Universitext

Editorial Board
(North America):

S. Axler
F.W. Gehring
K.A. Ribet

Springer Science+Business Media, LLC

Universitext

Editors (North America): S. Axler, F.W. Gehring, and K.A. Ribet

Aksoy/Khamsi: Nonstandard Methods in Fixed Point Theory
Andersson: Topics in Complex Analysis
Aupetit: A Primer on Spectral Theory
Bachman/Narici/Beckenstein: Fourier and Wavelet Analysis
Bdescu: Algebraic Surfaces
Balakrishnan/Ranganathan: A Textbook of Graph Theory
Balser: Formal Power Series and Linear Systems of Meromorphic Ordinary Differential Equations
Bapat: Linear Algebra and Linear Models (2nd ed.)
Berberian: Fundamentals of Real Analysis
Boltyanski/Efremovich: Intuitive Combinatorial Topology. Translated by Abe Shenitzer
Booss/Bleecker: Topology and Analysis
Borkar: Probability Theory: An Advanced Course
Böttcher/Silbermann: Introduction to Large Truncated Toeplitz Matrices
Carleson/Gamelin: Complex Dynamics
Cecil: Lie Sphere Geometry: With Applications to Submanifolds
Chae: Lebesgue Integration (2nd ed.)
Charlap: Bieberbach Groups and Flat Manifolds
Chern: Complex Manifolds Without Potential Theory
Cohn: A Classical Invitation to Algebraic Numbers and Class Fields
Curtis: Abstract Linear Algebra
Curtis: Matrix Groups
DiBenedetto: Degenerate Parabolic Equations
Dimca: Singularities and Topology of Hypersurfaces
Edwards: A Formal Background to Mathematics I a/b
Edwards: A Formal Background to Mathematics II a/b
Farenick: Algebras of Linear Transformations
Foulds: Graph Theory Applications
Friedman: Algebraic Surfaces and Holomorphic Vector Bundles
Fuhrmann: A Polynomial Approach to Linear Algebra
Gardiner: A First Course in Group Theory
Gårding/Tambour: Algebra for Computer Science
Goldblatt: Orthogonality and Spacetime Geometry
Gustafson/Rao: Numerical Range: The Field of Values of Linear Operators and Matrices
Hahn: Quadratic Algebras, Clifford Algebras, and Arithmetic Witt Groups
Heinonen: Lectures on Analysis on Metric Spaces
Holmgren: A First Course in Discrete Dynamical Systems
Howe/Tan: Non-Abelian Harmonic Analysis: Applications of $SL(2, \mathbb{R})$
Howes: Modern Analysis and Topology
Hsieh/Sibuya: Basic Theory of Ordinary Differential Equations
Humi/Miller: Second Course in Ordinary Differential Equations
Hurwitz/Kritikos: Lectures on Number Theory
Jennings: Modern Geometry with Applications
Jones/Morris/Pearson: Abstract Algebra and Famous Impossibilities
Kannan/Krueger: Advanced Analysis
Kelly/Matthews: The Non-Euclidean Hyperbolic Plane

(continued after index)

Saunders Mac Lane
Ieke Moerdijk

Sheaves in Geometry and Logic

A First Introduction to Topos Theory

Springer

Saunders Mac Lane
Department of Mathematics
University of Chicago
Chicago, IL 60637
USA

Ieke Moerdijk
Mathematical Institute
University of Utrecht
3508 TA Utrecht
The Netherlands

Editorial Board
(North America)

S. Axler
Mathematics Department
San Francisco State University
San Francisco, CA 94132
USA

F.W. Gehring
Mathematics Department
East Hall
University of Michigan
Ann Arbor, MI 48109-1109
USA

K.A. Ribet
Mathematics Department
University of California at Berkeley
Berkeley, CA 94720-3840
USA

Mathematics Subject Classification (2000): 03G30, 18B25, 14F20, 18F20, 54B40

Library of Congress Cataloging-in-Publication Data
Mac Lane, Saunders
Sheaves in geometry and logic: a first introduction to topos
theory/Saunders Mac Lane, Ieke Moerdijk.
p. cm. — (Universitext)
Includes bibliographical references and index.
ISBN 978-0-387-97710-2 **ISBN 978-1-4612-0927-0 (eBook)**
DOI 10.1007/978-1-4612-0927-0
1. Toposes. I. Moerdijk, Ieke. II. Title.
QA169.M335 1992
512′.55—dc20 91-33709

ISBN 978-0-387-97710-2 Printed on acid-free paper.

© 1992 Springer Science+Business Media New York
Originally published by Springer-Verlag New York, Inc. in 1992
All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher Springer Science+Business Media, LLC,
except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.
The use of general descriptive names, trade names, trademarks, etc., in this publication, even if the former are not especially identified, is not to be taken as a sign that such names, as understood by the Trade Marks and Merchandise Marks Act, may accordingly be used freely by anyone.

9 8 7 6 5 4 SPIN 10992173

springeronline.com

Dedicated to the memory of J. Frank Adams

Preface

We dedicate this book to the memory of J. Frank Adams. His clear insights have inspired many mathematicians, including both of us. In January 1989, when the first draft of our book had been completed, we heard the sad news of his untimely death. This has cast a shadow on our subsequent work.

Our views of topos theory, as presented here, have been shaped by continued study, by conferences, and by many personal contacts with friends and colleagues—including especially O. Bruno, P. Freyd, J.M.E. Hyland, P.T. Johnstone, A. Joyal, A. Kock, F.W. Lawvere, G.E. Reyes, R. Solovay, R. Swan, R.W. Thomason, M. Tierney, and G.C. Wraith. Our presentation combines ideas and results from these people and from many others, but we have not endeavored to specify the various original sources. Moreover, a number of people have assisted in our work by providing helpful comments on portions of the manuscript. In this respect, we extend our hearty thanks in particular to P. Corazza, K. Edwards, J. Greenlees, G. Janelidze, G. Lewis, and S. Schanuel.

Our work on the book has been supported by the Netherlands Science Foundation (NWO) and by the Department of Mathematics at the University of Chicago. Earlier related work was enthusiastically encouraged by Frank Adams during a visit by one of us (S.M.) to Cambridge University in 1972. Moreover, lectures by S.M. on the occasion of several visits to the University of Heidelberg were encouraged by A. Dold and D. Puppe. We gratefully note that, from the beginning of our joint project in May of 1988, Peter May has made effective arrangements for many visits by I.M. to Chicago. We are likewise grateful to M.C. Pedicchio, who arranged for our joint visit in 1990 to The University of Trieste, where we wrote Chapter VIII.

Our special thanks go to Walter Carlip, who typed up the entire manuscript with verve and understanding, and to Springer-Verlag for the smooth production of the resulting book.

Saunders Mac Lane, Ieke Moerdijk
Chicago and Utrecht, June 1991

Contents

Preface		vii
Prologue		1
Categorical Preliminaries		10
I.	**Categories of Functors**	24
	1. The Categories at Issue	24
	2. Pullbacks	29
	3. Characteristic Functions of Subobjects	31
	4. Typical Subobject Classifiers	35
	5. Colimits	39
	6. Exponentials	44
	7. Propositional Calculus	48
	8. Heyting Algebras	50
	9. Quantifiers as Adjoints	57
	Exercises	62
II.	**Sheaves of Sets**	64
	1. Sheaves	65
	2. Sieves and Sheaves	69
	3. Sheaves and Manifolds	73
	4. Bundles	79
	5. Sheaves and Cross-Sections	83
	6. Sheaves as Étale Spaces	88
	7. Sheaves with Algebraic Structure	95
	8. Sheaves are Typical	97
	9. Inverse Image Sheaf	99
	Exercises	103
III.	**Grothendieck Topologies and Sheaves**	106
	1. Generalized Neighborhoods	106
	2. Grothendieck Topologies	109
	3. The Zariski Site	116
	4. Sheaves on a Site	121

5. The Associated Sheaf Functor 128
6. First Properties of the Category of Sheaves 134
7. Subobject Classifiers for Sites 140
8. Subsheaves 145
9. Continuous Group Actions 150
Exercises 155

IV. First Properties of Elementary Topoi 161
1. Definition of a Topos 161
2. The Construction of Exponentials 167
3. Direct Image 171
4. Monads and Beck's Theorem 176
5. The Construction of Colimits 180
6. Factorization and Images 184
7. The Slice Category as a Topos 190
8. Lattice and Heyting Algebra Objects in a Topos 198
9. The Beck-Chevalley Condition 204
10. Injective Objects 210
Exercises 213

V. Basic Constructions of Topoi 218
1. Lawvere-Tierney Topologies 219
2. Sheaves 223
3. The Associated Sheaf Functor 227
4. Lawvere-Tierney Subsumes Grothendieck 233
5. Internal Versus External 235
6. Group Actions 237
7. Category Actions 240
8. The Topos of Coalgebras 247
9. The Filter-Quotient Construction 256
Exercises 263

VI. Topoi and Logic 267
1. The Topos of Sets 268
2. The Cohen Topos 277
3. The Preservation of Cardinal Inequalities 284
4. The Axiom of Choice 291
5. The Mitchell-Bénabou Language 296
6. Kripke-Joyal Semantics 302
7. Sheaf Semantics 315
8. Real Numbers in a Topos 318
9. Brouwer's Theorem: All Functions are Continuous 324
10. Topos-Theoretic and Set-Theoretic Foundations 331
Exercises 343

VII. Geometric Morphisms 347
1. Geometric Morphisms and Basic Examples 348
2. Tensor Products 353
3. Group Actions 361
4. Embeddings and Surjections 366
5. Points 378
6. Filtering Functors 384
7. Morphisms into Grothendieck Topoi 390
8. Filtering Functors into a Topos 394
9. Geometric Morphisms as Filtering Functors 399
10. Morphisms Between Sites 407
Exercises 414

VIII. Classifying Topoi 419
1. Classifying Spaces in Topology 420
2. Torsors 423
3. Classifying Topoi 432
4. The Object Classifier 434
5. The Classifying Topos for Rings 437
6. The Zariski Topos Classifies Local Rings 445
7. Simplicial Sets 450
8. Simplicial Sets Classify Linear Orders 455
Exercises 466

IX. Localic Topoi 470
1. Locales 471
2. Points and Sober Spaces 473
3. Spaces from Locales 475
4. Embeddings and Surjections of Locales 480
5. Localic Topoi 487
6. Open Geometric Morphisms 491
7. Open Maps of Locales 500
8. Open Maps and Sites 506
9. The Diaconescu Cover and Barr's Theorem 511
10. The Stone Space of a Complete Boolean Algebra 514
11. Deligne's Theorem 519
Exercises 521

X. Geometric Logic and Classifying Topoi 526
1. First-Order Theories 527
2. Models in Topoi 530
3. Geometric Theories 533
4. Categories of Definable Objects 539

5. Syntactic Sites 553
6. The Classifying Topos of a Geometric Theory 559
7. Universal Models 566
Exercises 569

Appendix: Sites for Topoi 572
1. Exactness Conditions 572
2. Construction of Coequalizers 575
3. The Construction of Sites 578
4. Some Consequences of Giraud's Theorem 587

Epilogue 596

Bibliography 603

Index of Notation 613

Index 617

Prologue

A startling aspect of topos theory is that it unifies two seemingly wholly distinct mathematical subjects: on the one hand, topology and algebraic geometry, and on the other hand, logic and set theory. Indeed, a topos can be considered both as a "generalized space" and as a "generalized universe of sets". These different aspects arose independently around 1963: with A. Grothendieck in his reformulation of sheaf theory for algebraic geometry, with F. W. Lawvere in his search for an axiomatization of the category of sets and that of "variable" sets, and with Paul Cohen in the use of forcing to construct new models of Zermelo–Frænkel set theory.

The study of cohomology for generalized spaces led Grothendieck to define his notion of a topos. The cohomology was to be one with variable coefficients—for example, varying under the action of the fundamental group, as in N. E. Steenrod's work in algebraic topology, or, more generally, varying in a sheaf. The notion of a sheaf has its origins in the analytic continuation of functions, as initiated in the 19th century and then formulated rigorously in H. Weyl's famous book on the "idea" of the Riemann surface. For several complex variables the study of domains of holomorphy and of the Cousin problems gradually led H. Cartan and K. Oka in the 1940's to study ideals on a domain. They were in effect sheaves; thus Cartan in 1944 spoke of "coherent systems of punctual ideals", while in 1949 Oka discussed "ideals with indeterminate domain". Then shortly after World War II, J. Leray published the first general and explicit definition of a sheaf on a space, described in terms of the closed sets of that space. H. Cartan, building on the ideas of Leray, rephrased the definition of sheaves in terms of open sets in his seminars of 1948–49 and 1950–51; in the course of these seminars, Lazard introduced the equivalent definition of a sheaf on a space X as an étale map into X. The subtle equivalence between these two notions is a central motivation of topos theory.

Roughly speaking, a sheaf A of abelian groups on a topological space X is a family of abelian groups A_x, parametrized by the points $x \in X$ in a suitably "continuous" way. This means in particular that the disjoint union $\coprod A_x$ of all these groups is a space, so topologized that the projection of this space into X (sending each group A_x to the point x) is continuous and also étale, in the sense that the topology on the

disjoint union is "horizontal" to match the topology of X, while the algebra (the abelian group structure of the various A_x) is vertical, along each fiber A_x. For each open set $U \subseteq X$ one can then consider the "sections" s over U of the sheaf A: each section is a function which selects—again in a suitably continuous way—for each point $x \in U$ an element $s(x)$ in the corresponding abelian group A_x. Thus, given a smaller open set $V \subseteq U$, each section s over U can be restricted to the smaller V. And conversely, the whole section s over U can be recovered by collating the restrictions of s to each of the smaller open sets V_i in some covering $U = \bigcup V_i$. For example, two sections s over V and s' over V' yield a new section $s \cup s'$ over $V \cup V'$, provided s coincides with s' on the overlap $V \cap V'$. Then the sheaf A can be described wholly in terms of all these sections s for all open sets U, together with these operations of restriction and collation. With this development of the notion of sheaf it became possible to define a corresponding cohomology of a topological space with sheaf coefficients.

Then J. P. Serre and others realized that such sheaves could be used not only in topology but also in algebraic geometry, and that the construction of a sheaf on a space X could proceed from sections s defined on objects U which were not necessarily subsets $U \subseteq X$, but simply mappings $U \to X$ from some other space U into X. Thus, ideas from category theory entered, even though the tradition of Bourbaki proscribed the use of such terms. They led Grothendieck to define sheaves in a general context, replacing the partially ordered collection of open subsets of a space by objects of a category $\mathbf{C}$ in which suitable families of maps $U_i \to X$ (for $i \in I$) form "covers" of objects X in $\mathbf{C}$. Then for such a "Grothendieck topology" a sheaf became something—indeed became a functor—which could be suitably collated over each such cover. With this general notion of sheaf, various cohomologies could be formulated in a long range attack on the Weil conjectures about solutions of polynomial equations. In the early sixties, these remarkably general ideas were rapidly developed by A. Grothendieck and his collaborators—J. L. Verdier, M. Artin, M. Giraud, M. Hakim, L. Illusie, and others. The results, initially recorded in a semi-secret document, SGA IV, later appeared as an expanded "SGA IV" in three volumes of the Springer Lecture Notes, for a total of 1623 pages. They were widely influential on the whole structure of algebraic geometry and in particular eventually led to the solution of the Weyl conjectures by P. Deligne in 1974.

For Grothendieck, topology became the study of (the cohomology of) sheaves, and the sheaves "sited" on a given Grothendieck topology formed a topos—subsequently called a Grothendieck topos. Since the very notion of sheaf is thus central to topos theory, Chapter II will develop properties of sheaves on a topological space, so as to introduce the intuition of sheaves, both in terms of fibers pasted together and in

terms of sections which can be restricted and collated. The equivalence between these two notions will be discussed, as well as the way in which a continuous map between spaces leads to a geometric morphism—a suitable pair of adjoint functors between the sheaf categories. Chapter III introduces the more general notion of coverings in a category (a Grothendieck topology), the resulting "sites", as well as the topos formed as the category of all the sheaves of sets on such a site. Furthermore, it will be explained how any functor on a site can be transformed, in two steps, into a sheaf; this process of "sheafification" provides another basic example of a pair of adjoint functors.

Thus, categories (from 1945) and adjoint functors between them, as revealed by D. M. Kan in 1957, form a language indispensable for the organization and understanding of our subject. Our categorical preliminaries (before Chapter I) may serve to remind the reader of such indispensable notions as "pullback", "adjoint functor", etc., while Chapter I will *inter alia* introduce a ubiquitous adjunction which will later provide a remarkable wealth of tensor products (Chapter VII).

Categories, initially a convenient way of formulating exact sequences, diagram chasing, and axiomatic homology for topology, acquired independent life in the work of Ehresmann and his students in France, and in the United States in the work of Kan and Mac Lane, and in a group around Eilenberg at Columbia, which included in particular Barr, Freyd, Gray, Lawvere, Linton, and Tierney. Then in 1963 Lawvere embarked on the daring project of a purely categorical foundation of all mathematics, beginning with an appropriate axiomatization of the category of sets, thus replacing set membership by the composition of functions. When Lawvere heard of the properties of Grothendieck topoi, he soon observed that such a topos admits basic operations of set theory such as the formation of sets Y^X of functions (all functions from X to Y) and of power sets $P(X)$ (all subsets of X). At about the same time M. Tierney saw that Grothendieck's work could lead to an axiomatic study of sheaves. Subsequently, Lawvere and Tierney, working together at Dalhousie University, discovered an effective axiomatization of categories of sheaves of sets (and, in particular, of the category of sets) via an appropriate formulation of set-theoretic properties. Thus they defined in an elementary way, free of all set-theoretic assumptions, the notion of an "elementary topos". The early definition underwent several changes and modifications to yield a final axiomatization of a beautiful and amazing simplicity: an elementary topos is a category with finite limits, function objects Y^X (defined as adjoints) for any two objects X and Y, and a power object $P(X)$ for each object X; these are required to satisfy some simple basic axioms, much like the first-order properties of ordinary function sets and power sets in naive set theory. Chapter I will begin our exposition by exhibiting the construction of these function

objects and power objects in several concrete examples of elementary topoi—notably in categories of set-valued functors (presheaves).

Every Grothendieck topos is an elementary topos, but not conversely: the axiomatization by Lawvere and Tierney is both elementary (first-order logic, with no reference to set theory) and more inclusive than Grothendieck's. Nonetheless, many of the basic properties of sets and functions, and of sheaf categories, can be developed on the basis of the Lawvere–Tierney axioms, as shown in Chapter IV. Furthermore, Grothendieck's definition of topology in terms of coverings can be reformulated for any elementary topos in terms of "coverings" of subobjects, giving rise to a theory of sheaves and sheafification relative to a topos, as described in detail (among other things) in Chapter V.

Lawvere's basic idea, as noted above, was that a topos is a "universe of sets". In Chapter VI, we will take up this idea and compare it with some of the developments in set theory. Our first example is a topos-theoretic presentation of Cohen's work on the independence of the Continuum Hypothesis. The Continuum Hypothesis goes back to G. Cantor and can be formulated thus: any infinite subset $B \subseteq \mathbf{R}$ of the real line has either the same cardinality as the real line itself, or is denumerable (i.e., has the cardinality of the set $\mathbf{N}$ of natural numbers). Gödel had already shown in 1938 that the Continuum Hypothesis does not contradict the usual (Zermelo–Frænkel) axioms of set theory, but for a long time it was unclear whether or not the Continuum Hypothesis follows from these Zermelo–Frænkel axioms. In 1963, Paul Cohen showed that this was not the case; his method was what is now called "Cohen forcing".

Since the cardinality of the set $\mathbf{R}$ of reals is the same as that of the powerset $P(\mathbf{N})$ of the set of natural numbers, Cohen's problem can be phrased as follows: find a set B and injective functions

$$\mathbf{N} \rightarrowtail B \rightarrowtail P(\mathbf{N}) \tag{1}$$

such that there exists no surjection $\mathbf{N} \twoheadrightarrow B$, and no surjection $B \twoheadrightarrow P(\mathbf{N})$; thus the cardinality of the set B lies strictly between the cardinality of $\mathbf{N}$ and that of $P(\mathbf{N})$. To do this, Cohen considered a "universe" of sets (a model of set theory) and then expanded this universe by "forcing" an altogether new set B of subsets of $\mathbf{N}$ into this expanded universe, so that in this new universe the cardinality of B is strictly between the cardinalities of $\mathbf{N}$ and $P(\mathbf{N})$. This technique of expanding the "universe" by forcing was later rephrased by R. Solovay and D. S. Scott in terms of Boolean-valued models, where the truth predicate takes values not just "true" and "false", but all values in an arbitrary Boolean algebra.

Shortly after this, Lawvere and Tierney made the remarkable discovery that Cohen's forcing technique could be explained in terms of topoi: indeed, using exactly Cohen's constructions, one obtains a topos

(a category of sheaves), in which there exists a sheaf B which lies strictly between the sheaf of natural numbers $\mathbf{N}$ and its power sheaf $P(\mathbf{N})$, as in (1). Chapter VI will extensively discuss this sheaf-theoretic version of Cohen's proof.

Cohen also used similar methods to show that the Axiom of Choice cannot be derived from the usual Zermelo–Frænkel axioms of set theory. More recently, P. Freyd gave an elegant and noticeably simpler sheaf-theoretic proof of this fact. Freyd's construction will be presented in Chapter VI.

Around the same time as Cohen, but evidently independently, the logician S. Kripke discovered semantical interpretations, first of modal logic and shortly after of intuitionistic logic, which bore a striking similarity to some aspects of Cohen's forcing technique. Sheaf theory also explains the relation between Cohen's forcing and Kripke's models for intuitionistic logic.

Intuitionistic logic, and the mathematics based on it, originated with Brouwer's work on the foundations of mathematics, at the beginning of this century. He defined real numbers by choice sequences and insisted that all proofs be constructive. This meant that he did not allow proof by contradiction and hence that he excluded the classical *tertium non datur* (for all p, either p, or not p). His approach was not formal or axiomatic, but subsequently Heyting and others introduced formal systems of intuitionistic logic, weaker than classical logic. This may suggest that intuitionistic mathematics is a proper part of ordinary mathematics, but this is not so: for example, in intuitionistic mathematics a suitable description of real numbers $\mathbf{R}$ leads to the result that *all* functions $\mathbf{R} \to \mathbf{R}$ are continuous, as was already shown by Brouwer.

In a topological space the complement of an open set U is closed but not usually open, so among the open sets the "negation" of U should be the interior of its complement. This has the consequence that the double negation of U is not necessarily equal to U. Thus, as observed first by Stone and Tarski, the algebra of open sets is not Boolean, but instead follows the rules of the intuitionistic propositional calculus. Since these rules were first formulated explicitly by A. Heyting, such an algebra is called a Heyting algebra. The "truth values" of any topos constitute such a Heyting algebra. The basic properties of these Heyting algebras are formulated in Chapter I.

From this point of view it is not surprising that subobjects in a category of sheaves have a negation operator which belongs to a Heyting algebra. Moreover, there are quantifier operations on sheaves, defined by adjunction, which have exactly the properties of the corresponding quantifiers in intuitionistic logic. This leads to the remarkable result, foreshadowed by the observation of Stone and Tarski as well as by Scott's topological models, that the "intrinsic" logic of a topos is in general

intuitionistic. However, there can be particular sheaf categories, such as those constructed by Cohen and by Freyd, where the intuitionistic logic becomes ordinary (classical) logic.

Kripke's semantics for intuitionistic logic can also be viewed as a description of truth for the language of a suitable topos. And as a further illustration of the way in which topos theory incorporates Brouwer's ideas we will present at the end of Chapter VI a particular topos together with its real numbers $\mathbf{R}$ in which all functions $\mathbf{R} \to \mathbf{R}$ are indeed continuous.

Together with the notion of a topos, there is the notion of a map—or a "geometric morphism"—between two topoi, defined as a pair of adjoint functors having certain additional exactness properties.

A more familiar example of a pair of adjoint functors as a map comes from ring theory. If R and S are commutative rings, consider a left R- and right S-module M. For each left R-module B, the module M then yields by "homming" a left S-module $\mathrm{Hom}_R(M, B)$. In the other direction, each left S-module A yields by tensor product a left R-module $M \otimes_S A$. The "tensor product" functor

$$M \otimes_S - : (S - \mathbf{Mod}) \to (R - \mathbf{Mod})$$

between module categories is actually a left adjoint to the Hom-functor

$$\mathrm{Hom}_R(M, -) : (R - \mathbf{Mod}) \to (S - \mathbf{Mod})$$

because of the familiar isomorphism (left adjoint on the left)

$$\mathrm{Hom}_R(M \otimes_S A, B) \cong \mathrm{Hom}_S(A, \mathrm{Hom}_R(M, B))$$

between the corresponding Hom-sets. Moreover, when M is flat as an S-module (for example, when the ring S is a field) the functor $M \otimes_S -$ preserves kernels and hence exact sequences, so is an "exact" functor.

For topoi, the definition of geometric morphisms is modeled on the case of a continuous map between topological spaces. Indeed, such a map $f : Y \to X$ induces operations in both directions on sheaves. Thus, if we regard a sheaf A on the codomain space X as a family A_x of sets parametrized by the points x of X, then f induces a family $A_{f(y)}$ of sets parametrized by the points $y \in Y$. The resulting sheaf on Y is called the "inverse image" of A and denoted by f^*A. On the other hand, a sheaf B on Y, regarded as a family of sections s over open sets V of Y, yields a new family of sections over open sets U of X by composition with f: the sections over such a U are exactly the composites $s \circ f$ where s is a section of B over the open set $f^{-1}(U) \subseteq Y$. These new sections

over open sets in X form a sheaf f_*B over the space X called the direct image of B. These two operations of inverse and direct image constitute a pair of adjoint functors, for which the inverse image f^* is left adjoint to the direct image f_*. Moreover, the left adjoint f^* is (left) exact, in the sense that it preserves finite limits.

By definition, a "geometric morphism" $f\colon \mathcal{F} \to \mathcal{E}$ between any two topoi $\mathcal{E}$ and $\mathcal{F}$ is such a pair of adjoint functors

$$f^*\colon \mathcal{E} \to \mathcal{F}, \qquad f_*\colon \mathcal{F} \to \mathcal{E},$$

where the left adjoint is required to be left exact. Such geometric morphisms arise not only from continuous maps between topological spaces, but also in many seemingly quite different contexts, as will be demonstrated in Chapter VII. This chapter will also show that any geometric morphism between (Grothendieck) topoi can be viewed as a Hom-tensor adjunction, very similar to the familiar such adjunction for modules as just described. For topoi, exactness of the left adjoint (the tensor product) will again be analyzed in terms of a notion of "flatness".

In topology, continuous maps lead naturally to the construction of classifying spaces. For example, there is a classifying space B for (complex) vector bundles. This means that there is a standard ("universal") vector bundle E over B such that, for any space X, maps from X into B correspond via E to vector bundles on X: the standard bundle E over B "pulls back" along any map $X \to B$ to produce a bundle over X, and every vector bundle over X is such a pullback of the standard bundle E. There is a similar "classifying space" for cohomology: For any integer $n \geq 0$ and any abelian group π the Eilenberg–Mac Lane space $K(\pi, n)$ classifies cohomology, in the sense that for any space X (homotopy classes of) maps from X into $K(\pi, n)$ correspond to elements in the (singular) cohomology group $H^n(X, \pi)$.

In a similar way, many sorts of mathematical structures can be "classified" by a suitable topos. For example, since a topos has products of objects, one can readily describe ring-objects in a topos. There is a special topos $\mathcal{R}$, with a "universal" ring-object R in $\mathcal{R}$, which is a classifying topos for ring-objects in topoi. This means that geometric morphisms $f\colon \mathcal{E} \to \mathcal{R}$ correspond exactly to ring-objects S in $\mathcal{E}$: the inverse image of such a morphism will carry the universal ring R in $\mathcal{R}$ to a ring $f^*(R)$ in $\mathcal{E}$, and any ring-object S in $\mathcal{E}$ is of the form $f^*(R)$ for a suitable geometric morphism f.

As an introduction to the properties of "classifying topoi", we will present this example of the classifying topos $\mathcal{R}$ in Chapter VIII. In the discussion of this and other examples, the construction of the required geometric morphisms makes use of the general Hom-tensor adjunction of

Chapter VII. This adjunction makes another appearance in Chapter X, which provides a general existence theorem for classifying topoi. It is shown that for any mathematical structure, which can be described by "geometric" axioms in a suitable language, there is a classifying topos. The proof makes use of models of the language in various topoi, and relates to earlier uses (in Chapter VI) of formal languages in the context of topoi.

A topos is, in a suitable sense, a generalized space, so should have (generalized!) points. Indeed, at a given point x of an ordinary topological space X, one can erect each set A as a sort of "skyscraper" sheaf A_x on X concentrated around the point x. The resulting mapping from the category of sets into that of sheaves on X is, in fact, the direct image of a geometric morphism $\mathbf{Sets} \to \mathrm{Sh}(X)$. But an arbitrary topos $\mathcal{E}$ may not have enough "points" $\mathbf{Sets} \to \mathcal{E}$ in this sense. In order to develop an adequate comparison between topoi and spaces, it is useful to alter the definition of a space by describing a space not in terms of its points, but in terms of its open sets. The objects so defined by a lattice of open sets are called locales. Since sheaves can be described in terms of coverings by open sets, one can construct a topos $\mathrm{Sh}(X)$ consisting of all the sheaves of sets on such a locale X. Moreover, any "continuous" map $Y \to X$ between locales gives rise to a geometric morphism $\mathrm{Sh}(Y) \to \mathrm{Sh}(X)$ between such sheaf topoi.

These locales are introduced in Chapter IX. There we show that every (Grothendieck) topos $\mathcal{E}$ has an underlying locale $\mathrm{Loc}(\mathcal{E})$. Moreover, every topos is a "quotient" of the sheaf topos for some locale. More explicitly, from any topos $\mathcal{E}$ one can construct, by a method of Diaconescu, first a locale X, then the topos $\mathrm{Sh}(X)$ of sheaves on that locale, and finally a geometric morphism $\mathrm{Sh}(X) \to \mathcal{E}$. This morphism is both a surjection (like a map onto a space) and open (in a suitable sense). Thus, $\mathcal{E}$ is indeed a quotient of its "Diaconescu cover" X.

Those topoi which are "finitely generated" in an appropriate sense are said to be coherent. Deligne's theorem in Chapter IX states that each coherent topos $\mathcal{E}$ has "enough" points. More explicitly, the underlying locale $\mathrm{Loc}(\mathcal{E})$ is an ordinary topological space, and the Diaconescu cover of $\mathcal{E}$ can be replaced by an ordinary topological space X which "covers" $\mathcal{E}$ by way of a surjection from the topos $\mathrm{Sh}(X)$ of sheaves onto the coherent topos $\mathcal{E}$.

At the end of the book, the reader will find an Appendix which discusses how various different sites can represent the same topos, and an Epilogue which provides some suggestions for further reading on the subject, beyond the "First Introduction" which this book is meant to provide.

A reference to III.6.(11) is to equation (11) in section 6 of Chapter III, and similarly for theorems.

One major correction has been made in this second printing. Dr. E. Vitale discovered that our proof of Theorem VII.9.1 was incomplete. We have provided a different proof, hopefully correct.

Categorical Preliminaries

Before embarking on the actual topic of this book, we wish to review briefly the basic notions that will be used from category theory. Many readers will be familiar with these preliminaries; they should immediately start with Chapter I, referring back to these preliminaries whenever necessary. On the other hand, these preliminaries do not present sufficiently many examples and are by no means enough to constitute a proper introduction to category theory, and the reader who lacks sufficient categorical background is advised to first read some of the relevant parts of Mac Lane's [**CWM—Categories for the Working Mathematician**, 1971] (or some other such text), perhaps using the following pages as a guideline.

A *category* $\mathbf{C}$ consists of a collection of *objects* (often denoted by capital letters, $A, B, C, \dots, X, \dots$), a collection of *morphisms* (or *maps* or *arrows*) $(f, g, \dots)$, and four operations; two of these operations associate with each morphism f of $\mathbf{C}$ its *domain* $\mathrm{dom}(f)$ or $\mathrm{d}_0(f)$ and its *codomain* $\mathrm{cod}(f)$ or $\mathrm{d}_1(f)$, respectively, both of which are objects of $\mathbf{C}$. One writes $f\colon C \to D$ or $C \xrightarrow{f} D$ to indicate that f is a morphism of $\mathbf{C}$ with domain C and codomain D, and one says that f is a morphism *from C to D*. The other two operations are an operation which associates with each object C of $\mathbf{C}$ a morphism 1_C (or id_C) of $\mathbf{C}$ called the *identity morphism* of C and an operation of *composition* which associates to any pair (f, g) of morphisms of $\mathbf{C}$ such that $\mathrm{d}_0(f) = \mathrm{d}_1(g)$ another morphism $f \circ g$, their composite. These operations are required to satisfy the following axioms:

(i) $\mathrm{d}_0(1_C) = C = \mathrm{d}_1(1_C)$,
(ii) $\mathrm{d}_0(f \circ g) = \mathrm{d}_0(g), \quad \mathrm{d}_1(f \circ g) = \mathrm{d}_1(f)$,
(iii) $1_D \circ f = f, \quad f \circ 1_C = f$,
(iv) $(f \circ g) \circ h = f \circ (g \circ h)$.

In (ii)–(iv), we assume that the compositions make sense; thus, (ii) is required to hold for any pair of arrows f and g with $\mathrm{d}_0(f) = \mathrm{d}_1(g)$, and (iii) is required to hold for any two objects C and D of $\mathbf{C}$ and any morphism f from C to D, etc.

For example, there is a category **Sets** whose objects are sets and whose morphisms are functions with the usual composition. Similarly, topological spaces and continuous maps between them form a category, as do groups and homomorphisms, or vector spaces (over $\mathbf{R}$ say) and linear maps. Any partially ordered set $(P, \leq)$ gives rise to a category, with the elements of P as objects, and with precisely one morphism from p to q iff $p \leq q$; in other words, the morphisms are pairs (p, q) such that $p \leq q$, and the domain and codomain operations on a pair are given by the first and second projections. Thus, the composition operation for P is uniquely determined by the transitivity of the order relation $\leq$. We mention in particular the categories $\mathbf{0}, \mathbf{1}, \mathbf{2}, \ldots$ coming from the ordered sets $\emptyset, \{0\}, \{0, 1\}, \ldots$ of natural numbers with their usual ordering.

An example of a different nature is obtained from a group G. Such a group can be regarded as a category with only one object, call it $*$, and with the elements of the group G as morphisms, where the multiplication of the group is used as the composition operation of the corresponding category.

In an arbitrary category $\mathbf{C}$, a morphism $f\colon C \to D$ in $\mathbf{C}$ is called an *isomorphism* if there exists a morphism $g\colon D \to C$ such that $f \circ g = 1_D$ and $g \circ f = 1_C$. (This defines g uniquely, and g is called the *inverse* of f.) If such a morphism f exists, one says that C is isomorphic to D, and one writes $f\colon C \xrightarrow{\sim} D$ and $C \cong D$. The example of a category coming from a group G, as mentioned above, shows that a group is the same thing as a category with only one object in which each morphism is an isomorphism.

A morphism $f\colon C \to D$ is called an *epi(morphism)* if for any object E and any two parallel morphisms $g, h\colon D \rightrightarrows E$ in $\mathbf{C}$, $gf = hf$ implies $g = h$; one writes $f\colon C \twoheadrightarrow D$ to indicate that f is an epimorphism. Dually, $f\colon C \to D$ is called a *mono(morphism)* if for any object B and any two parallel morphisms $g, h\colon B \rightrightarrows C$ in $\mathbf{C}$, $fg = fh$ implies $g = h$; in this case, one writes $f\colon C \rightarrowtail D$. Two monomorphisms $f\colon A \rightarrowtail D$ and $g\colon B \rightarrowtail D$ with a common codomain D are called *equivalent* if there exists an isomorphism $h\colon A \xrightarrow{\sim} B$ with $gh = f$. A *subobject* of D is an equivalence class of monomorphisms into D. The collection $\mathrm{Sub}_{\mathbf{C}}(D)$ of subobjects of D carries a natural partial order defined by $[f] \leq [g]$ iff there is an $h\colon A \to B$ such that $f = gh$, where $[f]$ and $[g]$ are the classes of $f\colon A \rightarrowtail D$ and $g\colon B \rightarrowtail D$.

For **Sets** (and other familiar categories) this definition matches the usual notion of subset (or subspace, etc.).

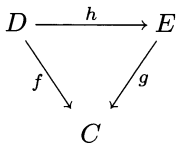
If $\mathbf{C}$ is a category, we sometimes write $\mathbf{C}_0$ for its collection of objects and $\mathbf{C}_1$ for its collection of morphisms. For two objects C and D, the collection of morphisms with domain C and codomain D is denoted by one of the following three symbols,

$$\mathrm{Hom}_{\mathbf{C}}(C, D), \qquad \mathrm{Hom}(C, D), \qquad \mathbf{C}(C, D).$$

In general, we shall not be very explicit about set-theoretical foundations, and we shall tacitly assume we are working in some fixed universe U of sets. Members of U are then called *small* sets, whereas a collection of members of U which does not itself belong to U will sometimes be referred to as a *large* set. Given such an ambient universe U, a category $\mathbf{C}$ is *locally small* if for any two objects C and D of $\mathbf{C}$ the hom-set $\mathrm{Hom}_{\mathbf{C}}(C, D)$ is a small set, while $\mathbf{C}$ is called *small* if both $\mathbf{C}_0$ and $\mathbf{C}_1$ are small sets. Of the categories mentioned above, the categories of small sets, of small topological spaces, of small vector spaces, and of small groups are all locally small but not small. The categories coming from a small poset $(P, \leq)$ or a small group G in the universe U are both small.

Given a category $\mathbf{C}$, one can form a new category $\mathbf{C}^{\mathrm{op}}$, called the *opposite* or *dual* category of $\mathbf{C}$, by taking the same objects but reversing the direction of all the morphisms and the order of all compositions. In other words, an arrow $C \to D$ in $\mathbf{C}^{\mathrm{op}}$ is the same thing as an arrow $D \to C$ in $\mathbf{C}$ (see [**CWM**, p. 33]).

Given a category $\mathbf{C}$ and an object C of $\mathbf{C}$, one can construct the *comma category* or the *slice category* $\mathbf{C}/C$ (read: $\mathbf{C}$ over C): objects of $\mathbf{C}/C$ are morphisms of $\mathbf{C}$ with codomain C, and morphisms in $\mathbf{C}/C$ from one such object $f\colon D \to C$ to another $g\colon E \to C$ are commutative triangles in $\mathbf{C}$



i.e., $gh = f$. (In [**CWM**], the notation $\mathbf{C} \downarrow C$ is used instead of $\mathbf{C}/C$; cf. [**CWM**, p. 46].) The composition in $\mathbf{C}/C$ is defined from the composition in $\mathbf{C}$, in the obvious way (paste triangles side by side).

Categories are compared by using functors. Given two categories $\mathbf{C}$ and $\mathbf{D}$, a *functor* from $\mathbf{C}$ to $\mathbf{D}$ is an operation F which assigns to each object C of $\mathbf{C}$ an object $F(C)$ of $\mathbf{D}$, and to each morphism f of $\mathbf{C}$ a morphism $F(f)$ of $\mathbf{D}$, in such a way that F respects the domain and codomain as well as the identities and the composition: $F(\mathrm{d}_0(f)) = \mathrm{d}_0(F(f))$, $F(\mathrm{d}_1(f)) = \mathrm{d}_1(F(f))$, $F(1_C) = 1_{F(C)}$, and also $F(f \circ g) = F(f) \circ F(g)$, whenever this makes sense. One writes $F\colon \mathbf{C} \to \mathbf{D}$ or $\mathbf{C} \xrightarrow{F} \mathbf{D}$. For example, there is a functor from the category of topological spaces and continuous maps to the category of sets and functions, given by sending a space to the "underlying" set of its points. If $\mathbf{C}$ is an arbitrary category and C is an object of $\mathbf{C}$, then the domain operation gives a functor

$$F\colon \mathbf{C}/C \to \mathbf{C}$$

then F is an equivalence (and G is sometimes called a *quasi-inverse* for F). Conversely, using a sufficiently strong axiom of choice, every equivalence F has a quasi-inverse; see [**CWM**, pp. 91–92].

Next, we recall several "universal" constructions. For example, in the category **Sets** of (small) sets and functions between them, there is the construction of the cartesian product $A \times B$ of two sets A and B. It comes together with two projections $\pi_1 \colon A \times B \to A$ and $\pi_2 \colon A \times B \to B$. Usually, $A \times B$ is constructed as the set of all pairs (a, b) with $a \in A$ and $b \in B$. However, up to isomorphism, the product $A \times B$ can also be described purely in terms of objects and morphisms in the category of sets, as follows. We say that an object X equipped with morphisms $\pi_1 \colon X \to A$ and $\pi_2 \colon X \to B$ is a *product* of A and B if for any other object Y and any two maps $f \colon Y \to A$ and $g \colon Y \to B$ there is a *unique* map $h \colon Y \to X$ such that $\pi_1 \circ h = f$ and $\pi_2 \circ h = g$. [This unique map is then denoted by $(f, g) \colon Y \to X$ or sometimes by $\langle f, g \rangle$ with pointed brackets.] This definition makes sense in any category and determines the object X (if it exists) to within isomorphism. It is common to denote a product of two objects A and B in an arbitrary category, if it exists, by $A \times B$. Iteration then yields products of three or more factors. In an arbitrary category, the product of two objects may or may not exist; for instance, in the category of topological spaces, the product of two spaces always exists, and it may be constructed as the cartesian product of the underlying sets, equipped with the familiar product topology. A product of an I-indexed family A_i is written $\Pi_i A_i$. For a poset $(P, \leq)$, viewed as a category in the way explained above, the product of two objects p and q is their infimum (greatest lower bound), which may or may not exist.

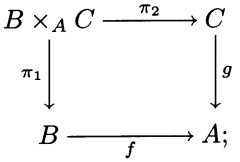
Other special constructions of sets may also be characterized purely in categorical terms, i.e., in terms of objects and morphisms. For example, the singleton set $\{*\}$ is the set S, unique up to isomorphism, for which there is exactly one morphism $A \to S$ from any other set A into S. In an arbitrary category $\mathbf{C}$, an object C with the property that for any other object D of $\mathbf{C}$ there is one and only one morphism from D to C is called a *terminal object* of $\mathbf{C}$. If it exists, it is, like $\{*\}$, unique up to isomorphism; it is often denoted by 1 or by $1_{\mathbf{C}}$ if necessary.

A construction which plays a central role in this book is that of a *pullback*, or *fibered product*. Given two functions $f \colon B \to A$ and $g \colon C \to A$ between sets, one may construct their fibered product as the set

$$B \times_A C = \{ (b, c) \in B \times C \mid f(b) = g(c) \}.$$

Thus, $B \times_A C$ is a subset of the product, and, therefore, comes equipped with two *projections* $\pi_1 \colon B \times_A C \to B$ and $\pi_2 : B \times_A C \to C$ which fit

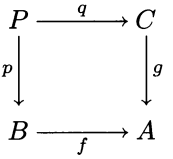
into a commutative diagram



(1)

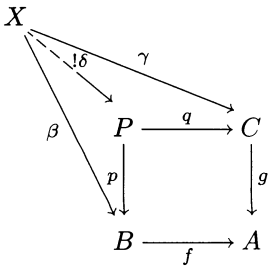
i.e., $f\pi_1 = g\pi_2$. This diagram has the property that given any other set X and functions $\beta\colon X \to B$ and $\gamma\colon X \to C$ such that $f\beta = g\gamma$, there is a unique function $\delta\colon X \to B \times_A C$ with $\pi_1\delta = \beta$ and $\pi_2\delta = \gamma$ [namely, the function $\delta(x) = (\beta(x), \gamma(x))$]. This property determines the set $B \times_A C$ up to isomorphism. One says that (1) is the "universal" commutative square on the data f and g. If $A = \{*\}$, $B \times_A C$ is the product $B \times C$.

In a general category $\mathbf{C}$, one says that a commutative square



(2)

is a *pullback (square)*, or a *fibered product*, if it has the property just described for sets: given any object X of $\mathbf{C}$ and morphisms $\beta\colon X \to B$ and $\gamma\colon X \to C$ with $f\beta = g\gamma$, there is a unique $\delta\colon X \to P$ such that $p\delta = \beta$ and $q\delta = \gamma$. [This unique map δ



is usually denoted by (β, γ).] Given $f\colon B \to A$ and $g\colon C \to A$, the pullback P with its projections p and q is uniquely determined up to isomorphism (if it exists at all), and one usually writes $B \times_A C$ for this pullback. If, for given f and g, the pullback (2) exists, one also says

that the arrow p is the pullback of g *along* f (and symmetrically that q is the pullback of f along g). Notice that p is a monomorphism if g is. One says that monomorphisms are preserved by pullback along an arbitrary morphism. Incidentally, the notion of monomorphism can be described in terms of pullbacks. A morphism $f\colon B \to A$ in a category $\mathbf{C}$ is a monomorphism iff the pullback of f along itself is an isomorphism, iff the square

$$\begin{array}{ccc} B & \xrightarrow{1} & B \\ {\scriptstyle 1}\downarrow & & \downarrow{\scriptstyle f} \\ B & \xrightarrow[f]{} & A \end{array} \tag{3}$$

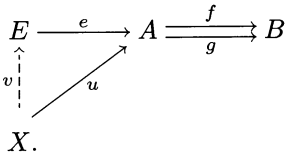
is a pullback.

There is an important "pasting-lemma" for pullback squares ([**CWM**, p. 72]). Given a commutative diagram of the form

$$\begin{array}{ccccc} Q & \longrightarrow & P & \longrightarrow & D \\ \downarrow & & \downarrow & & \downarrow \\ C & \longrightarrow & B & \longrightarrow & A \end{array} \tag{4}$$

in an arbitrary category $\mathbf{C}$, the outer rectangle is a pullback if both inner squares are pullbacks; and conversely, if the outer rectangle as well as the right-hand square are pullbacks, then so is the left-hand square.

Equalizers also deserve mention. For two parallel arrows $f\colon A \to B$ and $g\colon A \to B$ in a category $\mathbf{C}$, the *equalizer* of f and g is a morphism $e\colon E \to A$ such that $fe = ge$ and which is universal with this property; that is, given any other morphism $u\colon X \to A$ in $\mathbf{C}$ such that $fu = gu$, there is a unique $v\colon X \to E$ such that $ev = u$:



Equalizers need not always exist. However, in the category of sets the equalizer of any pair of functions $f, g\colon A \rightrightarrows B$ exists, and can be constructed as the set

$$E = \{\, a \in A \mid f(a) = g(a) \,\}$$

where $e\colon E \to A$ is set inclusion.

To each of the categorical notions of product, terminal object, pullback, and equalizer, there is a corresponding dual notion, namely, that of a *coproduct*, an *initial object*, a *pushout*, and a *coequalizer*. For example, a diagram $A \xrightarrow{u} X \xleftarrow{v} B$ in a category $\mathbf{C}$ is said to be a *coproduct* of A and B if the corresponding diagram $A \leftarrow X \rightarrow B$ in $\mathbf{C}^{\mathrm{op}}$ is a product. In other words, $A \xrightarrow{u} X \xleftarrow{v} B$ is a coproduct iff for any two morphisms $f\colon A \rightarrow Y$ and $g\colon B \rightarrow Y$ into another object Y, there is a unique morphism $h\colon X \rightarrow Y$ with $hu = f$ and $hv = g$. The coproduct of A and B, if it exists, is unique up to isomorphism and is denoted by $A + B$ or $A \amalg B$. The maps $A \xrightarrow{u} A \amalg B$ and $B \xrightarrow{v} A \amalg B$ are called the coproduct *inclusions*. For example, the disjoint sum (= disjoint union) defines a coproduct in the category of sets. A coproduct of a family A_i for $i \in I$ is written as $\amalg_i A_i$.

Similarly, an object C of $\mathbf{C}$ is an *initial object* of $\mathbf{C}$ if it is a terminal object in $\mathbf{C}^{\mathrm{op}}$. An initial object of $\mathbf{C}$, if it exists, is unique up to isomorphism and is usually denoted by 0.

Given morphisms $A \xrightarrow{f} B$ and $A \xrightarrow{g} C$ in $\mathbf{C}$, a diagram

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle g}\downarrow & & \downarrow{\scriptstyle p} \\ C & \xrightarrow[q]{} & P \end{array}$$

is called a *pushout* if the corresponding diagram in $\mathbf{C}^{\mathrm{op}}$ is a pullback. The pushout of f and g, if it exists, is unique up to isomorphism and is denoted by $P = B \amalg_A C$.

Finally, a morphism $B \xrightarrow{c} C$ in $\mathbf{C}$ is said to be a *coequalizer* of a given parallel pair of morphisms $f, g\colon A \rightrightarrows B$ iff the corresponding diagram $C \rightarrow B \rightrightarrows A$ in $\mathbf{C}^{\mathrm{op}}$ is an equalizer.

We now come to the central notion of adjoint functor, discussed at length in Chapter IV of [**CWM**]. Consider two categories $\mathbf{A}$ and $\mathbf{X}$ and two functors between them in opposite directions, say

$$F\colon \mathbf{X} \rightarrow \mathbf{A} \qquad\qquad G\colon \mathbf{A} \rightarrow \mathbf{X}. \tag{5}$$

One says that G is *right adjoint* to F (and that F is *left adjoint* to G, notation: $F \dashv G$) when for any two objects X from $\mathbf{X}$ and A from $\mathbf{A}$ there is a natural bijection between morphisms

$$\frac{X \xrightarrow{f} GA}{FX \xrightarrow{h} A}, \tag{6}$$

in the sense that each morphism f, as displayed, uniquely determines a morphism h, and conversely. This bijection is to be natural in the

following sense: given any morphisms $\alpha\colon A \to A'$ in $\mathbf{A}$ and $\xi\colon X' \to X$ in $\mathbf{X}$, and corresponding arrows f and h as in (6) the (inevitable) composites also correspond under the bijection (6):

$$\frac{X' \xrightarrow{\xi} X \xrightarrow{f} GA \xrightarrow{G\alpha} GA'}{FX' \xrightarrow{F\xi} FX \xrightarrow{h} A \xrightarrow{\alpha} A'}. \tag{7}$$

If we write this bijective correspondence as

$$\theta\colon \mathrm{Hom}_{\mathbf{X}}(X, GA) \xrightarrow{\sim} \mathrm{Hom}_{\mathbf{A}}(FX, A), \tag{8}$$

then this naturality condition can be expressed by the equation

$$\theta(G(\alpha) \circ f \circ \xi) = \alpha \circ \theta(f) \circ F(\xi). \tag{9}$$

Examples of adjoints abound in mathematics. The reader may find a list of examples on pp. 85–86 of [**CWM**], and will encounter numerous other examples in the course of this book.

With an adjunction as above, there are associated certain so-called unit and counit morphisms. Given θ as in (8), and an object X in $\mathbf{X}$, setting $A = FX$ gives a unique map

$$\eta = \eta_X\colon X \to GFX \tag{10}$$

such that $\theta(\eta_X) = 1_{F(X)}$. This map η_X is called the *unit* of the adjunction (at X). If one takes $\xi = 1_X$, $A = FX$, $f = \eta$, $\alpha = h$, and $A' = A$ in (7), the bottom composite is simply $h\colon FX \to A$, and it corresponds to the top composite $X \xrightarrow{\eta} GFX \xrightarrow{Gh} GA$. In short, η determines the adjunction, since h corresponds to $G(h) \circ \eta_X$ under the correspondence (6). This means that each f determines uniquely an h which makes the following triangle commute:

$$\begin{array}{ccc} X & \xrightarrow{\eta} & GFX \\ & \searrow^{f} & \downarrow{\scriptstyle Gh} \\ & & GA \end{array} \qquad \begin{array}{c} FX \\ \downarrow{\scriptstyle h} \\ A. \end{array} \tag{11}$$

One expresses this by saying that $\eta = \eta_X$ is *universal* among arrows from X to an object of the form GA. This also implies that when the functor G is given, the object FX is uniquely determined up to isomorphism. In other words, given a functor G, its left adjoint F (if it exists) is unique up to natural isomorphism. Also, given G and a universal arrow from each object X to some object of the form GA, the left adjoint must exist. The naturality condition (7) [or (9)] also implies that the unit

morphisms $\eta_X \colon X \to GFX$, for varying objects X of $\mathbf{X}$, constitute a natural transformation from the identity functor on $\mathbf{X}$ to the composite functor $GF \colon \mathbf{X} \to \mathbf{X}$ for $\theta(\eta_X \circ \xi) = F(\xi) = \theta(GF\xi \circ \eta_{X'})$.

Dual to the unit of an adjunction is the *counit*. In the correspondence (6), take $X = GA$ and f the identity on GA. The corresponding h is written

$$\epsilon \quad \text{or} \quad \epsilon_A \colon FGA \to A.$$

This defines a natural transformation from FG to the identity functor on $\mathbf{A}$. Its universal property is this: to every $h \colon FX \to A$ there exists a unique f which makes the following triangle commute:

$$\begin{array}{ccc} X & & FX \\ {\scriptstyle f}\,\vdots\,\downarrow & & {\scriptstyle Ff}\,\vdots\,\downarrow \quad \searrow^{h} \\ GA & & FGA \xrightarrow[\epsilon]{} A. \end{array} \tag{12}$$

In other words, ϵ is universal among arrows from an object in the image of F to A. As for the unit, this implies that given F, its right adjoint G (if it exists) is determined uniquely up to isomorphism.

In the diagram (12), one may take h to be the identity on $A = FX$. The corresponding map f is then the unit η_X of the adjunction, and we obtain a commutative triangle

$$\begin{array}{ccc} F & & G \xrightarrow{\eta G} GFG \\ {\scriptstyle F\eta}\downarrow \quad \searrow^{1} & & {\scriptstyle 1}\searrow \quad \downarrow{\scriptstyle G\epsilon} \\ FGF \xrightarrow[\epsilon F]{} F & & G \end{array} \tag{13}$$

as on the left of (13). Its dual is the right-hand triangle. Conversely, two natural transformations $\epsilon \colon FG \to \mathrm{id}$ and $\eta \colon \mathrm{id} \to GF$ which satisfy these two *triangular identities* (13) serve to make F a left adjoint of G [**CWM**, pp. 80-81].

As an example, consider the product category $\mathbf{C} \times \mathbf{C}$ of a given category $\mathbf{C}$ with itself. ($\mathbf{C} \times \mathbf{C}$ may also be viewed as a functor category $\mathbf{C}^{\mathbf{2}}$, where $\mathbf{2}$ is the category with two distinct objects 0 and 1, and identity morphisms only.) If the product $A \times B$ of any pair of objects A and B in $\mathbf{C}$ exists, this gives a functor $\times \colon \mathbf{C} \times \mathbf{C} \to \mathbf{C}$ which is right adjoint to the *diagonal* functor $\mathbf{C} \to \mathbf{C} \times \mathbf{C}$ sending A to (A, A). This follows immediately from the definition of the product.

Suppose products exist in $\mathbf{C}$. For a fixed object A of $\mathbf{C}$, one may consider the functor

$$A \times - \; : \mathbf{C} \to \mathbf{C}. \tag{14}$$

If this functor has a right adjoint (necessarily unique up to isomorphism), this adjoint is denoted by

$$(-)^A \colon \mathbf{C} \to \mathbf{C}. \tag{15}$$

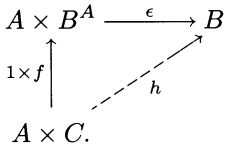
In this case A is said to be an *exponentiable* object of the category $\mathbf{C}$, while the value B^A of (15) for an object B of $\mathbf{C}$ is called the *exponential* of A and B. That $(-)^A$ is right adjoint to $A \times -$ means that for any objects B and C of $\mathbf{C}$ there is a bijective correspondence

$$\frac{C \to B^A}{A \times C \to B}, \tag{16}$$

natural in B and C. It follows from the various uniqueness properties that B^A is also a (contravariant!) functor of A (at least on those A which are exponentiable), and that (16) is also natural in A. The counit of this adjunction $A \times (-) \dashv (-)^A$ is a map

$$\epsilon \colon A \times B^A \to B \tag{17}$$

with the property that for any map $h \colon A \times C \to B$ there is a unique $f \colon C \to B^A$ such that $\epsilon \circ (1 \times f) = h$:



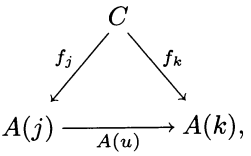
In this special case, the counit map is called the *evaluation* and denoted by e or $\mathrm{ev} \colon A \times B^A \to B$.

A category $\mathbf{C}$ is called *cartesian closed* if it has finite products (i.e., a terminal object and binary products) and if all objects of $\mathbf{C}$ are exponentiable. For example, the category of sets is cartesian closed: the exponential B^A of two sets A and B is simply the set of all functions from A to B, and the bijective correspondence (16) is the familiar process of turning a function $f \colon A \times C \to B$ of two variables into a function of a single variable in C with values in B^A.

We now turn to a brief discussion of limits and colimits. Let $\mathbf{C}$ be a fixed category. For a small category $\mathbf{J}$ (the "indexing category") we consider the functor category $\mathbf{C}^\mathbf{J}$. An object of $\mathbf{C}^\mathbf{J}$ is also called a *diagram* in $\mathbf{C}$ of type $\mathbf{J}$. For example, each object C of $\mathbf{C}$ determines a constant diagram $\Delta_\mathbf{J}(C)$ which has the same value C for all $j \in \mathbf{J}$; this defines the *diagonal* functor

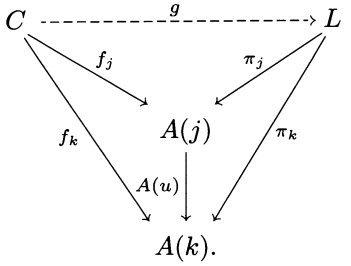
$$\Delta_\mathbf{J} \colon \mathbf{C} \to \mathbf{C}^\mathbf{J}. \tag{18}$$

A natural transformation π from the constant diagram $\Delta_{\mathbf{J}}(C)$ to some other diagram A of $\mathbf{C}^{\mathbf{J}}$ then consists of maps $f_j\colon C \to A_j$, one for each "index" j in $\mathbf{J}$, all such that the triangle



$$u\colon j \to k, \tag{19}$$

commutes for every arrow u of $\mathbf{J}$. Such a natural transformation is called a *cone* $f\colon C \to A$ on the diagram A with vertex C. In particular, a cone $\pi\colon L \to A$ with vertex L is universal to A when to every cone $f\colon C \to A$ there is a unique map $g\colon C \to L$ in $\mathbf{C}$ with $\pi_j \circ g = f_j$, for each j of $\mathbf{J}$, as in the commutative diagram



$$u\colon j \to k.$$

This universal cone $\pi\colon L \to A$ (or, less accurately, its vertex $L = \varprojlim_{\mathbf{J}} A$) is called the *limit* of the diagram A. If every diagram A in $\mathbf{C}^{\mathbf{J}}$ has a limit in this sense, then the diagonal functor $\Delta_{\mathbf{J}}$ has a right adjoint

$$\varprojlim_{\mathbf{J}}\colon \mathbf{C}^{\mathbf{J}} \to \mathbf{C}. \tag{20}$$

Indeed the counit of this adjunction is precisely the universal cone, which can be viewed as a natural transformation

$$\pi\colon \Delta_{\mathbf{J}}(L) = \Delta_{\mathbf{J}}(\varprojlim_{\mathbf{J}} A) \to A.$$

For example, if $\mathbf{J} = \mathbf{2} = \{0, 1\}$ is the discrete category with two objects 0 and 1 and only identity arrows, then a diagram in $\mathbf{C}^{\mathbf{J}}$ is just a pair of objects of $\mathbf{C}$ and a limit of that diagram is just a product of these objects. The product is thus a special case of limit. In the same way, a terminal object, a pullback, or an equalizer, as discussed above, are all special cases of limits (when $\mathbf{J}$ is the empty category, or the category $\to \bullet \leftarrow$, or $\bullet \rightrightarrows \bullet$, respectively, where we have indicated only the nonidentity morphisms).

The dual of the notion of limit is that of *colimit*. A *cocone* with vertex C on a diagram $A\colon \mathbf{J} \to \mathbf{C}$ is a map $A \to \Delta_{\mathbf{J}}(\mathbf{C})$ in the functor category $\mathbf{C}^{\mathbf{J}}$. The universal cocone on A, if it exists, is called the *colimit* of the diagram A, and its vertex is denoted by $\varinjlim_{\mathbf{J}} A$. If the colimit of any diagram of type $\mathbf{J}$ in $\mathbf{C}$ exists, this gives a functor

$$\varinjlim_{\mathbf{J}} : \mathbf{C}^{\mathbf{J}} \to \mathbf{C} \tag{21}$$

which is left adjoint to the diagonal $\Delta_{\mathbf{J}}\colon \mathbf{C} \to \mathbf{C}^{\mathbf{J}}$.

Now suppose $G\colon \mathbf{C} \to \mathbf{D}$ is a functor. If $\mathbf{J}$ is a small index category, G induces a functor $G^{\mathbf{J}}\colon \mathbf{C}^{\mathbf{J}} \to \mathbf{D}^{\mathbf{J}}$ of diagrams in the obvious way. If limits of type $\mathbf{J}$ exist in $\mathbf{C}$ and $\mathbf{D}$, one obtains a square of categories and functors

$$\begin{array}{ccc} \mathbf{C}^{\mathbf{J}} & \xrightarrow{\varprojlim_{\mathbf{J}}} & \mathbf{C} \\ {\scriptstyle G^{\mathbf{J}}}\downarrow & & \downarrow{\scriptstyle G} \\ \mathbf{D}^{\mathbf{J}} & \xrightarrow[\varprojlim_{\mathbf{J}}]{} & \mathbf{D}. \end{array} \tag{22}$$

The universal property of limits implies that there is a canonical natural transformation

$$\alpha_{\mathbf{J}}\colon G \circ \varprojlim_{\mathbf{J}} \to \varprojlim_{\mathbf{J}} \circ G^{\mathbf{J}}. \tag{23}$$

One says that G *preserves limits* (of type $\mathbf{J}$) if $\alpha_{\mathbf{J}}$ is a natural isomorphism. A basic property is that G preserves limits of any type if G has a left adjoint. Or briefly, *right adjoints preserve limits* (see [**CWM**, p. 114]). Dually, one defines preservation of colimits. The corresponding basic fact is that *left adjoints preserve colimits*.

We conclude these preliminaries by mentioning an important fact about limits and colimits in functor categories, namely, that these are computed *pointwise*. More precisely, let $\mathbf{C}$ and $\mathbf{D}$ be categories and consider the functor category $\mathbf{C}^{\mathbf{D}}$. If $\mathbf{J}$ is a small category such that limits of type $\mathbf{J}$ exist in $\mathbf{C}$, then the same is true in $\mathbf{C}^{\mathbf{D}}$, and the evaluation functor $(-)_D : \mathbf{C}^{\mathbf{D}} \to \mathbf{C}$, for any given object D of $\mathbf{D}$, preserves such limits. In other words, for a diagram $A\colon \mathbf{J} \to \mathbf{C}^{\mathbf{D}}$ of type $\mathbf{J}$ in $\mathbf{C}^{\mathbf{D}}$, one obtains a diagram $A_D\colon \mathbf{J} \to \mathbf{C}$ for each object D of $\mathbf{D}$, by setting

$$A_D(j) = A(j)(D).$$

If every such diagram A_D has a limit $L_D = \varprojlim_{\mathbf{J}} A_D$ in $\mathbf{C}$ then these limits fit together to give a functor $L\colon \mathbf{D} \to \mathbf{C}$, which is a limit for the diagram A. So for each D in $\mathbf{D}$

$$(\varprojlim_{\mathbf{J}} A)(D) \cong \varprojlim_{\mathbf{J}} A_D, \tag{24}$$

where the limit on the left is taken in $\mathbf{C}^{\mathbf{D}}$ and that on the right in $\mathbf{C}$.

The corresponding fact for colimits also holds, and gives an isomorphism

$$(\varinjlim_{\mathbf{J}} A)(D) \cong \varinjlim_{\mathbf{J}} A_D \tag{25}$$

analogous to (24). For details, we refer the reader to [**CWM**, p. 112].

I
Categories of Functors

Many constructions on various mathematical objects depend not just on the elements of those objects but also on the morphisms between them. Such constructions can thus be effectively formulated in the corresponding category of objects. A "topos" is a category in which a number of the most basic such constructions (product, pullback, exponential, characteristic function, ...) are always possible. With these constructions available, many other properties can be efficiently developed. Superficially quite different categories, arising in geometry, topology, algebraic geometry, group representations, and set theory, all turn out to satisfy the axioms defining such a topos.

1. The Categories at Issue

Our exposition starts by describing a number of specific categories which are topoi, exhibiting in each one several of the basic constructions required. These examples will pave the way to the formulation of the axioms for a topos.

In the following list of many such categories the most important examples are those numbered (i), (viii), (x), and (xi): sets, functor categories (presheaves), sheaves, and group actions.

Here is the list of examples of topoi:

(i) **Sets**, the category of all (small) sets S, T, and functions $S \to T$ between them.

(ii) **Sets** $\times$ **Sets**, the category of all pairs of sets, with morphisms pairs of functions.

(iii) $\mathbf{Sets}^n$, the category of all n-tuples of sets with morphisms all n-tuples of functions. Here n is a fixed natural number.

(iv) $\mathbf{B}G$, or G-**Sets**, the category of all representations of a fixed group G; where a representation of G consists of a set X together with a right action $\mu\colon X \times G \to X$ of G on X. This action is usually denoted simply by a dot, as in $\mu(x, g) = x \cdot g$; one requires μ to satisfy the identities $x \cdot 1 = x$ and $(x \cdot g) \cdot h = x \cdot (gh)$, for all $x \in X$ and g, $h \in G$. A morphism between two such representations (X, μ) and (Y, ν)

is a function $f\colon X \to Y$ which respects the action [in the sense that $f(x \cdot g) = f(x) \cdot g$ for all $x \in X$ and $g \in G$].

(v) $\mathbf{B}M$, or M-**Sets**, the category of all representations $X \times M \to X$ of a fixed monoid M on a variable set X; as in (iv), a morphism of $\mathbf{B}M$ is a function which respects the action.

(vi) $\mathbf{Sets}^{\mathbf{2}}$, the category whose objects are all functions $\sigma\colon X \to X'$ from one set X to a second set X', with the evident arrows (commutative squares) between these objects.

(vii) $\mathbf{Sets}^{\mathbf{N}}$, the category whose objects are all sequences X,

$$X_0 \to X_1 \to X_2 \to \cdots$$

of sets X_n and functions $X_n \to X_{n+1}$, with the evident arrows $X \to Y$. It has been suggested (Lawvere) that such a sequence X be considered as a "set through time", where each X_n is regarded as the state of the variable set X at the (discrete) time n. The exponent $\mathbf{N}$ here is the linearly ordered set of natural numbers, to be regarded as a category, so that a sequence X_n is a functor $\mathbf{N} \to \mathbf{Sets}$.

(viii) $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$, where $\mathbf{C}$ is a fixed small category and $\mathbf{C}^{\mathrm{op}}$ its opposite. This is the usual functor category, with objects all functors $P : \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$ and arrows $P \to P'$ all natural transformations $\theta\colon P \to P'$ between such functors. Recall that such a θ assigns to each object C of $\mathbf{C}$ a function $\theta_C\colon P(C) \to P'(C)$, in such a way that all diagrams

$$\begin{array}{ccc} PC & \xrightarrow{Pf} & PD \\ {\scriptstyle\theta_C}\downarrow & & \downarrow{\scriptstyle\theta_D} \\ P'C & \xrightarrow[P'f]{} & P'D, \end{array}$$

for $f\colon D \to C$ an arrow in $\mathbf{C}$, are commutative. Each object P in this category is a contravariant set-valued functor on $\mathbf{C}$; in anticipation of Example (x) below, such a P is also called a *presheaf* on $\mathbf{C}$. In the notation of the French school,

$$\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}} = \widehat{\mathbf{C}} \tag{1}$$

is the category of all presheaves on $\mathbf{C}$. If P is a presheaf on $\mathbf{C}$ and $x \in P(C)$, the value $P(f)(x)$ for an arrow $f\colon D \to C$ in $\mathbf{C}$ is called the *restriction* of x *along* f, and is often denoted by $|$ or by a dot:

$$P(f)(x) = x|f = x \cdot f. \tag{2}$$

Here f is written after x, because the contravariant character of P is then expressed for a composite $f \circ g$ as $x \cdot (f \circ g) = (x \cdot f) \cdot g$.

Each object C of $\mathbf{C}$ gives rise to a presheaf $\mathbf{y}(C)$ on $\mathbf{C}$, defined on an object D of $\mathbf{C}$ by

$$\mathbf{y}(C)(D) = \mathrm{Hom}_{\mathbf{C}}(D, C) \tag{3}$$

and on a morphism $D' \xrightarrow{\alpha} D$, for $u\colon D \to C$, by

$$\begin{aligned} \mathbf{y}(C)(\alpha)\colon \mathrm{Hom}_{\mathbf{C}}(D, C) &\to \mathrm{Hom}_{\mathbf{C}}(D', C) \\ \mathbf{y}(C)(\alpha)(u) &= u \circ \alpha; \end{aligned} \tag{4}$$

or briefly, $\mathbf{y}(C) = \mathrm{Hom}_{\mathbf{C}}(-, C)$ is the contravariant Hom-functor. Presheaves which, up to isomorphism, are of this form are called *representable presheaves* or *representable functors*. If $f\colon C_1 \to C_2$ is a morphism in $\mathbf{C}$, there is a natural transformation $\mathbf{y}(C_1) \to \mathbf{y}(C_2)$ obtained by composition with f. This makes $\mathbf{y}$ into a functor

$$\mathbf{y}\colon \mathbf{C} \to \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}, \qquad C \mapsto \mathrm{Hom}_{\mathbf{C}}(-, C) \tag{5}$$

from $\mathbf{C}$ to the *contravariant* functors on $\mathbf{C}$ (hence the exponent $\mathbf{C}^{\mathrm{op}}$). It is called the *Yoneda embedding*. The Yoneda embedding is a full and faithful functor. This fact is a special case of the so-called *Yoneda lemma*, which asserts for an arbitrary presheaf P on C that there is a bijective correspondence between natural transformations $\mathbf{y}(C) \to P$ and elements of the set $P(C)$:

$$\theta\colon \mathrm{Hom}_{\widehat{\mathbf{C}}}(\mathbf{y}(C), P) \xrightarrow{\sim} P(C), \tag{6}$$

defined for $\alpha\colon \mathbf{y}(C) \to P$ by $\theta(\alpha) = \alpha_C(1_C)$ (see [**CWM**, p. 61]).

(ix) $\mathbf{Sets}/J$, the *comma category* or *slice category*, with objects all sets over the fixed set J. Here, a *set over* J is by definition a function $h\colon X \to J$ from a (variable) set X to J, and with arrows commuting triangles as in (8) below. We also think of X (via h) as a set over the "base" J.

(x) $\mathbf{Sh}(X)$, the category of all sheaves of sets over a fixed topological space X. This important example will be explained in Chapter II below.

(xi) Let G be a topological group. The category $\mathbf{B}G$ of *continuous G-sets* has as objects sets X equipped with a right action $\mu\colon X \times G \to X$, as in (iv), with the additional requirement that this action be *continuous* when X is equipped with the discrete topology. The morphisms are the same as those described in (iv).

(xii) Simplicial sets: a *simplicial object* S in a category $\mathbf{C}$ is a family S_n for $n \geq 0$ of objects of $\mathbf{C}$, together with for each n two families of morphisms of $\mathbf{C}$

$$d_i\colon S_n \to S_{n-1}, \qquad s_i\colon S_n \to S_{n+1}, \qquad i = 0, \ldots, n$$

(and with $n > 0$ in the case of d_i) which satisfy the identities

$$\begin{aligned} d_i d_j &= d_{j-1} d_i, && i < j, \\ s_i s_j &= s_{j+1} s_i, && i \leq j, \\ d_i s_j &= s_{j-1} d_i, && i < j, \\ &= 1, && i = j \text{ and } i = j+1, \\ &= s_j d_{i-1}, && i > j+1. \end{aligned} \tag{7}$$

In particular, a simplicial object in **Sets** is called a *simplicial set.* If Δ_n is a "standard" affine n-simplex, a continuous map $f\colon \Delta_n \to X$ into a topological space X is called a *singular simplex* for X. Such a simplex f has $n+1$ faces $d_i f\colon \Delta_{n-1} \to X$ determined by restricting f to the i^{th} face of Δ_n; also collapsing vertex i to vertex $i+1$ gives $n+1$ maps $\Delta_{n+1} \to \Delta_n$: composed with f they yield $n+1$ "degenerate" singular simplices $s_i f$. Taking d_i to be the i^{th} face and s_i to be the i^{th} degeneracy makes the collection of all such singular simplices of X into a simplicial set—one from which the homology, cohomology, and homotopy of X can be computed. For many purposes, the category of topological spaces may be replaced by the category of simplicial sets.

(xiii) **FinSets**, the category of all finite sets and functions between them.

(xiv) $\mathbf{FinSets}^{\mathbf{C}^{\text{op}}}$, the category of all functors from $\mathbf{C}$ (a fixed finite category) to **FinSets**.

In this list the decisive types of examples are (i), (viii), and (x): **Sets**, set-valued functors (presheaves), and sheaves. These correspond to the major thrusts of our subject, toward the foundation of sets, the manipulation of functor categories, and the properties of sheaf cohomology. As a matter of fact, each of the categories (i)–(vii) above is a special case of a functor category $\mathbf{Sets}^{\mathbf{C}^{\text{op}}}$—and in each case, the arrows of the category in question are precisely the natural transformations of functors—for the following choices of the category:

(i) $\mathbf{C}$ is the category $\mathbf{1}$ with one object and one (identity) arrow.

(ii) $\mathbf{C} = \mathbf{1}+\mathbf{1}$ is the discrete category with two objects (and therefore with exactly two arrows, the identity arrows of these objects).

(iii) $\mathbf{C}$ is the discrete category with n objects.

(iv) $\mathbf{C}$ is the group G, regarded as a category with one object, with arrows the elements of G, and with composition the product in the group G.

(v) $\mathbf{C}$ is the monoid M, regarded in the same way as a one-object category.

(vi) $\mathbf{C} = \mathbf{2}$ is the "arrow category": The category with exactly two objects 0 and 1 and one nonidentity arrow $0 \to 1$.

(vii) $\mathbf{C} = \mathbf{N}^{\mathrm{op}}$ is the category whose objects are the natural numbers $n = 0, 1, 2, \ldots$, and whose arrows $n \to m$ are exactly the pairs $\langle n, m \rangle$ with $n \geq m$.

Extending (vi) and (vii), recall that any ordered or preordered set P will yield a category P with objects the elements $p \in P$ and arrows the pairs $\langle p, q \rangle$ with $p \leq q$ in the given preorder. For example, the ordered set $\mathbf{R}$ of real numbers yields in this way the functor category $\mathbf{Sets}^{\mathbf{R}}$ whose objects are "sets seen through (real) time" [cf. (vii) above].

The comma category described in (ix) is almost a functor category. An object of this comma category is an arrow $h\colon X \to J$ of **Sets**, while an arrow $f\colon h \to h'$ between two such objects is an arrow $f\colon X \to X'$ of **Sets** such that the triangular diagram

$$\begin{array}{ccc} X & \xrightarrow{\ f\ } & X' \\ & {\scriptstyle h}\searrow \quad \swarrow{\scriptstyle h'} & \\ & J & \end{array} \tag{8}$$

commutes. Each such object $h\colon X \to J$ over J determines a J-indexed family $\{ H_j \mid j \in J \}$ of sets, consisting of the sets

$$H_j = h^{-1}\{j\} = \{ x \mid x \in X \text{ and } hx = j \},$$

and then each arrow $f\colon h \to h'$ as in (8) determines a J-indexed family of functions $f_j\colon H_j \to H'_j$, $j \in J$. If we regard the set J here as a discrete category (with objects all elements $j \in J$ and arrows only the identity arrows $j \to j$), then each J-indexed family of sets is just a functor $H\colon J \to \mathbf{Sets}$ and each J-indexed family of functions f_j is just a natural transformation $F\colon H \to H'$ between these functors. In other words, the assignments $h \mapsto \{H_j\}$, $f \mapsto \{f_j\}$ constitute a functor

$$L\colon \mathbf{Sets}/J \to \mathbf{Sets}^J, \qquad h \mapsto \{H_j\}$$

from the comma category to the functor category.

Reciprocally, each functor $H\colon J \to \mathbf{Sets}$ determines a set $h\colon X \to J$ over J, with X the disjoint union (the coproduct) $X = \coprod H_j$ of the sets H_j for $j \in J$ and h the function which sends each $x \in X$ into its "index" (that $j \in J$ with $x \in H_j$). These two reciprocal constructions amount to constructing two functors L and M

$$\mathbf{Sets}/J \underset{M}{\overset{L}{\rightleftarrows}} \mathbf{Sets}^J \tag{9}$$

with both LM and ML naturally isomorphic to the respective identity functors. Therefore, these two constructions provide an equivalence of

the comma category $\mathbf{Sets}/J$ to the functor category $\mathbf{Sets}^J$. This equivalence is *not* an isomorphism of categories because the composite functor ML is not the identity, on account of the choice available in the formation of the disjoint union involved in the construction of M. (That is, many different sets over J correspond under L to the same functor on J.)

The category of simplicial sets is also a functor category:

$$\mathbf{SimpSets} = \mathbf{Sets}^{\Delta^{\mathrm{op}}}, \tag{10}$$

where Δ is the category whose objects are all the finite ordered sets $[n] = \{0, 1, \ldots, n\}$ and whose morphisms $[n] \to [m]$ are those maps $\phi\colon [n] \to [m]$ which preserve the order (i.e., $i \leq j$ implies $\phi i \leq \phi j$). The isomorphism (10)

comes about because the object $[n]$ in Δ can be regarded as the ordered set of vertices $0, 1, \ldots, n$ of the standard n-simplex Δ_n with i^{th} face spanned by $0, 1, \ldots, \widehat{i}, \ldots, n$, omitting i. Details may be found in many sources, for example in [**Mac Lane**, Homology 1963, p. 233; see also VIII §7].

2. Pullbacks

We will make extensive use of pullbacks. Recall that a *pullback* for a diagram $X \xrightarrow{f} B \xleftarrow{g} Y$ in a category $\mathbf{C}$ is a commutative square, with vertex P, on the edges f and g, as below, which is universal among such squares: To any other such commutative square, with vertex Q, on these edges, as in (1),

$$\begin{array}{ccc} P & \xrightarrow{f'} & Y \\ {\scriptstyle g'}\downarrow & & \downarrow{\scriptstyle g} \\ X & \xrightarrow[f]{} & B, \end{array} \qquad \begin{array}{ccc} Q & \xrightarrow{f_0} & Y \\ {\scriptstyle g_0}\downarrow & & \downarrow{\scriptstyle g} \\ X & \xrightarrow[f]{} & B \end{array} \tag{1}$$

there is a unique arrow $h\colon Q \to P$ with $f_0 = f'h$ and $g_0 = g'h$. As usual, this universality characterizes the pullback square (when it exists) up to isomorphism; its vertex P is called the *fibered product*, $P = X \times_B Y$ of X and Y (relative to f and g).

In **Sets**, the pullback P in (1) always exists and is (isomorphic to) the set of all those ordered pairs $\langle x, y \rangle$ of elements $x \in X$, $y \in Y$ with $fx = gy$ in B. In particular, if Y is a subset of B and $g\colon Y \to B$ the inclusion, the pullback P is (isomorphic to) the inverse image f^{-1} of Y in X. If both f and g are inclusions of subsets of B, their pullback P "is" the intersection of these subsets. If the set $g\colon Y \to B$ over B is regarded as a B-indexed set $\{G_b\}$, its pullback P along f is the X-indexed set

$\{ G'_x \}$ with $G'_x = G_{fx}$; in pictures: over each point $x \in X$ put a copy G'_x of the set given over fx. Similarly, if g is a fiber bundle of some sort, g' is (in a suitable category) the familiar induced fiber bundle.

The pullback exists in any functor category $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ and is constructed "pointwise" (as pointed out in the preliminaries): If X, Y, and B in (1) are functors to **Sets**, with f and g natural transformations, then $P\colon \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$ is that functor which sends each $C \in \mathbf{C}$ to the set PC which is the pullback in **Sets** of $XC \to BC$ and $YC \to BC$. In other words, $(X \times_B Y)(C) \cong X(C) \times_{B(C)} Y(C)$.

Consider the pullback P of f with itself [$f = g$ in (1)]. In **Sets**, P is the set of all pairs $\langle x, y \rangle$ of elements in X with $fx = fy$; in other words, $P \subset X \times X$ is the equivalence relation which f induces on its domain X. In any category, the pullback P of f with f, when it exists, is a parallel pair of arrows $P \rightrightarrows X$ called the *kernel pair* of f. In particular, an arrow f is *monic* (= left cancelable) precisely when, up to isomorphism, both arrows in its kernel pair are the identity $X \to X$. In particular, any functor preserving pullbacks preserves monics.

In any category, a pullback g' of a monic g along any arrow is itself monic; this may be proved by a simple formal argument. In **Sets**, it is also true that the pullback g' of an epi is always epi (epi = right cancelable arrow); this is evident from the description of the pullback in **Sets** by elements, but there is no simple formal (categorical) argument. This property does also hold in all Examples (i)–(xiv), but the common reason, as we shall see in Chapter IV, is deeper.

The one-point set $\{*\}$ may be characterized (up to isomorphism again!) as a *terminal* object in **Sets**: To every set X, there is a unique function $X \to \{*\}$. In the same way, each of our categories (ii)–(xiv) has a terminal object, call it 1; for example, the terminal object in $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ is the functor whose value at every object C is $\{*\}$.

Pullbacks and terminal objects are limits; specifically, a pullback in a category C is a limit of a functor $(\bullet \to \bullet \leftarrow \bullet) \to \mathbf{C}$, while a terminal object in $\mathbf{C}$ is a limit of a (the) functor from the empty category $\mathbf{0}$ into $\mathbf{C}$. Recall that a *finite limit* in $\mathbf{C}$ means a limit of a functor $J \to \mathbf{C}$ where J is a finite category. A category $\mathbf{C}$ with a terminal object 1 and with all pullbacks has all finite limits: It has binary products, since the product $X \times Y$ can be constructed as the pullback $X \to 1 \leftarrow Y$, it has products of no factors (the terminal 1), and hence it has all finite products. The equalizer e of a pair $f, g\colon X \rightrightarrows Y$ can also be constructed as a pullback, namely, that of the map $(f, g)\colon X \to Y \times Y$ and the

diagonal Δ:

$$\begin{array}{ccc} E & \xrightarrow{fe=ge} & Y \\ {\scriptstyle e}\downarrow & & \downarrow{\scriptstyle \Delta} \\ X & \xrightarrow[(f,g)]{} & Y \times Y. \end{array}$$

This shows that a category with pullbacks and a terminal object has all finite limits because these can be constructed from finite products and equalizers [**CWM**, p. 109].

All the categories in our list (i)–(xiv) have finite limits: We have seen how to construct the terminal object and pullbacks in **Sets**, so **Sets** has finite limits. Consequently, so does a functor category $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$, since limits in $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ can be computed pointwise, as just pointed out [cf. also (24) of the preliminaries]. In particular, this takes care of Examples (i)–(ix), and (xii), (xiii), and (xiv). Finite limits in the category of sheaves on a topological space X [Example (x) of §1] will be treated in Chapter II. This leaves the Example (xi) of the category $\mathbf{B}G$ of continuous G-sets, which is not a functor category. But if $a\colon X \to B$ and $b\colon Y \to B$ are maps of continuous G-sets, then their pullback $X \times_B Y$ in **Sets** has an obvious coordinatewise action by G:

$$(x, y) \cdot g = (x \cdot g, y \cdot g)$$

for $x \in X$ and $y \in Y$ with $a(x) = b(y)$, and $g \in G$. This action is continuous if the actions on X and Y are each continuous. So $X \times_B Y$ is again a continuous G-set, and this is easily seen to define the pullback in the category $\mathbf{B}G$. Since $\mathbf{B}G$ has a terminal object (the one-point set with its unique action by G), it follows that $\mathbf{B}G$ has all finite limits and that these limits can all be constructed as limits of the underlying sets. In other words, the "forgetful" functor $U\colon \mathbf{B}G \to \mathbf{Sets}$ which forgets the action, $U(X, \mu) = X$, preserves all finite limits.

3. Characteristic Functions of Subobjects

In **Sets**, a subset $S \subset X$ may be described in two very different ways: As the monic function $S \rightarrowtail X$ given by inclusion or as a characteristic function ϕ_S defined as usual for elements $x \in X$ by

$$\phi_S(x) = \begin{cases} 0, & x \in S, \\ 1, & x \notin S. \end{cases}$$

Here we take the values of ϕ_S in the typical 2-point set $\{0, 1\}$; it is the set of "truth values", where we have chosen 0 as the value "true". It is

convenient to regard true as the following subobject (the monic arrow) from 1 to 2:

$$\text{true}\colon 1 = \{0\} \rightarrowtail 2 = \{0, 1\}, \qquad 0 \mapsto 0. \tag{1}$$

With this notation, each subset S can evidently be recovered (up to equivalence) from its characteristic function ϕ_S as the pullback of true along ϕ_S:

$$\begin{array}{ccc} S & \longrightarrow & 1 \\ {\scriptstyle m}\downarrow & & \downarrow{\scriptstyle \text{true}} \\ X & \xrightarrow[\phi_S]{} & 2. \end{array} \tag{2}$$

In this diagram, $S \to 1$ is the unique function from S to the terminal object (one-point set) 1, and $1 \to 2$ is the fixed monic defined in (1); given the monic m, there is a unique ϕ (namely, the characteristic function) such that the diagram (2) is a pullback.

In Section 4, we will see that subobjects in our other typical categories have similar characteristic functions, which take values not in 2, but in a suitable object Ω of "truth-values".

Definition. *In a category* $\mathbf{C}$ *with finite limits, a subobject classifier is a monic,* $\text{true}\colon 1 \to \Omega$, *such that to every monic* $S \rightarrowtail X$ *in* $\mathbf{C}$ *there is a unique arrow* ϕ *which, with the given monic, forms a pullback square*

$$\begin{array}{ccc} S & \longrightarrow & 1 \\ \downarrow & & \downarrow{\scriptstyle \text{true}} \\ X & \xrightarrow[\phi]{} & \Omega. \end{array} \tag{3}$$

In other words, every subobject is uniquely a pullback of a "universal" monic true.

This property amounts to saying that the subobject functor is representable (i.e., isomorphic to a Hom-functor). In detail, a subobject of an object X in any category $\mathbf{C}$ is an equivalence class of monics $m\colon S \rightarrowtail X$ to X (cf. the preliminaries). By a familiar abuse of language, we say that the subobject *is* S or *is* m, meaning always the equivalence class of m. Then, $\text{Sub}_{\mathbf{C}} X$ is the set of all subobjects of X in the category $\mathbf{C}$; this set is partially ordered under inclusion. The category $\mathbf{C}$ is said to be *well-powered* when $\text{Sub}_{\mathbf{C}} X$ is isomorphic to a small set for all X; all of our typical categories are well-powered. Now given an arrow $f\colon Y \to X$ in $\mathbf{C}$, the pullback of any monic $m\colon S \rightarrowtail X$ along f is a monic $m'\colon S' \rightarrowtail Y$, and the assignment $m \mapsto m'$ defines a function $\text{Sub}_{\mathbf{C}} f\colon \text{Sub}_{\mathbf{C}} X \to \text{Sub}_{\mathbf{C}} Y$; when $\mathbf{C}$ is well-powered, this makes $\text{Sub}_{\mathbf{C}}\colon \mathbf{C}^{\text{op}} \to \mathbf{Sets}$ a functor to $\mathbf{Sets}$. Briefly, Sub is a functor "by pullback".

Proposition 1. *A category* $\mathbf{C}$ *with finite limits and small Hom-sets has a subobject classifier if and only if there is an object* Ω *and an isomorphism*

$$\theta_X\colon \mathrm{Sub}_{\mathbf{C}}(X) \cong \mathrm{Hom}_{\mathbf{C}}(X, \Omega), \tag{4}$$

natural for $X \in \mathbf{C}$. *When this holds,* $\mathbf{C}$ *is well-powered.*

Proof: Given a subobject classifier as in (3), the correspondence θ_X sending the equivalence class of each monic $S \rightarrowtail X$ to its (unique) "characteristic function" $\phi\colon X \to \Omega$ is a bijection for each X, as required for (4). Now $\mathrm{Sub}_{\mathbf{C}}(X)$ is a (contravariant) functor of X by pullback (= inverse image); so to prove this bijection natural, we must show that pullback along $f\colon Y \to X$ in $\mathrm{Sub}_{\mathbf{C}}(-)$ corresponds to composition with f in $\mathrm{Hom}_{\mathbf{C}}(-, \Omega)$. This is immediate by the elementary fact that two pullback squares placed side by side, as in

$$\begin{array}{ccccc} S' & \longrightarrow & S & \longrightarrow & 1 \\ \downarrow & & \downarrow & & \downarrow{\scriptstyle \mathrm{true}} \\ Y & \xrightarrow[f]{} & X & \longrightarrow & \Omega, \end{array}$$

yield a pullback (rectangle). Since the Hom-sets are all small, the bijection (4) also proves $\mathbf{C}$ well-powered.

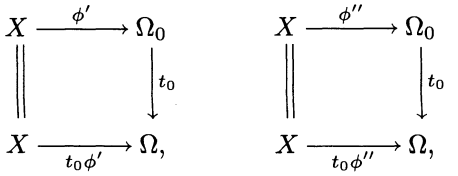
Conversely, suppose that (4) is a bijection, natural in X. This states that $\mathrm{Sub}_{\mathbf{C}} : \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$ is naturally isomorphic to $\mathrm{Hom}_{\mathbf{C}}(-, \Omega)$; that is, that the functor $\mathrm{Sub}_{\mathbf{C}}$ is *representable* [cf. §1(5)], with representing-object Ω. As for any such representation, some subobject $t_0\colon \Omega_0 \rightarrowtail \Omega$ of Ω corresponds to the identity $1\colon \Omega \rightarrowtail \Omega$, while each subobject $S \rightarrowtail X$ of X corresponds to an arrow $\phi : X \to \Omega$. By naturality of θ, the diagram

$$\begin{array}{ccc} \mathrm{Sub}(\Omega) & \xrightarrow{\sim} & \mathrm{Hom}(\Omega, \Omega) \\ {\scriptstyle \mathrm{Sub}(\phi)}\downarrow & & \downarrow{\scriptstyle \mathrm{Hom}(\phi, \Omega)} \\ \mathrm{Sub}(X) & \xrightarrow{\sim} & \mathrm{Hom}(X, \Omega), \end{array} \qquad \begin{array}{ccc} \Omega_0 & \longmapsto & 1 \\ \downarrow & & \downarrow \\ S & \longmapsto & \phi \end{array}$$

must commute; this states that $S = \mathrm{Sub}(\phi)\Omega_0$, and hence that each subobject S is the pullback of Ω_0 along a unique "characteristic function" ϕ, as in

$$\begin{array}{ccc} S & \xrightarrow{\phi'} & \Omega_0 \\ \downarrow & & \downarrow{\scriptstyle t_0} \\ X & \xrightarrow[\phi]{} & \Omega. \end{array} \tag{5}$$

This diagram is like the definition (3) of a subobject classifier, but it remains to show that Ω_0 is actually a terminal object in $\mathbf{C}$. But taking $S \to X$ in (5) to be the identity $X \to X$ gives a map $\phi' \colon X \to \Omega_0$. If there were two maps ϕ', $\phi'' \colon X \to \Omega_0$ then, since t_0 is monic, both the squares



would trivially be pullbacks. Therefore, by the uniqueness of ϕ in (5), $t_0\phi' = t_0\phi''$. As t_0 is monic, this gives $\phi' = \phi''$. Hence each object X has a unique map $\phi' \colon X \to \Omega_0$, so Ω_0 is terminal.

As with any representation of a functor, this result proves that the subobject classifier of a category, if it exists, is unique up to an isomorphism.

The idea of a "subobject classifier" is modeled on other "classifying" ideas in topology. A decisive example is that of the classifying bundle for a Lie group G. As we will subsequently indicate, a G-bundle over X is a suitable continuous map $\pi \colon E \to X$ of spaces for which G acts on the right on E in such a way that $\pi e = \pi e'$ for e and e' in E iff $e = e'g$ for a unique element g in the group G. If G-**Bund**(X) is the (suitably defined) set of all G-bundles over some space X, then pullback of a G-bundle $E \to X$ along a continuous map $f \colon Y \to X$ yields a G-bundle over Y, and this makes G-**Bund** a contravariant functor of X (again "by pullback"). Then a bundle $V \to B$ is said to be a universal G-bundle (and B is the "classifying space" for G) if every G-bundle $E \to X$ can be obtained from $V \to B$ by pullback along some $X \to B$. For $G = \mathbf{O}_k$ the real orthogonal group in k variables, there is a famous such classifying bundle V. For large n, it is the Stiefel manifold $V_{n+k+1,k}$ whose points are all orthonormal k-frames of vectors $v_1, \ldots, v_k$ in $\mathbf{R}^{n+k+1}$. The orthogonal group $\mathbf{O}_k$ acts continuously on these frames $v_1, \ldots, v_k$ in the evident way, so that two frames are equivalent under the action of $\mathbf{O}_k$ if and only if they span the same k-plane. Therefore the projection $p \colon V \to V/\mathbf{O}_k$ of the Stiefel manifold on its quotient by this action is an $\mathbf{O}_k$-bundle, and its base space $V/\mathbf{O}_k$ is precisely the Grassmann manifold $M_{n+k+1,k}$ of all k-planes in $\mathbf{R}^{n+k+1}$. A standard argument [**Steenrod**, 1951] shows that p is a classifying bundle for principal $\mathbf{O}_k$-bundles over n-complexes K, in the sense that any such bundle can be obtained from p by pullback along a continuous map $f \colon K \to V/\mathbf{O}_k$ which is unique *up to homotopy.*

Classifying bundles have played a major role in topology; we shall see that classifying subobjects play a similarly decisive role in category theory. Later on in this book, we shall see that the analogous idea of a classifying topos is central in our subject.

4. Typical Subobject Classifiers

Each of our typical categories (i)–(xiv) has a subobject classifier. We will now explicitly construct these classifiers in order to exemplify the general notion.

The classifier true: $1 \to 2$ for **Sets** is evidently also a subobject classifier for **FinSets**; indeed, the usual characteristic functions are still effective for finite sets.

In **Sets** $\times$ **Sets**, an arrow is a pair of functions $f\colon Y \to X$, $f'\colon Y' \to X'$. The pair of subsets $(1 \subset 2, 1 \subset 2)$ is a subobject classifier, and the characteristic arrow of any subobject $(S \subset X, S' \subset X')$ is evidently just the pair of characteristic functions $(\phi_S\colon X \to 2, \phi_{S'}\colon X' \to 2)$ from the category **Sets**. Thus, there are, in 2×2, four "truth-values". The corresponding subobject classifier for $\mathbf{Sets}^n$ has 2^n truth-values; as we shall see, it is the Boolean algebra of all 2^n subsets of n.

In the category $\mathbf{B}G = G$-**Sets** of representations of G [Example (iv) of §1], an object is an action $X \times G \to X$ of the fixed group G on some set X, and a subobject is just a subset $S \subset X$ closed under this action (i.e., $s \cdot g \in S$ whenever $s \in S$ and $g \in G$). The complement of S in X is thus also invariant under this action, so we can still use the ordinary characteristic function $\phi_S : X \to 2$ of S, where the subobject classifier is the usual map true: $1 \to 2$, with G acting trivially on both sets 1 and 2. Exactly the same argument applies in the case where G is a topological group [Example (xi) of §1].

For $\mathbf{B}M$ [Example (v) of §1], an object is again a right action $X \times M \to X$ of the fixed monoid M on some set X, and a subobject is again just a subset $S \subset X$ closed under this action, but the previous characteristic function will not do because the complement of S need not be closed under this action. Instead, we may define a function ϕ_S sending each $x \in X$ to the set L of all those $\ell \in M$ with $x \cdot \ell \in S$. This set L is a "right ideal" of M (a subset of M mapped into itself by the right action of M on itself, via right multiplication). Therefore, take $\Omega = \Omega_M$ to be the set of all right ideals L of M with action $\Omega \times M \to \Omega$ defined by $L \cdot m = \{\, k \in M \mid m \cdot k \in L \,\}$ for $L \in \Omega$ and $m \in M$. Then, the function ϕ_S above is an arrow $\phi_S\colon X \to \Omega$; in particular, it determines the given S as the inverse image of the right ideal M. Therefore, the subobject classifier is the function $\mathrm{true}_M\colon 1 \rightarrowtail \Omega_M$ which sends the one point of the object 1 to the "maximal" right ideal $M \in \Omega_M$.

In case M is a group G the only right ideals are G and $\emptyset$, so this Ω_G reduces to the previous set 2 with trivial G-action. In case M is the additive monoid of natural numbers, the right ideals are the empty set and the sets of numbers larger than some fixed number n.

For the arrow category $\mathbf{2}$ and $\mathbf{Sets}^{\mathbf{2}}$, a subset $(S_0 \xrightarrow{\sigma} S_1) \rightarrowtail (X_0 \xrightarrow{\sigma} X_1)$ is a pair of subsets $S_0 \subset X_0$, $S_1 \subset X_1$ with $\sigma S_0 \subset S_1$. Relative to this subset S there are three sorts of elements x of X_0: Those x in S_0,

those $x \notin S_0$ with $\sigma x \in S_1$, and those x with $\sigma x \notin S_1$. Define $\phi_0 x = 0$, 1, or 2 accordingly. Then, ϕ_0 on S_0, with the usual characteristic function ϕ_1 of $S_1 \subset X_1$, is an arrow $\phi = (\phi_0, \phi_1)$ to the object Ω displayed below,

$$\begin{array}{lccc} X\colon & X_0 & \xrightarrow{\ \sigma\ } & X_1 \\ {\scriptstyle\phi}\downarrow & {\scriptstyle\phi_0}\downarrow & & \downarrow{\scriptstyle\phi_1} \\ \Omega\colon & \{0,1,2\} & \xrightarrow[\ \sigma\ ]{} & \{0,1\}, \end{array} \qquad \sigma 0 = 0,\ \sigma 1 = 0,\ \sigma 2 = 1,$$

in $\mathbf{Sets}^{\mathbf{2}}$, and $S_0 \to S_1$ is the inverse image of $(\{0\} \xrightarrow{1} \{0\}) = 1 \rightarrowtail \Omega$.

In brief, this characteristic function $\phi = \langle \phi_0, \phi_1 \rangle$ is that arrow which specifies whether "x is in S" is "true" always, only at 1, or never. One may say that ϕ gives the "time till truth".

For $\mathbf{Sets}^{\mathbf{N}}$, a subobject of X has the form of a sequence S of subsets

$$\begin{array}{lccccccc} S\colon & S_0 & \longrightarrow & S_1 & \longrightarrow & S_2 & \longrightarrow & S_3 \ldots \\ \downarrow & \downarrow & & \downarrow & & \downarrow & & \downarrow \\ X & X_0 & \xrightarrow[\ \sigma\ ]{} & X_1 & \xrightarrow[\ \sigma\ ]{} & X_2 & \xrightarrow[\ \sigma\ ]{} & X_3 \ldots \end{array}$$

with $\sigma S_k \subset S_{k+1}$; for example, if X_k is constant and each $\sigma = 1$, this S is a monotone increasing sequence of subsets. For any $x \in X_k$ we can then measure the "time till truth" (the time till inclusion in S) by the function ϕ_k on X_k defined as

$$\begin{aligned} \phi_k x &= \text{ the least } n \text{ with } \sigma^n x \in S_{k+n}, \text{ if such exists,} \\ &= \infty \text{ otherwise.} \end{aligned}$$

Then $\phi_k\colon X_k \to \mathbf{N} + \{\infty\}$, so the sequence of these maps ϕ_k is an arrow to the sequence of sets

$$\Omega\colon \mathbf{N} + \{\infty\} \xrightarrow{\ \tau\ } \mathbf{N} + \{\infty\} \xrightarrow{\ \tau\ } \mathbf{N} + \{\infty\} \longrightarrow \cdots \tag{1}$$

where each τ has $\tau(0) = 0$, $\tau(n+1) = n$ for $n \neq 0$ and $\tau(\infty) = \infty$. Then, $\Omega \in \mathbf{Sets}^{\mathbf{N}}$ has $1\colon \{0\} \to \{0\} \to \{0\} \to \cdots$ as subobject, and the given S is the pullback of 1 along ϕ. In brief, "time till truth" provides a subobject classifier Ω.

For an arbitrary small category $\mathbf{C}$, a *subfunctor* of $P\colon \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$ is defined to be another functor $Q\colon \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$ with each QC a subset of PC and each $Qf\colon QD \to QC$ a restriction of Pf, for all arrows $f\colon C \to D$ of $\mathbf{C}$. The inclusion $Q \to P$ is then a monic arrow in the functor category $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$, so that each subfunctor Q is a subobject. Conversely, all subobjects are given by subfunctors; if a natural

transformation $\theta\colon R \rightarrowtail P$ is monic in the functor category, then each function $\theta C\colon RC \to PC$ is an injection (monics, like limits in the functor category, are taken pointwise). For each C let QC be the image of $RC \rightarrowtail PC$; thus Q is manifestly a subfunctor of P, and the given R is equivalent (as a subobject) to Q.

For an arbitrary presheaf category $\widehat{\mathbf{C}} = \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$, if there is a subobject classifier Ω, it must, in particular, classify the subobjects of each representable presheaf $\mathbf{y}C = \mathrm{Hom}_{\mathbf{C}}(-, C)\colon \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$. Therefore,

$$\mathrm{Sub}_{\widehat{\mathbf{C}}}(\mathrm{Hom}_{\mathbf{C}}(-, C)) \cong \mathrm{Hom}_{\widehat{\mathbf{C}}}(\mathrm{Hom}_{\mathbf{C}}(-, C), \Omega) = \mathrm{Nat}(\mathrm{Hom}_{\mathbf{C}}(-, C), \Omega).$$

By the Yoneda lemma [see §1(6) above], the set on the right is (up to isomorphism) $\Omega(C)$. Thus the subobject classifier Ω, if it exists, must be the functor $\Omega\colon \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$ with object function

$$\begin{aligned} \Omega(C) &= \mathrm{Sub}_{\widehat{\mathbf{C}}}(\mathrm{Hom}_{\mathbf{C}}(-, C)) \\ &= \{\, S \mid S \text{ a subfunctor of } \mathrm{Hom}_{\mathbf{C}}(-, C) \,\}, \end{aligned} \tag{2}$$

and with a suitable mapping function.

To understand this it is customary and useful to use an alternative terminology for subfunctors of a representable functor $\mathrm{Hom}(-, C)$. Given an object C in the category $\mathbf{C}$, a *sieve* on C (in French, a "crible" on C) is a set S of arrows with codomain C such that

$$f \in S \text{ and the composite } fh \text{ is defined implies } fh \in S.$$

If we think of the arrows $f \in S$ as those paths which are "allowed to get through" to C, this definition means that any path to some other B followed by an allowed path from B to C is allowed. For example, if the category $\mathbf{C}$ is a monoid M, a sieve is just a right ideal in M; if the category $\mathbf{C}$ is a partially ordered set regarded as a category, a sieve on $C \in \mathbf{C}$ is a set S of elements $B \leq C$ such that $A \leq B \in S$ implies $A \in S$: If B "goes through" the sieve, so does anything smaller: a sieve is a "downwards closed" subset.

Now if $Q \subset \mathrm{Hom}_{\mathbf{C}}(-, C)$ is a subfunctor, the set

$$S = \{\, f \mid \text{ for some object } A,\ f\colon A \to C \text{ and } f \in Q(A) \,\}$$

is clearly a sieve on C. Conversely, given a sieve S on C, the definition

$$Q(A) = \{\, f \mid f\colon A \to C \text{ and } f \in S \,\} \subseteq \mathrm{Hom}_{\mathbf{C}}(A, C)$$

yields a functor $Q\colon \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$ which is a subfunctor of the Hom-functor $\mathrm{Hom}_{\mathbf{C}}(-, C)$. The passages S to Q and Q to S are reciprocal;

hence, we can identify sieves and subfunctors in any locally small category $\mathbf{C}$. Thus,

$$\text{Sieve on } C = \text{Subfunctor of } \mathrm{Hom}_{\mathbf{C}}(-, C). \tag{3}$$

Moreover, for any arrow $g\colon B \to C$, a subobject Q of the functor $\mathrm{Hom}_{\mathbf{C}}(-, C)$ determines a subobject of $\mathrm{Hom}_{\mathbf{C}}(-, B)$ by pullback along g, and similarly each sieve S on C determines the following sieve on B:

$$S \cdot g = \{\, h \mid g \circ h \in S \,\}.$$

With this motivation, the proposed subobject classifier Ω for the functor category $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ is defined on objects by

$$\Omega(C) = \{\, S \mid S \text{ is a sieve on } C \text{ in } \mathbf{C} \,\} \tag{4}$$

and on arrows $g\colon C' \to C$ by

$$(-) \cdot g\colon \Omega(C) \to \Omega(C'), \qquad S \cdot g = \{\, h \mid g \circ h \in S \,\}. \tag{5}$$

For an object C of $\mathbf{C}$, the set $t(C)$ of *all* arrows into C is a sieve, called the *maximal sieve* on C. These maximal sieves patch together to give a morphism (natural transformation)

$$\mathrm{true}\colon 1 \to \Omega \tag{6}$$

in the presheaf category $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$.

To see that (6) defines a subobject classifier in $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$, consider any subfunctor Q of a given functor $P\colon \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$. Then each morphism $f\colon A \to C$ in $\mathbf{C}$ determines a function $P(f)\colon P(C) \to P(A)$ in $\mathbf{Sets}$ which may or may not take a given $x \in P(C)$ into $Q(A) \subseteq P(A)$. For a given $x \in P(C)$ set

$$\phi_C(x) = \{\, f \mid x \cdot f \in Q(\mathrm{dom}(f)) \,\}, \tag{7}$$

where f ranges over all morphisms in $\mathbf{C}$ with codomain C. Then $\phi_C(x)$ is a sieve on C, and $\phi\colon P \to \Omega$ is natural. Moreover, $\phi_C(x)$ is the maximal sieve $t(C)$ iff $x \in Q(C)$, so the given subfunctor $Q \subseteq P$ is the pullback along ϕ of the map "true" defined in (6) above.

$$\begin{array}{ccc} Q & \longrightarrow & 1 \\ \downarrow & & \downarrow{\scriptstyle \mathrm{true}} \\ P & \xrightarrow[\phi]{} & \Omega. \end{array} \tag{8}$$

This shows that ϕ is indeed a possible characteristic map for the subfunctor Q. But this ϕ is also the unique natural transformation $\theta\colon P \to \Omega$ making this diagram into a pullback. Indeed, given $x \in P(C)$ and $f\colon A \to C$, the pullback condition means that $x \cdot f \in Q(A)$ iff $\theta_A(x \cdot f) = \text{true}_A$; by naturality of θ, this is equivalent to $\theta_C(x) \cdot f = \text{true}_A$ and this, in turn, by the definition (5), means that $f \in \theta_C(x)$. The elements f of $\theta_C(x)$ are thus exactly those f with $x \cdot f \in Q(A)$, i.e., those $f \in \phi_C(x)$ as defined in (7). Thus, the definition (7) of ϕ is forced upon us if (8) is to be a pullback. Hence, we have shown that the mono $\text{true}\colon 1 \rightarrowtail \Omega$ defined in (6) provides a subobject classifier for the presheaf category $\widehat{\mathbf{C}} = \mathbf{Sets}^{\mathbf{C}^{\text{op}}}$.

Intuitively, the sieve $\phi_C(x)$ considered in (7) is the set of all those paths f to C which translate the element x of $P(C)$ into the subfunctor Q. As the set of "paths to truth", it clearly agrees with the characteristic arrows we have already constructed for the special functor categories $\mathbf{Sets}^{\mathbf{2}}$, $\mathbf{Sets}^{M}$, and $\mathbf{Sets}^{\mathbf{N}}$.

We have assumed $\mathbf{C}$ small because we must. Were $\mathbf{C}$ large—say the ordered set of all small ordinal numbers—the number of paths to truth would not in general be small, hence not an object of **Sets**.

The exhibition of subobject classifiers for our typical categories is completed by noting, for any set J, that the projection $J \times 2 \to J$ is a classifier for the category $\mathbf{Sets}/J$, while in $\mathbf{FinSets}^{\mathbf{C}^{\text{op}}}$ with $\mathbf{C}$ finite, the set $\Omega(C)$ of sieves on C is again finite so provides a suitable subobject classifier Ω.

Observe, however, that there are many "reasonable" categories with no such subobject classifier. The category $(\mathbf{FinSets})^{\mathbf{N}}$ provides an immediate such example, because in the linearly ordered set $\mathbf{N}^{\text{op}}$, the number of sieves on each object n is infinite. Another example is the category **Ab** of all (small) abelian groups. For, the terminal object 1 in **Ab** is the zero-group, so the group homomorphism $\text{true}\colon 1 \rightarrowtail \Omega$ must send 0 to $0 \in \Omega$, and thus its pullback along any $\phi\colon A \to \Omega$ is the subgroup $S = \text{Ker}\,\phi = \phi^{-1}(0)$ of A. This implies that the proposed subobject classifier must be an abelian group which contains a copy of every quotient group A/S of every group A, an absurdity.

5. Colimits

Each of our typical categories has all finite colimits. To show this it suffices (as in the dual case of finite limits discusssed in §2) to observe that each has an initial object 0 and pushouts (or cocartesian squares, as they are sometimes called). In **Sets**, the empty set $\emptyset$ is an initial object because there is for each set X exactly one function $\emptyset \to X$; in a functor category $\mathbf{Sets}^{\mathbf{C}^{\text{op}}}$ the constantly empty functor is initial. And for a topological group G, the empty set (with its unique action by G)

is clearly initial. So all the examples of §1 have initial objects [we leave to the next chapter the discussion of sheaves, listed as (x) in §1].

Next, we consider pushouts. In **Sets** or in **FinSets** the pushout of two functions f and g with a common domain X, as in the diagram

$$\begin{array}{ccc} X & \xrightarrow{\ g\ } & Z \\ {\scriptstyle f}\downarrow & & \vdots \\ Y & \dashrightarrow & Q \end{array} \qquad Q = Y \coprod Z / \{ f(x) = g(x) \}, \tag{1}$$

is the set Q which is obtained from the disjoint union of Y and Z by identifying the elements $f(x)$ and $g(x)$, for all $x \in X$. This quotient set has the usual universal property of a pushout. By this universal property, if in (1) X, Y, and Z are sets over some fixed set J (as in Example (ix) of §1), so is their pushout; therefore, this same construction yields pushouts in the comma category $\mathbf{Sets}/J$. And similarly, if f and g in Diagram (1) are maps of G-sets for a group or monoid G [Examples (iv) and (v) of §1], or maps of continuous G-sets for a topological group G [Example (xi)], then the quotient Q can again be equipped with an action by G making Q into the pushout in the category of G-sets or continuous G-sets. Furthermore, if f and g in (1) are natural transformations of functors $X, Y, Z\colon \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$, then the pointwise pushouts $Q(C)$ for each object C of $\mathbf{C}$ form (again) by universality a functor $Q\colon \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$, which is the pushout in the functor category $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$. This is a special case of the fact that colimits in functor categories can be computed pointwise: If $H\colon J \to \mathbf{A}^{\mathbf{C}}$ for categories J, $\mathbf{A}$, and $\mathbf{C}$, then its colimit $\varinjlim_J H$ in the functor category $\mathbf{A}^{\mathbf{C}}$ is given by

$$(\varinjlim_J H)(C) = \varinjlim_J H(C) \qquad (C \in \mathbf{C}), \tag{2}$$

where $H(C)\colon J \to \mathbf{A}$ is the functor obtained from H by evaluating at the object C of $\mathbf{C}$ [cf. (25) of the preliminaries]. This applies also when J is finite, as in our Example (xiv).

This shows that our typical categories have finite colimits. [In fact, except for (xiii) and (xiv), they all have arbitrary small colimits, and in this sense are cocomplete, but that need not concern us here.] Our typical categories have many other common formal properties. For example, each morphism f has an epi-mono factorization $f = m \cdot e$. However, these other common properties will all be deduced (in Chapter IV) from the ones we have previously examined. In the deduction, limits will play a much more important role than colimits.

To conclude this section, we wish to mention a useful fact concerning colimits in functor categories:

Proposition 1. *In a functor category* $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$, *any object* P *is the colimit of a diagram of representable objects, in a canonical way.*

This proposition asserts, in other words, that given a functor $P\colon \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$, there is a canonical way of constructing a small "index" category J and a diagram $A\colon J \to \mathbf{C}$ in $\mathbf{C}$ of type J, such that P is isomorphic to the colimit $\varinjlim_J (\mathbf{y} \circ A)$ of the diagram $J \xrightarrow{A} \mathbf{C} \xrightarrow{\mathbf{y}} \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$, obtained by composition with the Yoneda embedding described in (1.5).

Given P, the index category J which serves to prove the proposition is the so-called *category of elements* of P, denoted by $\int_{\mathbf{C}} P$ or, more briefly, $\int P$. Its objects are all pairs (C, p) where C is an object of $\mathbf{C}$ and p is an element $p \in P(C)$. Its morphisms $(C', p') \to (C, p)$ are those morphisms $u\colon C' \to C$ of $\mathbf{C}$ for which $pu = p'$; in other words, u must take the chosen element p in $P(C)$ "back" into p' in $P(C')$:

$$(C', p') \to (C, p) \qquad \text{by } u\colon C' \to C \text{ with } pu = p'. \tag{3}$$

These morphisms are composed by composing the underlying arrows u of $\mathbf{C}$. This category has an evident projection functor

$$\pi_P\colon \int_{\mathbf{C}} P \to \mathbf{C}, \qquad (C, p) \mapsto C. \tag{4}$$

Colimits over the category of elements can be used to construct a pair of adjoint functors which will have many uses, as follows.

Theorem 2. *If* $A\colon \mathbf{C} \to \mathcal{E}$ *is a functor from a small category* $\mathbf{C}$ *to a cocomplete category* $\mathcal{E}$, *the functor* R *from* $\mathcal{E}$ *to presheaves given by*

$$R(E)\colon C \mapsto \mathrm{Hom}_{\mathcal{E}}(A(C), E) \tag{5}$$

has a left adjoint $L\colon \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}} \to \mathcal{E}$ *defined for each presheaf* P *in* $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ *as the colimit*

$$L(P) = \mathrm{Colim}\Big(\int P \xrightarrow{\pi_P} \mathbf{C} \xrightarrow{A} \mathcal{E}\Big). \tag{6}$$

In other words, there is a pair of adjoint functors $L \dashv R$, as in

$$L\colon \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}} \rightleftarrows \mathcal{E} : R, \tag{7}$$

where we place the left adjoint L on the left.

Proof: A natural transformation $\tau\colon P \to R(E)$ is just a family $\{\tau_C\}$ indexed by objects C of $\mathbf{C}$ for which each τ_C is a map

$$\tau_C\colon P(C) \to \mathrm{Hom}_{\mathcal{E}}(A(C), E)$$

of sets which is natural in C, in the sense that the diagram of sets

$$\begin{array}{ccc} P(C) & \xrightarrow{\tau_C} & \mathrm{Hom}_{\mathcal{E}}(A(C), E) \\ {\scriptstyle P(u)}\downarrow & & \downarrow{\scriptstyle A(u)^*} \\ P(C') & \xrightarrow[\tau_{C'}]{} & \mathrm{Hom}_{\mathcal{E}}(A(C'), E) \end{array} \tag{8}$$

commutes for each morphism $u\colon C' \to C$ of $\mathbf{C}$. But such a τ may also be considered as a family of arrows of $\mathcal{E}$

$$\{\, \tau_C(p)\colon A(C) \to E \,\}_{(C,p)} \tag{9}$$

indexed by objects (C, p) of the category $\int P$ of elements of P. In this view, the condition (8) then means that the following diagram

$$\begin{array}{ccccc} A(C) & = & A\pi_P(C,p) & & \\ {\scriptstyle A(u)}\uparrow & & \uparrow{\scriptstyle u_*} & \searrow^{\tau_C(p)} & \\ & & & & E \\ A(C') & = & A\pi_P(C',p') & \nearrow_{\tau'_C(p')} & \end{array} \tag{10}$$

commutes for each arrow u. This visibly means that the arrows $\tau_C(p)$ constitute a cocone from the functor $A\pi_P$ to the object E. By the definition of a colimit, each such cocone comes by composing the colimiting cocone (to $\varinjlim$) with a unique arrow from the colimit LP to the object E. In other words, there is a bijection

$$\mathrm{Nat}(P, R(E)) \cong \mathrm{Hom}_{\mathcal{E}}(LP, E). \tag{11}$$

Since this bijection is clearly natural in P and in E, it asserts that L is a left adjoint to R, just as claimed.

Corollary 3 (= Proposition 1). *Every presheaf is a colimit of representable presheaves.*

Proof: In the theorem, take $\mathcal{E}$ to be the presheaf category and A to be the Yoneda embedding

$$A = \mathbf{y}\colon \mathbf{C} \to \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}} = \widehat{\mathbf{C}} = \mathcal{E}.$$

By the Yoneda lemma the definition (5) of the right adjoint R for any $E = P$ is then

$$R_A(E)(C) = \mathrm{Hom}_{\widehat{\mathbf{C}}}(\mathbf{y}(C), E) \cong E(C);$$

this states that R_A is isomorphic to the identity functor of $\widehat{\mathbf{C}}$. By the uniqueness up to isomorphism of adjoints, its left adjoint L must then also be isomorphic to the identity functor, so that the definition (6) of L gives for any presheaf P

$$P \cong \mathrm{Colim}(\int P \xrightarrow{\pi_P} \mathbf{C} \xrightarrow{\mathbf{y}} \widehat{\mathbf{C}}). \tag{12}$$

One may prove the result directly, by constructing a colimiting cone from $L(P)$ in (6), with $A = \mathbf{y}$, to P.

The Yoneda embedding of $\mathbf{C}$ in the presheaf category $\widehat{\mathbf{C}}$ is the "universal" way of making $\mathbf{C}$ cocomplete, in the sense that every functor A from $\mathbf{C}$ to a cocomplete category $\mathcal{E}$ factors through $\mathbf{y}$ by a unique colimit preserving functor from presheaves to $\mathcal{E}$, as follows.

Corollary 4. *For each functor $A\colon \mathbf{C} \to \mathcal{E}$ from a small category $\mathbf{C}$ to a cocomplete category $\mathcal{E}$ there exists a colimit preserving functor $L\colon \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}} \to \mathcal{E}$ for which the following diagram (with the Yoneda embedding $\mathbf{y}$) commutes*

$$\begin{array}{ccc} \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}} & \overset{L}{\dashrightarrow} & \mathcal{E} \\ {\scriptstyle \mathbf{y}}\uparrow & \nearrow_{A} & \\ \mathbf{C}. & & \end{array} \tag{13}$$

The functor L with these properties is unique up to isomorphism, and can be defined as in (6) *by a colimit.*

Proof: It will suffice to prove that L of (6) makes the diagram (13) commute; then L, as a left adjoint, preserves colimits. Moreover, since every presheaf P is a colimit of representable presheaves, L is unique up to isomorphism, as asserted in the corollary.

To prove that (13) commutes, note that when $P = \mathrm{Hom}(-, C) = \mathbf{y}C$ is representable the corresponding category of elements $\int P$ has a terminal object—the element $1\colon C \to C$ of $P(C)$. Therefore the colimit of the composite $A \circ \pi_P$ will be just the value of $A \circ \pi_P$ on the terminal object. Hence,

$$L\mathbf{y}(C) \cong A\pi_P(C, 1_C) = A(C)$$

so the diagram does commute.

The process $\mathbf{C} \mapsto \widehat{\mathbf{C}}$ is a functor from **Cat**, the (large) category of all small categories, to **Cocomp** the ("superlarge") category of "all" cocomplete categories, with morphisms all colimit preserving functors. This corollary states in effect that the Yoneda embedding provides universal arrows and so, like universal arrows generally, constitute the units

of an adjunction—an adjunction in which $\mathbf{C} \mapsto \widehat{\mathbf{C}}$ is left adjoint to the forgetful functor **Cocomp** $\to$ **Cat**—forget cocompleteness. This suggestive formulation stumbles on the fact that cocomplete categories are hardly ever small [**CWM**, p. 110], so do not become small by forgetting the colimits!

It is convenient to picture the category $\int P$ of elements of P in terms of its projection π_P by a diagram

$$\begin{array}{ccc} \int_{\mathbf{C}} P & & p' \dashrightarrow p \in P(C) \\ {\scriptstyle \pi_P}\downarrow & & \vdots \qquad\qquad \vdots \\ \mathbf{C} \xrightarrow[P]{} \mathbf{Sets}, & & C' \xrightarrow[u]{} C. \end{array} \tag{14}$$

(Here the object function of the functor P is a set indexed by the set $\mathbf{C}_0$ of objects of $\mathbf{C}$, and the objects of $\int P$ form the corresponding set "over" $\mathbf{C}_0$, in the sense described in §1: J-indexed sets $\sim$ sets over J.) In this diagram (14) the inverse image under π of an object C of $\mathbf{C}_0$ is the set of all $p \in P(C)$, while the projection π has the property that for each such $x \in P(C)$ and each $u\colon C' \to C$ there is a unique pair p', $u'\colon (C', p') \to (C, p)$ with $\pi p' = C'$, $\pi(u') = u$. Any functor $\pi\colon \mathbf{E} \to \mathbf{C}$ with this latter property is called a *fibration* of categories; in geometric terminology, given a point p over C each arrow in the base with target C lifts uniquely to an arrow "upstairs" with target p.

A construction similar to (14) may be carried out when **Sets** is replaced by the category **Cat** of all small categories; it is often called the Grothendieck construction, but the case of $\int P$ above was first done by Yoneda and developed by Mac Lane well before Grothendieck.

6. Exponentials

The basic arithmetic operations on numbers and on sets are $+$, times, and exponent. We have already described $+$ and times categorically as coproduct and product, respectively; we now consider the exponent Z^X. In **Sets**, this Z^X is the usual "function set", consisting of all functions $h\colon X \to Z$. It may be described by the familiar bijection

$$\mathrm{Hom}(Y \times X, Z) \to \mathrm{Hom}(Y, Z^X) \tag{1}$$

which sends each function $f\colon Y \times X \to Z$ of two variables into the function $f'\colon Y \to Z^X$ where, for each $y \in Y$, $f'y \in Z^X$ is the function with $(f'y)(x) = f(y, x) \in Z$. This bijection (1) completely determines Z^X up to isomorphism; for setting $Y = 1$ yields $Z^X \cong \mathrm{Hom}(1, Z^X) \cong$

$\mathrm{Hom}(X, Z)$. The bijection (1) is natural in Y, X, and Z and so states that the functor $(-)^X$ is the right adjoint of $- \times X$: **Sets** $\to$ **Sets**.

Now consider any category $\mathbf{C}$ with finite products. Then for each object X of $\mathbf{C}$ the evident assignment $Y \mapsto Y \times X$ determines a functor $- \times X$: $\mathbf{C} \to \mathbf{C}$, called "product with X". When this functor has a right adjoint, written $Z \mapsto Z^X$, we say that $\mathbf{C}$ has an exponential for X; this means that there is a bijection (1) natural in the objects Y and Z of $\mathbf{C}$. When this holds for all objects X, it implies, by the "parameter theorem" for adjunctions [**CWM**, p. 100], that $\langle X, Z \rangle \mapsto Z^X$ is a functor $\mathbf{C}^{\mathrm{op}} \times \mathbf{C} \to \mathbf{C}$ called the *exponential* for the category $\mathbf{C}$.

The existence of the adjunction (1) can be stated in elementary terms (i.e., without using Hom-sets). For, set $Y = Z^X$ in (1); the identity arrow 1: $Z^X \to Z^X$ on the right in (1) then corresponds under the bijection to an arrow $e = e_{Z,X}$,

$$e\colon Z^X \times X \to Z \tag{2}$$

called *evaluation*; in **Sets** this arrow e is the actual evaluation $e(h, x) = h(x)$ of the function h: $X \to Z$ at the argument $x \in X$. The bijection $f \mapsto f'$ of (1), by naturality, now becomes the statement that to each f: $Y \times X \to Z$ there is a unique f': $Y \to Z^X$ such that the diagram

$$\begin{array}{ccc} Y \times X & & \\ {\scriptstyle f' \times 1}\downarrow & \searrow^{f} & \\ Z^X \times X & \xrightarrow[e]{} & Z \end{array}$$

commutes. One also says that e is *universal* from $- \times X$ to Z. Also, e is the *counit* of the adjunction (1); the adjunction can also be described in terms of its unit η: $Y \to (Y \times X)^X$, as explained in the preliminaries.

The evaluation $e = e_{Z,X}$ of (2) is natural in Z and "dinatural" in X. The latter notion applies because the domain of e is a bifunctor in X contravariant in the first X and covariant in the second; dinaturality means (see [**CWM**, p. 214]) that for every arrow t: $X \to X'$ the diagram

$$\begin{array}{ccc} Z^{X'} \times X & \xrightarrow{1^t \times 1} & Z^X \times X \\ {\scriptstyle 1 \times t}\downarrow & & \downarrow{\scriptstyle e} \\ Z^{X'} \times X' & \xrightarrow[e]{} & Z \end{array}$$

commutes. One also has natural isomorphisms $Z^{X \times Y} \cong (Z^Y)^X$ and $Z^1 \cong Z$, with the evident values of the respective evaluations.

Recall that a category $\mathbf{C}$ is *cartesian closed* when it has a terminal object 1 and a binary product $X \times Y$ for any two objects X and Y, as well as exponentials Y^X (with their evaluations, for all objects X and Y). For example, the category **Sets** of all small sets and the category **Cat** of all small categories are both cartesian closed. Also, any product of cartesian closed categories is cartesian closed: both finite products and exponentials can be computed "termwise" in such a product category.

In any cartesian closed category, there are natural isomorphisms

$$1^X \cong 1, \qquad X^1 \cong X, \tag{3}$$

$$(Y \times Z)^X \cong Y^X \times Z^X, \qquad X^{(Y \times Z)} \cong (X^Y)^Z. \tag{4}$$

The last equation of (4) follows from the definition of the exponential and the associativity of the product; the first of (4) holds because $(-)^X$ has a left adjoint, hence preserves products.

All our typical categories are cartesian closed. The case of sheaves will be discussed in the next chapter. The following proposition takes care of most of the other cases.

Proposition 1. *For any small category* $\mathbf{C}$, *the functor category* $\mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ *is cartesian closed.*

To prove this proposition, recall that the product of two functors P and $Q\colon \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$ is their pointwise product. However, we cannot use a "pointwise" exponential $Q^P(C) = \mathrm{Hom}(PC, QC)$ because the right-hand side here is not a functor of C in any reasonable way.

To find a formula for the exponential, we first assume that an exponential Q^P exists, so that $\mathrm{Hom}(R \times P, Q) \cong \mathrm{Hom}(R, Q^P)$ for all R. In particular, for each representable functor $R = \mathrm{Hom}_{\mathbf{C}}(-, C) = \mathbf{y}C$, this isomorphism composed with the Yoneda isomorphism gives

$$\begin{aligned} Q^P(C) &\cong \mathrm{Hom}_{\widehat{\mathbf{C}}}(\mathbf{y}(C), Q^P) \\ &\cong \mathrm{Hom}_{\widehat{\mathbf{C}}}(\mathbf{y}(C) \times P, Q). \end{aligned}$$

Now drop the assumption that Q^P exists, but use this result to *define* Q^P as the functor

$$Q^P(C) = \mathrm{Hom}_{\widehat{\mathbf{C}}}(\mathbf{y}(C) \times P, Q); \tag{5}$$

i.e., $Q^P(C)$ is the set of all natural transformations θ from $\mathrm{Hom}_{\mathbf{C}}(-, C) \times P$ to Q. This clearly defines a functor $Q^P\colon \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$.

Associate with this definition (5) a putative evaluation map $e\colon Q^P \times P \to Q$ with components

$$e_C(\theta, y) = \theta_C(1_C, y) \in Q(C) \tag{6}$$

for $C \in \mathbf{C}$, $\theta\colon \mathrm{Hom}_{\mathbf{C}}(-, C) \times P \to Q$ and $y \in P(C)$. It follows that e is a natural transformation. Moreover, to every natural transformation $\phi\colon R \times P \to Q$ one can find a (unique) $\phi'\colon R \to Q^P$ such that the diagram of natural transformations

$$\begin{array}{ccc} R \times P & & \\ {\scriptstyle \phi' \times 1}\downarrow & \searrow^{\phi} & \\ Q^P \times P & \xrightarrow[e]{} & Q \end{array} \tag{7}$$

is commutative. Specifically, for $C \in \mathbf{C}$ and $u \in RC$, we must define an element $\phi'_C(u) \in Q^P(C)$, that is, a natural transformation $\phi'_C(u)\colon \mathrm{Hom}_{\mathbf{C}}(-, C) \times P \to Q$. We define the components $(\phi'_C(u))_D$ for $f\colon D \to C$ and $x \in P(D)$ in terms of ϕ by

$$\begin{aligned} (\phi'_C(u))_D\colon \mathrm{Hom}_{\mathbf{C}}(D, C) \times P(D) &\to Q(D), \\ (f, x) &\mapsto \phi_D(u \cdot f, x). \end{aligned} \tag{8}$$

The ϕ' so defined is clearly natural in D. Moreover, by the definition (6) of the evaluation e, for $u \in R(C)$ and $y \in P(C)$,

$$\begin{aligned} e_C(\phi'_C(u), y) &= (\phi'_C(u))_C(1_C, y) \\ &= \phi_C(u, y), \qquad \text{by (8)}. \end{aligned}$$

This means that the triangle (7) commutes, and that this condition plus naturality forces our definition of ϕ'. Therefore Q^P is the required adjoint and Proposition 1 is proved.

A somewhat different description of the same exponential appears as Exercise 8. The meaning of the formula (5) for the exponential Q^P is also illuminated by the special case in which the category $\mathbf{C}$ is a monoid or a group (Exercise 5).

It follows from Proposition 1 that most of our typical categories are cartesian closed. [Examples (xiii) and (xiv) of §1 are similar to (i) and (viii).] Besides the case of sheaves, which will be discussed in the next chapter, this only leaves the case of continuous G-sets for a topological group G; a construction of exponentials in this category is outlined in Exercise 6.

A *global section* γ of a functor $P\colon \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$ is defined to be a function γ which assigns to each object C of $\mathbf{C}$ an element $\gamma_C \in P(C)$ in such a way that the equation

$$\gamma_C \cdot f = \gamma_D, \qquad f\colon D \to C, \tag{9}$$

holds for every arrow f of $\mathbf{C}$. Thus γ is just a natural transformation $\gamma\colon 1 \to P$, where 1 is the constant functor 1 on $\mathbf{C}^{\mathrm{op}}$. (The geometric

origins of the term "global" will appear in Chapter II below.) The set $\Gamma(P)$ of all global sections γ of P yields a functor

$$\Gamma\colon \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}} \to \mathbf{Sets}.$$

In the opposite direction, the *constant presheaf* functor Δ assigns to each set S the functor ΔS with $(\Delta S)(C) = S$ and every $(\Delta S)(f)$ the identity. For each S and P there is a natural isomorphism

$$\mathrm{Hom}_{\widehat{\mathbf{C}}}(\Delta S, P) \cong \mathrm{Hom}_{\mathbf{Sets}}(S, \Gamma P) \tag{10}$$

since a natural transformation $\Delta S \to P$ simply assigns to each element $s \in S$ its image, a global section $1 \to P$ of P. Therefore, the functor Δ is left adjoint to the global sections functor Γ. This adjunction (where the left adjoint Δ is left exact) is a first instance of what will later be called a "geometric morphism" (Chapter VII). Also, a natural transformation $\Delta S \to P$ is just a cone from the set S to the functor P to **Sets**. Hence (10) states also that ΓP is exactly $\varprojlim P$.

We can now summarize the common properties of our typical categories. They are categories $\mathcal{E}$ with the following properties

(i) $\mathcal{E}$ has all finite limits and colimits,
(ii) $\mathcal{E}$ has exponentials,
(iii) $\mathcal{E}$ has a subobject classifier $1 \to \Omega$.

A category $\mathcal{E}$ with these properties will be called an *elementary topos*; in brief a topos (plural: topoi). Each topos is, in particular, a cartesian closed category.

7. Propositional Calculus

The propositional calculus considers "Propositions" $p, q, r, \ldots$ combined under the operations "and", "or", "implies", and "not", often written as $p \wedge q$, $p \vee q$, $p \Rightarrow q$, and $\neg p$. Alternatively, if $P, Q, R, \ldots$ are subsets of some fixed set U with elements u, each proposition p may be replaced by the proposition $u \in P$ for some subset $P \subset U$; the propositional connectives above then become operations on subsets; intersection $\wedge$, union $\vee$, implication ($P \Rightarrow Q$ is $\neg P \wedge Q$), and complement of subsets. These four operations satisfy various identities, so that the subsets of U under these operations constitute a Boolean algebra. In this way a Boolean algebra is the algebraic correlate of the classical propositional calculus. If, instead, one takes the intuitionistic propositional calculus, as formalized by Heyting, one obtains a different algebraic system on the same operations $\wedge$, $\vee$, $\Rightarrow$, $\neg$; such a system is known as a Heyting algebra. The typical model is not the set of *all* subsets of some set, but

the set of all *open* subsets of some topological space X; in the model, the operations $\wedge$ and $\vee$ still correspond to intersection and union, respectively, but the other two operations must be reinterpreted (in order to give open sets). Thus $U \Rightarrow V$ is the largest open set W such that $W \wedge U \subset V$, while $\neg U$ is the interior of the complement of U (the largest open set disjoint from U).

We now formulate the exact definitions, beginning with the notion of a lattice. A *lattice* L is a partially ordered set which, considered as a category, has all binary products and all binary coproducts. If we write x, y, and z for objects of L, then $x \leq y$ if and only if there is a (unique) arrow $x \to y$, the coproduct of x and y is the least upper bound (or sup) $x \vee y$ and the product is the greatest lower bound (or inf) $x \wedge y$. If a lattice L has elements 0 and 1 such that $0 \leq x \leq 1$ for all $x \in L$, then 0 and 1 are the (unique) initial and terminal objects, respectively, of L, considered as a category. Thus a lattice with 0 and 1 is a partially ordered set which, considered as a category, has all finite limits and all finite colimits.

A lattice with 0 and 1 can also be defined equationally, as a set with two distinguished elements 0 and 1 and two binary operations $\vee$ and $\wedge$, both of which are both associative and commutative and which satisfy the added identities

$$\begin{aligned} x \wedge x = x, \qquad & x \vee x = x, \\ 1 \wedge x = x, \qquad & 0 \vee x = x, \\ x \wedge (y \vee x) = x = & (x \wedge y) \vee x. \end{aligned} \tag{1}$$

These equations on the operations

$$\wedge \colon L \times L \to L, \quad \vee \colon L \times L \to L, \quad 0, 1 \colon 1 \to L$$

can be used to define a "lattice object" L in any category $\mathbf{C}$ with finite products. Here they follow from the above definitions of $\wedge$ and $\vee$ in terms of the partial order. And, given these equations, the partial order can be recovered because $x \leq y$ holds in L if and only if $x = x \wedge y$ (or, equivalently, $y = x \vee y$).

A *distributive lattice* L is a lattice in which the identity

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z) \tag{2}$$

holds for all x, y, and z. This identity implies the dual distributive law

$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z). \tag{3}$$

For, the right-hand side of (3) expands by (2) and then (1) to give, by associativity,

$$\begin{aligned} (x \vee y) \wedge (x \vee z) &= [(x \vee y) \wedge x] \vee [(x \vee y) \wedge z] \\ &= x \vee [(x \wedge z) \vee (y \wedge z)] \\ &= [x \vee (x \wedge z)] \vee (y \wedge z) = x \vee (y \wedge z). \end{aligned}$$

A *complement* for an element x in a lattice L with 0 and 1 is an element $a \in L$ such that

$$x \wedge a = 0, \qquad x \vee a = 1. \tag{4}$$

In a distributive lattice a complement a, if it exists, is unique. For let b be another complement to x. Then

$$\begin{aligned} b = b \wedge 1 = b \wedge (x \vee a) &= (b \wedge x) \vee (b \wedge a) \\ &= (x \wedge a) \vee (b \wedge a) \\ &= (x \vee b) \wedge a = a. \end{aligned}$$

We denote the unique complement a of x, when it exists, by $a = \neg x$.

A *Boolean algebra* B is a distributive lattice with 0 and 1 in which every element x has a complement $\neg x$; thus,

$$x \wedge \neg x = 0, \qquad x \vee \neg x = 1. \tag{5}$$

One may readily verify the additional properties

$$\neg(x \vee y) = \neg x \wedge \neg y, \quad \neg(x \wedge y) = (\neg x) \vee (\neg y), \tag{6}$$

$$\neg\neg x = x. \tag{7}$$

The identities (6) are called the DeMorgan laws.

The partially ordered set of all the subsets of a given set is always a Boolean algebra. A basic theorem due to M. H. Stone asserts that every Boolean algebra is isomorphic to an algebra of some of the subsets of some set U. We will use this theorem in Chapter IX.

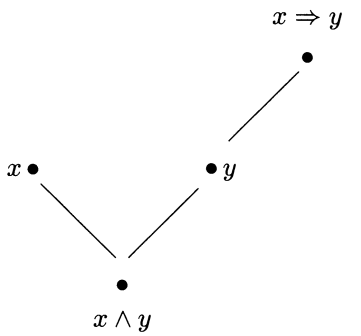
8. Heyting Algebras

A *Heyting algebra* H (also called a Brouwerian lattice) is a poset with all finite products and coproducts which is cartesian closed (as a category with products). In other words, a Heyting algebra is a lattice with 0 and 1 which has to each pair of elements x, y an exponential y^x. This exponential is usually written as $x \Rightarrow y$; by its definition it is characterized by the adjunction

$$z \leq (x \Rightarrow y) \quad \text{if and only if} \quad z \wedge x \leq y. \tag{1}$$

In other words, $x \Rightarrow y$ is a least upper bound for all those elements z with $z \wedge x \leq y$; in particular, then, $y \leq (x \Rightarrow y)$. Thus, in the usual

picture of a partially ordered set, $x \Rightarrow y$ lies above y,



but only so far above that its intersection with x is still $x \wedge y$. For example, for any topological space X the set $\mathrm{Open}(X)$ of all open sets in X is a Heyting algebra: It is a lattice (under inclusion) because binary unions and intersections of open sets are open, as are the sets $\emptyset$ and X. For two open sets U and V the exponential $U \Rightarrow V$ can be defined, as suggested by (1), as the union $\bigcup W_i$ of all those open sets W_i for which $W_i \cap U \subset V$. Then, because intersection is distributive over arbitrary unions,

$$(\bigcup W_i) \cap U = \bigcup (W_i \cap U) \subset V.$$

Therefore, $\bigcup W_i = (U \Rightarrow V)$.

A similar argument will show that any complete and (infinitely) distributive lattice is a Heyting algebra. Here a lattice is said to be *complete* when, regarded as a category, it has all small limits and small colimits, i.e., all small products and coproducts.

In a Boolean algebra, for all x, y, and z,

$$z \leq (\neg x \vee y) \quad \text{if and only if} \quad z \wedge x \leq y.$$

Proof, only if: $z \wedge x \leq (\neg x \vee y) \wedge x \leq y \wedge x \leq y$; **if:** $z = z \wedge 1 = z \wedge (\neg x \vee x) = (z \wedge \neg x) \vee (z \wedge x) \leq \neg x \vee y$.

Hence a Boolean algebra has exponentials given by

$$(x \Rightarrow y) = \neg x \vee y.$$

This is the classical definition of material implication $\Rightarrow$ (for propositions, $p \Rightarrow q$ is "not p or q"). Therefore, every Boolean algebra is a Heyting algebra. The converse does not hold; for example, the open sets in the real line form a Heyting algebra which is not Boolean (because the complement of an open set need not be open).

Now we examine the identities valid in any Heyting algebra. For any cartesian closed category with objects X, Y, the unit and the counit of

the basic adjunction defining the exponential are natural transformations

$$X \to (X \times Y)^Y, \qquad Y \times X^Y \to X.$$

For a Heyting algebra these become the inclusions

$$x \leq (y \Rightarrow (x \wedge y)), \qquad y \wedge (y \Rightarrow x) \leq x. \tag{2}$$

The properties $1^X \cong 1$ and $X^1 \cong X$ of the exponential become

$$(x \Rightarrow 1) = 1, \qquad (1 \Rightarrow x) = x. \tag{3}$$

Since the functor $x \Rightarrow (-)$ is a right adjoint, it preserves products, so

$$(x \Rightarrow (y \wedge z)) = ((x \Rightarrow y) \wedge (x \Rightarrow z)), \tag{4}$$

while the associative law for the product of objects implies $X^{Y \times Z} \cong (X^Y)^Z$ which becomes

$$((y \wedge z) \Rightarrow x) = (z \Rightarrow (y \Rightarrow x)). \tag{5}$$

Also $- \wedge y$ is a left adjoint, so must preserve coproducts, as in

$$((x \vee z) \wedge y) = ((x \wedge y) \vee (z \wedge y)). \tag{6}$$

This means that the underlying lattice of any Heyting algebra is distributive.

Intersection in any lattice is commutative. For a Heyting algebra H this means that, for all x, y, and z,

$$z \leq (x \Rightarrow y) \text{ iff } z \wedge x \leq y \text{ iff } x \wedge z \leq y \text{ iff } x \leq (z \Rightarrow y).$$

Now $- \Rightarrow y$ (like any exponential) is a contravariant functor in the argument $-$. Hence, in this display we may regard $- \Rightarrow y$ on the right as a functor from H to H^{op} and on the left as a functor from H^{op} to H; the equivalence then asserts that the first $- \Rightarrow y$ is left adjoint to the second $- \Rightarrow y$. Since any left adjoint preserves coproducts, this means that $- \Rightarrow y$ carries coproducts to products (coproducts in H^{op}), as in the identity

$$((x \vee z) \Rightarrow y) = ((x \Rightarrow y) \wedge (z \Rightarrow y)). \tag{7}$$

If one interprets x, y, and z as propositions, with $\wedge$ as "and" and $\vee$ as "or", all the equations (2) through (7) become familiar properties of the implication relation $\Rightarrow$ for propositions.

In any Heyting algebra we define the negation of x as

$$\neg x = (x \Rightarrow 0). \tag{8}$$

Thus "not x" means "x implies falsity" or "x implies absurdity". In view of the definition of $\Rightarrow$, this can be rewritten as

$$y \leq \neg x \quad \text{iff} \quad y \wedge x = 0. \tag{9}$$

In other words, in a complete lattice $\neg x$ is the union of all those y which meet x in 0. For example, in the case of the Heyting algebra $\mathrm{Open}(X)$ of all open subsets of a topological space X, the negation $\neg U$ is the union of all open subsets of X which do not meet U, so is the interior of the set-theoretic complement of U; that is, the set-theoretic complement of the closure of U. Thus $\neg\neg U$ is the interior of the closure of U, which may be larger than U, as for example when U is a suitable open subset of the line or the plane. This example shows that $\neg\neg x$ need not equal x. Moreover, $\neg x$ is not necessarily a complement of x; though $x \wedge \neg x = 0$, it may not be the case that $x \vee \neg x = 1$.

Some of the familiar properties of negation still apply, as follows.

Proposition 1. *In any Heyting algebra H,*

$$x \leq \neg\neg x, \quad x \leq y \text{ implies } \neg y \leq \neg x, \tag{10}$$

$$\neg x = \neg\neg\neg x, \tag{11}$$

$$\neg\neg(x \wedge y) = \neg\neg x \wedge \neg\neg y. \tag{11'}$$

Proof: Since $x \wedge \neg x = \neg x \wedge x = 0$, the first follows by (9). The second of (10) states that $\neg\colon H \to H^{\mathrm{op}}$ is a functor. Explicitly, if $x \leq y$, then $x \wedge \neg y \leq y \wedge \neg y = 0$, so $\neg y \leq \neg x$, again by (9). This result and $x \leq \neg\neg x$ gives $\neg\neg\neg x \leq \neg x$, while $x \leq \neg\neg x$ holds for all x, hence for $\neg x$, and so gives $\neg x \leq \neg\neg\neg x$. Hence, (11) holds. Furthermore, two applications of (10) to the inequality $x \wedge y \leq x$ yield $\neg\neg(x \wedge y) \leq \neg\neg x$. Similarly, one derives $\neg\neg(x \wedge y) \leq \neg\neg y$; therefore, $\neg\neg(x \wedge y) \leq \neg\neg x \wedge \neg\neg y$. To show the converse inequality $\neg\neg x \wedge \neg\neg y \leq \neg\neg(x \wedge y)$ we use the commutativity and associativity of the meet $\wedge$, together with (9) and (11), as follows:

$$\begin{aligned}
\neg\neg x \wedge \neg\neg y \leq \neg\neg(x \wedge y) \quad &\text{iff} \quad \neg\neg x \wedge \neg\neg y \wedge \neg(x \wedge y) = 0 && \text{by (9)},\\
&\text{iff} \quad \neg\neg y \wedge \neg(x \wedge y) \leq \neg\neg\neg x && \text{by (9)},\\
&\text{iff} \quad \neg\neg y \wedge \neg(x \wedge y) \leq \neg x && \text{by (11)},\\
&\text{iff} \quad \neg\neg y \wedge \neg(x \wedge y) \wedge x = 0 && \text{by (9)},\\
&\text{iff} \quad \neg(x \wedge y) \wedge x \leq \neg\neg\neg y = \neg y && \text{by (9) and (11)},\\
&\text{iff} \quad \neg(x \wedge y) \wedge x \wedge y = 0 && \text{by (9)}.
\end{aligned}$$

But in any Heyting algebra, the identity $\neg z \wedge z = 0$ holds, as an immediate consequence of the definitions (1) and (8).

Proposition 2. *In a Heyting algebra, if an element x has a complement, that complement must be $\neg x$.*

For this reason, $\neg x$ is sometimes called the *pseudo-complement* of x.

Proof: Suppose that x has a complement a, with $x \wedge a = 0$ and $x \vee a = 1$. By the first of these equations, $a \leq \neg x$. By the second and the distributive law,

$$\neg x = \neg x \wedge (x \vee a) = \neg x \wedge a;$$

hence $\neg x \leq a$. Combining these results, $a = \neg x$, as asserted.

Boolean algebras can be defined equationally, specifically by the associative and commutative laws and the equations (7.1), (7.2), and (7.5) on the operations $\wedge$, $\vee$, $\neg$ and the elements 0 and 1. A corresponding result holds, in a more subtle way, for Heyting algebras; this will be used later to define Heyting algebra objects.

Proposition 3. *In a Heyting algebra H the implication, $\Rightarrow$, satisfies the following identities, for all elements x, y, $z \in H$:*

$$(x \Rightarrow x) = 1, \tag{12}$$

$$x \wedge (x \Rightarrow y) = x \wedge y, \qquad y \wedge (x \Rightarrow y) = y, \tag{13}$$

$$x \Rightarrow (y \wedge z) = (x \Rightarrow y) \wedge (x \Rightarrow z). \tag{14}$$

Conversely, in any lattice L with 0 and 1 a binary operation $\Rightarrow$ satisfying these identities must be the implication of a Heyting algebra structure on the lattice L.

Proof: Since $y \wedge x \leq x$ for all y, the definition of $x \Rightarrow x$ shows that (12) must hold. By the definition of $\Rightarrow$ again, $x \wedge y \leq (x \Rightarrow y)$, while, by evaluation, $x \wedge (x \Rightarrow y) \leq y$; hence, the first of (13) holds. By the definition of $\Rightarrow$ once more, $y \leq (x \Rightarrow y)$, which gives the second of (13), while (14) is just (4), the fact that $x \Rightarrow -$ preserves products.

The equations (12), (13), and (14) represent familiar properties of implication. Also (14) with $z = x$, when combined with $(x \Rightarrow x) = 1$ from (12), gives

$$(x \Rightarrow (y \wedge x)) = x \Rightarrow y. \tag{15}$$

Since (14) states that the operation $(x \Rightarrow -)$ preserves products, it also shows that $(x \Rightarrow -)$ preserves inequalities, so that

$$a \leq b \quad \text{implies} \quad (x \Rightarrow a) \leq (x \Rightarrow b). \tag{16}$$

For the converse of the proposition, we must show for any lattice that a binary operation $\Rightarrow$ satisfying equations (12), (13), and (14) necessarily

satisfies the definition of an exponential; in other words, that (12), (13), and (14) imply

$$z \leq (x \Rightarrow y) \quad \text{if and only if} \quad z \wedge x \leq y \tag{17}$$

for all x, y, and z. Given the left-hand inequality, the first part of (13) yields

$$z \wedge x \leq (x \Rightarrow y) \wedge x = x \wedge y \leq y,$$

which is the right-hand inequality of (17). Conversely, given that inequality, one has

$$\begin{aligned} z = z \wedge (x \Rightarrow z) &\leq x \Rightarrow z, && \text{by (13)},\\ z &\leq [x \Rightarrow (z \wedge x)], && \text{by (15)},\\ [x \Rightarrow (z \wedge x)] &\leq (x \Rightarrow y), && \text{by assumption and (16)},\\ z &\leq (x \Rightarrow y), && \text{by transitivity}. \end{aligned}$$

This completes the proof of (17).

Proposition 4. *A Heyting algebra is Boolean if and only if* $\neg\neg x = x$ *for all* $x \in H$, *or, if and only if* $x \vee \neg x = 1$ *for all* x.

Proof: Since the complement is unique in a Boolean algebra, x is the complement of $\neg x$, so the equation $\neg\neg x = x$ holds there. Conversely, in any Heyting algebra, by (8) and (7),

$$\neg(x \vee y) = (x \vee y) \Rightarrow 0 = (x \Rightarrow 0) \wedge (y \Rightarrow 0) = (\neg x) \wedge (\neg y).$$

This is one of the DeMorgan laws (7.6). Now if also $\neg\neg x = x$ for all x, one has by this law

$$\begin{aligned} x \vee \neg x = \neg\neg(x \vee \neg x) &= \neg(\neg x \wedge \neg\neg x)\\ &= \neg 0\\ &= 1. \end{aligned}$$

Since always $x \wedge \neg x = 0$, this shows that $\neg x$ is a complement of x, and hence that H is indeed Boolean.

The analogous characterization of Boolean algebras by $x \vee \neg x = 1$ is immediate. The assertion $x \vee \neg x = 1$ is the famous "tertium non datur" of classical logic, doubted by intuitionists and constructivists.

As already observed, negation is a functor $\neg\colon H \to H^{\mathrm{op}}$ and also $H^{\mathrm{op}} \to H$. Since $x \leq \neg y$ iff $y \leq \neg x$, this functor is adjoint to itself. Thus, a Heyting algebra is Boolean iff this adjunction is an equivalence (actually, an isomorphism).

The relation between Heyting algebras and our typical categories discussed in §1–6 lies in the fact that the poset of subobjects of a given object in any such typical category is always a Heyting algebra (and sometimes a Boolean algebra). Most cases are taken care of by the following proposition.

Proposition 5. *Consider the functor category* $\widehat{\mathbf{C}} = \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$ *of a given small category* $\mathbf{C}$. *For any object* P *of* $\widehat{\mathbf{C}}$, *the partially ordered set* $\mathrm{Sub}_{\widehat{\mathbf{C}}}(P)$ *of subobjects of* P *is a Heyting algebra.*

Proof: Under pointwise operations, the set $\mathrm{Sub}(P)$ of all subfunctors of P is a complete lattice, satisfying the infinite distributive law. Hence (as for open sets above), it is a Heyting algebra. We list the explicit description of the operations $\wedge$, $\vee$, 0, 1, $\Rightarrow$, and $\neg$, and leave further verification to the reader. If S and T are two given subfunctors of P, then their least upper bound $S \vee T$ and their greatest lower bound $S \wedge T$ may be defined pointwise, as the functors

$$(S \vee T)(C) = S(C) \cup T(C),$$
$$(S \wedge T)(C) = S(C) \cap T(C),$$

since $S(C)$ and $T(C)$ are both subsets of $P(C)$. The implication $S \Rightarrow T$ is defined for C in $\mathbf{C}$ by

$$(S \Rightarrow T)(C) = \{\, x \in P(C) \mid \text{for all } f\colon D \to C \text{ in } \mathbf{C}\text{: if } x \cdot f \in S(D) \text{ then } x \cdot f \in T(D)\} \tag{18}$$

(The pointwise definition doesn't work, because it doesn't give a subfunctor.) The largest and smallest subfunctors of P are respectively the functor P itself, and the empty functor 0 [with $0(C) = \emptyset$ for all C]. Consequently, negation can be described explicitly for a subfunctor S as

$$(\neg S)(C) = \{\, x \in P(C) \mid \text{for all } f\colon D \to C \text{ in } \mathbf{C},\ x \cdot f \notin S(D)\}\,. \tag{19}$$

From the description of negation, it is clear that the identity $\neg S \vee S = P$ need not hold in general (e.g., take $\mathbf{C} = \mathbf{2}$).

In the particular case of $\mathbf{B}G = G\text{-}\mathbf{Sets}$ for a group G [Example (iv) of §1], a subobject of a given G-set X is just a subset S of X which is closed under the action by G (i.e., $x \in S$ implies $x \cdot g \in S$ for all $g \in G$ and $x \in X$). If S and T are two such subsets closed under the action, then so are $S \cup T$, $S \cap T$, and $X - S = \neg S$. So, in this case, the Heyting algebra structure of $\mathrm{Sub}(X)$ is just the restriction of the usual structure on the power set of X. In particular, $\mathrm{Sub}_{\mathbf{B}G}(X)$ is a Boolean algebra. The same reasoning applies to the category $\mathbf{B}G$ of continuous G-sets for a topological group G.

A poset P is *complete* iff every subset of P has an l.u.b. (a "sup" or a join) and a g.l.b. (an "inf" or a meet); actually it suffices to require the existence of all l.u.b.s. Thus P is complete as a poset iff P as a category has all limits and all colimits. A complete poset is necessarily

a lattice with 0 and 1. We will often have to do with complete Heyting algebras (**cHa**'s) and complete Boolean algebras (**cBa**'s); they are, of course, Heyting or Boolean algebras which are complete as posets. The algebra $\mathrm{Open}(X)$ of all open subsets U of a topological space is a **cHa** with the usual sups. However, the inf of a family $\{ U_i \mid i \in I \}$ of open sets is just the largest open set contained in all the U_i and so is not usually the set-theoretic intersection of the sets U_i.

9. Quantifiers as Adjoints

Our discussion of the propositional calculus has indicated that the basic operations $\wedge$, $\vee$, and $\Rightarrow$ of this calculus can all be described as adjoints. We turn now to the more subtle question of interpreting the quantifiers of the usual predicate calculus as adjoints, too.

Consider quantifiers $(\forall x)$ and $(\exists x)$—that is, "for all x" and "there exists an x"—as applied to a predicate $S(x, y)$, where x and y are elements of sets X and Y, respectively. If we regard S as the subset $S \subset X \times Y$ of those pairs $\langle x, y \rangle$ for which $S(x, y)$ is true, then $(\forall x)S(x, y)$ similarly denotes a related subset $T \subset Y$ consisting of all those y with every pair $\langle x, y \rangle \in S$. Writing $p\colon X \times Y \to Y$ for the usual projection, we will denote this subset T, corresponding to $(\forall x)S(x, y)$, as $\forall_p S$; similarly $\exists_p S$ denotes the subset corresponding to $(\exists x)S(x, y)$. Now let $\mathcal{P}Y$ be the Boolean algebra of all subsets $T \subset Y$ and $\mathcal{P}(X \times Y)$ the Boolean algebra of all S. Then $\mathcal{P}Y$ and $\mathcal{P}(X \times Y)$ can be viewed as categories, while the functions $\forall_p$ and $\exists_p$, since they preserve the inclusion relation $S \subset S'$ between subsets, are functors

$$\forall_p, \exists_p\colon \mathcal{P}(X \times Y) \to \mathcal{P}(Y). \tag{1}$$

Theorem 1. *For the projection $p : X \times Y \to Y$, the functors $\exists_p$ and $\forall_p$ are respectively left and right adjoints to the functor $p^*\colon \mathcal{P}(Y) \to \mathcal{P}(X \times Y)$ which sends each subset $T \subset Y$ to its inverse image p^*T under p.*

As usual, the inverse image is $p^*T = \{ \langle x, y \rangle \mid y \in T \}$; it may also be described as the pullback of $T \rightarrowtail Y$ along p.

Proof: For subsets $S \subset X \times Y$ and $T \subset Y$ one evidently has

$$\begin{aligned} p^*T \subset S \quad &\text{if and only if} \quad T \subset \forall_p S, \\ S \subset p^*T \quad &\text{if and only if} \quad \exists_p S \subset T. \end{aligned}$$

Since $p^*T \subset S$ means exactly that the set $\mathrm{Hom}(p^*T, S)$ in the category $\mathcal{P}(X \times Y)$ is nonempty (with one element) and so on, these equivalences give precisely the asserted adjunctions.

Much the same argument applies when the projection p is replaced by an arbitrary function f.

Theorem 2. *For any function $f\colon Z \to Y$ between sets Z and Y the inverse image functor $f^*\colon \mathcal{P}Y \to \mathcal{P}Z$ between subsets has left and right adjoints, $\exists_f$ and $\forall_f$.*

Proof: The left adjoint $\exists_f$ assigns to each $S \subset Z$ its image $\exists_f S \subset Y$, which may be described with a "there exists" as

$$\exists_f S = \{\, y \mid \text{there exists a } z \text{ with } fz = y \text{ and } z \in S \,\}.$$

The right adjoint $\forall_f$ assigns to each S the set

$$\forall_f S = \{\, y \mid \text{for all } z, \text{ if } fz = y, \text{ then } z \in S \,\},$$

described with a "for all". The proof that these are adjoints, as stated, is immediate. The notations $\exists_f$ and $\forall_f$ are chosen to match the special case of a projection $f = p$, when these adjoints correspond to ordinary quantifiers.

The result (left adjoint on the left) is the diagram

$$\begin{array}{ccc} Z & & \mathcal{P}Z \\ {\scriptstyle f}\downarrow & & \exists_f \downarrow \uparrow f^* \uparrow \downarrow \forall_f \\ Y, & & \mathcal{P}Y. \end{array} \tag{2}$$

This construction has provided adjoints for the operation f^* of pulling back a subobject of a set Y.

More generally, such adjoints exist not just for subsets S of a set Z but for arbitrary sets B over a given set Z. Indeed, for each function $f\colon Z \to Y$, as in Theorem 2, consider the pullback functor

$$f^*\colon \mathbf{Sets}/Y \to \mathbf{Sets}/Z \tag{3}$$

defined for any set A over Y by

$$f^*(A \xrightarrow{u} Y) = A \times_Y Z \xrightarrow{\pi_2} Z. \tag{4}$$

If we identify a set A over Y with a Y-indexed family $\{\, A_y \mid y \in Y \,\}$ of sets, as in (1.8), the pullback functor (2) corresponds to "reindexing" via f: it sends a Y-indexed family $\{\, A_y \mid y \in Y \,\}$ to the Z-indexed family $\{\, A_{f(z)} \mid z \in Z \,\}$.

Theorem 3. *For any function $f\colon Z \to Y$ between sets, the pullback functor $f^*\colon \mathbf{Sets}/Y \to \mathbf{Sets}/Z$ has both a left and a right adjoint.*

Proof: By the equivalences $\mathbf{Sets}/Y \cong \mathbf{Sets}^Y$ and $\mathbf{Sets}/Z = \mathbf{Sets}^Z$, we may as well prove that the reindexing functor

$$f^*\colon \mathbf{Sets}^Y \to \mathbf{Sets}^Z, \quad f^*(\{ A_y \mid y \in Y \}) = \{ A_{f(z)} \mid z \in Z \}$$

has both adjoints. The left adjoint is

$$\Sigma_f\colon \mathbf{Sets}^Z \to \mathbf{Sets}^Y,$$

defined for a Z-indexed family $B = \{ B_z \mid z \in Z \}$ by

$$(\Sigma_f(B))_y = \sum_{f(z)=y} B_z, \tag{5}$$

where Σ denotes the coproduct (disjoint union) of the sets B_z. The right adjoint

$$\Pi_f\colon \mathbf{Sets}^Z \to \mathbf{Sets}^Y$$

is defined by a cartesian product Π as

$$(\Pi_f(B))_y = \prod_{f(z)=y} B_z. \tag{6}$$

In words, given an indexed set $B = \{ B_z \mid z \in Z \}$, the indexed set $\Sigma_f(B)$ has at index y the coproduct of all the sets B_z for which $f(z) = y$. This set B_z is the "fiber" over z. Dually, $\Pi_f B$ is the product over the fibers. As for the proof that Σ_f is left adjoint to f^*, just observe that an indexed family of maps $h_y\colon (\Sigma_f B)_y \to A_y$ $(y \in Y)$ is the same thing as an indexed family of maps $B_z \to A_{f(z)}$ $(z \in Z)$. The proof that Π_f is right adjoint to the pullback f^* is similar.

More generally, suppose $\mathbf{C}$ is an arbitrary category with pullbacks. Then for each morphism of objects $f\colon B' \to B$ in $\mathbf{C}$, pulling back along f induces a functor between the corresponding slice categories

$$f^*\colon \mathbf{C}/B \to \mathbf{C}/B' \tag{7}$$

(the functor, of course, depends on the particular choice of the pullbacks). $\mathbf{C}/B$ is also called the category of objects over the "base" object B, and a functor of the form (7) is also called a *change of base* functor. Theorem 3 is a special case of the following result:

Theorem 4. *Let $\mathbf{C}$ be a category with pullbacks, and let B be an object of $\mathbf{C}$. For each $f\colon B' \to B$, the change of base functor $f^*\colon \mathbf{C}/B \to \mathbf{C}/B'$ has a left adjoint; moreover, if $\mathbf{C}/B$ is cartesian closed, each such f^* also has a right adjoint.*

The left adjoint is given by composition with f.

With Σ_f the left adjoint and Π_f the right, the result is

$$\begin{array}{ccc} B' & & \mathbf{C}/B' \\ {\scriptstyle f}\downarrow & & \Sigma_f \downarrow \; f^* \; \uparrow \downarrow \Pi_f \\ B, & & \mathbf{C}/B. \end{array} \tag{8}$$

We prove the theorem first in the case when $B = 1$ is the terminal object of $\mathbf{C}$, so that $\mathbf{C}/1$ is just (isomorphic to) $\mathbf{C}$, while pullback along the unique arrow $B' \to 1$ is just the functor

$$- \times B' \colon \mathbf{C} \to \mathbf{C}/B',$$

sending each object X to the object $X \times B' \to B'$ over B' (by projection). Take any object $h \colon Y \to B'$ in $\mathbf{C}/B'$. An arrow from this object $Y \to B'$ to $X \times B' \to B'$ in $\mathbf{C}/B'$ is then just an arrow from Y to X in $\mathbf{C}$; hence, a left adjoint to $- \times B'$ is the forgetful functor $\Sigma \colon \mathbf{C}/B' \to \mathbf{C}$ given by $\Sigma(Y \to B') = Y$.

On the other hand, an arrow from $X \times B' \to B'$ to $h \colon Y \to B'$ in $\mathbf{C}/B'$ is just an arrow $t \colon X \times B' \to Y$ in $\mathbf{C}$ such that ht is the projection $X \times B' \to B'$. By exponential adjunction, these arrows t correspond to those arrows $t' \colon X \to Y^{B'}$ for which $h^{B'} \circ t'$ is the composite $X \to 1 \xrightarrow{j} B'^{B'}$, where j arises by exponential adjunction from the identity $B' \to B'$. These arrows t' in turn correspond by pullback exactly to the arrows $t'' \colon X \to \Gamma h$, where Γh is the pullback in the square

$$\begin{array}{ccc} X & \xrightarrow{t''} & \Gamma h \longrightarrow Y^{B'} \\ & & \downarrow \qquad\quad \downarrow h^{B'} \\ & & 1 \xrightarrow[j]{} B'^{B'}. \end{array}$$

Therefore, Γh, the pullback of $h^{B'}$ along j, is the desired right adjoint to $- \times B'$.

Note that, if $\mathbf{C} = \mathbf{Sets}$, this pullback Γh is just the set of those functions on B' to Y whose composite with $h \colon Y \to B'$ is the identity of B'; that is, the set of cross sections of the map h. Hence, in general, we might call Γh the object of "cross sections" of the arrow h.

Now return to the general case of any $f \colon B' \to B$. This arrow f is also an object (f) in the slice category $\mathbf{C}/B$; moreover, an object over (f) is just a commutative square

$$\begin{array}{ccc} X & \longrightarrow & B \\ \downarrow & & \| \\ B' & \xrightarrow[f]{} & B \end{array}$$

and this square is determined by $X \to B'$; that is, by an object in $\mathbf{C}/B'$. This correspondence is an isomorphism of slice categories

$$(\mathbf{C}/B)/(f) \cong \mathbf{C}/B',$$

and pullback along $f^*\colon \mathbf{C}/B \to \mathbf{C}/B' = (\mathbf{C}/B)/(f)$ is reduced to the previous case of an object [that is, the object (f)] in the cartesian closed category $\mathbf{C}/B$. This proves Theorem 4.

From this theorem we can conclude that pullbacks preserve colimits, in the following sense.

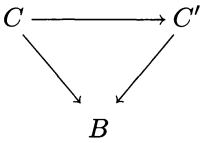
Proposition 5. *If B and B' are objects in a complete category $\mathbf{C}$ with pullbacks such that all the categories $\mathbf{C}$, $\mathbf{C}/B$, and $\mathbf{C}/B'$ are cartesian closed, then pullback along any arrow $f\colon B' \to B$ preserves all colimits which exist in $\mathbf{C}/B$.*

Proof: Since pullback is a functor with a right adjoint, by Theorem 4, it must, like all left adjoints, preserve colimits.

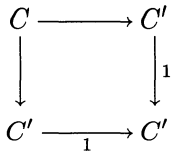
When colimits are preserved, as in this case, by pullbacks, we say that they are *stable* under pullback.

Notice, incidentally, for a category $\mathbf{C}$ with products, that a cocone in the category $\mathbf{C}/B$ is a colimit there iff the corresponding cocone in $\mathbf{C}$ (obtained by forgetting the arrows to B) is a colimit in $\mathbf{C}$. Indeed, the forgetful functor $U\colon \mathbf{C}/B \to \mathbf{C}$ has a right adjoint $B^*\colon \mathbf{C} \to \mathbf{C}/B$ (product with B), hence preserves colimits. Conversely, the fact that if the cocone yields a colimit in $\mathbf{C}$ then it was already a colimit in $\mathbf{C}/B$, follows immediately from the universal property of the colimit.

As a consequence, note also that a map



in $\mathbf{C}/B$ is an epi there iff $C \to C'$ is an epi in $\mathbf{C}$. Indeed, the square



(9)

consists of maps over B, and is a pushout in $\mathbf{C}/B$ iff it is a pushout in $\mathbf{C}$. But (9) is a pushout in either category iff $C \to C'$ is an epi there.

Exercises

1. Show that pullbacks of epis are epi for categories of each of the types (i)–(ix).
2. Prove that $\mathbf{FinSets}^{\mathbf{N}}$ has no subobject classifier.
3. For R a ring, prove that the category $R-\mathbf{Mod}$ of left R-modules has no subobject classifier.
4. If $\mathbf{A} \to \mathbf{B}$ is an equivalence of categories, prove that a subobject classifier for $\mathbf{A}$ yields one for $\mathbf{B}$, and that $\mathbf{A}$ cartesian closed implies $\mathbf{B}$ cartesian closed.
5. (a) In $\mathbf{B}M = \mathbf{Sets}^{M^{\mathrm{op}}}$ for M a monoid observe that an object X is a right action $X \times M \to X$ of M on a set X and that, Y being another object, $\mathrm{Hom}(X, Y)$ is the set of equivariant maps $e\colon X \to Y$ [maps with $e(xm) = (ex)m$ for all $x \in X$, $m \in M$]. Prove that the exponent Y^X is the set $\mathrm{Hom}(M \times X, Y)$ of equivariant maps $e\colon M \times X \to Y$, where M is the set M with right action by M, with the action $e \mapsto ek$ of $k \in M$ on e defined by $(ek)(g, x) = e(kg, x)$.

   (b) For objects X, Y in $\mathbf{Sets}^{G^{\mathrm{op}}}$, for G a group, show that the exponent Y^X can be described as the set of all functions $f\colon X \to Y$, with the right action of $g \in G$ on such a function defined by $(fg)x = [f(xg^{-1})]g$ for $x \in X$.

6. Let G be a topological group and $\mathbf{B}G$ the category of continuous G-sets. Let G^δ be the same group G with the discrete topology. So $\mathbf{B}G^\delta = \mathbf{Sets}^{(G^\delta)^{\mathrm{op}}}$ is a category as considered in the previous exercise. Let $i_G\colon \mathbf{B}G \to \mathbf{B}G^\delta$ be the inclusion functor.

   (a) Prove that a G-set $(X, \mu\colon X \times G \to X)$ is in the image of i_G, i.e., that μ is continuous, iff for each $x \in X$ its *isotropy* subgroup

   $$I_x = \{\, g \in G \mid x \cdot g = x \,\}$$

   is an open subgroup of G.

   (b) Prove that, for a G^δ-set (X, μ) as above, the set $r_G(X) = \{\, x \in X \mid I_x \text{ is open} \,\}$ is closed under the action by G, and that r_G defines a functor $\mathbf{B}G^\delta \to \mathbf{B}G$ which is right adjoint to the inclusion functor i_G.

   (c) Observe that i_G preserves products, and conclude from (b) that $\mathbf{B}G$ is cartesian closed since $\mathbf{B}G^\delta$ is. [Hint: define the exponential Y^X in $\mathbf{B}G$ by $r_G(i_G(Y)^{i_G(X)})$.]

7. In Exercise 6, show that the forgetful functor $U\colon \mathbf{B}G \to \mathbf{Sets}$ need not preserve infinite limits.
8. Consider a small category $\mathbf{C}$. For each object B of $\mathbf{C}$ there is a functor $D_B\colon \mathbf{C}/B \to \mathbf{C}$ defined by taking the domain of each arrow to B. Hence, each $T\colon \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$ yields $T_B = T \circ D^{\mathrm{op}}\colon (\mathbf{C}/B)^{\mathrm{op}} \to \mathbf{Sets}$. Define an exponential T^S by

$$T^S(B) = \mathrm{Hom}_{\widehat{(\mathbf{C}/B)}}(S_B, T_B),$$

with the evident evaluations $e_B\colon T^S(B) \times S(B) \to T(B)$. Show that T^S with this evaluation e is indeed the exponential in the functor category $\widehat{\mathbf{C}} = \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$.
9. Let $\mathbf{Q}$ be the (linearly) ordered set of all rational numbers considered as a category, while $\mathbf{R}^+$ is the set of reals with a symbol ∞ adjoined. In $\mathbf{Sets}^{\mathbf{Q}}$, prove that the subobject classifier Ω has $\Omega(q) = \{r \mid r \in \mathbf{R}^+, r \geq q\}$.
10. Generalize Theorem 2 of Section 9 to presheaf categories. More precisely, prove that for a morphism (i.e., a natural transformation) $f\colon Z \to Y$ in $\widehat{\mathbf{C}} = \mathbf{Sets}^{\mathbf{C}^{\mathrm{op}}}$, the pullback functor

$$f^*\colon \mathrm{Sub}_{\widehat{\mathbf{C}}}(Y) \to \mathrm{Sub}_{\widehat{\mathbf{C}}}(Z)$$

has both a left adjoint $\exists_f$ and a right adjoint $\forall_f$. [Hint: the left adjoint can be constructed by taking the pointwise image. Define the right adjoint $\forall_f$ on a subfunctor S of Z by $\forall_f(S)(C) = \{y \in Y(C) \mid$ for all $u\colon D \to C$ in $\mathbf{C}$ and $z \in Z(D)$, $z \in S(D)$ whenever $f_D(z) = yu\}$.]
11. Prove Proposition 5.1, that every functor P to sets is representable, by constructing for each $P\colon \mathbf{C}^{\mathrm{op}} \to \mathbf{Sets}$ a coequalizer

$$\coprod_{\substack{C' \xrightarrow{u} C \\ p \in P(C)}} \mathbf{y}(C') \underset{\tau}{\overset{\theta}{\rightrightarrows}} \coprod_{\substack{C \in \mathbf{C} \\ p \in P(C)}} \mathbf{y}(C) \xrightarrow{\epsilon} P,$$

where $\coprod$ denotes the coproduct and for each object B the maps are defined for each $v\colon B \to C$ or C' as follows

$$\epsilon_B(C, p; v) = P(v)p, \quad \theta_B(u, p; v) = (C, p; uv) \quad \tau_B(u, p; v) = (C', pu; v).$$

(Hint: For each B, this gives a split coequalizer, as defined in [**CWM**, p. 146].)

II
Sheaves of Sets

This chapter starts with the notion of a sheaf F on a topological space X. Such a sheaf is a way of describing a class of functions on X—especially classes of "good" functions, such as the functions on (parts of) X which are continuous or which are differentiable. The description tells the way in which a function f defined on an open subset U of X can be restricted to functions $f|_V$ on open subsets $V \subset U$ and then can be recovered by piecing together (collating) the restrictions to the open subsets V_i of a covering of U. This restriction-collation description applies not just to functions, but also to other mathematical structures defined "locally" on a space X.

Alternatively, a sheaf F on X can be described as a rule which assigns to each point x of the space a set F_x consisting of the "germs" at x of the functions to be considered, as defined in neighborhoods of the point x. The sets F_x for all x can then be "pasted" together by a suitable topology so as to form a space (or bundle) projected onto X; an individual "good" function (for this sheaf) is then a "cross section" of the projection of this bundle. Viewed in this way, the sheaf F is a set F_x which "varies" (with the point x) over the space X.

The letter F is often used for a sheaf because in French the word for "sheaf" is "faisceau".

We will show that the category $\mathrm{Sh}(X)$ of all sheaves of sets on a given space X has all the properties listed in Chapter I for our "typical" categories (i.e., for topoi). Much of the subsequent development of the properties of topoi from the axioms is motivated by geometrical considerations from this case of sheaf theory. This chapter is intended to develop some of the sheaf-theoretic intuition behind this development.

Readers familiar with sheaf theory might wish to skip this chapter; they should then note that we emphasize sheaves of *sets*, and not just those of abelian groups or of modules, and that a sheaf is defined here to be a suitable contravariant functor on open sets, and not the associated (étale) space of the sheaf, as described in §5 below.

1. Sheaves

A topology on a set X serves to define the continuous functions there; for example, the continuous functions from the space X to the reals $\mathbf{R}$, or from any open set U in X to $\mathbf{R}$. The continuity of each $f: U \to \mathbf{R}$ can be determined "locally". This means two things:

(i) If $f: U \to \mathbf{R}$ is continuous and $V \subset U$ is open, then the function f restricted to V is continuous, $f|_V: V \to \mathbf{R}$.

(ii) If U is covered by open sets U_i, and the functions $f_i: U_i \to \mathbf{R}$ are continuous for all $i \in I$, then there is at most one continuous $f: U \to \mathbf{R}$ with restrictions $f|_{U_i} = f_i$ for all i; moreover, such an f exists if and only if the various given f_i "match" on all the overlaps $U_i \cap U_j$, in the sense that $f_i x = f_j x$ for all $x \in U_i \cap U_j$ and all i, j in I.

Property (ii) states that continuous functions are uniquely "collatable".

Many other structures on a space X are "determined locally" in much the same sense. These properties (i) and (ii) can be conveniently expressed in terms of the function C which assigns to each open $U \subset X$ the set of all real-valued continuous functions on U,

$$C(U) = CU = \{ f \mid f: U \to \mathbf{R} \text{ continuous} \}. \tag{1}$$

For $V \subset U$, the operation of (i) restricting each f to the subset V, written as $f \mapsto f|_V$, is a function $CU \to CV$, while if $W \subset V \subset U$ are three nested open sets, restriction is transitive, in that $(f|_V)|_W = f|_W$. These two statements mean that the assignments

$$U \mapsto CU, \quad \{V \subset U\} \mapsto \{CU \to CV \text{ by } f \mapsto f|_V\}$$

define a functor $C: \mathcal{O}(X)^{\mathrm{op}} \to \mathbf{Sets}$. Here $\mathcal{O}(X)$ is the category with objects all open subsets U of X and arrows $V \to U$ the inclusions $V \subset U$. The statement that C is such a functor expresses property (i) above.

As for property (ii) for an open covering $U = \bigcup U_i$, an I-indexed family of functions $f_i: U_i \to \mathbf{R}$, $i \in I$, is an element of the product set $\prod_i CU_i$, while the assignments $\{f_i\} \mapsto \{ f_i|_{U_i \cap U_j} \}$ and $\{f_i\} \mapsto \{ f_j|_{U_i \cap U_j} \}$ define two maps p and q of I-indexed sets to $(I \times I)$-indexed sets, as in the diagram

$$CU \xrightarrow{e} \prod_i CU_i \underset{q}{\overset{p}{\rightrightarrows}} \prod_{i,j} C(U_i \cap U_j). \tag{2}$$

Then property (ii) above states that the map e given by $f \mapsto \{f|_{U_i}\}$ is the equalizer of the maps p and q (i.e., is the universal map e with

$pe = qe$). A sheaf will be defined below to be a functor C such that (2) is an equalizer for all coverings $U = \bigcup U_i$.

We have described this particular C as a sheaf of sets; it is actually a sheaf of algebras over the field $\mathbf{R}$ or a sheaf of $\mathbf{R}$-modules, because each set CU is an algebra over $\mathbf{R}$ under pointwise sum, product, and scalar multiple, while the maps p, q and e of (2) are $\mathbf{R}$-linear morphisms of rings. Hence, in this case, the statement that (2) is an equalizer is equivalent to the statement that the sequence of $\mathbf{R}$-modules

$$0 \longrightarrow CU \xrightarrow{e} \prod_i CU_i \xrightarrow{p-q} \prod_{i,j} C(U_i \cap U_j) \tag{3}$$

is left exact (i.e., that e is the kernel of $p - q$).

There are many other examples of sheaves on a space X; for example, the functor D with each $D(U)$ the set of all functions, continuous or not, on U to $\mathbf{R}$, or $I(U)$, the set of all continuous functions on U to the unit interval I in $\mathbf{R}$. However, the set $B(U)$ of all *bounded* functions on U to $\mathbf{R}$ is a functor of U but not a sheaf, because the collation of functions which are bounded may yield an unbounded function.

For the Euclidean n-space $X = \mathbf{R}^n$ there are a number of examples of sheaves. For U open in $\mathbf{R}^n$ let $C^k U$ be the set of all $f: U \to \mathbf{R}$ which have continuous partial derivatives of all orders up to order k inclusive. Then C^k is a functor $C^k: \mathcal{O}(X)^{\mathrm{op}} \to \mathbf{Sets}$ with values in **Sets** or in $\mathbf{R}$-Mod, and (2)

with C replaced by C^k is again an equalizer because differentiability is local. Thus, each C^k is a sheaf on $\mathbf{R}^n$. This leads to a nested sequence of subsheaves on $\mathbf{R}^n$:

$$C^\infty \subset \cdots \subset C^k \subset C^{k-1} \subset \cdots \subset C^1 \subset C^0 = C.$$

We will regard a sheaf as a functor, that is, as a special kind of a *presheaf*. Here a *presheaf* of sets P on a topological space X is defined to be a functor $P: \mathcal{O}(X)^{\mathrm{op}} \to \mathbf{Sets}$; that is, a presheaf on X is the same thing as a presheaf on the category $\mathcal{O}(X)$, as defined in §I.1. This means that each inclusion $V \subset U$ of open sets in X determines a function

$$P(V \subset U): PU \to PV, \tag{4}$$

which we will often write for each $t \in PU$ as $t \mapsto t|_V$, just as if it were restriction of an actual function t. Moreover, $(t|_V)|_W = t|_W$ whenever $W \subset V \subset U$.

Definition. *A sheaf of sets F on a topological space X is a functor $F: \mathcal{O}(X)^{\mathrm{op}} \to \mathbf{Sets}$ such that each open covering $U = \bigcup_i U_i$, $i \in I$, of an open set U of X yields an equalizer diagram*

$$FU \overset{e}{\dashrightarrow} \prod_i FU_i \underset{q}{\overset{p}{\rightrightarrows}} \prod_{i,j} F(U_i \cap U_j), \tag{5}$$

where for $t \in FU$, $e(t) = \{ t|_{U_i} \mid i \in I \}$ *and for a family* $t_i \in FU_i$,

$$p\{t_i\} = \{t_i|_{(U_i \cap U_j)}\}, \quad q\{t_i\} = \{t_j|_{(U_i \cap U_j)}\}.$$

A morphism $F \to G$ of sheaves is a natural transformation of functors. $\mathrm{Sh}(X)$ will denote the category of all sheaves F of sets on X, with these morphisms as arrows; so, by definition, $\mathrm{Sh}(X)$ is a full subcategory of the functor category $\widehat{\mathcal{O}(X)} = \mathbf{Sets}^{\mathcal{O}(X)^{\mathrm{op}}}$. A *separated presheaf* is a functor F, as above, such that the map e in (5) is injective, (i.e., a monic in **Sets**), though not necessarily the equalizer of p and q. For example, the functor B, with each $B(U)$ the bounded real-valued continuous functions on U, is a separated presheaf but not a sheaf.

Since an arrow into a product is determined by its components (its composites with the projections of the product), the maps e, p, and q of the diagram (2) are the unique maps which make the diagrams below

$$\begin{array}{ccccc} & & FU_i & \xrightarrow{F(U_i \cap U_j \subset U_i)} & F(U_i \cap U_j) \\ & \nearrow & \uparrow & & \uparrow \\ FU & \xrightarrow{e} & \prod_i FU_i & \underset{q}{\overset{p}{\rightrightarrows}} & \prod_{i,j} F(U_i \cap U_j) \\ & \searrow & \downarrow & & \downarrow \\ & & FU_j & \xrightarrow[F(U_i \cap U_j \subset U_j)]{} & F(U_i \cap U_j) \end{array} \tag{6}$$

commute for all i, $j \in I$, where the vertical maps are the (various) projections of the products in question. This categorical description of the equalizer diagram means that our definition applies with **Sets** replaced by other suitable categories, and so defines sheaves $F\colon \mathcal{O}(X)^{\mathrm{op}} \to \mathbf{C}$ of $\mathbf{C}$-objects on a space X, where $\mathbf{C}$ is any category with all small products. The classically useful cases are sheaves of abelian groups, of rings, and of R-modules and R-algebras, for various rings R. Sheaves of modules are important as coefficients for the cohomology of a space.

Note that the definition of a sheaf implies that every sheaf F must send the empty set $\emptyset$ onto a one-point set $\{*\}$. For, in any space X the empty open set $\emptyset$ has an empty cover ($I = \emptyset$); since a product $\prod_i$ over an empty index set I is the one-point set $\{*\}$, the equalizer (5) becomes $F(\emptyset) \to \{*\} \rightrightarrows \{*\}$, so $F(\emptyset) = \{*\}$, as asserted.

A *subsheaf* of a sheaf F on X is defined to be a subfunctor of F which is itself a sheaf. The local character of a sheaf is exhibited by the following description of a subsheaf:

Proposition 1. *If* F *is a sheaf on* X, *then a subfunctor* $S \subset F$ *is a subsheaf if and only if, for every open set* U *and every element* $f \in FU$, *and every open covering* $U = \bigcup U_i$, *one has* $f \in SU$ *if and only if* $f|_{U_i} \in SU_i$ *for all* i.

Proof: The stated condition is clearly necessary for S to be a sheaf. Conversely, consider the commutative diagram

$$\begin{array}{ccccc} SU & \longrightarrow & \prod SU_i & \rightrightarrows & \prod S(U_i \cap U_j) \\ \downarrow & & \downarrow & & \downarrow \\ FU & \longrightarrow & \prod FU_i & \rightrightarrows & \prod F(U_i \cap U_j), \end{array}$$

with vertical maps monic and bottom row an equalizer. The last condition of the proposition states precisely that the left-hand square is a pullback. It follows by a diagram chase that the top row is an equalizer.

If $f\colon X \to Y$ is a continuous map of spaces, then each sheaf F on X yields a sheaf f_*F on Y defined, for V open in Y, by $(f_*F)V = F(f^{-1}V)$; that is, f_*F is defined as the composite functor

$$\mathcal{O}(Y)^{\mathrm{op}} \xrightarrow{f^{-1}} \mathcal{O}(X)^{\mathrm{op}} \xrightarrow{F} \mathbf{Sets}.$$

This sheaf f_*F is called the *direct image* of F under f. The map f_* so defined is clearly a functor

$$f_*\colon \mathrm{Sh}(X) \to \mathrm{Sh}(Y).$$

Also $(fg)_* = f_*g_*$, so the definition $\mathrm{Sh}(f) = f_*$ makes Sh a functor on the category of all small topological spaces. In particular, if $f\colon X \to Y$ is a homeomorphism, f_* gives an isomorphism of categories between sheaves on X and sheaves on Y.

Let U be an open set in the space X. Any sheaf F on X, restricted to open subsets of U, is clearly a sheaf $F|_U$ on U. In this way, $U \mapsto \mathrm{Sh}(U)$ and $U \supset V \mapsto (F|_U \mapsto F|_V)$ define a contravariant functor on $\mathcal{O}(X)$. In fact, since the notion of a sheaf is "local", this functor is itself almost a sheaf:

Theorem 2. *If $X = \bigcup W_k$ is an open covering of the space X, and if, for each k, F_k is a sheaf of sets on W_k such that*

$$F_k|_{(W_k \cap W_\ell)} = F_\ell|_{(W_k \cap W_\ell)} \tag{7}$$

for all indices k and ℓ, then there exists a sheaf F on X, unique up to isomorphism, with isomorphisms $F|_{W_k} \cong F_k$ for all indices k, which match on the equation (7).

Proof: Write $F_{k\ell}$ for the sheaf (7) on $W_k \cap W_\ell$. If the desired sheaf F exists, then for each open U one must have an equalizer

$$FU \longrightarrow \prod_k F_k(U \cap W_k) \rightrightarrows \prod_{k,\ell} F_{k\ell}(U \cap W_k \cap W_\ell). \tag{8}$$

Take this as the definition of each set FU. If $U \supset V$, a comparison with the equalizer definition of FV gives a unique map $FU \to FV$, and, with these maps as the restrictions, F is a functor on $\mathcal{O}(X)^{\mathrm{op}}$, so is a presheaf. To prove it a sheaf, consider any covering U_i of U and construct the commutative 3×3 diagram with first column $FU \to \prod FU_i \rightrightarrows \prod F(U_i \cap U_j)$ and with rows the definitions, like (8), of FU and FU_i. Then all these rows and (F_k being a sheaf) the last two columns are equalizers. A simple diagram chase resembling that for the 3×3 lemma (see [**Mac Lane**, 1963, Lemma XII.3.3]) then proves that the left-hand column is an equalizer, so that F is indeed a sheaf. Uniqueness up to isomorphism is evident from the similar uniqueness of the equalizer (8).

Many explicit sheaves can be constructed from the local pieces F_k, according to the instructions contained in the proof of this theorem. A more liberal version of Theorem 2 is stated in Exercise 8.

Another method of constructing sheaves on a space X is given by Theorem 3 below. Let $\mathcal{B} \subset \mathcal{O}(X)$ be a basis for the topology on X. So for any point $x \in X$ and any open set U containing x, there is a basic open set $B \in \mathcal{B}$ with $x \in B \subset U$. Moreover, we shall assume that $\mathcal{B}$ is closed under finite intersections (but this is not strictly necessary, cf. the Appendix, §4). $\mathcal{B}$ can be viewed as a full subcategory of $\mathcal{O}(X)$, so it makes sense to speak of presheaves on $\mathcal{B}$, i.e., functors $F\colon \mathcal{B}^{\mathrm{op}} \to \mathbf{Sets}$. Such a functor F is called a *sheaf* on $\mathcal{B}$ if for any basic open set $B \in \mathcal{B}$ and any open cover $B = \bigcup_{i \in I} B_i$ of B by basic open sets $B_i \in \mathcal{B}$, the diagram

$$F(B) \rightarrowtail \prod_i F(B_i) \rightrightarrows \prod_{i,j} F(B_i \cap B_j), \tag{9}$$

analogous to (5), is an equalizer diagram. A morphism of sheaves on $\mathcal{B}$ is a natural transformation; so one obtains a category $\mathrm{Sh}(\mathcal{B})$ of sheaves on $\mathcal{B}$. Clearly, any sheaf $F\colon \mathcal{O}(X)^{\mathrm{op}} \to \mathbf{Sets}$ on X restricts to a sheaf on $\mathcal{B}$, and this process defines a functor $\mathbf{r}\colon \mathrm{Sh}(X) \to \mathrm{Sh}(\mathcal{B})$.

Theorem 3. *For a basis $\mathcal{B}$ of the topology on a space X, the restriction functor $\mathbf{r}\colon \mathrm{Sh}(X) \to \mathrm{Sh}(\mathcal{B})$ is an equivalence of categories.*

More informally, this theorem says that a sheaf F on X, or a map $\tau\colon F \to G$ between sheaves on X, may equivalently be defined by specifying the values $F(U)$, or the components τ_U, only for basic open sets U. This theorem is a special case of a result in the Appendix. We leave a direct proof to the reader as Exercise 4.

2. Sieves and Sheaves

On any space X, each open set U determines a presheaf $\mathrm{Hom}(-, U)$ defined, for each open set V, by