep and surprising connections.

1970s Applications were already appearing in computer science, linguistics, cognitive science, philosophy, and many other areas.

One very striking thing about the field is that it has such wide-ranging applications. In fact, it turns out to be a kind of universal mathematical language like set theory. As a result of these various applications, category theory also tends to reveal certain connections between different fields—like logic and geometry. For example, the important notion of an *adjoint functor* occurs in logic as the existential quantifier and in topology as the image operation along a continuous function. From a categorical point of view these turn out to be essentially the same operation.

The concept of adjoint functor is in fact one of the main things that the reader should take away from the study of this book. It is a strictly category-theoretical notion that has turned out to be a conceptual tool of the first magnitude—on par with the idea of a continuous function.

In fact, just as the idea of a topological space arose in connection with continuous functions, so also the notion of a category arose in order to define that of a functor, at least according to one of the inventors. The notion of a functor arose—so the story goes on—in order to define natural transformations. One might as well continue that natural transformations serve to define adjoints:

Category
Functor
Natural transformation
Adjunction

Indeed, that gives a pretty good outline of this book.

Before getting down to business, let us ask why it should be that category theory has such far-reaching applications. Well, we said that it is the abstract theory of functions, so the answer is simply this:

Functions are everywhere!

And everywhere that functions are, there are categories. Indeed, the subject might better have been called *abstract function theory*, or, perhaps even better: *archery*.

1.2 Functions of sets

We begin by considering functions between sets. I am not going to say here what a function is, anymore than what a set is. Instead, we will assume a working knowledge of these terms. They can in fact be *defined* using category theory, but that is not our purpose here.

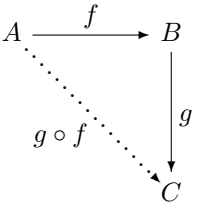
Let f be a function from a set A to another set B, we write

$$f : A \to B.$$

To be explicit, this means that f is defined on all of A and all the values of f are in B. In set theoretic terms,

$$\text{range}(f) \subseteq B.$$

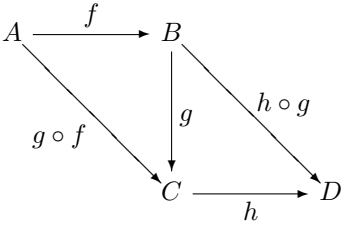
Now suppose we also have a function $g : B \to C$,



then there is a composite function $g \circ f : A \to C$, given by

$$(g \circ f)(a) = g(f(a)) \qquad a \in A. \tag{1.1}$$

Now this operation "$\circ$" of composition of functions is associative, as follows. If we have a further function $h : C \to D$



and form $h \circ g$ and $g \circ f$ then we can compare $(h \circ g) \circ f$ and $h \circ (g \circ f)$ as indicated in the above diagram. It turns out that these two functions are always identical,

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

since for any $a \in A$, we have

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

using (1.1).

By the way, this is of course what it means for two functions to be equal: for every argument, they have the same value.

Finally, note that every set A has an identity function

$$1_A : A \to A$$

given by

$$1_A(a) = a.$$

These identity functions act as "units" for the operation $\circ$ of composition, in the sense of abstract algebra. That is to say,

$$f \circ 1_A = f = 1_B \circ f$$

for any $f : A \to B$.

These are all the properties of set functions that we want to consider for the *abstract* notion of function—composition and identities. Thus, we now want to "abstract away" everything else, so to speak. That is what is accomplished by the following definition.

1.3 Definition of a category

Definition 1.1. A *category* consists of the following data:

- *Objects*: $A, B, C, \ldots$
- *Arrows*: $f, g, h, \ldots$
- For each arrow f there are given objects:

  $$\mathrm{dom}(f), \qquad \mathrm{cod}(f)$$

  called the *domain* and *codomain* of f. We write:

  $$f : A \to B$$

  to indicate that $A = \mathrm{dom}(f)$ and $B = \mathrm{cod}(f)$.
- Given arrows $f : A \to B$ and $g : B \to C$, that is, with:

  $$\mathrm{cod}(f) = \mathrm{dom}(g)$$

  there is given an arrow:

  $$g \circ f : A \to C$$

  called the *composite* of f and g.
- For each object A there is given an arrow:

  $$1_A : A \to A$$

  called the *identity arrow* of A.

These data are required to satisfy the following laws:

- Associativity:
$$h \circ (g \circ f) = (h \circ g) \circ f$$
for all $f : A \to B$, $g : B \to C$, $h : C \to D$.
- Unit:
$$f \circ 1_A = f = 1_B \circ f$$
for all $f : A \to B$.

A category is *anything* that satisfies this definition—and we will have plenty of examples very soon. For now I want to emphasize that, unlike in the previous section, the objects do not have to be sets and the arrows need not be functions. In this sense, a category is an *abstract* algebra of functions, or "arrows" (sometimes also called "morphisms"), with the composition operation "$\circ$" as primitive. If you are familiar with groups, you may think of a category as a sort of generalized group.

1.4 Examples of categories

1. We have already encountered the category **Sets** of sets and functions. There is also the category
$$\mathbf{Sets}_{\text{fin}}$$
of all finite sets and functions between them.

   Indeed, there are many categories like this, given by restricting the sets that are to be the objects and the functions that are to be the arrows. For example, take finite sets as objects and injective (i.e., "1 to 1") functions as arrows. Since injective functions compose to give an injective function, and since the identity functions are injective, this also gives a category.

   What if we take sets as objects and as arrows, those $f : A \to B$ such that for all $b \in B$, the subset
$$f^{-1}(b) \subseteq A$$
has at most two elements (rather than one)? Is this still a category? What if we take the functions such that $f^{-1}(b)$ is finite? infinite? There are lots of such restricted categories of sets and functions.

2. Another kind of example one often sees in mathematics is categories of *structured sets*, that is, sets with some further "structure" and functions which "preserve it," where these notions are determined in some independent way. Examples of this kind you may be familiar with are:

   - groups and group homomorphisms,

- vector spaces and linear mappings,
- graphs and graph homomorphisms,
- the real numbers $\mathbb{R}$ and continuous functions $\mathbb{R} \to \mathbb{R}$,
- open subsets $U \subseteq \mathbb{R}$ and continuous functions $f : U \to V \subseteq \mathbb{R}$ defined on them,
- topological spaces and continuous mappings,
- differentiable manifolds and smooth mappings,
- the natural numbers $\mathbb{N}$ and all recursive functions $\mathbb{N} \to \mathbb{N}$, or as in the example of continuous functions, one can take partial recursive functions defined on subsets $U \subseteq \mathbb{N}$.
- posets and monotone functions.

Do not worry if some of these examples are unfamiliar to you. Later on, we will take a closer look at some of them. For now, let us just consider the last of the above examples in more detail.

3. A partially ordered set or *poset* is a set A equipped with a binary relation $a \leq_A b$ such that the following conditions hold for all $a, b, c \in A$:

   reflexivity: $a \leq_A a$,
   transitivity: if $a \leq_A b$ and $b \leq_A c$, then $a \leq_A c$,
   antisymmetry: if $a \leq_A b$ and $b \leq_A a$, then $a = b$.

   For example, the real numbers $\mathbb{R}$ with their usual ordering $x \leq y$ form a poset that is also *linearly* ordered: either $x \leq y$ or $y \leq x$ for any x, y.

   An arrow from a poset A to a poset B is a function

   $$m : A \to B$$

   that is *monotone*, in the sense that, for all $a, a' \in A$,

   $$a \leq_A a' \quad \text{implies} \quad m(a) \leq_B m(a').$$

   What does it take for this to be a category? We need to know that $1_A : A \to A$ is monotone, but that is clear since $a \leq_A a'$ implies $a \leq_A a'$. We also need to know that if $f : A \to B$ and $g : B \to C$ are monotone, then $g \circ f : A \to C$ is monotone. This also holds, since $a \leq a'$ implies $f(a) \leq f(a')$ implies $g(f(a)) \leq g(f(a'))$ implies $(g \circ f)(a) \leq (g \circ f)(a')$. So we have the category **Pos** of posets and monotone functions.

4. The categories that we have been considering so far are examples of what are sometimes called *concrete categories*. Informally, these are categories in which the objects are sets, possibly equipped with some structure, and the arrows are certain, possibly structure-preserving, functions (we shall see later on that this notion is not entirely coherent; see Remark 1.7). Let's now take a look at a few examples that are not of this sort.

Let **Rel** be the following category: take sets as objects and take binary relations as arrows. That is, an arrow $f : A \to B$ is a subset $f \subseteq A \times B$. The identity arrow on a set A is the identity relation.

$$1_A = \{(a, a) \in A \times A \mid a \in A\} \subseteq A \times A.$$

Given $R \subseteq A \times B$ and $S \subseteq B \times C$, define composition $S \circ R$ by

$$(a, c) \in S \circ R \quad \text{iff} \quad \exists b.\ (a, b) \in R \ \&\ (b, c) \in S$$

that is, the "relative product" of S and R. We leave it as an exercise to show that **Rel** is in fact a category. (What needs to be done?)

For another example of a category in which the arrows are not "functions," let the objects be finite sets A, B, C and an arrow $F : A \to B$ is a rectangular matrix $F = (n_{ij})_{i<a, j<b}$ of natural numbers with $a = |A|$ and $b = |B|$, where $|C|$ is the number of elements in a set C. The composition of arrows is by the usual matrix multiplication, and the identity arrows are the usual unit matrices.

5. *Finite categories*
Of course, the objects of a category do not have to be sets, either. Here are some very simple examples:

- The category **1** looks like this:

  $*$

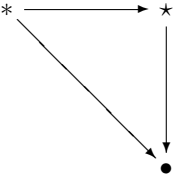
  It has one object and its identity arrow, which we do not draw.

- The category **2** looks like this:

  $* \longrightarrow \star$

  It has two objects, their required identity arrows, and exactly one arrow between the objects.

- The category **3** looks like this:



  It has three objects, their required identity arrows, exactly one arrow from the first to the second object, exactly one arrow from the second to the third object, and exactly one arrow from the first to the third object (which is therefore the composite of the other two).

- The category **0** looks like this:

  It has no objects or arrows.

As above, we will omit the identity arrows in drawing categories from now on.

It is fairly easy to specify finite categories—just take some objects and start putting arrows between them, but make sure to put in the necessary identities and composites, as required by the axioms for a category. Also, if there are any loops, then they need to be cut off by equations to keep the category finite. For example, consider the following specification:

$$A \underset{g}{\overset{f}{\rightleftarrows}} B$$

Unless we stipulate an equation like $gf = 1_A$, we will end up with infinitely many arrows gf, $gfgf$, $gfgfgf$, ... This is still a category, of course, but it is not a *finite* category. We will come back to this situation when we discuss free categories later in this chapter.

6. One important slogan of category theory is,

It's the arrows that really matter!

So we should also look at the arrows or "mappings" between categories. A "homomorphism of categories" is called a functor.

Definition 1.2. A *functor*

$$F : \mathbf{C} \to \mathbf{D}$$

between categories $\mathbf{C}$ and $\mathbf{D}$ is a mapping of objects to objects and arrows to arrows, in such a way that:

(a) $F(f : A \to B) = F(f) : F(A) \to F(B)$,
(b) $F(g \circ f) = F(g) \circ F(f)$,
(c) $F(1_A) = 1_{F(A)}$.

Now, one can check that functors compose in the expected way, and that every category $\mathbf{C}$ has an identity functor $1_{\mathbf{C}} : \mathbf{C} \to \mathbf{C}$. So we have another example of a category, namely **Cat**, the category of all categories and functors.

7. A *preorder* is a set P equipped with a binary relation $p \leq q$ that is both reflexive and transitive: $a \leq a$, and if $a \leq b$ and $b \leq c$, then $a \leq c$. Any preorder P can be regarded as a category by taking the objects to be the elements of P and taking a unique arrow,

$$a \to b \quad \text{if and only if} \quad a \leq b. \tag{1.2}$$

The reflexive and transitive conditions on $\leq$ ensure that this is a category.

Going in the other direction, any category with at most one arrow between any two objects determines a preorder, simply by defining a binary relation $\leq$ on the objects by (1.2).

8. A poset is evidently a preorder satisfying the additional condition of antisymmetry: if $a \leq b$ and $b \leq a$, then $a = b$. So, in particular, a poset is also a category. Such *poset categories* are very common; for example, for any set X, the powerset $P(X)$ is a poset under the usual inclusion relation $U \subseteq V$ between the subsets U, V of X.

   What is a functor $F : P \to Q$ between poset categories P and Q? It must satisfy the identity and composition laws.... Clearly, these are just the monotone functions already considered above.

   It is often useful to think of a category as a kind of *generalized poset*, one with "more structure" than just $p \leq q$. Thus, one can also think of a functor as a generalized monotone map.

9. *An example from logic:* Given a deductive system of logic, there's an associated category, where the objects are formulas:

$$\varphi, \psi, \ldots$$

   An arrow from φ to ψ is a deduction of ψ from the assumption φ. Composition of arrows is given by putting together deductions in the obvious way, which is clearly associative. (What are the identity arrows 1_φ?) Observe that there can be many different arrows

$$p : \varphi \to \psi$$

   since there may be many different proofs. This category turns out to have a very rich structure, which we will consider later in connection with the lambda-calculus.

10. *An example from computer science:* Given a functional programming language L, there is an associated category, where the objects are the data types of L, and the arrows are the computable functions of L ("processes," "procedures," "programs"). The composition of two such programs $X \xrightarrow{f} Y \xrightarrow{g} Z$ is given by applying g to the output of f, sometimes also written

$$g \circ f = f; g.$$

    The identity is the "do nothing" program.

    Categories such as this are basic to the idea of denotational semantics of programming languages. For example, if $\mathbf{C}(L)$ is the category just defined, then the denotational semantics of the language L in a category $\mathbf{D}$ of, say, Scott domains is simply a functor

$$S : \mathbf{C}(L) \to \mathbf{D}$$

since S assigns domains to the types of L and continuous functions to the programs. Both this example and the previous one are related to the notion of "cartesian closed category" that is considered later.

11. Let X be a set. We can regard X as a category $\mathbf{Dis}(X)$ by taking the objects to be the elements of X and taking the arrows to be just the identity arrows, one for each $x \in X$. Such categories are called *discrete*. Note that discrete categories are just very special posets.

12. A *monoid* (sometimes called a *semigroup with unit*) is a set M equipped with a binary operation $\cdot : M \times M \to M$ and a distinguished "unit" element $u \in M$ such that for all $x, y, z \in M$,

$$x \cdot (y \cdot z) = (x \cdot y) \cdot z$$

and

$$u \cdot x = x = x \cdot u.$$

Equivalently, a monoid is a category with just one object. The arrows of the category are the elements of the monoid. In particular, the identity arrow is the unit element u. Composition of arrows is the binary operation $m \cdot n$ of the monoid.

Monoids are very common: there are the monoids of numbers like $\mathbb{N}$, $\mathbb{Q}$ or $\mathbb{R}$ with addition and 0, or multiplication and 1. But also for any set X, the set of functions from X to X, written

$$\mathrm{Hom}_{\mathbf{Sets}}(X, X)$$

is a monoid under the operation of composition. More generally, for any object C in any category $\mathbf{C}$, the set of arrows from C to C, written as $\mathrm{Hom}_{\mathbf{C}}(C, C)$, is a monoid under the composition operation of $\mathbf{C}$.

Since monoids are structured sets, there is a category **Mon** whose objects are monoids and whose arrows are functions that preserve the monoid structure. In detail, a homomorphism from a monoid M to a monoid N is a function $h : M \to N$ such that for all $m, n \in M$,

$$h(m \cdot_M n) = h(m) \cdot_N h(n)$$

and

$$h(u_M) = u_N.$$

The reader should check that a monoid homomorphism from M to N is the same thing as a functor from M regarded as a category to N regarded as a category. In this sense, categories are also generalized monoids, and functors are generalized homomorphisms.

1.5 Isomorphisms

Definition 1.3. In any category $\mathbf{C}$, an arrow $f : A \to B$ is called an *isomorphism* if there is an arrow $g : B \to A$ in $\mathbf{C}$ such that

$$g \circ f = 1_A \quad \text{and} \quad f \circ g = 1_B.$$

Since inverses are unique (proof!), we write $g = f^{-1}$. We say that A is *isomorphic* to B, written $A \cong B$, if there exists an isomorphism between them.

The definition of isomorphism is our first example of an *abstract*, category theoretic definition of an important notion. It is abstract in the sense that it makes use only of the category theoretic notions, rather than some additional information about the objects and arrows. It has the advantage over other possible definitions that it applies in any category. For example, one sometimes defines an isomorphism of sets (groups, etc.) as a *bijective* function (resp. homomorphism), that is, one that is "1-1 and onto." This is *equivalent* to our definition in some cases. But note that, for example in **Pos**, the category theoretic definition gives the right notion, while there are "bijective homomorphisms" between non-isomorphic posets. Moreover, in many cases *only* the abstract definition makes sense, for example, in the case of a monoid.

Definition 1.4. A *group* G is a monoid with an inverse g^{-1} for every element g. Thus G is a category with one object, in which every arrow is an isomorphism.

The natural numbers $\mathbb{N}$ do not form a group under either addition or multiplication, but the integers $\mathbb{Z}$ and the positive rationals $\mathbb{Q}^+$, respectively, do. For any set X, we have the group $\mathrm{Aut}(X)$ of automorphisms (or "permutations") of X, that is, isomorphisms $f : X \to X$. (Question: why is this closed under "$\circ$"?) A *group of permutations* is a subgroup $G \subseteq Aut(X)$ for some set X, that is, a group of automorphisms of X. Thus G must satisfy the following:

1. $1_X \in G$.
2. If $g, g' \in G$, then $g \circ g' \in G$.
3. If $g \in G$, then $g^{-1} \in G$.

A *homomorphism* of groups $h : G \to H$ is just a homomorphism of monoids, which then necessarily also preserves the inverses (proof!).

Theorem (Cayley). *Every group G is isomorphic to a group of permutations.*

Proof. (sketch)

1. First, define the Cayley representation $\bar{G}$ of G to be the following group of permutations: the underlying set of $\bar{G}$ is just G, and for each $g \in G$, we have the permutation $\bar{g}$, defined for all $h \in G$ by:

$$\bar{g}(h) = g \cdot h.$$

Now check that $\bar{g} = \bar{h}$ implies $g = h$.

2. Next define homomorphisms $i : G \to \bar{G}$ by $i(g) = \bar{g}$, and $j : \bar{G} \to G$ by $j(\bar{g}) = g$.
3. Finally show that $i \circ j = 1_{\bar{G}}$ and $j \circ i = 1_G$.

□

Warning 1.5. Note the two different levels of isomorphisms that occur in the proof of Cayley's theorem. There are permutations of the set of elements of G, which are isomorphisms in **Sets**, and there is the isomorphism between G and $\bar{G}$, which is in the category **Groups** of groups and group homomorphisms.

Cayley's theorem says that any abstract group can be represented as a "concrete" one, that is, a group of permutations of a set. The theorem can in fact be generalized to show that any category that is not "too big" can be represented as one that is "concrete," that is, a category of sets and functions. (There is a technical sense of not being "too big" that will be introduced in Section 1.8.)

Theorem 1.6. *Every category* $\mathbf{C}$ *with a set of arrows is isomorphic to one in which the objects are sets and the arrows are functions.*

Proof. (sketch) Define the Cayley representation $\bar{\mathbf{C}}$ of $\mathbf{C}$ to be the following concrete category:

- objects are sets of the form
$$\bar{C} = \{f \in \mathbf{C} \mid \text{cod}(f) = C\}$$
for all $C \in \mathbf{C}$,
- arrows are functions
$$\bar{g} : \bar{C} \to \bar{D}$$
for $g : C \to D$ in $\mathbf{C}$, defined by $\bar{g}(f) = g \circ f$.

□

Remark 1.7. This shows us what is *wrong* with the naive notion of a "concrete" category of sets and functions: while not every category has special sets and functions as its objects and arrows, every category is isomorphic to such a one. Thus, the only special properties such categories can possess are ones that are categorically irrelevant, such as features of the objects that do not affect the arrows in any way (like the difference between the real numbers constructed as Dedekind cuts or as Cauchy sequences). A better attempt to capture what is intended by the rather vague idea of a "concrete" category is that arbitrary arrows $f : C \to D$ are completely determined by their composites with arrows $x : T \to C$ from some "test object" T, in the sense that $fx = gx$ for all such x implies $f = g$. This amounts to considering a particular representation of the category, determined by T. A category is then said to be "concrete" when this

condition holds for T a "terminal object," in the sense of Section 2.2; but there are also good reasons for considering other objects T, as we see in the next chapter.

Note that the condition that $\mathbf{C}$ have a *set* of arrows is needed to ensure that the collections $\{f \in \mathbf{C} \mid \mathrm{cod}(f) = C\}$ really are *sets*—we return to this point in Section 1.8.

1.6 Constructions on categories

Now that we have a stock of categories to work with, we consider some constructions that produce new categories from old.

1. The *product* of two categories $\mathbf{C}$ and $\mathbf{D}$, written

$$\mathbf{C} \times \mathbf{D}$$

has objects of the form (C, D), for $C \in \mathbf{C}$ and $D \in \mathbf{D}$, and arrows of the form

$$(f, g) : (C, D) \to (C', D')$$

for $f : C \to C' \in \mathbf{C}$ and $g : D \to D' \in \mathbf{D}$. Composition and units are defined componentwise; that is,

$$(f', g') \circ (f, g) = (f' \circ f, g' \circ g)$$
$$1_{(C,D)} = (1_C, 1_D).$$

There are two obvious *projection functors*

$$\mathbf{C} \xleftarrow{\pi_1} \mathbf{C} \times \mathbf{D} \xrightarrow{\pi_2} \mathbf{D}$$

defined by $\pi_1(C, D) = C$ and $\pi_1(f, g) = f$, and similarly for π_2.

The reader familiar with groups will recognize that for groups G and H, the product category $G \times H$ is the usual (direct) product of groups.

2. The *opposite* (or "dual") category $\mathbf{C}^{\mathrm{op}}$ of a category $\mathbf{C}$ has the same objects as $\mathbf{C}$, and an arrow $f : C \to D$ in $\mathbf{C^{op}}$ is an arrow $f : D \to C$ in C. That is $\mathbf{C}^{\mathrm{op}}$ is just $\mathbf{C}$ with all of the arrows formally turned around.

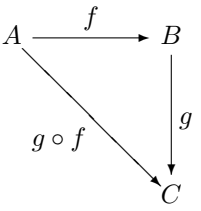
It is convenient to have a notation to distinguish an object (resp. arrow) in $\mathbf{C}$ from the same one in $\mathbf{C^{op}}$. Thus, let us write

$$\bar{f} : \bar{D} \to \bar{C}$$

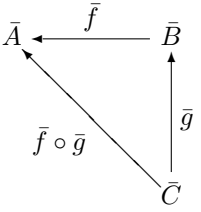
in $\mathbf{C}^{\mathrm{op}}$ for $f : C \to D$ in $\mathbf{C}$. With this notation we can define composition and units in $\mathbf{C}^{\mathrm{op}}$ in terms of the corresponding operations in $\mathbf{C}$, namely,

$$1_{\bar{C}} = \bar{1_C}$$
$$\bar{f} \circ \bar{g} = \overline{g \circ f}.$$

Thus, a diagram in $\mathbf{C}$

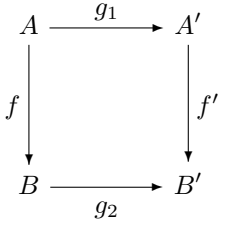


looks like this in $\mathbf{C}^{\mathrm{op}}$



Many "duality" theorems of mathematics express the fact that one category is (a subcategory of) the opposite of another. An example of this sort which we will prove later is that **Sets** is dual to the category of complete, atomic Boolean algebras.

3. The *arrow category* $\mathbf{C}^{\to}$ of a category $\mathbf{C}$ has the arrows of $\mathbf{C}$ as objects, and an arrow g from $f : A \to B$ to $f' : A' \to B'$ in $\mathbf{C}^{\to}$ is a "commutative square"



where g_1 and g_2 are arrows in $\mathbf{C}$. That is, such an arrow is a pair of arrows $g = (g_1, g_2)$ in $\mathbf{C}$ such that

$$g_2 \circ f = f' \circ g_1.$$

The identity arrow 1_f on an object $f : A \to B$ is the pair $(1_A, 1_B)$. Composition of arrows is done componentwise:

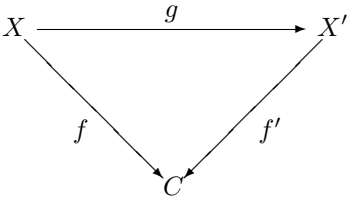
$$(h_1, h_2) \circ (g_1, g_2) = (h_1 \circ g_1, h_2 \circ g_2)$$

The reader should verify that this works out by drawing the appropriate commutative diagram.

Observe that there are two functors:

$$\mathbf{C} \xleftarrow{\textbf{dom}} \mathbf{C}^{\to} \xrightarrow{\textbf{cod}} \mathbf{C}$$

4. The *slice category* $\mathbf{C}/C$ of a category $\mathbf{C}$ over an object $C \in \mathbf{C}$ has:
   - objects: all arrows $f \in \mathbf{C}$ such that $\mathrm{cod}(f) = C$,
   - arrows: g from $f : X \to C$ to $f' : X' \to C$ is an arrow $g : X \to X'$ in $\mathbf{C}$ such that $f' \circ g = f$, as indicated in



   We leave it to the reader to work out the identity arrows and composites.
   If $\mathbf{C} = \mathbf{P}$ is a poset category and $p \in \mathbf{P}$, then

$$\mathbf{P}/p \cong {\downarrow}(p)$$

   the slice category $\mathbf{P}/p$ is just the "principal ideal" ${\downarrow}(p)$ of elements $q \in \mathbf{P}$ with $q \leq p$. We will have more examples of slice categories soon.

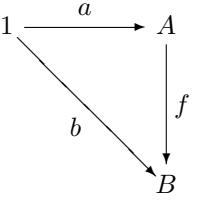
   There is an obvious functor $U : \mathbf{C}/C \to \mathbf{C}$ that "forgets about the base object C." Can you find a functor $F : \mathbf{C}/C \to \mathbf{C}^{\to}$ such that $\mathbf{dom} \circ F = U$? The *coslice* category $C/\mathbf{C}$ of a category $\mathbf{C}$ under an object C of $\mathbf{C}$ has as objects all arrows f of $\mathbf{C}$ such that $\mathrm{dom}(f) = C$, and an arrow from $f : C \to X$ to $f' : C \to X'$ is an arrow $h : X \to X'$ such that $h \circ f = f'$. The reader should now carry out the rest of the definition of the coslice category by analogy with the definition of the slice category.

   How can the coslice category be defined in terms of the slice category and the opposite construction?

Example 1.8. The category $\mathbf{Sets}_*$ of *pointed sets* consists of sets A with a distinguished element $a \in A$, and arrows $f : (A, a) \to (B, b)$ are functions $f : A \to B$ that preserves the "points," $f(a) = b$. This is isomorphic to the coslice category,

$$\mathbf{Sets}_* \cong 1\backslash\mathbf{Sets}$$

of Sets "under" any singleton $1 = \{*\}$. Indeed, functions $a : 1 \to A$ correspond uniquely to elements, $a(*) = a \in A$, and arrows $f : (A, a) \to (B, b)$ correspond exactly to commutative triangles:



1.7 Free categories

Free monoid. Start with an "alphabet" A of "letters" (a set)

$$A = \{a, b, c, \ldots\}.$$

A *word* over A is a finite sequence of letters:

$$thisword, \quad categoriesarefun, \quad asddjbnzzfj, \ldots$$

We write "-" for the empty word. The "Kleene closure" of A is defined to be the set

$$A^* = \{\text{words over } A\}.$$

Define a binary operation "$*$" on A^* defined for $w, w' \in A^*$ by $w * w' = ww'$. Thus, "$*$" is just *concatenation*. The operation "$*$" is associative, and the empty word "-" is a unit. Thus, A^* is a monoid—called the *free monoid* on the set A. The elements $a \in A$ can be regarded as words of length one, so we have a function

$$i : A \to A^*$$

defined by $i(a) = a$, and called the "insertion of generators." The elements of A "generate" the free monoid, in the sense that every $w \in A^*$ is a $*$-product of a's, that is, $w = a_1 * a_2 * \cdots * a_n$ for some $a_1, a_2, ..., a_n$ in A.

Now what does "free" mean here? Any guesses?

One sometimes sees definitions in "baby algebra" books along the following lines:

A monoid M is *freely generated* by a subset A of M, if the following conditions hold.

1. Every element $m \in M$ can be written as a product of elements of A

$$m = a_1 \cdot_M \ldots \cdot_M a_n \ , \quad a_i \in A.$$

2. No "nontrivial" relations hold in M, that is, if $a_1 \ldots a_j = a'_1 \ldots a'_k$, then this is required by the axioms for monoids.

The first condition is sometimes called "no junk," while the second condition is sometimes called "no noise." Thus, the free monoid on A is a monoid containing A and having no junk and no noise. What do you think of this definition of a free monoid?

I would object to the reference in the second condition to "provability," or something. This must be made more precise for this to succeed as a definition. In category theory, we give a precise definition of "free"—capturing what is meant in the above—which avoids such vagueness.

First, every monoid N has an underlying set $|N|$, and every monoid homomorphism $f : N \to M$ has an underlying function $|f| : |N| \to |M|$. It is easy to see that this is a functor, called the "forgetful functor." The free monoid $M(A)$

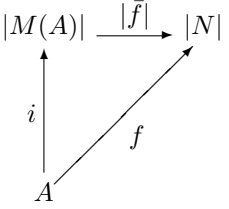
on a set A is by definition "the" monoid with the following so called *universal mapping property*, or UMP!

Universal Mapping Property of $M(A)$
There is a function $i : A \to |M(A)|$, and given any monoid N and any function $f : A \to |N|$, there is a *unique* monoid homomorphism $\bar{f} : M(A) \to N$ such that $|\bar{f}| \circ i = f$, all as indicated in the following diagram:

in **Mon***:*

$$M(A) \overset{\bar{f}}{\dashrightarrow} N$$

in **Sets***:*



Proposition 1.9. *A^* has the UMP of the free monoid on A.*

Proof. Given $f : A \to |N|$, define $\bar{f} : M(A) \to N$ by

$$\begin{aligned} \bar{f}(-) &= u_N, \quad \text{the unit of } N \\ \bar{f}(a_1 \dots a_i) &= f(a_1) \cdot_N \dots \cdot_N f(a_i). \end{aligned}$$

Then $\bar{f}$ is clearly a homomorphism with

$$\bar{f}(a) = f(a) \quad \text{for all } a \in A.$$

If $g : M(A) \to N$ also satisfies $g(a) = f(a)$ for all $a \in A$, then for all $a_1 \dots a_i \in A^*$:

$$\begin{aligned} g(a_1 \dots a_i) &= g(a_1) \cdot_N \dots \cdot_N g(a_i) \\ &= f(a_1) \cdot_N \dots \cdot_N f(a_i) \\ &= \bar{f}(a_1) \cdot_N \dots \cdot_N \bar{f}(a_i) \\ &= \bar{f}(a_1 \dots a_i). \end{aligned}$$

So, $g = \bar{f}$, as required. □

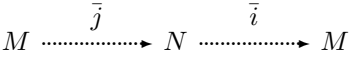
Think about why the above UMP captures precisely what is meant by "no junk" and "no noise." Specifically, the existence part of the UMP captures the vague notion of "no noise," while the uniqueness part makes precise the "no junk" idea.

Using the UMP, it is easy to show that the free monoid $M(A)$ is determined uniquely up to isomorphism, in the following sense.

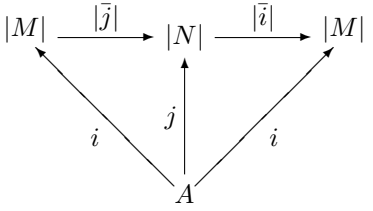
Proposition 1.10. *Given monoids M and N with functions $i : A \to |M|$ and $j : A \to |N|$, each with the UMP of the free monoid on A, there is a (unique) monoid isomorphism $h : M \cong N$ such that $|h|i = j$ and $|h^{-1}|j = i$.*

Proof. From j and the UMP of M, we have $\bar{j} : M \to N$ with $|\bar{j}|i = j$ and from i and the UMP of N, we have $\bar{i} : N \to M$ with $|\bar{i}|j = i$. Composing gives a homomorphism $\bar{i} \circ \bar{j} : M \to M$ such that $|\bar{i} \circ \bar{j}|i = i$. Since $1_M : M \to M$ also has this property, by the uniqueness part of the UMP of M, we have $\bar{i} \circ \bar{j} = 1_M$. Exchanging the roles of M and N shows $\bar{j} \circ \bar{i} = 1_N$;

in **Mon***:*



in **Sets***:*

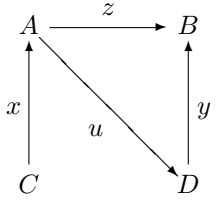


□

For example, the free monoid on a set with a single element is isomorphic to the monoid of natural numbers $\mathbb{N}$ under addition (the "generator" is 1). Thus, as a monoid, $\mathbb{N}$ is uniquely determined up to isomorphism by the UMP of free monoids.

Free category. Now, we want to do the same thing for categories in general (not just monoids). Instead of underlying sets, categories have underlying graphs, so let us review these first.

A *directed graph* consists of vertices and edges, each of which is directed, that is, each edge has a "source" and a "target" vertex.



We draw graphs just like categories, but there is no composition of edges, and there are no identities.

A graph thus consists of two sets, E (edges) and V (vertices), and two functions, $s : E \to V$ (source) and $t : E \to V$ (target).

Now, every graph G "generates" a category $\mathbf{C}(G)$, the *free category* on G. It is defined by taking the vertices of G as objects, and the *paths* in G as arrows,

where a path is a finite sequence of edges $e_1, \ldots, e_n$ such that $t(e_i) = s(e_{i+1})$, for all $i = 1 \ldots n$. We'll write the arrows of $\mathbf{C}(G)$ in the form $e_n e_{n-1} \ldots e_1$.

Put

$$\begin{aligned} \mathrm{dom}(e_n \ldots e_1) &= s(e_1) \\ \mathrm{cod}(e_n \ldots e_1) &= t(e_n) \end{aligned}$$

and define composition by concatenation:

$$e_n \ldots e_1 \circ e'_m \ldots e'_1 = e_n \ldots e_1 e'_m \ldots e'_1.$$

For each vertex v, we have an "empty path" denoted 1_v, which is to be the identity arrow at v.

Note that if G has only one vertex, then $\mathbf{C}(G)$ is just the free monoid on the set of edges of G. Also note that if G has only vertices (no edges), then $\mathbf{C}(G)$ is the discrete category on the set of vertices of G.

Later on, we will have a general definition of "free." For now, let us see that $\mathbf{C}(G)$ also has a UMP.

First, define a "forgetful functor"

$$U : \mathbf{Cat} \to \mathbf{Graphs}$$

in the obvious way: the underlying graph of a category $\mathbf{C}$ has as edges the arrows of $\mathbf{C}$, and as vertices the objects, with $s = \mathrm{dom}$ and $t = cod$. The action of U on functors is equally clear, or at least it will be, once we have defined the arrows in **Graphs**.

A homomorphism of graphs is of course a "functor without the conditions on identities and composition," that is, a mapping of edges to edges and vertices to vertices that preserves sources and targets. We will describe this from a slightly different point of view, that will be useful later on. First, observe that we can describe a category $\mathbf{C}$ with a diagram like this:

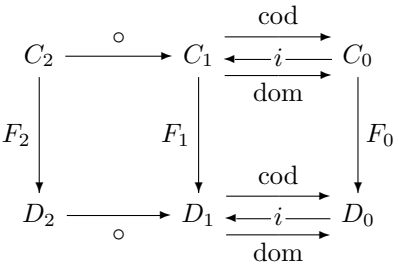
$$C_2 \xrightarrow{\ \circ\ } C_1 \underset{\mathrm{dom}}{\overset{\mathrm{cod}}{\leftarrow\!\!-i-\!\!\rightarrow}} C_0$$

where C_0 is the collection of objects of $\mathbf{C}$, C_1 the arrows, and C_2 is the collection $\{(f, g) \in C_1 \times C_1 : \mathrm{cod}(f) = \mathrm{dom}(g)\}$.

Then a functor $F : \mathbf{C} \to \mathbf{D}$ from $\mathbf{C}$ to another category $\mathbf{D}$ is a pair of functions

$$\begin{aligned} F_0 &: C_0 \to D_0 \\ F_1 &: C_1 \to D_1 \end{aligned}$$

such that each similarly labeled square in the following diagram commutes:

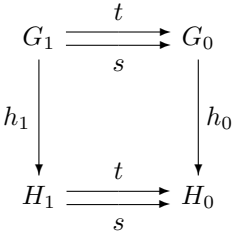


where $F_2(f, g) = (F_1(f), F_1(g))$.

Now let us describe a *homomorphism of graphs*,

$$h : G \to H.$$

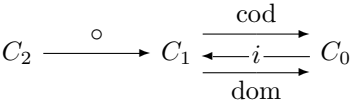
We need a pair of functions $h_0 : G_0 \to H_0$, $h_1 : G_1 \to H_1$ making the two squares (once with t's, once with s's) in the following diagram commute:



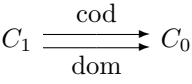
In these terms, we can easily describe the forgetful functor,

$$U : \mathbf{Cat} \to \mathbf{Graphs}$$

as sending the category



to the underlying graph:



And similarly for functors, the effect of U is described by erasing some parts of the diagrams (which is easier to demonstrate with chalk!).

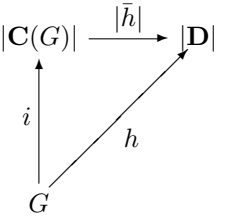
The free category on a graph has the following UMP:

Universal Mapping Property of $\mathbf{C}(G)$
There is a graph homomorphism $i : G \to |\mathbf{C}(G)|$, and given any category $\mathbf{D}$ and any graph homomorphism $h : G \to |\mathbf{D}|$, there is a *unique* functor $\bar{h} : \mathbf{C}(G) \to \mathbf{D}$ with $|\bar{h}| \circ i = h$.

in **Cat***:*

$$\mathbf{C}(G) \overset{\bar{h}}{\dashrightarrow} \mathbf{D}$$

in **Graph***:*



The free category on a graph with just one vertex is just a free monoid on the set of edges. The free category on a graph with two vertices and one edge between them is the finite category **2**. The free category on a graph of the form:

$$A \underset{f}{\overset{e}{\rightleftarrows}} B$$

has (in addition to the identity arrows) the infinitely many arrows:

$$e, f, ef, fe, efe, fef, efef, ...$$

1.8 Foundations: large, small, and locally small

Let us begin by distinguishing between the following things:

categorical foundations for mathematics,
mathematical foundations for category theory.

As for the first: one sometimes hears it said that category theory can be used to provide "foundations for mathematics," as an alternative to set theory. That is in fact the case, but it is not what we are doing here. In set theory, one often begins with existential axioms such as "every set has a powerset" and derives further sets, building up a universe of mathematical objects (namely sets), which in principle suffice for "all of mathematics." Our axiom that every arrow has a domain and a codomain is not to be understood in the same way as set theory's axiom that every set has a powerset! The difference is that in set theory—at least as usually conceived—the axioms are to be regarded as referring to (or determining) a single universe of sets. In category theory, by contrast, the axioms are a *definition* of something, namely of categories. This is just like in group theory or topology, where the axioms serve to define the objects under investigation. These, in turn, are assumed to exist in some "background" or "foundational" system, like set theory. That theory of sets could itself, in turn, be determined using category theory, or in some other way.

This brings us to the second point: we assume that our categories are comprised of sets and functions, in one way or another, like most mathematical objects, and taking into account the remarks just made about the possibility of categorical foundations. But in category theory, we sometimes run into difficulties with set theory as usually practiced. Mostly these are questions of size; some categories are "too big" to be handled comfortably in conventional set theory. We already encountered this issue when we considered the Cayley representation in Section 1.5. There we had to require that the category under consideration had (no more than) a set of arrows. We would certainly not want to impose this restriction in general, however (as one usually does for, say, groups); for then even the "category" **Sets** would fail to be a proper category, as would many other categories that we definitely want to study.

There are various formal devices for addressing these issues, and they are discussed in the book by Mac Lane. For our immediate purposes, the following distinction will be useful:

Definition 1.11. A category $\mathbf{C}$ is called *small* if both the collection $\mathbf{C}_0$ of objects of $\mathbf{C}$ and the collection $\mathbf{C}_1$ of arrows of $\mathbf{C}$ are sets. Otherwise, $\mathbf{C}$ is called *large*.

For example, all finite categories are clearly small, as is the category $\mathbf{Sets}_{\text{fin}}$ of finite sets and functions. On the other hand, the category **Pos** of posets, the category **Groups** of groups, and the category **Sets** of sets are all large. We let **Cat** be the category of all *small categories*, which itself is a large category. In particular, then, **Cat** is not an object of itself, which may come as a relief to some readers.

This does not really solve all of our difficulties. Even for large categories like **Groups** and **Sets** we will want to also consider constructions like the category of all functors from one to the other (we will define this "functor category" later). But if these are not small, conventional set theory does not provide the means to do this directly (these categories would be "too large"). So, one needs a more elaborate theory of "classes" to handle such constructions. We will not worry about this when it is just a matter of technical foundations (Mac Lane I.6 addresses this issue). However, one very useful notion in this connection is the following:

Definition 1.12. A category $\mathbf{C}$ is called *locally small* if for all objects X, Y in $\mathbf{C}$, the collection $\text{Hom}_{\mathbf{C}}(X,Y) = \{f \in \mathbf{C}_1 |\ f : X \to Y\ \}$ is a *set* (called a *hom-set*).

Many of the large categories we want to consider are in fact locally small. **Sets** is locally small since $\text{Hom}_{\mathbf{Sets}}(X,Y) = Y^X$, the *set* of all functions from X to Y. Similarly, **Pos**, **Top**, and **Group** are all locally small (is **Cat**?), and, of course, any small category is locally small.

1.9 Exercises

1. The objects of **Rel** are sets, and an arrow $f : A \to B$ is a relation from A to B, that is, $f \subseteq A \times B$. The identity relation $\{\langle a, a\rangle \in A \times A |\ a \in A\}$ is the identity arrow on a set A. Composition in **Rel** is to be given by

$$g \circ f = \{\langle a, c\rangle \in A \times C \mid \exists b\, (\langle a, b\rangle \in f \ \&\ \langle b, c\rangle \in g)\}$$

for $f \subseteq A \times B$ and $g \subseteq B \times C$.
Show that **Rel** is a category.

2. Consider the following isomorphisms of categories and determine which hold.
   (a) $\mathbf{Rel} \cong \mathbf{Rel}^{\mathrm{op}}$
   (b) $\mathbf{Sets} \cong \mathbf{Sets}^{\mathrm{op}}$
   (c) For a fixed set X with powerset $P(X)$, as poset categories $P(X) \cong P(X)^{\mathrm{op}}$ (the arrows in $P(X)$ are subset inclusions $A \subseteq B$ for $A, B \subseteq X$).

3. (a) Show that in **Sets**, the isomorphisms are exactly the bijections.
   (b) Show that in **Monoids**, the isomorphisms are exactly the bijective homomorphisms.
   (c) Show that in **Posets**, the isomorphisms are *not* the same as the bijective homomorphisms.

4. Construct the "coslice category" $C/\mathbf{C}$ of a category $\mathbf{C}$ under an object C from the slice category $\mathbf{C}/C$ and the "dual category" operation $-^{\mathrm{op}}$.

5. How many free categories on graphs are there which have exactly six arrows? Draw the graphs that generate these categories.

6. Prove the UMP for free categories on graphs:
Let $\mathbf{C}(G)$ be the free category on the graph G and $i : G \to U(\mathbf{C}(G))$ the graph homomorphism taking vertices and edges to themselves, regarded as objects and arrows in $\mathbf{C}(G)$. For any category $\mathbf{D}$ and graph homomorphism $f : G \to U(\mathbf{D})$, there is a unique functor

$$\bar{h} : \mathbf{C}(G) \to \mathbf{D}$$

with

$$U(\bar{h}) \circ i = h,$$

where $U : \mathbf{Cat} \to \mathbf{Graph}$ is the underlying graph functor.

This page intentionally left blank

2

ABSTRACT STRUCTURES

Let me begin with some remarks about category-theoretical definitions. By this I mean characterizations of properties of objects and arrows in a category in terms of other objects and arrows only, that is, in the language of category theory. Such definitions may be said to be abstract, structural, operational, relational, or external (as opposed to internal). The idea is that objects and arrows are determined by the role they play in the category via their relations to other objects and arrows, that is, by their position in a structure and not by what they "are" or "are made of" in some absolute sense. We will see many more examples of this kind of thing later; for now we start with some very simple ones. Let me call them *abstract characterizations*. We will see that one of the basic ways of giving such an abstract characterization is via a Universal Mapping Property or *UMP*.

2.1 Epis and monos

Recall that in **Sets**, a function $f : A \to B$ is called

injective if $f(a) = f(a')$ implies $a = a'$ for all $a, a' \in A$,
surjective if for all $b \in B$ there is some $a \in A$ with $f(a) = b$.

We have the following abstract characterizations of these properties:

Definition 2.1. In any category $\mathbf{C}$, an arrow

$$f : A \to B$$

is called a

monomorphism, if given any $g, h : C \to A$, $fg = fh$ implies $g = h$,

$$C \underset{h}{\overset{g}{\rightrightarrows}} A \xrightarrow{f} B$$

epimorphism, if given any $i, j : B \to D$, $if = jf$ implies $i = j$,

$$A \xrightarrow{f} B \underset{j}{\overset{i}{\rightrightarrows}} D$$

We often write $f : A \rightarrowtail B$ if f is a monomorphism and $f : A \twoheadrightarrow B$ if f is an epimorphism.

Proposition 2.2. *A function $f : A \to B$ between sets is monic just in case it is injective.*

Proof. Suppose $f : A \rightarrowtail B$. Let $a, a' \in A$ such that $a \neq a'$, and let $\{x\}$ be any given one-element set. Consider the functions

$$\bar{a}, \bar{a'} : \{x\} \to A$$

where

$$\bar{a}(x) = a, \qquad \bar{a'}(x) = a'.$$

Since $\bar{a} \neq \bar{a'}$, it follows, since f is a monomorphism, that $f\bar{a} \neq f\bar{a'}$. Thus, $f(a) = (f\bar{a})(x) \neq (f\bar{a'})(x) = f(a')$. Whence f is injective.

Conversely, if f is injective and $g, h : C \to A$ are functions such that $g \neq h$, then for some $c \in C$, $g(c) \neq h(c)$. Since f is injective, it follows that $f(g(c)) \neq f(h(c))$, whence $fg \neq fh$. □

Example 2.3. In many categories of "structured sets" like monoids, the monos are exactly the "injective homomorphisms." More precisely, a homomorphism $h : M \to N$ of monoids is monic just if the underlying function $|h| : |M| \to |N|$ is monic, that is, injective by the foregoing. To prove this, suppose h is monic and take two different "elements" $m, m' : 1 \to |M|$, where $1 = \{*\}$ is any one-element set. By the UMP of the free monoid $M(*)$ there are distinct corresponding homomorphisms $\bar{m}, \bar{m'} : M(*) \to M$, with distinct composites $h \circ \bar{m}, h \circ \bar{m'} : M(*) \to M \to N$, since h is monic. Thus, the corresponding "elements" $hm, hm' : 1 \to N$ of N are also distinct, again by the UMP of $M(*)$. Conversely, if $|h| : |M| \to |N|$ is monic and $f, g : X \to M$ are any distinct homomorphisms, then $|f|, |g| : |X| \to |M|$ are distinct functions, and so $|h| \circ |f|, |h| \circ |g| : |X| \to |M| \to |N|$ are distinct, since $|h|$ is monic. Since therefore $|h \circ f| = |h| \circ |f| \neq |h| \circ |g| = |h \circ g|$, we also must have $h \circ f \neq h \circ g$.

A completely analogous situation holds, for example, for groups, rings, vector spaces, and posets. We shall see that this fact follows from the presence, in each of these categories, of certain objects like the free monoid $M(*)$.

Example 2.4. In a fixed poset $\mathbf{P}$, every arrow $p \leq q$ is both monic and epic. Why?

While the epis in **Sets** are exactly the surjective functions (exercise!), epis in other categories are not always surjective homomorphisms, as the following example shows.

Example 2.5. In the category **Mon** of monoids and monoid homomorphisms, there is a monic homomorphism

$$\mathbb{N} \rightarrowtail \mathbb{Z}$$

where $\mathbb{N}$ is the additive monoid $(N, +, 0)$ of natural numbers and $\mathbb{Z}$ is the additive monoid $(Z, +, 0)$ of integers. We will show that this map, given by the inclusion $N \subset Z$ of sets, is also epi in **Mon** by showing that the following holds:

Given any monoid homomorphisms $f, g : (\mathbb{Z}, +, 0) \to (M, *, u)$, if the restrictions to N are equal, $f \mid_N = g \mid_N$, then $f = g$.

Note first that:

$$\begin{aligned} f(-n) &= f((-1)_1 + (-1)_2 + \cdots + (-1)_n) \\ &= f(-1)_1 * f(-1)_2 * \cdots * f(-1)_n \end{aligned}$$

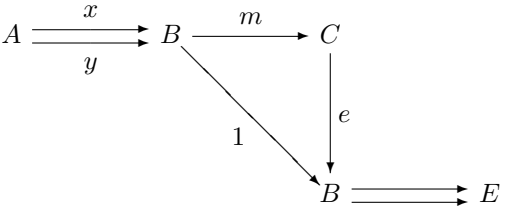
and similarly for g. It therefore suffices to show that $f(-1) = g(-1)$. But

$$\begin{aligned} f(-1) &= f(-1) * u \\ &= f(-1) * g(0) \\ &= f(-1) * g(1-1) \\ &= f(-1) * g(1) * g(-1) \\ &= f(-1) * f(1) * g(-1) \\ &= f(-1+1) * g(-1) \\ &= f(0) * g(-1) \\ &= u * g(-1) \\ &= g(-1). \end{aligned}$$

Note that a morphism e is epi if and only if e cancels on the right: $xe = ye$ implies $x = y$. Dually, m is mono if and only if m cancels on the left: $mx = my$ implies $x = y$.

Proposition 2.6. *Every iso is mono and epi.*

Proof. Consider the following diagram:



If m is an isomorphism with inverse e, then $mx = my$ implies $x = emx = emy = y$. Thus, m is monic. Similarly, e cancels on the right, and thus e is epic. □

In **Sets** the converse of the foregoing also holds: every mono-epi is iso. But this is not in general true, as shown by the example in monoids above.

2.2 Initial and terminal objects

We now consider abstract characterizations of the empty set and the one-element sets in the category **Sets** and structurally similar objects in general categories.

Definition 2.7. In any category $\mathbf{C}$, an object

0 is *initial* if for any object C there is a unique morphism

$$0 \to C,$$

1 is *terminal* if for any object C there is a unique morphism

$$C \to 1.$$

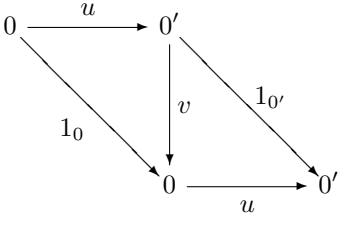
As in the case of monos and epis, note that there is a kind of "duality" in these definitions. Precisely, a terminal object in $\mathbf{C}$ is exactly an initial object in $\mathbf{C}^{\mathrm{op}}$. We will consider this duality systematically later.

Since the notions of initial and terminal object are simple UMPs, such objects are uniquely determined up to isomorphism, just like the free monoids were.

Proposition 2.8. *Initial (terminal) objects are unique up to isomorphism.*

Proof. In fact, if C and C' are both initial (terminal), then there is a *unique* isomorphism $C \to C'$.

Suppose that 0 and $0'$ are both initial objects, the following diagram makes it clear that 0 and $0'$ are uniquely isomorphic.



For terminal objects, apply the foregoing to $\mathbf{C}^{\mathrm{op}}$. □

Example 2.9.

1. In **Sets** the empty set is initial and any singleton set is terminal. Observe that **Sets** has just one initial object but many terminal objects (answering the question of whether $\mathbf{Sets} \cong \mathbf{Sets}^{\mathrm{op}}$).
2. In **Cat** the category **0** (no objects and no arrows) is initial and the category **1** (one object and its identity arrow) is terminal.

3. In **Groups**, the one-element group is *both* initial and terminal (similarly for the category of vector spaces and linear transformations, as well as the category of monoids and monoid homomorphisms). But in **Rings** (commutative with unit), the ring $\mathbb{Z}$ of integers is initial (the one-element ring with $0 = 1$ is still terminal).
4. A *Boolean algebra* is a poset B equipped with distinguished elements $0, 1$, binary operations $a \vee b$ of "join" and $a \wedge b$ of "meet," and a unary operation $\neg b$ of "complementation." These are required to satisfy the conditions

$$\begin{array}{rcrcl} & & & & 0 \leq a \\ & & & & a \leq 1 \\ a \leq c & \text{and} & b \leq c & \text{iff} & a \vee b \leq c \\ c \leq a & \text{and} & c \leq b & \text{iff} & c \leq a \wedge b \\ & & a \leq \neg b & \text{iff} & a \wedge b = 0 \\ & & & & \neg\neg a = a \end{array}$$

   There is also an equivalent, fully equational characterization not involving the ordering. A typical example of a Boolean algebra is the powerset $P(X)$ of all subsets $A \subseteq X$ of a set X, ordered by inclusion $A \subseteq B$, and with the Boolean operations being the empty set $0 = \emptyset$, the whole set $1 = X$, union and intersection of subsets as join and meet, and the relative complement $X - A$ as $\neg A$. A familiar special case is the two-element Boolean algebra $\mathbf{2} = \{0, 1\}$, sometimes also regarded as "truth values" with the logical operations of disjunction, conjunction, and negation as the Boolean operations. It is an initial object in the category of Boolean algebras, which has as arrows the Boolean homomorphisms. These are functors $h : B \to B'$ that preserve the additional structure, in the sense that $h(0) = 0$, $h(a \vee b) = h(a) \vee h(b)$, etc. The one-element Boolean algebra is terminal.
5. In a poset, an object is plainly initial iff it is the least element, and terminal iff it is the greatest element. Clearly, a category *need not* have either an initial object or a terminal object, for example, the poset $(\mathbb{Z}, \leq)$ has neither an initial object nor a terminal object.
6. For any category $\mathbf{C}$ and any object $X \in \mathbf{C}$, the identity arrow $1_X : X \to X$ is a terminal object in $\mathbf{C}/X$ and an initial object in $X/\mathbf{C}$.

2.3 Generalized elements

Let us consider arrows into and out of initial and terminal objects. Clearly only certain of these will be of interest, but those are often especially significant.

A set A has an arrow into the initial object $A \to 0$ just if it is itself empty and the same is true for posets. In monoids and groups, by contrast, every object has a unique arrow to the initial object, since it is also terminal.

In the category **Bool** of Boolean algebras, however, the situation is quite different. The maps $p : B \to \mathbf{2}$ into the initial Boolean algebra $\mathbf{2}$ correspond uniquely to the so-called *ultrafilters* U in B. A *filter* in a Boolean algebra B is a nonempty subset $F \subseteq B$ that is closed upward and under meets:

$$a \in F \text{ and } a \leq b \quad \text{implies} \quad b \in F$$
$$a \in F \text{ and } b \in F \quad \text{implies} \quad a \wedge b \in F$$

A filter F is *maximal* if the only strictly larger filter $F \subset F'$ is all of B. An *ultrafilter* is a maximal filter. It is not hard to see that a filter F is an ultrafilter just if for every element $b \in B$, either $b \in F$ or $\neg b \in F$, and not both (exercise!). Now if $p : B \to \mathbf{2}$, let $U_p = p^{-1}(1)$ to get an ultrafilter $U_p \subset B$. And given an ultrafilter $U \subset B$, define $p_U(b) = 1$ iff $b \in U$ to get a Boolean homomorphism $p_U : B \to \mathbf{2}$. This is easy to check, as is the fact that these operations are mutually inverse. Boolean homomorphisms $B \to \mathbf{2}$ are also used in forming the "truth tables" one meets in logic. Indeed, a row of a truth table corresponds to such a homomorphism on a Boolean algebra of formulas.

Ring homomorphisms $A \to \mathbb{Z}$ into the initial ring $\mathbb{Z}$ play an analogous and equally important role in algebraic geometry.

Now let us consider arrows from terminal objects. For any set X, for instance, we have an isomorphism

$$X \cong \mathrm{Hom}_{\mathbf{Sets}}(1, X)$$

between elements $x \in X$ and arrows $\bar{x} : 1 \to X$, determined by $\bar{x}(*) = x$, from a terminal object $1 = \{*\}$. We have already used this correspondence several times. A similar situation holds in posets (and in topological spaces), where the arrows $1 \to P$ correspond to elements of the underlying set of a poset (or space) P. In any category with a terminal object 1, such arrows $1 \to A$ are called *global elements*, or *points*, or *constants* of A. In sets, posets, and spaces, the general arrows $A \to B$ are determined by what they do to the points of A, in the sense that two arrows $f, g : A \to B$ are equal if for every point $a : 1 \to A$ the composites are equal, $fa = ga$.

But be careful; this is not always the case! How many points are there of an object M in the category of monoids? That is, how many arrows of the form $1 \to M$ for a given monoid M? Just one! And how many points does a Boolean algebra have?

Because, in general, an object is not determined by its points it is convenient to introduce the device of *generalized elements*. These are arbitrary arrows,

$$x : X \to A$$

(any domain X), which can be regarded as *generalized* or *variable elements* of A. Computer scientists and logicians sometimes think of arrows $1 \to A$ as constants or closed terms and general arrows $X \to A$ as (arbitrary) terms.

Example 2.10.

1. Consider arrows $f, g : X \to Y$ in **Pos**. Then $f = g$ iff for all $x : 1 \to X$, we have $fx = gx$. In this sense, posets "have enough points" to separate the arrows.
2. By contrast, in **Mon**, for homomorphisms $h, j : M \to N$ we always have $hx = jx$, for all $x : 1 \to M$, since there's just one such point x. Thus monoids do not "have enough points."
3. But in any category $\mathbf{C}$, and for any arrows $f, g : C \to C'$, we always have $f = g$ iff for all $x : D \to C$, it holds that $fx = gx$ (why?). Thus, every object has enough generalized elements.
4. In fact, it often happens that it is enough to consider generalized elements of just a certain form $T \to A$, that is, for certain "test" objects T. We shall consider this presently.

Generalized elements are also good for "testing" for various conditions. Consider, for instance, the following diagram.

$$X \underset{x'}{\overset{x}{\rightrightarrows}} A \xrightarrow{f} B$$

The arrow f is monic iff $x \neq x'$ implies $fx \neq fx'$, that is, just if f is "injective on generalized elements."

Similarly, in any category $\mathbf{C}$, to test whether a square commutes

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle g}\downarrow & & \downarrow{\scriptstyle \alpha} \\ D & \xrightarrow[\beta]{} & C \end{array}$$

we shall have $\alpha f = \beta g$ just if $\alpha f x = \beta g x$ for all generalized elements $x : X \to A$ (just take $x = 1_A : A \to A$!).

Example 2.11. Generalized elements can be used to "reveal more structure" than do the constant elements. For example, consider the following posets X and A:

$$A = \{a \leq b \leq c\}$$
$$X = \{x \leq y, x \leq z\}$$

There is an order-preserving bijection $f : X \to A$ defined by

$$f(x) = a, \qquad f(y) = b, \qquad f(z) = c.$$

It is easy to see that f is both monic and epic in the category **Pos**; however, it is clearly not an iso. One would like to say that X and A are "different structures," and indeed, their being non-isomorphic says just this. But now, how to *prove* that they are *not* isomorphic (say, via some other $X \to A$)? In general, this can be quite difficult.

One way to prove that two objects are not isomorphic is to use "invariants": attributes that are preserved by isomorphisms. If two objects differ by an invariant they cannot be isomorphic. Generalized elements provide an easy way to define invariants. For instance, the number of global elements of X and A is the same, namely the three elements of the sets. But consider instead the **2**-elements $\mathbf{2} \to X$. Then X has 5 such, and A has 6. Since these numbers are invariants, the posets cannot be isomorphic. In more detail, we can define for any poset P the numerical invariant

$$|\operatorname{Hom}(2, P)| \quad = \quad \text{number of elements of } \operatorname{Hom}(2, P).$$

Then if $P \cong Q$, it is easy to see that $|\operatorname{Hom}(2, P)| = |\operatorname{Hom}(2, Q)|$, since any isomorphism

$$P \underset{g}{\overset{f}{\rightleftarrows}} Q$$

also gives an iso

$$\operatorname{Hom}(2, P) \underset{g_*}{\overset{f_*}{\rightleftarrows}} \operatorname{Hom}(2, Q)$$

by composition:

$$f_*(h) = fh$$
$$g_*(k) = gk$$

for all $h : 2 \to P$ and $k : 2 \to Q$.

Example 2.12. As in the foregoing example, it is often the case that generalized elements $t : T \to A$ "based at" certain objects T are especially "revealing." We can think of such elements geometrically as "figures of shape T in A," just as an arrow $\mathbf{2} \to P$ in posets is a figure of shape $p \leq p'$ in P. For instance, as we have already seen, in the category of monoids, the arrows from the terminal monoid 1 are entirely uninformative, but those from the free monoid $M(*)$ on one generator suffice to distinguish homomorphisms, in the sense that two homomorphisms $f, g : M \to N$ are equal if their composites with all such arrows are equal. In fact, for monoids the underlying set $|M|$ is plainly (isomorphic to) $\operatorname{Hom}_{\mathbf{Mon}}(M(*), M)$.

2.4 Sections and retractions

We already noted that any iso is both monic and epic. More generally, if an arrow

$$f : A \to B$$

has a left inverse

$$g : B \to A, \quad gf = 1_A$$

then f must be monic and g epic, by an easy exercise.

Definition 2.13. A *split* mono (epi) is an arrow with a left (right) inverse.

Terminology: Given arrows $e : X \to A$ and $s : A \to X$ such that $es = 1_A$, then s is called a *section* or *splitting* of e, and e is called a *retraction* of s. The object A is called a *retract* of X.

Remark 2.14. Since functors preserve identities, they also preserve *split* epis and *split* monos. Compare the example above in **Mon** where the forgetful functor

$$\mathbf{Mon} \to \mathbf{Set}$$

did not preserve the epi $\mathbb{N} \to \mathbb{Z}$.

Example 2.15. In **Sets**, every mono splits except those of the form

$$\emptyset \rightarrowtail A.$$

The condition that *every epi splits* is the categorical version of the axiom of choice. Indeed, consider an epi

$$e : E \twoheadrightarrow X.$$

We have the family of nonempty sets

$$E_x = e^{-1}\{x\}, \quad x \in X.$$

A splitting of e is exactly a choice function for this family $(E_x)_{x\in X}$, that is, a function $s : X \to E$ such that $es = 1_X$, since this means that $s(x) \in E_x$ for all $x \in X$.

Conversely, given a family of nonempty sets,

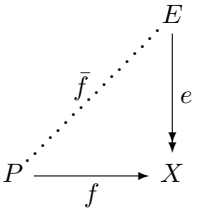
$$(E_x)_{x\in X}$$

take $E = \{(x, y) \mid x \in X,\ y \in E_x\}$ and define the epi $e : E \twoheadrightarrow X$ by $(x, y) \mapsto x$. A splitting s of e then determines a choice function for the family.

The idea that a "family of objects" $(E_x)_{x\in X}$ can be represented by a single arrow $e : E \to X$ by using the "fibers" $e^{-1}(x)$ has much wider application than this, and will be considered further in Section 7.9.

A notion related to the existence of "choice functions" is that of being "projective": an object P is said to be *projective* if for any epi $e : E \twoheadrightarrow X$ and arrow

$f : P \to X$ there is some (not necessarily unique) arrow $\bar{f} : P \to E$ such that $e \circ \bar{f} = f$, as indicated in the following diagram:



One says that f *lifts across* e. Projective objects may be thought of as having "less structure," thus permitting "more arrows."

The axiom of choice implies that all sets are projective, and it follows that free objects in many (but not all!) categories of algebras then are also projective. The reader should show that, in any category, any retract of a projective object is also projective.

2.5 Products

Next we are going to see the categorical definition of a product of two objects in a category. This was first given by Mac Lane in 1950, and it is probably the earliest example of category theory being used to define a fundamental mathematical notion.

By "define" here I mean an abstract characterization, in the sense already used, in terms of objects and arrows in a category. And as before, we do this by giving a UMP, which determines the structure at issue up to isomorphism, as usual in category theory. Later in this chapter, we will have several other examples of such characterizations.

Let us begin by considering products of sets. Given sets A and B the *cartesian product* of A and B is the set of ordered pairs

$$A \times B = \{(a, b) \mid a \in A,\ b \in B\}.$$

Observe that there are two "coordinate projections"

$$A \xleftarrow{\pi_1} A \times B \xrightarrow{\pi_2} B$$

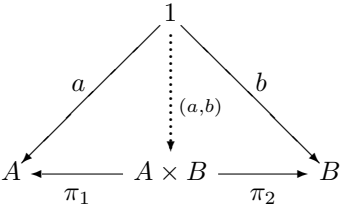
with

$$\pi_1(a, b) = a, \qquad \pi_2(a, b) = b.$$

And indeed, given any element $c \in A \times B$ we have

$$c = (\pi_1 c, \pi_2 c).$$

The situation is captured concisely in the following diagram:



Replacing elements by generalized elements, we get the following definition.

Definition 2.16. In any category **C**, a *product diagram* for the objects A and B consists of an object P and arrows

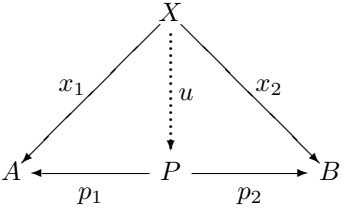
$$A \xleftarrow{p_1} P \xrightarrow{p_2} B$$

satisfying the following UMP:

Given any diagram of the form

$$A \xleftarrow{x_1} X \xrightarrow{x_2} B$$

there exists a unique $u : X \to P$, making the diagram



commute, that is, such that $x_1 = p_1 u$ and $x_2 = p_2 u$.

Remark 2.17. As in other UMPs, there are two parts:

Existence: There is some $u : X \to U$ such that $x_1 = p_1 u$ and $x_2 = p_2 u$.
Uniqueness: Given any $v : X \to U$, if $p_1 v = x_1$ and $p_2 v = x_2$, then $v = u$.

Proposition 2.18. *Products are unique up to isomorphism.*

Proof. Suppose

$$A \xleftarrow{p_1} P \xrightarrow{p_2} B$$

and

$$A \xleftarrow{q_1} Q \xrightarrow{q_2} B$$

are products of A and B. Then there is a unique $i : P \to Q$ such that $q_1 \circ i = p_1$ and $q_2 \circ i = p_2$. Similarly, there is a unique $j : Q \to P$ such that $p_1 \circ j = q_1$

and $p_2 \circ j = q_2$. Thus, $p_1 \circ j \circ i = p_1$ and $p_2 \circ j \circ i = p_2$. Since $p_1 \circ 1_P = p_1$ and $p_2 \circ 1_P = p_2$, it follows from the uniqueness condition that $j \circ i = 1_P$. Similarly, $i \circ j = 1_Q$. Thus, $i : P \to Q$ is an isomorphism. □

If A and B have a product, we write

$$A \xleftarrow{p_1} A \times B \xrightarrow{p_2} B$$

for one such product. Then given X, x_1, x_2 as in the definition, we write

$$\langle x_1, x_2 \rangle \text{ for } u : X \to A \times B.$$

Note, however, that a pair of objects may have many different products in a category. For example, given a product $A \times B, p_1, p_2$, and any iso $h : A \times B \to Q$, the diagram $Q, p_1 \circ h, p_2 \circ h$ is also a product of A and B.

Now an arrow *into* a product

$$f : X \to A \times B$$

is "the same thing" as a pair of arrows

$$f_1 : X \to A, \qquad f_2 : X \to B.$$

So we can essentially forget about such arrows, in that they are uniquely determined by pairs of arrows. But something useful *is* gained if a category has products; namely, consider arrows *out of* the product,

$$g : A \times B \to Y.$$

Such a g is a "function in two variables"; given any two generalized elements $f_1 : X \to A$ and $f_2 : X \to B$, we have an element $g\langle f_1, f_2 \rangle : X \to Y$. Such arrows $g : A \times B \to Y$ are not "reducible" to anything more basic, the way arrows into products were (to be sure, they are related to the notion of an "exponential" Y^B, via "currying" $\lambda f : A \to Y^B$; we discuss this further in chapter 6).

2.6 Examples of products

1. We have already seen cartesian products of sets. Note that if we choose a different definition of ordered pairs $\langle a, b \rangle$ we get different sets

   $$A \times B \quad \text{and} \quad A \times' B$$

   each of which is (part of) a product, and so are isomorphic. For instance, we could set:

   $$\langle a, b \rangle = \{\{a\}, \{a, b\}\}$$
   $$\langle a, b \rangle' = \langle a, \langle a, b \rangle \rangle$$

2. Products of "structured sets" like monoids or groups can often be constructed as products of the underlying sets with *componentwise* operations: If G and H are groups, for instance, $G \times H$ can be constructed by taking the underlying set of $G \times H$ to be the set $\{\langle g, h\rangle \mid g \in G,\ h \in H\}$ and defining the binary operation by

$$\langle g, h\rangle \cdot \langle g', h'\rangle = \langle g \cdot g', h \cdot h'\rangle$$

the unit by

$$u = \langle u_G, u_H\rangle$$

and inverses by

$$\langle a, b\rangle^{-1} = \langle a^{-1}, b^{-1}\rangle.$$

The projection homomorphisms $G \times H \to G$ (or H) are the evident ones $\langle g, h\rangle \mapsto g$ (or h).

3. Similarly, for categories $\mathbf{C}$ and $\mathbf{D}$, we already defined the category of pairs of objects and arrows,

$$\mathbf{C} \times \mathbf{D}.$$

Together with the evident projection functors, this is indeed a product in **Cat** (when $\mathbf{C}$ and $\mathbf{D}$ are small).

As a special case, we also get products of posets and of monoids as products of categories.

4. Let P be a poset and consider a product of elements $p, q \in P$. We must have projections

$$p \times q \leq p$$
$$p \times q \leq q$$

and if for any element x,

$$x \leq p, \quad \text{and} \quad x \leq q$$

then we need

$$x \leq p \times q.$$

Do you recognize this operation $p \times q$? It is just what is usually called the *greatest lower bound*: $p \times q = p \wedge q$. Many other order-theoretic notions are also special cases of categorical ones, as we shall see later.

5. (For those who know something about Topology.) Let us show that the product of two *topological spaces* X, Y, as usually defined, really is a product

in **Top**, the category of spaces and continuous functions. Thus suppose we have spaces X and Y and the product spaces $X \times Y$ with its projections

$$X \xleftarrow{p_1} X \times Y \xrightarrow{p_2} Y$$

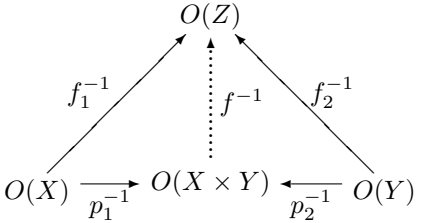
Recall that $O(X \times Y)$ is generated by basic open sets of the form $U \times V$ where $U \in O(X)$ and $V \in O(Y)$, so every $W \in O(X \times Y)$ is a union of such basic opens.

- Clearly p_1 is continuous, since $p_1^{-1}U = U \times Y$.
- Given any continuous $f_1 : Z \to X, f_2 : Z \to Y$, let $f : Z \to X \times Y$ be the function $f = \langle f_1, f_2 \rangle$. We just need to see that f is continuous.
- Given any $W = \bigcup_i (U_i \times V_i) \in O(X \times Y)$, $f^{-1}(W) = \bigcup_i f^{-1}(U_i \times V_i)$, so it suffices to show $f^{-1}(U \times V)$ is open. But

$$\begin{aligned} f^{-1}(U \times V) &= f^{-1}((U \times Y) \cap (X \times V)) \\ &= f^{-1}(U \times Y) \cap f^{-1}(X \times V) \\ &= f^{-1} \circ p_1^{-1}(U) \cap f^{-1} \circ p_2^{-1}(V) \\ &= (f_1)^{-1}(U) \cap (f_2)^{-1}(V) \end{aligned}$$

where $(f_1)^{-1}(U)$ and $(f_2)^{-1}(V)$ are open, since f_1 and f_2 are continuous.

The following diagram concisely captures the situation at hand:



6. (For those familiar with type theory.) Let us consider the *category of types* of the (simply typed) λ-calculus. The λ–calculus is a formalism for the specification and manipulation of functions, based on the notions of "binding of variables" and functional evaluation. For example, given the real polynomial expression $x^2 + 2y$, in the λ-calculus one writes $\lambda y.x^2 + 2y$ for the function $y \mapsto x^2 + 2y$ (for each fixed value x), and $\lambda x \lambda y.x^2 + 2y$ for the function-valued function $x \mapsto (y \mapsto x^2 + 2y)$.
Formally, the λ-calculus consists of:

- Types: $A \times B,\ A \to B, \ldots$ (generated from some basic types)
- Terms:

$$\begin{aligned} x, y, z, \ldots : A \quad &\text{(variables for each type } A\text{)} \\ a : A,\ b : B, \ldots \quad &\text{(possibly some typed constants)} \end{aligned}$$

$$\begin{aligned}
\langle a, b\rangle : A \times B &\quad (a : A,\ b : B)\\
\mathrm{fst}(c) : A &\quad (c : A \times B)\\
\mathrm{snd}(c) : B &\quad (c : A \times B)\\
ca : B &\quad (c : A \to B,\ a : A)\\
\lambda x.b : A \to B &\quad (x : A,\ b : B)
\end{aligned}$$

- Equations:

$$\begin{aligned}
\mathrm{fst}(\langle a, b\rangle) &= a\\
\mathrm{snd}(\langle a, b\rangle) &= b\\
\langle \mathrm{fst}(c), \mathrm{snd}(c)\rangle &= c\\
(\lambda x.b)a &= b[a/x]\\
\lambda x.cx &= c \quad (\text{no } x \text{ in } c)
\end{aligned}$$

The relation $a \sim b$ (usually called *$\beta\eta$-equivalence*) on terms is defined to be the equivalence relation generated by the equations, and renaming of bound variables:

$$\lambda x.b = \lambda y.b[y/x] \quad (\text{no } y \text{ in } b)$$

The category of types $\mathbf{C}(\lambda)$ is now defined as follows:

- objects: the types,
- arrows $A \to B$: closed terms $c : A \to B$, identified if $c \sim c'$,
- identities: $1_A = \lambda x.x$ (where $x : A$),
- composition: $c \circ b = \lambda x.c(bx)$.

Let us verify that this is a well-defined category:
Unit laws:

$$\begin{aligned}
c \circ 1_B &= \lambda x(c((\lambda y.y)x)) = \lambda x(cx) = c\\
1_C \circ c &= \lambda x((\lambda y.y)(cx)) = \lambda x(cx) = c
\end{aligned}$$

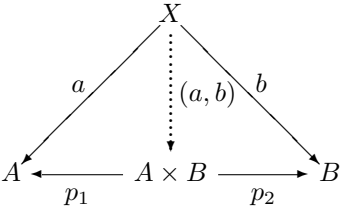
Associativity:

$$\begin{aligned}
c \circ (b \circ a) &= \lambda x(c((b \circ a)x))\\
&= \lambda x(c((\lambda y.b(ay))x))\\
&= \lambda x(c(b(ax)))\\
&= \lambda x(\lambda y(c(by))(ax))\\
&= \lambda x((c \circ b)(ax))\\
&= (c \circ b) \circ a
\end{aligned}$$

This category has binary products. Indeed, given types A and B, let

$$p_1 = \lambda z.\mathrm{fst}(z), \quad p_2 = \lambda z.\mathrm{snd}(z) \quad (z : A \times B).$$

And given a and b as in



let

$$(a, b) = \lambda x.\langle ax, bx\rangle.$$

Then

$$\begin{aligned} p_1 \circ (a, b) &= \lambda x(p_1((\lambda y.\langle ay, by\rangle)x)) \\ &= \lambda x(p_1\langle ax, bx\rangle) \\ &= \lambda x(ax) \\ &= a. \end{aligned}$$

Similarly, $p_2 \circ (a, b) = b$.

Finally, if $c : X \to A \times B$ also has

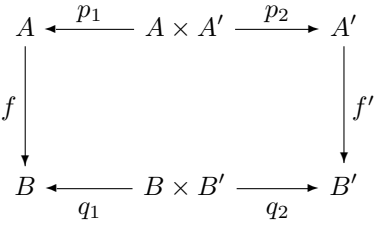
$$p_1 \circ c = a, \qquad p_2 \circ c = b$$

then

$$\begin{aligned} (a, b) &= \lambda x.\langle ax, bx\rangle \\ &= \lambda x.\langle (p_1 \circ c)x, (p_2 \circ c)x\rangle \\ &= \lambda x.\langle (\lambda y(p_1(cy)))x, (\lambda y(p_2(cy)))x\rangle \\ &= \lambda x.\langle (\lambda y((\lambda z.\mathrm{fst}(z))(cy)))x, (\lambda y((\lambda z.\mathrm{snd}(z))(cy)))x\rangle \\ &= \lambda x.\langle \lambda y(\mathrm{fst}(cy))x, \lambda y(\mathrm{snd}(cy))x\rangle \\ &= \lambda x.\langle \mathrm{fst}(cx), \mathrm{snd}(cx)\rangle \\ &= \lambda x.(cx) \\ &= c. \end{aligned}$$

2.7 Categories with products

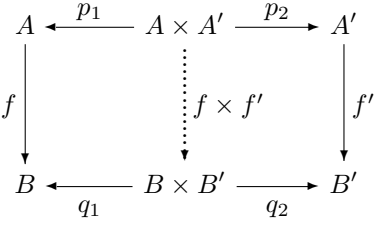
Let $\mathbf{C}$ be a category that has a product diagram for every pair of objects. Suppose we have objects and arrows



with indicated products. Then we write

$$f \times f' : A \times A' \to B \times B'$$

for $f \times f' = \langle f \circ p_1, f' \circ p_2 \rangle$. Thus, both squares in the following diagram commute.



In this way, if we choose a product for each pair of objects, we get a functor

$$\times : \mathbf{C} \times \mathbf{C} \to \mathbf{C}$$

as the reader can easily check, using the UMP of the product. A category which has a product for every pair of objects is said to *have binary products.*

We can also define ternary products

$$A_1 \times A_2 \times A_3$$

with an analogous UMP (there are three projections $p_i : A_1 \times A_2 \times A_3 \to A_i$, and for any object X and three arrows $x_i : X \to A_i$, there is a unique arrow $u : X \to A_1 \times A_2 \times A_3$ such that $p_i u = x_i$ for each of the three i's.) Plainly, such a condition can be formulated for any number of factors.

It is clear, however, that if a category has binary products, then it has all finite products with two or more factors; for instance, one could set

$$A \times B \times C = (A \times B) \times C$$

to satisfy the UMP for ternary products. On the other hand, one could instead have taken $A \times (B \times C)$ just as well. This shows that the binary product operation

$A \times B$ is associative up to isomorphism, for we must have

$$(A \times B) \times C \cong A \times (B \times C)$$

by the UMP of ternary products.

Observe also that a terminal object is a "null-ary" product, that is, a product of no objects:

> Given no objects, there is an object 1 with no maps, and given any other object X and no maps, there is a unique arrow
>
> $$! : X \to 1$$
>
> making nothing further commute.

Similarly, any object A is the *unary product* of A with itself one time.

Finally, one can also define the product of a family of objects $(C_i)_{i \in I}$ indexed by *any* set I, by giving a UMP for "I-ary products" analogous to those for nullary, unary, binary, and n-ary products. We leave the precise formulation of this UMP as an exercise.

Definition 2.19. A category $\mathbf{C}$ is said to *have all finite products* if it has a terminal object and all binary products (and therewith products of any finite cardinality). The category $\mathbf{C}$ *has all (small) products* if every set of objects in $\mathbf{C}$ has a product.

2.8 Hom-sets

In this section, we assume that all categories are locally small.

Recall that in any category $\mathbf{C}$, given any objects A and B, we write

$$\mathrm{Hom}(A, B) = \{f \in \mathbf{C} \mid f : A \to B\}$$

and call such a set of arrows a *Hom-set*. Note that any arrow $g : B \to B'$ in $\mathbf{C}$ induces a function:

$$\mathrm{Hom}(A, g) : \mathrm{Hom}(A, B) \to \mathrm{Hom}(A, B')$$
$$(f : A \to B) \mapsto (g \circ f : A \to B \to B')$$

Thus, $\mathrm{Hom}(A, g) = g \circ f$; one sometimes writes g_* instead of $\mathrm{Hom}(A, g)$, so

$$g_*(f) = g \circ f.$$

Let us show that this determines a functor

$$\mathrm{Hom}(A, -) : \mathbf{C} \to \mathbf{Sets}$$

called the (covariant) *representable functor* of A.

We need to show that

$$\mathrm{Hom}(A, 1_X) = 1_{\mathrm{Hom}(A,X)}$$

and that

$$\mathrm{Hom}(A, g \circ f) = \mathrm{Hom}(A, g) \circ \mathrm{Hom}(A, f).$$

Taking an argument $x : A \to X$, we clearly have

$$\begin{aligned} \mathrm{Hom}(A, 1_X)(x) &= 1_X \circ x \\ &= x \\ &= 1_{\mathrm{Hom}(A,X)}(x) \end{aligned}$$

and

$$\begin{aligned} \mathrm{Hom}(A, g \circ f)(x) &= (g \circ f) \circ x \\ &= g \circ (f \circ x) \\ &= \mathrm{Hom}(A, g)(\mathrm{Hom}(A, f)(x)). \end{aligned}$$

We will study such representable functors much more carefully later. For now we just want to see how one can use Hom-sets to give another definition of products.

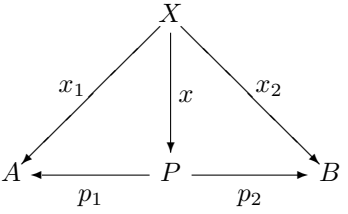
An object P with arrows $p_1 : P \to A$ and $p_2 : P \to B$ is an element (p_1, p_2) of the set

$$\mathrm{Hom}(P, A) \times \mathrm{Hom}(P, B).$$

And similarly for any set X in place of P. Now, given any arrow

$$x : X \to P$$

composing with p_1 and p_2 gives a pair of arrows $x_1 = p_1 \circ x : X \to A$ and $x_2 = p_2 \circ x : X \to B$, as indicated in the following diagram.



In this way, we have a function

$$\vartheta_X = \langle \mathrm{Hom}(X, p_1), \mathrm{Hom}(X, p_2) \rangle : \mathrm{Hom}(X, P) \to \mathrm{Hom}(X, A) \times \mathrm{Hom}(X, B)$$

defined by

$$\vartheta_X(x) = (x_1, x_2) \tag{2.1}$$

This function ϑ_X can be used to express concisely the condition of being a product as follows.

Proposition 2.20. *A diagram of the form*

$$A \xleftarrow[p_1]{} P \xrightarrow[p_2]{} B$$

is a product for A and B iff for every object X, the canonical function ϑ_X given in (2.1) *is an isomorphism,*

$$\vartheta_X : \mathrm{Hom}(X, P) \cong \mathrm{Hom}(X, A) \times \mathrm{Hom}(X, B).$$

Proof. Examine the UMP of the product: it says exactly that for every element $(x_1, x_2) \in \mathrm{Hom}(X, A) \times \mathrm{Hom}(X, B)$, there is a unique $x \in \mathrm{Hom}(X, P)$ such that $\vartheta_X(x) = (x_1, x_2)$, that is, ϑ_X is bijective. □

Definition 2.21. Let $\mathbf{C}$, $\mathbf{D}$ be categories with binary products. A functor $F : \mathbf{C} \to \mathbf{D}$ is said to *preserve binary products* if it takes every product diagram

$$A \xleftarrow[p_1]{} A \times B \xrightarrow[p_2]{} B \qquad \text{in } \mathbf{C}$$

to a product diagram

$$FA \xleftarrow[Fp_1]{} F(A \times B) \xrightarrow[Fp_2]{} FB \qquad \text{in } \mathbf{D}.$$

It follows that F preserves products just if

$$F(A \times B) \cong FA \times FB \quad \text{(canonically)}$$

that is, iff the canonical "comparison arrow"

$$\langle Fp_1, Fp_2 \rangle : F(A \times B) \to FA \times FB$$

is an iso.

For example, the forgetful functor $U : \mathbf{Mon} \to \mathbf{Sets}$ preserves binary products.

Corollary 2.22. *For any object X in a category $\mathbf{C}$ with products, the (covariant) representable functor*

$$\mathrm{Hom}_{\mathbf{C}}(X, -) : \mathbf{C} \to \mathbf{Sets}$$

preserves products.

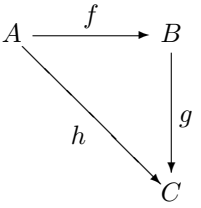
Proof. For any $A, B \in \mathbf{C}$, the foregoing proposition says that there is a canonical isomorphism:

$$\mathrm{Hom}_{\mathbf{C}}(X, A \times B) \cong \mathrm{Hom}_{\mathbf{C}}(X, A) \times \mathrm{Hom}_{\mathbf{C}}(X, B)$$

□

2.9 Exercises

1. Show that a function between sets is surjective if it is an epimorphism in **Sets**.
2. With regard to a commutative triangle,



   in any category **C**, show

   a. if f and g are isos (resp. monos, resp. epis), so is h;

   b. if h is monic, so is f;

   c. if h is epic, so is g;

   d. (by example) if h is monic, g need not be.

3. Show that all sets are projective (use the axiom of choice). Show that the epis among posets are the surjections (on objects), and that the one-element poset **1** is projective. Finally, show that in any category, any retract of a projective object is also projective.
4. Let A be a set. Define an *A-monoid* to be a monoid M equipped with a function $m : A \to U(M)$ (to the underlying set of M). A morphism $h : (M, m) \to (N, n)$ of A-monoids is to be a monoid homomorphism $h : M \to N$ such that $U(h) \circ m = n$ (a commutative triangle). Together with the evident identities and composites, this defines a category $A - \mathbf{Mon}$ of A-monoids.

   Show that an initial object in A-**Mon** is the same thing as a free monoid $M(A)$ on A. (Hint: compare their respective UMPs.)
5. Show that for any Boolean algebra B, Boolean homomorphisms $h : B \to \mathbf{2}$ correspond exactly to ultrafilters in B.
6. In any category with binary products, show directly that:

$$A \times (B \times C) \cong (A \times B) \times C.$$

7. (a) For any index set I, define the product $\prod_{i \in I} X_i$ of an I-indexed family of objects $(X_i)_{i \in I}$ in a category, by giving a UMP generalizing that for binary products (the case $I = 2$).

   (b) Show that in **Sets**, for any set X the set X^I of all functions $f : I \to X$ has this UMP, with respect to the "constant family" where $X_i = X$ for all $i \in I$, and thus

$$X^I \cong \prod_{i \in I} X$$

This page intentionally left blank

3

DUALITY

We have seen a few examples of definitions and statements which exhibit a kind of "duality," like initial and terminal object and epimorphisms and monomorphisms. We now want to consider this duality more systematically. Despite its rather trivial first impression, it is indeed a deep and powerful aspect of the categorical approach.

3.1 The duality principle

First, let us look again at the formal definition of a category: There are two kinds of things, objects A, B, C and $\ldots$, arrows $f, g, h, \ldots$; four operations $\mathrm{dom}(f)$, $\mathrm{cod}(f)$, 1_A, $g \circ f$; and these satisfy the following seven axioms:

$$\begin{aligned} \mathrm{dom}(1_A) = A, \qquad & \mathrm{cod}(1_A) = A \\ f \circ 1_{\mathrm{dom}(f)} = f, \qquad & 1_{\mathrm{cod}(f)} \circ f = f \\ \mathrm{dom}(g \circ f) = \mathrm{dom}(f), \qquad & \mathrm{cod}(g \circ f) = \mathrm{cod}(g) \\ h \circ (g \circ f) = (h \circ g) \circ f & \end{aligned}$$

The operation "$g \circ f$" is only defined where

$$\mathrm{dom}(g) = \mathrm{cod}(f),$$

so a suitable form of this should occur as a condition on each equation containing $\circ$, as in $\mathrm{dom}(g) = \mathrm{cod}(f) \Rightarrow \mathrm{dom}(g \circ f) = \mathrm{dom}(f)$.

Now, given any sentence Σ in the elementary language of category theory, we can form the "dual statement" Σ^* by making the following replacements:

$$g \circ f \quad \text{for} \quad f \circ g$$

$$\mathrm{cod} \quad \text{for} \quad \mathrm{dom}$$

$$\mathrm{dom} \quad \text{for} \quad \mathrm{cod}.$$

It is easy to see that then Σ^* will also be a well-formed sentence. Next, suppose we have shown a statement Σ to entail one Δ, without using any of the category axioms, then clearly $\Sigma^* \vdash \Delta^*$, since the substituted terms are treated as mere

undefined constants. But now observe that the axioms for category theory CT are themselves "self-dual," in the sense that we have:

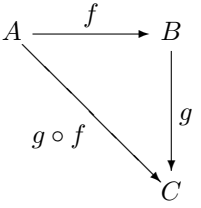
$$\mathrm{CT}^* = \mathrm{CT}$$

We therefore have the following *duality principle*:

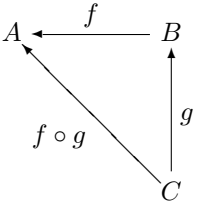
Proposition 3.1 (Formal duality). *For any statement* Σ *in the language of category theory, if* Σ *follows from the axioms for categories, then so does* Σ^**:*

$$\mathrm{CT} \vdash \Sigma \textit{ implies } \mathrm{CT} \vdash \Sigma^*$$

Taking a more conceptual point of view, note that if Σ involves some diagram of objects and arrows,



then Σ^* involves the diagram obtained from it by reversing direction and the order of compositions of arrows.



Recalling the opposite category $\mathbf{C}^{\mathrm{op}}$ of a category $\mathbf{C}$, we see that an interpretation of a statement Σ in $\mathbf{C}$ automatically gives an interpretation of Σ^* in $\mathbf{C}^{\mathrm{op}}$.

Now, since for every category $\mathbf{C}$,

$$(\mathbf{C}^{\mathrm{op}})^{\mathrm{op}} = \mathbf{C} \tag{3.1}$$

the conceptual form of the duality principle then results similarly in

Proposition 3.2 (Conceptual duality). *For any statement* Σ *about categories, if* Σ *holds for all categories, then so does the dual statement* Σ^**.*

Proof. If Σ holds for all categories $\mathbf{C}$, then it also holds in all categories $\mathbf{C}^{\mathrm{op}}$, but then Σ^* holds in all categories $(\mathbf{C}^{\mathrm{op}})^{\mathrm{op}}$, thus in all categories $\mathbf{C}$. □

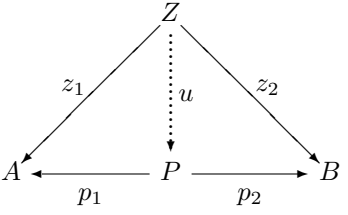
It may seem that only very simple or trivial properties such as "having a terminal object" are going to be subject to this sort of duality, but in fact this is far from

being so. Categorical duality turns out to be a very powerful and far-reaching phenomenon, as we shall see later. One way this occurs is that, rather than considering statements about all categories, we can also consider the dual of an abstract definition or property of objects and arrows, like "being a product diagram." The dual property is arrived at by reversing the order of composition and the words "dom" and "cod." Equivalently, it results from interpreting the original property in the opposite category. The next section provides an example of this procedure.

3.2 Coproducts

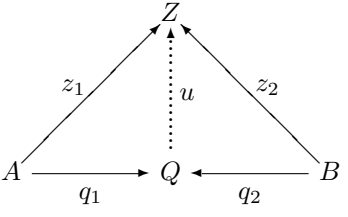
Let us consider the example of products and see what the dual notion must be. First, recall the definition of a product.

Definition 3.3. A diagram $A \overset{p_1}{\leftarrow} P \overset{p_2}{\rightarrow} B$ is a *product* of A and B, if for any Z and $A \overset{z_1}{\leftarrow} Z \overset{z_2}{\rightarrow} B$ there is a unique $u : Z \to P$ with $p_i \circ u = z_i$, all as indicated in



Now what is the dual statement?

A diagram $A \overset{q_1}{\rightarrow} Q \overset{q_2}{\leftarrow} B$ is a "dual-product" of A and B if for any Z and $A \overset{z_1}{\rightarrow} Z \overset{z_2}{\leftarrow} B$ there is a unique $u : Q \to Z$ with $u \circ q_i = z_i$, all as indicated in



Actually, these are called *coproducts*; the convention is to use the prefix "co-" to indicate the dual notion. We usually write $A \overset{i_1}{\rightarrow} A + B \overset{i_2}{\leftarrow} B$ for the coproduct and $[f, g]$ for the uniquely determined arrow $u : A + B \to Z$. The "coprojections" $i_1 : A \to A + B$ and $i_2 : B \to A + B$ are usually called *injections*, even though they need not be "injective" in any sense.

A coproduct of two objects is therefore exactly their product in the opposite category. Of course, this immediately gives lots of examples of coproducts. But what about some more familiar ones?

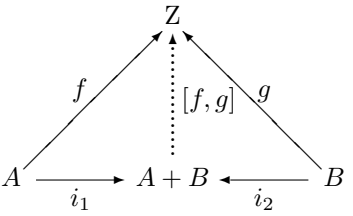
Example 3.4. In **Sets**, the coproduct $A + B$ of two sets is their disjoint union,

$$A + B = \{(a, 1) \mid a \in A\} \cup \{(b, 2) \mid b \in B\}$$

with evident coproduct injections

$$i_1(a) = (a, 1), \qquad i_2(b) = (b, 2).$$

Given any functions f and g as in:



we define

$$[f, g](x, \delta) = \begin{cases} f(x) & \delta = 1 \\ g(x) & \delta = 2. \end{cases}$$

If, in addition, $h \circ i_1 = f$ and $h \circ i_2 = g$, then for any $(x, \delta) \in A + B$, we must have

$$h(x, \delta) = [f, g](x, \delta)$$

as can be easily calculated.

Note that in **Sets**, every finite set A is a coproduct:

$$A \cong 1 + 1 + \cdots + 1 \ \ (n\text{-times})$$

for $n = \mathrm{card}(A)$. This is because a function $f : A \to Z$ is uniquely determined by its values $f(a)$ for all $a \in A$. So we have

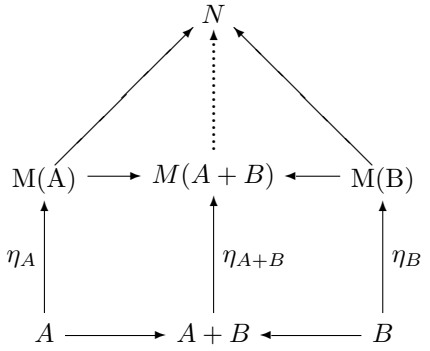
$$\begin{aligned} A &\cong \{a_1\} + \{a_n\} + \cdots + \{a_n\} \\ &\cong 1 + 1 + \cdots + 1 \ \ (n\text{-times}). \end{aligned}$$

In this spirit, we often write simply $2 = 1 + 1, 3 = 1 + 1 + 1$, etc.

Example 3.5. If $M(A)$ and $M(B)$ are free monoids on sets A and B, then in **Mon** we can construct their coproduct as

$$M(A) + M(B) \cong M(A + B).$$

One can see this directly by considering words over $A + B$, but it also follows abstractly by using the diagram



in which the η's are the respective insertions of generators. The UMPs of $M(A)$, $M(B)$, $A + B$, and $M(A + B)$ then imply that the last of these has the required UMP of $M(A) + M(B)$.

It follows that the free monoid functor $M : \mathbf{Sets} \to \mathbf{Mon}$ preserves coproducts. This is an instance of a much more general phenomenon, which we will consider later, related to the fact we have already seen that the forgetful functor $U : \mathbf{Mon} \to \mathbf{Sets}$ is representable and so preserves products.

Example 3.6. In **Top** the coproduct of two spaces

$$X + Y$$

is their disjoint union with the topology $O(X + Y) \cong O(X) \times O(Y)$. Note that this follows the pattern of discrete spaces, for which $O(X) = P(X) \cong 2^X$. Thus, for discrete spaces we indeed have

$$O(X + Y) \cong 2^{X+Y} \cong 2^X \times 2^Y \cong O(X) \times O(Y).$$

Coproducts of posets are similarly constructed from the coproducts of the underlying sets, by "putting them side by side." What about "rooted" posets, that is, posets with a distinguished initial element 0? In the category Pos_0 of such posets and monotone maps that preserve 0, one constructs the coproduct of two such posets A and B from the coproduct $A + B$ in the category Pos of posets, by identifying the two different 0s,

$$A +_{\mathrm{Pos}_0} B = (A +_{\mathrm{Pos}} B)/(0_A = 0_B).$$

Example 3.7. In a fixed poset P, what is a coproduct of two elements $p, q \in P$? We have

$$p \leq p + q \quad \text{and} \quad q \leq p + q$$

and if

$$p \leq z \quad \text{and} \quad q \leq z$$

then

$$p + q \leq z.$$

So $p + q = p \vee q$ is the *join*, or "least upper bound," of p and q.

Example 3.8. Sum types in the λ-calculus as usually formulated using case terms are coproducts in the *category of types* defined in subsection 2.6.

Example 3.9. *Coproduct of monoids*

Two monoids A, B have a coproduct of the form

$$A + B = M(|A| + |B|)/\sim$$

where, as before, the free monoid $M(|A| + |B|)$ is strings (words) over the disjoint union $|A| + |B|$ of the underlying sets—the elements of A and B—and the equivalence relation $v \sim w$ is the least one containing the following equations

$$u^A = (-) = u^B$$
$$(\ldots aa' \ldots) = (\ldots a \cdot a' \ldots)$$
$$(\ldots bb' \ldots) = (\ldots b \cdot b' \ldots).$$

(If you need a refresher on quotienting a set by an equivalence relation, skip ahead and read the beginning of Section 3.4 now.) The unit is of course the equivalence class $[-]$ of the empty word. Multiplication of equivalence classes is also as expected, namely

$$[x \ldots y] \cdot [x' \ldots y'] = [x \ldots y x' \ldots y'].$$

The coproduct injections $i_A : A \to A + B$ and $i_B : B \to A + B$ are simply:

$$i_A(a) = [a], \qquad i_B(b) = [b].$$

Given any homomorphism $f : A \to M$ and $g : B \to M$ into a monoid M, the unique homomorphism

$$[f, g] : A + B \longrightarrow M$$

is defined by first lifting the function $[|f|, |g|] : |A| + |B| \to |M|$ to the free monoid $M(|A| + |B|)$, and then observing that if $v \sim w$ in $M(|A| + |B|)$, then $[|f|, |g|](v) = [|f|, |g|](w)$. Why is this homomorphism the *unique* one

$$h : M(|A| + |B|)/\sim \longrightarrow M$$

with $hi_A = f$ and $hi_B = g$?

This construction also works to give coproducts in **Groups**, where it is usually called the *free product* of A and B and written $A \oplus B$.

Example 3.10. For *abelian groups* A, B, the free product $A \oplus B$ need not be abelian. One could, of course, take a quotient of $A \oplus B$ to get a coproduct in the

category Ab of abelian groups, but there is a more convenient (and important) presentation, which we now consider.

Since the words in the free product $A \oplus B$ must be forced to satisfy the further commutativity conditions

$$(a_1 b_1 b_2 a_2 \ldots) \sim (a_1 a_2 \ldots b_1 b_2 \ldots)$$

we can shuffle all the a's to the front, and the b's to the back, of the words. But, furthermore, we already have

$$(a_1 a_2 \ldots b_1 b_2 \ldots) \sim (a_1 + a_2 + \cdots + b_1 + b_2 + \cdots).$$

Thus, we in effect have pairs of elements (a, b). So we take the product set as the underlying set of the coproduct

$$|A + B| = |A \times B|.$$

As inclusions, we use the homomorphisms

$$i_A(a) = (a, 0_B)$$
$$i_B(b) = (0_A, b).$$

Then given any homomorphisms $A \xrightarrow{f} X \xleftarrow{g} B$, we let $[f, g] : A + B \to X$ be defined by

$$[f, g](a, b) = f(a) +_X g(b)$$

which can easily be seen to do the trick (exercise!).

Proposition 3.11. *In the category* Ab *of abelian groups, there is a canonical isomorphism between the binary coproduct and product,*

$$A + B \cong A \times B.$$

Proof. To define an arrow $\vartheta : A + B \to A \times B$ we need one $A \to A \times B$ (and one $B \to A \times B$), so we need arrows $A \to A$ and $A \to B$ (and $B \to A$ and $B \to B$). For these we take $1_A : A \to A$ and the zero homomorphism $0_B : A \to B$ (and $0_A : B \to A$ and $1_B : B \to B$). Thus, all together we get

$$\vartheta = [\langle 1_A, 0_B \rangle, \langle 0_A, 1_B \rangle] : A + B \to A \times B.$$

Then given any $(a, b) \in A + B$, we have

$$\begin{aligned}
\vartheta(a, b) &= [\langle 1_A, 0_B \rangle, \langle 0_A, 1_B \rangle](a, b) \\
&= \langle 1_A, 0_B \rangle(a) + \langle 0_A, 1_B \rangle(b) \\
&= (1_A(a), 0_B(a)) + (0_A(b), 1_B(b)) \\
&= (a, 0_B) + (0_A, b) \\
&= (a + 0_A, 0_B + b) \\
&= (a, b).
\end{aligned}$$

□

This fact was first observed by Mac Lane, and it was shown to lead to a binary operation of addition on parallel arrows $f, g : A \to B$ between abelian groups (and related structures like modules and vector spaces). In fact, the group structure of a particular abelian group A can be recovered from this operation on arrows into A. More generally, the existence of such an addition operation on arrows can be used as the basis of an abstract description of categories like **Ab**, called "abelian categories," which are suitable for axiomatic homology theory.

Just as with products, one can consider the empty coproduct, which is an initial object 0, as well as coproducts of several factors, and the coproduct of two arrows,

$$f + f' : A + A' \to B + B'$$

which leads to a coproduct functor $+ : \mathbf{C} \times \mathbf{C} \to \mathbf{C}$ on categories $\mathbf{C}$ with binary coproducts. All of these facts follows simply by duality; that is, by considering the dual notions in the opposite category. Similarly, we have the following proposition.

Proposition 3.12. *Coproducts are unique up to isomorphism.*

Proof. Use duality and the fact that the dual of "isomorphism" is "isomorphism." □

In just the same way, one shows that binary coproducts are associative up to isomorphism, $(A + B) + C \cong A + (B + C)$.

In this way, in the future it will thus suffice to introduce new notions and then simply observe that the dual notions have analogous (but dual) properties. The next two sections give another example of this sort.

3.3 Equalizers

In this section, we consider another abstract characterization; this time a generalization of the idea of the kernel of a homomorphism or an equationally defined "variety," like the set of zeros of a real-valued function.

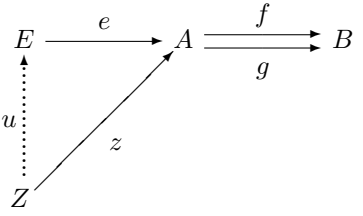
Definition 3.13. In any category $\mathbf{C}$, given parallel arrows

$$A \underset{g}{\overset{f}{\rightrightarrows}} B$$

an *equalizer* of f and g consists of E and $e : E \to A$, universal such that

$$f \circ e = g \circ e.$$

That is, given $z : Z \to A$ with $f \circ z = g \circ z$ there is a *unique* $u : Z \to E$ with $e \circ u = z$, all as in the diagram



Let us consider some simple examples.

Example 3.14. In **Sets**, given functions $f, g : A \rightrightarrows B$, their equalizer is the inclusion into A of the equationally defined subset

$$i : \{x \in A | f(x) = g(x)\} \hookrightarrow A.$$

Since if $fh(z) = gh(z)$ for some $h : Z \to A$, then $h(z) \in \{x \in A | f(x) = g(x)\}$ for all $z \in Z$, whence h "factors through" the inclusion function i, in the sense that there is a function $\bar{h} : Z \to \{x \in A | f(x) = g(x)\}$ such that $i \circ \bar{h} = h$. Observe that $\bar{h}$ is necessarily unique with this property, since i is monic.

Let us pause here to note that in fact, every subset $U \subseteq A$ is of this "equational" form, that is, is an equalizer for some pair of functions. Indeed, one can do this in a very canonical way; first let us put

$$2 = \{\top, \bot\}.$$

Then consider the *characteristic function*

$$\chi_U : A \to 2$$

defined for $x \in A$ by

$$\chi_U(x) = \begin{cases} \top & x \in U \\ \bot & x \notin U. \end{cases}$$

Thus we have

$$U = \{x \in A \mid \chi_U(x) = \top\}.$$

So the following is an equalizer

$$U \longrightarrow A \underset{\chi_U}{\overset{\top!}{\rightrightarrows}} 2$$

where $\top! = \top \circ ! : U \xrightarrow{!} 1 \xrightarrow{\top} 2$.

Moreover, for every function,

$$\varphi : A \to 2$$

we can form the "variety" (i.e. equational subset)

$$V_\varphi = \{x \in A \mid \varphi(x) = \top\}$$

as an equalizer, in the same way.

Now, it is easy to see that these operations χ_U and V_φ are mutually inverse:

$$\begin{aligned} V_{\chi_U} &= \{x \in A | \chi_U(x) = \top\} \\ &= \{x \in A | x \in U\} \\ &= U \end{aligned}$$

and given $\varphi : A \to 2$,

$$\begin{aligned} \chi_{V_\varphi}(x) &= \begin{cases} \top & x \in V_\varphi \\ \bot & x \notin V_\varphi \end{cases} \\ &= \begin{cases} \top & \varphi(x) = \top \\ \bot & \varphi(x) = \bot \end{cases} \\ &= \varphi(x). \end{aligned}$$

Therefore, we have an isomorphism

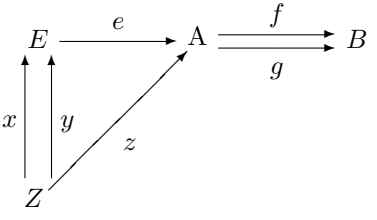
$$\mathrm{Hom}(A, 2) \cong P(A)$$

via the maps V and χ.

The fact that equalizers of functions can be taken as subsets is a special case of a more general phenomenon:

Proposition 3.15. *In any category, if $e : E \to A$ is an equalizer of some pair of arrows, then e is monic.*

Proof. Consider the diagram:



in which we assume e is the equalizer of f and g. Supposing $ex = ey$, we want to show $x = y$. Put $z = ex = ey$. Then $fz = fex = gex = gz$, so there is a *unique* $u : Z \to E$ such that $eu = z$. But from $ex = z$ and $ey = z$, it follows that $x = u = y$. □

Example 3.16. In many other categories, such as posets and monoids, the equalizer of a parallel pair of arrows $f, g : A \rightrightarrows B$ can be constructed by taking the equalizer of the underlying functions as above, that is, the subset $A(f = g) \subseteq A$

of elements $x \in A$ where f and g agree, $f(x) = g(x)$, and then restricting the structure of A to $A(f = g)$. For instance, in posets one takes the ordering from A restricted to this subset $A(f = g)$, and in topological spaces one takes the subspace topology.

In monoids, the subset $A(f = g)$ is then also a monoid with the operations from A, and the inclusion is therefore a homomorphism, because $f(u_A) = u_B = g(u_A)$, and if $f(a) = g(a)$ and $f(a') = g(a')$, then $f(a \cdot a') = f(a) \cdot f(a') = g(a) \cdot g(a') = g(a \cdot a')$.

In abelian groups, one has an alternate description of the equalizer, using the fact that,

$$f(x) = g(x) \quad \text{iff} \quad (f - g)(x) = 0.$$

Thus the equalizer of f and g is the same as that of the homomorphism $(f - g)$ and the zero homomorphism $0 : A \to B$, so it suffices to consider equalizers of the special form $A(h, 0) \rightarrowtail A$ for arbitrary homomorphisms $h : A \to B$. This subgroup of A is called the *kernel* of h, written $\ker(h)$. Thus we have the equalizer:

$$\ker(f - g) \hookrightarrow A \underset{g}{\overset{f}{\rightrightarrows}} B$$

The kernel of a homomorphism is of fundamental importance in the study of groups.

3.4 Coequalizers

A coequalizer is a generalization of a quotient by an equivalence relation, so let us begin by reviewing that notion. Recall first that an *equivalence relation* on a set X is a binary relation $x \sim y$ which is

reflexive: $x \sim x$,
symmetric: $x \sim y$ implies $y \sim x$,
transitive: $x \sim y$ and $y \sim z$ implies $x \sim z$.

Given such a relation, define the *equivalence class* $[x]$ of an element $x \in X$ by

$$[x] = \{y \in X |\ x \sim y\}.$$

The various different equivalence classes $[x]$ then form a *partition* of X, in the sense that every element y is in exactly one of them, namely $[y]$ (prove this!).

One sometimes thinks of an equivalence relation as arising from the equivalent elements having some property in common (like being the same color). One can then regard the equivalence classes $[x]$ as the properties and in that sense as "abstract objects" (the colors red, blue, etc., themselves). This is sometimes known as "definition by abstraction," and it describes the way that the real

numbers can be constructed from Cauchy sequences of rationals or the finite cardinal numbers from finite sets.

The set of all equivalence classes

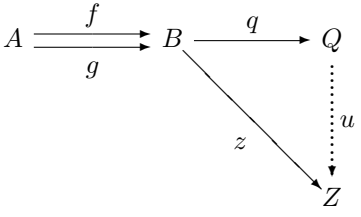
$$X/\sim \;=\; \{[x] \mid x \in X\}$$

may be called the *quotient* of X by $\sim$. It is used in place of X when one wants to "abstract away" the difference between equivalent elements $x \sim y$, in the sense that in $X/\sim$ such elements (and only such) are identified, since

$$[x] = [y] \quad \text{iff} \quad x \sim y.$$

Now let us consider the notion dual to that of equalizer, namely that of a coequalizer.

Definition 3.17. For any parallel arrows $f, g : A \to B$ in a category $\mathbf{C}$, a *coequalizer* consists of Q and $q : B \to Q$, universal with the property $qf = qg$, as in:

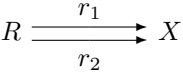


That is, given any Z and $z : B \to Z$, if $zf = zg$, then there exists a unique $u : Q \to Z$ such that $uq = z$.

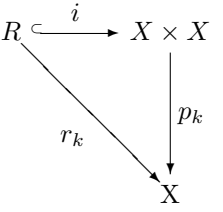
First observe that by duality, we know that such a coequalizer q in a category $\mathbf{C}$ is an equalizer in $\mathbf{C}^{\mathrm{op}}$, hence monic by the last proposition, and so q is epic in $\mathbf{C}$.

Proposition 3.18. *If $q : B \to Q$ is a coequalizer of some pair of arrows, then q is epic.*

Example 3.19. Let $R \subseteq X \times X$ be an equivalence relation on a set X, and consider the diagram



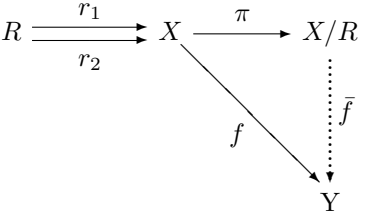
where the r's are the two projections of the inclusion $i : R \subseteq X \times X$, as indicated in the diagram



The canonical projection

$$\pi : X \longrightarrow X/R$$

defined by $x \mapsto [x]$ is then a coequalizer of r_1 and r_2. For given an $f : X \to Y$ as in:



there exists a function $\bar{f}$ such that

$$\bar{f}\pi(x) = f(x)$$

just in case f "respects R" in the sense that

$$(x, x') \in R \quad \text{implies} \quad f(x) = f(x').$$

But this condition just says that $fr_1 = fr_2$, since $fr_1(x, x') = f(x)$ and $fr_2(x, x') = f(x')$ for all $(x, x') \in R$. Moreover, such a function $\bar{f}$, if it exists, is then necessarily unique, since π is an epimorphism.

The coequalizer in **Sets** of an arbitrary parallel pair of functions $f, g : A \rightrightarrows B$ can be constructed by quotienting B by the equivalence relation generated by the equations $f(x) = g(x)$ for all $x \in A$. We leave the details as an exercise.

Example 3.20. *Presentations of algebras*
Consider a category of "algebras"—say, monoids or groups—that has free algebras for all sets and coequalizers for all parallel pairs of arrows (see the exercises for a proof that monoids have coequalizers). We can use these to determine the notion of a *presentation* of an algebra by *generators* and *relations*.

For example, suppose we are given:

Generators: $x,\ y,\ z$
Relations: $xy = z,\ y^2 = 1$

To build an algebra on these generators and satisfying these relations, start with

$$F(3) = F(x, y, z)$$

and then "force" the relation $xy = z$ by taking a coequalizer of the maps

$$F(1) \underset{z}{\overset{xy}{\rightrightarrows}} F(3) \xrightarrow{\quad q \quad} Q$$

We use the fact that maps $F(1) \to A$ correspond to elements $a \in A$ by $v \mapsto a$, where v is the single generator of $F(1)$. Now similarly, for the equation $y^2 = 1$, take the coequalizer:

$$F(1) \underset{q(1)}{\overset{q(y^2)}{\rightrightarrows}} Q \longrightarrow Q'$$

These two steps can actually be done simultaneously; let

$$F(2) = F(1) + F(1)$$

$$F(2) \underset{g}{\overset{f}{\rightrightarrows}} F(3)$$

where $f = [xy, y^2]$ and $g = [z, 1]$. The coequalizer $q : F(3) \to Q$ of f and g then "forces" both equations to hold, in the sense that in Q we have

$$q(x)q(y) = q(z), \quad q(y)^2 = 1.$$

Moreover, no other relations among the generators hold in Q except those required to hold by the stipulated equations. For, given any algebra A and any three elements $a, b, c \in A$ such that $ab = c$ and $b^2 = 1$, by the UMP of Q there is a unique homomorphism $u : Q \to A$ such that

$$u(x) = a, \qquad u(y) = b, \qquad u(z) = c.$$

Thus any other equation that holds among the generators will also hold in any other algebra in which the stipulated equations hold, since the homomorphism u also preserves equations. In this sense, Q is the "universal" algebra with three generators satisfying the stipulated equations; as may be written suggestively in the form

$$Q \cong F(x, y, z)/(xy = z,\ y^2 = 1).$$

Generally, given a finite presentation:

Generators: $g_1, \ldots, g_n$
Relations: $l_1 = r_1, \ldots, l_m = r_m$

the algebra with that presentation is the coequalizer

$$F(m) \underset{r}{\overset{l}{\rightrightarrows}} F(n) \longrightarrow Q = F(n)/(l = r)$$

where $l = [l_1, \ldots, l_m]$ and $r = [r_1, \ldots, r_m]$. Such algebras are said to be *finitely presented*.

Warning 3.21. Presentations are not unique. One may well have two different presentations $F(n)/(l = r)$ and $F(n')/(l' = r')$ by generators and relations of the same algebra,

$$F(n)/(l = r) \;\cong\; F(n')/(l' = r').$$

For instance, given $F(n)/(l = r)$ add a new generator g_{n+1} and the new relation $g_n = g_{n+1}$. In general, there are many different ways of presenting a given algebra, just like there are many ways of axiomatizing a logical theory.

We did not really make use of the finiteness condition in the foregoing considerations. Indeed, *any* sets of generators G and relations R give rise to an algebra in the same way, by taking the coequalizer

$$F(R) \underset{r_2}{\overset{r_1}{\rightrightarrows}} F(G) \longrightarrow F(G)/(r_1 = r_2).$$

In fact, every "algebra" can be "presented" by generators and relations in this way, given a suitable notion of an "algebra." More precisely, we have the following proposition for monoids, an analogous version of which also holds for groups, and many related structures.

Proposition 3.22. *For every monoid M there are sets G and R and a coequalizer diagram,*

$$F(R) \underset{r_2}{\overset{r_1}{\rightrightarrows}} F(G) \longrightarrow M$$

with $F(G)$ and $F(R)$ free, thus $M \cong F(G)/(r_1 = r_2)$.

Proof. For any monoid N, let us write $TN = M(|N|)$ for the free monoid on the set of elements of N (and note that T is therefore a functor). There is a homomorphism,

$$\pi : TN \to N$$
$$\pi(x_1, \ldots, x_n) = x_1 \cdot \ldots \cdot x_n$$

induced by the identity $1_{|N|} : |N| \to |N|$ on the generators. (Here we are writing the elements of TN as tuples $(x_1, \ldots, x_n)$ rather than strings $x_1 \ldots x_n$ for clarity.)

Applying this construction twice to a monoid M results in the arrows π and ε in the following diagram,

$$T^2M \underset{\mu}{\overset{\varepsilon}{\rightrightarrows}} TM \xrightarrow[\pi]{} M \tag{3.2}$$

where $T^2M = TTM$ and $\mu = T\pi$. Explicitly, the elements of T^2M are tuples of tuples of elements of M, say $((x_1, \ldots, x_n), \ldots, (z_1, \ldots, z_m))$, and the homomorphisms ε and μ have the effect:

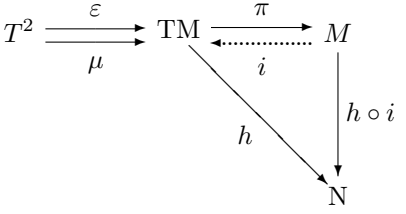
$$\varepsilon((x_1, \ldots, x_n), \ldots, (z_1, \ldots, z_m)) = (x_1, \ldots, x_n, \ldots, z_1, \ldots, z_m)$$
$$\mu((x_1, \ldots, x_n), \ldots, (z_1, \ldots, z_m)) = (x_1 \cdot \ldots \cdot x_n, \ldots, z_1 \cdot \ldots \cdot z_m)$$

Briefly, ε uses the multiplication in TM and μ uses that in M.

We claim that (3.2) is a coequalizer of monoids. To that end, suppose we have a monoid N and a homomorphism $h : TM \to N$ with $h\varepsilon = h\mu$. Then for any tuple $(x, \ldots, z)$ we have

$$\begin{aligned} h(x, \ldots, z) &= h\varepsilon((x, \ldots, z)) \\ &= h\mu((x, \ldots, z)) \\ &= h(x \cdot \cdots \cdot z). \end{aligned} \tag{3.3}$$

Now define $\bar{h} = h \circ i$, where $i : |M| \to |TM|$ is the insertion of generators as indicated in the following:



We then have:

$$\begin{aligned} \bar{h}\pi(x, \ldots, z) &= hi\pi(x, \ldots, z) \\ &= h(x \cdot \ldots \cdot z) \\ &= h(x, \ldots, z) \qquad \text{by (3.3)} \end{aligned}$$

We leave it as an easy exercise for the reader to show that $\bar{h}$ is a homomorphism. □

3.5 Exercises

1. (a) In any category $\mathbf{C}$, show that

$$A \xrightarrow[c_1]{} C \xleftarrow[c_2]{} B$$

is a coproduct diagram just if for every object Z, the map

$$\begin{aligned} \mathrm{Hom}(C, Z) &\longrightarrow \mathrm{Hom}(A, Z) \times \mathrm{Hom}(B, Z) \\ f &\longmapsto \quad \langle f \circ c_1,\ f \circ c_2 \rangle \end{aligned}$$

is an isomorphism. If you do this by using duality, you may take the corresponding fact about products as given.

   (b) If you proved the first part directly, prove the corresponding fact about products by using duality.

2. Show that the category **Ab** of abelian groups ($xy = yx$) has all equalizers.
3. In the proof of Proposition 3.22 in the text it is shown that any monoid M has a specific presentation $T^2M \rightrightarrows TM \to M$ as a coequalizer of free monoids. Show that coequalizers of this particular form are preserved by the forgetful functor **Mon** $\to$ **Sets**.
4. Prove that **Sets** has all coequalizers by constructing the coequalizer of a parallel pair of functions,

$$A \underset{g}{\overset{f}{\rightrightarrows}} B \longrightarrow Q = B/(f = g)$$

by quotienting B by a suitable equivalence relation R on B, generated by the pairs $(f(x), g(x))$ for all $x \in A$. (Define R to be the intersection of all equivalence relations on B containing all such pairs.)

5. Show that the category of monoids has all coequalizers.
6. Consider the category of sets.

   (a) Given a function $f : A \to B$, describe the equalizer of the functions $f \circ p_1, f \circ p_2 : A \times A \to B$ as a (binary) relation on A and show that it is an equivalence relation (called the *kernel* of f).

   (b) Show that the kernel of the quotient $A \to A/R$ by an equivalence relation R is R itself.

(c) Given *any* binary relation $R \subseteq A \times A$, let $\langle R \rangle$ be the equivalence relation on A generated by R (the least equivalence relation on A containing R). Show that the quotient $A \to A/\langle R \rangle$ is the coequalizer of the two projections $R \rightrightarrows A$.

(d) Using the foregoing, show that for any binary relation R on a set A, one can characterize the equivalence relation $\langle R \rangle$ generated by R as the kernel of the coequalizer of the two projections of R.

4

GROUPS AND CATEGORIES

This chapter is devoted to some of the various connections between groups and categories. If you already know the basic group theory covered here, then this will give you some insight into the categorical constructions we have learned so far; and if you do not know it yet, then you will learn it now as an application of category theory. We will focus on three different aspects of the relationship between categories and groups:

1. groups in a category,
2. the category of groups,
3. groups as categories.

4.1 Groups in a category

As we have already seen, the notion of a group arises as an abstraction of the automorphisms of an object. In a specific, concrete case, a group G may thus consist of certain arrows $g : X \to X$ for some object X in a category $\mathbf{C}$,

$$G \subseteq \mathrm{Hom}_{\mathbf{C}}(X, X)$$

But the abstract group concept can also described directly as an object in a category, equipped with a certain structure. This more subtle notion of a "group in a category" also proves to be quite useful.

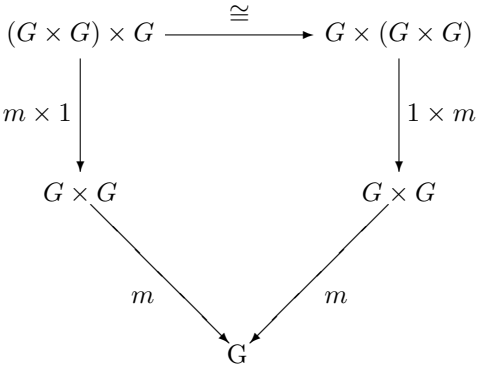
Let $\mathbf{C}$ be a category with finite products. The notion of a group in $\mathbf{C}$ essentially generalizes the usual notion of a group in **Sets**.

Definition 4.1. A *group* in $\mathbf{C}$ consists of objects and arrows as so:

$$\begin{array}{ccccc} G \times G & \xrightarrow{m} & G & \xleftarrow{i} & G \\ & & \uparrow{\scriptstyle u} & & \\ & & 1 & & \end{array}$$

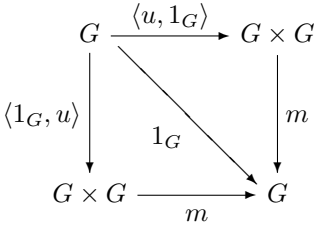
satisfying the following conditions:

1. m is associative, that is, the following commutes:



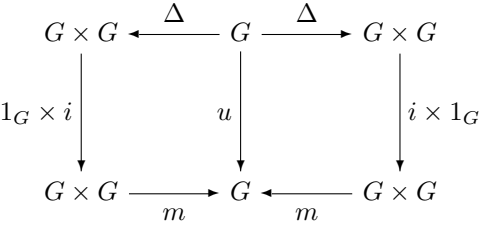
   where $\cong$ is the canonical associativity isomorphism for products.

2. u is a unit for m, that is, both triangles in the following commute:



   where we write u for the "constant arrow" $u! : G \xrightarrow{!} 1 \xrightarrow{u} G$.

3. i is an inverse with respect to m, that is, both sides of this commute:



   where $\Delta = \langle 1_G, 1_G \rangle$.

Note that the requirement that these diagrams commute is equivalent to the more familiar condition that, for all (generalized) elements,

$$x, y, z \colon\ Z \to\ G$$

the following equations hold:

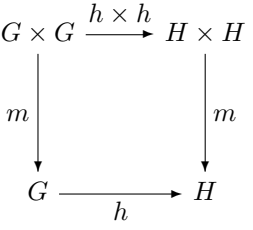
$$m(m(x, y), z) = m(x, m(y, z))$$
$$m(x, u) = x = m(u, x)$$
$$m(x, ix) = u = m(ix, x)$$

Definition 4.2. A *homomorphism* $h : G \to H$ of groups in $\mathbf{C}$ consists of an arrow in $\mathbf{C}$,

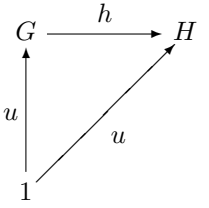
$$h : G \to H$$

such that

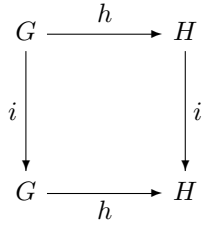
1. h preserves m:



2. h preserves u:



3. h preserves i:



With the evident identities and composites, we thus have a category of groups in $\mathbf{C}$, denoted:

$$\mathrm{Group}(\mathbf{C})$$

Example 4.3. The idea of a group in a category captures the familiar notion of a group with additional structure.

- A group in the usual sense is a group in the category **Sets**.
- A topological group is a group in **Top**, the category of topological spaces.
- A (partially) ordered group is a group in the category **Pos** of posets (in this case the inverse operation is usually required to be order-*reversing*, that is, of the form $i : G^{\mathrm{op}} \to G$).

For example, the real numbers $\mathbb{R}$ under addition are a topological and an ordered group, since the operations of addition $x + y$ and additive inverse $-x$ are continuous and order-preserving (resp. reversing). They are a topological "semigroup" under multiplication $x \cdot y$ as well, but the multiplicative inverse operation $1/x$ is not continuous (or even defined!) at 0.

In logical terms, according to this point of view one can "model the theory of groups" in *any* category with finite products, not just **Sets**. Of course the same is true for other theories—like monoids and rings—given by operations and equations. Thus, for instance, one can also define the notion of a *group in the lambda-calculus*, since the category of types of the lambda-calculus also has finite products. Theories involving other logical operations like quantifiers can be modeled in categories having more structure than just finite products. Here we have a glimpse of so-called *categorical semantics*. Such semantics can be useful for theories that are not complete with respect to models in **Sets**, such as certain theories in intuitionistic logic.

4.2 The category of groups

Let G and H be groups (in **Sets**), and let

$$h : G \to H$$

be a group homomorphism. The *kernel* of h is defined by the equalizer

$$\ker(h) = \{g \in G \mid h(g) = u\} \longrightarrow G \underset{u}{\overset{h}{\rightrightarrows}} H$$

where, again, we write $u : G \to H$ for the constant homomorphism

$$u! = G \xrightarrow{!} 1 \xrightarrow{u} H.$$

We have already seen that this specification makes the above an equalizer diagram.

Observe that $\ker(h)$ is a *subgroup*. Indeed, it is a *normal subgroup*, in the sense that for any $k \in \ker(h)$, we have (using multiplicative notation)

$$g \cdot k \cdot g^{-1} \in \ker(h) \quad \text{for all } g \in G.$$

Now if $N \overset{i}{\rightarrowtail} G$ is *any* normal subgroup, we can construct the *coequalizer*

$$N \underset{u}{\overset{i}{\rightrightarrows}} G \xrightarrow{\pi} G/N$$

sending $g \in G$ to u iff $g \in N$ ("killing off N"), as follows: the elements of G/N are the "cosets of N," that is, equivalence classes of the form $[g]$ for all $g \in G$, where we define

$$g \sim h \quad \text{iff} \quad g \cdot h^{-1} \in N.$$

(Prove that this is an equivalence relation!) The multiplication on the *factor group* G/N is then given by

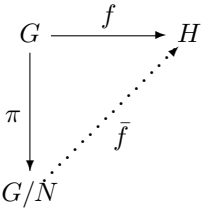
$$[g] \cdot [g'] = [g \cdot g']$$

which is well defined since N is normal (proof!).

Let us show that the diagram above really is a coequalizer. First, it is clear that

$$\pi \circ i = \pi \circ u!$$

since $n \cdot u = n$ implies $[n] = [u]$. Suppose we have $f : G \to H$ killing N, that is, $f(n) = u$ for all $n \in N$. We then propose a "factorization" $\bar{f}$, as indicated in



to be defined by

$$\bar{f}[g] = f(g).$$

This will be well defined if $x \sim y$ implies $f(x) = f(y)$. But, since $x \sim y$ implies $f(x \cdot y^{-1}) = u$, we have

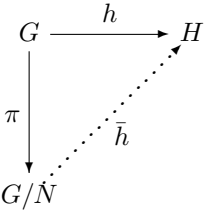
$$f(x) = f(x \cdot y^{-1} \cdot y) = f(x \cdot y^{-1}) \cdot f(y) = u \cdot f(y) = f(y).$$

Moreover, $\bar{f}$ is unique with $\pi \bar{f} = f$, since π is epic. Thus, we've shown most of the following classical *Homomorphism Theorem for Groups*.

Theorem 4.4. *Every group homomorphism $h : G \to H$ has a kernel* $\ker(h) = h^{-1}(u)$*, which is a normal subgroup of G with the property that, for any normal subgroup $N \subseteq G$*

$$N \subseteq \ker(h)$$

iff there is a (necessarily unique) homomorphism $\bar{h} : G/N \to H$ with $\bar{h} \circ \pi = h$, as indicated in:



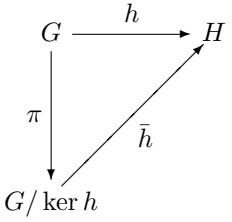
Proof. It only remains to show that if such a factorization $\bar{h}$ exists, then $N \subseteq \ker(h)$. But this is clear, since $\pi(N) = \{[u_G]\}$. So $h(n) = \bar{h}\pi(n) = \bar{h}([n]) = u_H$. □

Finally, putting $N = \ker(h)$ in the theorem and taking any $[x], [y] \in G/\ker(h)$, we have

$$\begin{aligned} \bar{h}[x] = \bar{h}[y] &\Rightarrow h(x) = h(y) \\ &\Rightarrow h(xy^{-1}) = u \\ &\Rightarrow xy^{-1} \in \ker(h) \\ &\Rightarrow x \sim y \\ &\Rightarrow [x] = [y]. \end{aligned}$$

Thus, $\bar{h}$ is injective and we have

Corollary 4.5. *Every group homomorphism $h : G \to H$ factors as a quotient followed by an injective homomorphism,*



Thus $\bar{h} : G/\ker(h) \xrightarrow{\sim} \mathrm{im}(h) \subseteq H$ is an isomorphism onto the subgroup $\mathrm{im}(h)$ that is the image of h.

In particular, therefore, a homomorphism h is injective if and only if its kernel is "trivial," in the sense that $\ker(h) = \{u\}$.

4.3 Groups as categories

First, let us recall that a group is a category. In particular, a group is a category with one object, in which every arrow is an iso.

If G and H are groups, regarded as categories, then we can consider arbitrary functors between them

$$f : G \to H.$$

It is easy to see that a functor between groups is exactly the same thing as a group homomorphism.

What is a functor $R : G \to \mathbf{C}$ from a group G to another category $\mathbf{C}$ that is not necessarily a group? If $\mathbf{C}$ is the category of vector spaces and linear transformations, then such a functor is just what the group theorist calls a "linear

representation" of G. In general, such a functor $R : G \to \mathbf{C}$ may be regarded as a representation of G in $\mathbf{C}$.

We will now generalize the notions of kernel of a homomorphism, and quotient or factor group by a normal subgroup, from groups to arbitrary categories, and then give the analogous homomorphism theorem for categories.

Definition 4.6. A *congruence* on a category $\mathbf{C}$ is an equivalence relation $f \sim g$ on arrows such that:

1. $f \sim g$ implies $\mathrm{dom}(f) = \mathrm{dom}(g)$ and $\mathrm{cod}(f) = \mathrm{cod}(g)$,

$$\bullet \underset{g}{\overset{f}{\rightrightarrows}} \bullet$$

2. $f \sim g$ implies $bfa \sim bga$ for all arrows $a : A \to X$ and $b : Y \to B$, where $\mathrm{dom}(f) = X = \mathrm{dom}(g)$ and $\mathrm{cod}(f) = Y = \mathrm{cod}(g)$,

$$\bullet \xrightarrow{a} \bullet \underset{g}{\overset{f}{\rightrightarrows}} \bullet \xrightarrow{b} \bullet$$

Let $\sim$ be a congruence on the category $\mathbf{C}$, and define the *congruence category* $\mathbf{C}^{\sim}$ by:

$$\begin{aligned} (\mathbf{C}^{\sim})_0 &= \mathbf{C}_0 \\ (\mathbf{C}^{\sim})_1 &= \{\langle f, g\rangle | f \sim g\} \\ \overset{\sim}{1}_C &= \langle 1_C, 1_C\rangle \\ \langle f', g'\rangle \circ \langle f, g\rangle &= \langle f'f, g'g\rangle \end{aligned}$$

One easily checks that this composition is well defined, using the congruence conditions.

There are two evident projection functors:

$$\mathbf{C}^{\sim} \underset{p_2}{\overset{p_1}{\rightrightarrows}} \mathbf{C}$$

We build the *quotient category* $\mathbf{C}/\!\sim$ as follows:

$$\begin{aligned} (\mathbf{C}/\!\sim)_0 &= \mathbf{C}_0 \\ (\mathbf{C}/\!\sim)_1 &= (\mathbf{C}_1)/\!\sim \end{aligned}$$

The arrows have the form $[f]$ where $f \in \mathbf{C}_1$, and we can put $1_{[C]} = [1_C]$, and $[g] \circ [f] = [g \circ f]$, as is easily checked, again using the congruence conditions.

There is an evident quotient functor $\pi : \mathbf{C} \to \mathbf{C}/\!\sim$. It then makes the following a coequalizer of categories:

$$\mathbf{C}^{\sim} \underset{p_2}{\overset{p_1}{\rightrightarrows}} \mathbf{C} \xrightarrow{\pi} \mathbf{C}/\!\sim$$

This is proved much as for groups.

An exercise shows how to use this construction to make coequalizers for certain functors. Let us show how to use it to prove an analogous "homomorphism theorem for categories." Suppose we have categories $\mathbf{C}$ and $\mathbf{D}$ and a functor

$$F : \mathbf{C} \to \mathbf{D}.$$

Then F determines a congruence $\sim_F$ on $\mathbf{C}$ by setting:

$$f \sim_F g \quad \text{iff} \quad \mathrm{dom}(f) = \mathrm{dom}(g),\ \mathrm{cod}(f) = \mathrm{cod}(g),\ F(f) = F(g)$$

That this *is* a congruence is easily checked.

Let us write

$$\ker(F) = \mathbf{C}^{\sim_F} \rightrightarrows \mathbf{C}$$

for the congruence category, and call this the *kernel category* of F.

The quotient category

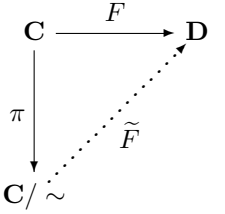
$$\mathbf{C}/\sim_F$$

then has the following UMP:

Theorem 4.7. *Every functor $F : \mathbf{C} \to \mathbf{D}$ has a kernel category $\ker(F)$, determined by a congruence $\sim_F$ on $\mathbf{C}$ such that given any congruence $\sim$ on $\mathbf{C}$ one has:*

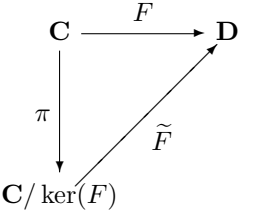
$$f \sim g \Rightarrow f \sim_F g$$

if and only if there is a factorization $\widetilde{F} : \mathbf{C}/\sim \longrightarrow \mathbf{D}$, as indicated in:



Just as in the case of groups, applying the theorem to the case $\mathbf{C}^{\sim} = \ker(F)$ gives a factorization theorem:

Corollary 4.8. *Every functor $F : \mathbf{C} \to \mathbf{D}$ factors as $F = \widetilde{F} \circ \pi$,*



*where π is bijective on objects and surjective on Hom-sets, and $\widetilde{F}$ is injective on Hom-sets (i.e. "*faithful*"):*

$$\widetilde{F}_{A,B} : \mathrm{Hom}(A, B) \rightarrowtail \mathrm{Hom}(FA, FB) \quad \textit{for all } A, B \in \mathbf{C}/\ker(F)$$

4.4 Finitely presented categories

Finally, let us consider categories presented by generators and relations.

We begin with the free category $\mathbf{C}(G)$ on some finite graph G, and then consider a finite set Σ of relations of the form

$$(g_1 \circ \ldots \circ g_n) = (g'_1 \circ \ldots \circ g'_m)$$

with all $g_i \in G$, and $\mathrm{dom}(g_n) = \mathrm{dom}(g'_m)$ and $\mathrm{cod}(g_1) = \mathrm{cod}(g'_1)$. Such a relation identifies two "paths" in $\mathbf{C}(G)$ with the same "endpoints" and "direction." Next, let $\sim_\Sigma$ be the smallest congruence $\sim$ on $\mathbf{C}$ such that $f \sim f'$ for each equation $g = g'$ in Σ. Such a congruence exists simply because the intersection of a family of congruences is again a congruence. Taking the quotient by this congruence we have a notion of a finitely presented category:

$$\mathbf{C}(G, \Sigma) = \mathbf{C}(G)/\sim_\Sigma$$

This is completely analogous to the notion of a finite presentation for groups, and indeed specializes to that notion in the case of a graph with only one vertex. The UMP of $\mathbf{C}(G, \Sigma)$ is then an obvious variant of that already given for groups.

Specifically, in $\mathbf{C}(G, \Sigma)$ there is a "diagram of type G," that is, a graph homomorphism $i : G \to |\mathbf{C}(G, \Sigma)|$, satisfying all the conditions $i(g) = i(g')$, for all $g = g' \in \Sigma$. Moreover, given any category $\mathbf{D}$ with a diagram of type G, say $h : G \to |\mathbf{D}|$, that satisfies all the conditions $h(g) = h(g')$, for all $g = g' \in \Sigma$, there is a unique functor $\bar{h} : \mathbf{C}(G, \Sigma) \to \mathbf{D}$ with $|\bar{h}| \circ i = h$. The reader should draw the associated diagram of graphs and categories.

Just as in the case of presentations of groups, one can describe the construction of $\mathbf{C}(G, \Sigma)$ as a coequalizer for two functors.

Indeed, suppose we have arrows $f, f' \in \mathbf{C}$. Take the least congruence $\sim$ on $\mathbf{C}$ with $f \sim f'$. Consider the diagram

$$\mathbf{C}(\mathbf{2}) \underset{f'}{\overset{f}{\rightrightarrows}} \mathbf{C} \xrightarrow{\ q\ } \mathbf{C}/\sim$$

where $\mathbf{2}$ is the graph with two vertices and an edge between them, f and f' are the unique functors taking the generating edge to the arrows by the same names, and q is the canonical functor to the quotient category. Then q is a coequalizer of f and f'. To show this, take any $d : \mathbf{C} \to \mathbf{D}$ with

$$df = df'.$$

Since $\mathbf{C}(2)$ is free on $\cdot \xrightarrow{x} \cdot$, and $f(x) = f$ and $f'(x) = f'$, we have

$$d(f) = d(f(x)) = d(f'(x)) = d(f').$$

Thus, $\langle f, f' \rangle \in \ker(d)$, so $\sim \subseteq \ker(d)$ (since $\sim$ is minimal with $f \sim f'$). So there is a functor $\bar{d} : \mathbf{C}/\sim \to \mathbf{D}$ such that $d = \bar{d} \circ q$ by the homomorphism theorem.

Example 4.9. The category with two uniquely isomorphic objects is not free on *any* graph, since it's finite, but has "loops" (cycles). But it *is* finitely presented with graph

$$A \underset{g}{\overset{f}{\rightleftarrows}} B$$

and relations

$$gf = 1_A, \qquad fg = 1_B.$$

Similarly, there are finitely presented categories with just one non-identity arrow $f : \cdot \to \cdot$ and either

$$f \circ f = 1 \quad \text{or} \quad f \circ f = f.$$

In the first case we have the group $\mathbb{Z}/2\mathbb{Z}$. In the second case an "idempotent" (but not a group).

Indeed, any of the cyclic groups

$$\mathbb{Z}_n \cong \mathbb{Z}/\mathbb{Z}n$$

occur in this way, with the graph

$$\star \xrightarrow{\quad f \quad} \star$$

and the relation

$$f^n = 1.$$

4.5 Exercises

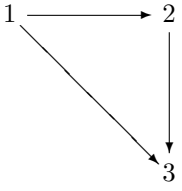
1. Regarding a group G as a category with one object and every arrow an isomorphism, show that a categorical congruence $\sim$ on G is the same thing as (the equivalence relation on G determined by) a normal subgroup $N \subseteq G$, that is, show that the two kinds of things are in isomorphic correspondence.

   Show further that the quotient category $G/\sim$ and the factor group G/N coincide. Conclude that the homomorphism theorem for groups is a special case of the one for categories.

2. Consider the definition of a group in a category as applied to the category $\mathbf{Sets}/I$ of sets sliced over a set I. Show that such a group G determines an

I-indexed family of (ordinary) groups G_i by setting $G_i = G^{-1}(i)$ for each $i \in I$. Show that this determines a functor $\mathbf{Groups}(\mathbf{Sets}/I) \to \mathbf{Groups}^I$ into the category of I-indexed families of groups and I-indexed families of homomorphisms.

3. Give four different presentations by generators and relations of the category **3**, pictured:



Is **3** free?

4. Given a congruence $\sim$ on a category $\mathbf{C}$ and arrows in $\mathbf{C}$ as follows,

$$A \underset{f'}{\overset{f}{\rightrightarrows}} B \underset{g'}{\overset{g}{\rightrightarrows}} C$$

show that $f \sim f'$ and $g \sim g'$ implies $g \circ f \sim g' \circ f'$.

5. Given functors $F, G : \mathbf{C} \to \mathbf{D}$ such that for all $C \in \mathbf{C}$, $FC = GC$, define a congruence on $\mathbf{D}$ by the condition:

$$\begin{aligned} f \sim g \quad \text{iff} \quad & \mathrm{dom}(f) = \mathrm{dom}(g) \\ & \& \ \mathrm{cod}(f) = \mathrm{cod}(g) \\ & \& \ \forall \mathbf{E}, H : \mathbf{D} \to \mathbf{E}.\ HF = HG \Rightarrow H(f) = H(g) \end{aligned}$$

Prove that this is a congruence.

Prove that $\mathbf{D}/\sim$ is the coequalizer of F and G.

This page intentionally left blank

5

LIMITS AND COLIMITS

In this chapter we briefly discuss some topics relating to the definitions that we already have, rather than pushing on to new ones. This is partly in order to see how these are used, but also because we will need this material soon enough. After that, and after a brief look at one more elementary notion, we shall go on to what may be called "higher category theory."

5.1 Subobjects

We have seen that every subset $U \subseteq X$ of a set X occurs as an equalizer and that equalizers are always monomorphisms. So it is natural to regard monos as generalized subsets. That is, a mono in **Groups** can be regarded as a subgroup, a mono in **Top** as a subspace, and so on.

The rough idea is this: given a monomorphism,

$$m : M \rightarrowtail X$$

in a category **G** of structured sets of some sort—call them "gadgets"—the image sub*set*

$$\{m(y) \mid y \in M\} \subseteq X$$

which may be written $m(M)$, is often a sub-gadget of X to which M is isomorphic via m.

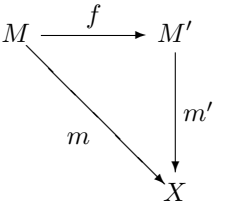
$$m : M \xrightarrow{\sim} m(M) \subseteq X$$

More generally, we can think of the mono $m : M \rightarrowtail X$ itself as determining a "part" of X even in categories that do not have underlying functions.

Definition 5.1. A *subobject* of an object X in a category **C** is a mono

$$m : M \rightarrowtail X$$

Given subobjects m and m' of X, a morphism $f : m \to m'$ is an arrow in $\mathbf{C}/X$, as in



Thus we have a category,

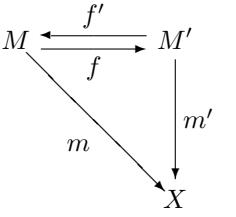
$$\mathrm{Sub}_{\mathbf{C}}(X)$$

of subobjects of X in $\mathbf{C}$.

In this definition, since m' is monic, there is at most one f as in the diagram above, so that $\mathrm{Sub}_{\mathbf{C}}(X)$ is a preorder category. We define the relation of *inclusion* of subobjects by:

$$m \subseteq m' \quad \text{iff} \quad \text{there exists some } f : m \to m'$$

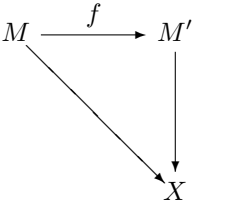
Finally, we say that m and m' are *equivalent*, written $m \equiv m'$, if and only if they are isomorphic as subobjects, that is, $m \subseteq m'$ and $m' \subseteq m$. This holds just if there are f and f' making both triangles below commute.



Observe that, in the above diagram, $m = m'f = mf'f$, and since m is monic, $f'f = 1_M$ and similarly $ff' = 1_{M'}$. So $M \cong M'$ via f. Thus we see that *equivalent subobjects have isomorphic domains.*

Remark 5.2. We sometimes abuse notation and language by calling M the subobject when the mono $m : M \rightarrowtail X$ is clear.

Note that if $M \subseteq M'$ then the arrow f which makes this so in



is also monic, so also M is a subobject of M'. In fact, we have a functor

$$\mathrm{Sub}(M') \to \mathrm{Sub}(X)$$

defined by composition (since the composite of monos is monic).

In terms of generalized elements of an object X,

$$z : Z \to X$$

one can define a *local membership relation*,

$$z \in_X M$$

between these and subobjects $m : M \rightarrowtail X$ by

$$z \in_X M \text{ iff there exists } f : Z \to M \text{ such that } z = mf$$

Since m is monic, if z factors through it then it does so uniquely.

Example 5.3. An equalizer

$$E \longrightarrow A \underset{g}{\overset{f}{\rightrightarrows}} B$$

is a subobject of A with the property

$$z \in_A E \quad \text{iff } f(z) = g(z)$$

Thus, we can regard E as the subobject of generalized elements $z : Z \to A$ such that $f(z) = g(z)$.

$$E = \{z \in Z \mid f(z) = g(z)\} \subseteq A$$

just as was the case for global elements in **Sets**. In categorical logic, one develops a way of making this intuition even more precise by giving a calculus of such subobjects.

Remark 5.4. It is often convenient to pass from the preorder

$$\mathrm{Sub}_{\mathbf{C}}(X)$$

to the *poset* given by factoring out the equivalence relation "$\equiv$". Then a subobject is an equivalence class of monos under mutual inclusion.

In **Sets**, under this notion of subobject, one then has an isomorphism,

$$\mathrm{Sub}_{\mathbf{Sets}}(X) \cong P(X)$$

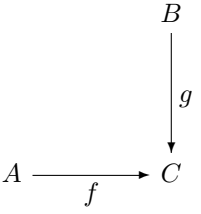
that is, every subobject is represented by a unique subset. We shall use both notions of subobject, making clear when monos are intended, and when equivalence classes thereof are intended.

5.2 Pullbacks

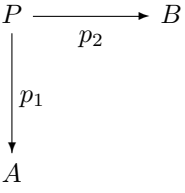
The notion of a pullback, like that of a product, is one that comes up very often in mathematics and logic. It is a generalization of both intersection and inverse image.

We begin with the definition,

Definition 5.5. In any category $\mathbf{C}$, a *pullback* of arrows f, g with $\mathrm{cod}(f) = \mathrm{cod}(g)$



consists of arrows

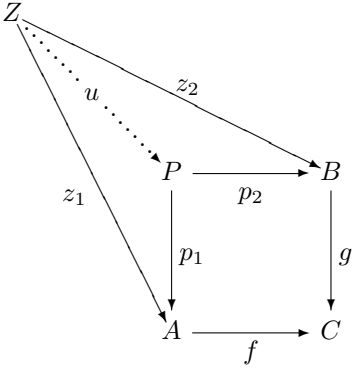


such that $fp_1 = gp_2$ and universal with this property. That is, given any $z_1 : Z \to A$ and $z_2 : Z \to B$ with $fz_1 = gz_2$, there exists a unique

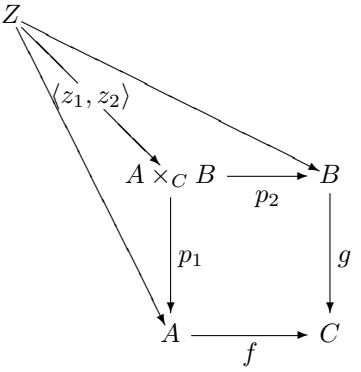
$$u : Z \to P$$

with $z_1 = p_1 u$ and $z_2 = p_2 u$.

Here is the picture:

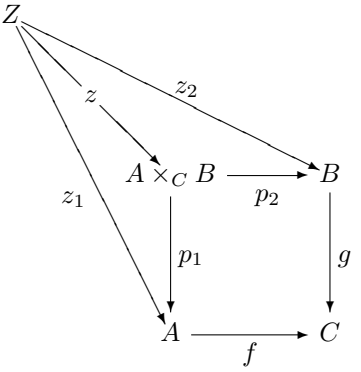


Remark 5.6. One sometimes uses product-style notation for pullbacks.



Pullbacks are clearly unique up to isomorphism since they are given by an UMP (universal mapping property). Here this means that given two pullbacks of a given pair of arrows, the uniquely determined maps between the pullbacks are mutually inverse.

In terms of generalized elements, any $z \in A \times_C B$, can be written uniquely as $z = \langle z_1, z_2 \rangle$ with $fz_1 = gz_2$.

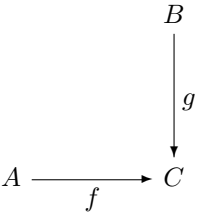


This makes

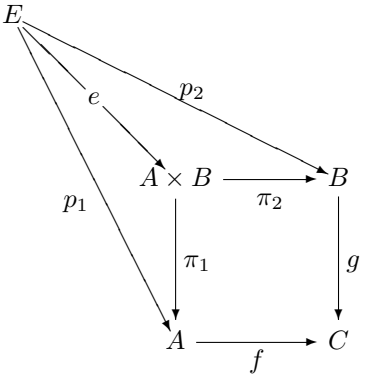
$$A \times_C B = \{\langle z_1, z_2 \rangle \in A \times B \mid fz_1 = gz_2\}$$

look like a subobject of $A \times B$, determined as an equalizer of $f \circ \pi_1$ and $g \circ \pi_2$. In fact, this is so.

Proposition 5.7. *In a category with products and equalizers, given a corner of arrows*



Consider the diagram

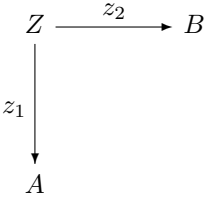


in which e is an equalizer of $f\pi_1$ and $g\pi_2$ and $p_1 = \pi_1 e$, $p_2 = \pi_2 e$. Then E, p_1, p_2 is a pullback of f and g. Conversely, if E, p_1, p_2 are given as such a pullback, then the arrow

$$e = \langle p_1, p_2 \rangle : E \to A \times B$$

is an equalizer of $f\pi_1$ and $g\pi_2$.

Proof. Take



with $fz_1 = gz_2$. We have $\langle z_1, z_2 \rangle : Z \to A \times B$ so

$$f\pi_1\langle z_1, z_2 \rangle = g\pi_2\langle z_1, z_2 \rangle.$$

Thus, there is a $u : Z \to E$ to the equalizer with $eu = \langle z_1, z_2 \rangle$. Then

$$p_1 u = \pi_1 eu = \pi_1\langle z_1, z_2 \rangle = z_1$$

and

$$p_2 u = \pi_2 eu = \pi_2\langle z_1, z_2 \rangle = z_2.$$

If also $u' : Z \to E$ has $p_i u' = z_i, i = 1, 2$, then $\pi_i eu' = z_i$ so $eu' = \langle z_1, z_2 \rangle = eu$ whence $u' = u$ since e in monic. The converse is similar. □

Corollary 5.8. *If a category* **C** *has binary products and equalizers, then it has pullbacks.*

The foregoing gives an explicit construction of a pullback in **Sets** as a subset of the product:

$$\{\langle a, b\rangle \mid fa = gb\} = A \times_C B \hookrightarrow A \times B$$

Example 5.9. In **Sets**, take a function $f : A \to B$ and a subset $V \subseteq B$. Let, as usual,

$$f^{-1}(V) = \{a \in A \mid f(a) \in V\} \subseteq A$$

and consider

$$\begin{array}{ccc} f^{-1}(V) & \xrightarrow{\bar{f}} & V \\ {\scriptstyle j}\downarrow & & \downarrow{\scriptstyle i} \\ A & \xrightarrow[f]{} & B \end{array}$$

where i and j are the canonical inclusions and $\bar{f}$ is the evident factorization of the restriction of f to $f^{-1}(V)$ (since $a \in f^{-1}(V) \Rightarrow f(a) \in V$).

This diagram is a pullback (observe that $z \in f^{-1}(V) \Leftrightarrow fz \in V$ for all $z : Z \to A$). Thus, the inverse image

$$f^{-1}(V) \subseteq A$$

is *determined uniquely up to isomorphism* as a pullback.

As suggested by the previous example, we can use pullbacks to *define* inverse images in categories other than **Sets**. Indeed, given a pullback in any category:

$$\begin{array}{ccc} A \times_B M & \longrightarrow & M \\ {\scriptstyle m'}\downarrow & & \downarrow{\scriptstyle m} \\ A & \xrightarrow[f]{} & B \end{array}$$

if m is monic, then m' is monic. (Exercise!)

Thus we see that, for fixed $f : A \to B$, taking pullbacks induces a map

$$f^{-1} : \mathrm{Sub}(B) \to \mathrm{Sub}(A)$$

$$m \mapsto m'$$

We will show that f^{-1} also respects equivalence of subobjects

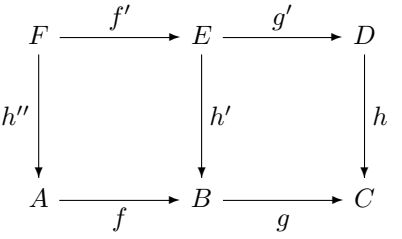
$$M \equiv N \Rightarrow f^{-1}(M) \equiv f^{-1}(N)$$

by showing that f^{-1} is a functor; that is our next goal.

5.3 Properties of pullbacks

We start with the following simple lemma, which seems to come up all the time.

Lemma 5.10. *(Two-pullbacks) Consider the commutative diagram below in a category with pullbacks:*



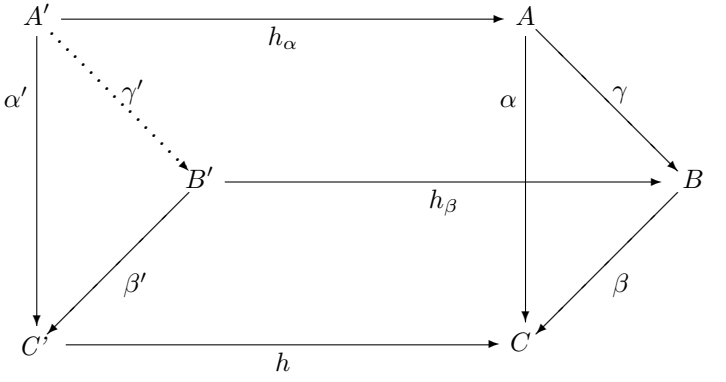
1. *If the two squares are pullbacks, so is the outer rectangle. Thus,*

$$A \times_B (B \times_C D) \cong A \times_C D$$

2. *If the right square and the outer rectangle are pullbacks, so is the left square.*

Proof. Diagram chase. □

Corollary 5.11. *The pullback of a commutative triangle is a commutative triangle. Specifically, given a commutative triangle as on the right end of the following "prism diagram"*



for any $h : C' \to C$, if one can form the pullbacks α' and β' as on the left end, then there exists a unique γ' as indicated, making the left end a commutative triangle, and the upper face a commutative rectangle, and indeed a pullback.

Proof. Apply the two-pullbacks lemma. □

Proposition 5.12. *Pullback is a functor. That is, for fixed $h : C' \to C$ in a category $\mathbf{C}$ with pullbacks, there is a functor*

$$h^* : \mathbf{C}/C \to \mathbf{C}/C'$$

defined by

$$(A \xrightarrow{\alpha} C) \mapsto (C' \times_C A \xrightarrow{\alpha'} C')$$

where α' is the pullback of α along h, and the effect on an arrow $\gamma : \alpha \to \beta$ is given by the foregoing corollary.

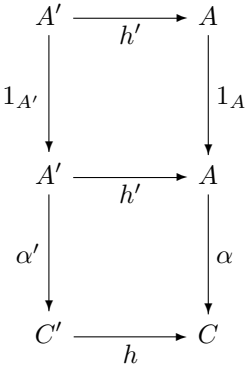
Proof. One must check that

$$h^*(1_X) = 1_{h^*X}$$

and

$$h^*(g \circ f) = h^*(g) \circ h^*(f)$$

These can easily be verified by repeated applications of the two-pullbacks lemma. For example, for the first condition, consider

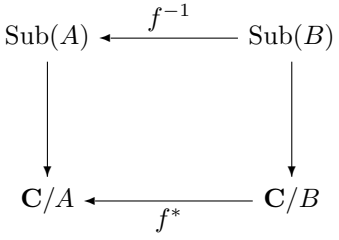


If the lower square is a pullback, then plainly so is the outer rectangle, whence the upper square is, too, and we have

$$h^*1_X = 1_{X'} = 1_{h^*X}.$$

□

Corollary 5.13. *Let $\mathbf{C}$ be a category with pullbacks. For any arrow $f : A \to B$ in $\mathbf{C}$ we have the following diagram of categories and functors:*



This commutes simply because f^{-1} *is defined to be the restriction of* f^* *to the subcategory* $\mathrm{Sub}(B)$*. Thus, in particular,* f^{-1} *is functorial:*

$$M \subseteq N \;\Rightarrow\; f^{-1}(M) \subseteq f^{-1}(N)$$

It follows that $M \equiv N$ *implies* $f^{-1}(M) \equiv f^{-1}(N)$*, so that* f^{-1} *is also defined on equivalence classes.*

$$f^{-1}/\equiv \,:\, \mathrm{Sub}(B)/\equiv \longrightarrow \mathrm{Sub}(A)/\equiv$$

Example 5.14. Consider a pullback in **Sets**:

$$\begin{array}{ccc} E & \xrightarrow{f'} & B \\ {\scriptstyle g'}\downarrow & & \downarrow{\scriptstyle g} \\ A & \xrightarrow[f]{} & C \end{array}$$

We saw that

$$E = \{\langle a, b\rangle \mid f(a) = g(b)\}$$

can be constructed as an equalizer

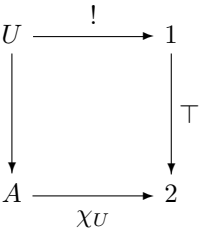
$$E \xrightarrow{\langle f', g'\rangle} A \times B \underset{g\pi_2}{\overset{f\pi_1}{\rightrightarrows}} C$$

Now let $B = 1$, $C = 2 = \{\top, \bot\}$, and $g = \top : 1 \to 2$. Then the equalizer

$$E \longrightarrow A \times 1 \underset{\top\pi_2}{\overset{f\pi_1}{\rightrightarrows}} 2$$

is how we already described the "extension" of the "propositional function" $f : A \to 2$. Therefore we can rephrase the correspondence between subsets

$U \subseteq A$ and their characteristic functions $\chi_U : A \to 2$ in terms of pullbacks:



Precisely, the isomorphism,

$$2^A \cong P(A)$$

given by taking a function $\varphi : A \to 2$ to its "extension"

$$V_\varphi = \{x \in A \mid \varphi(x) = \top\}$$

can be described as a pullback.

$$V_\varphi = \{x \in A \mid \varphi(x) = \top\} = \varphi^{-1}(\top)$$

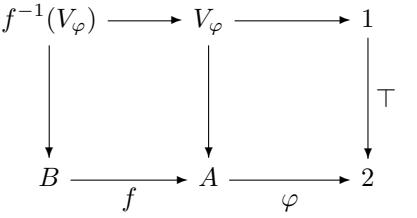
Now suppose we have any function

$$f : B \to A$$

and consider the induced inverse image operation

$$f^{-1} : P(A) \to P(B)$$

given by pullback, as in example 5.9 above. Taking some extension $V_\varphi \subseteq A$, consider the two-pullback diagram



We therefore have (by the two-pullbacks lemma)

$$f^{-1}(V_\varphi) = f^{-1}(\varphi^{-1}(\top)) = (\varphi f)^{-1}(\top) = V_{\varphi f}$$

which from a logical point of view expresses the fact that the substitution of a term f for the variable x in the propositional function φ is modeled by taking the pullback along f of the corresponding extension

$$f^{-1}(\{x \in A \mid \varphi(x) = \top\}) = \{y \in B \mid \varphi(f(y)) = \top\}.$$

Note that we have shown that for any function $f : B \to A$ the following square commutes

$$\begin{array}{ccc} 2^A & \xrightarrow{\cong} & P(A) \\ {\scriptstyle 2^f}\downarrow & & \downarrow{\scriptstyle f^{-1}} \\ 2^B & \xrightarrow[\cong]{} & P(B) \end{array}$$

where $2^f : 2^A \to 2^B$ is precomposition $2^f(g) = g \circ f$. In a situation like this, one says that the isomorphism

$$2^A \cong P(A)$$

is *natural* in A, which is obviously a much stronger condition than just having isomorphisms at each object A. We will consider such "naturality" systematically later. It was in fact one of the phenomena that originally gave rise to category theory.

Example 5.15. Let I be an index set, and consider an I-indexed family of sets:

$$(A_i)_{i \in I}$$

Given any function $\alpha : J \to I$, there is a J-indexed family

$$(A_{\alpha(j)})_{j \in J} \ ,$$

obtained by "reindexing along α." This reindexing can also be described as a pullback. Specifically, for each set A_i take the constant, i-valued function $p_i : A_i \to I$ and consider the induced map on the coproduct

$$p = [p_i] : \coprod_{i \in I} A_i \to I$$

The reindexed family $(A_{\alpha(j)})_{j \in J}$ can be obtained by taking a pullback along α, as indicated in the following diagram:

$$\begin{array}{ccc} \coprod_{j \in J} A_{\alpha(j)} & \longrightarrow & \coprod_{i \in I} A_i \\ {\scriptstyle q}\downarrow & & \downarrow{\scriptstyle p} \\ J & \xrightarrow[\alpha]{} & I \end{array}$$

where q is the indexing projection for $(A_{\alpha(j)})_{j \in J}$ analogous to p. In other words, we have

$$J \times_I (\coprod_{i \in I} A_i) \ \cong \ \coprod_{j \in J} A_{\alpha(j)}$$

The reader should work out the details as an instructive exercise.

5.4 Limits

We have already seen that the notions of product, equalizer, and pullback are not independent; the precise relation between them is this.

Proposition 5.16. *A category has finite products and equalizers iff it has pullbacks and a terminal object.*

Proof. The "only if" direction has already been done. For the other direction, suppose $\mathbf{C}$ has pullbacks and a terminal object 1.

- For any objects A, B we clearly have $A \times B \cong A \times_1 B$, as indicated in the following:

$$\begin{array}{ccc} A \times B & \longrightarrow & B \\ \downarrow & & \downarrow \\ A & \longrightarrow & 1 \end{array}$$

- For any arrows $f, g : A \to B$, the equalizer $e : E \to A$ is constructed as the following pullback:

$$\begin{array}{ccc} E & \xrightarrow{h} & B \\ {\scriptstyle e}\downarrow & & \downarrow{\scriptstyle \Delta = \langle 1_B, 1_B \rangle} \\ A & \xrightarrow[\langle f, g \rangle]{} & B \times B \end{array}$$

In terms of generalized elements,

$$E = \{(a, b) \mid \langle f, g \rangle(a) = \Delta b\}$$

where $\langle f, g \rangle(a) = \langle fa, ga \rangle$ and $\Delta(b) = \langle b, b \rangle$. So,

$$\begin{aligned} E &= \{\langle a, b \rangle \mid f(a) = b = g(a)\} \\ &\cong \{a \mid f(a) = g(a)\} \end{aligned}$$

which is just what we want. An easy diagram chase shows that

$$E \xrightarrow{e} A \underset{g}{\overset{f}{\rightrightarrows}} B$$

is indeed an equalizer. □

Product, terminal object, pullback, and equalizer, are all special cases of the general notion of a *limit*, which we will consider now. First, we need some preliminary definitions.

Definition 5.17. Let **J** and **C** be categories. A *diagram* of *type* **J** in **C** is a functor.

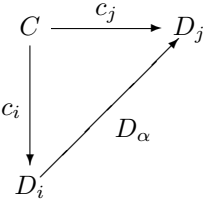
$$D : \mathbf{J} \to \mathbf{C}.$$

We will write the objects in the "index category" **J** lower case, $i, j, \ldots$ and the values of the functor $D : \mathbf{J} \to \mathbf{C}$ in the form D_i, D_j, etc.

A *cone* to a diagram D consists of an object C in **C** and a family of arrows in **C**,

$$c_j : C \to D_j$$

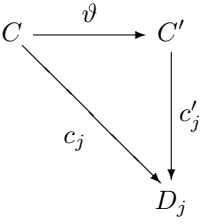
one for each object $j \in J$, such that for each arrow $\alpha : i \to j$ in **J**, the following triangle commutes.



A *morphism* of cones

$$\vartheta : (C, c_j) \to (C', c'_j)$$

is an arrow ϑ in **C** making each triangle,



commute. That is, such that $c_j = c'_j \circ \vartheta$ for all $j \in \mathbf{J}$. Thus, we have an apparent category

$$\mathbf{Cone}(D)$$

of cones to D.

We are here thinking of the diagram D as a "picture of **J** in **C**." A cone to such a diagram D is then imagined as a many-sided pyramid over the "base" D and a morphism of cones is an arrow between the apexes of such pyramids. (The reader should draw some pictures at this point!)

Definition 5.18. A *limit* for a diagram $D : \mathbf{J} \to \mathbf{C}$ is a terminal object in $\mathbf{Cone}(D)$. A *finite limit* is a limit for a diagram on a finite index category **J**.

We often denote a limit in the form

$$p_i : \varprojlim_j D_j \to D_i.$$

Spelling out the definition, the limit of a diagram D has the following UMP: given any cone (C, c_j) to D, there is a unique arrow $u : C \to \varprojlim_j D_j$ such that for all j,

$$p_j \circ u = c_j.$$

Example 5.19. Take $\mathbf{J} = \{1, 2\}$ the discrete category with two objects and no nonidentity arrows. A diagram $D : \mathbf{J} \to \mathbf{C}$ is a pair of objects $D_1, D_2 \in \mathbf{C}$. A cone on D is an object of $\mathbf{C}$ equipped with arrows

$$D_1 \xleftarrow{c_1} C \xrightarrow{c_2} D_2.$$

And a limit of D is a terminal such cone, that is, a *product* in $\mathbf{C}$ of D_1 and D_2,

$$D_1 \xleftarrow{p_1} D_1 \times D_2 \xrightarrow{p_2} D_2.$$

Thus, in this case,

$$\varprojlim_j D_j \;\cong\; D_1 \times D_2.$$

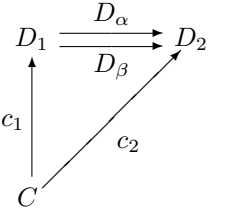
Example 5.20. Take $\mathbf{J}$ to be the following category:

$$\cdot \underset{\beta}{\overset{\alpha}{\rightrightarrows}} \cdot$$

A diagram of type $\mathbf{J}$ looks like

$$D_1 \underset{D_\beta}{\overset{D_\alpha}{\rightrightarrows}} D_2$$

and a cone is a pair of arrows

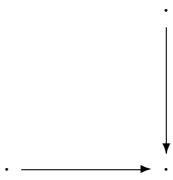


such that $D_\alpha c_1 = c_2$ and $D_\beta c_1 = c_2$; thus, $D_\alpha c_1 = D_\beta c_1$. A limit for D is therefore an *equalizer* for D_α, D_β.

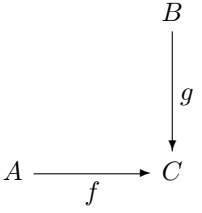
Example 5.21. If $\mathbf{J}$ is empty, there is just one diagram $D : \mathbf{J} \to \mathbf{C}$, and a limit for it is thus a *terminal object* in $\mathbf{C}$,

$$\varprojlim_{j \in \mathbf{0}} D_j \;\cong\; 1.$$

Example 5.22. If $\mathbf{J}$ is the finite category



we see that a limit for a diagram of the form



is just a pullback of f and g,

$$\varprojlim_{j} D_j \;\cong\; A \times_C B.$$

Thus, we have shown half of the following:

Proposition 5.23. *A category has all finite limits iff it has finite products and equalizers (resp. pullbacks and a terminal object by the last proposition).*

Here a category $\mathbf{C}$ is said to *have all finite limits* if every finite diagram $D : \mathbf{J} \to \mathbf{C}$ has a limit in $\mathbf{C}$.

Proof. We need to show that any finite limit can be constructed from finite products and equalizers. Take a finite diagram

$$D : \mathbf{J} \to \mathbf{C}.$$

Consider the finite products

$$\prod_{i \in J_0} D_i \quad \text{and} \quad \prod_{(\alpha : i \to j) \in J_1} D_j.$$

Define two arrows

$$\prod_i D_i \underset{\psi}{\overset{\phi}{\rightrightarrows}} \prod_{\alpha : i \to j} D_j$$

by taking their composites with the projections π_α from the second product to be, respectively:

$$\pi_\alpha \circ \phi = \phi_\alpha = \pi_{\mathrm{cod}(\alpha)}$$
$$\pi_\alpha \circ \psi = \psi_\alpha = D_\alpha \circ \pi_{\mathrm{dom}(\alpha)}$$

where $\pi_{\mathrm{cod}(\alpha)}$ and $\pi_{\mathrm{dom}(\alpha)}$ are projections from the first product.

Now we take the equalizer:

$$E \xrightarrow{\ e\ } \prod_i D_i \underset{\psi}{\overset{\phi}{\rightrightarrows}} \prod_{\alpha : i \to j} D_j$$

We will show that (E, e_i) is a limit for D, where $e_i = \pi_i \circ e$. To that end, take any arrow $c : C \to \prod_i D_i$, and write $c = \langle c_i \rangle$ for $c_i = \pi_i \circ c$. Observe that the family of arrows $(c_i : C \to D_i)$ is a cone to D if and only iff $\phi c = \psi c$. Indeed,

$$\phi\langle c_i \rangle = \psi\langle c_i \rangle$$

iff for all α,

$$\pi_\alpha \phi\langle c_i \rangle = \pi_\alpha \psi\langle c_i \rangle.$$

But,

$$\pi_\alpha \phi\langle c_i \rangle = \phi_\alpha \langle c_i \rangle = \pi_{\mathrm{cod}(\alpha)} \langle c_i \rangle = c_j$$

and

$$\pi_\alpha \psi\langle c_i \rangle = \psi_\alpha \langle c_i \rangle = D_\alpha \circ \pi_{\mathrm{dom}(\alpha)} \langle c_i \rangle = D_\alpha \circ c_i.$$

Whence $\phi c = \psi c$ iff for all $\alpha : i \to j$ we have $c_j = D_\alpha \circ c_i$ thus, iff $(c_i : C \to D_i)$ is a cone, as claimed. It follows that (E, e_i) is a cone, and that any cone $(c_i : C \to D_i)$ gives an arrow $\langle c_i \rangle : C \to \prod_i D_i$ with $\phi\langle c_i \rangle = \psi\langle c_i \rangle$, thus there is a unique factorization $u : C \to E$ of $\langle c_i \rangle$ through E, which is clearly a morphism of cones. □

The same proof yields the following:

Corollary 5.24. *A category has all limits of some cardinality iff it has all equalizers and products of that cardinality, where* **C** *has limits (resp. products) of cardinality* κ *iff* **C** *has a limit for every diagram* $D : J \to C$ *where* $\mathrm{card}(J_1) \leq \kappa$ *(resp.* **C** *has all products of* κ *many objects).*

The notions cones and limits of course dualize to give those of *cocones* and *colimits*. One then has the following dual theorem.

Theorem 5.25. *A category* **C** *has finite colimits iff it has finite coproducts and coequalizers (resp. iff it has pushouts and an initial object).* **C** *has all colimits of size* κ *iff it has coequalizers and coproducts of size* κ.

5.5 Preservation of limits

Here is an application of limits by products and equalizers.

Definition 5.26. A functor $F : \mathbf{C} \to \mathbf{D}$ is said to *preserve limits of type* $\mathbf{J}$ if, whenever $p_j : L \to D_j$ is a limit for a diagram $D : \mathbf{J} \to \mathbf{C}$; the cone $Fp_j : FL \to FD_j$ is then a limit for the diagram $FD : \mathbf{J} \to \mathbf{D}$. Briefly

$$F(\varprojlim D_j) \cong \varprojlim F(D_j).$$

A functor that preserves all limits is said to be *continuous*.

For example, let $\mathbf{C}$ be locally small and recall the representable functor

$$\mathrm{Hom}_{\mathbf{C}}(C, -) : \mathbf{C} \to \mathbf{Sets}$$

for any object $C \in \mathbf{C}$, taking $f : X \to Y$ to

$$f_* : \mathrm{Hom}(C, X) \to \mathrm{Hom}(C, Y)$$

where $f_*(g : C \to X) = f \circ g$.

Proposition 5.27. *Representable functors preserve all limits.*

It suffices to show that $\mathrm{Hom}(C, -)$ preserves products and equalizers.

- Suppose $\mathbf{C}$ has a terminal object 1. Then,

$$\mathrm{Hom}_{\mathbf{C}}(C, 1) = \{!_C\} \cong 1.$$

- Consider a binary product $X \times Y$ in $\mathbf{C}$. Then we already know that,

$$\mathrm{Hom}(C, X \times Y) \cong \mathrm{Hom}(C, X) \times \mathrm{Hom}(C, Y)$$

  by composing any $f : C \to X \times Y$ with the two product projections $p_1 : X \times Y \to X$, and $p_2 : X \times Y \to Y$.

- For arbitrary products $\prod_{i \in I} X_i$ one has analogously:

$$\mathrm{Hom}_{\mathbf{C}}(C, \prod_i X_i) \cong \prod_i \mathrm{Hom}_{\mathbf{C}}(C, X_i)$$

- Given an equalizer in $\mathbf{C}$,

$$E \xrightarrow[e]{} X \underset{g}{\overset{f}{\rightrightarrows}} Y$$

  consider the resulting diagram,

$$\mathrm{Hom}(C, E) \xrightarrow[e_*]{} \mathrm{Hom}(C, X) \underset{g_*}{\overset{f_*}{\rightrightarrows}} \mathrm{Hom}(C, Y).$$

  To show this is an equalizer in **Sets**, let $h : C \to X \in \mathrm{Hom}(C, X)$ with $f_*h = g_*h$. Then $fh = gh$, so there is a unique $u : C \to E$ such that

$eu = h$. Thus, we have a unique $u \in \mathrm{Hom}(C, E)$ with $e_* u = eu = h$. So $e_* : \mathrm{Hom}(C, E) \to \mathrm{Hom}(C, X)$ is indeed the equalizer of f_* and g_*.

Definition 5.28. A functor of the form $F : \mathbf{C}^{\mathrm{op}} \to \mathbf{D}$ is called a *contravariant functor* on $\mathbf{C}$. Explicitly, such a functor takes $f : A \to B$ to $F(f) : F(B) \to F(A)$ and $F(g \circ f) = F(f) \circ F(g)$.

A typical example of a contravariant functor is a representable functor of the form,

$$\mathrm{Hom}_{\mathbf{C}}(-, C) : \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$$

for any $C \in \mathbf{C}$ (where $\mathbf{C}$ is any locally small category). Such a contravariant representable functor takes $f : X \to Y$ to

$$f^* : \mathrm{Hom}(Y, C) \to \mathrm{Hom}(X, C)$$

by $f^*(g : X \to C) = g \circ f$.

The dual version of the foregoing proposition is then this:

Corollary 5.29. *Contravariant representable functors map all colimits to limits.*

For example, given a coproduct $X + Y$ in any locally small category $\mathbf{C}$, there is a canonical isomorphism,

$$\mathrm{Hom}(X + Y, C) \;\cong\; \mathrm{Hom}(X, C) \times \mathrm{Hom}(Y, C) \tag{5.1}$$

given by precomposing with the two coproduct inclusions.

From an example in Section 2.3 we can therefore conclude that the ultrafilters in a coproduct $A + B$ of Boolean algebras correspond exactly to pairs of ultrafilters (U, V), with U in A and V in B. This follows because we showed there that the ultrafilter functor $\mathrm{Ult} : \mathbf{Bool}^{\mathrm{op}} \to \mathbf{Sets}$ is representable:

$$\mathrm{Ult}(B) \;\cong\; \mathrm{Hom}_{\mathbf{Bool}}(B, 2).$$

Another case of the above iso (5.1) is the familiar law of exponents for sets:

$$C^{X+Y} \cong C^X \times C^Y$$

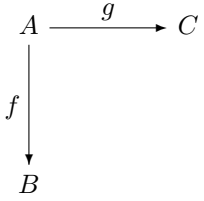
The arithmetical law of exponents $k^{m+n} = k^n \cdot k^m$ is actually a special case of this!

5.6 Colimits

Let us briefly discuss some special colimits, since we did not really say much about them in the foregoing section.

First, we consider *pushouts* in **Sets**.

Suppose we have two functions



We can construct the pushout of f and g like this. Start with the coproduct (disjoint sum):

$$B \longrightarrow B + C \longleftarrow C$$

Now identify those elements $b \in B$ and $c \in C$ such that, for some $a \in A$,

$$f(a) = b \quad \text{and} \quad g(a) = c$$

That is, we take the equivalence relation $\sim$ on $B+C$ generated by the conditions $f(a) \sim g(a)$ for all $a \in A$.

Then we take the quotient by $\sim$ to get the pushout

$$B +_A C \;\cong\; (B + C)/\sim$$

which can be imagined as B placed next to C, with the respective parts that are images of A overlapping. This construction follows simply by dualizing the one for pullbacks by products and equalizers.

In general, a colimit for a diagram $D : \mathbf{J} \to \mathbf{C}$ is of course an initial object in the category of *cocones*. Explicitly, a *cocone from the base* D consists of an object C (the vertex) and arrows $c_j : D_j \to C$ for each $j \in \mathbf{J}$, such that for all $\alpha : i \to j$ in $\mathbf{J}$,

$$c_j \circ D(\alpha) = c_i$$

A morphism of cocones $f : (C, (c_j)) \to (C', (c_j{}'))$ is an arrow $f : C \to C'$ in $\mathbf{C}$ such that $f \circ c_j = c_j{}'$ for all $j \in \mathbf{J}$. An initial cocone is the expected thing: one that maps uniquely *to* any other cocone from D.

We write such a colimit in the form:

$$\varinjlim_{j \in \mathbf{J}} D_j$$

Now let us consider some examples of a particular kind of colimit that comes up quite often. Our first example is what is sometimes called a *direct limit* of a sequence of algebraic objects, say groups. A similar construction will work for any sort of algebras (but non-equational conditions are not always preserved by direct limits).

Example 5.30. *Direct limit of groups.* Suppose we're given a sequence,

$$G_0 \underset{g_0}{\longrightarrow} G_1 \underset{g_1}{\longrightarrow} G_2 \underset{g_2}{\longrightarrow} \cdots$$

of groups and homomorphisms, and we want a "colimiting" group G_∞ with homomorphisms

$$u_n : G_n \to G_\infty$$

satisfying $u_{n+1} \circ g_n = u_n$. Moreover, G_∞ should be "universal" with this property. I think you can see the colimit setup here:

- the index category is the ordinal number $\omega = (\mathbb{N}, \leq)$, regarded as a poset category,
- the sequence

$$G_0 \underset{g_0}{\longrightarrow} G_1 \underset{g_1}{\longrightarrow} G_2 \underset{g_2}{\longrightarrow} \cdots$$

  is a diagram of type ω in the category **Groups**,
- the colimiting group is the colimit of the sequence:

$$G_\infty \cong \varinjlim_{n \in \omega} G_n$$

This group always exists, and can be constructed as follows. Begin with the coproduct (disjoint sum) of sets

$$\coprod_{n \in \omega} G_n.$$

Then make identifications $x_n \sim y_m$, where $x_n \in G_n$ and $y_m \in G_m$, to ensure in particular that

$$x_n \sim g_n(x_n)$$

for all $x_n \in G_n$ and $g_n : G_n \to G_{n+1}$.

This means, specifically, that the elements of G_∞ are equivalence classes of the form

$$[x_n], \quad x_n \in G_n$$

for any n, and $[x_n] = [y_m]$ iff for some $k \geq m, n$,

$$g_{n,k}(x_n) = g_{m,k}(y_m)$$

where, generally, if $i \leq j$, we define

$$g_{i,j} : G_i \to \cdots \to G_j$$

by composing consecutive g's as in $g_{i,j} = g_{j-1} \circ \ldots \circ g_i$. The reader can easily check that this is indeed the equivalence relation generated by all the conditions $x_n \sim g_n(x_n)$.

The operations on G_∞ are now defined by

$$[x] \cdot [y] = [x' \cdot y']$$

where $x \sim x'$, $y \sim y'$, and $x', y' \in G_n$ for n sufficiently large. The unit is just $[u_0]$, and we take,

$$[x]^{-1} = [x^{-1}].$$

One can easily check that these operations are well defined, and determine a group structure on G_∞, which moreover makes all the evident functions

$$u_n : G_n \to G_\infty , \qquad u_n(x) = [x]$$

into homomorphisms.

The universality of G_∞ and the u_n results from the fact that the construction is essentially a colimit in **Sets**, equipped with an induced group structure. Indeed, given any group H and homomorphisms $h_n : G_n \to H$ with $h_{n+1} \circ g_n = h_n$ define $h_\omega : G_\omega \to H$ by $g_\omega([x_n]) = g_n(x_n)$. This is easily seen to be well defined and indeed a homomorphism. Moreover, it is the unique *function* that commutes with all the u_n.

The fact that the ω-colimit G_∞ of groups can be constructed as the colimit of the underlying sets is a case of a general phenomenon, expressed by saying that the forgetful functor $U :$ **Groups** $\to$ **Sets** "creates ω-colimits."

Definition 5.31. A functor $F : \mathbf{C} \to \mathbf{D}$ is said to *create limits of type* $\mathbf{J}$ if for every diagram $C : \mathbf{J} \to \mathbf{C}$ and limit $p_j : L \to FC_j$ in $\mathbf{D}$ there is a unique cone $\overline{p_j} : \overline{L} \to C_j$ in $\mathbf{C}$ with $F(\overline{L}) = L$ and $F(\overline{p_j}) = p_j$, which, furthermore, is a limit for C. Briefly, every limit in $\mathbf{D}$ is the image of a unique cone in $\mathbf{C}$, which is a limit there. The notion of *creating colimits* is defined analogously.

In these terms, then, we have the following proposition, the remaining details of which are left as an exercise.

Proposition 5.32. *The forgetful functor* $U :$ **Groups** $\to$ **Sets** *creates* ω*-colimits. It also creates all limits.*

The same fact holds quite generally for other categories of algebraic objects, that is, sets equipped with operations satisfying some equations. Observe that not *all* colimits are created in this way. For instance, we have already seen that the coproduct of two abelian groups has their product as underlying set, since $G + H \cong G \times H$.

Example 5.33. *Cumulative hierarchy.* Another example of this kind is the "cumulative hierarchy" construction encountered in set theory. Let us set

$$\begin{aligned} V_0 &= \emptyset \\ V_1 &= P(\emptyset) \\ &\vdots \\ V_{n+1} &= P(V_n) \end{aligned}$$

Then there is a sequence of subset inclusions,

$$\emptyset = V_0 \subseteq V_1 \subseteq V_2 \subseteq \cdots$$

since, generally, $A \subseteq B$ implies $P(A) \subseteq P(B)$ for any sets A and B. The colimit of the sequence

$$V_\omega = \varinjlim_n V_n$$

is called the *cumulative hierarchy* of rank ω. One can of course continue this construction through higher ordinals $\omega + 1, \omega + 2, \ldots$.

More generally, let us start with some set A (of "atoms"), and let

$$V_0(A) = A$$

and then put

$$V_{n+1}(A) = A + P(V_n(A))$$

There is a sequence $V_0(A) \to V_1(A) \to V_2(A) \to \ldots$ as follows. Let

$$v_0 : V_0(A) = A \to A + P(A) = V_1(A)$$

be the left coproduct inclusion. Given $v_{n-1} : V_{n-1}(A) \to V_n(A)$, let $v_n : V_n(A) \to V_{n+1}(A)$ be defined by

$$v_n = 1_A + P_!(v_{n-1}) : A + P(V_{n-1}(A)) \to A + P(V_n(A))$$

where $P_!$ denotes the *covariant* powerset functor, taking a function $f : X \to Y$ to the "image under f" operation $P_!(f) : P(X) \to P(Y)$ defined by taking $U \subseteq X$ to

$$P_!(f)(U) = \{f(u) \mid u \in U\} \subseteq Y$$

The idea behind the sequence is that we start with A, then add all the subsets of A, then add all the *new* subsets that can be formed from all of those, and so on. The colimit of the sequence

$$V_\omega(A) = \varinjlim_n V_n(A)$$

is called the *cumulative hierarchy* (of rank ω) over A. Of course, $V_\omega = V_\omega(\emptyset)$.

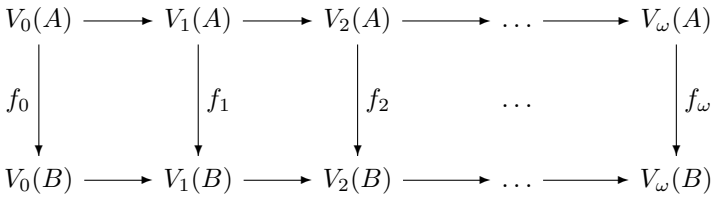
Now suppose we have some function

$$f : A \to B$$

Then there is a map

$$V_\omega(f) : V_\omega(A) \to V_\omega(B)$$

determined by the colimit description of V_ω, as indicated in the following diagram.



Here the f_n are defined by

$$\begin{aligned} f_0 &= f : A \to B \\ f_1 &= f + P_!(f) : A + P(A) \to B + P(B) \\ &\vdots \\ f_{n+1} &= f + P_!(f_n) : A + P(V_n(A)) \to B + P(V_n(B)) \end{aligned}$$

Since all the squares clearly commute, we have a cocone on the diagram $V_n(A)$ with vertex $V_\omega(B)$, and there is thus a unique $f_\omega : V_\omega(A) \to V_\omega(B)$ that completes the diagram.

Thus we see that the cumulative hierarchy is functorial.

Example 5.34. *ωCPOs.* An *ωCPO* is a poset that is "ω-cocomplete," meaning it has all colimits of type $\omega = (N, \leq)$. Specifically, a poset D is an ωCPO if for every diagram $d : \omega \to D$, i.e. every chain of elements of D,

$$d_0 \leq d_1 \leq d_2 \leq \cdots$$

we have a colimit $d_\omega = \varinjlim d_n$. This is an element such that:

1. $d_n \leq d_\omega$ for all $n \in \omega$;
2. for all $x \in D$; if $d_n \leq x$ for all $n \in \omega$, then also $d_\omega \leq x$.

A monotone map of ωCPOs

$$h : D \to E$$

is usually called *continuous* if it preserves colimits of type ω, that is,

$$h(\varinjlim d_n) = \varinjlim h(d_n).$$

An application of these notions is the following:

Proposition 5.35. *If D is an ωCPO with initial element 0 and*

$$h : D \to D$$

is continuous, then h has a fixed point

$$h(x) = x$$

which, moreover, is least among all fixed points.

Proof. We use "Newton's method," which can be used, for example, to find fixed points of monotone, continuous functions $f : [0,1] \to [0,1]$. Consider the sequence $d : \omega \to D$, defined by

$$d_0 = 0$$
$$d_{n+1} = h(d_n)$$

Since $0 \leq d_0$, repeated application of h gives $d_n \leq d_{n+1}$. Now take the colimit $d_\omega = \varinjlim_{n\in\omega} d_n$. Then

$$\begin{aligned} h(d_\omega) &= h(\varinjlim_{n\in\omega} d_n) \\ &= \varinjlim_{n\in\omega} h(d_n) \\ &= \varinjlim_{n\in\omega} d_{n+1} \\ &= d_\omega. \end{aligned}$$

The last step follows because the first term $d_0 = 0$ of the sequence is trivial.

Moreover, if x is also a fixed point, $h(x) = x$, then we have

$$\begin{aligned} d_0 &= 0 \leq x \\ d_1 &= h(0) \leq h(x) = x \\ &\vdots \\ d_{n+1} &= h(d_n) \leq h(x) = x. \end{aligned}$$

So also $d_\omega \leq x$, since d_ω is a colimit. □

Finally, here is an example of how (co)limits depend on the ambient category; we consider colimits *of* posets and ωCPOs, rather than *in* them.

Let us define the finite ωCPOs

$$\omega_n = \{k \leq n \mid k \in \omega\}$$

then we have continuous inclusion maps:

$$\omega_0 \to \omega_1 \to \omega_2 \to \cdots$$

In **Pos**, the colimit exists, and is ω, as can be easily checked. *But* ω itself is not ω-complete. Indeed, the sequence

$$0 \leq 1 \leq 2 \leq \cdots$$

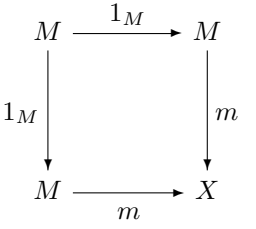
has no colimit. So the colimit of the ω_n in the category of ωCPOs, if it exists, must be something else. In fact it is $\omega + 1$.

$$0 \leq 1 \leq 2 \leq \cdots \leq \omega$$

For then any bounded sequence has a colimit in the bounded part and any unbounded one has ω as colimit.

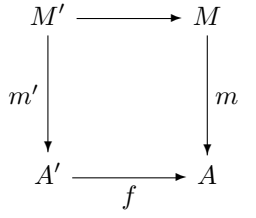
5.7 Exercises

1. Show that an arrow $m : M \to C$ in any category is monic if and only if the diagram below is a pullback.



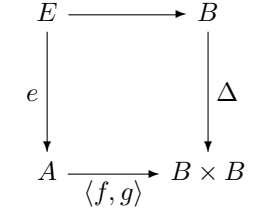
   Conclude that representable functors $\mathrm{Hom}(C, -)$ preserve monos.

2. Show that in any category, given a pullback square



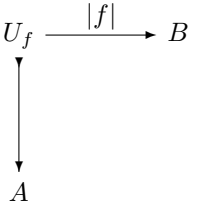
   if m is monic, then so is m'.

3. (Equalizers by pullbacks and products) Show that a category with pullbacks and products has equalizers as follows: given arrows $f, g : A \to B$, take the pullback indicated below, where $\Delta = \langle 1_B, 1_B \rangle$:

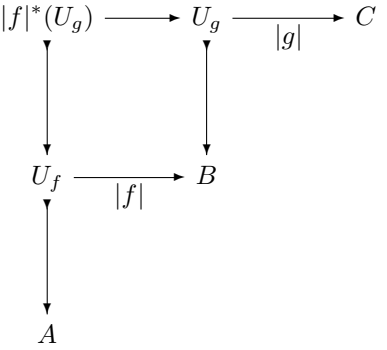


   Show that $e : E \to A$ is the equalizer of f and g.

4. (Partial maps) For any category $\mathbf{C}$ with pullbacks, define the category $\mathbf{Par}(\mathbf{C})$ of partial maps in $\mathbf{C}$ as follows: the objects are the same as those of $\mathbf{C}$, but an arrow $f : A \to B$ is a pair $(|f|, U_f)$ where $U_f \rightarrowtail A$ is a subobject (an equivalence class of monomorphisms) and $|f| : U_f \to B$ (take a suitably defined equivalence class of arrows), as indicated in the diagram



Composition of $(|f|, U_f) : A \to B$ and $(|g|, U_g) : B \to C$ is given by taking a pullback and then composing to get $(|g \circ f|, |f|^*(U_g))$, as suggested by the follow diagram.



Check to see that this really does define a category.

5. (Pushouts)
   (a) Dualize the definition of a pullback to define the "copullback" (usually called the "pushout") of two arrows with common domain.
   (b) Indicate how to construct pushouts using coproducts and coequalizers (proof "by duality").

This page intentionally left blank

6

EXPONENTIALS

We have now managed to unify most of the universal mapping properties (UMP) that we have seen so far with the notion of limits (or colimits). Of course, the free algebras are an exception to this. In fact, it will turn out that there is a common source of UMP's, but it lies somewhat deeper, in the notion of *adjoints*, which unify free algebras, limits, and other universals of various kinds.

Next we are going to look at one more elementary universal structure, which is also an example of a universal that is not a limit. This important structure is called an "exponential" and it can be thought of as a categorical notion of a "function space." As we'll see it subsumes much more than just that, however.

6.1 Exponential in a category

Let us start by considering a function of sets,

$$f(x, y) : A \times B \to C$$

written using variables x over A and y over B. If we now hold $a \in A$ fixed, we have a function

$$f(a, y) : B \to C$$

and thus an element

$$f(a, y) \in C^B$$

of the set of all such functions.

Letting a vary over A then gives a map, which I will write like this:

$$\tilde{f} : A \to C^B$$

defined by $a \mapsto f(a, y)$.

The map $\tilde{f} : A \to C^B$ takes the "parameter" a to the function $f_a(y) : B \to C$. It's uniquely determined by the equation

$$\tilde{f}(a)(b) = f(a, b).$$

Indeed, *any* map

$$\phi : A \to C^B$$

is uniquely of the form

$$\phi = \tilde{f}$$

for some $f : A \times B \to C$. For we can set

$$f(a, b) := \phi(a)(b).$$

What this means, in sum, is that we have an isomorphism of Hom-sets:

$$\mathrm{Hom}_{\mathbf{Sets}}(A \times B, C) \cong \mathrm{Hom}_{\mathbf{Sets}}(A, C^B)$$

That is, there is a bijective correspondence between functions of the form $f : A \times B \to C$ and those of the form $\tilde{f} : A \to C^B$. Moreover, this bijection is mediated by a certain operation of *evaluation*, which we have indicated in the foregoing by using variables. In order to generalize the indicated bijection to other categories, we're going to need to make this evaluation operation explicit, too.

In **Sets**, it is the function

$$\mathrm{eval} : C^B \times B \to C$$

defined by $(g, b) \mapsto g(b)$, that is,

$$\mathrm{eval}(g, b) = g(b).$$

This evaluation function has the following UMP: given any set A and any function

$$f : A \times B \to C$$

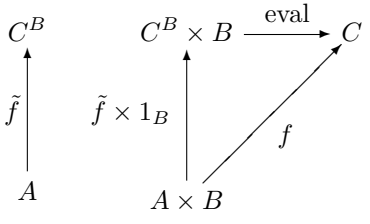
there is a unique function

$$\tilde{f} : A \to C^B$$

such that $\mathrm{eval} \circ (\tilde{f} \times 1_B) = f$. That is,

$$\mathrm{eval}(\tilde{f}(a), b) = f(a, b). \tag{6.1}$$

Here is the diagram:



You can read the equation (6.1) off from this diagram by taking a pair of elements $(a, b) \in A \times B$ and chasing them around both ways, using the fact that $(\tilde{f} \times 1_B)(a, b) = (\tilde{f}(a), b)$.

Now, the property just stated of the evaluation function and the set C^B of functions $B \to C$ is one that will make sense in any category having binary products.

So, in that form, we can use it to define the notion we seek.

Definition 6.1. Let the category $\mathbf{C}$ have binary products. An *exponential* of objects B and C of $\mathbf{C}$ consists of an object

$$C^B$$

and an arrow

$$\epsilon : C^B \times B \to C$$

such that, for any object Z and arrow

$$f : Z \times B \to C$$

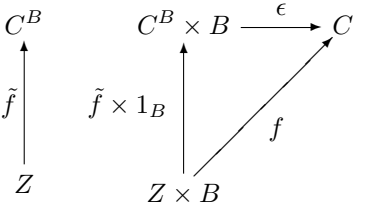
there is a unique arrow

$$\tilde{f} : Z \to C^B$$

such that

$$\epsilon \circ (\tilde{f} \times 1_B) = f$$

all as in the diagram



Here is some terminology

- $\epsilon : C^B \times B \to C$ is called *evaluation*.
- $\tilde{f} : Z \to C^B$ is called the (exponential) *transpose* of f.
- Given an arrow

  $$g : Z \to C^B$$

  we write

  $$\bar{g} := \epsilon(g \times 1_B) : Z \times B \to C$$

  and also call $\bar{g}$ the *transpose* of g. By the uniqueness clause of the definition, then

  $$\bar{\bar{g}} = g$$

and for any $f : Z \times B \to C$,

$$\bar{\bar{f}} = f.$$

Briefly, transposition of transposition is the identity.

Thus transposition provides the desired isomorphism,

$$\mathrm{Hom}_{\mathbf{C}}(Z \times B, C) \;\cong\; \mathrm{Hom}_{\mathbf{C}}(Z, C^B)$$

where $f \mapsto \tilde{f}$ and $g \mapsto \bar{g}$.

6.2 Cartesian closed categories

Definition 6.2. A category is called *cartesian closed* if it has all finite products and exponentials.

Example 6.3. We already have **Sets** as one example, but note that also $\mathbf{Sets}_{fin}$ is cartesian closed, since for finite sets M, N, the set of functions N^M has cardinality

$$|N^M| = |N|^{|M|}$$

and so is also finite.

Example 6.4. Recall that the category **Pos** of posets has as arrows $f : P \to Q$ the monotone functions, $p \leq p'$ implies $fp \leq fp'$. Given posets P and Q, the poset $P \times Q$ has pairs (p, q) as elements, and is partially ordered by

$$(p, q) \leq (p', q') \quad \text{iff} \quad p \leq p' \text{ and } q \leq q'.$$

Thus the evident projections

$$P \xleftarrow[\pi_1]{} P \times Q \xrightarrow[\pi_2]{} Q$$

are monotone, as is the pairing $\langle f, g \rangle : X \to P \times Q$ if $f : X \to P$ and $g : X \to Q$ are monotone.

For the exponential Q^P, we take the set of monotone functions,

$$Q^P = \{ f : P \to Q \mid f \text{ monotone } \}$$

ordered *pointwise*, that is,

$$f \leq g \quad \text{iff} \quad fp \leq gp \text{ for all } p \in P.$$

The evaluation

$$\epsilon : Q^P \times P \to Q$$

and transposition

$$\tilde{f} : X \to Q^P$$

of a given arrow

$$f : X \times P \to Q$$

are the usual ones of the underlying functions. Thus we need only show that these are monotone.

To that end, given $(f, p) \leq (f', p')$ in $Q^P \times P$ we have

$$\begin{aligned} \epsilon(f, p) &= f(p) \\ &\leq f(p') \\ &\leq f'(p') \\ &= \epsilon(f', p') \end{aligned}$$

so ϵ is monotone. Now take $f : X \times P \to Q$ monotone and let $x \leq x'$. We need to show

$$\tilde{f}(x) \leq \tilde{f}(x') \quad \text{in } Q^P$$

which means

$$\tilde{f}(x)(p) \leq \tilde{f}(x')(p) \quad \text{for all } p \in P.$$

But

$$\begin{aligned} \tilde{f}(x)(p) &= f(x, p) \\ &\leq f(x', p) \\ &\leq f'(x', p) \\ &= \tilde{f}(x')(p). \end{aligned}$$

Example 6.5. Now let us consider what happens if we restrict to the category of ωCPOs (see example 5.34). Given two ωCPOs P and Q, we will take as an exponential the subset,

$$Q^P = \{f : P \to Q \mid f \text{ monotone and } \omega\text{-continuous}\}.$$

Then take evaluation $\epsilon : Q^P \times P \to Q$ and transposition as before, for functions. Then, since we know that the required equations are satisfied, we just need to check the following:

- Q^P is an ωCPO
- ϵ is ω-continuous
- $\tilde{f}$ is ω-continuous if f is

We leave this as an exercise!

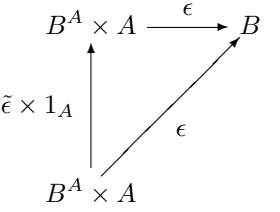
We now derive some of the basic facts about exponentials and cartesian closed categories. First, let us ask, what is the transpose of evaluation?

$$\epsilon : B^A \times A \to B$$

It must be an arrow $\tilde{\epsilon} : B^A \to B^A$ such that

$$\epsilon(\tilde{\epsilon} \times 1_A) = \epsilon$$

that is, making the following diagram commute:



Since $1_{B^A} \times 1_A = 1_{(B^A \times A)}$ clearly has this property, we must have

$$\tilde{\epsilon} = 1_{B^A}$$

and so we also know that $\epsilon = \overline{(1_{B^A})}$.

Now let us show that the operation $X \mapsto X^A$ on a CCC is *functorial*.

Proposition 6.6. *In any cartesian closed category* $\mathbf{C}$*, exponentiation by a fixed object* A *is a functor,*

$$-^A : \mathbf{C} \to \mathbf{C}.$$

Toward the proof, consider first the case of sets. Given some function

$$\beta : B \to C$$

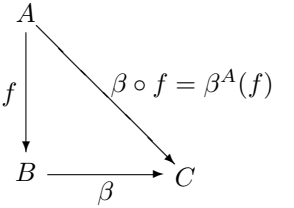
we put

$$\beta^A : B^A \to C^A$$

defined by

$$f \mapsto \beta \circ f.$$

That is,



This assignment is functorial, because: for any $\alpha : C \to D$

$$\begin{aligned}(\alpha \circ \beta)^A(f) &= \alpha \circ \beta \circ f \\ &= \alpha \circ \beta^A(f) \\ &= \alpha^A \circ \beta^A(f).\end{aligned}$$

Whence $(\alpha \circ \beta)^A = \alpha^A \circ \beta^A$. Also

$$\begin{aligned}(1_B)^A(f) &= 1_B \circ f \\ &= f \\ &= 1_{B^A}(f).\end{aligned}$$

So $(1_B)^A = 1_{B^A}$.

In a general CCC then, given $\beta : B \to C$, we define

$$\beta^A : B^A \to C^A$$

by

$$\beta^A := \widetilde{(\beta \circ \epsilon)}.$$

That is, we take the transpose of the composite

$$B^A \times A \xrightarrow{\epsilon} B \xrightarrow{\beta} C$$

giving

$$\beta^A : B^A \to C^A.$$

It is easier to see in the form

$$\begin{array}{ccc} C^A & & C^A \times A \xrightarrow{\epsilon} C \\ \uparrow{\scriptstyle \beta^A} & & \uparrow{\scriptstyle \beta^A \times 1_A} \quad \uparrow{\scriptstyle \beta} \\ B^A & & B^A \times A \xrightarrow[\epsilon]{} B \end{array}$$

Now, clearly,

$$(1_B)^A = 1_{B^A} : B^A \to B^A$$

by examining

$$\begin{array}{ccc} B^A \times A & \longrightarrow & B \\ \uparrow{\scriptstyle 1_{(B^A \times A)} = 1_{B^A} \times 1_A} & & \uparrow{\scriptstyle 1_B} \\ B^A \times A & \xrightarrow[\epsilon]{} & B \end{array}$$

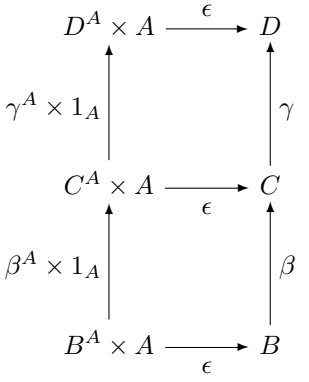
Quite similarly, given

$$B \xrightarrow{\beta} C \xrightarrow{\gamma} D$$

we have

$$\gamma^A \circ \beta^A = (\gamma \circ \beta)^A.$$

This follows from considering the commutative diagram



We use the fact that

$$(\gamma^A \times 1_A) \circ (\beta^A \times 1_A) = ((\gamma^A \circ \beta^A) \times 1_A).$$

The result follows by the uniqueness of transposes.

This suggests looking for another "universal" arrow, namely the transpose of the identity $1_{A\times B} : A \times B \to A \times B$,

$$\tilde{1}_{A\times B} : A \to (A \times B)^B.$$

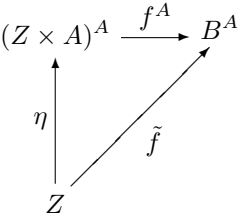
In **Sets** it has the values $\tilde{1}_{A\times B}(a)(b) = (a, b)$. Let us denote this map by $\eta = \tilde{1}_{A\times B}$, so that

$$\eta(a)(b) = (a, b).$$

The map η lets us compute $\tilde{f}$ from the functor $-^A$. Indeed, given $f : Z \times A \to B$ take

$$f^A : (Z \times A)^A \to B^A$$

and precompose with $\eta : Z \to (Z \times A)^A$, as indicated in



This gives the useful equation

$$\tilde{f} = f^A \circ \eta$$

which the reader should prove.

6.3 Heyting algebras

Any Boolean algebra B, regarded as a poset category, has finite products 1 and $a \wedge b$. We can also define the exponential in B by

$$b^a = (\neg a \vee b)$$

which we will also write $a \Rightarrow b$. The evaluation arrow is

$$(a \Rightarrow b) \wedge a \leq b.$$

This always holds since

$$(\neg a \vee b) \wedge a = (\neg a \wedge a) \vee (b \wedge a) = 0 \vee (b \wedge a) = b \wedge a \leq b.$$

To show that $a \Rightarrow b$ is indeed an exponential in B, we just need to verify that if $a \wedge b \leq c$ then $a \leq b \Rightarrow c$, that is, transposition. But if $a \wedge b \leq c$, then

$$\neg b \vee (a \wedge b) \leq \neg b \vee c = b \Rightarrow c.$$

But we also have

$$a \leq \neg b \vee a \leq (\neg b \vee a) \wedge (\neg b \vee b) = \neg b \vee (a \wedge b).$$

This example suggests generalizing the notion of a Boolean algebra to that of a cartesian closed poset. Indeed, consider first the following useful notion.

Definition 6.7. A *Heyting algebra* is a poset with

1. finite meets: 1 and $p \wedge q$,
2. finite joins: 0 and $p \vee q$,
3. exponentials: for each a, b, an element $a \Rightarrow b$ such that

$$a \wedge b \leq c \quad \text{iff} \quad a \leq b \Rightarrow c.$$

The stated condition on exponentials $a \Rightarrow b$ is equivalent to the UMP in the case of posets. Indeed, given the condition, the transpose of $a \wedge b \leq c$ is $a \leq b \Rightarrow c$ and the evaluation $(a \Rightarrow b) \wedge a \leq b$ follows immediately from $a \Rightarrow b \leq a \Rightarrow b$ (the converse is just as simple).

Every Heyting algebra is a *distributive lattice*, since we have

$$\begin{aligned}(a \vee b) \wedge c \leq z \text{ iff } & a \vee b \leq c \Rightarrow z \\ \text{iff } & a \leq c \Rightarrow z \text{ and } b \leq c \Rightarrow z \\ \text{iff } & a \wedge c \leq z \text{ and } b \wedge c \leq z \\ \text{iff } & (a \wedge c) \vee (b \wedge c) \leq z.\end{aligned}$$

Now pick $z = (a \vee b) \wedge c$, respectively $z = (a \wedge c) \vee (b \wedge c)$.

One may well wonder whether all distributive lattices are Heyting algebras. The answer is in general, no; but certain ones always are.

Definition 6.8. A poset is *(co)complete* if it is so as a category, thus if it has all set-indexed meets $\bigwedge_{i\in I} a_i$ (resp. joins $\bigvee_{i\in I} a_i$). For posets, completeness and cocompleteness are equivalent (exercise!). A lattice, Heyting algebra, Boolean algebra, etc. is called *complete* if it is so as a poset.

Proposition 6.9. *A complete lattice is a Heyting algebra iff it satisfies the infinite distributive law*

$$a \wedge \left(\bigvee_i b_i \right) = \bigvee_i (a \wedge b_i).$$

Proof. One shows that Heyting algebra implies distributivity just as in the finite case. To show that the infinite distributive law implies Heyting algebra, set

$$a \Rightarrow b = \bigvee_{x \wedge a \leq b} x.$$

Then, if

$$y \wedge a \leq b$$

then $y \leq \bigvee_{x\wedge a\leq b} x = a \Rightarrow b$. And conversely, if $y \leq a \Rightarrow b$ then $y \wedge a \leq (\bigvee_{x\wedge a\leq b} x) \wedge a = \bigvee_{x\wedge a\leq b}(x \wedge a) \leq \bigvee b = b$. □

Example 6.10. For any set A, the powerset $P(A)$ is a complete Heyting algebra with unions and intersections as joins and meets, since it satisfies the infinite distributive law. More generally, the lattice of open sets of a topological space is also a Heyting algebra, since the open sets are closed under finite intersections and arbitrary unions.

Of course, every Boolean algebra is a Heyting algebra with $a \Rightarrow b = \neg a \vee b$, as we already showed. But in general, a Heyting algebra is not Boolean. Indeed, we can define a proposed negation by,

$$\neg a = a \Rightarrow 0$$

as must be the case, since in a Boolean algebra $\neg a = \neg a \vee 0 = a \Rightarrow 0$. Then $a \leq \neg\neg a$ since $a \wedge (a \Rightarrow 0) \leq 0$. But, conversely, $\neg\neg a \leq a$ need not hold in a Heyting algebra. Indeed, in a topological space X, the negation $\neg U$ of an open subset U is the *interior* of the complement $X - U$. Thus, for example, in the real interval $[0, 1]$ we have $\neg\neg(0, 1) = [0, 1]$.

Moreover, the law,

$$1 \leq a \vee \neg a$$

also need not hold in general. In fact, the concept of a Heyting algebra is the algebraic equivalent of the *intuitionistic* propositional calculus, in the same sense that Boolean algebras are an algebraic formulation of the *classical* propositional calculus.

To make this more precise, we first give a system of rules for the intuitionistic propositional calculus (IPC) in terms of entailments $p \vdash q$ between formulas p and q:

1. $\vdash$ is reflexive and transitive
2. $p \vdash \top$
3. $\bot \vdash p$
4. $p \vdash q$ and $p \vdash r$ iff $p \vdash q \wedge r$
5. $p \vdash r$ and $q \vdash r$ iff $p \vee q \vdash r$
6. $p \wedge q \vdash r$ iff $p \vdash q \Rightarrow r$

This is a complete system for IPC, equivalent to the more standard presentations the reader may have seen. To show this, note first that we have an "evaluation" entailment by reflexivity and (6):

$$\begin{aligned} p \Rightarrow q &\vdash p \Rightarrow q \\ (p \Rightarrow q) \wedge p &\vdash q \end{aligned}$$

We therefore have the rule of "modus ponens" by (4) and transitivity:

$$\begin{aligned} &\top \vdash p \Rightarrow q \quad \text{and} \quad \top \vdash p \\ &\top \vdash (p \Rightarrow q) \wedge p \\ &\top \vdash q \end{aligned}$$

Moreover, by (4) there are "projections":

$$\begin{aligned} p \wedge q &\vdash p \wedge q \\ p \wedge q &\vdash p \quad (\text{resp. } q) \end{aligned}$$

from which it follows that $p \dashv\vdash \top \wedge p$. Thus, we get one of the usual axioms for products:

$$\begin{aligned} p \wedge q &\vdash p \\ \top \wedge (p \wedge q) &\vdash p \\ \top &\vdash (p \wedge q) \Rightarrow p \end{aligned}$$

Now let us derive the usual axioms for $\Rightarrow$, namely:

1. $p \Rightarrow p$,
2. $p \Rightarrow (q \Rightarrow p)$,
3. $(p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r))$.

The first two are almost immediate:

$$\begin{array}{r} p \vdash p \\ \top \wedge p \vdash p \\ \top \vdash p \Rightarrow p \end{array}$$

$$\begin{array}{r} p \wedge q \vdash p \\ p \vdash q \Rightarrow p \\ \top \wedge p \vdash (q \Rightarrow p) \\ \top \vdash p \Rightarrow (q \Rightarrow p) \end{array}$$

For the third one, we shall use the fact that $\Rightarrow$ distributes over $\wedge$ on the right:

$$a \Rightarrow (b \wedge c) \dashv\vdash (a \Rightarrow b) \wedge (a \Rightarrow c)$$

This is a special case of the exercise:

$$(B \times C)^A \;\cong\; B^A \times C^A$$

We also use the following simple fact, which will be recognized as a special case of proposition 6.6:

$$a \vdash b \quad \text{implies} \quad p \Rightarrow a \vdash p \Rightarrow b \tag{6.2}$$

Then we have,

$$\begin{array}{rl} (q \Rightarrow r) \wedge q \vdash r & \\ p \Rightarrow ((q \Rightarrow r) \wedge q) \vdash p \Rightarrow r & \\ (p \Rightarrow (q \Rightarrow r)) \wedge (p \Rightarrow q) \vdash p \Rightarrow r & \text{by (6.3)} \\ (p \Rightarrow (q \Rightarrow r)) \vdash (p \Rightarrow q) \Rightarrow (p \Rightarrow r) & \\ \top \vdash (p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \wedge (p \Rightarrow r)). & \end{array}$$

The "positive" fragment of IPC, involving only the logical operations

$$\top,\ \wedge,\ \Rightarrow$$

corresponds to the notion of a cartesian closed poset. We then add $\bot$ and disjunction $p \vee q$ on the logical side and finite joins on the algebraic side to arrive at a correspondence between IPC and Heyting algebras. The exact correspondence is given by mutually inverse constructions between Heyting algebras and intuitionistic propositional calculi. We briefly indicate the interesting direction, leaving the more routine one to the reader's ingenuity.

Given any intuitionistic propositional calculus $\mathcal{L}$, consisting of propositional formulas $p, q, r, \ldots$ over some set of variables $x, y, z, \ldots$ together with the rules

of inference stated above, and perhaps some distinguished formulas $a, b, c, \ldots$ as axioms, one constructs from $\mathcal{L}$ a Heyting algebra $\mathrm{HA}(\mathcal{L})$, called the *Lindenbaum-Tarski algebra*, consisting of equivalence classes $[p]$ of formulas p, where:

$$[p] = [q] \quad \text{iff} \quad p \dashv\vdash q \tag{6.3}$$

The ordering in $\mathrm{HA}(\mathcal{L})$ is given by:

$$[p] \leq [q] \quad \text{iff} \quad p \vdash q \tag{6.4}$$

This is clearly well defined on equivalence classes, in the sense that if $p \vdash q$ and $[p] = [p']$ then $p' \vdash q$, and similarly for q. The operations in $\mathrm{HA}(\mathcal{L})$ are then induced in the expected way by the logical operations:

$$\begin{aligned} 1 &= [\top] \\ 0 &= [\bot] \\ [p] \wedge [q] &= [p \wedge q] \\ [p] \vee [q] &= [p \vee q] \\ [p] \Rightarrow [q] &= [p \Rightarrow q] \end{aligned}$$

Again, these operations are easily seen to be well defined on equivalence classes, and they satisfy the laws for a Heyting algebra because the logical rules evidently imply them.

By (6.3) this Heyting algebra then has the property that a formula p is provable in $\mathcal{L}$ just if $[p] = 1$. We, therefore, have the following logical *completeness theorem for IPC*.

Proposition 6.11. *The intuitionistic propositional calculus is complete with respect to models in Heyting algebras.*

Proof. Suppose a formula p is true in all Heyting algebras. Then in particular, it is so in $\mathrm{HA}(\mathcal{L})$. Thus $p = 1$ in $\mathrm{HA}(\mathcal{L})$, and so $\top \vdash p$. □

In sum, then, a particular instance $\mathcal{L}$ of intuitionistic propositional calculus can be regarded as a way of specifying (and reasoning about) a particular Heyting algebra $\mathrm{HA}(\mathcal{L})$, that is, it is a presentation by generators and relations, in just the way that we have already seen other algebraic objects like monoids have such presentations. Indeed, the Heyting algebra $\mathrm{HA}(\mathcal{L})$ even has a UMP with respect to $\mathcal{L}$ that is entirely analogous to the UMP of a monoid that is finitely presented by generators and relations. Specifically, if, for instance, $\mathcal{L}$ is generated by the two elements a, b and the single "axiom" $a \vee b \Rightarrow a \wedge b$, then in $\mathrm{HA}(\mathcal{L})$ the elements $[a]$ and $[b]$ satisfy $[a] \vee [b] \leq [a] \wedge [b]$ (which is of course equivalent to $([a] \vee [b] \Rightarrow [a] \wedge [b]) = 1$), and given any Heyting algebra A with two elements x and y satisfying $x \vee y \leq x \wedge y$, there is a unique Heyting homomorphism $h : \mathrm{HA}(\mathcal{L}) \to A$ with $h([a]) = x$ and $h([b]) = y$.

6.4 Equational definition

The following description of CCCs in terms of operations and equations on a category is sometimes useful. The proof is entirely routine and left to the reader.

Proposition 6.12. *A category* $\mathbf{C}$ *is a CCC iff it has the following structure:*

- *A distinguished object* 1*, and for each object* C *there is given an arrow*

$$!_C : C \to 1$$

such that for each arrow $f : C \to 1$*,*

$$f = !_C.$$

- *For each pair of objects* A, B*, there is given an object* $A \times B$ *and arrows,*

$$p_1 : A \times B \to A \quad \text{and} \quad p_2 : A \times B \to B$$

and for each pair of arrows $f : Z \to A$ *and* $g : Z \to B$*, there is given an arrow,*

$$\langle f, g \rangle : Z \to A \times B$$

such that:

$$\begin{aligned} p_1\langle f, g\rangle &= f \\ p_2\langle f, g\rangle &= g \\ \langle p_1 h, p_2 h\rangle &= h \quad \text{for all } h : Z \to A \times B. \end{aligned}$$

- *For each pair of objects* A, B*, there is given an object* B^A *and an arrow,*

$$\epsilon : B^A \times A \to B$$

and for each arrow $f : Z \times A \to B$ *there is given an arrow*

$$\tilde{f} : Z \to B^A$$

such that:

$$\epsilon \circ (\tilde{f} \times 1_A) = f$$

and

$$\widetilde{(\epsilon \circ (g \times 1_A))} = g$$

for all $g : Z \to B^A$*, where,*

$$g \times 1_A = \langle g p_1, p_2 \rangle : Z \times A \to B^A \times A.$$

It is sometimes easier to check these equational conditions than to verify the corresponding UMPs. The next section provides an example of this sort.

6.5 λ-calculus

We have seen that the notions of a cartesian closed poset with finite joins (i.e. a Heyting algebra) and intuitionistic propositional calculus are essentially the same,

$$\text{HA} \sim \text{IPC}.$$

These are two different ways of describing one and the same structure; whereby, to be sure, the logical description contains some superfluous data in the choice of a particular presentation, not required by the algebraic description.

We now want to consider another, very similar, correspondence between logic and categories, involving more general CCC's. Indeed, the foregoing correspondence was the poset case of the following general one between CCCs and λ-calculus,

$$\text{CCC} \ \sim \ \lambda\text{-calculus}.$$

These notions are also essentially equivalent, in a sense that we will now sketch (a more detailed treatment can be found in the book by Lambek and Scott). They are two different ways of representing the same idea, namely that of a collection of objects and functions, with operations of *pairing*, *projection*, *application*, and *transposition* (or "*currying*").

First, recall the notion of a (typed) λ-calculus from Chapter 2. It consists of:

- Types: $A \times B,\ A \to B, \ldots$ (and some basic types)
- Terms: $x, y, z, \ldots : A$ (variables for each type A)
  $a : A,\ b : B, \ldots$ (possibly some typed constants)

$$\begin{aligned}
\langle a, b\rangle &: A \times B && (a : A,\ b : B)\\
\text{fst}(c) &: A && (c : A \times B)\\
\text{snd}(c) &: B && (c : A \times B)\\
ca &: B && (c : A \to B,\ a : A)\\
\lambda x.b &: A \to B && (x : A,\ b : B)
\end{aligned}$$

- Equations:

$$\begin{aligned}
\text{fst}(\langle a, b\rangle) &= a\\
\text{snd}(\langle a, b\rangle) &= b\\
\langle \text{fst}(c), \text{snd}(c)\rangle &= c\\
(\lambda x.b)a &= b[a/x]\\
\lambda x.cx &= c \quad (\text{no } x \text{ in } c)
\end{aligned}$$

Given a particular such λ-calculus $\mathcal{L}$, the associated *category of types* $\mathbf{C}(\mathcal{L})$ was then defined as follows:

- objects: the types,
- arrows $A \to B$: equivalence classes of closed terms $[c] : A \to B$, identified according to,

$$[a] = [b] \quad \text{iff } \mathcal{L} \vdash a = b \tag{6.5}$$

- identities: $1_A = \lambda x.x$ (where $x : A$),
- composition: $c \circ b = \lambda x.c(bx)$.

We have already seen that this is a well-defined category, and that it has binary products. It is a simple matter to add a terminal object. Now let us use the equational characterization of CCCs to show that it is cartesian closed. Given any objects A, B, we set $B^A = A \to B$, and as the evaluation arrow we take,

$$\epsilon = \lambda z.\mathrm{fst}(z)\mathrm{snd}(z) : B^A \times A \to B \quad (z : Z).$$

Then for any arrow $f : Z \times A \to B$, we take as the transpose,

$$\tilde{f} = \lambda z \lambda x.f\langle z, x\rangle : Z \to B^A \quad (z : Z,\ x : A).$$

It is now a straightforward λ-calculus calculation to verify the two required equations, namely,

$$\epsilon \circ (\tilde{f} \times 1_A) = f,$$
$$\widetilde{(\epsilon \circ (g \times 1_A))} = g.$$

In detail, for the first one recall that

$$\alpha \times \beta = \lambda w.\langle \alpha\mathrm{fst}(w), \beta\mathrm{snd}(w)\rangle.$$

So we have

$$\begin{aligned}
\epsilon \circ (\tilde{f} \times 1_A) &= (\lambda z.\mathrm{fst}(z)\mathrm{snd}(z)) \circ [(\lambda y \lambda x.f\langle y, x\rangle) \times \lambda u.u] \\
&= \lambda v.(\lambda z.\mathrm{fst}(z)\mathrm{snd}(z))[(\lambda y \lambda x.f\langle y, x\rangle) \times \lambda u.u]v \\
&= \lambda v.(\lambda z.\mathrm{fst}(z)\mathrm{snd}(z))[\lambda w.\langle(\lambda y \lambda x.f\langle y, x\rangle)\mathrm{fst}(w), (\lambda u.u)\mathrm{snd}(w)\rangle]v \\
&= \lambda v.(\lambda z.\mathrm{fst}(z)\mathrm{snd}(z))[\lambda w.\langle(\lambda x.f\langle\mathrm{fst}(w), x\rangle), \mathrm{snd}(w)\rangle]v \\
&= \lambda v.(\lambda z.\mathrm{fst}(z)\mathrm{snd}(z))[\langle(\lambda x.f\langle\mathrm{fst}(v), x\rangle), \mathrm{snd}(v)\rangle] \\
&= \lambda v.(\lambda x.f\langle\mathrm{fst}(v), x\rangle)\mathrm{snd}(v) \\
&= \lambda v.f\langle\mathrm{fst}(v), \mathrm{snd}(v)\rangle \\
&= \lambda v.fv \\
&= f.
\end{aligned}$$

The second equation is proved similarly.

Let us call a set of basic types and terms, together with a set of equations between terms, a *theory* in the λ-calculus. Given such a theory $\mathcal{L}$, the cartesian closed category $\mathbf{C}(\mathcal{L})$ built from the λ-calculus over $\mathcal{L}$ is the CCC presented by the generators and relations stated by $\mathcal{L}$. Just as in the poset case of IPC and Heyting algebras, there is a logical completeness theorem that follows from this fact. To state it, we first require the notion of a model of a theory $\mathcal{L}$ in the λ-calculus in an arbitrary cartesian closed category $\mathbf{C}$. Roughly, a model is an assignment of the types and terms of $\mathcal{L}$ to objects and arrows of $\mathbf{C}$:

$$\begin{array}{rcl} X \text{ basic type} & \rightsquigarrow & [\![X]\!] \text{ object} \\ b : A \to B \text{ basic term} & \rightsquigarrow & [\![b]\!] : [\![A]\!] \to [\![B]\!] \text{ arrow} \end{array}$$

This assignment is then extended to all types and terms in such a way that the λ-calculus operations are taken to the corresponding CCC ones:

$$\begin{array}{c} [\![A \times B]\!] = [\![A]\!] \times [\![B]\!] \\ [\![\langle f, g \rangle]\!] = \langle [\![f]\!], [\![g]\!] \rangle \\ \text{etc.} \end{array}$$

Finally, it is required that all the equations of $\mathcal{L}$ are satisfied, in the sense that:

$$\mathcal{L} \vdash [a] = [b] : A \to B \quad \text{implies} \quad [\![a]\!] = [\![b]\!] : [\![A]\!] \to [\![B]\!] \tag{6.6}$$

This is what is sometimes called "denotational semantics" for the λ-calculus. It is essentially the conventional, set-theoretic semantics for first-order logic, but extended to higher types, restricted to equational theories, and generalized to CCCs.

For example, let $\mathcal{L}$ be the theory with one basic type X, two basic terms,

$$\begin{array}{c} u : X \\ m : X \times X \to X \end{array}$$

and the usual equations for associativity and units,

$$\begin{array}{c} m\langle u, x \rangle = x \\ m\langle x, u \rangle = x \\ m\langle x, m\langle y, z \rangle \rangle = m\langle m\langle x, y \rangle, z \rangle. \end{array}$$

Thus $\mathcal{L}$ is just the usual equational theory of monoids. Then a model of $\mathcal{L}$ in a cartesian closed category $\mathbf{C}$ is nothing but a monoid in $\mathbf{C}$, that is, an object $M = [\![X]\!]$ equipped with a distinguished point

$$[\![u]\!] : 1 \to M$$

and a binary operation

$$[\![m]\!] : M \times M \to M$$

satisfying the unit and associativity laws.

Note that by (6.5) and (6.6), there is a model of $\mathcal{L}$ in $\mathbf{C}(\mathcal{L})$ with the property that $[\![a]\!] = [\![b]\!] : X \to Y$ if and only if $a = b$ is provable in $\mathcal{L}$. In this way, one can prove the following *CCC completeness theorem for λ-calculus.*

Proposition 6.13. *For any theory $\mathcal{L}$ in the λ-calculus, and any terms a, b in $\mathcal{L}$,*

$$\mathcal{L} \vdash a = b \quad \textit{if for all models } M \textit{ in CCCs,} \quad [\![a]\!]_M = [\![b]\!]_M .$$

This proposition says that the λ-calculus is *deductively complete* for models in CCCs. It is worth emphasizing that this is not true if one restricts attention to models in the single category **Sets**; indeed there are many examples of theories in λ-calculus in which equations holding for all models in **Sets** are still not provable (see the exercises for an example).

The cartesian closed category $\mathbf{C}(\mathcal{L})$ has the following UMP, analogous to the one for any algebra presented by generators and relations. Given any model M of $\mathcal{L}$ in any cartesian closed category $\mathbf{C}$, there is a unique functor,

$$[\![-]\!]_M : \mathbf{C}(\mathcal{L}) \to \mathbf{C}$$

preserving the CCC structure, given by,

$$[\![X]\!]_M = M$$

for the basic type X, and similarly for the other basic types and terms of $\mathcal{L}$. In this precise sense, the theory $\mathcal{L}$ is a presentation of the cartesian closed category $\mathbf{C}(\mathcal{L})$ by generators and relations.

Finally, let us note that the notions of λ-calculus and CCC are essentially "equivalent," in the sense that any cartesian closed category $\mathbf{C}$ also gives rise to a λ-calculus $\mathcal{L}(\mathbf{C})$,

$$\mathcal{L} \rightsquigarrow \mathbf{C}(\mathcal{L})$$

and this construction is essentially inverse to the one just sketched.

Briefly, given $\mathbf{C}$, we define $\mathcal{L}(\mathbf{C})$ by:

- basic types: the objects of $\mathbf{C}$
- basic terms: $a : A \to B$ for each $a : A \to B$ in $\mathbf{C}$
- equations: many equations identifying the λ-calculus operations with the corresponding category and CCC structure on $\mathbf{C}$, for example:

$$\begin{aligned} \lambda x.\mathrm{fst}(x) &= p_1 \\ \lambda x.\mathrm{snd}(x) &= p_2 \\ \lambda y.f(x,y) &= \tilde{f}(x) \\ g(f(x)) &= (g \circ f)(x) \\ \lambda y.y &= 1_A \end{aligned}$$

This suffices to ensure that there is an isomorphism of categories,

$$\mathbf{C}(\mathcal{L}(\mathbf{C})) \cong \mathbf{C}.$$

Moreover, the theories $\mathcal{L}$ and $\mathcal{L}(\mathbf{C}(\mathcal{L}))$ will also be "equivalent" in a suitable sense, involving the kinds of considerations typical of comparing different presentations of algebras. We refer the reader to the book by Lambek and Scott (1986) for further details.

6.6 Exercises

1. Show that for any three objects A, B, C in a cartesian closed category, there are isomorphisms:
   (a) $(A \times B)^C \cong A^C \times B^C$
   (b) $(A^B)^C \cong A^{B \times C}$
2. Is the category of monoids cartesian closed?
3. Show that for any objects A, B in a cartesian closed category, there is a bijective correspondence between points of the exponential $1 \to B^A$ and arrows $A \to B$.
4. Show that the category of ωCPOs is cartesian closed, but that the category of *strict* ωCPOs is not (the strict ωCPOs are the ones with initial object $\perp$, and the continuous maps between them are supposed to preserve $\perp$).
5. In the λ-calculus, consider the theory (due to Dana Scott) of a *reflexive domain*: there is one basic type D, two constants s and r of types $s : (D \to D) \to D$ and $r : D \to (D \to D)$, and two equations,

$$srx = x \quad (x : D)$$
$$rsy = y \quad (y : D \to D).$$

   Prove that, up to isomorphism, this theory has only one model M in **Sets**, and that *every* equation holds in M.

This page intentionally left blank

7

FUNCTORS AND NATURALITY

We now want to start considering categories and functors more systematically, developing the "category theory" of category theory itself, rather than of other mathematical objects, like groups, or formulas in a logical system. Let me emphasize that, while some of this may look a bit like "abstract nonsense," the idea behind it is that when one has a particular application at hand, the theory can then be specialized to that concrete case. The notion of a functor is a case in point; developing its general theory makes it a clarifying, simplifying, and powerful tool in its many instances.

7.1 Category of categories

We begin by reviewing what we know about the category **Cat** of categories and functors and tying up some loose ends.

We have already seen that **Cat** has finite coproducts $\mathbf{0}$, $\mathbf{C} + \mathbf{D}$; and finite products $\mathbf{1}$, $\mathbf{C} \times \mathbf{D}$. It is very easy to see that there are also all small coproducts and products, constructed analogously. We can therefore show that **Cat** has all limits by constructing equalizers. Thus let categories $\mathbf{C}$ and $\mathbf{D}$ and parallel functors F and G be given, and define the category $\mathbf{E}$ and functor E,

$$\mathbf{E} \xrightarrow{E} \mathbf{C} \underset{G}{\overset{F}{\rightrightarrows}} \mathbf{D}$$

as follows (recall that for a category $\mathbf{C}$, we write $\mathbf{C}_0$ and $\mathbf{C}_1$ for the collections of objects and arrows respectively):

$$\mathbf{E}_0 = \{C \in \mathbf{C}_0 \mid F(C) = G(C)\}$$
$$\mathbf{E}_1 = \{f \in \mathbf{C}_1 \mid F(f) = G(f)\}$$

and let $E : \mathbf{E} \to \mathbf{C}$ be the evident inclusion. This is then an equalizer, as the reader can easily check.

The category $\mathbf{E}$ is an example of a *subcategory*, that is, a monomorphism in **Cat** (recall that equalizers are monic). Often, by a subcategory of a category $\mathbf{C}$ one means specifically a collection $\mathbf{U}$ of some of the objects and arrows, $\mathbf{U}_0 \subseteq \mathbf{C}_0$ and $\mathbf{U}_1 \subseteq \mathbf{C}_1$), that is closed under the operations $\mathrm{dom}, \mathrm{cod}, \mathrm{id}$, and $\circ$. There is

then an evident inclusion functor

$$i : \mathbf{U} \to \mathbf{C}$$

which is clearly monic.

In general, coequalizers of categories are more complicated to describe—indeed, even for posets, determining the coequalizer of a pair of monotone maps can be quite involved, as the reader should consider.

There are various properties of functors other than being monic and epic, which turn out to be quite useful in **Cat**. A few of these are given by the following:

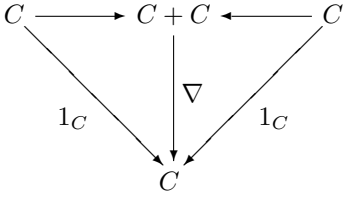
Definition 7.1. A functor $F : \mathbf{C} \to \mathbf{D}$ is said to be:

- *injective on objects* if the object part $F_0 : \mathbf{C}_0 \to \mathbf{D}_0$ is injective, it is *surjective on objects* if F_0 is surjective.
- Similarly F is *injective* (resp. *surjective*) *on arrows* if the arrow part $F_1 : \mathbf{C}_1 \to \mathbf{D}_1$ is injective (resp. surjective).
- F is *faithful* if for all $A, B \in \mathbf{C}_0$, the map

  $$F_{A,B} : \mathrm{Hom}_{\mathbf{C}}(A, B) \to \mathrm{Hom}_{\mathbf{D}}(FA, FB)$$

  defined by $f \mapsto F(f)$ is injective.
- Similarly, F is *full* if $F_{A,B}$ is always surjective.

What is the difference between being faithful and being injective on arrows? Consider for example the "codiagonal functor" $\nabla : \mathbf{C} + \mathbf{C} \to \mathbf{C}$, as indicated in the following:



∇ is faithful, but not injective on arrows.

A *full subcategory*

$$\mathbf{U} \rightarrowtail \mathbf{C}$$

consists of some objects of $\mathbf{C}$ and *all* of the arrows between them (thus satisfying the closure conditions for a subcategory). For example, the inclusion functor $\mathbf{Sets}_{fin} \rightarrowtail \mathbf{Sets}$ is full and faithful, but the forgetful functor $\mathbf{Groups} \rightarrowtail \mathbf{Sets}$ is faithful but not full.

Example 7.2. There is another "forgetful" functor for groups, namely to the category **Cat** of categories,

$$G : \mathbf{Groups} \to \mathbf{Cat}.$$

Observe that this functor is full and faithful, since a functor between groups $F : G(A) \to G(B)$ is exactly the same thing as a group homomorphism.

And exactly the same situation holds for monoids.

For posets, too, there is a full and faithful, forgetful functor

$$P : \mathbf{Pos} \to \mathbf{Cat}$$

again because a functor between posets $F : P(A) \to P(B)$ is exactly a monotone map. And the same thing holds for the "discrete category" functor $S : \mathbf{Sets} \to \mathbf{Cat}$.

Thus **Cat** provides a setting for *comparing structures of many different kinds*. For instance, one can have a functor $R : G \to \mathbf{C}$ from a group G to a category $\mathbf{C}$ that is not a group. If $\mathbf{C}$ is a poset, then any such functor must be trivial (why?). But if $\mathbf{C}$ is, say, the category of finite dimensional, real vector spaces and linear maps, then a functor R is exactly a *linear representation* of the group G, representing every element of G as an invertible matrix of real numbers and the group multiplication as matrix multiplication.

What is a functor $g : P \to G$ from a poset to a group? Since G has only one object $*$, it will have $g(p) = * = g(q)$ for all $p, q \in P$. For each $p \leq q$, it picks an element $g_{p,q}$ in such a way that

$$\begin{aligned} g_{p,p} &= u \qquad \text{(the unit of } G\text{)} \\ g_{q,r} \cdot g_{p,q} &= g_{p,r}. \end{aligned}$$

For example, take $P = (\mathbb{R}, \leq)$ to be the ordered real numbers and $G = (\mathbb{R}, +)$ the additive group of reals, then *subtraction* is a functor,

$$g : (\mathbb{R}, \leq) \to (\mathbb{R}, +)$$

defined by,

$$g_{x,y} = (y - x).$$

Indeed, we have,

$$\begin{aligned} g_{x,x} &= (x - x) = 0 \\ g_{y,z} \cdot g_{x,y} &= (z - y) + (y - x) = (z - x) = g_{x,z}. \end{aligned}$$

7.2 Representable structure

Let $\mathbf{C}$ be a *locally small category*, so that we have the *representable functors*,

$$\mathrm{Hom}_{\mathbf{C}}(C, -) : \mathbf{C} \to \mathbf{Sets}$$

for all objects $C \in \mathbf{C}$. This function is evidently faithful if the object C has the property that for any objects X and Y and arrows $f, g : X \rightrightarrows Y$, if $f \neq g$ there is an arrow $x : C \to X$ such that $fx \neq gx$. That is, the arrows in the category

are distinguished by their effect on generalized elements based at C. Such an object C is called a *generator* for $\mathbf{C}$.

In the category of sets, for example, the terminal object 1 is a generator. In groups, as we have already discussed, the free group $F(1)$ on one element is a generator. Indeed, the functor represented by $F(1)$ is isomorphic to the forgetful functor $U : \mathbf{Grp} \to \mathbf{Sets}$,

$$\mathrm{Hom}(F(1), G) \cong U(G). \tag{7.1}$$

This isomorphism not only holds for each group G, but also respects group homomorphisms, in the sense that for any such $h : G \to H$ there is a commutative square,

$$\begin{array}{ccc} G & \mathrm{Hom}(F(1), G) & \xrightarrow{\cong} U(G) \\ \downarrow{\scriptstyle h} & \downarrow{\scriptstyle h_*} & \downarrow{\scriptstyle U(h)} \\ H & \mathrm{Hom}(F(1), H) & \xrightarrow[\cong]{} U(H) \end{array}$$

One says that the isomorphism (7.1) is "natural in G." In a certain sense, this "explains" why the forgetful functor U preserves all limits, since representable functors necessarily do. The related fact that the forgetful functor is faithful is a precise way to capture the vague idea, which we initially used for motivation, that the category of groups is "concrete."

Recall that there are also *contravariant* representable functors

$$\mathrm{Hom}_{\mathbf{C}}(-, C) : \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$$

taking $f : A \to B$ to $f^* : \mathrm{Hom}_{\mathbf{C}}(B, C) \to \mathrm{Hom}_{\mathbf{C}}(A, C)$ by $f^*(h) = h \circ f$ for $h : B \to C$.

Example 7.3. Given a group G in a (locally small) category $\mathbf{C}$, the contravariant representable functor $\mathrm{Hom}_{\mathbf{C}}(-, G)$ actually has a group structure, giving a functor

$$\mathrm{Hom}_{\mathbf{C}}(-, G) : \mathbf{C}^{\mathrm{op}} \to \mathbf{Grp}.$$

In **Sets**, for example, for each set X we can define the operations on the group $\mathrm{Hom}(X, G)$ pointwise,

$$\begin{aligned} u(x) &= u \quad \text{(the unit of } G\text{)} \\ (f \cdot g)(x) &= f(x) \cdot g(x) \\ f^{-1}(x) &= f(x)^{-1}. \end{aligned}$$

In this case, we have an isomorphism

$$\mathrm{Hom}(X, G) \cong \Pi_{x \in X} G$$

with the product group. Functoriality in X is given simply by precomposition; thus, for any function $h : Y \to X$, one has

$$\begin{aligned} h^*(f \cdot g)(y) &= (f \cdot g)(h(y)) \\ &= f(h(y)) \cdot g(h(y)) \\ &= h^*(f)(y) \cdot h^*(g)(y) \\ &= (h^*(f) \cdot h^*(g))(y) \end{aligned}$$

and similarly for inverses and the unit. Indeed, it is easy to see that this construction works just as well for any other algebraic structure defined by operations and equations.

For instance, in topology one has the ring $\mathbb{R}$ of real numbers and, for any space X, the ring

$$\mathcal{C}(X) = \mathrm{Hom}_{\mathrm{Top}}(X, \mathbb{R})$$

of real-valued, continuous functions on X. Just as in the previous case, if

$$h : Y \to X$$

is any continuous function, we then get a ring homomorphism

$$h^* : \mathcal{C}(X) \to \mathcal{C}(Y)$$

by precomposing with h. The recognition of $\mathcal{C}(X)$ as representable ensures that this "ring of real-valued functions" construction is functorial,

$$\mathcal{C} : \mathbf{Top}^{\mathrm{op}} \to \mathbf{Rings}.$$

In passing from $\mathbb{R}$ to $\mathrm{Hom}_{\mathbf{Top}}(X, \mathbb{R})$, all the algebraic structure of $\mathbb{R}$ is retained, but properties determined by conditions that are not strictly equational are not necessarily preserved. For instance, $\mathbb{R}$ is not only a ring, but also a *field*, meaning that every nonzero real number r has a multiplicative inverse r^{-1}; formally,

$$\forall x (x = 0 \vee \exists y.\ y \cdot x = 1).$$

To see that this condition fails in for example $\mathcal{C}(\mathbb{R})$, consider the continuous function $f(x) = x^2$. For any argument $y \neq 0$, the multiplicative inverse must be $g(y) = 1/y^2$. But if this function were to be continuous, at 0 it would have to be $\lim_{y \to 0} 1/y^2$ which does not exist in $\mathbb{R}$.

Example 7.4. A very similar situation occurs in the category **BA** of Boolean algebras. Given the Boolean algebra **2** with the usual (truth-table) operations $\wedge, \vee, \neg, 0, 1$, for any set X, we make the set

$$\mathrm{Hom}_{\mathbf{Sets}}(X, \mathbf{2})$$

into a Boolean algebra with the pointwise operations:

$$0(x) = 0$$
$$1(x) = 1$$
$$(f \wedge g)(x) = f(x) \wedge g(x)$$
$$\text{etc.}$$

When we define the operations in this way in terms of those on $\mathbf{2}$ we see immediately that $\mathrm{Hom}(X, \mathbf{2})$ is a Boolean algebra too, and that precomposition is a contravariant functor,

$$\mathrm{Hom}(-, \mathbf{2}) : \mathbf{Sets}^{\mathrm{op}} \to \mathbf{BA}$$

into the category $\mathbf{BA}$ of Boolean algebras and their homomorphisms.

Now observe that for any set X, the familiar isomorphism

$$\mathrm{Hom}(X, \mathbf{2}) \cong \mathcal{P}(X)$$

between characteristic functions $\phi : X \to \mathbf{2}$ and subsets $V_\phi = \phi^{-1}(1) \subseteq X$, relates the pointwise Boolean operations in $\mathrm{Hom}(X, \mathbf{2})$ to the subset operations of intersection, union, etc. in $\mathcal{P}(X)$:

$$V_{\phi \wedge \psi} = V_\phi \cap V_\psi$$
$$V_{\phi \vee \psi} = V_\phi \cup V_\psi$$
$$V_{\neg \phi} = X - V_\phi$$
$$V_1 = X$$
$$V_0 = \emptyset$$

In this sense, the set-theoretic Boolean operations on $\mathcal{P}(X)$ are induced by those on $\mathbf{2}$, and the powerset $\mathcal{P}$ is seen to be a contravariant functor to the category of Boolean algebras,

$$\mathcal{P}^{\mathrm{BA}} : \mathbf{Sets}^{\mathrm{op}} \to \mathrm{BA}.$$

As was the case for the covariant representable functor $\mathrm{Hom}_{\mathbf{Grp}}(F(1), -)$ and the forgetful functor U from groups to sets, here the contravariant functors $\mathrm{Hom}_{\mathbf{Sets}}(-, \mathbf{2})$ and $\mathcal{P}^{\mathrm{BA}}$ from sets to Boolean algebras can also be seen to be *naturally* isomorphic, in the sense that for any function $f : Y \to X$, the following square of Boolean algebras and homomorphisms commutes:

$$\begin{array}{ccc} X & \mathrm{Hom}(X, \mathbf{2}) & \xrightarrow{\cong} P(X) \\ \uparrow f & \downarrow f^* & \downarrow f^{-1} \\ Y & \mathrm{Hom}(Y, \mathbf{2}) & \xrightarrow[\cong]{} P(Y) \end{array}$$

7.3 Stone duality

Before considering the topic of naturality more systematically, let us take a closer look at the foregoing example of powersets and Boolean algebras.

Recall that an *ultrafilter* in a Boolean algebra B is a proper subset $U \subset B$ such that:

- $1 \in U$
- $x, y \in U$ implies $x \wedge y \in U$
- $x \in U$ and $x \leq y$ implies $y \in U$
- if $U \subset U'$ and U' is a filter, then $U' = B$

The maximality condition on U is equivalent to the condition that for every $x \in B$, either $x \in U$ or $\neg x \in U$ but not both (exercise!).

We already know that there is an isomorphism between the set $\mathrm{Ult}(B)$ of ultrafilters on B and the Boolean homomorphisms $B \to \mathbf{2}$,

$$\mathrm{Ult}(B) \cong \mathrm{Hom}_{\mathbf{BA}}(B, \mathbf{2}).$$

This assignment $\mathrm{Ult}(B)$ is functorial and contravariant, and the displayed isomorphism above is natural in B. Indeed, given a Boolean homomorphism $h : B \to B'$, let

$$\mathrm{Ult}(h) = h^{-1} : \mathrm{Ult}(B') \to \mathrm{Ult}(B).$$

Of course, we have to show that the inverse image $h^{-1}(U) \subset B$ of an ultrafilter $U \subset B'$ is an ultrafilter in B. But since we know that $U = \chi_U^{-1}(1)$ for some $\chi_U : B' \to 2$, we have

$$\begin{aligned} \mathrm{Ult}(h)(U) &= h^{-1}(\chi_U^{-1}(1)) \\ &= (\chi_U \circ h)^{-1}(1). \end{aligned}$$

So $\mathrm{Ult}(h)(U)$ is also an ultrafilter. Thus, we have a contravariant functor of ultrafilters

$$\mathrm{Ult} : \mathbf{BA}^{\mathrm{op}} \to \mathbf{Sets}$$

as well as the contravariant powerset functor coming back

$$\mathcal{P}^{\mathbf{BA}} : \mathbf{Sets}^{\mathrm{op}} \to \mathbf{BA}.$$

The constructions,

$$\mathbf{BA}^{\mathrm{op}} \underset{\mathrm{Ult}}{\overset{(\mathcal{P}^{\mathbf{BA}})^{\mathrm{op}}}{\rightleftarrows}} \mathbf{Sets}$$

are *not* mutually inverse, however. For in general, $\mathrm{Ult}(\mathcal{P}(X))$ is much larger than X, since there are many ultrafilters in $\mathcal{P}(X)$ that are not "principal," that is, of

the form $\{U \subseteq X \mid x \in U\}$ for some $x \in X$. (But what if X is finite?) Instead, there is a more subtle relation between these functors which we consider in more detail later; namely, these are an example of adjoint functors.

For now, consider the following observations. Let

$$\mathcal{U} = \mathrm{Ult} \circ (\mathcal{P}^{\mathbf{BA}})^{\mathrm{op}} : \mathbf{Sets} \to \mathbf{BA}^{\mathrm{op}} \to \mathbf{Sets}$$

so that,

$$\mathcal{U}(X) = \{U \subseteq \mathcal{P}(X) \mid U \text{ is an ultrafilter}\}$$

is a *covariant* functor on **Sets**. Now observe that for any set X there is a function,

$$\eta : X \to \mathcal{U}(X)$$

taking each element $x \in X$ to the principal ultrafilter

$$\eta(x) = \{U \subseteq X \mid x \in U\}.$$

This map is "natural" in X, that is, for any function $f : X \to Y$, the following diagram commutes.

$$\begin{array}{ccc} X & \xrightarrow{\eta_X} & \mathcal{U}(X) \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle \mathcal{U}(f)} \\ Y & \xrightarrow[\eta_Y]{} & \mathcal{U}(Y) \end{array}$$

This is so because, for any ultrafilter $\mathcal{V}$ in $\mathcal{P}(X)$,

$$\mathcal{U}(f)(\mathcal{V}) = \{U \subseteq Y \mid f^{-1}(U) \in \mathcal{V}\}.$$

So in the case of the principal ultrafilters $\eta(x)$, we have

$$\begin{aligned} (\mathcal{U}(f) \circ \eta_X)(x) &= \mathcal{U}(f)(\eta_X(x)) \\ &= \{V \subseteq Y \mid f^{-1}(V) \in \eta_X(x)\} \\ &= \{V \subseteq Y \mid x \in f^{-1}(V)\} \\ &= \{V \subseteq Y \mid fx \in V\} \\ &= \eta_Y(fx) \\ &= (\eta_Y \circ f)(x). \end{aligned}$$

Finally, observe that there is an analogous natural map at the "other side" of this situation, in the category of Boolean algebras. Specifically, for every Boolean algebra B there is a homomorphism similar to the function η,

$$\phi_B : B \to \mathcal{P}(\mathrm{Ult}(B))$$

given by

$$\phi_B(b) = \{\mathcal{V} \in \mathrm{Ult}(B) \mid b \in \mathcal{V}\}.$$

It is not hard to see that ϕ_B is always injective. For, given any distinct elements $b, b' \in B$, the Boolean prime ideal theorem implies that there is an ultrafilter $\mathcal{V}$ such that $b \in \mathcal{V}$ but not $b' \in \mathcal{V}$. The Boolean algebra $\mathcal{P}(\mathrm{Ult}(B))$, together with the homomorphism ϕ_B, is called the *Stone representation* of B. It presents the arbitrary Boolean algebra B as an algebra of subsets. For the record, we thus have the following very special case of the far-reaching *Stone Duality Theorem*:

Proposition 7.5. *Every Boolean algebra B is isomorphic to one consisting of subsets of some set X, equipped with the set-theoretical Boolean operations.*

7.4 Naturality

A natural transformation is a morphism of functors. That's right: for fixed categories $\mathbf{C}$ and $\mathbf{D}$ we can regard the functors $\mathbf{C} \to \mathbf{D}$ as the *objects* of a new category, and the arrows between these objects are what we are going to call natural transformations. They are to be thought of as different ways of "relating" functors to each other, in a sense that we will now explain.

Let us begin by considering a certain kind of situation that often arises: we have some "construction" on a category $\mathbf{C}$ and some other "construction," and we observe that these two "constructions" are related to each other in a way that is independent of the specific objects and arrows involved. That is, the relation is really between the constructions themselves. To give a simple example, suppose $\mathbf{C}$ has products and consider, for objects $A, B, C \in \mathbf{C}$,

$$(A \times B) \times C \quad \text{and} \quad A \times (B \times C).$$

Regardless of what objects A, B, and C are, we have an isomorphism

$$h : (A \times B) \times C \xrightarrow{\sim} A \times (B \times C).$$

What does it mean that this isomorphism does not really depend on the particular objects A, B, C? One way to explain it is this:

Given any $f : A \to A'$, we get a commutative square

$$\begin{array}{ccc} (A \times B) \times C & \xrightarrow{h_A} & A \times (B \times C) \\ \downarrow & & \downarrow \\ (A' \times B) \times C & \xrightarrow[h_{A'}]{} & A' \times (B \times C) \end{array}$$

So what we really have is an isomorphism between the "constructions"

$$(- \times B) \times C \quad \text{and} \quad - \times (B \times C)$$

without regard to what is in the argument-place of these.

Now, by a "construction" we of course just mean a functor, and by a "relation between constructors" we mean a *morphism of functors* (which is what we are about to define). In the example, it is an isomorphism

$$(- \times B) \times C \;\cong\; - \times (B \times C)$$

of functors $\mathbf{C} \to \mathbf{C}$. In fact, we can of course consider the functors of three arguments

$$F = (-_1 \times -_2) \times -_3 : \mathbf{C}^3 \to \mathbf{C}$$

and

$$G = -_1 \times (-_2 \times -_3) : \mathbf{C}^3 \to \mathbf{C}$$

and there is an analogous isomorphism

$$F \cong G.$$

But an *isomorphism* is a special morphism, so let us define the general notion first.

Definition 7.6. For categories $\mathbf{C}, \mathbf{D}$ and functors

$$F, G : \mathbf{C} \to \mathbf{D}$$

a *natural transformation* $\vartheta : F \to G$ is a family of arrows in $\mathbf{D}$

$$(\vartheta_C : FC \to GC)_{C \in \mathbf{C}_0}$$

such that, for any $f : C \to C'$ in $\mathbf{C}$, one has $\vartheta_{C'} \circ F(f) = G(f) \circ \vartheta_C$, that is, the following commutes:

$$\begin{array}{ccc} FC & \xrightarrow{\vartheta_C} & GC \\ {\scriptstyle Ff}\downarrow & & \downarrow{\scriptstyle Gf} \\ FC' & \xrightarrow[\vartheta_{C'}]{} & GC' \end{array}$$

Given such a natural transformation $\vartheta : F \to G$, the $\mathbf{D}$-arrow $\vartheta_C : FC \to GC$ is called the *component of* ϑ at C.

If you think of a functor $F : \mathbf{C} \to \mathbf{D}$ as "picture" of $\mathbf{C}$ in $\mathbf{D}$, then you can think of a natural transformation $\vartheta_C : FC \to GC$ as a "cylinder" with such a picture at each end.

7.5 Examples of natural transformations

We have already seen several examples of natural transformations in previous sections, namely the isomorphisms

$$\begin{aligned} \mathrm{Hom}_{\mathbf{Grp}}(F(1), G) &\cong U(G) \\ \mathrm{Hom}_{\mathbf{Sets}}(X, \mathbf{2}) &\cong \mathcal{P}(X) \\ \mathrm{Hom}_{\mathbf{BA}}(B, \mathbf{2}) &\cong \mathrm{Ult}(X). \end{aligned}$$

There were also the maps from Stone duality,

$$\begin{aligned} \eta_X &: X \to \mathrm{Ult}(\mathcal{P}(X)) \\ \phi_B &: B \to \mathcal{P}(\mathrm{Ult}(B)). \end{aligned}$$

We now consider some further examples.

Example 7.7. Consider the free monoid $M(X)$ on a set X and define a natural transformation $\eta : 1_{\mathrm{Sets}} \to UM$, such that each component $\eta_X : X \to UM(X)$ is given by the "insertion of generators" taking every element x to itself, considered as a word.

$$\begin{array}{ccc} X & \xrightarrow{\eta_X} & UM(X) \\ {\scriptstyle f}\downarrow & & \vdots\,{\scriptstyle UM(f)} \\ Y & \xrightarrow[\eta_Y]{} & UM(Y) \end{array}$$

This is natural, because the homomorphism $M(f)$ on the free monoid $M(X)$ is completely determined by what f does to the generators.

Example 7.8. Let $\mathbf{C}$ be a category with products, and $A \in \mathbf{C}$ fixed. A natural transformation from the functor $A \times - : \mathbf{C} \to \mathbf{C}$ to $1_{\mathbf{C}} : \mathbf{C} \to \mathbf{C}$ is given by taking the component at C to be the second projection

$$\pi_2 : A \times C \to C.$$

From this, together with the pairing operation $\langle -, - \rangle$, one can build up the isomorphism,

$$h : (A \times B) \times C \xrightarrow{\sim} A \times (B \times C).$$

For another such example in more detail, consider the functors

$$\begin{aligned} \times &: \mathbf{C}^2 \to \mathbf{C} \\ \bar{\times} &: \mathbf{C}^2 \to \mathbf{C} \end{aligned}$$

where $\bar{\times}$ is defined on objects by

$$A \mathbin{\bar{\times}} B = B \times A$$

and on arrows by

$$\alpha \mathbin{\bar{\times}} \beta = \beta \times \alpha.$$

Define a "twist" natural transformation $t : \times \to \bar{\times}$ by

$$t_{(A,B)}\langle a, b\rangle = \langle b, a\rangle.$$

To check that the following commutes,

$$\begin{array}{ccc} A \times B & \xrightarrow{t_{(A,B)}} & B \times A \\ {\scriptstyle \alpha \times \beta}\downarrow & & \downarrow{\scriptstyle \beta \times \alpha} \\ A' \times B' & \xrightarrow[t_{(A',B')}]{} & B' \times A' \end{array}$$

observe that for any generalized elements $a : Z \to A$ and $b : Z \to B$,

$$\begin{aligned} (\beta \times \alpha) t_{(A,B)} \langle a, b\rangle &= (\beta \times \alpha)\langle b, a\rangle \\ &= \langle \beta b, \alpha a\rangle \\ &= t_{(A',B')}\langle \alpha a, \beta b\rangle \\ &= t_{(A',B')} \circ (\alpha \times \beta)\langle a, b\rangle. \end{aligned}$$

Thus $t : \times \to \bar{\times}$ is natural. In fact, each component $t_{(A,B)}$ is an isomorphism with inverse $t_{(B,A)}$. This is a simple case of an isomorphism of functors.

Definition 7.9. The *functor category* $\mathrm{Fun}(\mathbf{C}, \mathbf{D})$ has

objects: functors $F : \mathbf{C} \to \mathbf{D}$,

arrows: natural transformations $\vartheta : F \to G$.

For each object F, the natural transformation 1_F has components

$$(1_F)_C = 1_{FC} : FC \to FC$$

and the composite natural transformation of $F \xrightarrow{\vartheta} G \xrightarrow{\phi} H$ has components

$$(\phi \circ \vartheta)_C = \phi_C \circ \vartheta_C.$$

Definition 7.10. A *natural isomorphism* is a natural transformation

$$\vartheta : F \to G$$

which is an isomorphism in the functor category $\mathrm{Fun}(\mathbf{C}, \mathbf{D})$.

Lemma 7.11. *A natural transformation $\vartheta : F \to G$ is a natural isomorphism iff each component $\vartheta_C : FC \to GC$ is an isomorphism.*

Proof. Exercise! □

In our first example we can therefore say that the isomorphism

$$\vartheta_A : (A \times B) \times C \cong A \times (B \times C)$$

is *natural in* A, meaning that the functors

$$F(A) = (A \times B) \times C$$
$$G(A) = A \times (B \times C)$$

are naturally isomorphic.

Here is a classical example of a natural isomorphism.

Example 7.12. Consider the category

$$\mathrm{Vect}(\mathbb{R})$$

of real vector spaces and linear transformations

$$f : V \to W.$$

Every vector space V has a *dual space*

$$V^* = \mathrm{Vect}(V, \mathbb{R})$$

of linear transformations. And every linear transformation

$$f : V \to W$$

gives rise to a *dual linear transformation*

$$f^* : W^* \to V^*$$

defined by precomposition, $f^*(A) = A \circ f$ for $A : W \to \mathbb{R}$. In brief $(-)^* = \mathrm{Vect}(-, \mathbb{R}) : \mathbf{Vect}^{\mathrm{op}} \to \mathbf{Vect}$ is the *contravariant representable* functor endowed with vector space structure, just like the examples already considered in Section 7.2.

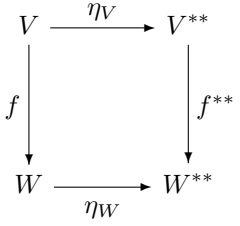
As in those examples, there is a canonical linear transformation from each vector space to its double dual,

$$\eta_V : \; V \to V^{**}$$
$$x \mapsto (\mathrm{ev}_x : V^* \to \mathbb{R})$$

where $\mathrm{ev}_x(A) = A(x)$ for every $A : V \to \mathbb{R}$. This map is the component of a natural transformation,

$$\eta : 1_{\mathbf{Vect}} \to **$$

since the following always commutes



in **Vect**. Indeed, given any $v \in V$ and $A : W \to \mathbb{R}$ in W^*, we have

$$\begin{aligned}(f^{**} \circ \eta_V)(v)(A) &= f^{**}(\mathrm{ev}_v)(A)\\ &= \mathrm{ev}_v(f^*(A))\\ &= \mathrm{ev}_v(A \circ f)\\ &= (A \circ f)(v)\\ &= A(fv)\\ &= \mathrm{ev}_{fv}(A)\\ &= (\eta_W \circ f)(v)(A).\end{aligned}$$

Now, it is a well-known fact in linear algebra that every *finite dimensional* vector space V is isomorphic to its dual space $V \cong V^*$ just for reasons of dimension. However, there is no "natural" way to choose such an isomorphism. On the other hand, the natural transformation,

$$\eta_V : V \to V^{**}$$

is a natural isomorphism when V is finite dimensional.

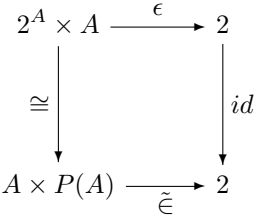
Thus, the formal notion of naturality captures the informal fact that $V \cong V^{**}$ "naturally," unlike $V \cong V^*$.

A similar situation occurs in **Sets**. Here we take 2 instead of $\mathbb{R}$, and the dual A^* of a set A then becomes

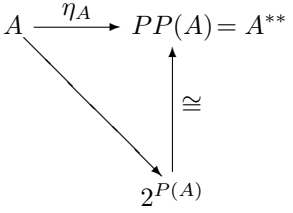
$$A^* = \mathcal{P}(A) \cong \mathbf{Sets}(A, 2)$$

while the dual of a map $f : A \to B$ is the inverse image $f^* : \mathcal{P}(B) \to \mathcal{P}(A)$.

Note that the exponential evaluation corresponds to (the characteristic function of) the membership relation on $A \times \mathcal{P}(A)$.



Transposing again gives a map



which is described by

$$\eta_A(a) = \{U \subseteq A \mid a \in U\}.$$

In **Sets**, one always has A strictly smaller than $\mathcal{P}(A)$, so $\eta_A : A \to A^{**}$ is *never* an isomorphism. Nonetheless, $\eta : 1_{\mathbf{Sets}} \to **$ is a natural transformation, which the reader should prove.

7.6 Exponentials of categories

We now want to show that the category **Cat** of (small) categories and functors is cartesian closed, by showing that any two categories $\mathbf{C}, \mathbf{D}$ have an exponential $\mathbf{D}^{\mathbf{C}}$. Of course, we will take $\mathbf{D}^{\mathbf{C}} = \mathrm{Fun}(\mathbf{C}, \mathbf{D})$, the category of functors and natural transformations, for which we need to prove the required universal mapping property (UMP).

Proposition 7.13. Cat *is cartesian closed, with the exponentials*

$$\mathbf{D}^{\mathbf{C}} = \mathrm{Fun}(\mathbf{C}, \mathbf{D}).$$

Before giving the proof, let us note the following. Since exponentials are unique up to isomorphism, this gives us a way to verify that we have found the "*right*" definition of a morphism of functors. For the notion of a natural transformation is completely determined by the requirement that it makes the set $\mathrm{Hom}(\mathbf{C}, \mathbf{D})$ into an exponential category. This is an example of how category theory can serve as a conceptual tool for discovering new concepts. Before giving the proof, we need the following.

Lemma 7.14 (Bifunctor lemma). *Given categories* $\mathbf{A}, \mathbf{B}$ *and* $\mathbf{C}$*, a map of arrows and objects,*

$$F_0 : \mathbf{A}_0 \times \mathbf{B}_0 \to \mathbf{C}_0$$
$$F_1 : \mathbf{A}_1 \times \mathbf{B}_1 \to \mathbf{C}_1$$

is a functor $F : \mathbf{A} \times \mathbf{B} \to \mathbf{C}$ *iff:*

1. *F is functorial in each argument: $F(A,-) : \mathbf{B} \to \mathbf{C}$ and $F(-,B) : \mathbf{A} \to \mathbf{C}$ are functors for all $A \in \mathbf{A}_0$ and $B \in \mathbf{B}_0$.*
2. *F satisfies the following "interchange law:" Given $\alpha : A \to A' \in \mathbf{A}$ and $\beta : B \to B' \in \mathbf{B}$, the following commutes.*

$$\begin{array}{ccc} F(A,B) & \xrightarrow{F(A,\beta)} & F(A,B') \\ {\scriptstyle F(\alpha,B)}\downarrow & & \downarrow{\scriptstyle F(\alpha,B')} \\ F(A',B) & \xrightarrow[F(A',\beta)]{} & F(A',B') \end{array}$$

that is, $F(A',\beta) \circ F(\alpha,B) = F(\alpha,B') \circ F(A,\beta)$ in $\mathbf{C}$.

Proof. (of Lemma) In $\mathbf{A} \times \mathbf{B}$ any arrow

$$\langle \alpha, \beta \rangle : \langle A, B \rangle \to \langle A', B' \rangle$$

factors as

$$\begin{array}{ccc} \langle A,B\rangle & \xrightarrow{\langle 1_A,\beta\rangle} & \langle A,B'\rangle \\ {\scriptstyle \langle \alpha,1_B\rangle}\downarrow & & \downarrow{\scriptstyle \langle \alpha,1_{B'}\rangle} \\ \langle A',B\rangle & \xrightarrow[\langle 1_{A'},\beta\rangle]{} & \langle A',B'\rangle \end{array}$$

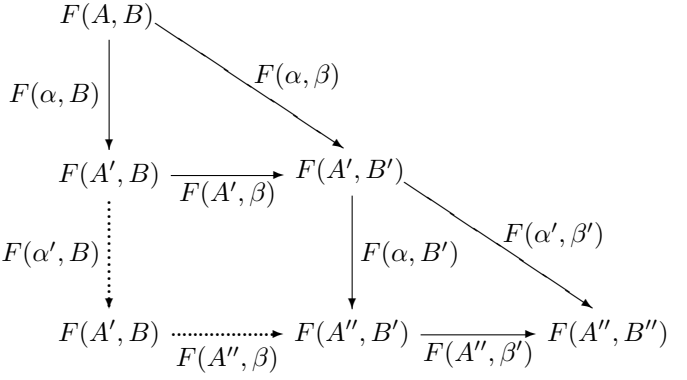
So (1) and (2) are clearly necessary. To show that they are also sufficient, we can define the (proposed) functor:

$$\begin{aligned} F(\langle A,B\rangle) &= F(A,B) \\ F(\langle \alpha,\beta\rangle) &= F(A',\beta) \circ F(\alpha,B) \end{aligned}$$

The interchange law, together with functoriality in each argument, then ensures that

$$F(\alpha',\beta') \circ F(\alpha,\beta) = F(\langle \alpha',\beta'\rangle \circ \langle \alpha,\beta\rangle)$$

as can be read off from the following diagram.



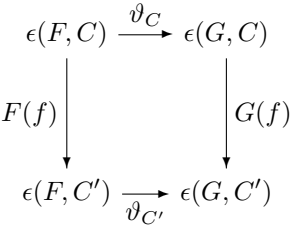
□

Proof. (of Proposition) We need to show:

1. $\epsilon = \text{eval} : \text{Fun}(\mathbf{C}, \mathbf{D}) \times \mathbf{C} \to \mathbf{D}$ is functorial.
2. For any category $\mathbf{X}$ and functor
$$F : \mathbf{X} \times \mathbf{C} \to \mathbf{D}$$
there is a functor
$$\tilde{F} : \mathbf{X} \to \text{Fun}(\mathbf{C}, \mathbf{D})$$
such that $\epsilon \circ (\tilde{F} \times 1_{\mathbf{C}}) = F$.
3. Given any functor
$$G : \mathbf{X} \to \text{Fun}(\mathbf{C}, \mathbf{D})$$
one has $(\epsilon \circ \widetilde{(G \times 1_{\mathbf{C}})}) = G$.

(1) Using the bifunctor lemma, we show that ϵ is functorial.
Fix $F : \mathbf{C} \to \mathbf{D}$ and consider $\epsilon(F, -) = F : \mathbf{C} \to \mathbf{D}$. This is clearly functorial!
Next, fix $C \in \mathbf{C}_0$ and consider $\epsilon(-, C) : \text{Fun}(\mathbf{C}, \mathbf{D}) \to \mathbf{D}$ defined by
$$(\vartheta : F \to G) \mapsto (\vartheta_C : FC \to GC).$$
This is also clearly functorial.

For the interchange law, consider any $\vartheta : F \to G \in \text{Fun}(\mathbf{C}, \mathbf{D})$ and $(f : C \to C') \in \mathbf{C}$, then we need the following to commute:



But this holds because $\epsilon(F, C) = F(C)$ and ϑ is a natural transformation.

The conditions (2) and (3) are now routine. For example, for (2), given

$$F : \mathbf{X} \times \mathbf{C} \to \mathbf{D}$$

let

$$\tilde{F} : \mathbf{X} \to \mathrm{Fun}(\mathbf{C}, \mathbf{D})$$

be defined by

$$\tilde{F}(X)(C) = F(X, C).$$

□

7.7 Functor categories

Let us consider some particular functor categories.

Example 7.15. First, clearly $\mathbf{C^1} = \mathbf{C}$ for the terminal category $\mathbf{1}$. Next, what about $\mathbf{C^2}$, where $\mathbf{2} = \cdot \to \cdot$ is the single arrow category? This is just the arrow category of $\mathbf{C}$ that we already know,

$$\mathbf{C^2} = \mathbf{C}^{\to}.$$

Consider instead the discrete category, $2 = \{0, 1\}$. Then clearly,

$$\mathbf{C}^2 \cong \mathbf{C} \times \mathbf{C}.$$

Similarly for any set I (regarded as a discrete category) we have

$$\mathbf{C}^I \cong \prod_{i \in I} \mathbf{C}.$$

Example 7.16. "Transcendental deduction of natural transformations"
Given the *possibility* of functor categories $\mathbf{D^C}$, we can determine what the objects and arrows therein must be as follows:

objects: these correspond uniquely to functors of the form

$$\mathbf{1} \to \mathbf{D^C}$$

and hence to functors

$$\mathbf{C} \to \mathbf{D}$$

arrows: by the foregoing example, these correspond uniquely to functors of the form

$$\mathbf{1} \to (\mathbf{D^C})^{\mathbf{2}}$$

thus to functors of the form

$$\mathbf{2} \to \mathbf{D}^{\mathbf{C}}$$

and hence to functors

$$\mathbf{C} \times \mathbf{2} \to \mathbf{D}$$

respectively

$$\mathbf{C} \to \mathbf{D}^{\mathbf{2}}.$$

But a functor from $\mathbf{C}$ into the arrow category $\mathbf{D}^{\mathbf{2}}$ (respectively a functor into $\mathbf{D}$ from the cylinder category $\mathbf{C} \times \mathbf{2}$) is exactly a natural transformation between two functors from $\mathbf{C}$ into $\mathbf{D}$, as the reader can see by drawing a picture of the functor's image in $\mathbf{D}$.

Example 7.17. Recall that a (directed) graph can be regarded as a pair of sets and a pair of functions,

$$G_1 \underset{s}{\overset{t}{\rightrightarrows}} G_0$$

where G_1 is the set of edges, G_0 is the set of vertices, and s and t are the source and target operations.

A homomorphism of graphs $h : G \to H$ is a map that preserves sources and targets. In detail, this is a pair of functions $h_1 : G_1 \to H_1$ and $h_0 : G_0 \to H_0$ such that for all edges $e \in G$, we have $sh_1(e) = h_0 s(e)$ and similarly for t as well. But this amounts exactly to saying that the following two diagrams commute.

$$\begin{array}{ccc} G_1 & \xrightarrow{h_1} & H_1 \\ {\scriptstyle s_G}\downarrow & & \downarrow{\scriptstyle s_H} \\ G_0 & \xrightarrow[h_0]{} & H_0 \end{array} \qquad \begin{array}{ccc} G_1 & \xrightarrow{h_1} & H_1 \\ {\scriptstyle t_G}\downarrow & & \downarrow{\scriptstyle t_H} \\ G_0 & \xrightarrow[h_0]{} & H_0 \end{array}$$

Now consider the category Γ, pictured as follows:

$$\cdot \rightrightarrows \cdot$$

It has exactly two objects and two distinct, parallel, non-identity arrows. A graph G is then exactly a functor,

$$G : \Gamma \to \mathbf{Sets}$$

and a homomorphism of graphs $h : G \to H$ is exactly a natural transformation between these functors. Thus, the category of graphs is a functor category,

$$\mathbf{Graphs} = \mathbf{Sets}^{\Gamma}.$$

As we see later, it follows from this fact that **Graphs** is cartesian closed.

Example 7.18. Given a product $\mathbf{C} \times \mathbf{D}$ of categories, take the first product projection

$$\mathbf{C} \times \mathbf{D} \to \mathbf{C}$$

and transpose it to get a functor

$$\Delta : \mathbf{C} \to \mathbf{C}^{\mathbf{D}}.$$

For $C \in \mathbf{C}$, the functor $\Delta(C)$ is the "constant C-valued functor,"

- $\Delta(C)(X) = C$ for all $X \in \mathbf{D}_0$
- $\Delta(x) = 1_C$ for all $x \in \mathbf{D}_1$.

Moreover, $\Delta(f) : \Delta(C) \to \Delta(C')$ is the natural transformation, each component of which is f.

Now suppose we have any functor

$$F : \mathbf{D} \to \mathbf{C}$$

and a natural transformation

$$\vartheta : \Delta(C) \to F.$$

Then the components of ϑ all look like

$$\vartheta_D : C \to F(D)$$

since $\Delta(C)(D) = C$. Moreover, for any $d : D \to D'$ in $\mathbf{D}$, the usual naturality square becomes a triangle, since $\Delta(C)(d) = 1_C$ for all $d : D \to D'$.

$$\begin{array}{ccc} C & \xrightarrow{\vartheta_D} & FD \\ {\scriptstyle 1_C}\downarrow & & \downarrow{\scriptstyle Fd} \\ C & \xrightarrow[\vartheta_{D'}]{} & FD' \end{array}$$

Thus, such a natural transformation $\vartheta : \Delta(C) \to F$ is exactly a cone to the base F (with vertex C). Similarly, a map of cones $\vartheta \to \varphi$ is a constant natural transformation, that is, one of the form $\Delta(h)$ for some $h : C \to D$, making a commutative triangle

$$\begin{array}{ccccc} \Delta(C) & & \xrightarrow{\Delta(h)} & & \Delta(D) \\ & {\scriptstyle \vartheta}\searrow & & \swarrow{\scriptstyle \varphi} & \\ & & F & & \end{array}$$

Example 7.19. Take posets P, Q and consider the functor category,

$$Q^P.$$

The functors $Q \to P$, as we know, are just monotone maps, but what is a natural transformation?

$$\vartheta : f \to g$$

For each $p \in P$ we must have

$$\vartheta_p : fp \leq gp$$

and if $p \leq q$, then there must be a commutative square involving $fp \leq fq$ and $gp \leq gq$, which, however, is automatic. Thus, the only condition is that $fp \leq gp$ for all p, that is, $f \leq g$ *pointwise*. Since this is just the usual ordering of the poset Q^P, the exponential poset agrees with the functor category. Thus, we have the following:

Proposition 7.20. *The inclusion functor,*

$$\mathbf{Pos} \to \mathbf{Cat}$$

preserves CCC structure.

Example 7.21. What happens if we take the functor *category* of two groups G and H?

$$H^G$$

Do we get an exponential of groups? Let us first ask, what is a natural transformation between two group homomorphisms $f, g : G \to H$? Such a map $\vartheta : f \to g$ would be an element $h \in H$ such that for every $x \in G$, we have

$$f(x) \cdot h = h \cdot g(x)$$

or, equivalently,

$$h^{-1} \cdot f(x) \cdot h = g(x).$$

Therefore, a natural transformation $\vartheta : f \to g$ is an *inner automorphism* $y \mapsto h^{-1} \cdot y \cdot h$ of H (called *conjugation by* h), that takes f to g. Clearly, every such arrow $\vartheta : f \to g$ has an inverse $\vartheta^{-1} : g \to f$ (conjugation by h^{-1}). But H^G is still not usually a group, simply because there may be many different homomorphisms $G \to H$, so the functor category H^G has more than one object.

This suggests enlarging the category of groups to include also categories with more than one object, but still having inverses for all arrows. Such categories are called *groupoids*, and have been studied by topologists (they occur as the paths between different points in a topological space). A groupoid can thus be regarded as a generalized group, in which the domains and codomains of elements x and y must match up, as in any category, for the multiplication $x \cdot y$ to be defined.

It is clear that if G and H are any groupoids, then the functor category H^G is also a groupoid. Thus we have the following proposition, the detailed proof of which is left as an exercise.

Proposition 7.22. *The category* **Grpd** *of groupoids is cartesian closed and the inclusion functor*

$$\mathbf{Grpd} \to \mathbf{Cat}$$

preserves the CCC structure.

7.8 Equivalence of categories

Before examining some particular functor categories in more detail, we consider one very special application of the idea of natural isomorphism. Consider first the following situation.

Example 7.23. Let $\mathbf{Ord}_{\text{fin}}$ be the category of finite ordinal numbers. Thus, the *objects* are the sets $0, 1, 2, \ldots$, where $0 = \emptyset$ and $n = \{0, \ldots, n-1\}$, while the *arrows* are *all* functions between these sets. Now suppose that for each finite set A we select an ordinal $|A|$ that is its cardinal and an isomorphism,

$$A \cong |A|.$$

Then for each function $f : A \to B$ of finite sets we have a function $|f|$ by completing the square

$$\begin{array}{ccc} A & \xrightarrow{\cong} & |A| \\ {\scriptstyle f}\downarrow & & \vdots\,{\scriptstyle |f|} \\ B & \xrightarrow[\cong]{} & |B| \end{array} \tag{7.2}$$

This clearly gives us a functor

$$|-| : \mathbf{Sets}_{\text{fin}} \to \mathbf{Ord}_{\text{fin}}.$$

Actually, all the maps in the above square are in $\mathbf{Sets}_{\text{fin}}$; so we should also make the inclusion functor

$$i : \mathbf{Ord}_{\text{fin}} \to \mathbf{Sets}_{\text{fin}}$$

explicit. Then we have the selected isos,

$$\vartheta_A : A \xrightarrow{\sim} i|A|$$

and we know by (7.2) that,

$$i(|f|) \circ \vartheta_A = \vartheta_B \circ f.$$

This of course says that we have a natural isomorphism

$$\vartheta : 1_{\mathbf{Sets}_{\text{fin}}} \to i \circ |-|$$

between two functors of the form

$$\mathbf{Sets}_{\text{fin}} \to \mathbf{Sets}_{\text{fin}}.$$

On the other hand, if we take an ordinal and take its ordinal, we get nothing new,

$$|i(-)| = 1_{\mathbf{Ord}_{\text{fin}}} : \mathbf{Ord}_{\text{fin}} \to \mathbf{Ord}_{\text{fin}}.$$

This is so because, for any cardinal n,

$$|i(n)| = n$$

and we can assume that we take $\vartheta_n = 1_n : n \to |i(n)|$, so that also,

$$|i(f)| = f : n \to m.$$

In sum, then, we have a situation where two categories are very similar; but they are *not* the same and they are *not even isomorphic* (why?). This kind of correspondence is what is captured by the notion of equivalence of categories.

Definition 7.24. An *equivalence of categories* consists of a pair of functors

$$E : \mathbf{C} \to \mathbf{D}$$
$$F : \mathbf{D} \to \mathbf{C}$$

and a pair of natural isomorphisms

$$\alpha : 1_{\mathbf{C}} \xrightarrow{\sim} F \circ E \qquad \text{in } \mathbf{C}^{\mathbf{C}}$$
$$\beta : 1_{\mathbf{D}} \xrightarrow{\sim} E \circ F \qquad \text{in } \mathbf{D}^{\mathbf{D}}.$$

In this situation, the functor F is called a *pseudo-inverse* of E. The categories $\mathbf{C}$ and $\mathbf{D}$ are then said to be *equivalent*, written $\mathbf{C} \simeq \mathbf{D}$.

Observe that equivalence of categories is a generalization of isomorphism. Indeed, two categories $\mathbf{C}, \mathbf{D}$ are isomorphic if there are functors.

$$E : \mathbf{C} \to \mathbf{D}$$
$$F : \mathbf{D} \to \mathbf{C}$$

such that

$$1_{\mathbf{C}} = GF$$
$$1_{\mathbf{D}} = FG.$$

In the case of equivalence $\mathbf{C} \simeq \mathbf{D}$, we replace the identity natural transformations by natural isomorphisms.

Experience has shown that the mathematically significant properties of objects are those that are invariant under isomorphisms, and in category theory, identity of objects is a much less important relation than isomorphism. So it is really equivalence of categories that is the more important notion of "similarity" for categories. One can think of equivalence of categories as "isomorphism up to isomorphism."

In the foregoing example $\mathbf{Sets}_{\text{fin}} \simeq \mathbf{Ord}_{\text{fin}}$, we see that every set is isomorphic to an ordinal, and the maps between ordinals are just the maps between them *as sets*. Thus we have:

1. for every set A, there is some ordinal n with $A \cong i(n)$,
2. for any ordinals n, m, there is an isomorphism

$$\text{Hom}_{\mathbf{Ord}_{\text{fin}}}(n, m) \cong \text{Hom}_{\mathbf{Sets}_{\text{fin}}}(i(n), i(m))$$

where $i : \mathbf{Ord}_{\text{fin}} \to \mathbf{Sets}_{\text{fin}}$ is the inclusion functor.

In fact, these conditions are characteristic of equivalences, as the following proposition shows.

Proposition 7.25. *The following conditions on a functor $F : \mathbf{C} \to \mathbf{D}$ are equivalent:*

1. *F is (part of) an equivalence of categories.*
2. *F is full and faithful and "essentially surjective" on objects: for every $D \in \mathbf{D}$ there's some $C \in \mathbf{C}$ such that $FC \cong D$.*

Proof. (1 implies 2) Take $E : \mathbf{D} \to \mathbf{C}$, and

$$\alpha : 1_{\mathbf{C}} \xrightarrow{\sim} EF$$

$$\beta : 1_{\mathbf{D}} \xrightarrow{\sim} FE.$$

In $\mathbf{C}$, for any C, we then have $\alpha_C : C \xrightarrow{\sim} EF(C)$, and

$$\begin{array}{ccc} C & \xrightarrow{\alpha_C} & EF(C) \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle EF(f)} \\ C' & \xrightarrow[\alpha_{C'}]{} & EF(C') \end{array}$$

commutes for any $f : C \to C'$.

Thus, if $F(f) = F(f')$, then $EF(f) = EF(f')$, so $f = f'$. So F is faithful. Note that by symmetry, E is also faithful.

Now take any arrow

$$h : F(C) \to F(C') \text{ in } \mathbf{D},$$

and consider

$$\begin{array}{ccc} C & \xrightarrow{\cong} & EF(C) \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle E(h)} \\ C' & \xrightarrow[\cong]{} & EF(C') \end{array}$$

where $f = (\alpha_{C'})^{-1} \circ E(h) \circ \alpha_C$. Then we have also $F(f) : F(C) \to F(C')$ and

$$EF(f) = E(h) : EF(C) \to EF(C')$$

by the square

$$\begin{array}{ccc} C & \xrightarrow{\alpha_C} & EF(C) \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle EF(f)} \\ C' & \xrightarrow[\alpha_{C'}]{} & EF(C') \end{array}$$

Since E is faithful, $F(f) = h$. So F is also *full*.

Finally, for any object $D \in \mathbf{D}$, we have

$$\beta : 1_{\mathbf{D}} \xrightarrow{\sim} FE$$

so

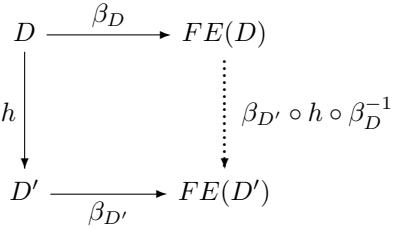
$$\beta_D : D \cong F(ED), \quad \text{for } ED \in \mathbf{C}_0.$$

(2 implies 1) We need to define $E : \mathbf{D} \to \mathbf{C}$ and natural transformations,

$$\alpha : 1_{\mathbf{C}} \xrightarrow{\sim} EF$$
$$\beta : 1_{\mathbf{D}} \xrightarrow{\sim} FE.$$

Since F is essentially surjective, for each $D \in \mathbf{D}_0$ we can pick some $E(D) \in \mathbf{C}_0$ such that there's some $\beta_D : D \xrightarrow{\sim} FE(D)$. That gives E on objects and the proposed components of $\beta : 1_{\mathbf{D}} \to FE$.

Given $h : D \to D'$ in $\mathbf{D}$, consider



Since $F : \mathbf{C} \to \mathbf{D}$ is full and faithful, we can find a unique arrow

$$E(h) : E(D) \to E(D')$$

with $FE(h) = \beta_{D'} \circ h \circ \beta_D^{-1}$. It is easy to see that then $E : \mathbf{D} \to \mathbf{C}$ is a functor and $\beta : 1_{\mathbf{D}} \xrightarrow{\sim} FE$ is clearly a natural isomorphism.

To find $\alpha : 1_{\mathbf{C}} \to EF$, apply F to C and consider $\beta_{FC} : F(C) \to FEF(C)$. Since F is full and faithful, the preimage of β_{FC} is an isomorphism,

$$\alpha_C = F^{-1}(\beta_{FC}) : C \xrightarrow{\sim} EF(C)$$

which is easily seen to be natural, since β is. □

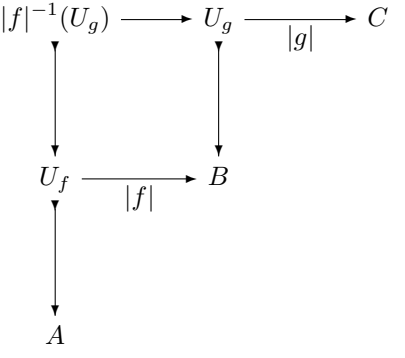
7.9 Examples of equivalence

Example 7.26. *Pointed sets and partial maps*

Let **Par** be the category of sets and partial functions. An arrow

$$f : A \rightharpoonup B$$

is a function $|f| : U_f \to B$ for some $U_f \subseteq A$. Identities in **Par** are the same as those in **Sets**, that is, 1_A is the *total* identity function on A. The composite of $f : A \rightharpoonup B$ and $g : B \rightharpoonup C$ is given as follows: Let $U_{(g \circ f)} := f^{-1}(U_g) \subseteq A$, and $|g \circ f| : U_{(g \circ f)} \to C$ is the horizontal composite indicated in the following diagram, in which the square is a pullback.



It is easy to see that composition is associative and that the identities are units.

The category of *pointed sets*,

$$\mathbf{Sets}_*$$

has as objects, sets A equipped with a distinguished "point" $a \in A$, that is, pairs,

$$(A, a) \text{ with } a \in A.$$

Arrows are functions that preserve the point, that is, an arrow $f : (A, a) \to (B, b)$ is a function $f : A \to B$ such that $f(a) = b$.

Now we show:

Proposition 7.27. $\mathbf{Par} \cong \mathbf{Sets}_*$

The functors establishing the equivalence are as follows:

$$F : \mathbf{Par} \to \mathbf{Sets}_*$$

and is defined on an object A by $F(A) = (A \cup \{*\}, *)$ where $*$ is a new point that we add to A. We also write $A_* = A \cup \{*\}$. For arrows, given $f : A \rightharpoonup B$, $F(f) : A_* \to B_*$ is defined by

$$f_*(x) = \begin{cases} f(x) & \text{if } x \in U_f \\ * & \text{otherwise.} \end{cases}$$

Then clearly $f_*(*_A) = *_B$, so in fact $f_* : A_* \to B_*$ as required.

Coming back, the functor

$$G : \mathbf{Sets}_* \to \mathbf{Par}$$

is defined on an object (A, a) by $G(A, a) = A - \{a\}$ and for an arrow $f : (A, a) \to (B, b)$

$$G(f) : A - \{a\} \rightharpoonup B - \{b\}$$

is the function with domain

$$U_{G(f)} = A - f^{-1}(b)$$

defined by $G(f)(x) = f(x)$ for every $f(x) \neq b$.

Now $F \circ G$ is the identity on $\mathbf{Par}$, because we are just adding a new point and then throwing it away. But $G \circ F$ is only naturally isomorphic to $1_{\mathbf{Sets}_*}$, since we have

$$(A, a) \cong ((A - \{a\}) \cup \{*\}, *).$$

These sets are not equal, since $a \neq *$. It still needs to be checked, of course, that F and G are functorial, and that the comparison $(A, a) \cong ((A - \{a\}) \cup \{*\}, *)$ is natural, but we leave these easy verifications to the reader.

Example 7.28. *Slice categories and indexed families*

For any set I, the functor category $\mathbf{Sets}^I$ is the category of I-indexed sets. The objects are I-indexed families of sets

$$(A_i)_{i \in I}$$

and the arrows are I-indexed families of functions,

$$(f_i : A_i \to B_i)_{i \in I} : (A_i)_{i \in I} \longrightarrow (B_i)_{i \in I}.$$

This category has an equivalent description that is often quite useful: it is equivalent to the slice category of **Sets** over I, consisting of arrows $\alpha : A \to I$ and "commutative triangles" over I (see Section 1.6),

$$\mathbf{Sets}^I \;\simeq\; \mathbf{Sets}/I.$$

Indeed, define functors

$$\Phi : \mathbf{Sets}^I \;\longrightarrow\; \mathbf{Sets}/I$$

$$\Psi : \mathbf{Sets}/I \;\longrightarrow\; \mathbf{Sets}^I$$

on objects as follows:

$$\Phi((A_i)_{i \in I}) = \pi : \sum_{i \in I} A_i \to I \quad \text{(the indexing projection)}$$

$$\Psi(\alpha : A \to I) = (\alpha^{-1}\{i\})_{i \in I}$$

and analogously on arrows. We leave it as an exercise to show that these are mutually pseudo-inverse functors. (Why are they *not* inverses?)

The equivalent description of $\mathbf{Sets}^I$ as $\mathbf{Sets}/I$ has the advantage that the slice category $\mathcal{E}/E$ makes sense for any object E in a general category $\mathcal{E}$, where the functor category $\mathcal{E}^E$ usually does not (for instance, in topological spaces). This allows the intuition of an "E-indexed family of objects of $\mathcal{E}$" to be made precise. If $\mathcal{E}$ has pullbacks, reindexing along an arrow $f : E \to F$ in $\mathcal{E}$ is then represented by the pullback functor $f^* : \mathcal{E}/F \to \mathcal{E}/E$. This is also worked out in the exercises.

Example 7.29. *Stone duality*

A class of examples of equivalences of categories are given by what are called "dualities." Often, classical duality theorems are not of the form $\mathbf{C} \cong \mathbf{D}^{\mathrm{op}}$ (much less $\mathbf{C} = \mathbf{D}^{\mathrm{op}}$), but rather $\mathbf{C} \simeq \mathbf{D}^{\mathrm{op}}$, that is, $\mathbf{C}$ is equivalent to the opposite

(or "dual") category of **D**. This is because the duality is established by a construction which returns the original thing only up to isomorphism, not "on the nose." Here is a simple example, which is a very special case of the far-reaching *Stone-Duality* theorem.

Proposition 7.30. *The category of finite Boolean algebras is equivalent to the opposite of the category of finite sets,*

$$\mathbf{BA}_{\text{fin}} \simeq \mathbf{Sets}_{\text{fin}}^{\text{op}}.$$

Proof. The functors involved here are the contravariant powerset functor

$$\mathcal{P}^{\mathbf{BA}} : \mathbf{Sets}_{\text{fin}}^{\text{op}} \to \mathbf{BA}_{\text{fin}}$$

on one side (the powerset of a finite set is finite!). Going back, we will use the functor,

$$A : \mathbf{BA}_{\text{fin}}^{\text{op}} \to \mathbf{Sets}_{\text{fin}}$$

taking the set of *atoms* of a Boolean algebra,

$$A(\mathcal{B}) = \{a \in \mathcal{B} \mid 0 < a \text{ and } (b < a \Rightarrow b = 0)\}.$$

In the finite case, this is isomorphic to the ultrafilter functor that we have already studied (see Section 7.3).

Lemma 7.31. *For any finite Boolean algebra $\mathcal{B}$, there is an isomorphism between atoms a in $\mathcal{B}$ and ultrafilters $U \subseteq \mathcal{B}$, given by*

$$U \mapsto \bigwedge_{b \in U} b$$

and

$$a \mapsto \uparrow(a).$$

Proof. If a is an atom then $\uparrow(a)$ is an ultrafilter, since for any b either $a \wedge b = a$ and then $b \in \uparrow(a)$ or $a \wedge b = 0$ and so $\neg b \in \uparrow(a)$.

If $U \subseteq \mathcal{B}$ is an ultrafilter then $0 < \bigwedge_{b \in U} b$, because, since U is finite and closed under intersections, we must have $\bigwedge_{b \in U} b \in U$. If $0 \neq b_0 < \bigwedge_{b \in U} b$ then b_0 is not in U, so $\neg b_0 \in U$. But then $b_0 < \neg b_0$ and so $b_0 = 0$.

Plainly $\uparrow(\bigwedge_{b \in U} b) \subseteq U$ since $b \in U$ implies $\bigwedge_{b \in U} b \subseteq b$. Now let $\bigwedge_{b \in U} b \leq a$ for some a not in U. Then $\neg a \in U$ implies that also $\bigwedge_{b \in U} b \leq \neg a$, and so $\bigwedge_{b \in U} b \leq a \wedge \neg a = 0$, which is impossible. □

Since we know that the set of ultrafilters $\text{Ult}(\mathcal{B})$ is contravariantly functorial (it is represented by the Boolean algebra **2**, see Section 7.3), we therefore also have a contravariant functor of atoms $A \cong \text{Ult}$. The explicit description of this functor is this: if $h : \mathcal{B} \to \mathcal{B}'$ and $a' \in A(\mathcal{B}')$, then it follows from the lemma that there's

a *unique* atom $a \in \mathcal{B}$ such that $a' \leq h(b)$ iff $a \leq b$ for all $b \in \mathcal{B}$. To find this atom a, take the intersection over the ultrafilter $h^{-1}(\uparrow (a'))$. Thus we can set $A(h)(a') = a$, to get a function

$$A(h) : A(\mathcal{B}') \to A(\mathcal{B}).$$

Of course, we must still check that this is a pseudo-inverse for $\mathcal{P}^{\mathbf{BA}} : \mathbf{Sets}^{\mathrm{op}}_{\mathrm{fin}} \to \mathbf{BA}_{\mathrm{fin}}$. The required natural isomorphisms,

$$\alpha_X : X \to A(\mathcal{P}(X))$$
$$\beta_{\mathcal{B}} : \mathcal{B} \to \mathcal{P}(A(\mathcal{B}))$$

are explicitly described as follows.

The atoms in a finite powerset $\mathcal{P}(X)$ are just the singletons $\{x\}$ for $x \in X$, thus $\alpha_X(x) = \{x\}$ is clearly an isomorphism.

To define $\beta_{\mathcal{B}}$, let

$$\beta_{\mathcal{B}}(b) = \{a \in A(\mathcal{B}) \mid a \leq b\}.$$

To see that $\beta_{\mathcal{B}}$ is also iso, consider the proposed inverse,

$$(\beta_{\mathcal{B}})^{-1}(B) = \bigvee_{a \in B} a \qquad \text{for } B \subseteq A(\mathcal{B}).$$

The isomorphism then follows from the following lemma, the proof of which is routine.

Lemma 7.32. *For any finite Boolean algebra $\mathcal{B}$,*

1. $b = \bigvee\{a \in A(\mathcal{B}) \mid a \leq b\}$.
2. *If a is an atom and $a \leq b \vee b'$, then $a \leq b$ or $a \leq b'$.*

Of course, one must still check that α and β really are natural transformations. This is left to the reader. □

Finally, we remark that the duality

$$\mathbf{BA}_{\mathrm{fin}} \simeq \mathbf{Sets}^{\mathrm{op}}_{\mathrm{fin}}$$

extends to one between *all* sets on the one side and the complete, atomic Boolean algebras, on the other,

$$\mathbf{caBA} \simeq \mathbf{Sets}^{\mathrm{op}}.$$

where, a Boolean algebra $\mathcal{B}$ is *complete* if every subset $U \subseteq \mathcal{B}$ has a join $\bigvee U \in \mathcal{B}$ and a *complete homomorphism* preserves these joins and $\mathcal{B}$ is *atomic* if every nonzero element $0 \neq b \in \mathcal{B}$ has some $a \leq b$ with a an atom.

Moreover, this is just the discrete case of the full Stone Duality Theorem, which states an equivalence between the category of *all* Boolean algebras

and the opposite of a certain category of topological spaces, called "Stone spaces," and all continuous maps between them. See the book by Johnstone for details.

7.10 Exercises

1. Consider the (covariant) composite functor,

$$\mathcal{F} = \mathcal{P}^{\mathbf{BA}} \circ \mathrm{Ult}^{\mathrm{op}} : \mathbf{BA} \to \mathbf{Sets}^{\mathrm{op}} \to \mathbf{BA}$$

taking each Boolean algebra B to the powerset algebra of sets of ultrafilters in B. Note that

$$\mathcal{F}(B) \cong \mathrm{Hom}_{\mathbf{Sets}}(\mathrm{Hom}_{\mathbf{BA}}(B, \mathbf{2}), 2)$$

is a sort of "double-dual" Boolean algebra. There is always a homomorphism,

$$\phi_B : B \to \mathcal{F}(B)$$

given by $\phi_B(b) = \{\mathcal{V} \in \mathrm{Ult}(B) \mid b \in \mathcal{V}\}$. Show that for any Boolean homomorphism $h : A \to B$, the following square commutes.

$$\begin{array}{ccc} A & \xrightarrow{\phi_A} & \mathcal{F}(A) \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle \mathcal{F}(h)} \\ B & \xrightarrow[\phi_B]{} & \mathcal{F}(B) \end{array}$$

2. Consider the forgetful functors

$$\mathbf{Groups} \xrightarrow{U} \mathbf{Monoids} \xrightarrow{V} \mathbf{Sets}$$

Say whether each is faithful, full, injective on arrows, surjective on arrows, injective on objects, and surjective on objects.

3. Show that a natural transformation is a natural isomorphism just if each of its components is an isomorphism. Is the same true for monomorphisms?

4. Prove that for any *finite* Boolean algebra B, the Stone representation

$$\phi : B \to \mathcal{P}\mathrm{Ult}(B)$$

is an isomorphism of Boolean algebras. Note the similarity to the case of finite dimensional vector spaces.

5. Show that a functor category $\mathbf{D}^{\mathbf{C}}$ has binary products if $\mathbf{D}$ does (construct the product of two functors F and G "objectwise": $(F \times G)(C) = F(C) \times G(C)$).
6. Show that the map of sets

$$\begin{aligned} \eta_A : A &\longrightarrow PP(A) \\ a &\longmapsto \{U \subseteq A | a \in U\} \end{aligned}$$

is the component at A of a natural transformation $\eta : 1_{\mathbf{Sets}} \to PP$, where $P : \mathbf{Sets}^{\mathrm{op}} \to \mathbf{Sets}$ is the (contravariant) powerset functor.
7. Let $\mathbf{C}$ be a locally small category. Show that there is a functor

$$\mathrm{Hom} : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{Sets}$$

such that for each object C of $\mathbf{C}$,

$$\mathrm{Hom}(C, -) : \mathbf{C} \to \mathbf{Sets}$$

is the covariant representable functor and

$$\mathrm{Hom}(-, C) : \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$$

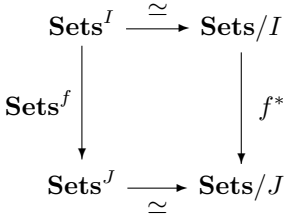
is the contravariant one. (*Hint: use the Bifunctor Lemma*)
8. What sorts of properties of categories do *not* respect equivalence? Find one that respects isomorphism, but not equivalence.
9. A category is *skeletal* if isomorphic objects are always identical. Show that every category is equivalent to a skeletal subcategory. (Every category has a "skeleton.")
10. Complete the proof that, for any set I, the category of I-indexed families of sets, regarded as the functor category $\mathbf{Sets}^I$, is equivalent to the slice category $\mathbf{Sets}/I$ of sets over I,

$$\mathbf{Sets}^I \simeq \mathbf{Sets}/I.$$

Show that reindexing of families along a function $f : J \to I$, given by precomposition,

$$\mathbf{Sets}^f((A_i)_{i \in I}) = (A_{f(j)})_{j \in J}$$

is represented by pullback, in the sense that the following diagram of categories and functors commutes up to natural isomorphism.



Here $f^* : \mathbf{Sets}/J \to \mathbf{Sets}/I$ is the pullback functor along $f : J \to I$.

This page intentionally left blank

8

CATEGORIES OF DIAGRAMS

In this chapter we will prove a very useful technical result called the Yoneda Lemma. It is probably the single most used result in category theory. Indeed, it is amazing how often it comes up, especially in view of the fact that it is a straightforward generalization of facts that we were able to show fairly easily about monoids and posets.

8.1 Set-valued functor categories

We are going to focus on special functor categories of the form,

$$\mathbf{Sets}^{\mathbf{C}}$$

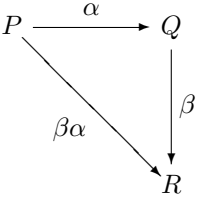
where the category $\mathbf{C}$ is locally small. Thus the objects are set-valued functors,

$$P, Q : \mathbf{C} \to \mathbf{Sets}$$

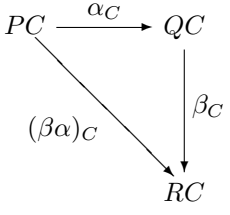
(sometimes called "diagrams on $\mathbf{C}$"), and the arrows are natural transformations

$$\alpha, \beta : P \to Q.$$

Remember that, for each object $C \in \mathbf{C}$, we can evaluate any commutative diagram in $\mathbf{Sets}^{\mathbf{C}}$,



at any object C to get a commutative diagram in **Sets**,



(8.1)

Thus, for each object C there is an evaluation functor

$$\mathrm{ev}_C : \mathbf{Sets}^{\mathbf{C}} \to \mathbf{Sets}.$$

Moreover, naturality means that if we have any arrow $f : D \to C$, we get a "cylinder" over the diagram (8.1) in **Sets**. One way of thinking about such functor categories that was already considered in Section 7.7 is suggested by considering the case where $\mathbf{C}$ is the category Γ pictured as:

$$1 \rightrightarrows 0$$

Then a set-valued functor $G : \Gamma \to \mathbf{Sets}$ is just a graph and a natural transformation $\alpha : G \to H$ is a graph homomorphism. Thus, for this case,

$$\mathbf{Sets}^{\Gamma} = \mathbf{Graphs}.$$

This suggests regarding an arbitrary category of the form $\mathbf{Sets}^{\mathbf{C}}$ as a generalized "category of structured sets" and their "homomorphisms"; indeed, this is a very useful way of thinking of such functors and their natural transformations.

8.2 The Yoneda embedding

Among the objects of $\mathbf{Sets}^{\mathbf{C}}$ are certain very special ones, namely the (covariant) representable functors,

$$\mathrm{Hom}_{\mathbf{C}}(C, -) : \mathbf{C} \to \mathbf{Sets}.$$

Observe that for each $h : C \to D$ in $\mathbf{C}$, we have a natural transformation

$$\mathrm{Hom}_{\mathbf{C}}(h, -) : \mathrm{Hom}_{\mathbf{C}}(D, -) \to \mathrm{Hom}_{\mathbf{C}}(C, -)$$

(note the direction!) where the component at X is defined by precomposition:

$$(f : D \to X) \mapsto (f \circ h : C \to X).$$

Thus, we have a *contravariant* functor

$$k : \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}^{\mathbf{C}}$$

defined by $k(C) = \mathrm{Hom}_{\mathbf{C}}(C, -)$. Of course, this functor k is just the exponential transpose of the bifunctor

$$\mathrm{Hom}_{\mathbf{C}} : \mathbf{C}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{Sets}$$

which was shown as an exercise to be functorial.

If we instead transpose $\mathrm{Hom}_{\mathbf{C}}$ with respect to its other argument, we get a *covariant* functor,

$$y : \mathbf{C} \to \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$$

from $\mathbf{C}$ to a category of *contravariant* set-valued functors, sometimes called "presheaves." (Or, what amounts to the same thing, we can put $\mathbf{D} = \mathbf{C}^{\mathrm{op}}$ and apply the previous considerations to $\mathbf{D}$ in place of $\mathbf{C}$.) More formally:

Definition 8.1. The *Yoneda embedding* is the functor $y : \mathbf{C} \to \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ taking $C \in \mathbf{C}$ to the contravariant representable functor,

$$yC = \mathrm{Hom}_{\mathbf{C}}(-, C) : \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$$

and taking $f : C \to D$ to the natural transformation,

$$yf = \mathrm{Hom}_{\mathbf{C}}(-, f) : \mathrm{Hom}_{\mathbf{C}}(-, C) \to \mathrm{Hom}_{\mathbf{C}}(-, D).$$

A functor $F : \mathbf{C} \to \mathbf{D}$ is called an *embedding* if it is full, faithful, and injective on objects. We will soon show that y really *is* an embedding; this is a corollary of the Yoneda Lemma.

One should thus think of the Yoneda embedding y as a "representation" of $\mathbf{C}$ in a category of set-valued functors and natural transformations on *some* index category. Compared to the Cayley representation considered in Section 1.5, this one has the virtue of being *full*: any map $\vartheta : yC \to yD$ in $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ comes from a unique map $h : C \to D$ in $\mathbf{C}$ as $yh = \vartheta$. Indeed, recall that the Cayley representation of a group G was an injective group homomorphism

$$G \rightarrowtail \mathrm{Aut}(|G|) \subseteq |G|^{|G|}$$

where each $g \in G$ is represented as an automorphism $\tilde{g}$ of the set $|G|$ of elements (i.e. a "permutation"), by letting it "act on the left,"

$$\widetilde{g}(x) = g \cdot x$$

and the group multiplication is represented by composition of permutations,

$$\widetilde{g \cdot h} = \tilde{g} \circ \tilde{h}.$$

We also showed a generalization of this representation to arbitrary categories. Thus for any monoid M, there is an analogous representation

$$M \rightarrowtail \mathrm{End}(|M|) \subseteq |M|^{|M|}$$

by left action, representing the elements of M as endomorphisms of $|M|$.

Similarly, any poset P can be represented as a poset of subsets and inclusions by considering the poset $\mathrm{Low}(P)$ of "lower sets" $A \subseteq P$, that is, subsets that are "closed down" in the sense that $a' \leq a \in A$ implies $a' \in A$, ordered by inclusion. Taking the "principal lower set"

$$\downarrow(p) = \{q \in P \mid q \leq p\}$$

of each element $p \in P$ determines a monotone injection

$$\downarrow : P \rightarrowtail \mathrm{Low}(P) \subseteq \mathcal{P}(|P|)$$

such that $p \leq q$ iff $\downarrow(p) \subseteq \downarrow(q)$.

The representation given by the Yoneda embedding is "better" than these in that it cuts down the arrows in the codomain category to just those in the image of the representation functor $y : \mathbf{C} \to \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ (since y is full). Indeed, there may be many automorphisms $\alpha : G \to G$ of a group G that are not left actions by an element, but if we require α to commute with all *right* actions $\alpha(x \cdot g) = \alpha(x) \cdot g$, then α must itself be a left action. This is what the Yoneda embedding does in general; it adds enough "structure" to the objects yA in the image of the representation that the only "homomorphisms" $\vartheta : yA \to yB$ between those objects are the representable ones $\vartheta = yh$ for some $h : A \to B$. In this sense, the Yoneda embedding y represents the objects and arrows of $\mathbf{C}$ as certain "structured sets" and (*all of*) their "homomorphisms."

8.3 The Yoneda Lemma

Lemma 8.2 (Yoneda). *Let* $\mathbf{C}$ *be locally small. For any object* $C \in \mathbf{C}$ *and functor* $F \in \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ *there is an isomorphism*

$$\mathrm{Hom}(yC, F) \cong FC$$

which, moreover, is natural in both F *and* C*.*

Here:

(1) the Hom is $\mathrm{Hom}_{\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}}$,

(2) naturality in F means that, given any $\vartheta : F \to G$, the following diagram commutes:

$$\begin{array}{ccc} \mathrm{Hom}(yC, F) & \xrightarrow{\cong} & FC \\ \downarrow{\scriptstyle \mathrm{Hom}(yC, \vartheta)} & & \downarrow{\scriptstyle \vartheta_C} \\ \mathrm{Hom}(yC, G) & \xrightarrow[\cong]{} & GC \end{array}$$

(3) naturality in C means that, given any $h : C \to D$, the following diagram commutes:

$$\begin{array}{ccc} \mathrm{Hom}(yC, F) & \xrightarrow{\cong} & FC \\ \uparrow{\scriptstyle \mathrm{Hom}(yh, F)} & & \uparrow{\scriptstyle Fh} \\ \mathrm{Hom}(yD, F) & \xrightarrow[\cong]{} & FD \end{array}$$

Proof. To define the desired isomorphism,

$$\eta_{C,F} : \mathrm{Hom}(yC, F) \xrightarrow{\cong} FC$$

take $\vartheta : yC \to F$ and let

$$\eta_{C,F}(\vartheta) = \vartheta_C(1_C)$$

which we shall also write as

$$x_\vartheta = \vartheta_C(1_C)$$

where $\vartheta_C : \mathbf{C}(C, C) \to FC$ and so $\vartheta_C(1_C) \in FC$.

Conversely, given any $a \in FC$ we define the natural transformation $\vartheta_a : yC \to F$ as follows. Given any C' we define the component

$$(\vartheta_a)_{C'} : \mathrm{Hom}(C', C) \to FC'$$

by setting

$$(\vartheta_a)_{C'}(h) = F(h)(a)$$

for $h : C' \to C$.

To show that ϑ_a is natural, take any $f : C'' \to C'$ and consider the diagram:

$$\begin{array}{ccc} \mathrm{Hom}(C'', C) & \xrightarrow{(\vartheta_a)_{C''}} & FC'' \\ {\scriptstyle \mathrm{Hom}(f,C)}\uparrow & & \uparrow{\scriptstyle F(f)} \\ \mathrm{Hom}(C', C) & \xrightarrow[(\vartheta_a)_{C'}]{} & FC' \end{array}$$

We then calculate, for any $h \in yC(C')$

$$\begin{aligned} (\vartheta_a)_{C''} \circ \mathrm{Hom}(f, C)(h) &= (\vartheta_a)_{C''}(h \circ f) \\ &= F(h \circ f)(a) \\ &= F(f) \circ F(h)(a) \\ &= F(f)(\vartheta_a)_{C'}(h). \end{aligned}$$

So ϑ_a is indeed natural.

Now to show that ϑ_a and x_ϑ are mutually inverse, let us calculate ϑ_{x_ϑ} for a given $\vartheta : yC \to F$. First, just from the definitions, we have that for any $h : C' \to C$,

$$(\vartheta_{(x_\vartheta)})_{C'}(h) = F(h)(\vartheta_C(1_C)).$$

But since ϑ is natural, the following commutes:

$$\begin{array}{ccc} yC(C) & \xrightarrow{\vartheta_C} & FC \\ {\scriptstyle yC(h)}\downarrow & & \downarrow{\scriptstyle Fh} \\ yC(C') & \xrightarrow[\vartheta_{C'}]{} & FC' \end{array}$$

So, continuing,

$$\begin{aligned} (\vartheta_{(x_\vartheta)})_{C'}(h) &= F(h)(\vartheta_C(1_C)) \\ &= \vartheta_{C'} \circ yC(h)(1_C) \\ &= \vartheta_{C'}(h). \end{aligned}$$

Therefore $\vartheta_{(x_\vartheta)} = \vartheta$.

Going the other way around, for any $a \in FC$ we have

$$\begin{aligned} x_{\vartheta_a} &= (\vartheta_a)_C(1_C) \\ &= F(1_C)(a) \\ &= 1_{FC}(a) \\ &= a. \end{aligned}$$

Thus, $\mathrm{Hom}(yC, F) \cong FC$, as required.

The naturality claims are also easy: given $\phi : F \to F'$, taking $\vartheta \in \mathrm{Hom}(yC, F)$, and chasing around the diagram

$$\begin{array}{ccc} (yC, F) & \xrightarrow{\eta_{C,F}} & FC \\ {\scriptstyle (yC,\phi)}\downarrow & & \downarrow{\scriptstyle \phi_C} \\ (yC, F') & \xrightarrow[\eta_{C,F'}]{} & F'C \end{array}$$

we get

$$\begin{aligned} \phi_C(x_\vartheta) &= \phi_C(\vartheta_C(1_C)) \\ &= (\phi\vartheta)_C(1_C) \\ &= x_{(\phi\vartheta)} \\ &= \eta_{C,F'}((yC, \phi)(\vartheta)). \end{aligned}$$

For naturality in C, take some $f : C' \to C$. We then have

$$\begin{aligned}
\eta_{C'}(yf)^*(\vartheta) &= \eta_{C'}(\vartheta \circ yf) \\
&= (\vartheta \circ yf)_{C'}(1_{C'}) \\
&= \vartheta_{C'} \circ (yf)_{C'}(1_{C'}) \\
&= \vartheta_{C'}(f \circ 1_{C'}) \\
&= \vartheta_{C'}(f) \\
&= \vartheta_{C'}(1_C \circ f) \\
&= \vartheta_{C'} \circ (yC)(f)(1_C) \\
&= F(f) \circ \vartheta_C(1_C) \\
&= F(f)\eta_C(\vartheta).
\end{aligned}$$

The penultimate equation is by the naturality square,

$$\begin{array}{ccc}
yC(C) & \xrightarrow{\vartheta_C} & F(C) \\
\Big\downarrow{\scriptstyle yC(f)} & & \Big\downarrow{\scriptstyle F(f)} \\
yC(C') & \xrightarrow[\vartheta_{C'}]{} & F(C')
\end{array}$$

Therefore, $\eta_{C'} \circ (yf)^* = F(f) \circ \eta_C$.
We leave the naturality in F as an exercise. □

The Yoneda Lemma is used to prove our first "theorem."

Theorem 8.3. *The Yoneda embedding $y : \mathbf{C} \to \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ is full and faithful.*

Proof. For any objects $C, D \in \mathbf{C}$ we have an isomorphism

$$\mathrm{Hom}_{\mathbf{C}}(C, D) = yD(C) \cong \mathrm{Hom}_{\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}}(yC, yD).$$

And this isomorphism is indeed induced by the functor y, since it takes $h : C \to D$ to the natural transformation $\vartheta_h : yC \to yD$ given by

$$\begin{aligned}
(\vartheta_h)_{C'}(f : C' \to C) &= yD(f)(h) \\
&= \mathrm{Hom}_{\mathbf{C}}(f, D)(h) \\
&= h \circ f \\
&= (yh)_{C'}(f).
\end{aligned}$$

So $\vartheta_h = y(h)$. □

Remark 8.4. Note the following:

- If $\mathbf{C}$ is small, then $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ is locally small, and so $\mathrm{Hom}(yC, P)$ in $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ is a set.
- If $\mathbf{C}$ is locally small, then $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ need not be locally small. In this case, the Yoneda Lemma tells us that $\mathrm{Hom}(yC, P)$ is always a set.
- If $\mathbf{C}$ is not locally small, then $y : \mathbf{C} \to \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ will not even be defined, so the Yoneda Lemma does not apply.

Finally, observe that the Yoneda embedding $y : \mathbf{C} \to \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ is also injective on objects. For, given objects A, B in $\mathbf{C}$, if $yA = yB$ then $1_C \in \mathrm{Hom}(C, C) = yC(C) = yD(C) = \mathrm{Hom}(C, D)$ implies $C = D$.

8.4 Applications of the Yoneda Lemma

One frequent sort of application of the Yoneda Lemma is of the following form: given objects A, B in a category $\mathbf{C}$, to show that $A \cong B$ it suffices to show that $yA \cong yB$ in $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$. This results from the foregoing theorem and the fact that, if $F : \mathbf{C} \to \mathbf{D}$ is any full and faithful functor, then $FA \cong FB$ clearly implies $A \cong B$. We record this as the following.

Corollary 8.5. *Given objects A and B in any locally small category $\mathbf{C}$,*

$$yA \cong yB \qquad \textit{implies} \qquad A \cong B.$$

A typical such case is this. In any cartesian closed category $\mathbf{C}$, there is an isomorphism,

$$(A^B)^C \cong A^{(B \times C)}.$$

Recall how involved it was to prove this directly, using the compound universal mapping property (UMP). Now, we just need to show that

$$y((A^B)^C) \;\cong\; y(A^{(B \times C)}).$$

To that end, take any object $X \in \mathbf{C}$, then we have isomorphisms

$$\begin{aligned} \mathrm{Hom}(X, (A^B)^C) &\cong \mathrm{Hom}(X \times C, A^B) \\ &\cong \mathrm{Hom}((X \times C) \times B, A) \\ &\cong \mathrm{Hom}(X \times (B \times C), A) \\ &\cong \mathrm{Hom}(X, A^{(B \times C)}). \end{aligned}$$

Of course, it must be checked that these isomorphisms are natural in X, but that's left as an easy exercise.

Let us do another example of this kind.

Proposition 8.6. *If the cartesian closed category $\mathbf{C}$ has coproducts, then $\mathbf{C}$ is "distributive," that is, there is a canonical isomorphism,*

$$(A \times B) + (A \times C) \cong A \times (B + C).$$

Proof. As in the previous proposition, we check that

$$\begin{aligned}\mathrm{Hom}(A \times (B+C), X) &\cong \mathrm{Hom}(B+C, X^A)\\ &\cong \mathrm{Hom}(B, X^A) \times \mathrm{Hom}(C, X^A)\\ &\cong \mathrm{Hom}(A \times B, X) \times \mathrm{Hom}(A \times C, X)\\ &\cong \mathrm{Hom}((A \times B) + (A \times C), X).\end{aligned}$$

Finally, one sees easily that these isos are all natural in X. □

We have already seen a simple logical example of the Yoneda Lemma: to show that in the propositional calculus one has $\varphi \dashv\vdash \psi$ for some formulas φ, ψ, it suffices to show that for any formula ϑ, one has $\vartheta \vdash \varphi$ iff $\vartheta \vdash \psi$.

More generally, given any objects A, B in a locally small category $\mathbf{C}$, to find an arrow $h : A \to B$ it suffices to give one $\vartheta : yA \to yB$ in $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$, for then there is a unique h with $\vartheta = yh$. Why should it be easier to give an arrow $yA \to yB$ than one $A \to B$? The key difference is that in general $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ has much more structure to work with than does $\mathbf{C}$; as we will see, it is complete, cocomplete, cartesian closed, and more. So one can use various "higher-order" tools, from limits to λ-calculus; and if the result is an arrow of the form $yA \to yB$, then it comes from a unique one $A \to B$—despite the fact that $\mathbf{C}$ may not admit the "higher-order" constructions. In that sense, the category $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ is like an extension of $\mathbf{C}$ by "ideal elements" that permit calculations which can not be done in $\mathbf{C}$. This is something like passing to the complex numbers to solve equations in the reals, or adding higher types to an elementary logical theory.

8.5 Limits in categories of diagrams

Recall that a category $\mathcal{E}$ is said to be *complete* if it has all small limits; that is, for any small category J and functor $F : J \to \mathcal{E}$ there is a limit $L = \varprojlim_{j \in J} Fj$ in $\mathcal{E}$ and a "cone" $\eta : \Delta L \to F$ in $\mathcal{E}^J$, universal among arrows from constant functors ΔE. Here the constant functor $\Delta : \mathcal{E} \to \mathcal{E}^J$ is the transposed projection $\mathcal{E} \times J \to \mathcal{E}$.

Proposition 8.7. *For any locally small category* $\mathbf{C}$*, the functor category* $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ *is complete. Moreover, for every object* $C \in C$ *the evaluation functor*

$$\mathrm{ev}_C : \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}} \to \mathbf{Sets}$$

preserves all limits.

Proof. Suppose we have J small and $F : J \to \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$. The limit of F, if it exists, is an object in $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$, hence is a functor,

$$(\varprojlim_{i \in J} F_i) : \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}.$$

By the Yoneda Lemma, if we had such a functor, then for each object $C \in \mathbf{C}$ we'd have a natural isomorphism,

$$(\varprojlim F_i)(C) \cong \mathrm{Hom}(yC, \varprojlim F_i).$$

But then it would be the case that,

$$\begin{aligned} \mathrm{Hom}(yC, \varprojlim F_i) &\cong \varprojlim \mathrm{Hom}(yC, F_i) && \text{in } \mathbf{Sets} \\ &\cong \varprojlim F_i(C) && \text{in } \mathbf{Sets}. \end{aligned}$$

where the first isomorphism is because representable functors preserve limits, and the second is Yoneda again. Thus we simply *define* the limit $\varprojlim_{i \in J} F_i$ to be

$$(\varprojlim_{i \in J} F_i)(C) = \varprojlim_{i \in J} (F_i C) \tag{8.2}$$

that is, the *pointwise limit* of the functors F_i. The reader can easily work out how $\varprojlim F_i$ acts on $\mathbf{C}$-arrows, and what the universal cone is. Direct verification shows that it is a limit in $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$.

Finally, the preservation of limits by evaluation functors is stated by (8.2). □

8.6 Colimits in categories of diagrams

The notion of *cocompleteness* is of course the dual of completeness: a category is cocomplete if it has all (small) colimits. Like the foregoing proposition about the completeness of $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$, its cocompleteness actually follows simply from the fact that **Sets** is cocomplete. We leave the proof of the following as an exercise.

Proposition 8.8. *Given any categories* $\mathbf{C}$ *and* $\mathbf{D}$*, if* $\mathbf{D}$ *is cocomplete, then so is the functor category* $\mathbf{D}^{\mathbf{C}}$*, and the colimits in* $\mathbf{D}^{\mathbf{C}}$ *are "computed pointwise," in the sense that for every* $C \in \mathbf{C}$*, the evaluation functor*

$$\mathrm{ev}_C : \mathbf{D}^{\mathbf{C}} \to \mathbf{D}$$

preserves colimits. Thus, for any small index category J *and functor* $A : J \to \mathbf{D}^{\mathbf{C}}$*, and for each* $C \in \mathbf{C}$ *there is a canonical isomorphism,*

$$(\varinjlim_{j \in J} A_j)(C) \;\cong\; \varinjlim_{j \in J} (A_j C).$$

Proof. Exercise. □

Corollary 8.9. *For any locally small* $\mathbf{C}$*, the functor category* $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ *is cocomplete, and colimits there are computed pointwise.*

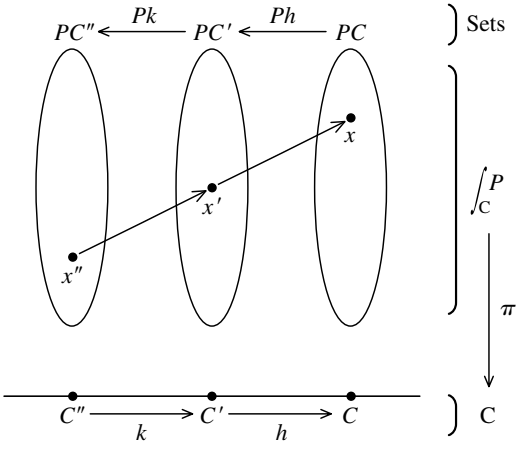


Figure 8.1: Category of elements

Proposition 8.10. *For any small category* $\mathbf{C}$*, every object* P *in the functor category* $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ *is a colimit of representable functors,*

$$\varinjlim_{j \in J} yA_j \;\cong\; P.$$

More precisely, there is a canonical choice of an index category J *and a functor* $A : J \to \mathbf{C}$ *such that there is a natural isomorphism* $\varinjlim_J y \circ A \;\cong\; P$*.*

Proof. Given $P : \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$, the index category we need is the so-called *category of elements* of P, written,

$$\int_{\mathbf{C}} P$$

and defined as follows.

objects: pairs (x, C) where $C \in \mathbf{C}$ and $x \in PC$,

arrows: an $h : (x', C') \to (x, C)$ is an arrow $h : C' \to C$ in $\mathbf{C}$ such that

$$P(h)(x) = x' \tag{8.3}$$

actually, the arrows are triples of the form $(h, (x', C'), (x, C))$ satisfying (8.3).

The reader can easily work out the obvious identities and composites. See Figure 8.1.

Note that $\int_{\mathbf{C}} P$ is a small category since $\mathbf{C}$ is small. There is a "projection" functor,

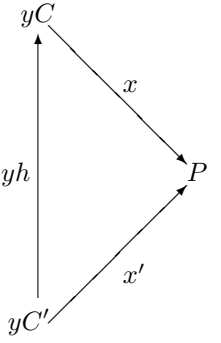
$$\pi : \int_{\mathbf{C}} P \to \mathbf{C}$$

defined by $\pi(x, C) = C$ and $\pi(h, (x', C'), (x, C)) = h$.

To define the cocone of the form $y \circ \pi \to P$, take an object $(x, C) \in \int_{\mathbf{C}} P$ and observe that (by the Yoneda Lemma) there is a natural, bijective correspondence between,

$$\frac{x \in P(C)}{x : yC \to P}$$

which we simply identify notationally. Moreover, given any arrow $h : (x', C') \to (x, C)$ naturality in C implies that there is a commutative triangle

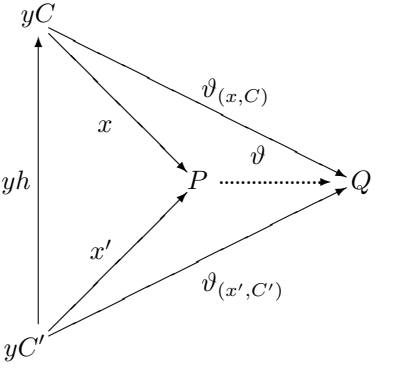


Indeed, the category $\int_{\mathbf{C}} P$ is equivalent to the full subcategory

$$y/P \;\rightarrowtail\; \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}/P$$

of the slice category over P on the objects (i.e. arrows in $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$) with representable domains.

We can therefore take the component of the desired cocone $y\pi \to P$ at (x, C) to be simply $x : yC \to P$. To see that this is a colimiting cocone, take any cocone $y\pi \to Q$ with components $\vartheta_{(x,C)} : yC \to Q$ and we require a unique natural transformation $\vartheta : P \to Q$ as indicated in the following diagram.



We can define $\vartheta_C : PC \to QC$ by setting

$$\vartheta_C(x) = \vartheta_{(x,C)}$$

where we again identify,

$$\frac{\vartheta_{(x,C)} \in Q(C)}{\vartheta_{(x,C)} : yC \to Q}$$

This assignment is easily seen to be natural in C, again by the natural isomorphism of the Yoneda Lemma. We leave the remaining verification of uniqueness to the reader as well. □

We include the following because it fits naturally here, but defer the proof to the next chapter, where a neat proof can be given using adjoint functors. The reader may wish to prove it at this point using the materials at hand, which is also quite doable.

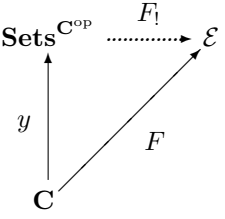
Proposition 8.11. *For any small category* $\mathbf{C}$*, the Yoneda embedding*

$$y : \mathbf{C} \to \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$$

is the "free cocompletion" of $\mathbf{C}$*, in the following sense. Given any cocomplete category* $\mathcal{E}$ *and functor* $F : \mathbf{C} \to \mathcal{E}$*, there is a colimit preserving functor* $F_! : \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}} \to \mathcal{E}$*, unique up to natural isomorphism with the property*

$$F_! \circ y \cong A$$

as indicated in the following diagram.



Proof. (Sketch, see Proposition 9.16.) Given $F : \mathbf{C} \to \mathcal{E}$, define $F_!$ as follows. For any $P \in \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$, let

$$\varinjlim_{j \in J} yA_j \;\cong\; P$$

be the canonical presentation of P as a colimit of representables with $J = \int_{\mathbf{C}} P$, the category of elements of P. Then set,

$$F_!(P) = \varinjlim_{j \in J} F(A_j)$$

which exists since $\mathcal{E}$ is cocomplete. □

8.7 Exponentials in categories of diagrams

As an application, let us consider exponentials in categories of the form $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ for small $\mathbf{C}$. We will need the following lemma.

Lemma 8.12. *For any small index category J, functor $A : J \to \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ and diagram $B \in \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$, there is a natural isomorphism*

$$\varinjlim_j (A_j \times B) \;\cong\; (\varinjlim_j A_j) \times B. \tag{8.4}$$

Briefly, the functor $- \times B : \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}} \to \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ preserves colimits.

Proof. To specify the canonical natural transformation mentioned in (8.4), start with the cocone,

$$\vartheta_j : A_j \to \varinjlim_j A_j, \qquad j \in J$$

apply the functor $- \times B$ to get a cocone,

$$\vartheta_j \times B : A_j \times B \to (\varinjlim_j A_j) \times B, \qquad j \in J$$

and so there is a "comparison arrow" from the colimit,

$$\vartheta : \varinjlim_j (A_j \times B) \to (\varinjlim_j A_j) \times B$$

which we claim is a natural isomorphism.

By a past exercise, it suffices to show that each component,

$$\vartheta_C : (\varinjlim_j (A_j \times B))(C) \to ((\varinjlim_j A_j) \times B)(C)$$

is iso. But since the limits and colimits involved are all computed pointwise, it therefore suffices to show (8.4) under the assumption that the A_j and B are just sets. To that end, take any set X and consider the following isomorphisms in **Sets**,

$$\begin{aligned}
\mathrm{Hom}(\varinjlim_j (A_j \times B), X) &\cong \varprojlim_j \mathrm{Hom}(A_j \times B, X) \\
&\cong \varprojlim_j \mathrm{Hom}(A_j, X^B) && (\mathbf{Sets} \text{ is CCC}) \\
&\cong \mathrm{Hom}(\varinjlim_j A_j, X^B) \\
&\cong \mathrm{Hom}((\varinjlim_j A_j) \times B, X).
\end{aligned}$$

Since these are natural in X, the claim follows by Yoneda. □

Now suppose we have functors P and Q and we want Q^P. The reader should try to construct the exponential "pointwise,"

$$Q^P(C) \stackrel{?}{=} Q(C)^{P(C)}$$

to see that it *does not* work (it is not functorial).

If we had such an exponential Q^P, we could compute its value at any object $C \in \mathbf{C}$ by Yoneda:

$$Q^P(C) \cong \mathrm{Hom}(yC, Q^P)$$

And if it is to be an exponential, then we must also have

$$\mathrm{Hom}(yC, Q^P) \cong \mathrm{Hom}(yC \times P, Q).$$

But this latter set *does* exist. Thus, we can just *define*

$$Q^P(C) = \mathrm{Hom}(yC \times P, Q) \tag{8.5}$$

with the action on $h : C' \to C$ being

$$Q^P(h) = \mathrm{Hom}(yh \times 1_P, Q).$$

This is clearly a contravariant, set-valued functor on $\mathbf{C}$. Let us now check that it indeed gives an exponential of P and Q.

Proposition 8.13. *For any objects X, P, Q in $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$, there is an isomorphism, natural in X,*

$$\mathrm{Hom}(X, Q^P) \cong \mathrm{Hom}(X \times P, Q).$$

Proof. By Proposition 8.10, for a suitable index category J, we can write X as a colimit of representables,

$$X \cong \varinjlim_{j \in J} yC_j.$$

Thus we have isomorphisms,

$$\begin{aligned}
\mathrm{Hom}(X, Q^P) &\cong \mathrm{Hom}(\varinjlim_j yC_j, Q^P) \\
&\cong \varprojlim_j \mathrm{Hom}(yC_j, Q^P) \\
&\cong \varprojlim_j Q^P(C_j) && \text{(by Yoneda)} \\
&\cong \varprojlim \mathrm{Hom}(yC_j \times P, Q) && \text{(by 8.5)} \\
&\cong \mathrm{Hom}(\varinjlim_j (yC_j \times P), Q) \\
&\cong \mathrm{Hom}(\varinjlim_j (yC_j) \times P, Q) && \text{(Lemma 8.12)} \\
&\cong \mathrm{Hom}(X \times P, Q).
\end{aligned}$$

And these isos are clearly natural in X. □

Theorem 8.14. *For any small category* $\mathbf{C}$*, the category of diagrams* $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ *is cartesian closed. Moreover, the Yoneda embedding*

$$y : \mathbf{C} \to \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$$

preserves all products and exponentials that exist in $\mathbf{C}$*.*

Proof. In light of the foregoing proposition, it only remains to show that y preserves products and exponentials. We leave this as an exercise. □

8.8 Topoi

Since we are now so close to it, we might as well introduce the important notion of a "topos," even though this is not the place to develop that theory, as appealing as it is. First we require the following generalization of characteristic functions of subsets.

Definition 8.15. Let $\mathcal{E}$ be a category with all finite limits. A *subobject classifier* in $\mathcal{E}$ consists of an object Ω together with an arrow $t : 1 \to \Omega$ that is a "universal subobject," in the following sense:

> Given any object E and any subobject $U \rightarrowtail E$, there is a unique arrow $u : E \to \Omega$ making the following diagram a pullback.

$$\begin{array}{ccc} U & \longrightarrow & 1 \\ \downarrow & & \downarrow{\scriptstyle t} \\ E & \xrightarrow[u]{} & \Omega \end{array} \tag{8.6}$$

The arrow u is called the *classifying* arrow of the subobject $U \rightarrowtail E$; it can be thought of as taking exactly the part of E that is U to the "point" t of Ω. The most familiar example of a subobject classifier is of course the set $2 = \{0, 1\}$ with a selected element as $t : 1 \to 2$. The fact that every subset $U \subseteq S$ of any set S has a unique characteristic function $u : S \to 2$ is then exactly the subobject classifier condition.

It is easy show that a subobject classifier is unique up to isomorphism: the pullback condition is clearly equivalent to requiring the contravariant subobject functor,

$$\mathrm{Sub}_{\mathcal{E}}(-) : \mathcal{E}^{\mathrm{op}} \to \mathbf{Sets}$$

(which acts by pullback) to be representable,

$$\mathrm{Sub}_{\mathcal{E}}(-) \;\cong\; \mathrm{Hom}_{\mathcal{E}}(-, \Omega).$$

The required isomorphism is just the pullback condition stated in the definition of a subobject classifier.

Definition 8.16. A *topos* is a category $\mathcal{E}$ such that

1. $\mathcal{E}$ has all finite limits,
2. $\mathcal{E}$ has a subobject classifier,
3. $\mathcal{E}$ has all exponentials.

This compact definition proves to be amazingly rich in consequences: it can be shown for instance that topoi also have all finite colimits and that every slice category of a topos is again a topos. We refer the reader to the books by Mac Lane and Moerdijk, Johnstone, and McLarty for information on topoi, and here just give an example (albeit one that covers a very large number of cases).

Proposition 8.17. *For any small category* $\mathbf{C}$*, the category of diagrams* $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ *is a topos.*

Proof. Since we already know that $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ has all limits, and we know that it has exponentials by the foregoing section, we just need to find a subobject classifier. To that end, we define a *sieve* on an object C of $\mathbf{C}$ to be a set S of arrows $f : \cdot \to C$ (with arbitrary domain) that is closed under precomposition, that is, if $f : D \to C$ is in S then so is $f \circ g : E \to D \to C$ for every $g : E \to D$ (think of a sieve as a common generalization of a "lower set" in a poset and an "ideal" in a ring). Then let

$$\Omega(C) \;=\; \{S \subseteq \mathbf{C}_1 \mid S \text{ is a sieve on } C\}$$

and given $h : D \to C$ let

$$h^* : \Omega(C) \to \Omega(D)$$

be defined by

$$h^*(S) = \{g : \cdot \to D \mid h \circ g \in S\}.$$

This clearly defines a presheaf $\Omega : \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$, with a distinguished point,

$$t : 1 \to \Omega$$

namely, at each C, the "total sieve"

$$t_C = \{f : \cdot \to C\}.$$

We claim that $t : 1 \to \Omega$ so defined is a subobject classifier for $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$. Indeed, given any object E and subobject $U \rightarrowtail E$, define $u : E \to \Omega$ at any object $C \in \mathbf{C}$ by:

$$u_C(e) = \{f : D \to C \mid f^*(e) \in U(D) \rightarrowtail E(D)\}$$

for any $e \in E(C)$. That is, $u_C(e)$ is the sieve of arrows into C that take $e \in E(C)$ back into the subobject U.

The reader should verify that this specification does indeed determine a unique classifying morphism for $U \rightarrowtail E$. □

Remark 8.18. One of the most fascinating aspects of topoi is their relation to logic. In virtue of the association of subobjects $U \rightarrowtail E$ with arrows $u : E \to \Omega$, the subobject classifier Ω can be regarded as an object of "propositions" or "truth-values," with $t = \text{true}$. An arrow $u : E \to \Omega$ is then a "propositional function" of which $U \rightarrowtail E$ is the "extension." For, by the pullback condition (8.6), a generalized element $x : X \to E$ is "in" U (i.e. factors through $U \rightarrowtail E$) just if $ux = \text{true}$,

$$x \in_E U \qquad \text{iff} \qquad ux = \text{true}$$

so that, again in the notation of Section 5.1,

$$U = \{x \in E \mid ux = \text{true}\}.$$

This permits an interpretation of first-order logic in any topos, since topoi also have a way of modeling the logical quantifiers $\exists$ and $\forall$ as adjoints to pullbacks (as will be described in Section 9.5).

Since topoi are also cartesian closed, they have an internal type theory described by the λ-calculus (see Section 6.5). Combining this with the first-order logic and subobject classifier Ω provides a natural interpretation of *higher-order* logic, essentially employing the exponential Ω^E as a "power object" $P(E)$ of subobjects of E. This logical aspect of topoi is also treated in the books already mentioned.

8.9 Exercises

1. If $F : \mathbf{C} \to \mathbf{D}$ is full and faithful, then $C \cong C'$ iff $FC \cong FC'$.
2. Let $\mathbf{C}$ be a locally small, cartesian closed category. Use the Yoneda embedding to show that for any objects A, B, C in $\mathbf{C}$

   $$(A \times B)^C \cong A^C \times B^C$$

   (cf. Problem 1., Chapter 6). If $\mathbf{C}$ also has binary coproducts, show that also

   $$A^{(B+C)} \cong A^B \times A^C.$$

3. For any locally small category $\mathbf{C}$, the functor category $\mathbf{Sets}^{\mathbf{C}^{\text{op}}}$ has binary products $F \times G$, and these are computed pointwise:

   $$(F \times G)(C) \cong FC \times GC$$

   for all $C \in \mathbf{C}_0$.

4. Let $\mathbf{C}$ be a locally small category with binary products, and show that the Yoneda embedding

$$y\colon \mathbf{C} \to \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$$

preserves them. (*Hint*: this involves only a few lines of calculation.)
If $\mathbf{C}$ also has exponentials, show that y also preserves them.

5. (a) Explicitly determine the subobject classifiers for the topoi $\mathbf{Sets}^{\mathbf{2}}$ and $\mathbf{Sets}^{\omega}$, where as always $\mathbf{2}$ is the poset $0 < 1$ and ω is the poset of natural numbers $0 < 1 < 2 < \cdots$.

   (b) Show that $(\mathbf{Sets}_{\mathrm{fin}})^{\mathbf{2}}$ is a topos.

This page intentionally left blank

9

ADJOINTS

This chapter represents the high point of this book, the goal toward which we have been working steadily. The notion of adjoint functor applies everything that we have learned up to now to unify and subsume all the different universal mapping properties that we have encountered, from free groups to limits to exponentials. But more importantly, it also captures an important mathematical phenomenon that is invisible without the lens of category theory. Indeed, I will make the admittedly provocative claim that adjointness is a concept of fundamental logical and mathematical importance that is not captured elsewhere in mathematics. pt

Many of the most striking applications of category theory involve adjoints and many important and fundamental mathematical notions are instances of adjoint functors. As such, they share the common behavior and formal properties of all adjoints and in many cases this fact alone accounts for all of their essential features.

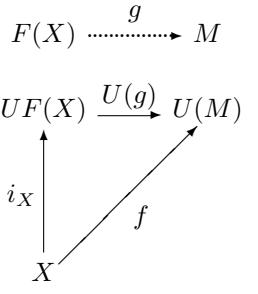
9.1 Preliminary definition

We begin by recalling the universal mapping property (UMP) of free monoids: every monoid M has an underlying set $U(M)$, and every set X has a free monoid $F(X)$, and there is a function

$$i_X : X \to UF(X)$$

with the following UMP.

> For every monoid M and every function $f : X \to U(M)$ there is a unique homomorphism $g : F(X) \to M$ such that $f = U(g) \circ i_X$, all as indicated in the diagram:



Now consider the following map

$$\phi : \mathrm{Hom}_{\mathbf{Mon}}(F(X), M) \to \mathrm{Hom}_{\mathbf{Sets}}(X, U(M))$$

defined by

$$g \mapsto U(g) \circ i_X.$$

The UMP above says exactly that ϕ is an isomorphism,

$$\mathrm{Hom}_{\mathbf{Mon}}(F(X), M) \;\cong\; \mathrm{Hom}_{\mathbf{Sets}}(X, U(M)). \tag{9.1}$$

This bijection (9.1) can also be written schematically as a *2-way rule*

$$\frac{F(X) \longrightarrow M}{X \longrightarrow U(M)}$$

where one gets from an arrow g of the upper form to one $\phi(g)$ of the lower form by the recipe

$$\phi(g) = U(g) \circ i_X.$$

We pattern our *preliminary* definition of adjunction on this situation. It is preliminary because it really only gives half of the picture, as it were; in the next section an equivalent definition will emerge as both more convenient and conceptually clearer.

Definition 9.1 (preliminary). An *adjunction* between categories $\mathbf{C}$ and $\mathbf{D}$ consists of functors

$$F : \mathbf{C} \rightleftarrows \mathbf{D} : \mathbf{U}$$

and a natural transformation

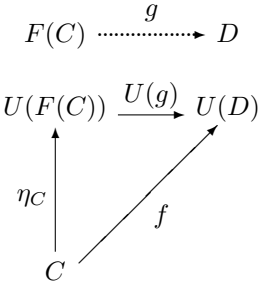
$$\eta : 1_{\mathbf{C}} \to U \circ F$$

with the property:

(*) For any $C \in \mathbf{C}$, $D \in \mathbf{D}$, and $f : C \to U(D)$, there exists a unique $g : FC \to D$ such that

$$f = U(g) \circ \eta_C$$

as indicated in



Terminology and notation:

- F is called the *left adjoint*, U is called the *right adjoint*, and η is called the *unit* of the adjunction.
- One sometimes writes $F \dashv U$ for "F is left and U right adjoint."
- The statement (*) is the UMP of the unit η.

Note that the situation $F \dashv U$ is a generalization of equivalence of categories, in that a pseudo-inverse is an adjoint. In that case, however, it is the relation between categories that one is interested in. Here, one is concerned with the relation between specific functors. Thus, it is not the relation on categories "there exists an adjunction," but rather "this functor has an adjoint" that we are concerned with.

Suppose now that we have an adjunction,

$$\mathbf{C} \underset{F}{\overset{U}{\leftrightarrows}} \mathbf{D}.$$

Then, as in the example of monoids, take $C \in \mathbf{C}$ and $D \in \mathbf{D}$ and consider the operation

$$\phi : \mathrm{Hom}_{\mathbf{D}}(FC, D) \to \mathrm{Hom}_{\mathbf{C}}(C, UD)$$

given by $\phi(g) = U(g) \circ \eta_C$. Since, by the UMP of η, every $f : C \to UD$ is $\phi(g)$ for a unique g, just as in our example we see that ϕ is an isomorphism

$$\mathrm{Hom}_{\mathbf{D}}(F(C), D) \cong \mathrm{Hom}_{\mathbf{C}}(C, U(D)) \tag{9.2}$$

which, again, can be displayed as the two-way rule:

$$\frac{F(C) \longrightarrow D}{C \longrightarrow U(D)}$$

Example 9.2. Consider the "diagonal" functor,

$$\Delta : \mathbf{C} \to \mathbf{C} \times \mathbf{C}$$

defined on objects by

$$\Delta(C) = (C, C)$$

and on arrows by

$$\Delta(f : C \to C') = (f, f) : (C, C) \to (C', C').$$

What would it mean for this functor to have a right adjoint? We would need a functor $R : \mathbf{C} \times \mathbf{C} \to \mathbf{C}$ such that for all $C \in \mathbf{C}$ and $(X, Y) \in \mathbf{C} \times \mathbf{C}$, there is a bijection:

$$\frac{\Delta C \longrightarrow (X, Y)}{C \longrightarrow R(X, Y)}$$

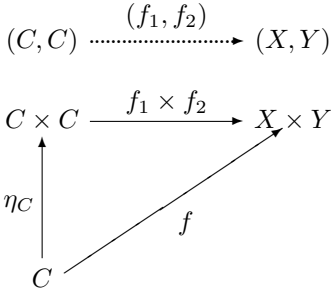
That is, we would have

$$\begin{aligned}\mathrm{Hom}_{\mathbf{C}}(C, R(X,Y)) &\cong \mathrm{Hom}_{\mathbf{C}\times\mathbf{C}}(\Delta C, (X,Y)) \\ &\cong \mathrm{Hom}_{\mathbf{C}}(C,X) \times \mathrm{Hom}_{\mathbf{C}}(C,Y).\end{aligned}$$

We therefore must have $R(X,Y) \cong X \times Y$, suggesting that Δ has as a right adjoint the product functor $\times : \mathbf{C} \times \mathbf{C} \to \mathbf{C}$,

$$\Delta \dashv \times.$$

The counit η would have the form $\eta_C : C \to C \times C$, so we propose the "diagonal arrow" $\eta_C = \langle 1_C, 1_C \rangle$, and we need to check the UMP indicated in the following.



Indeed, given any $f : C \to X \times Y$, we have unique f_1 and f_2 with $f = \langle f_1, f_2 \rangle$, for which we then have

$$\begin{aligned}(f_1 \times f_2) \circ \eta_C &= \langle f_1\pi_1, f_2\pi_2 \rangle \eta_C \\ &= \langle f_1\pi_1\eta_C, f_2\pi_2\eta_C \rangle \\ &= \langle f_1, f_2 \rangle \\ &= f.\end{aligned}$$

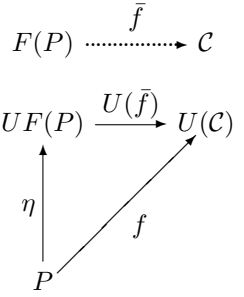
Thus in sum, the functor Δ has a right adjoint if and only if $\mathbf{C}$ has binary products.

Example 9.3. For an example of a different sort, consider the category **Pos** of posets and monotone maps and $\mathcal{C}$**Pos** of cocomplete posets and cocontinuous maps. A poset $\mathcal{C}$ is cocomplete just if it has a join $\bigvee_i c_i$ for every family of elements $(c_i)_{i\in I}$ indexed by a set I, and a monotone map $f : \mathcal{C} \to \mathcal{D}$ is cocontinuous if it preserves all such joins, $f(\bigvee_i c_i) = \bigvee_i f(c_i)$. There is an obvious forgetful functor

$$U : \mathcal{C}\mathbf{Pos} \to \mathbf{Pos}.$$

What would a left adjoint $F \dashv U$ be? There would have to be a monotone map $\eta : P \to UF(P)$ with the property: given any cocomplete poset $\mathcal{C}$ and monotone

$f : P \to U(\mathcal{C})$, there exists a unique cocontinuous $\bar{f} : F(P) \to \mathcal{C}$ such that $f = U(\bar{f}) \circ \eta_P$, as indicated in



In this precise sense, such a poset $F(P)$ would be a "free cocompletion" of P, and $\eta : P \to UF(P)$ a "best approximation" of P by a cocomplete poset.

We leave it to the reader to show that such a "cocompletion" always exists, namely the poset of *lower sets*,

$$\mathrm{Low}(P) = \{U \subseteq P \mid p' \leq p \in U \text{ implies } p' \in U\}.$$

9.2 Hom-set definition

The following proposition shows that the isomorphism (9.2) is in fact natural in both C and D.

Proposition 9.4. *Given categories and functors,*

$$\mathbf{C} \underset{F}{\overset{U}{\rightleftarrows}} \mathbf{D}$$

the following conditions are equivalent:

1. *F is left adjoint to U; that is, there is a natural transformation*

$$\eta : 1_{\mathbf{C}} \to U \circ F$$

that has the UMP of the unit:

> *For any $C \in \mathbf{C}$, $D \in \mathbf{D}$ and $f : C \to U(D)$ there exists a unique $g : FC \to D$ such that*
>
> $$f = U(g) \circ \eta_C.$$

2. *For any $C \in \mathbf{C}$ and $D \in \mathbf{D}$ there is an isomorphism,*

$$\phi : \mathrm{Hom}_{\mathbf{D}}(FC, D) \cong \mathrm{Hom}_{\mathbf{C}}(C, UD)$$

that is natural in both C and D.

Moreover, the two conditions are related by the formulas

$$\phi(g) = U(g) \circ \eta_C$$
$$\eta_C = \phi(1_{FC}).$$

Proof. (1 implies 2) The recipe for ϕ, given η is just the one stated and we have already observed it to be an isomorphism, given the UMP of the unit. For naturality in C, take $h : C' \to C$ and consider the following diagram:

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbf{D}}(FC, D) & \xrightarrow[\cong]{\phi_{C,D}} & \mathrm{Hom}_{\mathbf{C}}(C, UD) \\ \Big\downarrow{\scriptstyle (Fh)^*} & & \Big\downarrow{\scriptstyle h^*} \\ \mathrm{Hom}_{\mathbf{D}}(FC', D) & \xrightarrow[\phi_{C',D}]{\cong} & \mathrm{Hom}_{\mathbf{C}}(C', UD) \end{array}$$

Then for any $f : FC \to D$ we have

$$\begin{aligned} h^*(\phi_{C,D}(f)) &= (U(f) \circ \eta_C) \circ h \\ &= U(f) \circ UF(h) \circ \eta_{C'} \\ &= U(f \circ F(h)) \circ \eta_{C'} \\ &= \phi_{C',D}(F(h)^*(f)). \end{aligned}$$

For naturality in D, take $g : D \to D'$ and consider the diagram

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbf{D}}(FC, D) & \xrightarrow[\cong]{\phi_{C,D}} & \mathrm{Hom}_{\mathbf{C}}(C, UD) \\ \Big\downarrow{\scriptstyle g_*} & & \Big\downarrow{\scriptstyle U(g)_*} \\ \mathrm{Hom}_{\mathbf{D}}(FC, D') & \xrightarrow[\phi_{C',D}]{\cong} & \mathrm{Hom}_{\mathbf{C}}(C, UD') \end{array}$$

Then for any $f : FC \to D$ we have

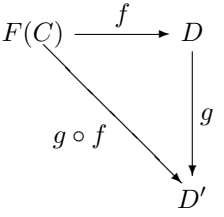
$$\begin{aligned} U(g)_*(\phi_{C,D}(f)) &= U(g) \circ (U(f) \circ \eta_C) \\ &= U(g \circ f) \circ \eta_C \\ &= \phi_{C',D}(g \circ f) \circ \eta_C \\ &= \phi_{C',D}(g_*(f)) \circ \eta_C. \end{aligned}$$

So ϕ is indeed natural.

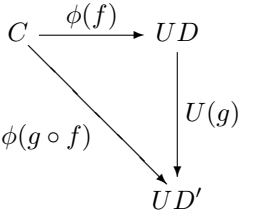
(2 implies 1) We are given a bijection ϕ,

$$\frac{F(C) \longrightarrow D}{C \longrightarrow U(D)} \tag{9.3}$$

for each C, D, that is natural in C and D. In detail, this means that given a commutative triangle



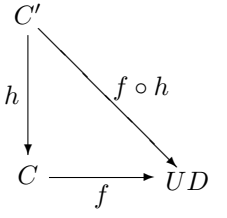
there are two ways to get an arrow of the form $C \to UD'$, namely



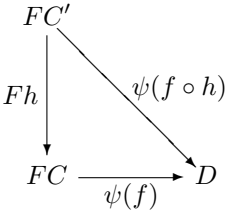
Naturality in D means that this diagram commutes,

$$\phi(g \circ f) = U(g) \circ \phi(f).$$

Dually, naturality in C means that given



and writing $\psi = \phi^{-1}$, the following commutes:



That is,

$$\psi(f \circ h) = \psi(f) \circ F(h).$$

Now, given such a natural bijection ϕ we want a natural transformation

$$\eta : 1_{\mathbf{C}} \to U \circ F$$

with the UMP of the unit. To find

$$\eta_C : C \to UFC$$

we put FC for D and $1_{FC} : FC \to FC$ in the adjoint schema (9.3) to get

$$\frac{1_{FC} : \quad FC \longrightarrow FC}{\eta_C : \quad C \longrightarrow UFC} \ \phi$$

That is, we define

$$\eta_C = \phi(1_{FC}).$$

Finally, to see that η has the required UMP of the unit, it clearly suffices to show that for all $g : FC \to D$, we have

$$\phi(g) = U(g) \circ \eta_C$$

since we are assuming that ϕ is iso. But

$$\begin{aligned} U_g \circ \eta_C &= U_g \circ \phi(1_{FC}) \\ &= \phi(g \circ 1_{FC}) \\ &= \phi(g). \end{aligned}$$

□

Note that the second condition in the proposition is symmetric, but the first condition is not. This implies that we also have following dual proposition.

Corollary 9.5. *Given categories and functors*

$$\mathbf{C} \underset{F}{\overset{U}{\leftarrow\!\!\rightarrow}} \mathbf{D}$$

the following conditions are equivalent:

1. *For any $C \in \mathbf{C}$, $D \in \mathbf{D}$ there is an isomorphism*

   $$\phi : \mathrm{Hom}_{\mathbf{D}}(FC, D) \cong \mathrm{Hom}_{\mathbf{C}}(C, UD)$$

   that is natural in C and D.

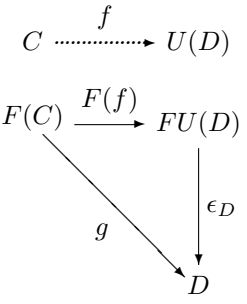
2. *There is a natural transformation*

   $$\epsilon : F \circ U \to 1_{\mathbf{D}}$$

   with the following UMP:

   > *For any $C \in \mathbf{C}$, $D \in \mathbf{D}$ and $g : F(C) \to D$, there exists a unique $f : C \to UD$ such that:*
   >
   > $$g = \epsilon_D \circ F(f)$$

as indicated in the diagram:



Moreover, the two conditions are related by the equations

$$\psi(f) = \epsilon_D \circ F(f)$$
$$\epsilon_D = \psi(1_{UD})$$

where $\psi = \phi^{-1}$.

Proof. Duality. □

We take the symmetric "Hom-set" formulation as our "official" definition of an adjunction.

Definition 9.6 "official". An *adjunction* consists of functors

$$F : \mathbf{C} \rightleftarrows \mathbf{D} : \mathbf{U}$$

and a natural isomorphism

$$\phi : \mathrm{Hom}_{\mathbf{D}}(FC, D) \cong \mathrm{Hom}_{\mathbf{C}}(C, UD) : \psi.$$

This definition has the virtue of being symmetric in F and U. The unit $\eta : 1_{\mathbf{C}} \to U \circ F$ and the *counit* $\epsilon : F \circ U \to 1_{\mathbf{D}}$ of the adjunction are then determined as:

$$\eta_C = \phi(1_{FC})$$
$$\epsilon_D = \psi(1_{UD})$$

9.3 Examples of adjoints

Example 9.7. Suppose $\mathbf{C}$ has binary products. Take a fixed object $A \in \mathbf{C}$, and consider the product functor

$$- \times A : \mathbf{C} \to \mathbf{C}$$

defined on objects by

$$X \mapsto X \times A$$

and on arrows by

$$(h : X \to Y) \mapsto (h \times 1_A : X \times A \longrightarrow Y \times A).$$

When does $- \times A$ have a right adjoint?

We would need a functor

$$U : \mathbf{C} \to \mathbf{C}$$

such that for all $X, Y \in \mathbf{C}$ there is a natural bijection

$$\frac{X \times A \longrightarrow Y}{X \longrightarrow U(Y)}$$

So let us try defining U by

$$U(Y) = Y^A$$

on objects and on arrows by

$$U(g : Y \to Z) = g^A : Y^A \longrightarrow Z^A.$$

Putting $U(Y)$ for X in the adjoint schema above then gives the counit:

$$\frac{Y^A \times A \xrightarrow{\epsilon} Y}{Y^A \xrightarrow[1]{} Y^A}$$

This is, therefore, an adjunction if ϵ has the following UMP:

> For any $f : X \times A \to Y$, there is a unique $\bar{f} : X \to Y^A$ such that $f = \epsilon \circ (\bar{f} \times 1_A)$.

But this is exactly the UMP of the exponential! Thus, we do indeed have an adjunction:

$$- \times A \ \dashv \ -^A$$

Example 9.8. Here is a much more simple example. For any category $\mathbf{C}$, consider the unique functor to the terminal category $\mathbf{1}$,

$$! : \mathbf{C} \to \mathbf{1}.$$

Now we ask, when does ! have a right adjoint? This would be an object $U : \mathbf{1} \to \mathbf{C}$ such that for any $C \in \mathbf{C}$, there is a bijective correspondence,

$$\frac{!C \longrightarrow *}{C \longrightarrow U(*)}$$

Such a U would have to be a terminal object in $\mathbf{C}$. So ! has a right adjoint iff $\mathbf{C}$ has a terminal object. What do you think a left adjoint would be?

This last example is a clear case of the following general fact:

Proposition 9.9. *Adjoints are unique up to isomorphism. Specifically, given a functor $F : \mathbf{C} \to \mathbf{D}$ and right adjoints $U, V : \mathbf{D} \to \mathbf{C}$,*

$$F \dashv U \quad \textit{and} \quad F \dashv V$$

we then have $U \cong V$.

Proof. Here is the easy way. For any $D \in \mathbf{D}$, and $C \in \mathbf{C}$ we have

$$\begin{aligned} \mathrm{Hom}_{\mathbf{C}}(C, UD) &\cong \mathrm{Hom}_{\mathbf{D}}(FC, D) \quad \text{naturally, since } F \dashv U \\ &\cong \mathrm{Hom}_{\mathbf{C}}(C, VD) \quad \text{naturally, since } F \dashv V. \end{aligned}$$

Thus, by Yoneda, $UD \cong VD$. But this isomorphism is natural in D, again by adjointness. □

This proposition implies that one can use the condition of being right or left adjoint to a given functor to define (uniquely characterize up to isomorphism) a new functor. This sort of characterization, like a UMP, determines an object or construction "intrinsically," in terms of its relation to some other given construction. Many important constructions turn out to be adjoints to particularly simple ones.

For example, what do you suppose would be a *left* adjoint to the diagonal functor

$$\Delta : \mathbf{C} \to \mathbf{C} \times \mathbf{C}$$

in the earlier example, where $\Delta(C) = (C, C)$ and we had $\Delta \dashv \times$? It would have to be functor $L(X, Y)$ standing in the correspondence

$$\frac{L(X,Y) \longrightarrow C}{(X,Y) \longrightarrow (C,C)}$$

Thus, it could only be the coproduct $L(X, Y) = X + Y$. Therefore, Δ has a left adjoint if and only if $\mathbf{C}$ has binary coproducts, and then

$$+ \dashv \Delta.$$

Next, note that $\mathbf{C} \times \mathbf{C} \cong \mathbf{C}^2$ where 2 is the discrete two-object category (i.e. any two-element set). Then $\Delta(C)$ is the constant C-valued functor for each $C \in \mathbf{C}$. Let us now replace 2 by any small index category $\mathbf{J}$ and consider possible adjoints to the corresponding diagonal functor

$$\Delta : \mathbf{C} \to \mathbf{C}^{\mathbf{J}}$$

with $\Delta(C)(j) = C$ for all $C \in \mathbf{C}$ and $j \in \mathbf{J}$. In this case, one has left and right adjoints

$$\varinjlim \ \dashv \ \Delta \ \dashv \ \varprojlim$$

if and only if **C** has colimits and limits, respectively, of type **J**. Thus all of the particular limits and colimits we met earlier, such as pullbacks and coequalizers are instances of adjoints. We leave the proof of the general fact as an exercise. What are the units and counits of these adjunctions?

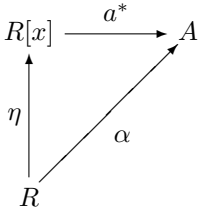
Example 9.10. *Polynomial rings:* Let R be a commutative ring ($\mathbb{Z}$ if you like) and consider the ring $R[x]$ of polynomials in one indeterminate x with coefficients in R. The elements of $R[x]$ all look like this:

$$r_0 + r_1 x + r_2 x^2 + \cdots + r_n x^n \tag{9.4}$$

with the coefficients $r_i \in R$. Of course, there may be some identifications between such expressions depending on the ring R.

There is an evident homomorphism $\eta : R \to R[x]$, taking elements r to constant polynomials $r = r_0$, and it has the following UMP:

> Given any ring A, homomorphism $\alpha : R \to A$, and element $a \in A$, there is a unique homomorphism
>
> $$a^* : R[x] \to A$$
>
> such that $a^*(x) = a$ and $a^*\eta = \alpha$.



Namely, for a^* we take the "formal evaluation at a"

$$a^*(r(x)) = \alpha(r)(a/x)$$

given by applying α to the coefficients r_i, substituting a for x, and evaluating the result in A,

$$a^*(r_0 + r_1 x + r_2 x^2 + \cdots + r_n x^n) = \alpha(r_0) + \alpha(r_1)a + \alpha(r_2)a^2 + \cdots + \alpha(r_n)a^n.$$

To describe this in terms of adjoints, define $\mathbf{Rings}_*$ to be the category of "pointed" rings, with objects of the form (A, a), where A is a ring and $a \in A$, and arrows $h : (A, a) \to (B, b)$ are homomorphisms $h : A \to B$ that preserve the distinguished point, $h(a) = b$.

The UMP just given says exactly that the functor

$$U : \mathbf{Rings}_* \to \mathbf{Rings}$$

that "forgets the point" $U(A, a) = A$ has as left adjoint the functor

$$[x] : \mathbf{Rings} \to \mathbf{Rings}_*$$

that "adjoins an indeterminate"

$$[x](R) = (R[x], x)$$

and $\eta : R \to R[x]$ is the unit of the adjunction. The reader should have no difficulty working out the details of this example. This provides a characterization of the polynomial ring $R[x]$ that does not depend on the somewhat vague description in terms of "formal polynomial expressions" like (9.4).

9.4 Order adjoints

Let P be a *preordered set*, that is, a category in which there is at most one arrow $x \to y$ between any two objects. A poset is a preorder that is skeletal. We define an ordering relation on the objects of P by

$$x \leq y \quad \text{iff there exists an arrow } x \to y.$$

Given another such preorder Q, suppose we have adjoint functors:

$$P \underset{U}{\overset{F}{\rightleftarrows}} Q \qquad\qquad F \dashv U$$

Then the correspondence $Q(Fa, x) \cong P(a, Ux)$ comes down to the simple condition $Fa \leq x$ iff $a \leq Ux$. Thus an adjunction on preorders consists simply of order preserving maps F, U satisfying the two-way rule or "bicondition":

$$\frac{Fa \leq x}{a \leq Ux}$$

For each $p \in P$, the unit is therefore an element $p \leq UFp$ that is least among all x with $p \leq Ux$. Dually, for each $q \in Q$ the counit is an element $FUq \leq q$ that is greatest among all y with $Fy \leq q$.

Such a set-up on preordered sets is sometimes called a *Galois connection*.

Example 9.11. A basic example is the interior operation on the subsets of a topological space X. Let $\mathcal{O}(X)$ be the set of open subsets of X and consider the operations of inclusion and interior:

$$\text{inc} : \mathcal{O}(X) \to \mathcal{P}(X)$$
$$\text{int} : \mathcal{P}(X) \to \mathcal{O}(X)$$

For any subset A and open subset U, the valid bicondition

$$\frac{U \subseteq A}{U \subseteq int(A)}$$

means that the interior operation is right adjoint to the inclusion of the open subsets among all the subsets:

$$\text{inc} \dashv \text{int}$$

The counit here is the inclusion $int(A) \subseteq A$, valid for all subsets A. The case of closed subsets and the closure operation is dual.

Example 9.12. A related example is the adjunction on powersets induced by any function $f : A \to B$, between the inverse image operation f^{-1} and the direct image $\mathrm{im}(f)$,

$$\mathcal{P}(A) \underset{\mathrm{im}(f)}{\overset{f^{-1}}{\leftarrow\!\!\!\rightarrow}} \mathcal{P}(B)$$

Here we have an adjunction $\mathrm{im}(f) \dashv f^{-1}$ as indicated by the bicondition

$$\frac{\mathrm{im}(f)(U) \subseteq V}{U \subseteq f^{-1}(V)}$$

which is plainly valid for all subsets $U \subseteq A$ and $V \subseteq B$.

The inverse image operation $f^{-1} : \mathcal{P}(B) \to \mathcal{P}(A)$ also has a *right* adjoint, given by

$$f_*(U) = \{b \in B \mid f^{-1}(b) \subseteq U\}$$

which we also leave for the reader to verify.

Note that if A and B are topological spaces and $f : A \to B$ is continuous, then f^{-1} restricts to the open sets $f^{-1} : \mathcal{O}(B) \to \mathcal{O}(A)$. Then the left adjoint $\mathrm{im}(f)$ need not exist (on opens), but the right adjoint f_* still does.

$$\mathcal{O}(A) \underset{f_*}{\overset{f^{-1}}{\leftarrow\!\!\!\rightarrow}} \mathcal{O}(B)$$

Example 9.13. Suppose we have a preorder P. Then, as we know, P has *meets* iff for all $p, q \in P$, there is an element $p \wedge q \in P$ satisfying the bicondition

$$\frac{r \leq p \wedge q}{r \leq p \text{ and } r \leq q}$$

Dually, P has *joins* if there is always an element $p \vee q \in P$ such that

$$\frac{p \vee q \leq r}{p \leq r \text{ and } q \leq r}$$

The *Heyting implication* $q \Rightarrow r$ is characterized as an exponential by the bicondition

$$\frac{p \wedge q \leq r}{p \leq q \Rightarrow r}$$

Finally, an initial object 0 and a terminal object 1 are determined by the conditions

$$0 \leq p$$

and

$$p \leq 1.$$

In this way, the notion of a Heyting algebra can be formulated entirely in terms of adjoints. Equivalently, the intuitionistic propositional calculus is neatly axiomatized by the "adjoint rules of inference" just given. Together with the reflexivity and transitivity of entailment $p \vdash q$, these rules are completely sufficient for the propositional logical operations. That is, they can serve as the rules of inference for a logical calculus of "sequents" $p \vdash q$ which is equivalent to the usual intuitionistic propositional calculus.

When we furthermore define negation by $\neg p = p \Rightarrow \bot$, we then get the derived rule

$$\frac{q \vdash \neg p}{p \wedge q \leq 0}$$

Finally, the classical propositional calculus results from adding the rule

$$\neg\neg p \vdash p.$$

Let us now consider how this analysis can be extended to all of first-order logic.

9.5 Quantifiers as adjoints

Traditionally, the main obstacle to the further development of *algebraic logic* has been the treatment of the quantifiers. Categorical logic solves this problem beautifully with the recognition (due to F.W. Lawvere in the 1960s) that they, too, are adjoint functors.

Let $\mathcal{L}$ be a first-order language. For any list $\bar{x} = x_1, \ldots, x_n$ of distinct variables let us denote the set of formulas with at most those variables free by

$$\mathrm{Form}(\bar{x}) = \{\phi(\bar{x}) \mid \phi(\bar{x}) \text{ has at most } \bar{x} \text{ free}\}.$$

Then $\mathrm{Form}(\bar{x})$ is a preorder under the entailment relation of first-order logic

$$\phi(\bar{x}) \vdash \psi(\bar{x}).$$

Now let y be a variable not in the list $\bar{x}$, and note that we have a trivial operation

$$* : \mathrm{Form}(\bar{x}) \to \mathrm{Form}(\bar{x}, y)$$

taking each $\phi(\bar{x})$ to itself; this is just a matter of observing that if $\phi(\bar{x}) \in \mathrm{Form}(\bar{x})$ then y cannot be free in $\phi(\bar{x})$. Of course, $*$ is a functor since,

$$\phi(\bar{x}) \vdash \psi(\bar{x}) \quad \text{ in } \mathrm{Form}(\bar{x})$$

implies,

$$*\phi(\bar{x}) \vdash *\psi(\bar{x}) \quad \text{in } \mathrm{Form}(\bar{x}, y).$$

Now since for any $\psi(\bar{x}, y) \in \mathrm{Form}(\bar{x}, y)$ there is, of course, no free y in the formula $\forall y.\psi(\bar{x}, y)$, there is a map

$$\forall y : \mathrm{Form}(\bar{x}, y) \to \mathrm{Form}(\bar{x}).$$

We claim that this map is right adjoint to $*$,

$$* \dashv \forall.$$

Indeed, the usual rules of universal introduction and elimination imply that the following two-way rule of inference holds:

$$\frac{*\phi(\bar{x}) \vdash \psi(\bar{x}, y) \qquad \mathrm{Form}(\bar{x}, y)}{\phi(\bar{x}) \vdash \forall y.\psi(\bar{x}, y) \qquad \mathrm{Form}(\bar{x})}$$

Observe that this derived rule, saying that the operation $\forall y$ which binds the variable y, is right adjoint to the trivial operation $*$, also takes account of the usual "book-keeping" side-conditions of the quantifier rules.

Conversely, we could instead take this adjoint rule as basic and derive the customary introduction and elimination rules from it. Indeed, the counit of the adjunction is the usual $\forall$-elimination "axiom"

$$\forall y.\psi(\bar{x}, y) \vdash \psi(\bar{x}, y).$$

It is now natural to wonder about the other quantifier; indeed, we have a further adjunction

$$\exists \dashv * \dashv \forall$$

since the following two-way rule also holds.

$$\frac{\exists y.\psi(\bar{x}, y) \vdash \phi(\bar{x})}{\psi(\bar{x}, y) \vdash \phi(\bar{x})}$$

Here the unit is the existential introduction "axiom"

$$\psi(\bar{x}, y) \vdash \exists y.\psi(\bar{x}, y).$$

It actually follows from these rules that $\exists y$ and $\forall y$ are in particular functors, that is, that $\psi \vdash \phi$ implies $\exists y.\psi \vdash \exists y.\phi$ and similarly for $\forall$.

The adjoint rules just given can therefore be used in place of the customary introduction and elimination rules, to give a complete system of deduction for quantificational logic. Many typical laws of predicate logic are just simple formal

manipulations of adjoints. For example:

$$\begin{aligned}
\psi(x,y) &\vdash \psi(x,y) \\
\psi(x,y) &\vdash \exists y.\psi(x,y) && (\text{unit of } \exists \dashv *) \\
\forall x.\psi(x,y) &\vdash \exists y.\psi(x,y) && (\text{counit of } * \dashv \forall) \\
\exists y \forall x.\psi(x,y) &\vdash \exists y.\psi(x,y) && (\exists \dashv *) \\
\exists y \forall x.\psi(x,y) &\vdash \forall x \exists y.\psi(x,y) && (* \dashv \forall)
\end{aligned}$$

The recognition of the quantifiers as adjoints also gives rise to the following *geometric interpretation*. Take any $\mathcal{L}$ structure M and consider a formula $\phi(x)$ in at most one variable x. It determines a subset,

$$[\phi(x)]^M = \{m \in M \mid M \models \phi(m)\} \subseteq M.$$

Similarly, a formula in several variables determines a subset of the cartesian product

$$[\psi(x_1, \ldots, x_n)]^M = \{(m_1, \ldots, m_n) \mid M \models \psi(m_1, \ldots, m_n)\} \subseteq M^n.$$

For instance $[x = y]^M$ is the diagonal subset $\{(m, m) \mid m \in M\} \subseteq M \times M$. Let us take two variables x, y and consider the effect of the $*$ operation on these subsets. The assignment $*[\phi(x)] = [*\phi(x)]$ determines a functor

$$* : \mathcal{P}(M) \to \mathcal{P}(M \times M).$$

Explicitly, given $[\phi(x)] \in \mathcal{P}(M)$, we have

$$*[\phi(x)] = \{(m_1, m_2) \in M \times M \mid M \models \phi(m_1)\} = \pi^{-1}([\phi(x)])$$

where $\pi : M \times M \to M$ is the first projection. Thus,

$$* = \pi^{-1}.$$

Similarly, the existential quantifier can be regarded as an operation on subsets by $\exists[\psi(x,y)] = [\exists y.\psi(x,y)]$,

$$\exists : \mathcal{P}(M \times M) \to \mathcal{P}(M).$$

Specifically, given $[\psi(x,y)] \subseteq M \times M$, we have

$$\begin{aligned}
\exists[\psi(x,y)] &= [\exists y.\psi(x,y)] \\
&= \{m \mid \text{ for some } y, M \models \psi(m,y)\} \\
&= \mathrm{im}(\pi)[\psi(x,y)].
\end{aligned}$$

Therefore,

$$\exists = \mathrm{im}(\pi).$$

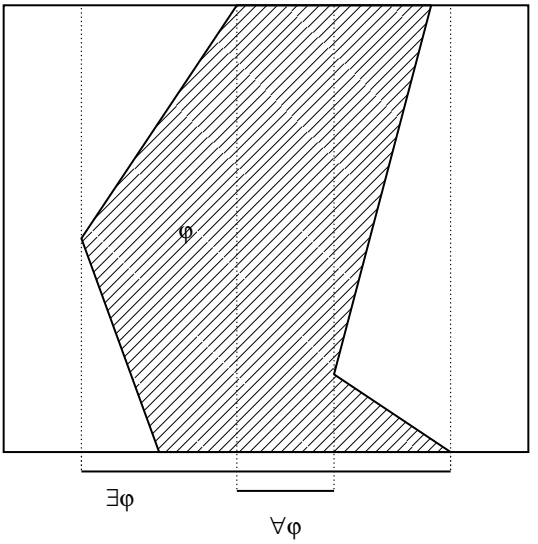


Figure 9.1: Quantifiers as adjoints

In this way, you can actually "see" the adjunction:

$$\frac{\exists y.\psi(x,y) \vdash \phi(x)}{\psi(x,y) \vdash \phi(x)}$$

It is essentially the one already considered (example 9.12) between direct and inverse images, applied to the case of a product projection $\pi : M \times M \to M$.

$$\mathrm{im}(\pi) \dashv \pi^{-1}$$

See Figure 9.1.

Finally, the universal quantifier can also be regarded as an operation of the form

$$\forall : \mathcal{P}(M \times M) \to \mathcal{P}(M)$$

by setting $\forall[\psi(x,y)] = [\forall y.\psi(x,y)]$. Then given $[\psi(x,y)] \subseteq M \times M$, we have

$$\begin{aligned} \forall[\psi(x,y)] &= [\forall y.\psi(x,y)] \\ &= \{m \mid \text{ for all } y, M \models \psi(m,y)\} \\ &= \{m \mid \pi^{-1}\{m\} \subseteq [\psi(x,y)]\} \\ &= \pi_*([\psi(x,y)]). \end{aligned}$$

Therefore,

$$\forall = \pi_*$$

so the universal quantifier is the right adjoint to pullback along the projection π. Again, in Figure 9.1 one can see the adjunction:

$$\frac{\phi(x) \leq \psi(x,y)}{\phi(x) \leq \forall y.\psi(x,y)}$$

by considering the corresponding operations induced on subsets.

9.6 RAPL

In addition to the conceptual unification achieved by recognizing constructions as different as existential quantifiers and free groups as instances of adjoints, there is the practical benefit that one then knows that these operations behave in certain ways that are common to all adjoints. We next consider one of the fundamental properties of adjoints: preservation of limits.

In the previous section we had a string of three adjoints,

$$\exists \dashv * \dashv \forall$$

and it is easy to find other such strings. For example, there is a string of four adjoints between **Cat** and **Sets**,

$$V \dashv F \dashv U \dashv R$$

where $U : \mathbf{Cat} \to \mathbf{Sets}$ is the forgetful functor to the set of objects

$$U(\mathbf{C}) = \mathbf{C}_0.$$

An obvious question in this kind of situation is "are there more?" That is, given a functor does it have an adjoint? A useful necessary condition which shows that, for example, the strings above stop is the following proposition, which is also important in its own right.

Proposition 9.14. *Right adjoints preserve limits (remember: "RAPL"!), and left adjoints preserve colimits.*

Proof. Here is the easy way: suppose we have an adjunction

$$\mathbf{C} \underset{U}{\overset{F}{\rightleftarrows}} \mathbf{D} \qquad\qquad F \dashv U$$

and we are given a diagram $C : J \to \mathbf{D}$ such that the limit $\varprojlim D_j$ exists in $\mathbf{D}$. Then for any $X \in \mathbf{C}$, we have

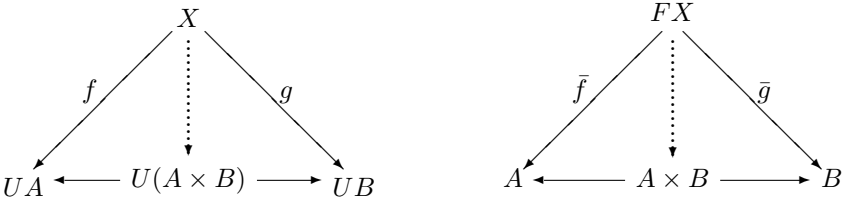
$$\begin{aligned} \mathrm{Hom}_{\mathbf{C}}(X, U(\varprojlim D_j)) &\cong \mathrm{Hom}_{\mathbf{D}}(FX, \varprojlim D_j) \\ &\cong \varprojlim \mathrm{Hom}_{\mathbf{D}}(FX, D_j) \\ &\cong \varprojlim \mathrm{Hom}_{\mathbf{C}}(X, UD_j) \\ &\cong \mathrm{Hom}_{\mathbf{C}}(X, \varprojlim UD_j) \end{aligned}$$

whence (by Yoneda), we have the required isomorphism

$$U(\varprojlim D_j) \cong \varprojlim UD_j.$$

It follows by duality that left adjoints preserve colimits. □

It is illuminating to work out what the above argument "really means" in a particular case, say binary products. Given a product $A \times B$ in $\mathbf{D}$, consider the following diagram, in which the part on the left is in $\mathbf{C}$ and that on the right in $\mathbf{D}$.



Then given any f and g as indicated, we get the required unique arrow $\langle f, g\rangle$ by adjointness as

$$\langle f, g\rangle = \overline{\langle \bar{f}, \bar{g}\rangle}$$

where we write $\bar{f}$, etc., for both directions of the adjoint correspondence.

For another example, recall that in the proof that $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ has exponentials we needed the following distributivity law for sets:

$$(\varinjlim_i X_i) \times A \;\cong\; \varinjlim_i (X_i \times A)$$

We now see that this is a consequence of the fact that the functor $(-) \times A$ is a left adjoint (namely to $(-)^A$) and therefore preserves colimits.

It also follows immediately for the propositional calculus (and in any Heyting algebra) that, for example,

$$p \Rightarrow (a \wedge b) \dashv\vdash (p \Rightarrow a) \wedge (p \Rightarrow b)$$

and

$$(a \vee b) \wedge p \dashv\vdash (a \wedge p) \vee (b \wedge p).$$

Similarly, for the quantifiers one has, for example,

$$\forall x(\phi(x) \wedge \psi(x)) \dashv\vdash \forall x\phi(x) \wedge \forall x\psi(x).$$

Since this does not hold for $\exists x$, it cannot be a right adjoint to some other "quantifier." Similarly

$$\exists x(\phi(x) \vee \psi(x)) \dashv\vdash \exists x\phi(x) \vee \exists x\psi(x).$$

And, as above, $\forall x$ cannot be a left adjoint, since it does not have this property.

The proposition gives an extremely important and useful property of adjoints. As in the foregoing examples, it can be used to show that a given functor does not have an adjoint by showing that it does not preserve (co)limits. But also, to show that a given functor preserves all (co)limits, sometimes the easiest way to

do so is to show that it has an adjoint. For example, it is very easy to recognize that the forgetful functor $U : \mathbf{Pos} \to \mathbf{Sets}$ from posets to sets has a left adjoint (what is it?). Thus, we know that limits of posets are limits of the underlying sets (suitably ordered). Dually, you may have shown "by hand" as an exercise that the coproduct of free monoids is the free monoid on the coproduct of their generating sets

$$F(A) + F(B) \cong F(A + B).$$

This now follows simply from the free $\dashv$ forgetful adjunction.

Example 9.15. Our final example of preservation of (co)limits by adjoints involves the UMP of categories of diagrams $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ mentioned in the foregoing chapter. For a small category $\mathbf{C}$, a contravariant functor $P : \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$ is often called a *presheaf* on $\mathbf{C}$, and the functor category $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ is accordingly called the *category of presheaves* on $\mathbf{C}$, sometimes written $\hat{\mathbf{C}}$. This cocomplete category is the "free cocompletion" of $\mathbf{C}$ in the following sense.

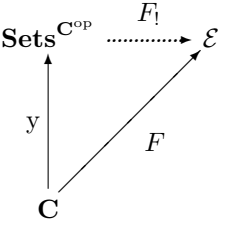
Proposition 9.16. *For any small category* $\mathbf{C}$*, the Yoneda embedding*

$$\mathrm{y} : \mathbf{C} \to \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$$

has the following universal mapping property: given any cocomplete category $\mathcal{E}$ *and functor* $F : \mathbf{C} \to \mathcal{E}$*, there is a colimit preserving functor* $F_! : \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}} \to \mathcal{E}$ *such that,*

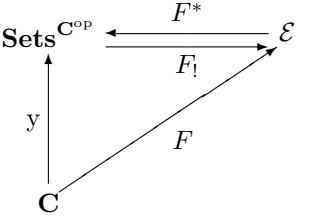
$$F_! \circ \mathrm{y} \cong F \tag{9.5}$$

as indicated in the following diagram.



Moreover, up to natural isomorphism, $F_!$ *is the unique cocontinuous functor with this property.*

Proof. We will show that there are adjoint functors,



$F_! \dashv F^*$

with $F_! \circ \mathrm{y} \cong F$. It then follows that $F_!$ preserves all colimits. To define $F_!$, take any presheaf $P \in \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ and write it as a canonical colimit of representables

$$\varinjlim_{j \in J} \mathrm{y}C_j \;\cong\; P$$

with $J = \int_{\mathbf{C}} P$ the category of elements of P, as in proposition 8.10. Then set

$$F_!(P) = \varinjlim_{j \in J} FC_j$$

with the colimit taken in $\mathcal{E}$, which is cocomplete. (We leave it to the reader to determine how to define $F_!$ on arrows.) Clearly, if $F_!$ is to preserve all colimits and satisfy (9.5), then up to isomorphism this must be its value for P. For F^*, take any $E \in \mathcal{E}$ and $C \in \mathbf{C}$ and observe that by (Yoneda and) the intended adjunction, for $F^*(E)(C)$ we must have

$$\begin{aligned} F^*(E)(C) &\cong \mathrm{Hom}_{\hat{C}}(\mathrm{y}C, F^*(E)) \\ &\cong \mathrm{Hom}_{\mathcal{E}}(F_!(\mathrm{y}C), E) \\ &\cong \mathrm{Hom}_{\mathcal{E}}(FC, E). \end{aligned}$$

Thus we simply set

$$F^*(E)(C) = \mathrm{Hom}_{\mathcal{E}}(FC, E)$$

which is plainly a presheaf on $\mathbf{C}$ (we use here that $\mathcal{E}$ is locally small). Now let us check that indeed $F_! \dashv F^*$. For any $E \in \mathcal{E}$ and $P \in \hat{\mathbf{C}}$, we have natural isomorphisms

$$\begin{aligned} \mathrm{Hom}_{\hat{\mathbf{C}}}(P, F^*(E)) &\cong \mathrm{Hom}_{\hat{\mathbf{C}}}(\varinjlim_{j \in J} \mathrm{y}C_j \,, F^*(E)) \\ &\cong \varprojlim_{j \in J} \mathrm{Hom}_{\hat{\mathbf{C}}}(\mathrm{y}C_j \,, F^*(E)) \\ &\cong \varprojlim_{j \in J} F^*(E)(C_j) \\ &\cong \varprojlim_{j \in J} \mathrm{Hom}_{\mathcal{E}}(FC_j, E) \\ &\cong \mathrm{Hom}_{\mathcal{E}}(\varinjlim_{j \in J} FC_j, E) \\ &\cong \mathrm{Hom}_{\mathcal{E}}(F_!(P), E). \end{aligned}$$

Finally, for any object $C \in \mathbf{C}$,

$$F_!(\mathrm{y}C) = \varinjlim_{j \in J} FC_j \cong\; FC$$

since the category of elements J of a representable $\mathrm{y}C$ has a terminal object, namely the element $1_C \in \mathrm{Hom}_{\mathbf{C}}(C, C)$. □

Corollary 9.17. *Let $f : \mathbf{C} \to \mathbf{D}$ be a functor between small categories. The precomposition functor*

$$f^* : \mathbf{Sets}^{\mathbf{D}^{\mathrm{op}}} \to \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$$

given by

$$f^*(Q)(C) = Q(fC)$$

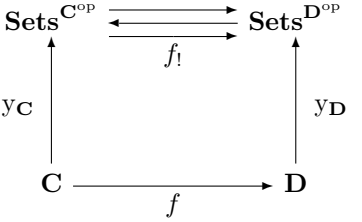
has both left and right adjoints

$$f_! \vdash f^* \vdash f_*$$

Moreover, there is a natural isomorphism

$$f_! \circ \mathrm{y}_{\mathbf{C}} \;\cong\; \mathrm{y}_{\mathbf{D}} \circ f$$

as indicated in the following diagram.

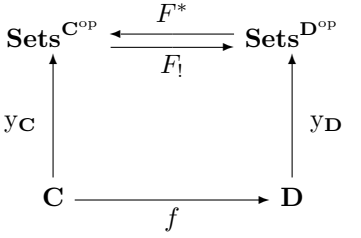


The induced functors $f_!$ and f_* are sometimes referred to in the literature as *(left and right) Kan extensions.*

Proof. First, define

$$F = \mathrm{y}_{\mathbf{D}} \circ f : \mathbf{C} \to \mathbf{Sets}^{\mathbf{D}^{\mathrm{op}}}.$$

Then by the foregoing proposition we have adjoints $F_!$ and F^* as indicated in



and we know that $F_! \circ \mathrm{y}_{\mathbf{C}} \cong \mathrm{y}_D \circ f$. We claim that $F^* \cong f^*$. Indeed, by the definition of F^* we have

$$F^*(Q)(C) = \mathrm{Hom}_{\hat{\mathbf{D}}}(FC, Q) \cong \mathrm{Hom}_{\hat{\mathbf{D}}}(\mathrm{y}(fC), Q) \cong Q(fC) = f^*(Q)(C).$$

This, therefore, gives the functors $f_! \dashv f^*$. For f_*, apply the proposition to the composite

$$f^* \circ \mathrm{y}_{\mathbf{D}} : \mathbf{D} \to \mathbf{Sets}^{\mathbf{D}^{\mathrm{op}}} \to \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}.$$

This gives an adjunction

$$(f^* \circ \mathrm{y}_{\mathbf{D}})_! \dashv (f^* \circ \mathrm{y}_{\mathbf{D}})^*$$

so we just need to show that

$$(f^* \circ \mathrm{y}_{\mathbf{D}})_! \cong f^*$$

in order to get the required right adjoint as $f_* = (f^* \circ \mathrm{y}_{\mathbf{D}})^*$. By the universal property of $\mathbf{Sets}^{\mathbf{D}^{\mathrm{op}}}$, it suffices to show that f^* preserves colimits. But for any colimit $\varinjlim_j Q_j$ in $\mathbf{Sets}^{\mathbf{D}^{\mathrm{op}}}$

$$\begin{aligned}(f^*(\varinjlim_j Q_j))(C) &\cong (\varinjlim_j Q_j)(fC) \\ &\cong \varinjlim_j (Q_j(fC)) \\ &\cong \varinjlim_j ((f^* Q_j)(C)) \\ &\cong (\varinjlim_j (f^* Q_j))(C).\end{aligned}$$

□

This corollary says that, in a sense, *every functor has an adjoint*! For, given any $f : \mathbf{C} \to \mathbf{D}$, we indeed have the right adjoint

$$f^* \circ \mathrm{y}_{\mathbf{D}} : \mathbf{D} \to \hat{\mathbf{C}}$$

except that its values are in the "ideal elements" of the cocompletion $\hat{\mathbf{C}}$.

9.7 Locally cartesian closed categories

A special case of the situation described by corollary 9.17 is the *change of base* for indexed families of sets along a "reindexing" function $\alpha : J \to I$. An arbitrary such function between sets gives rise, by that corollary, to a triple of adjoint functors:

$$\mathbf{Sets}^J \underset{\alpha_!}{\overset{\alpha_*}{\underset{\longleftarrow}{\overset{\longrightarrow}{\longrightarrow}}}} \mathbf{Sets}^I \qquad \alpha_! \dashv \alpha^* \dashv \alpha_*$$

Let us examine these functors more closely in this special case.

An object A of $\mathbf{Sets}^I$ is an I-indexed family of sets

$$(A_i)_{i\in I}.$$

Then $\alpha^*(A) = A \circ \alpha$ is the reindexing of A along α to a J-indexed family of sets

$$\alpha^*(A) = (A_{\alpha(j)})_{j\in J}.$$

Given a J-indexed family B, let is calculate $\alpha_!(B)$ and $\alpha_*(B)$.

Consider first the case $I = 1$ and $\alpha = !_J : J \to 1$. Then $(!_J)^* : \mathbf{Sets} \to \mathbf{Sets}^J$ is the "constant family" or diagonal functor $\Delta(A)(j) = A$, for which we know the adjoints:

$$\mathbf{Sets}^J \;\underset{\Sigma}{\overset{\Pi}{\underset{\longleftarrow}{\overset{\longrightarrow}{\Delta}}}}\; \mathbf{Sets}$$

$$\Sigma \dashv \Delta \dashv \Pi$$

These are, namely, just the (disjoint) sum and cartesian product of the sets in the family

$$\sum_{j\in J} B_j, \qquad \prod_{j\in J} B_j.$$

For recall that we have the adjunctions:

$$\frac{\vartheta_j : B_j \to A}{(\vartheta_j) : \sum_j B_j \to A}, \qquad \frac{\vartheta_j : A \to B_j}{\langle \vartheta_j \rangle : A \to \prod_j B_j}$$

By uniqueness of adjoints, it therefore follows that $(!_J)_! \cong \Sigma$ and $(!_J)_* \cong \Pi$.

A general reindexing $\alpha : J \to I$ gives rise to generalized sum and product operations along α

$$\Sigma_\alpha \dashv \alpha^* \dashv \Pi_\alpha$$

defined on J-indexed families (B_j) by

$$(\Sigma_\alpha(B_j))_i = \sum_{\alpha(j)=i} B_j$$

$$(\Pi_\alpha(B_j))_i = \prod_{\alpha(j)=i} B_j.$$

These operations thus assign to an element $i \in I$ the sum, respectively the product, over all the sets indexed by the elements j in the preimage of i under α.

Now let us recall from example 7.28 the equivalence between J-indexed families of sets and the slice category of "sets over J"

$$\mathbf{Sets}^J \;\simeq\; \mathbf{Sets}/J.$$

It takes a family $(A_j)_{j\in J}$ to the indexing projection $p : \sum_{j\in J} A_j \to J$ and a map $\pi : A \to J$ to the family $(\pi^{-1}(j))_{j\in J}$. We know, moreover, from an exercise in Chapter 7 that this equivalence respects reindexing, in the sense that for any $\alpha : J \to I$ the following square commutes up to natural isomorphism.

$$\begin{array}{ccc} J & & \mathbf{Sets}/J \xrightarrow[\simeq]{} \mathbf{Sets}^J \\ \alpha \downarrow & & \alpha^\sharp \uparrow \qquad\quad \uparrow \alpha^* \\ I & & \mathbf{Sets}/I \xrightarrow[\simeq]{} \mathbf{Sets}^I \end{array}$$

Here we write $\alpha^\sharp$ for the pullback functor along α. Since α^* has both right and left adjoints, we have the diagram of induced adjoints:

$$\begin{array}{ccc} J & & \mathbf{Sets}/J \xrightarrow[\simeq]{} \mathbf{Sets}^J \\ \alpha \downarrow & & \alpha_L \downarrow \alpha^\sharp \uparrow \alpha_\sharp \downarrow \qquad \alpha_! \uparrow \alpha^* \downarrow \alpha_* \\ I & & \mathbf{Sets}/I \xrightarrow[\simeq]{} \mathbf{Sets}^I \end{array}$$

Proposition 9.18. *For any function $\alpha : J \to I$, the pullback functor $\alpha^\sharp : \mathbf{Sets}/I \to \mathbf{Sets}/J$ has both left and right adjoints:*

$$\alpha_L \dashv \alpha^\sharp \dashv \alpha_\sharp$$

In particular, $\alpha^\sharp$ therefore preserves all limits and colimits.

Let us compute the functors explicitly. Given $\pi : A \to J$, let $A_j = \pi^{-1}(j)$ and recall that

$$\alpha_!(A)_i = \sum_{\alpha(j)=i} A_i.$$

But then we have

$$\begin{aligned} \alpha_!(A)_i &= \sum_{\alpha(j)=i} A_i \\ &= \sum_{i\in\alpha^{-1}(j)} A_i \\ &= \sum_{i\in\alpha^{-1}(j)} \pi^{-1}(j) \\ &= \pi^{-1} \circ \alpha^{-1}(j) \\ &= (\alpha \circ \pi)^{-1}(i). \end{aligned}$$

It follows that $\alpha_L(\pi : A \to J)$ is simply the composite $\alpha \circ \pi : A \to J \to I$,

$$\alpha_L(\pi : A \to J) = (\alpha \circ \pi : A \to J \to I).$$

Indeed, it is easy to check directly that composition along any function α is left adjoint to pullback along α.

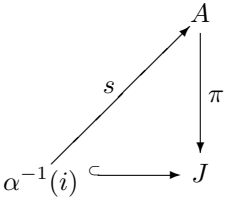
As for the right adjoint

$$\alpha_\sharp : \mathbf{Sets}/J \longrightarrow \mathbf{Sets}/I$$

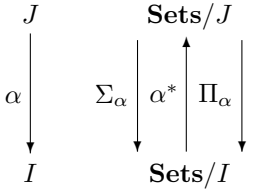
given $\pi : A \to J$, the result $\alpha_\sharp(\pi) : \alpha_\sharp(A) \to I$ can be described fiberwise by

$$(\alpha_\sharp(A))_i = \{s : \alpha^{-1}(i) \to A \mid \text{“}s \text{ is a partial section of } \pi\text{”}\}$$

where the condition “s is a partial section of π” means that the following triangle commutes with the canonical inclusion $\alpha^{-1}(i) \subseteq J$ at the base.

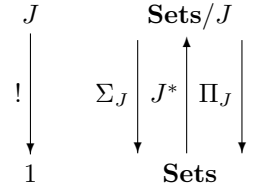


Henceforth, we shall also write these “change of base” adjoints along a map $\alpha : J \to I$ in the form:



$$\Sigma_\alpha \dashv \alpha^* \dashv \Pi_\alpha$$

Finally, let us reconsider the case $I = 1$, where these adjoints take the form:



$$\Sigma_J \dashv J^* \dashv \Pi_J$$

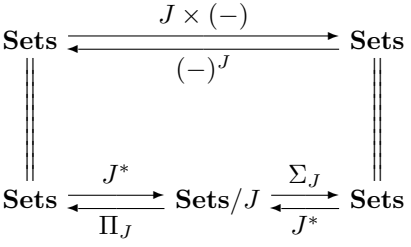
In this case we have

$$\begin{aligned} \Sigma_J(\pi : A \to J) &= A \\ J^*(A) &= (p_1 : J \times A \to J) \\ \Pi_J(\pi : A \to J) &= \{s : J \to A \mid \pi \circ s = 1\} \end{aligned}$$

as the reader can easily verify. Moreover, one therefore has

$$\begin{aligned} \Sigma_J J^*(A) &= J \times A \\ \Pi_J J^*(A) &= A^J. \end{aligned}$$

Thus the product $\dashv$ exponential adjunction can be factored as a composite of adjunctions as follows:



The following definition captures the notion of a category having this sort of adjoint structure. In such a category $\mathcal{E}$, the slice categories can be regarded as categories of abstract indexed families of objects of $\mathcal{E}$, and the reindexing of such families can be carried out, and it has associated adjoint operations of sum and product.

Definition 9.19. A category $\mathcal{E}$ is called *locally cartesian closed* if $\mathcal{E}$ has a terminal object and for every arrow $f : A \to B$ in $\mathcal{E}$, the composition functor

$$\Sigma_f : \mathcal{E}/A \to \mathcal{E}/B$$

has a right adjoint f^* which, in turn, has a right adjoint Π_f:

$$\Sigma_f \dashv f^* \dashv \Pi_f$$

The choice of name for such categories is explained by the following important fact.

Proposition 9.20. *For any category $\mathcal{E}$ with a terminal object, the following are equivalent.*

1. *$\mathcal{E}$ is locally cartesian closed.*
2. *Every slice category $\mathcal{E}/A$ of $\mathcal{E}$ is cartesian closed.*

Proof. Let $\mathcal{E}$ be locally cartesian closed. Since $\mathcal{E}$ has a terminal object, products and exponentials in $\mathcal{E}$ can be built as:

$$\begin{aligned} A \times B &= \Sigma_B B^* A \\ B^A &= \Pi_B B^* A \end{aligned}$$

So $\mathcal{E}$ is cartesian closed. But clearly every slice category $\mathcal{E}/X$ is also locally cartesian closed, since "a slice of a slice is a slice." Thus every slice of $\mathcal{E}$ is cartesian closed.

Conversely, suppose every slice of $\mathcal{E}$ is cartesian closed. Then $\mathcal{E}$ has pullbacks, since these are just binary products in a slice. Thus, we just need to construct the "relative product" functor $\Pi_f : \mathcal{E}/A \to \mathcal{E}/B$ along a map $f : A \to B$. First, change notation:

$$\begin{aligned} \mathcal{F} &= \mathcal{E}/B \\ F &= f : A \to B \\ \mathcal{F}/F &= \mathcal{E}/A \end{aligned}$$

Thus, we want to construct $\Pi_F : \mathcal{F}/F \to F$. Given an object $p : X \to F$ in $\mathcal{F}/F$, the object $\Pi_F(p)$ is constructed as the following pullback,

$$\begin{array}{ccc} \Pi_F(p) & \longrightarrow & X^F \\ \downarrow & & \downarrow{\scriptstyle p^F} \\ 1 & \xrightarrow[\widetilde{1_F}]{} & F^F \end{array}$$

where $\widetilde{1_F}$ is the exponential transpose of the composite arrow

$$1 \times F \cong F \xrightarrow{1} F.$$

We leave it to the reader to recognize that there is then a natural bijection of the form:

$$\frac{Y \to \Pi_F(p)}{F^*Y \to p}$$

□

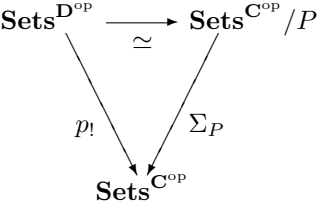
Remark 9.21. The reader should be aware that some authors do not require the existence of a terminal object in the definition of a locally cartesian closed category.

Example 9.22 (Presheaves). For any small category $\mathbf{C}$, the category $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ of presheaves on $\mathbf{C}$ is locally cartesian closed. This is a consequence of the following fact.

Lemma 9.23. *For any object $P \in \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$, there is a small category $\mathbf{D}$ and an equivalence of categories,*

$$\mathbf{Sets}^{\mathbf{D}^{\mathrm{op}}} \simeq \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}/P.$$

Moreover, there is also a functor $p : \mathbf{D} \to \mathbf{C}$ such that the following diagram commutes (up to natural isomorphism).



Proof. One can take:

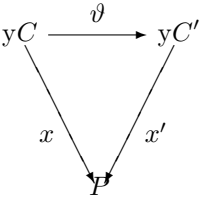
$$\mathbf{D} = \int_{\mathbf{C}} P$$

$$p = \pi : \int_{\mathbf{C}} P \to \mathbf{C}$$

Indeed, recall that by the Yoneda lemma, $\int_{\mathbf{C}} P$ can be described equivalently (isomorphically, in fact) as the category that we shall write suggestively as y/P, described as follows:

objects pairs (C, x) where $C \in \mathbf{C}$ and $x : \mathrm{y}C \to P$ in $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$

arrows all arrows between such objects in the slice category over P



Note that by Yoneda, each such arrow is of the form $\vartheta = \mathrm{y}h$ for a unique $h : C \to D$ in $\mathbf{C}$, which, moreover, is such that $P(h)(x') = x$.

Now let $I : \mathrm{y}/P \to \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}/P$ be the evident (full and faithful) inclusion functor, and define a functor

$$\Phi : \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}/P \to \mathbf{Sets}^{(\mathrm{y}/P)^{\mathrm{op}}}$$

by setting, for any $q : Q \to P$ and $(C, x) \in \mathrm{y}/P$

$$\Phi(q)(C, x) = \mathrm{Hom}_{\hat{\mathbf{C}}/P}(x, q).$$

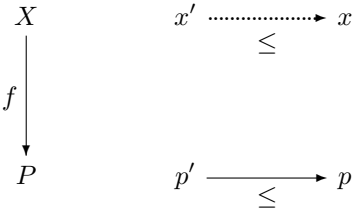
In other words, $\Phi(q) = I^*(\mathrm{y}q)$, which is plainly functorial. We leave it to the reader as an exercise to show that this functor establishes an equivalence of categories. □

Corollary 9.24. *For any small category* $\mathbf{C}$*, the category* $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ *of presheaves on* $\mathbf{C}$ *is locally cartesian closed.*

Example 9.25 Fibrations of posets. A monotone map of posets $f : X \to P$ is a (discrete) *fibration* if it has the following *lifting property*:

For every $x \in X$ and $p' \leq fx$ there is a unique $x' \leq x$ such that $f(x') = p'$.

One says that x "lies over" $p = f(x)$ and that any $p' \leq p$ "lifts" to a unique $x' \leq x$ lying over it, as indicated in the following diagram.



The identity morphism of a given poset P is clearly a fibration, and the composite of two fibrations is easily seen to be a fibration. Let **Fib** denote the (non-full) subcategory of posets and fibrations between them as arrows.

Lemma 9.26. *For any poset* P*, the slice category* $\mathbf{Fib}/P$ *is cartesian closed.*

Proof. The category $\mathbf{Fib}/P$ is equivalent to the category of presheaves on P,

$$\mathbf{Fib}/P \;\simeq\; \mathbf{Sets}^{P^{\mathrm{op}}}.$$

To get a functor $\Phi : \mathbf{Fib}/P \to \mathbf{Sets}^{P^{\mathrm{op}}}$, take a fibration $q : Q \to P$ to the presheaf defined on objects by

$$\Phi(q)(p) = q^{-1}(p) \qquad \text{for } p \in P.$$

The lifting property then determines the action on arrows $p' \leq p$. For the other direction $\Psi : \mathbf{Sets}^{P^{\mathrm{op}}} \to \mathbf{Fib}/P$ take a presheaf $Q : P^{\mathrm{op}} \to \mathbf{Sets}$ to (the indexing projection of) its category of elements,

$$\Psi(Q) = \int_P Q \xrightarrow{\pi} P.$$

These are easily seen to be quasi-inverses. □

The category **Fib** itself is *almost* locally cartesian closed; it only lacks a terminal object (why?). We can "fix" this simply by slicing it.

Corollary 9.27. *For any poset* P*, the slice category* $\mathbf{Fib}/P$ *is locally cartesian closed.*

This sort of case is not uncommon, which is why the notion "locally cartesian closed" is often formulated without requiring a terminal object.

9.8 Adjoint functor theorem

The question we now want to consider systematically is, when does a functor have an adjoint? Consider first the question, when does a functor of the form $\mathbf{C} \to \mathbf{Sets}$ have a left adjoint? If $U : \mathbf{C} \to \mathbf{Sets}$ has $F \dashv U$, then U is representable $U \cong \mathrm{Hom}(F1, -)$, since $U(C) \cong \mathrm{Hom}(1, UC) \cong \mathrm{Hom}(F1, C)$.

A related necessary condition that makes sense for categories other than **Sets** is preservation of limits. Suppose that $\mathbf{C}$ is complete and $U : \mathbf{C} \to \mathbf{X}$ preserves limits; then we can ask whether U has a left adjoint. The *Adjoint Functor Theorem* (AFT) gives a necessary and sufficient condition for this case.

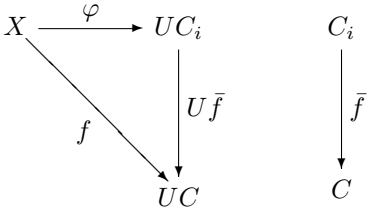
Theorem 9.28 (Freyd). *Let* $\mathbf{C}$ *be locally small and complete. Given any category* $\mathbf{X}$ *and functor*

$$U : \mathbf{C} \to \mathbf{X}$$

the following are equivalent:

1. *U has a left adjoint.*
2. *U preserves limits, and for each $X \in \mathbf{X}$ the functor U satisfies the following: Solution set condition: There exists a set of objects $(C_i)_{i \in I}$ in $\mathbf{C}$ such that for any object $C \in \mathbf{C}$ and arrow $f : X \to UC$, there exists an $i \in I$ and arrows $\varphi : X \to UC_i$ and $\bar{f} : C_i \to C$ such that:*

$$f = U(\bar{f}) \circ \varphi$$



Briefly: "every arrow $X \to UC$ factors through some object C_i in the solution set."

For the proof, we require the following.

Lemma 9.29. *Let* $\mathbf{D}$ *be locally small and complete. Then the following are equivalent:*

1. $\mathbf{D}$ *has an initial object.*
2. $\mathbf{D}$ *satisfies the following:*

   Solution set condition: There is a set of objects $(D_i)_{i \in I}$ in $\mathbf{D}$ such that for any object $D \in \mathbf{C}$ there is an arrow $D_i \to D$ for some $i \in I$.

Proof. If $\mathbf{D}$ has an initial object 0, then $\{0\}$ is obviously a solution set.

Conversely, suppose we have a solution set $(D_i)_{i\in I}$ and consider the object

$$W = \prod_{i\in I} D_i.$$

Then W is "weakly initial" in the sense that for any object D there is a (not necessarily unique) arrow $W \to D$, namely the composite

$$\prod_{i\in I} D_i \to D_i \to D$$

for a suitable product projection $\prod_{i\in I} D_i \to D_i$. Now take the joint equalizer of all endomorphisms $d : W \to W$ (which is a set, since $\mathbf{D}$ is locally small), as indicated in the diagram:

$$V \xrightarrow{\ h\ } W \underset{\langle d\rangle}{\overset{\Delta}{\rightrightarrows}} \prod_{d:W\to W} W$$

Here the arrows Δ and $\langle d\rangle$ have the d-projections $1_W : W \to W$ and $d : W \to W$, respectively. This equalizer then has the property that for any endomorphism $d : W \to W$,

$$d \circ h = h. \tag{9.6}$$

Note, moreover, that V is still weakly initial, since for any D there is an arrow $V \rightarrowtail W \to D$. Suppose that for some D there are two arrows $f, g : V \to D$. Take their equalizer $e : U \to V$, and consider the following diagram,

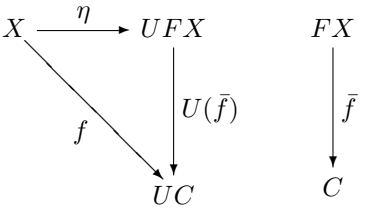
$$\begin{array}{ccccc} U & \xrightarrow{\ e\ } & V & \underset{g}{\overset{f}{\rightrightarrows}} & D \\ {\scriptstyle s}\uparrow & & \downarrow{\scriptstyle h} & & \\ W & \xrightarrow[\ hes\ ]{} & W & & \end{array}$$

in which the arrow s comes from W being weakly initial. So for the endomorphism hes by (9.6) we have

$$hesh = h.$$

Since h is monic, $esh = 1_V$. But then $eshe = e$, and so also $she = 1_U$ since e is monic. Therefore $U \cong V$, and so $f = g$. Thus V is an initial object. □

Proof. (Theorem) If U has a left adjoint $F \dashv U$, then $\{FX\}$ is itself a solution set for X, since we always have a factorization,

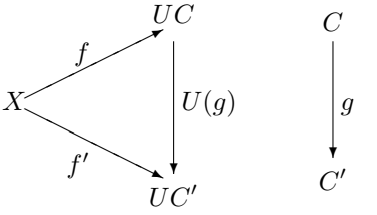


where $\bar{f} : FX \to C$ is the adjoint transpose of f and $\eta : X \to UFX$ the unit of the adjunction.

Conversely, consider the following so-called *comma-category* $(X|U)$, with:

objects: are pairs (C, f) with $f : X \to UC$

arrows: $g : (C, f) \to (C', f')$ are arrows $g : C \to C'$ with $f' = U(g)f$.



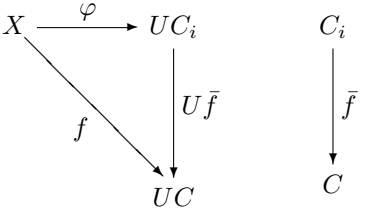
Clearly U has a left adjoint F iff for each object X this category $(X|U)$ has an initial object, $(FX, \eta : X \to UFX)$, which then has the UMP of the unit. Thus, to use the foregoing initial object lemma, we must check:

1. $(X|U)$ is locally small.
2. $(X|U)$ satisfies the solution set condition in the lemma.
3. $(X|U)$ is complete.

For (1), we just observe that $\mathbf{C}$ is locally small. For (2), the solution set condition of the theorem implies that there is a set of objects,

$$\{(C_i, \varphi : X \to UC_i) \mid i \in I\}$$

such that every object $(C, f : X \to UC)$ has an arrow $\bar{f} : (C_i, \varphi) \to (C, f)$.



Finally, to see that $(X|U)$ is complete, one can check directly that it has products and equalizers, because U preserves these. We leave this as an easy exercise for the reader. □

Remark 9.30. 1. The theorem simply does not apply if $\mathbf{C}$ is not complete. In that case, a given functor may have an adjoint, but the AFT will not tell us that.

2. It is essential that the solution set in the theorem be a *set* (and that $\mathbf{C}$ have all set-sized limits).
3. On the other hand, if $\mathbf{C}$ is itself *small* and complete, then we can plainly drop the solution set condition entirely. In that case we have the following.

Corollary 9.31. *If* $\mathbf{C}$ *is a small and complete category and* $U : \mathbf{C} \to \mathbf{X}$ *is a functor that preserves all limits, then* U *has a left adjoint.*

Example 9.32. For complete posets P, Q, a monotone function $f : P \to Q$ has a right adjoint $g : Q \to P$ iff f is *cocontinuous*, in the sense that $f(\bigvee_i p_i) = \bigvee_i f(p_i)$ for any set-indexed family of elements $(p_i)_{i \in I}$. (Of course, here we are using the dual formulation of the AFT.)

Indeed, we can let

$$g(q) = \bigvee_{f(x) \leq q} x.$$

Then for any $p \in P$ and $q \in Q$, if

$$p \leq g(q)$$

then

$$f(p) \leq fg(q) = f(\bigvee_{f(x) \leq q} x) = \bigvee_{f(x) \leq q} f(x) \leq q.$$

While, conversely, if

$$f(p) \leq q$$

then clearly

$$p \leq \bigvee_{f(x) \leq q} x = g(q).$$

As a further consequence of the AFT, we have the following characterization of representable functors on small complete categories.

Corollary 9.33. *If* $\mathbf{C}$ *is a small and complete category, then for any functor* $U : \mathbf{C} \to \mathbf{Sets}$ *the following are equivalent:*

1. *U preserves all limits.*
2. *U has a left adjoint.*
3. *U is representable.*

Proof. Immediate. □

These corollaries are, however, somewhat weaker than it may at first appear, in light of the following fact:

Proposition 9.34. *If* $\mathbf{C}$ *is small and complete, then* $\mathbf{C}$ *is a preorder.*

Proof. Suppose not, and take $C, D \in \mathbf{C}$ with $\mathrm{Hom}(C, D) \geq 2$. Let J be any set, and take the product

$$\prod_J D.$$

There are isomorphisms:

$$\mathrm{Hom}(C, \prod_J D) \cong \prod_J \mathrm{Hom}(C, D) \cong \mathrm{Hom}(C, D)^J$$

So, for the cardinalities of these sets, we have

$$|\,\mathrm{Hom}(C, \prod_J D)| = |\,\mathrm{Hom}(C, C')|^{|J|} \geq 2^{|J|} = |P(J)|.$$

And that is for *any* set J. On the other hand, clearly $|\mathbf{C}_1| \geq |\,\mathrm{Hom}(C, \prod_J D)|$. So taking $J = \mathbf{C}_1$ in the above gives a contradiction. □

Remark 9.35. An important special case of the AFT that often occurs "in nature" is that in which the domain category satisfies certain conditions that eliminate the need for the (rather unpleasant!) solution set condition entirely. Specifically, let $\mathbf{A}$ be a locally small, complete category satisfying the following conditions:

1. $\mathbf{A}$ is *well powered*: each object A has at most a *set* of subobjects $S \rightarrowtail A$.
2. $\mathbf{A}$ has a *cogenerating set*: there is a set of objects $\{A_i \mid i \in I\}$ (I some index *set*), such that for any A, X and $x \neq y : X \rightrightarrows A$ in $\mathbf{A}$ there is some $s : A \to A_i$ (for some i) that "separates" x and y, in the sense that $sx \neq sy$.

Then any functor $U : \mathbf{A} \to \mathbf{X}$ that preserves limits necessarily has a left adjoint. In this form (also originally proved by Freyd), the theorem is usually known as the Special Adjoint Functor Theorem ("SAFT"). We refer to Mac Lane, V.8 for the proof, and some sample applications.

Example 9.36. An important application of the AFT is that any *equational theory* T gives rise to a free $\dashv$ forgetful adjunction between sets and the category of models of the theory, or "T-algebras." In somewhat more detail, let T be a (finitary) equational theory, consisting of finitely many operation symbols, each of some finite arity (including 0-ary operations, that is, constant symbols), and a set of equations between terms built from these operations and variables. For instance, the theory of groups has a constant u (the group unit), a unary operation g^{-1} (the inverse), and a binary operation $g \cdot h$ (the group product), and

a handful of equations such as $g \cdot u = g$. The theory of rings has a further binary operation and some more equations. The theory of fields is not equational, however, because the condition $x \neq 0$ is required for an element x to have a multiplicative inverse. A T-algebra is a set equipped with operations (of the proper arities) corresponding to the operation symbols in T, and satisfying the equations of T. A homomorphism of T-algebras $h : A \to B$ is a function on the underlying sets that preserves all the operations, in the usual sense. Let T-**Alg** be the category of all such algebras and their homomorphisms. There is an evident forgetful functor

$$U : T\text{-}\mathbf{Alg} \to \mathbf{Sets}$$

The AFT implies that this functor has a left adjoint F, the "free algebra" functor.

Proposition 9.37. *For any equational theory T, the forgetful functor to* **Sets** *has a left adjoint.*

Rather than proving this general proposition (for which see Mac Lane, chapter V), it will be more illuminating to do a simple example.

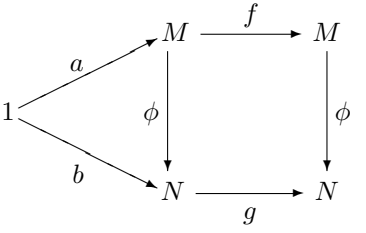
Example 9.38. Let T be the theory with one constant and one unary operation (no axioms). A T-algebra is a set M with the structure:

$$1 \xrightarrow{a} M \xrightarrow{f} M$$

If $1 \xrightarrow{b} N \xrightarrow{g} N$ is another such algebra, a homomorphism of T-algebras $\phi : (M, a, f) \to (N, b, g)$ is a function $\phi : M \to N$ that preserves the element and the operation, in the expected sense that

$$\phi a = b$$
$$\phi f = g\phi.$$

as indicated in the commutative diagram:



There is an evident forgetful functor (forget the structure):

$$U : T\text{-}\mathbf{Alg} \to \mathbf{Sets}$$

This functor is easily seen to create all limits, as is the case for algebras for any theory T. So in particular, T-**Alg** is complete and U preserves limits. Thus in order to apply the AFT we just need to check the solution set condition.

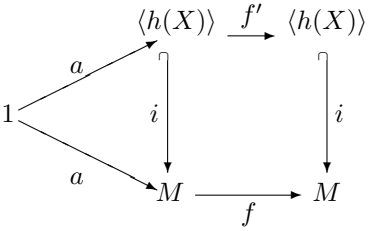
To that end, let X be any set and take any function

$$h : X \to M.$$

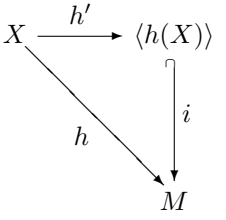
The image $h(X) \subseteq M$ generates a sub-T-model of (M, a, f) as follows. Define the set "generated by $h(X)$" to be

$$\langle h(X) \rangle = \{ f^n(z) \mid n \in \mathbb{N}, z = a \text{ or } z = h(x) \text{ for some } x \in X \}.$$

Then $a \in \langle h(X) \rangle$, and f restricts to $\langle h(X) \rangle$ to give a function $f' : \langle h(X) \rangle \to \langle h(X) \rangle$. Moreover, the inclusion $i : \langle h(X) \rangle \to M$ is a T-algebra homomorphism



Furthermore, since $h(X) \subseteq \langle h(X) \rangle$ there is a factorization h' of h, as indicated in the following diagram.



Now observe that, given X, the cardinality $|\langle h(X) \rangle|$ is *bounded*, that is, for a sufficiently large κ independent of h and M, we have

$$|\langle h(X) \rangle| \leq \kappa.$$

Indeed, we can take $\kappa = (|X| + 1) \times |\mathbb{N}|$.

To find a solution set for X, let us now take one representative N of each isomorphism class of T-algebras with cardinality at most κ. The set of all such algebras N is then a solution set for X and U. Indeed, as we just showed, any function $h : X \to M$ factors through an element of this set (namely an isomorphic copy N of $\langle h(X) \rangle$). By the AFT, there thus exists a free functor,

$$F : \mathbf{Sets} \to T\text{-}\mathbf{Alg}.$$

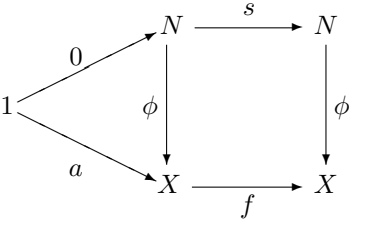
A similar argument works for any equational theory T.

Finally, let us consider the particular free model $F(\emptyset)$ in T-**Alg**. Since left adjoints preserve colimits, this is an initial object. It follows that $F(\emptyset)$ is a *natural numbers object*, in the following sense:

Definition 9.39. Let $\mathbf{C}$ be a category with a terminal object 1. A *natural numbers object* (NNO) in $\mathbf{C}$ is a structure of the form

$$1 \xrightarrow{0} N \xrightarrow{s} N$$

which is initial among all such structures. Precisely, given any $1 \xrightarrow{a} X \xrightarrow{f} X$ in $\mathbf{C}$ there is a unique arrow $\phi : N \to X$ such that the following commutes.



In other words, given any object X, a "starting point" $a \in X$ and an operation $x \mapsto f(x)$ on X, we can build up a unique $\phi : N \to X$ *recursively* by the equations:

$$\phi(0) = a$$
$$\phi(s(n)) = f(\phi(n)) \quad \text{for all } n \in N$$

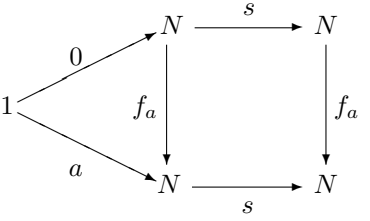
Thus, the UMP of a natural numbers object says precisely that such an object supports recursive definitions. It is easy to show that the set $\mathbb{N}$ of natural numbers with the canonical structure of 0 and the "successor function" $s(n) = n+1$ is an NNO, and that any NNO in **Sets** is isomorphic to it. The characterization of $\mathbb{N}$ in terms of the UMP of recursive definitions is therefore equivalent to the usual logical definition using the Peano axioms in **Sets**. But note that the notion of an NNO also makes sense in many categories where the Peano axioms do not make any sense.

Let us consider some simple examples of recursively defined functions using this universal mapping property.

Example 9.40. 1. Let $(N, 0, s)$ be an NNO in a category $\mathbf{C}$. Take any point $a : 1 \to N$, and consider the new structure:

$$1 \xrightarrow{a} N \xrightarrow{s} N$$

Then by the universal property of the NNO, there is a unique morphism $f_a : N \to N$ such that the following commutes.



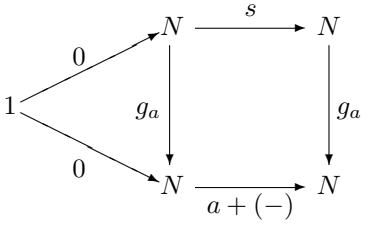
Thus we have the following "recursion equations":

$$f_a(0) = a$$
$$f_a(s(n)) = s(f_a(n))$$

If we write $f_a(n) = a + n$ then the above equations become the familiar recursive definition of addition:

$$a + 0 = a$$
$$a + (sn) = s(a + n)$$

2. Now take this arrow $a + (-) : N \to N$ together with $0 : 1 \to N$ to get another arrow $g_a : N \to N$, which is the unique one making the following commute.



We then have the recursion equations:

$$g_a(0) = 0$$
$$g_a(sn) = a + g_a(n)$$

So, writing $g_a(n) = a \cdot n$, the above equations become the familiar recursive definition of multiplication:

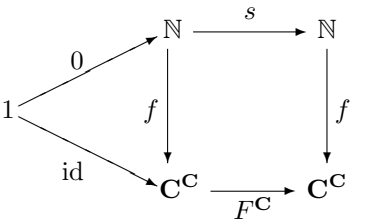
$$a \cdot 0 = 0$$
$$a \cdot (sn) = a + a \cdot n$$

3. For an example of a different sort, suppose we have a (small) category $\mathbf{C}$ and an endofunctor $F : \mathbf{C} \to \mathbf{C}$. Then there is a structure

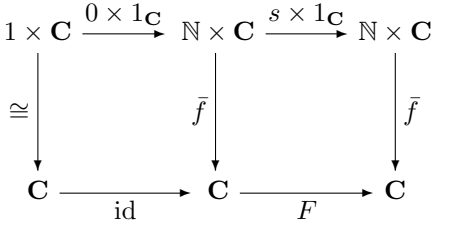
$$1 \xrightarrow{\text{id}} \mathbf{C}^{\mathbf{C}} \xrightarrow{F^{\mathbf{C}}} \mathbf{C}^{\mathbf{C}}$$

where $\text{id} : 1 \to \mathbf{C}^{\mathbf{C}}$ is the transpose of the identity $1_{\mathbf{C}} : \mathbf{C} \to \mathbf{C}$ (composed with the iso projection $1 \times \mathbf{C} \cong \mathbf{C}$). We therefore have a unique functor $f : \mathbb{N} \to \mathbf{C}^{\mathbf{C}}$ making the following diagram commute (we use the easy fact,

which the reader should check, that the discrete category $\mathbb{N}$ is an NNO in **Cat**).



Transposing gives the commutative diagram



from which we can read off the recursion equations:

$$\bar{f}(0, C) = C$$
$$\bar{f}(sn, C) = F(\bar{f}(n, C))$$

It follows that $\bar{f}(n, C) = F^{(n)}(C)$, that is, $f(n)$ is the nth iterate of F.

9.9 Exercises

1. Show that every monoid M admits a surjection from a free monoid $F(X) \to M$, by considering the counit of the *free* $\dashv$ *forgetful* adjunction.
2. Let 2 be any two-element set and consider the "diagonal functor"

$$\Delta : \mathbf{C} \to \mathbf{C}^2$$

for any category $\mathbf{C}$, that is, the exponential transpose of the first product projection

$$\mathbf{C} \times 2 \to \mathbf{C}.$$

Show that Δ has a right (resp. left) adjoint if and only if $\mathbf{C}$ has binary products (resp. coproducts).

Now let $\mathbf{C} = \mathbf{Sets}$ and replace 2 with an arbitrary small category $\mathbf{J}$. Determine both left and right adjoints for $\Delta : \mathbf{Sets} \to \mathbf{Sets}^{\mathbf{J}}$. (Hint: **Sets** is complete and cocomplete.)

3. Given a function $f : A \to B$ between sets, show that the direct image operation $\mathrm{im}(f) : P(A) \to P(B)$ is left adjoint to the inverse image $f^{-1} : P(B) \to P(A)$.
4. Show that the contravariant powerset functor $\mathcal{P} : \mathbf{Sets}^{\mathrm{op}} \to \mathbf{Sets}$ is self-adjoint.
5. Given an object C in a category $\mathbf{C}$ when does the evident forgetful functor from the slice category $\mathbf{C}/C$

$$U : \mathbf{C}/C \to \mathbf{C}$$

have a right adjoint? What about a left adjoint?

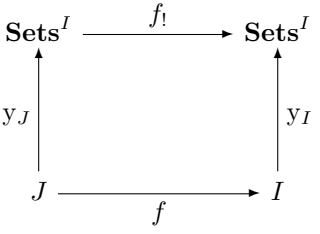
6. Let $\mathcal{P}$ be the category of propositions (i.e. the preorder category associated to the propositional calculus, say with countably many propositional variables $p, q, r, \ldots$, and a unique arrow $p \to q$ if and only if $p \vdash q$). Show that for any fixed object p, there is a functor

$$- \wedge p : \mathcal{P} \to \mathcal{P}$$

and that this functor has a right adjoint. What is the counit of the adjunction? (When) does $- \wedge p$ have a left adjoint?

7. (a) Given any set I, explicitly describe the Yoneda embedding $y : I \to \mathbf{Sets}^I$ of I into the category $\mathbf{Sets}^I$ of I-indexed sets.

   (b) Given any function $f : J \to I$ from another set J, prove directly that the following diagram commutes up to natural isomorphism.



   (c) Describe the result of composing the Yoneda embedding with the equivalence,

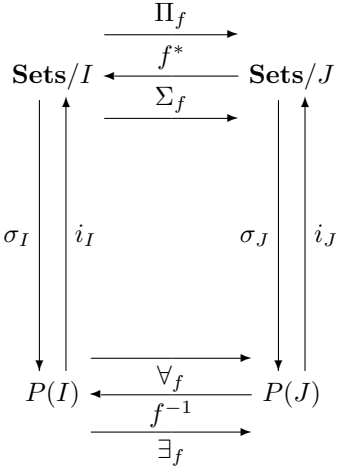
$$\mathbf{Sets}^I \simeq \mathbf{Sets}/I.$$

   (d) What does the commutativity of the above "change of base" square mean in terms of the categories $\mathbf{Sets}/I$ and $\mathbf{Sets}/J$?

   (e) Consider the inclusion functor $i : P(I) \to \mathbf{Sets}/I$ that takes a subset $U \subseteq I$ to its inclusion function $i(U) : U \to I$. Show that this is a functor and that is has a left adjoint

$$\sigma : \mathbf{Sets}/I \longrightarrow P(I).$$

(f) (Lawvere's Hyperdoctrine Diagram) In **Sets**, given any function $f : I \to J$, consider the following diagram of functors.



There are adjunctions $\sigma \dashv i$ (for both I and J), as well as $\Sigma_f \dashv f^* \dashv \Pi_f$ and $\exists_f \dashv f^{-1} \dashv \forall_f$, where $f^* : \mathbf{Sets}/J \to \mathbf{Sets}/I$ is pullback and $f^{-1} : P(J) \to P(I)$ is inverse image.

Consider which of the many possible squares commute.

8. Use the adjoint functor theorem to prove the following facts, which were shown by explicit constructions in Chapter 1:

   (a) Free monoids on sets exist.

   (b) Free categories on graphs exist.

9. Let $1 \to N \to N$ be a natural numbers object in a cartesian closed category $\mathbf{C}$. Show how to define the exponentiation operation m^n as an arrow $N \times N \to N$.

This page intentionally left blank

10

MONADS AND ALGEBRAS

In the foregoing chapter, the adjoint functor theorem was seen to imply that the category of algebras for an equational theory T always has a "free T-algebra" functor, left adjoint to the forgetful functor into **Sets**. This adjunction describes the notion of a T-algebra in a way that is independent of the specific syntactic description given by the theory T, the operations and equations of which are rather like a particular *presentation* of that notion. In a certain sense that we are about to make precise, it turns out that *every* adjunction describes, in a "syntax invariant" way, a notion of an "algebra" for an abstract "equational theory."

Toward this end, we begin with yet a third characterization of adjunctions. This one has the virtue of being entirely equational.

10.1 The triangle identities

Suppose we are given an adjunction,

$$F : \mathbf{C} \rightleftarrows \mathbf{D} : U.$$

with unit and counit,

$$\eta : 1_{\mathbf{C}} \to UF$$
$$\epsilon : FU \to 1_{\mathbf{D}}.$$

We can take any $f : FC \to D$ to $\phi(f) = U(f) \circ \eta_C : C \to UD$ and for any $g : C \to UD$ we have $\phi^{-1}(g) = \epsilon_D \circ F(g) : FC \to D$. This we know gives the isomorphism

$$\mathrm{Hom}_{\mathbf{D}}(FC, D) \cong_{\phi} \mathrm{Hom}_{\mathbf{C}}(C, UD).$$

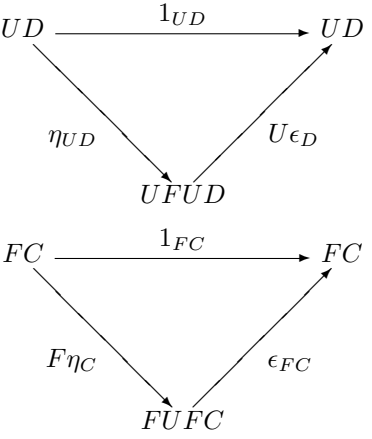
Now put $1_{UD} : UD \to UD$ in place of $g : C \to UD$ in the foregoing. We know that $\phi^{-1}(1_{UD}) = \epsilon_D$, and so

$$\begin{aligned} 1_{UD} &= \phi(\epsilon_D) \\ &= U(\epsilon_D) \circ \eta_{UD}. \end{aligned}$$

And similarly, $\phi(1_{FC}) = \eta_C$, so

$$\begin{aligned} 1_{FC} &= \phi^{-1}(\eta_C) \\ &= \epsilon_{FC} \circ F(\eta_C). \end{aligned}$$

Thus, the two diagrams below commute.



Indeed, one has the following equations of natural transformations:

$$U\epsilon \circ \eta_U = 1_U \tag{10.1}$$

$$\epsilon_F \circ F\eta = 1_F \tag{10.2}$$

These are called the "triangle identities."

Proposition 10.1. *Given categories, functors, and natural transformations*

$$F : \mathbf{C} \rightleftarrows \mathbf{D} : \mathbf{U}$$

$$\begin{aligned} \eta &: 1_{\mathbf{C}} \to U \circ F \\ \epsilon &: F \circ U \to 1_{\mathbf{D}} \end{aligned}$$

one has $F \dashv U$ with unit η and counit ϵ iff the triangle identities (10.1) *and* (10.2) *hold.*

Proof. We have already shown one direction. For the other, we just need a natural isomorphism,

$$\phi : \mathrm{Hom}_{\mathbf{D}}(FC, D) \cong \mathrm{Hom}_{\mathbf{C}}(C, UD).$$

As earlier, we put

$$\begin{aligned} \phi(f : FC \to D) &= U(f) \circ \eta_C \\ \vartheta(g : C \to UD) &= \epsilon_D \circ F(g). \end{aligned}$$

Then we check that these are mutually inverse:

$$\begin{aligned}(\vartheta(g)) &= \phi(\epsilon_D \circ F(g)) \\ &= U(\epsilon_D) \circ UF(g) \circ \eta_C \\ &= U(\epsilon_D) \circ \eta_{UD} \circ g \qquad \eta \text{ natural} \\ &= g \qquad (10.1)\end{aligned}$$

Similarly,

$$\begin{aligned}(\phi(f)) &= \vartheta(U(f) \circ \eta_C) \\ &= \epsilon_D \circ FU(f) \circ F\eta_C \\ &= f \circ \epsilon_{FC} \circ F\eta_C \qquad \epsilon \text{ natural} \\ &= f \qquad (10.2)\end{aligned}$$

Moreover, this isomorphism is easily seen to be natural. □

The triangle identities have the virtue of being entirely "algebraic"—no quantifiers, limits, Hom-sets, infinite conditions, etc. Thus anything defined by adjoints, such as free groups, product spaces, quantifiers, ... can be defined *equationally*. This is not only a matter of conceptual simplification; it also has important consequences for the existence and properties of the structures that are so determined.

10.2 Monads and adjoints

Next consider an adjunction $F \dashv U$ and the composite functor

$$U \circ F : \mathbf{C} \to \mathbf{D} \to \mathbf{C}.$$

Given *any* category $\mathbf{C}$ and endofunctor

$$T : \mathbf{C} \to \mathbf{C}$$

one can ask:

Question: When is $T = U \circ F$ for some adjoint functors $F \dashv U$ to and from another category $\mathbf{D}$?

Thus, we seek necessary and sufficient conditions on the given endofunctor $T : \mathbf{C} \to \mathbf{C}$ for recovering a category $\mathbf{D}$ and adjunction $F \dashv U$. Of course, not every T arises so, and we'll see that even if $T = U \circ F$ for *some* $\mathbf{D}$ and $F \dashv U$, we cannot always recover *that* adjunction. Thus a better way to ask the question would be, given an adjunction what sort of "trace" does it leave on a category and can we recover the adjunction from this?

First, suppose we have $\mathbf{D}$ and $F \dashv U$ and T is the composite functor $T = U \circ F$. We have then a natural transformation,

$$\eta : 1 \to T.$$

And from the counit ϵ at FC,

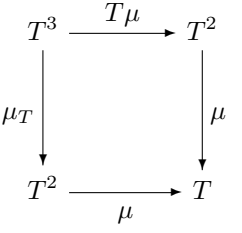
$$\epsilon_{FC} : FUFC \to FC$$

we have $U\epsilon_{FC} : UFUFC \to UFC$, which we'll call,

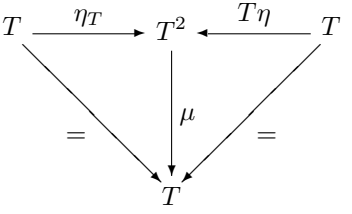
$$\mu : T^2 \to T.$$

In general, then, as a first step toward answering our question, if T arises from an adjunction, then it should have such a structure $\eta : 1 \to T$ and $\mu : T^2 \to T$.

Now, what can be said about the structure (T, η, μ)? Actually, quite a bit! Indeed, the triangle equalities give us the following commutative diagrams:

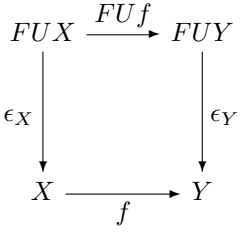


$$\mu \circ \mu_T = \mu \circ T\mu \tag{10.3}$$

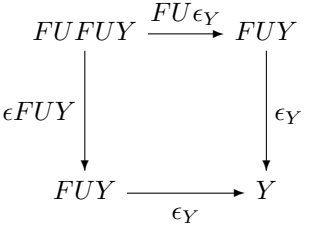


$$\mu \circ \eta_T = 1_T = \mu \circ T\eta \tag{10.4}$$

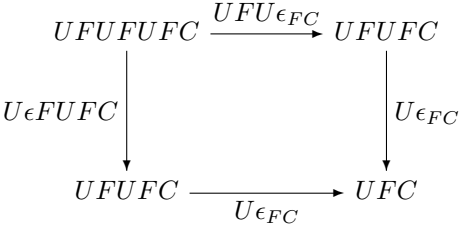
To prove the first one, for any $f : X \to Y$ in $\mathbf{D}$, the following square in $\mathbf{C}$ commutes, just since ϵ is natural.



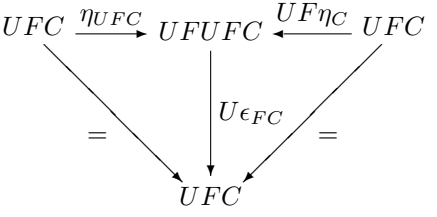
Now take $X = FUY$ and $f = \epsilon_Y$ to get the following:



Putting FC for Y and applying U therefore gives this



which has the required form (10.3). The equations (10.4) in the form



are simply the triangle identities, once taken at FC, and once under U. We record this data in the following:

Definition 10.2. A *monad* on a category $\mathbf{C}$ consists of an endofunctor $T : \mathbf{C} \to \mathbf{C}$, and natural transformations $\eta : 1_{\mathbf{C}} \to T$, and $\mu : T^2 \to T$ satisfying the two commutative diagrams above, that is,

$$\mu \circ \mu_T = \mu \circ T\mu$$
$$\mu \circ \eta_T = 1 = \mu \circ T\eta.$$

Note the formal analogy to the definition of a monoid. For this reason, the equations are called the *associativity* and *unit* laws, respectively. It is remarkable that the notion of a monad actually first arose independently of adjunctions.

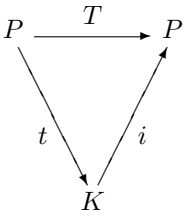
We have now shown the following:

Proposition 10.3. *Every adjoint pair $F \dashv U$ with $U : \mathbf{D} \to \mathbf{C}$, unit $\eta : UF \to 1_{\mathbf{C}}$ and counit $\epsilon : 1_{\mathbf{D}} \to FU$ gives rise to a monad (T, η, μ) on $\mathbf{C}$ with*

$$T = U \circ F : \mathbf{C} \to \mathbf{C}$$
$$\eta : 1 \to T \quad \textit{the unit}$$
$$\mu = U\epsilon_F : T^2 \to T.$$

Example 10.4. Let P be a poset. A monad on P is a monotone function $T : P \to P$ with $x \leq Tx$ and $T^2x \leq Tx$. But then $T^2 = T$, that is, T is *idempotent*. Such a T, that is both inflationary and idempotent, is sometimes called a *closure operation* and written $Tp = \bar{p}$, since it acts like the closure operation on the subsets of a topological space. The "possibility operator" $\diamond p$ in modal logic is another example.

In the poset case, we can easily recover an adjunction from the monad. First, let $K = \mathrm{im}(T)(P)$ (the fixed points of T), and let $i : K \to P$ be the inclusion. Then let t be the factorization of T through K, as indicated in:



Observe that since $TTp = Tp$, for any element $k \in K$ we then have, for some $p \in P$, the equation $itik = ititp = itp = ik$, whence $tik = k$ since i is monic. We therefore have:

$$p \leq ik \quad \text{implies} \quad tp \leq tik = k$$
$$tp \leq k \quad \text{implies} \quad p \leq itp \leq ik$$

So indeed $t \dashv i$.

Example 10.5. Consider the covariant powerset functor

$$\mathcal{P} : \mathbf{Sets} \to \mathbf{Sets}$$

which takes each function $f : X \to Y$ to the image mapping $\mathrm{im}(f) : P(X) \to P(Y)$. Let $\eta_X : X \to \mathcal{P}(X)$ be the singleton operation

$$\eta_X(x) = \{x\}$$

and let $\mu_X : \mathcal{P}\mathcal{P}(X) \to \mathcal{P}(X)$ be the union operation

$$\mu_X(\alpha) = \bigcup \alpha.$$

The reader should verify as an exercise that these operations are in fact natural in X and that this defines a monad $(\mathcal{P}, \{-\}, \bigcup)$ on **Sets**.

As we see in these examples, monads can, and often do, arise without coming from evident adjunctions. In fact, the notion of a monad originally did occur independently of adjunctions; they are also known as "triples," and "standard constructions." Despite this, the question, when does an endofunctor T arise from an adjunction, has the answer: just if it is the functor part of a monad.

10.3 Algebras for a monad

Proposition 10.6. *Every monad arises from an adjunction. More precisely, given a monad (T, η, μ) on the category $\mathbf{C}$, there exists a category $\mathbf{D}$ and an adjunction $F \dashv U$, $\eta : 1 \to UF$, $\epsilon : FU \to 1$ with $U : \mathbf{D} \to \mathbf{C}$ such that*

$$\begin{aligned} T &= U \circ F \\ \eta &= \eta \quad \textit{(the unit)} \\ \mu &= U\epsilon_F. \end{aligned}$$

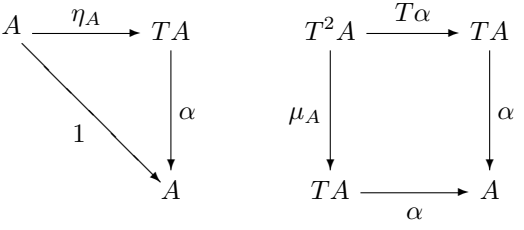
Proof. We will first define the important category $\mathbf{C}^T$ called the *Eilenberg-Moore category of* T. This will be our "$\mathbf{D}$." Then we need suitable functors

$$F : \mathbf{C} \rightleftarrows \mathbf{C}^T : U.$$

And, finally, we need natural transformations $\eta : 1 \to UF$ and $\epsilon : FU \to 1$ satisfying the triangle identities.

To begin, $\mathbf{C}^T$ has as *objects* the "T-algebras," which are pairs (A, α) of the form $\alpha : TA \to A$ in $\mathbf{C}$, such that

$$1_A = \alpha \circ \eta_A \quad \text{and} \quad \alpha \circ \mu_A = \alpha \circ T\alpha. \tag{10.5}$$



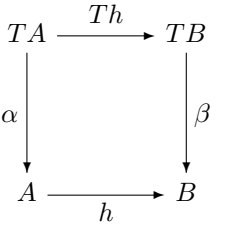
A *morphism* of T-algebras,

$$h : (A, \alpha) \to (B, \beta)$$

is simply an arrow $h : A \to B$ in $\mathbf{C}$, such that,

$$h \circ \alpha = \beta \circ T(h)$$

as indicated in the following diagram.



It is obvious that $\mathbf{C}^T$ is a category with the expected composites and identities coming from $\mathbf{C}$, and that T is a functor.

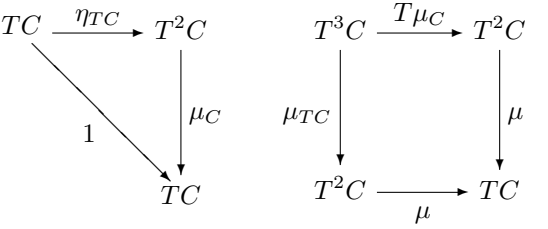
Now define the functors,

$$U : \mathbf{C}^T \to \mathbf{C}$$
$$U(A, \alpha) = A$$

and

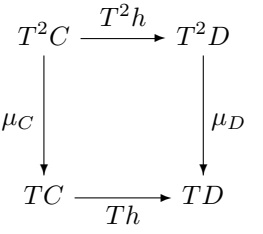
$$F : \mathbf{C} \to \mathbf{C}^T$$
$$FC = (TC, \mu_C).$$

We need to check that (TC, μ_C) is a T-algebra. The equations (10.5) for T-algebras in this case become:



But these come directly from the definition of a monad.

To see that F is a functor, given any $h : C \to D$ in $\mathbf{C}$, we have



since μ is natural. But this is a T-algebra homomorphism $FC \to FD$, so we can put

$$Fh = Th : TC \to TD$$

to get an arrow in $\mathbf{C}^T$.

Now we've defined the category $\mathbf{C}^T$ and the functors

$$\mathbf{C} \underset{U}{\overset{F}{\rightleftarrows}} \mathbf{C^T}$$

and we want to show that $F \dashv U$. Next, we need the unit and counit:

$$\begin{aligned} \bar{\eta} &: 1_{\mathbf{C}} \to U \circ F \\ \epsilon &: F \circ U \to 1_{\mathbf{C}^T} \end{aligned}$$

Given $C \in \mathbf{C}$, we have

$$UF(C) = U(TC, \mu_C) = TC.$$

So we can take $\bar{\eta} = \eta : 1_{\mathbf{C}} \to U \circ F$, as required.

Given $(A, \alpha) \in \mathbf{C}^T$,

$$FU(A, \alpha) = (TA, \mu_A)$$

and the definition of a T-algebra makes the following diagram commute:

$$\begin{array}{ccc} T^2A & \xrightarrow{T\alpha} & TA \\ {\scriptstyle \mu_A}\downarrow & & \downarrow{\scriptstyle \alpha} \\ TA & \xrightarrow[\alpha]{} & A \end{array}$$

But this is a morphism $\epsilon_{(A,\alpha)} : (TA, \mu_A) \to (A, \alpha)$ in $\mathbf{C}^T$. Thus we are setting

$$\epsilon_{(A,\alpha)} = \alpha.$$

And ϵ is *natural* by the definition of a morphism of T-algebras, as follows. Given any $h : (A, \alpha) \to (B, \beta)$, we need to show

$$h \circ \epsilon_{(A,\alpha)} = \epsilon_{(B,\beta)} \circ Th.$$

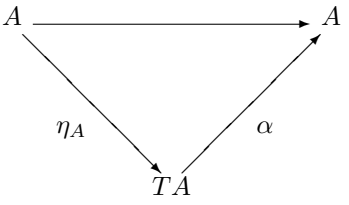
But by the definition of ϵ, that is, $h \circ \alpha = \beta \circ Th$, which holds since h is a T-algebra homomorphism.

Finally, the triangle identities now read as follows:

1. For (A, α) a T-algebra

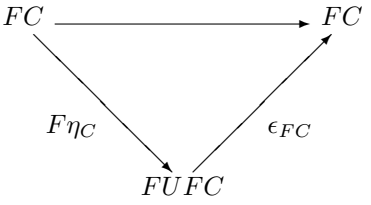
$$\begin{array}{ccc} U(A,\alpha) & \longrightarrow & U(A,\alpha) \\ & {\scriptstyle \eta_{U(A,\alpha)}} \searrow \quad \nearrow {\scriptstyle U\epsilon_{(A,\alpha)}} & \\ & UFU(A,\alpha) & \end{array}$$

which amounts to

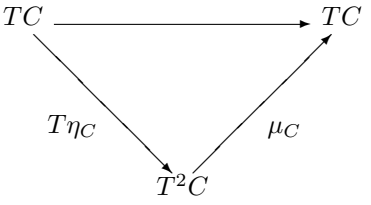


which holds since (A, α) is T-algebra.

2. For any $C \in \mathbf{C}$



which is



which holds by one of the unit laws for T.

Finally, note that we indeed have

$$T = U \circ F$$
$$\eta = \text{ unit of } F \dashv U.$$

And for the multiplication,

$$\bar{\mu} = U\epsilon F$$

we have, for any $C \in \mathbf{C}$,

$$\bar{\mu}_C = U\epsilon_{FC} = U\epsilon_{(TC,\mu_C)} = U\mu_C = \mu_C.$$

So $\bar{\mu} = \mu$ and we are done; the adjunction $F \dashv U$ via η and ϵ gives rise to the monad (T, η, μ). □

Example 10.7. Take the free monoid adjunction,

$$F : \mathbf{Sets} \rightleftarrows \mathbf{Mon} : \mathbf{U}.$$

The monad on **Sets** is then $T : \mathbf{Sets} \to \mathbf{Sets}$, where for any set X, $T(X) = UF(X) =$ "strings over X." The unit $\eta : X \to TX$ is the usual "string of length one" operation, but what is the multiplication?

$$\mu : T^2X \to TX$$

Here T^2X is the set of strings of strings,

$$[[x_{11}, \ldots, x_{1n}], [x_{21}, \ldots, x_{2n}], \ldots, [x_{m1}, \ldots, x_{mn}]].$$

And μ of such a string of strings is the string of their elements,

$$\mu([[x_{11}, \ldots, x_{1n}], [x_{21}, \ldots, x_{2n}], \ldots, [x_{m1}, \ldots, x_{mn}]]) = [x_{11}, \ldots, x_{mn}].$$

Now, what is a T-algebra in this case? By the equations for a T-algebra, it is a map,

$$\alpha : TA \to A$$

from strings over A to elements of A, such that

$$\alpha[a] = a$$

and

$$\alpha(\mu([[\ldots], [\ldots], \ldots, [\ldots]])) = \alpha(\alpha[\ldots], \alpha[\ldots], \ldots, \alpha[\ldots]).$$

If we start with a monoid, then we can get a T-algebra $m : TM \to M$ by:

$$m[m_1, \ldots, m_n] = m_1 \cdot \ldots \cdot m_n$$

This clearly satisfies the required conditions. Observe that we can even recover the monoid structure from m by $u = m(-)$ for the unit and $x \cdot y = m(x, y)$ for the multiplication. Indeed *every* T-algebra is of this form for a *unique* monoid (exercise!).

We have now given constructions back and forth between adjunctions and monads. And we know that if we start with a monad $T : \mathbf{C} \to \mathbf{C}$, and then take the adjunction,

$$F^T : \mathbf{C} \rightleftarrows \mathbf{C}^T : U^T$$

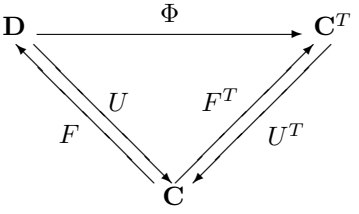
then we can get the monad back by $T = U^T \circ F^T$. Thus, in particular, every monoid arises from *some* adjunction. But are $\mathbf{C}^T, U^T, F^T$ unique with this property?

In general, the answer is *no*. There may be many different categories $\mathbf{D}$ and adjunctions $F \dashv U : \mathbf{D} \to \mathbf{C}$, all giving the same monad on $\mathbf{C}$. We have used the Eilenberg-Moore category $\mathbf{C}^T$, but there is also something called the "Kleisli category," which is in general different from $\mathbf{C}^T$, but also has an adjoint pair to $\mathbf{C}$ giving rise to the same monad (see the exercises).

If we start with an adjunction $F \dashv U$ and construct $\mathbf{C}^T$ for $T = U \circ F$, we then get a comparison functor $\Phi : \mathbf{D} \to \mathbf{C}^T$, with

$$U^T \circ \Phi \cong U$$

$$\Phi \circ F = F^T$$



In fact, Φ is *unique* with this property. A functor $U : \mathbf{D} \to \mathbf{C}$ is called *monadic* if it has a left adjoint $F \dashv U$, such that this comparison functor is an equivalence of categories,

$$\mathbf{D} \xrightarrow[\cong]{\Phi} \mathbf{C}^T$$

for $T = UF$.

Typical examples of monadic forgetful functors $U : \mathbf{C} \to \mathbf{Sets}$ are those from the "algebraic" categories arising as models for equational theories, like monoids, groups, rings, etc. Indeed, one can reasonably take monadicity as the *definition* of being "algebraic."

An example of a right adjoint that is *not* monadic is the forgetful functor from posets,

$$U : \mathbf{Pos} \to \mathbf{Sets}.$$

Its left adjoint F is the discrete poset functor. For any set X, therefore, one has as the unit the identity function $X = UF(X)$. The reader can easily show that the Eilenberg-Moore category for $T = 1_{\mathbf{Sets}}$ is then just **Sets** itself.

10.4 Comonads and coalgebras

By definition, a *comonad* on a category $\mathbf{C}$ is a monad on $\mathbf{C}^{\mathrm{op}}$. Explicitly, this consists of an endofunctor $G : \mathbf{C} \to \mathbf{C}$ and natural transformations,

$$\epsilon : G \to 1 \qquad \text{the counit}$$

$$\delta : G \to G^2 \quad \text{comultiplication}$$

satisfying the duals of the equations for a monad, namely

$$\delta_G \circ \delta \;=\; G\delta \circ \delta$$

$$\epsilon_G \circ \delta \;=\; 1_G \;=\; G\epsilon \circ \delta.$$

We leave it as an exercise in duality for the reader to verify that an adjoint pair $F \dashv U$ with $U : \mathbf{D} \to \mathbf{C}$ and $F : \mathbf{C} \to \mathbf{D}$ and $\eta : 1_{\mathbf{C}} \to UF$ and $\epsilon : FU \to 1_{\mathbf{D}}$ gives rise to a comonad (G, ϵ, δ) on $\mathbf{D}$, where

$$G = F \circ U : \mathbf{D} \to \mathbf{D}$$
$$\epsilon : G \to 1$$
$$\delta = F\eta_U : G \to G^2.$$

The notions of coalgebra for a comonad, and of a comonadic functor, are of course also precisely dual to the corresponding ones for monads. Why do we even bother to study these notions separately, rather than just considering their duals? As in other examples of duality, there are actually two distinct reasons:

1. We may be interested in a particular category with special properties not had by its dual. A comonad on $\mathbf{Sets}^{\mathbf{C}}$ is of course a monad on $(\mathbf{Sets}^{\mathbf{C}})^{\mathrm{op}}$, but as we now know, $\mathbf{Sets}^{\mathbf{C}}$ has many special properties that its dual does not have (e.g. it is a topos!). So we can profitably consider the notion of a comonad on such a category.

   A simple example of this kind is the comonad $G = \Delta \circ \varprojlim$ resulting from composing the "constant functor" functor $\Delta : \mathbf{Sets} \to \mathbf{Sets}^{\mathbf{C}}$ with the "limit" functor $\varprojlim : \mathbf{Sets}^{\mathbf{C}} \to \mathbf{Sets}$. It can be shown in general that the coalgebras for this comonad again form a topos. In fact, they are just the constant functors $\Delta(S)$ for sets S, and the category $\mathbf{Sets}$ is thus comonadic over $\mathbf{Sets}^{\mathbf{C}}$.

2. It may happen that both structures—monad and comonad—occur together, and interact. Taking the opposite category will not alter this situation! This happens for instance when a system of *three* adjoint functors are composed:

   $$L \dashv U \dashv R \qquad \mathbf{C} \underset{\longrightarrow}{\overset{R}{\underset{L}{\overset{\longleftarrow}{\longrightarrow}}}} \mathbf{D}$$

   resulting in a monad $T = U \circ L$ and a comonad $G = U \circ R$, both on $\mathbf{C}$. In such a case, T and U are then of course also adjoint $T \dashv G$.

   This arises, for instance, in the foregoing example with $R = \varprojlim$, and $U = \Delta$, and $L = \varinjlim$ the "colimit" functor. It also occurs in propositional modal logic, with $T = \Diamond$ "possibility" and $G = \Box$ "necessity," where the adjointness $\Diamond \dashv \Box$ is equivalent to the law known to modal logicians as "S5."

   A related example is given by the open and closed subsets of a topological space: the topological interior operation on arbitrary subsets is a comonad and closure is a monad. We leave the details as an exercise.

10.5 Algebras for endofunctors

Some very basic kinds of algebraic structures have a more simple description than as algebras for a monad, and this description generalizes to structures that are not algebras for any monad, but still have some algebra-like properties.

As a familiar example, consider first the underlying structure of the notion of a group. We have a set G equipped with operations as indicated in the following:

$$\begin{array}{ccccc} G \times G & \xrightarrow{m} & G & \xleftarrow{i} & G \\ & & \uparrow{\scriptstyle u} & & \\ & & 1 & & \end{array}$$

We do not assume, however, that these operations satisfy the group equations of associativity, etc. Observe that this description of what we will call a "group structure" can plainly be compressed into a single arrow of the form:

$$1 + G + G \times G \xrightarrow{[u,i,m]} G$$

Now let us define the functor $F : \mathbf{Sets} \to \mathbf{Sets}$ by:

$$F(X) \;=\; 1 + X + X \times X$$

Then a group structure is simply an arrow,

$$\gamma : F(G) \to G.$$

Moreover, a homomorphism of group structures in the conventional sense

$$h : G \to H,$$

$$\begin{aligned} h(u_G) &= u_H \\ h(i(x)) &= i(h(x)) \\ h(m(x,y)) &= m(h(x), h(y)) \end{aligned}$$

is then exactly a function $h : G \to H$ such that the following diagram commutes.

$$\begin{array}{ccc} F(G) & \xrightarrow{F(h)} & F(H) \\ \downarrow{\scriptstyle \gamma} & & \downarrow{\scriptstyle \vartheta} \\ G & \xrightarrow[h]{} & H \end{array}$$

where $\vartheta : F(H) \to H$ is the group structure on H. This observation motivates the following definition.

Definition 10.8. Given an endofunctor $P : \mathcal{S} \to \mathcal{S}$ on any category $\mathcal{S}$, a *P-algebra* consists of an object A of $\mathcal{S}$ and an arrow,

$$\alpha : PA \to A.$$

A *homomorphism* $h : (A, \alpha) \to (B, \beta)$ of P-algebras is an arrow $h : A \to B$ in $\mathcal{S}$ such that $h \circ \alpha = \beta \circ P(h)$, as indicated in the following diagram:

$$\begin{array}{ccc} P(A) & \xrightarrow{P(h)} & P(B) \\ \downarrow{\scriptstyle \alpha} & & \downarrow{\scriptstyle \beta} \\ A & \xrightarrow[h]{} & B \end{array}$$

The category of all such P-algebras and their homomorphisms will be denoted as:

$$P\text{-Alg}(\mathcal{S})$$

We will usually write more simply P-Alg when $\mathcal{S}$ is understood. Also, if there is a monad present, we will need to be careful to distinguish between algebras for the monad and algebras for the endofunctor (especially if P is the functor part of the monad!).

Example 10.9. 1. For the functor $P(X) = 1 + X + X \times X$ on **Sets** we have already seen that the category **GrpStr** of group structures is the same thing as the category of P-algebras,

$$P\text{-Alg} \;=\; \mathbf{GrpStr}.$$

2. Clearly, for any other algebraic structure of finite "signature," that is, consisting of finitely many, finitary operations, there is an analogous description of the structures of that sort as algebras for an associated endofunctor. For instance, a *ring structure*, with two nullary, one unary, and two binary operations is given by the endofunctor

$$R(X) = 2 + X + 2 \times X^2.$$

In general, a functor of the form

$$P(X) = C_0 + C_1 \times X + C_2 \times X^2 + \cdots + C_n \times X^n$$

with natural number coefficients C_k, is called a (finitary) *polynomial functor*, for obvious reasons. These functors present exactly the *finitary structures*. The same thing holds for finitary structures in any category $\mathcal{S}$

with finite products and coproducts; these can always be represented as algebras for a suitable endofunctor.

3. In a category such as **Sets** that is complete and cocomplete, there is an evident generalization to infinitary signatures by using generalized or "infinitary" polynomial functors, that is, ones with infinite sets C_k as coefficients (representing infinitely many operations of a given arity), infinite sets B_k as the exponents X^{B_k} (representing operations of infinite arity), or infinitely many terms (representing infinitely many different arities of operations), or some combination of these. The algebras for such an endofunctor

$$P(X) = \sum_{i \in I} C_i \times X^{B_i}$$

can then be naturally viewed as generalized "algebraic structures." Using locally cartesian closed categories, one can even present this notion without needing (co)completeness.

4. One can of course also consider algebras for an endofunctor $P : \mathcal{S} \to \mathcal{S}$ that is not polynomial at all, such as the covariant powerset functor $\mathcal{P} : \mathbf{Sets} \to \mathbf{Sets}$. This leads to a proper generalization of the notion of an "algebra," which however still shares some of the formal properties of conventional algebras, as shall be seen below.

Let $P : \mathbf{Sets} \to \mathbf{Sets}$ be a polynomial functor, say

$$P(X) = 1 + X^2$$

(what structure is this?). Then the notion of an *initial* P-algebra gives rise to a recursion property analogous to that of the natural numbers. Specifically, let

$$[o, m] : 1 + I^2 \to I$$

be an initial P-algebra, that is, an initial object in the category of P-algebras. Then, explicitly, we have the structure

$$o \in I, \qquad m : I \times I \to I$$

and for any set X with a distinguished element and a binary operation

$$a \in X, \qquad * : X \times X \to X$$

there is a unique function $u : I \to X$ such that the following diagram commutes.

$$\begin{array}{ccc} 1 + I^2 & \xrightarrow{P(u)} & 1 + X^2 \\ {\scriptstyle [o,m]}\downarrow & & \downarrow{\scriptstyle [a,*]} \\ I & \xrightarrow[u]{} & X \end{array}$$

This of course says that, for all $i, j \in I$,

$$u(o) = a$$
$$u(m(i,j)) = u(i) * u(j).$$

which is exactly a *definition by recursion* of the function $u : I \to X$. Indeed, the usual recursion property of the natural numbers $\mathbb{N}$ with zero $0 \in \mathbb{N}$ and successor $s : \mathbb{N} \to \mathbb{N}$ says precisely that $(\mathbb{N}, 0, s)$ is the initial algebra for the endofunctor,

$$P(X) = 1 + X : \mathbf{Sets} \to \mathbf{Sets}$$

as the reader should check.

We next briefly investigate the question: When does an endofunctor have an initial algebra? The existence is constrained by the fact that initial algebras, when they exist, must have the following noteworthy property.

Lemma 10.10 (Lambek). *Given any endofunctor $P : \mathcal{S} \to \mathcal{S}$ on an arbitrary category $\mathcal{S}$, if $i : P(I) \to I$ is an initial P-algebra, then i is an isomorphism,*

$$P(I) \cong I.$$

We leave the proof as an easy exercise.

In this sense, the initial algebra for an endofunctor $P : \mathcal{S} \to \mathcal{S}$ is a "least fixed point" for P. Such algebras are often used in computer science to model "recursive datatypes" determined by so-called "fixed point equations" $X = P(X)$.

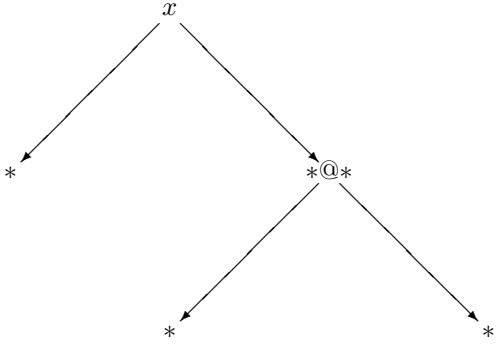
Example 10.11. 1. For the polynomial functor,

$$P(X) = 1 + X^2$$

(monoid structure!), let us "unwind" the initial algebra,

$$[*, @] : 1 + I \times I \cong I.$$

Given any element $x \in I$, it is thus either of the form $*$ or of the form $x_1 @ x_2$ for some elements $x_1, x_2 \in I$. Each of these x_i, in turn, is either of the form $*$ or of the form $x_{i1} @ x_{i2}$, and so on. Continuing in this way, we have a representation of x as a finite, binary tree. For instance, an element of the form $x = *@(*@*)$ looks like:



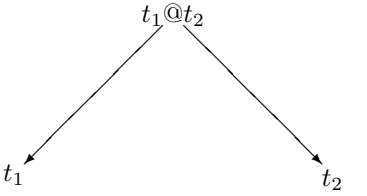
We can present the monoid structure explicitly by letting

$$I = \{t \mid t \text{ is a finite, binary tree}\}$$

with:

$$* = \text{“the empty tree”}$$
$$@(t_1, t_2) = t_1@t_2$$

$=$



The isomorphism,

$$[*, @] : 1 + I \times I \to I$$

here is plain to see.

2. Similarly, for any other polynomial functor,

$$P(X) = C_0 + C_1 \times X + C_2 \times X^2 + \cdots + C_n \times X^n$$

we can describe the initial algebra (in **Sets**),

$$P(I) \cong I$$

as a set of trees with branching types and labels determined by P. For instance, consider the polynomial

$$P(X) = 1 + A \times X$$

for some set A. What is the initial algebra? Since,

$$[*, @] : 1 + A \times I \cong I$$

we can unwind an element x as:

$$x = * \text{ or } a_1@x_1$$
$$x_1 = * \text{ or } a_2@x_2$$
$$\cdots$$

Thus we essentially have $x = a_1@a_2@\cdots@a_n$. So I can be represented as the set A-List of (finite) lists of elements $a_1, a_2, \ldots$ of A, with the structure:

$$* = \text{“the empty list”}$$
$$@(a, \ell) = a@\ell$$

The usual procedure of "recursive definition" follows from initiality. For example, the length function for lists length : A-List $\to \mathbb{N}$ is usually defined by:

$$\text{length}(*) = 0 \tag{10.6}$$

$$\text{length}(a@\ell) = 1 + \text{length}(\ell) \tag{10.7}$$

We can do this by equipping $\mathbb{N}$ with a suitable $P(X) = 1 + A \times X$ structure, namely,

$$[0, m] : 1 + A \times \mathbb{N} \to \mathbb{N}$$

where $m(a, n) = 1 + n$ for all $n \in \mathbb{N}$. Then by the UMP of the initial algebra we get a unique function length : A-List $\to \mathbb{N}$ making a commutative square:

$$\begin{array}{ccc} 1 + A \times A\text{-List} & \xrightarrow{1 + A \times \text{length}} & 1 + A \times \mathbb{N} \\ \downarrow{\scriptstyle [*, @]} & & \downarrow{\scriptstyle [0, m]} \\ A\text{-List} & \xrightarrow[\text{length}]{} & \mathbb{N} \end{array}$$

But this commutativity is, of course, precisely equivalent to the equations (10.6) and (10.7) above.

In virtue of Lambek's lemma, we at least know that not all endofunctors can have initial algebras. For, consider the covariant powerset functor $\mathcal{P} : \mathbf{Sets} \to \mathbf{Sets}$. An initial algebra for this would give us a set I with the property that $\mathcal{P}(I) \cong I$, which is impossible by the well-known theorem of Cantor!

The following proposition gives a useful sufficient condition for the existence of an initial algebra.

Proposition 10.12. *If the category $\mathcal{S}$ has an initial object 0 and colimits of diagrams of type ω (call them "ω-colimits"), and the functor*

$$P : \mathcal{S} \to \mathcal{S}$$

preserves ω-colimits, then P has an initial algebra.

Proof. Note that this generalizes a very similar result for posets already given above as Proposition 5.35. And even the proof by "Newton's method" is essentially the same! Take the ω-sequence

$$0 \to P0 \to P^2 0 \to \cdots$$

and let I be the colimit

$$I = \varinjlim_n P^n 0.$$

Then, since P preserves the colimit, there is an isomorphism

$$P(I) = P(\varinjlim_n P^n 0) \cong \varinjlim_n P(P^n 0) = \varinjlim_n P^n 0 = I$$

which is seen to be an initial algebra for P by an easy diagram chase. □

Since (as the reader should verify) every polynomial functor $P : \mathbf{Sets} \to \mathbf{Sets}$ preserves ω-colimits, we have:

Corollary 10.13. *Every polynomial functor $P : \mathbf{Sets} \to \mathbf{Sets}$ has an initial algebra.*

Finally, we ask, what is the relationship between algebras for endofunctors and algebras for monads? The following proposition, which is a sort of "folk theorem", gives the answer.

Proposition 10.14. *Let the category $\mathcal{S}$ have finite coproducts. Given an endofunctor $P : \mathcal{S} \to \mathcal{S}$, the following conditions are equivalent:*

1. *The P-algebras are the algebras for a monad. Precisely, there is a monad $(T : \mathcal{S} \to \mathcal{S}, \eta, \mu)$, and an equivalence:*

   $$P\text{-Alg}(\mathcal{S}) \simeq \mathcal{S}^T$$

   between the category of P-algebras and the category $\mathcal{S}^T$ of algebras for the monad. Moreover, this equivalence preserves the respective forgetful functors to $\mathcal{S}$.
2. *The forgetful functor $U : P\text{-Alg}(\mathcal{S}) \to \mathcal{S}$ has a left adjoint*

   $$F \vdash U.$$

3. *For each object A of $\mathcal{S}$, the endofunctor*

   $$P_A(X) = A + P(X) : \mathcal{S} \to \mathcal{S}$$

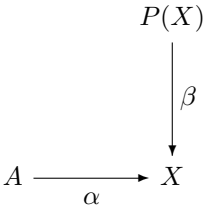
   has an initial algebra.

Proof. That (1) implies (2) is clear.

For (2) implies (3), suppose that U has a left adjoint $F : \mathcal{S} \to P\text{-Alg}$ and consider the endofunctor $P_A(X) = A + P(X)$. An algebra (X, γ) is a map $\gamma : A + P(X) \to X$. But there is clearly a unique correspondence between the

following three types of things:

$$\gamma : A + P(X) \to X$$



$$\alpha : A \to U(X, \beta)$$

Thus the P_A-algebras can be described equivalently as arrows of the form $\alpha : A \to U(X, \beta)$ for P-algebras (X, β). Moreover, a P_A-homomorphism $h : (\alpha, U(X, \beta)) \to (\alpha', U(X', \beta'))$ is just a P-homomorphism $h : (X, \beta) \to (X', \beta')$ making a commutative triangle with α and $\alpha' : A \to U(X', \beta')$. But an initial object in this category is given by the unit $\eta : A \to UFA$ of the adjunction $F \vdash U$, which shows (3).

Indeed, given just the forgetful functor $U : P\text{-Alg} \to \mathcal{S}$, the existence of initial objects in the respective categories of arrows $\alpha : A \to U(X, \beta)$, for each A, is exactly what is needed for the existence of a left adjoint F to U. So (3) also implies (2).

Before concluding the proof, it is illuminating to see how the free functor $F : \mathcal{S} \to P\text{-Alg}$ results from condition (3). For each object A in $\mathcal{S}$, consider the initial P_A-algebra $\alpha : A + P(I_A) \to I_A$. In the notation of recursive type theory,

$$I_A = \mu_X . A + P(X)$$

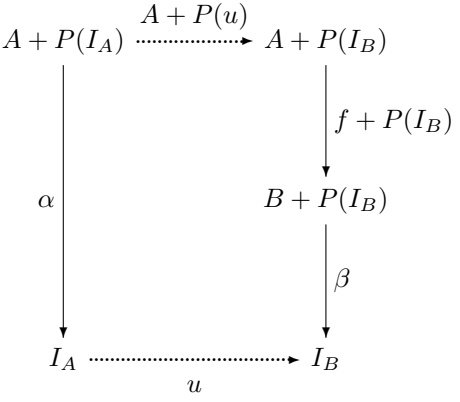
meaning it is the (least) solution to the "fixed point equation"

$$X = A + P(X).$$

Since α is a map on the coproduct $A + P(I_A)$, we have $\alpha = [\alpha_1, \alpha_2]$, and we can let:

$$F(A) = (I_A,\ \alpha_2 : P(I_A) \to I_A)$$

To define the action of F on an arrow $f : A \to B$, let $\beta : B + P(I_B) \to I_B$ be the initial P_B-algebra and consider the diagram



The right-hand vertical composite $\beta \circ (f + P(I_B))$ now makes I_B into a P_A-algebra. There is thus a unique P_A-homomorphism u as indicated, and we can set

$$F(f) = u.$$

Finally, to conclude, the fact that (2) implies (1) is an easy application of Beck's Precise Tripleability Theorem, for which we refer the reader to section VI.7 of Mac Lane's (1971) "*Categories Work*." □

10.6 Exercises

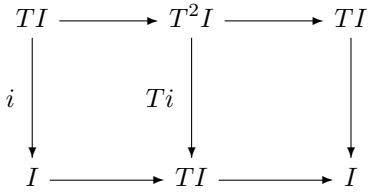
1. Let $\mathbb{T}$ be the equational theory with one constant symbol and one unary function symbol (no axioms). In any category with a terminal object, a natural numbers object (NNO) is just an initial $\mathbb{T}$-model. Show that the natural numbers

   $$(\mathbb{N}, 0 \in \mathbb{N}, n + 1 : \mathbb{N} \to \mathbb{N})$$

   is a NNO in **Sets**, and that any NNO is uniquely isomorphic to it (as a $\mathbb{T}$-model).

   Finally, show that $(\mathbb{N}, 0 \in \mathbb{N}, n + 1 : \mathbb{N} \to \mathbb{N})$ is uniquely characterized (up to isomorphism) as the initial algebra for the endofunctor $F(X) = X + 1$.

2. ("Lambek's Lemma") Show that for any endofunctor $T : \mathbf{C} \to \mathbf{C}$, if $i : TI \to I$ is an initial T-algebra, then i is an isomorphism.
   Hint: Consider a diagram of the following form, with suitable arrows.

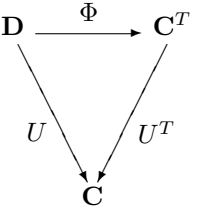


Conclude that for any NNO N in any category, there is an isomorphism $N + 1 \cong N$. Also, derive the usual recursion property of the natural numbers from initiality.

3. Assume given categories $\mathbf{C}$ and $\mathbf{D}$ and adjoint functors

$$F : \mathbf{C} \rightleftarrows \mathbf{D} : U$$

with unit $\eta : 1_{\mathbf{C}} \to UF$ and counit $\epsilon : FU \to 1_{\mathbf{D}}$. Show that every D in $\mathbf{D}$ determines a $T = UF$ algebra $U\epsilon : UFUD \to UD$, and that there is a "comparison functor" $\Phi : \mathbf{D} \to \mathbf{C}^T$ which, moreover, commutes with the "forgetful" functors $U : \mathbf{D} \to \mathbf{C}$ and $U^T : \mathbf{C}^T \to \mathbf{C}$.



4. Show that $(P, s, \cup)$ is a monad on **Sets**, where

- $P : \mathbf{Sets} \to \mathbf{Sets}$ is the covariant powerset functor, which takes each function $f : X \to Y$ to the image mapping

$$P(f) = im(f) : P(X) \to P(Y)$$

- for each set X, the component $s_X : X \to P(X)$ is the singleton mapping, with

$$s_X(x) = \{x\} \subseteq X$$

for each $x \in X$;

- for each set X, the component $\cup_X : PP(X) \to P(X)$ is the union operation, with

$$\cup_X(\alpha) = \{x \in X \mid \exists_{U \in \alpha}.\ x \in U\} \subseteq X$$

for each $\alpha \subseteq P(X)$.

5. Consider the free $\dashv$ forgetful adjunction

$$F : \mathbf{Sets} \rightleftarrows \mathbf{Mon} : \mathbf{U}$$

between sets and monoids, and let (T, η^T, μ^T) be the associated monad on **Sets**. Show that any T-algebra $\alpha : TA \to A$ for this monad comes from a monoid structure on A (exhibit the monoid multiplication and unit element).

6. (a) Show that an adjoint pair $F \dashv U$ with $U : \mathbf{D} \to \mathbf{C}$ and $\eta : UF \to 1_{\mathbf{C}}$ and $\epsilon : 1_{\mathbf{D}} \to FU$ also gives rise to a *comonad* (G, ϵ, δ) in $\mathbf{D}$, with

$$G = F \circ U : \mathbf{D} \to \mathbf{D}$$
$$\epsilon : G \to 1 \text{ the counit}$$
$$\delta = F\eta_U : G \to G^2$$

satisfying the duals of the equations for a monad.

(b) Define the notion of a *coalgebra* for a comonad, and show (by duality) that every comonad (G, ϵ, δ) on a category $\mathbf{D}$ "comes from" a (not necessarily unique) adjunction $F \dashv G$ such that $G = FU$ and ϵ is the counit.

(c) Let **End** be the category of sets equipped with an endomorphism, $e : S \to S$. Consider the functor $G : \mathbf{End} \to \mathbf{End}$ defined by

$$G(S, e) = \{x \in S \mid e^{(n+1)}(x) = e^{(n)}(x) \text{ for some } n\}$$

equipped with the restriction of e. Show that this is the functor part of a comonad on **End**.

7. Verify that the open and closed subsets of a topological space give rise to comonad and monad, respectively, on the powerset of the underlying point-set. Moreover, the categories of coalgebras and algebras are isomorphic.

8. (Kleisli category) Given a monad (T, η, μ) on a category $\mathbf{C}$, in addition to the Eilenberg-Moore category we can construct another category $\mathbf{C}_T$ and an adjunction $F \dashv U$, $\eta : 1 \to UF$, $\epsilon : FU \to 1$ with $U : \mathbf{C}_T \to \mathbf{C}$ such that:

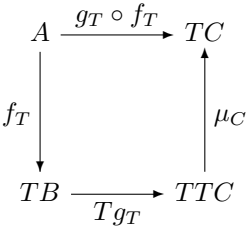
$$T = U \circ F$$
$$\eta = \eta \quad \text{(the unit)}$$
$$\mu = U\epsilon_F$$

This category $\mathbf{C}_T$ is called the *Kleisli Category* of the adjunction, and is defined as follows:

- the objects are the same as those of $\mathbf{C}$, but written $A_T, B_T, \ldots$,
- an arrow $f_T : A_T \to B_T$ is an arrow $f : A \to TB$ in $\mathbf{C}$,
- the identity arrow $1_{A_T} : A_T \to A_T$ is the arrow $\eta_A : A \to TA$ in $\mathbf{C}$,
- for composition, given $f_T : A_T \to B_T$ and $g_T : B_T \to C_T$, the composite $g_T \circ f_T : A_T \to C_T$ is defined to be

$$\mu_C \circ Tg_T \circ f_T$$

as indicated in the following diagram:



Verify that this indeed defines a category, and that there are adjoint functors $F : \mathbf{C} \to \mathbf{C}_T$ and $U : \mathbf{C}_T \to \mathbf{C}$ giving rise to the monad as $T = UF$, as claimed.

9. The notion of a *coalgebra* for an endofunctor $P : \mathcal{S} \to \mathcal{S}$ on an arbitrary category $\mathcal{S}$ is exactly dual to that of a P-algebra. Determine the *final* coalgebra for the functor

$$P(X) = 1 + A \times X$$

for a set A. (*Hint:* Recall that the initial algebra consisted of *finite* lists $a_1, a_2, \ldots$ of elements of A.)

This page intentionally left blank

REFERENCES

The following works are referred to in the text. They are also recommended for further reading.

1. Eilenberg, S. and S. Mac Lane (1945) "General theory of natural equivalences," *Transactions of the American Mathematical Society* 58, 231–294.
2. Johnstone, P.T. (1982) *Stone Spaces*, Cambridge: Cambridge University Press.
3. Johnstone, P.T. (2002) *Sketches of an Elephant: A Topos Theory Compendium*, 2 vols, Oxford: Oxford University Press.
4. Lambek, J. and P. Scott (1986) *Introduciton to Higher-Order Categorical Logic*, Cambridge: Cambridge University Press.
5. Lawvere, F.W. (1969) "Adjointness in Foundations," *Dialectica*, 23, 281–296.
6. Mac Lane, S. (1971) *Categories for the Working Mathematician*, Springer: Berlin Heidelberg New York, 2nd ed. 1998.
7. Mac Lane, S. and I. Moerdijk (1992) *Sheaves in Geometry and Logic: A First Introduction to Topos Theory*, Springer: Berlin Heidelberg New York.
8. McLarty, C. (1995) *Elementary Categories, Elementary Toposes*, Oxford: Oxford University Press.

This page intentionally left blank

INDEX

0, 7
1, 7
2, 7, 129
3, 7, 75

adjoint, 179–219
  and orders, 191–193
  cocompletion as, 182
  examples of, 187–193
  existence of, 197, 210
  exponential as, 187
  freeness as, 179
  image as, 192
  left/right, 181
  limit as, 188
  open interior as, 191
  polynomial ring as, 190
  in a poset, 192
  preservation of limits, 197
  product as, 181
  quantifiers as, 193
  rules of inference, 194
adjoint functor, 2
Adjoint functor theorem, 210
  special, 214
adjunction, 2, *see* adjoint
  counit of, 187
  Hom-set definition, *187*
  preliminary definition, *180*
  unit of, 181
AFT, *see* Adjoint functor theorem
algebra, 1, 4, 54
  boolean, *see* boolean algebra
  for a monad, 229
    and endofunctors, 242
  for an endofunctor, 236, *237*
    and monads, 242
  heyting, *see* heyting algebra
  initial, 238, 241
  Lindenbaum-Tarski, 117
  and logic, 193
  presentation, 59, 61, 223
arrow, *4*
  classifying, 174
associative law, *5*
atom
  in a Boolean algebra, 153
automorphism, 11
  inner, 145

BA, 129
Beck's theorem, 244
bifunctor, 139
Boolean algebra, 14, *29*, 30, 45, 113, 129, 131
  finite, 153

$\mathcal{C}(X)$, 129
Cantor, 241
Cat, 8, 22, 28, 125, 127
  cartesian closed, 139
category, 2, *4*
  algebraic, 234
  arrow, *14*
  of Boolean algebras, 129
  cartesian closed, 10, *108*, 108–112, 145
    and lambda calculus, 119
    equational, 118
  of categories, 8, 125
  cocomplete, 168
  complete, 167, 213
  concrete, 6, 12, 128
  congruence, 71
  constructions on, 13–15
  coslice, *15*, 23
  of diagrams, 159, 175
  discrete, 10, 91, 127
  dual, 13
  Eilenberg-Moore, 229
  of elements, 169, 208
  finite, 7
  finitely presented, 73–74
  free, 16–21, 23
  of functors, *see* functor category
  of graphs, 144
  of groups, 68–70
  homomorphism of, 8
  homomorphism theorem, 73
  Kleisli, 233, 246
  large, *22*
  with limits, 92
  locally cartesian closed, 202, *206*

category (*contd.*)
locally small, *22*
monoid, 10
opposite, *13*
of partial maps, 103
pointed sets, 15, 151
poset, 9, 15
preorder, 8
product, *13*
with products, *42*, 41–42
of propositions, 220
quotient, 71
skeletal, 156
slice, *15*, 152
and indexed families, 203
small, 13, *22*
of subobjects, 78
of types, 38, 120
category theory, 1–2, 8, 11, 21, 25, 34, 47, 125
and adjoints, 179
history of, 1
Cayley's theorem, 11–12, 161
CCC, *see* category, cartesian closed
change of base, 202
choice
axiom of, 33, 45
function, *see* function, choice
co-, *49*
coalgebra
for a comonad, 235
cocompletion of a category, 199
cocone, 93
codomain, *4*, 49
coequalizer, *58*, 57–62
of monoids, 62
of categories, 126
of posets, 63, 126
of sets, 63
cogenerating set, 214
colimit, 93, 95–102
creation of, 98
ω-colimit, 241
colimits
of posets, 102
comonad, *234*, 234–235
and adjoints, 235
component, *134*
composition, *4*
of functions, 3
of partial maps, 103
computer science, 2, 9, 31
concatination, 16
cone, *90*
as a natural transformation, 144
congruence, 71–73, 75
constant, 30
contravariant, 95
coproduct, *49*, 49–54
of categories, 125
in a CCC, 167
of groups, 52
injection, 49
of monoids, 50, 52
in a posets, 51
of posets, 51
of sets, 50, 97
of spaces, 51
covariant, 42
ωCPO, 100, 109, 123
map of, 100
strict, 123

data type, 9
deduction, 9, 115
transcendental, 142
denotational semantics, 9, 121
diagram
commutative, 3
as a functor, 159
of a group, 66
hyperdoctrine, 221
product, 35
pullback, *81*
of type, *90*
distributive law, 114
and adjoints, 198
domain, *4*, 49
duality, 14, 47–49, 58, 63, 235
conceptual, 48
and equivalence, 152
formal, 48
principle, *47*
in **Sets**, 138
stone, 131–133, 152
of vector spaces, 137

element
generalized, 29–32, 35, 79, 81
global, 30
variable, 31
endofunctor, 218, 225, 236–244
epi, *see* epimorphism
epimorphism, 25, 45, 58
of monoids, 26
split, 33
equalizer, *54*, 54–57
of abelian groups, 57

of functors, 125
of monoids, 57
of posets, 57
of spaces, 57
equivalence
of categories, *147*, 146–155
and isomorphism, 148
and duality, 152
class, 57
relation, 57, 63, 79
equivalent categories, 147
evaluation, *107*
exponent, 95
exponential, *107*, 105–123
of categories, 139, 142
of ωCPOs, 109
of diagrams, 172
of graphs, 143
of groups, 145
in a poset, 113
of posets, 108

Fib, 209
and presheaves, 209
fibration, 209
field, 129
filter, 30
maximal, 30
ultra-, *see* ultrafilter
fixed point, 101, 239, 243
forgetful functor, *see* functor, forgetful
foundations, 21–22
free, 16
category, 18
cocompletion, 171
monoid, 16, 51, 179
monad, 232
free algebra, 214
Freyd, 210, 214
Fun, 136
function, 1–3
bijective, 11
characteristic, 55, 87, 130, 174
choice, 33
continuous, 6
equality of, 3
identity, 4
injective, 5, 25
monotone, *6*
recursive, 6, 217
surjective, 25, 45
functor, 2, *8*, 33, 125–127
adjoint, *see* adjoint, *see* adjoint
cocontinuous, 199
codiagonal, 126
comonadic, 235
comparison for a monad, 234
composition of, 8
constant, 144, 167
continuous, *94*
contravariant, 95
coproduct, 54
embedding, 161
endo-, *see* endofunctor
exponential, 110
factoring, 156
faithful, 126
forgetful, 15, 16, 19, 44, 126
adjoint, 215
and limits, 98
full, 126
of groups, 70
identity, 8
injective, 126
monadic, 234
morphism of, 133
polynomial, *see* polynomial functor
of posets, 9
precomposition, 201
product, 41
product preserving, 44
projection, 13
representable, *see* representable functor
set-valued, 159
functor category, 136, 142–146, 159
cocompleteness of, 168
completeness of, 167
of groups, 145
of posets, 145

Galois connection, 191
generator, *128*
generators, 19, 59
for a category, 73, 75
for a CCC, 122
insertion of, 16, 135
graph, 6, 23
directed, *18*
as a functor, 144, 160
underlying, 19
Graphs, 144
Grothendieck, 2
group, 5, *11*, 13, 65–75, 127
abelian, 53, 63
abstract, 65
as a category, 70–74
factor, 69
homomorphism theorem, 69

group (*contd.*)
in a category, 65–68, 75
in lambda calculus, 68
normal subgroup, 68
ordered, 67
permutation, 11
representation, 71, 127
structure, 128, 236
topological, 67
Group, 22
group theory, 1
groupoid, *145*
groups
direct limit of, 97

heyting algebra, *113*, 113–118
and adjoints, 193
Hom-set, 10, *42*, 42–44
homomorphism
of algebras, 237
category, 8
complete, 154
graph, 19
group, 11, 67, 127
monoid, 10
ring, 30

ideal, 15
adjoint, 202
identity arrow, *4*
indexed family, 152
and adjoints, 202–206
$\int_{\mathbf{C}} P$, 169, 208
invariant, 32
IPC, *see* propositional calculus, intuitionistic
isomorphic, *11*
isomorphism, *11*, 11–13, 17, 23, 27
natural, 136

junk, 16, 17

Kan extension, 201
kernel
of a functor, 72
of a group homomorphism, 69
Kleene closure, 16

λ-calculus, *see* lambda calculus
lambda calculus, 9, 38, 68, 119–123, 167
and CCC, 122
completeness theorem, 121
theory in, 121
Lambek's lemma, 239, 244
large category, *see* category, large
lattice, 113
Lawvere, 2, 193, 221
LCCC, *see* category, locally cartesian closed
lifts, 34
$\varinjlim$, 97
$\overrightarrow{\lim}$, *91*
$\overleftarrow{\lim}$, *90*
creation of, 98
direct, 97
equalizer as, 91
product as, 91
pullback as, 92
terminal object as, 92
limits, 89–96
of categories, 125
of diagrams, 167
and monomorphisms, 103
preservation of, 94–96
preserved by adjoints, 197
by products and equalizers, 92
locally small category, *see* category, locally small
logic, 2, 9, 31, 61, 68, 119, 167, 220
and adjoints, 193
categorical, 79
first-order, 193
higher-order, 167, 176
intuitionistic, 114
modal, 228, 235
positive, 116
and topoi, 176

Mac Lane, 1, 34, 54
mapping, 6, 8
property, *see* universal mapping property
Mon, 10
monad, *227*, 223–247
algebra for, *229*
and adjoints, 223, 225, 229
Eilenberg-Moore category, 229
on a poset, 228
powerset, 228
monadic, 234
mono, *see* monomorphism
monoid, *10*, 26, 52, 61, 239
free, *17*, 16–18
and monad, 227
monomorphism, 25, 45, 56, 77
as a limit, 102
pullback of, 103
split, 33

morphism, *5*
  of cocones, 96
  of cones, *90*
  of T-algebras, 229

natural numbers object, *217*, 215–219, 239, 244
natural transformation, 2, *134*, 135–139, 142
naturality, 133–134
Newton's method, 101, 241
NNO, *see* natural numbers object
noise, 16, 17
nonsense
  abstract, 125
numbers, 10, 11, 18, 217
  ordinal, 146

object, *4*
  generating, *see* generator
  initial, 28–30
  natural numbers, *see* natural numbers object
  projective, 33, 45
  terminal, 28–30
objects
  family of, 33
Ord, 146

Par(**C**), 103
Par, 103, 150
partial map, 103, 150
Π, 203
point, 30, 151
points
  enough, 31
polynomial functor, 237
  generalized, 238
  and trees, 240
Pos, 6, 22
  not monadic, 234
poset, *6*, 26, 29, 51, 127
  complete, 114
  fibration of, 209
preorder, *8*, 79, 214
presheaf, *161*
product, *35*, 34–42
  cartesian, 34
  of categories, 37, 125
  of functors, 135
  of monoids, 37
  in a poset, 37
  of posets, 37
  of sets, 36
  of topological spaces, 37
programming language, 9
projective, *see* object, projective
propositional calculus, 114–118, 167
  and adjoints, 193
  completeness theorem, 117
  intuitionistic, 115
propositional function, 87
  in a topos, 176
pseudo-inverse, *147*
pullback, *80*, 80–88
  and characteristic functions, 87
  and extensions, 87
  functor, 85, 152
  inverse image as, 83
  by products and equalizers, 83
  properties of, 84–88
  and reindexing, 88
  of sets, 83
  two, lemma, 84
pushout, 103
  of sets, 95

quantifiers
  geometric interpretation, 195
quantifiers as adjoints, 176, 193–196
quotient, 57, *58*
  of sets, 63

RAPL, 197–202
recursion, 239, 241
reindexing, 203
Rel, 7, 23
relation, 7, 59
  composition, 7
  identity, 7
  local membership, 79
relative product, 7
representable functor, 42, 127, 160
  characterization of, 213
  and colimits, 95
  and limits, 94
  contravariant, 95, 128
representation, 11, 161
retract, *33*
retraction, 33–34
ring, 26
  of functions, 129
  pointed, 190
  polynomial, 190

SAFT, *see* Adjoint functor theorem, special
Scott, Dana, 123

section, 33–34
semigroup, 10
set, 3, 10, 13
  function, 106
  pointed, 15, 151
  power, 29, 79, 114, 130, 228, 238
    as an adjoint, 220
  structured, 5, 10, 26, 37, 162
    and functors, 160
set theory, 21
$\mathbf{Sets}_*$, 15, 151
sets
  cumulative hierarchy, 98
  family of, 88, 96, 152, 202
Sets, 5, 22
  cartesian closed, 108
$\mathbf{Sets}^{\mathbf{C}}$, 159
  and adjoint functors, 201
  cartesian closed, 174
  cocomplete, 168
  complete, 167
  locally cartesian closed, 209
  slice category of, 207
  topos, 175
$\mathbf{Sets}_{\mathrm{fin}}$, 5, 146, 153
$\mathbf{Sets}^I$, 152
$\mathbf{Sets}^I$, 202
sieve, *175*
Σ, 203
small category, *see* category, small
solution set, 210
space, *see* topological space
Stone duality, *see* duality, stone
Stone duality theorem, 133
Stone representation, 133
Stone space, 155
string, 233
structure, 1, 25, 31, 129, 226
  finitary, 237
  representable, 127–131
Sub, 78
subcategory, 125
  full, 126
subobject, *77*, 77–80
  classifier, 174, 177
    for presheaves, 175
  as an extension, 176

T-algebra, 214
T-algebra, 223
  for a monad, 229
test object, 13, 31
Top, 22
topological space, 6, 37, 51, 114, 129, 155, 235, 246
  interior, 191
topos, *175*, 174–176
  and logic, 176
transformation
  natural, *see* natural transformation
transpose, *107*
tree, 239, 240
triangle identities, *224*, 223–225
triple, *see* monad
twist, 136
type theory, 38–40, 119–123, *see* lambda calculus

ultrafilter, *30*, 45, 95, 131, 153
  principal, 132
UMP, *see* universal mapping property
unit law, *5*
universal, 17, 105
  subobject, 174
universal mapping property, 17, 19, 25, 34, 122
  and adjoints, 180
  of a congruence, 72
  of coequalizers, 58
  of coproducts, 49
  of equalizers, 54
  of exponentials, 107
  of free categories, 20
  of free monoids, 17
  of function sets, 106
  of initial objects, 28
  of limits, 91
  of natural numbers, 216
  of polynomials, 191
  of products, 35, 45
  of pullbacks, 80
  of terminal objects, 28
  of Yoneda, 171, 199

vector space, 6, 26, 137
  finite dimensional, 138

well powered, 214
word, 16, 53
  empty, 16

Yoneda embedding, *161*, 160–162, 165
Yoneda lemma, 162, 166, 208