Introduction to **Lattices and Order**

This new edition of *Introduction to Lattices and Order* presents a radical reorganization and updating of the content of the successful first (1990) edition. The primary aim of the original – to serve as 'a textbook devoted to ordered sets and lattices and to their contemporary applications' – is unchanged.

The explosive development of theoretical computer science in recent years has, in particular, influenced the book's evolution: a fresh treatment of fixpoint theorems testifies to this and Galois connections now feature prominently. Concept analysis, a methodology for data analysis, has been moved forward, so as to allow an early presentation of both a concrete foundation for the subsequent theory of complete lattices and an application of order theory which is of commercial value in social science.

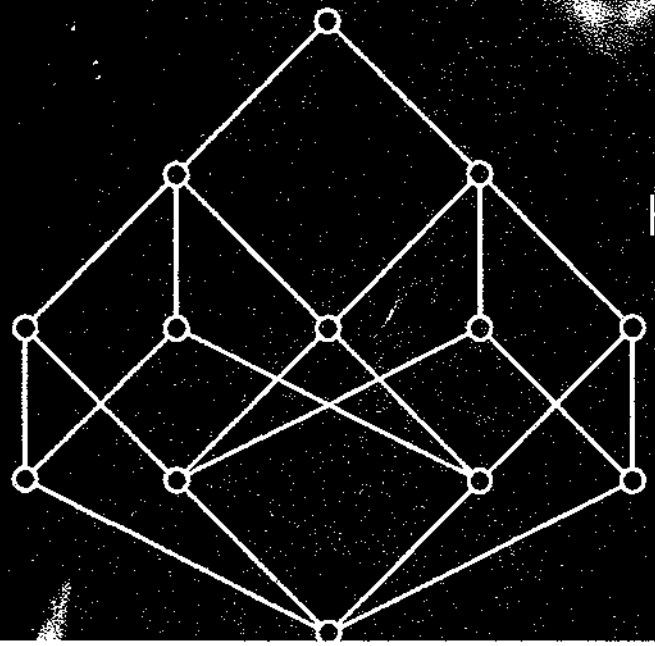
Classroom experience has led to numerous pedagogical improvements and many new exercises have been added. As before, exposure to elementary abstract algebra and the notation of set theory are the only prerequisites. And the level is suitable for advanced undergraduates and first-year graduate students. The book will also be a valuable resource for anyone who, in whatever context, meets ordered structures.

CAMBRIDGE
UNIVERSITY PRESS
www.cambridge.org
978-0-521-78451-

Davey & Priestley Introduction to **Lattices and Order** 2nd edition

Introduction to **Lattices and Order**

Second Edition



B.A. Davey
H.A. Priestley

Introduction to Lattices and Order

Second edition

B. A. Davey
La Trobe University

H. A. Priestley
University of Oxford

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521784511

First edition published 1990
Second edition published 2002
Fourth printing 2008

Printed in the United Kingdom at the University Press, Cambridge

A catalogue record for this publication is available from the British Library

ISBN 978-0-521-78451-1 paperback

Contents

Preface to the second edition

This new edition of *Introduction to Lattices and Order* is substantially different from the original one published in 1990. We believe that the revision greatly enhances the book's usefulness and topicality. Our overall aims however remain the same: to provide a textbook introduction which shows the importance of the concept of order in algebra, logic, computer science and other fields and which makes the basic theory accessible to undergraduate and beginning graduate students in mathematics and to professionals in adjacent areas.

In preparing the new edition we have drawn extensively on our teaching experience over the past 10 years and on helpful comments from colleagues. We have taken account of important developments in areas of application, in particular in computer science. Almost all the original material is included, but it has been completely re-organized. Some new material has been added, most notably on Galois connections and fixpoint calculus, and there are many new exercises.

Our objectives in re-arranging the material have been:

- to present elementary and motivational topics as early as possible, for pedagogical reasons;
- to arrange the chapters so that the first part of the book contains core material, suitable for a short, first course;
- to make it easy for particular interest groups to pick out just the sections they want.

Originally, we treated ordered sets first and began the algebraic theory of lattices only in Chapter 5. This meant that some quite sophisticated and specialized material appeared early on, in particular the treatment of CPOs, algebraic lattices and domains. We have now reversed this, and have also made the treatment of the latter topics more independent. We have moved forward the presentation of formal concept analysis so that it now provides a concrete, application-oriented introduction to complete lattices, to which the material on Galois connections and on completions is later linked. There are numerous more localized repackagings of individual topics too, giving a smoother presentation overall. Readers of the first edition who look at the new table of contents will appreciate how major the re-organization is.

Mathematical modelling in computer science has advanced extremely rapidly in the last decade, and this is reflected in the book. We draw attention in particular to:

- our acknowledgement of the importance of Galois connections in formal methods for program development and verification, and
- the revised presentation of fixpoint theorems.

Our debt to those who have pioneered these advances will be clear from the extent to which we have updated the appendix (now Appendix B) which gives suggestions for further reading. Many colleagues, in particular past and present members of the Oxford University Computing Laboratory, have assisted us, either by the insights their books and papers have provided or through their comments. They are too numerous for us to acknowledge their influence and their contributions individually here.

We are grateful to many readers of the first edition who drew our attention to typographical and other minor errors. They were rewarded with a Mars Bar for each misprint found and their corrections were incorporated into the 1994 printing. For the present edition, our sincere thanks go to a team of proof-readers based at La Trobe University: Jane Pitkethly, aided by Miroslav Haviar, Shamsun Naher and Rashed Talukder. They have done a very careful job of eradicating errors that crept into successive drafts, pinpointing obscurities and spotting a few typos from the first edition that were previously missed.

B.A.D. and H.A.P.

June 2001

Preface to the first edition

This is the first textbook devoted to ordered sets and lattices and to their contemporary applications. It acknowledges the increasingly major role order theory is playing on the mathematical stage and is aimed at students of mathematics and at professionals in adjacent areas, including logic, discrete mathematics and computer science.

Lattice theory has been taught to undergraduates at La Trobe University since 1975, and more recently at Oxford University. The notes for these courses were our starting point. The core of the book – Chapters 1, 2 and 5 to 8 – provides a basic introduction to ordered sets, lattices and Boolean algebras and is buttressed by exercises which have been classroom-tested over many years. In a proselytizing article, *Order: a theory with a view* [in *Klassifikation und Ordnung*, INDEKS, Frankfurt, 1989], Ivan Rival discusses the modern role of order. The pictorial philosophy he advocates is strongly evident in our approach: diagrams and diagrammatic arguments are stressed in both the text and the exercises.

Prerequisites are minimal. A reader who has taken a course in linear algebra, group theory or discrete mathematics should have sufficient background knowledge and be familiar with our vocabulary and with those symbols not listed in the notation index. To keep the treatment as elementary as possible, we have denied ourselves the formalism of category theory and of universal algebra. However, we have prepared the ground carefully for those who will progress to texts on general lattice theory or universal algebra and we have included, at the ends of Chapters 2 and 5 and in Chapters 3, 9 and 10, some material suitable for honours students or those beginning graduate work. Inevitably, there was not space for all the topics we should have liked to cover; hints of resisted temptations will be apparent in a few of the exercises. Within lattice theory we have placed the emphasis on distributive lattices. We thereby complement more advanced texts, in which modular and general lattices are already well treated. The study of finite distributive lattices (undertaken in Chapter 8) combines algebraic, order-theoretic and graph-theoretic ideas to provide results which are linked to the ordered set constructions presented in Chapter 1, are easily accessible to undergraduates and are complete in themselves. Our colleagues will doubtless not be surprised that we have also included the extension of the representation theory to the infinite case. To coax those wary of topology, this introduction to duality is accompanied by a self-contained primer containing the small number of topological results which we need.

Order has recently appeared, sometimes a little coyly, in many computational models. The thorough treatment of ordered sets in Chapter 1 (with examples foreshadowing applications in computation) and of intersection structures in Chapter 2 provides a firm foundation on which to build the theory of CPOs and domains. Chapter 3 studies these structures and relates them to Scott's information systems. Our account is necessarily brief. Collateral reading of specialized texts, in which the computer science applications are fully developed, may assist those meeting domain theory for the first time. Chapter 4 deals with fixpoint theory (and also discusses the order-theoretic roots of Zorn's lemma). Thus Chapters 1–4 serve as an introduction to order theory for computer scientists, and for mathematicians seeking to enter their world. In Chapter 11 we look outwards in a different direction and present the rudiments of formal concept analysis. This new field has already made an impact on lattice theory and has much to offer to social scientists concerned with data analysis. We acknowledge our debt to the authors of many unpublished notes and manuscripts on computer science and on concept analysis. In particular, course notes by Dana Scott, Samson Abramsky and Bill Roscoe enticed us into previously unfamiliar territory and Jeff Sanders' notes for the hardware course taught to Mathematics and Computation undergraduates in Oxford influenced our treatment of Boolean algebras.

The technological developments of the 1980s have made our collaboration possible. Our respective computers have faithfully worked many nocturnal hours of overtime. Electronic mail has enabled us to communicate almost daily and, in conjunction with TEX, to confer easily on fine points of presentation in a way that would have been impossible with conventional ('snail') mail. TEX has also allowed us to control the final shape of the text and has given the second author in particular innumerable hours of fun and frustration.

Many people deserve our thanks. We are grateful to David Tranah and the staff of Cambridge University Press for their patient assistance and support and to Dorothy Berridge for her help in typing TEX files. Generations of students have provided valuable consumer feedback and Oxford undergraduates Mark Joshi, Graham Pollitt and Andy Sanderson earn a special mention for their proof-reading. Thanks are due to the colleagues we have pestered to read the book in draft: in particular to Michael Albert, Ralph McKenzie and J.B. Nation. We must also thank Rudolf Wille and Bernhard Ganter for their advice on concept analysis. Notwithstanding electronic communication, we greatly benefited from the opportunity to spend a month discussing the book face to face. The second author gratefully acknowledges the financial assistance

of La Trobe University and the hospitality of its mathematics department. Finally, a very big thank you for their support and forbearance goes to the Davey family: wife Helen and children Evan, Owen and Caitlin.

B.A.D. and H.A.P.

September 1989

1

Ordered Sets

Order, order, order – it permeates mathematics, and everyday life, to such an extent that we take it for granted. It appears in many guises: first, second, third, ... ; bigger versus smaller; better versus worse. Notions of progression, precedence and preference may all be brought under its umbrella. Our first task is to crystallize these imprecise ideas and to formalize the relationship of 'less-than-or-equal-to'. Besides presenting examples and basic properties of ordered sets, this chapter also introduces the diagrams which make order theory such a pictorial subject and give it much of its character.

Ordered sets

What exactly do we mean by order? More mathematically, what do we mean by an ordered set?

1.1 Order. Each of the following miscellany of statements has something to do with order.

(a) $0 < 1$ and $1 < 10^{23}$.

(b) Two first cousins have a common grandfather.

(c) $22/7$ is a worse approximation to π than 3.141592654.

(d) The planets in order of increasing distance from the sun are Mercury, Venus, Earth, Mars, Jupiter, Saturn, Uranus, Neptune, Pluto.

(e) Neither of the sets $\{1, 2, 4\}$ and $\{2, 3, 5\}$ is a subset of the other, but $\{1, 2, 3, 4, 5\}$ contains both.

(f) Given any two distinct real numbers a and b, either a is greater than b or b is greater than a.

Order is not a property intrinsic to a single object. It concerns comparison between pairs of objects: 0 is smaller than 1; Mars is further from the sun than Earth; a seraphim ranks above an angel, etc. In mathematical terms, an ordering is a binary relation on a set of objects. In our examples, the relation may be taken to be 'less than' on $\mathbb{N}$ in (a), 'is a descendant of' on the set of all human beings in (b) and $\subseteq$ on the subsets of $\{1, 2, 3, 4, 5\}$ (or of $\mathbb{N}$) in (e).

What distinguishes an order relation from some other kind of relation? Firstly, ordering is transitive. From the facts that $0 < 1$ and $1 < 10^{23}$ we can deduce that $0 < 10^{23}$. Mars is nearer the sun than Saturn and Saturn is nearer than Neptune, so Mars is nearer than Neptune.

Secondly, order is antisymmetric: 5 is bigger than 3 but 3 is not bigger than 5. It is on these two properties – transitivity and antisymmetry – that the theory of order rests.

Order relations are of two types: strict and non-strict. Outside mathematics, the strict notion is more common. The statement 'Charles is taller than Bruce' is generally taken to mean 'Charles is strictly taller than Bruce', with the possibility that Charles is the same height as Bruce not included. Mathematicians usually allow equality and write, for instance, $3 \leqslant 3$ and $3 \leqslant 22/7$. We shall deal mainly with non-strict order relations.

Finally a comment about comparability. Statement (f) asserts that, for the ordering $<$ on the real numbers, any two distinct elements can be compared. This property is possessed by many familiar orderings, but it is not universal. For example, there certainly exist human beings A and B such that A is not a descendant of B and B is not a descendant of A. Non-comparability also arises in (e).

1.2 Definitions. Let P be a set. An **order** (or **partial order**) on P is a binary relation $\leqslant$ on P such that, for all $x, y, z \in P$,

(i) $x \leqslant x$,

(ii) $x \leqslant y$ and $y \leqslant x$ imply $x = y$,

(iii) $x \leqslant y$ and $y \leqslant z$ imply $x \leqslant z$.

These conditions are referred to, respectively, as **reflexivity**, **antisymmetry** and **transitivity**. A set P equipped with an order relation $\leqslant$ is said to be an **ordered set** (or **partially ordered set**). Some authors use the shorthand **poset**. Usually we shall be a little slovenly and say simply 'P is an ordered set'. Where it is necessary to specify the order relation overtly we write $\langle P; \leqslant \rangle$. On any set, $=$ is an order, the **discrete order**. A relation $\leqslant$ on a set P which is reflexive and transitive but not necessarily antisymmetric is called a **quasi-order** or, by some authors, a **pre-order**. An order relation $\leqslant$ on P gives rise to a relation $<$ of **strict inequality**: $x < y$ in P if and only if $x \leqslant y$ and $x \neq y$. It is possible to re-state conditions (i)–(iii) above in terms of $<$, and so to regard $<$ rather than $\leqslant$ as the fundamental relation; see Exercise 1.1.

Other notation associated with $\leqslant$ is predictable. We use $x \leqslant y$ and $y \geqslant x$ interchangeably, and write $x \not\leqslant y$ to mean '$x \leqslant y$ is false', and so on. Less familiar is the symbol $\parallel$ used to denote non-comparability: we write $x \parallel y$ if $x \not\leqslant y$ and $y \not\leqslant x$.

We later deal systematically with the construction of new ordered sets from existing ones. However, there is one such construction which it is convenient to have available immediately. Let P be an ordered set and let Q be a subset of P. Then Q inherits an order relation from P; given $x, y \in Q$, $x \leqslant y$ in Q if and only if $x \leqslant y$ in P. We say in these circumstances that Q has the **induced order**, or, when we wish to be more explicit, the order **inherited from** P.

1.3 Chains and antichains. Let P be an ordered set. Then P is a **chain** if, for all $x, y \in P$, either $x \leqslant y$ or $y \leqslant x$ (that is, if any two elements of P are comparable). Alternative names for a chain are **linearly ordered set** and **totally ordered set**. At the opposite extreme from a chain is an antichain. The ordered set P is an **antichain** if $x \leqslant y$ in P only if $x = y$. Clearly, with the induced order, any subset of a chain (an antichain) is a chain (an antichain).

Let P be the n-element set $\{0, 1, \ldots, n-1\}$. We write $\mathbf{n}$ to denote the chain obtained by giving P the order in which $0 < 1 < \cdots < n-1$ and $\overline{\mathbf{n}}$ for P regarded as an antichain. Any set S may be converted into an antichain $\overline{S}$ by giving S the discrete order.

1.4 Order-isomorphisms. We need to be able to recognize when two ordered sets, P and Q, are 'essentially the same'. We say that P and Q are (**order-**)**isomorphic**, and write $P \cong Q$, if there exists a map φ from P onto Q such that $x \leqslant y$ in P if and only if $\varphi(x) \leqslant \varphi(y)$ in Q. Then φ is called an **order-isomorphism**. Such a map φ faithfully mirrors the order structure. It is necessarily bijective (that is, one-to-one and onto): using reflexivity and antisymmetry of $\leqslant$, first in Q and then in P,

$$\begin{aligned} \varphi(x) = \varphi(y) &\iff \varphi(x) \leqslant \varphi(y) \ \&\ \varphi(y) \leqslant \varphi(x) \\ &\iff x \leqslant y \ \&\ y \leqslant x \\ &\iff x = y. \end{aligned}$$

On the other hand, not every bijective map between ordered sets is an order-isomorphism: consider, for example, $P = Q = \mathbf{2}$ and define φ by $\varphi(0) = 1$, $\varphi(1) = 0$.

Being a bijection, an order-isomorphism $\varphi\colon P \to Q$ has a well-defined inverse, $\varphi^{-1}\colon Q \to P$. It is easily seen that this is also an order-isomorphism.

We hinted in 1.1 at a variety of situations in which order is present. In 1.2 we developed the vocabulary for treating these examples systematically. We conclude this section by presenting formally the important orderings carried by some fundamental mathematical structures.

1.5 Number systems. The set $\mathbb{R}$ of real numbers, with its usual order, forms a chain. Each of $\mathbb{N}$ (the natural numbers $\{1, 2, 3, \ldots\}$), $\mathbb{Z}$ (the

integers) and $\mathbb{Q}$ (the rational numbers) also has a natural order making it a chain. In each case this order relation is compatible with the arithmetic structure in the sense that the sum and product of two elements strictly greater than zero is also greater than zero.

We denote the set $\mathbb{N} \cup \{0\}$ $(= \{0, 1, 2, \dots\})$ by $\mathbb{N}_0$. Endowed with the order in which $0 < 1 < 2 < \dots$, the set $\mathbb{N}_0$ becomes the chain known in set theory as ω. It is order-isomorphic to $\mathbb{N}$: the successor function $n \mapsto n^+ := n + 1$ from $\mathbb{N}_0$ to $\mathbb{N}$ is an order-isomorphism. A different order on $\mathbb{N}_0$ is defined as follows. Write $m \preccurlyeq n$ if and only if there exists $k \in \mathbb{N}_0$ such that $km = n$ (that is, m divides n). Then $\preccurlyeq$ is an order relation. Of course, $\langle \mathbb{N}_0; \preccurlyeq \rangle$ is not a chain. Yet another order on $\mathbb{N}_0$ is introduced in 1.22 for use in Chapters 8 and 9.

1.6 Families of sets. Let X be any set. The powerset $\wp(X)$, consisting of all subsets of X, is ordered by set inclusion: for $A, B \in \wp(X)$, we define $A \leqslant B$ if and only if $A \subseteq B$.

Any subset of $\wp(X)$ inherits the inclusion order. Such a family of sets might be specified set-theoretically. For example, it might consist of all finite subsets of an infinite set X. More commonly, families of sets arise where X carries some additional structure. For instance, X might have an algebraic structure – it might be a group, a vector space, or a ring. Each of the following is an ordered set under inclusion:

- the set of all subgroups of a group G (denoted $\operatorname{Sub} G$), and the set of all normal subgroups of G (denoted $\mathcal{N}\text{-}\operatorname{Sub} G$);
- the set of all subspaces of a vector space V (denoted $\operatorname{Sub} V$);
- the set of all subrings of a ring R, and the set of all ideals of R.

Families of sets also occur in other mathematical contexts. For example, let $(X; \mathcal{T})$ be a topological space. We may consider the families of open, closed, and clopen (meaning simultaneously closed and open) subsets of X as ordered sets under inclusion. Finally we note a more inbred member in this class of ordered sets which is of fundamental importance later. This is the family $\mathcal{O}(P)$ of down-sets of an ordered set P; it is introduced in 1.27.

Essentially the same ordered set as $\langle \wp(X); \subseteq \rangle$ manifests itself in a different form, as the set of predicates on X. A **predicate** is a statement taking value T (true) or value F (false). More precisely, a predicate on X is a function from X to $\{\mathsf{T}, \mathsf{F}\}$; here we don't distinguish between different ways of specifying the same function. For example, the map $p\colon \mathbb{R} \to \{\mathsf{T}, \mathsf{F}\}$ given by $p(x) = \mathsf{T}$ if $x \geqslant 0$ and $p(x) = \mathsf{F}$ if $x < 0$ is a predicate on $\mathbb{R}$, which can alternatively be specified by $p(x) = \mathsf{T}$ if $|x-1| \leqslant |x+1|$ and F otherwise. We write $\mathbb{P}(X)$ for the set of predicates on X and order it by implication: for $p, q \in \mathbb{P}(X)$,

$$p \Rrightarrow q \text{ if and only if } \{\, x \in X \mid p(x) = \mathsf{T} \,\} \subseteq \{\, x \in X \mid q(x) = \mathsf{T} \,\}.$$

Define a map $\varphi\colon \mathbb{P}(X) \to \wp(X)$ by $\varphi(p) := \{\, x \in X \mid p(x) = \mathsf{T} \,\}$. Then φ is an order-isomorphism between $\langle \mathbb{P}(X); \Rrightarrow \rangle$ and $\langle \wp(X); \subseteq \rangle$. The notion of a predicate is fundamental in logic and in computer science.

Examples from social science and computer science

Order and ordered structures enter into computer science, and also into social science, in many ways and on many different levels. Our aim in this section is to give a glimpse of *why* this should be so, rather than to explain in detail *how* order theory is employed in applications. This discussion supplies motivation for some of the theory we develop later on, but much of it is not used directly. We look first at ways in which ordered sets arise in social science.

1.7 Ordered sets in the humanities and social sciences. Below is a pot-pourri of examples to indicate how ordered sets occur in the social sciences and elsewhere. Each of these areas of application has led to the investigation of ordered sets of special types.

An **interval order** on a set X is an order relation such that there is a mapping φ of the points of X into subintervals of $\mathbb{R}$ such that, for $x < y$ in X, the right-hand endpoint of $\varphi(x)$ is less than the left-hand endpoint of $\varphi(y)$. Interval orders model, for example, the time spans over which animal species are found or the occurrence of styles of pottery in archaeological strata. A variant on the definition requires all the image intervals to be of the same length, with problems of inexact measurement in mind.

The problem of amalgamating the expressed preferences of a group of individuals to arrive at a consensus is of concern to selection committees, market researchers, psephologists and many others. More explicitly, given m objects and rankings of them by n individuals specified by n chains, how should a chain be constructed which best reflects the individuals' collective preferences? A **social choice function** assigns to any n–tuple of rankings a single ranking which defines a consensus, according to specified criteria. A famous theorem, due to K. Arrow, asserts that there is a set of criteria which are very natural but mutually incompatible. This paradoxical result set off an avalanche of research on social choice theory.

The problem of scheduling a collection of activities or events arises in many different contexts, such as manufacturing and conference planning. Many such problems involve precedence constraints. For example,

certain stages in the assembly of a car must precede others and a conference organizer is likely to have to schedule certain lectures before others. The computational complexity of a scheduling problem depends critically on the order relation which describes the precedence constraints.

Order enters into the classification of objects on two rather different levels. The first is illustrated by our introductory example of the arrangement of the planets into a hierarchical list according to their distance from the sun and by Figure 9.1 which classifies certain ordered sets according to various criteria. On a deeper level, the rather new discipline of **concept analysis** provides a powerful technique for classifying and for analysing complex sets of data. From a set of objects (to take a simple example, the planets) and a set of attributes (for the planets, perhaps large/small, moon/no moon, near sun/far from sun), concept analysis builds an ordered set which reveals inherent hierarchical structure and thence natural groupings and dependencies among the objects and the attributes. Chapter 3 gives a brief introduction to concept analysis.

We now turn to order-theoretic ideas relating to computer science. Our focus in this book is limited to certain aspects of this burgeoning subject in which ordered structures provide useful mathematical models and in this introductory chapter we concentrate on the description of models for some particularly important datatypes. In each case, a relation $\geqslant$ serves to capture the notion of 'is at least as informative as', with the precise interpretation depending on the context. But before presenting examples of such information orderings we need to clarify how computations are to be viewed.

1.8 Programs. Speaking simplistically, a program to perform a computation takes a certain input and, the user hopes, returns a corresponding output. The input and output data may come from many different datatypes, such as natural numbers, strings, lists, sets, and so forth. The term state is used to denote an assignment, to the variables used by a program, of values drawn from the appropriate datatypes. The program **terminates** if it transforms any given state before its execution to a state afterwards; the initial and final states may be regarded as incorporating the input and output data. Frequently, the result of a computation will be generated step by step, with additional information being gained at each stage. Non-termination of a program naturally arises where only partial information towards the solution is output in finite time. A program is **deterministic** if, starting from a given initial state, it will terminate in the same final state each time it is run. Non-determinism can occur where the program's specification allows for more than one valid solution. For example, a program to compute an integer y such that $y^2 = x$ might start in the state $x = 9$ and terminate in either the state $y = 3$ or $y = -3$.

We now give three examples of order relations on datatypes. In 1.12 we look at the features these examples have in common.

1.9 Binary strings. Let Σ^* be the set of all finite binary strings, that is, all finite sequences of zeros and ones; the empty string is included. Adding the infinite sequences, we get the set of all finite or infinite sequences, which we denote by Σ^{**}. We order Σ^{**} by putting $u \leqslant v$ if and only if $u = v$ or u is a finite initial substring (the technical term is **prefix**) of v. Thus, for example, $0100 < 010011$, $010 \parallel 100$ and $10101 < 101010\ldots$ (the infinite string of alternating ones and zeros). Strings may be thought of as information encoded in binary form: the longer the string, the greater the information content. Further, given any string v, we may think of elements u with $u < v$ as providing approximations to v. In particular, any infinite string is, in a sense we shall later need to make precise, the limit of its finite initial substrings. Obviously this example can be generalized by considering strings whose elements are drawn from an arbitrary alphabet of symbols.

1.10 Partial maps. Let X and Y be non-empty sets and $f\colon X \to Y$ a map. Then f may be regarded as a recipe which assigns a member $f(x)$ of Y to each $x \in X$. Alternatively, and equivalently, f is determined by its **graph**, namely $\operatorname{graph} f := \{(x, f(x)) \mid x \in X\}$, a subset of $X \times Y$. If the values of f are given on some subset S of X, we have partial information towards determining f. Formally, we define a **partial map** from X to Y to be a map $\sigma\colon S \to Y$, where $\operatorname{dom}\sigma$, the domain of σ, is a subset S of X; here $S = \varnothing$ is allowed. If $\operatorname{dom}\sigma = X$, then σ is a map (or, for emphasis, a **total map**) from X to Y. The set of partial maps from X to Y is denoted $(X \multimap Y)$; it contains all total maps from X to Y and all partial determinations of them. The elements of $(X \multimap X)$ are called partial maps on X. We order $(X \multimap Y)$ as follows: given $\sigma, \tau \in (X \multimap Y)$, define $\sigma \leqslant \tau$ if and only if $\operatorname{dom}\sigma \subseteq \operatorname{dom}\tau$ and $\sigma(x) = \tau(x)$ for all $x \in \operatorname{dom}\sigma$. Equivalently, $\sigma \leqslant \tau$ if and only if $\operatorname{graph}\sigma \subseteq \operatorname{graph}\tau$ in $\mathcal{P}(X \times Y)$. Note that a subset G of $X \times Y$ is the graph of a partial map if and only if

$$(\forall s \in X)\,(((s, y) \in G \;\&\; (s, y') \in G) \implies y = y').$$

In a (non-terminating) computation to determine a map $f\colon X \to Y$, we may think of f as being built up from tokens of information, each of which is an element σ of $(X \multimap Y)$ with finite domain and which partially specifies f and where $\sigma < f$ in the ordering defined above on

$(X \multimap Y)$. In the other direction, suppose we are given a collection $\mathcal{F}$ of elements of $(X \multimap Y)$. Is there a map f such that we have $\sigma \leqslant f$ for each $\sigma \in \mathcal{F}$? Clearly, the tokens must not supply conflicting messages about the putative f. For example, f cannot exist if $\mathcal{F}$ contains elements σ and τ such that, for some $x \in X$, we have $(x, y) \in \operatorname{graph}\sigma$ and $(x, y') \in \operatorname{graph}\tau$, where $y \neq y'$. We say that a subset $\mathcal{F}$ of $(X \multimap Y)$ is **consistent** if, for any finite subset $\mathcal{G}$ of $\mathcal{F}$, there exists $\rho \in (X \multimap Y)$ (but not necessarily in $\mathcal{F}$) such that $\sigma \leqslant \rho$ for all $\sigma \in \mathcal{G}$. It is easy to see that, so long as $\mathcal{F}$ is a consistent subset of $(X \multimap Y)$, there exists a map $f : X \to Y$ such that $\sigma \leqslant f$ for all $\sigma \in \mathcal{F}$. Consistency is treated in a more general setting in Chapter 9.

Let us now compare two programs P and Q having a common set X of initial states and set Y of final states. Suppose first that they are deterministic but do not necessarily terminate. As above, we may view them as given by partial maps σ_P and σ_Q on X. Assume that $\sigma_P \leqslant \sigma_Q$, so that $\operatorname{dom}\sigma_P \subseteq \operatorname{dom}\sigma_Q$ and $\sigma_P(x) = \sigma_Q(x)$ for all $x \in \operatorname{dom}\sigma_P$. Thus from any input state from which P terminates, Q does too and in the same final state that P does; additionally Q may terminate from initial states from which P fails to do so. Thus Q can achieve everything that P can (more if $\sigma_P < \sigma_Q$, since then Q terminates from at least one state from which P fails to terminate). We write $P \sqsubseteq Q$ if $\sigma_P \leqslant \sigma_Q$. Widening this to (possibly) non-deterministic programs P and Q we say that Q **refines** P, and write $P \sqsubseteq Q$, if 'Q is at least as good as P' in the sense that Q achieves, at least, what P does. A particular situation in which this may arise is the progressive unfolding of a while-loop. Refinement of a non-deterministic program may result in one which is deterministic, or closer to being deterministic. Refinement of a non-terminating program may yield one which terminates more often.

1.11 Intervals in $\mathbb{R}$ and exact real arithmetic. The statement that some computed quantity r equals 1.35 correct to 2 decimal places may be re-expressed as the assertion that r lies in a particular interval in $\mathbb{R}$. We may accordingly treat the collection of all intervals $[\underline{x}, \overline{x}]$ (where $-\infty \leqslant \underline{x} \leqslant \overline{x} \leqslant \infty$) as a set P of approximations to the real numbers, with a smaller interval giving a tighter bound than a larger one and so being more informative. The intervals for which $\underline{x} = \overline{x}$ correspond to exact values. The set P carries a very natural order: for $x = [\underline{x}, \overline{x}]$ and $y = [\underline{y}, \overline{y}]$ define $x \leqslant y$ if and only if $\underline{x} \leqslant \underline{y}$ and $\overline{y} \leqslant \overline{x}$. Then $x \leqslant y$ means that y represents (or contains) at least as much information as x.

Traditionally, a floating-point representation of real numbers has been used in numerical computation. But this has inherent disadvantages: rounding errors are endemic and errors in the input data, resulting from its inexact representation, get propagated. An important goal, therefore, is to implement in a suitable high-level programming language the datatype for $\mathbb{R}$, and the basic arithmetic operations and elementary functions on it, in an efficient way and without rounding errors. (More precisely, the objective is to accomplish this within the framework of effective computation, in the sense of computability theory.) Building on various partially successful attempts, A. Edelat has developed an approach starting from the observation that a real number can be viewed as (being determined by) a shrinking nested sequence of intervals with rational endpoints. The move to rational numbers here is, of course, motivated by the fact that exact calculations can be performed with rationals. A survey of Edelat's work on this and other computational models can be found in [18].

1.12 Information orderings. In each of Examples 1.9–1.11 the order relation captures a notion of 'is more informative than': $x \leqslant y$ has an interpretation such as 'y is more defined than x' or 'y is a better approximation than x'. In each case, we have a notion of a **total object** (a completely defined, or idealized, element). These total objects are the infinite binary strings in the first example, the total maps in the second and the 1-point intervals in the third. An important feature of these examples from a computational point of view is that in each case the total objects may be realized in a natural way as limits of partial objects. Further, in Σ^{**} and in $(\mathbb{N} \multimap \mathbb{N})$, for example, we have approximations by partial objects which are in some sense 'finite': respectively, finite strings or partial maps which have finite domain. In general, a finite object should be one which encodes a finite amount of information.

We conclude our informal introduction to the occurrence of order in computer science with a few general remarks, to widen the perspective and to hint at themes picked up in later chapters.

1.13 Semantics and semantic domains. Running through our discussion of ordered sets in computer science is the idea of a semantic domain: a mathematical structure through which one can describe, analyse and reason about the behaviour of entities such as datatypes, programs and specifications. This use here of 'semantic domain' is generic, and very broad. In Chapter 9, the term 'domain' acquires a narrower, more technical meaning, as an ordered set of a special sort. In a domain, certain elements are to be viewed as partial, or incompletely specified, and each element is required to be the limit (in an appropriate order-theoretic sense) of special elements, designated finite. The ordered structures presented in 1.9 and 1.10 are examples of semantic domains. In fact, both

these are domains in the sense of the formal definition in 9.7. Example 1.11 is a structure of a similar but more general type.

Domains can alternatively be viewed as logical structures. From this perspective, elements of domains are seen as being determined by assertions about them (or propositions they satisfy) and are modelled by sets of tokens of information. Defined in a precise way, either as a special class of ordered sets or, equivalently, as information systems, domains have a mathematical theory worthy of study in its own right, and not just for its computational significance.

In computer science, different styles of semantic modelling are favoured depending on what aspect of computing is being studied. As we hint in our discussion of Galois connections in Chapter 7, a semantics focussing on laws governing the actions of programs may be a useful view. In operational semantics, programs are modelled by the actions they perform on a computer (idealized rather than actual). With denotational semantics, by contrast, the emphasis is on *what* programs do, rather than on *how* a computer executes them. Programs are represented, for example, by functions or relations and can be studied through their representations without the distracting detail needed to describe their implementation.

Chapters 8 and 9 deal with classes of semantic domains rich enough to provide denotational models for complex computational processes, including recursion. The term **recursive** is used of an algorithm defined in terms of itself or a program which calls itself. An example is the specification of the factorial function on $\mathbb{N}_0$ via the recursive formula

$$\mathsf{fact}(k) = \begin{cases} 1 & \text{if } k = 0, \\ k\mathsf{fact}(k-1) & \text{if } k > 0. \end{cases}$$

Recursion is a very powerful computational tool and so a central issue in computer science has been the development of mathematical models of programming which can accommodate it. Specifically, what is required is a denotational framework within which theorems can be proved which assert the existence of computable solutions to (suitable) objects given by recursive definitions. The theory of fixpoints, which we introduce in a mathematical way in Chapter 8, and of domains, have in part been developed with this end in view.

Diagrams: the art of drawing ordered sets

One of the most useful and attractive features of ordered sets is that, in the finite case at least, they can be 'drawn'. To describe how to represent ordered sets diagrammatically, we need the idea of covering.

1.14 The covering relation. Let P be an ordered set and let $x, y \in P$. We say x is **covered by** y (or y **covers** x), and write $x \prec y$ or $y \succ x$, if $x < y$ and $x \leqslant z < y$ implies $z = x$. The latter condition is demanding that there be no element z of P with $x < z < y$.

Observe that, if P is finite, $x < y$ if and only if there exists a finite sequence of covering relations $x = x_0 \prec x_1 \prec \cdots \prec x_n = y$. Thus, in the finite case, the order relation determines, and is determined by, the covering relation.

Here are some simple examples.

- In the chain $\mathbb{N}$, we have $m \prec n$ if and only if $n = m + 1$.
- In $\mathbb{R}$, there are no pairs x, y such that $x \prec y$.
- In $\mathcal{P}(X)$, we have $A \prec B$ if and only if $B = A \cup \{b\}$, for some $b \in X \setminus A$.

1.15 Diagrams. Let P be a finite ordered set. We can represent P by a configuration of circles (representing the elements of P) and interconnecting lines (indicating the covering relation). The construction goes as follows.

(1) To each point $x \in P$, associate a point $p(x)$ of the Euclidean plane $\mathbb{R}^2$, depicted by a small circle with centre at $p(x)$.

(2) For each covering pair $x \prec y$ in P, take a line segment $\ell(x, y)$ joining the circle at $p(x)$ to the circle at $p(y)$.

(3) Carry out (1) and (2) in such a way that

  (a) if $x \prec y$, then $p(x)$ is 'lower' than $p(y)$ (that is, in standard Cartesian coordinates, has a strictly smaller second coordinate),

  (b) the circle at $p(z)$ does not intersect the line segment $\ell(x, y)$ if $z \neq x$ and $z \neq y$.

It is easily proved by induction on the size, $|P|$, of P that (3) can be achieved. A configuration satisfying (1)–(3) is called a **diagram** (or **Hasse diagram**) of P. In the other direction, a diagram may be used to define a finite ordered set; an example is given below. Of course, the same ordered set may have many different diagrams. Diagram-drawing is as much an art as a science, and, as will become increasingly apparent as we proceed, good diagrams can be a real asset to understanding and to theorem-proving.

Figure 1.1(i) shows two alternative diagrams for the ordered set $P = \{a, b, c, d\}$ in which $a < c$, $a < d$, $b < c$ and $b < d$. (When we specify an ordered set by a set of inequalities in this way, it is to be understood that no other pairs of distinct elements are comparable.) In

Figure 1.1(ii) we have drawings which are not legitimate diagrams for P; in the first, (3)(a) in 1.15 is violated and in the second, (3)(b) is.

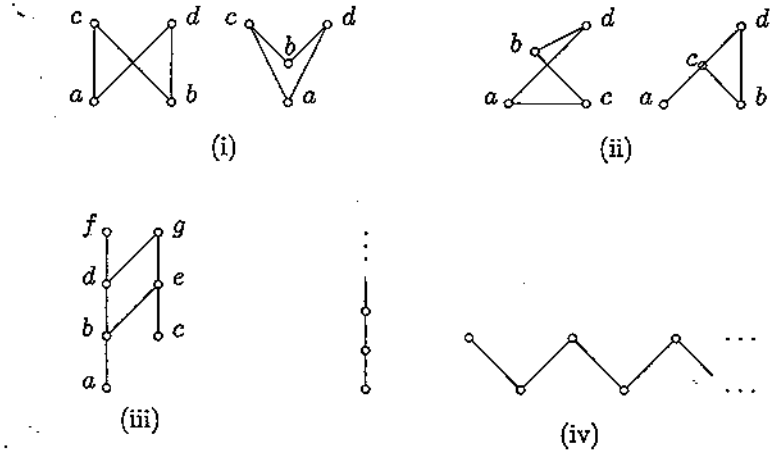


Figure 1.1

It is easy to tell from a diagram whether one element of an ordered set is less than another: $x < y$ if and only if there is a sequence of connected line segments moving upwards from x to y. Thus in the ordered set given in Figure 1.1(iii), $e \parallel f$ and $a < g$, for example.

We have only defined diagrams for finite ordered sets. It is not possible to represent the whole of an infinite ordered set by a diagram, but if its structure is sufficiently regular it can often be suggested diagrammatically, as indicated by the examples in Figure 1.1(iv).

1.16 Examples. Figure 1.2 contains diagrams for a variety of ordered sets. All possible ordered sets with three elements are presented in (i). In (ii) we have diagrams for **2**, **4** and $\overline{\mathbf{3}}$.

Figure 1.2(iii)(a) depicts $\wp(\{1,2,3\})$ (known as the **cube**). A less satisfying, but equally valid, diagram for the same ordered set is shown in Figure 1.2(iii)(b). In Figure 1.2(iv) are diagrams for $\mathrm{Sub}\, G$ for $G = \mathbf{V}_4$, the Klein 4-group, and $G = \mathbf{S}_3$, the symmetric group on 3 letters; in each case the subset $\mathcal{N}\text{-}\mathrm{Sub}\, G$ is shaded.

Figure 1.2(v) gives a diagram for the subset of Σ^* consisting of strings of length not more than 3.

The diagrammatic approach to finite ordered sets is made fully legitimate by Proposition 1.18, which follows easily from Lemma 1.17.

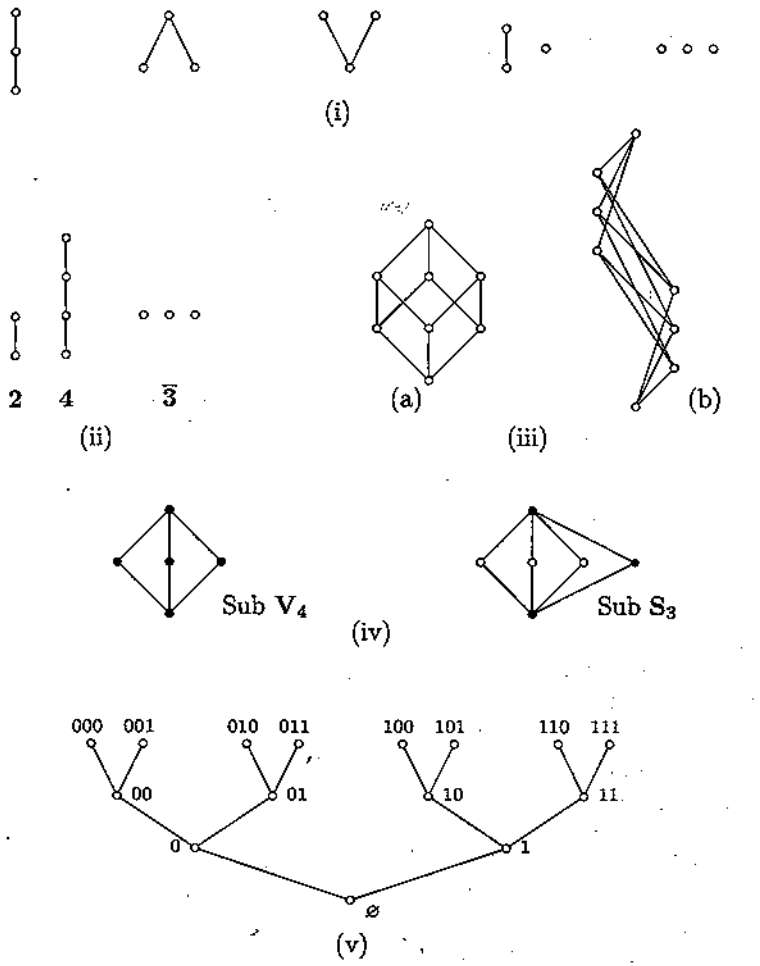


Figure 1.2

1.17 Lemma. *Let P and Q be finite ordered sets and let $\varphi\colon P \to Q$ be a bijective map. Then the following are equivalent:*

(i) *φ is an order-isomorphism;*

(ii) *$x < y$ in P if and only if $\varphi(x) < \varphi(y)$ in Q;*

(iii) *$x \prec y$ in P if and only if $\varphi(x) \prec \varphi(y)$ in Q.*

Proof. The equivalence of (i) and (ii) is immediate from the definitions.

Now assume (ii) holds and take $x \prec y$ in P. Then $x < y$, so $\varphi(x) < \varphi(y)$ in Q. Suppose there exists $w \in Q$ with $\varphi(x) < w < \varphi(y)$.

Since φ is onto, there exists $u \in P$ such that $w = \varphi(u)$. By (ii), $x < u < y$, $\not\downarrow$. Hence $\varphi(x) \prec \varphi(y)$. The reverse implication is proved in much the same way. Hence (iii) holds.

Now assume (iii) and let $x < y$ in P. Then there exist elements $x = x_0 \prec x_1 \prec \cdots \prec x_n = y$. By (iii), $\varphi(x_0) = \varphi(x) \prec \varphi(x_1) \prec \cdots \prec \varphi(x_n) = \varphi(y)$. Hence $\varphi(x) < \varphi(y)$. The reverse implication is proved similarly, using the fact that φ is onto. Hence (ii) holds. □

1.18 Proposition. *Two finite ordered sets P and Q are order-isomorphic if and only if they can be drawn with identical diagrams.*

Proof. Assume there exists an order-isomorphism $\varphi\colon P \to Q$. To show that the same diagram represents both P and Q, note that the diagram is determined by the covering relation and invoke 1.17 (i) ⇒ (iii). Conversely, assume P and Q can both be represented by the same diagram, D. Then there exist bijective maps f and g from P and Q onto the points of D. The composite map $\varphi = g^{-1} \circ f$ is bijective and satisfies condition (iii) in Lemma 1.17, so is an order-isomorphism. □

Constructing and de-constructing ordered sets

This section collects together a number of ways of constructing new ordered sets from existing ones. The other way round, it will often be helpful to analyse ordered sets by regarding them as built up from simpler components. Where we refer to diagrams, it is to be assumed that the ordered sets involved are finite.

1.19 The dual of an ordered set. Given any ordered set P we can form a new ordered set P^∂ (the **dual** of P) by defining $x \leqslant y$ to hold in P^∂ if and only if $y \leqslant x$ holds in P. For P finite, we obtain a diagram for P^∂ simply by 'turning upside down' a diagram for P. Figure 1.3 provides a simple illustration.

To each statement about the ordered set P there corresponds a statement about P^∂. For example, we can assert that in P in Figure 1.3 there exists a unique element covering just three other elements, while in P^∂ there exists a unique element covered by just three other elements. In general, given any statement Φ about ordered sets, we obtain the **dual statement** Φ^∂ by replacing each occurrence of $\leqslant$ by $\geqslant$ and vice versa.

Thus ordered set concepts and results hunt in pairs. This fact can often be used to give two theorems for the price of one or to reduce work (as, for example, in the proof of Theorem 2.9). The formal basis for this observation is the Duality Principle below; its proof is a triviality.

Figure 1.3

1.20 The Duality Principle. *Given a statement Φ about ordered sets which is true in all ordered sets, the dual statement Φ^∂ is also true in all ordered sets.*

We next introduce some important special elements.

1.21 Bottom and top. Let P be an ordered set. We say P has a bottom element if there exists $\bot \in P$ (called **bottom**) with the property that $\bot \leqslant x$ for all $x \in P$. Dually, P has a top element if there exists $\top \in P$ such that $x \leqslant \top$ for all $x \in P$. As a simple instance of the Duality Principle note that the true statement '$\bot$ is unique when it exists' has as its dual version the statement '$\top$ is unique when it exists'. (The uniqueness comes from the antisymmetry of $\leqslant$.)

In $\langle \wp(X); \subseteq \rangle$, we have $\bot = \varnothing$ and $\top = X$. A finite chain always has bottom and top elements, but an infinite chain need not have. For example, the chain $\mathbb{N}$ has bottom element 1, but no top, while the chain $\mathbb{Z}$ of integers possesses neither bottom nor top. Bottom and top do not exist in any antichain with more than one element.

In the context of information orderings, $\bot$ and $\top$ have the following interpretations: $\bot$ represents 'no information', while $\top$ corresponds to an over-determined, or contradictory, element. None of the ordered sets considered in 1.9–1.11 has a top element, except for the ordered set $(X \multimap Y)$ in the very special case that X has just one element. Each has a bottom element: this is the empty string for Σ^{**}, the partial map with empty domain for $(X \multimap Y)$ and $[-\infty, \infty]$ for interval approximations to real numbers. In each case $\bot$ is the least informative element. In modelling computations, a bottom element is also useful for representing and handling non-termination. Accordingly, computer scientists commonly choose models which have bottoms, but prefer them topless.

1.22 Lifting. In Chapters 8 and 9 almost all the results refer to ordered sets with $\bot$. Lack of a bottom element can be easily remedied by adding one. Given an ordered set P (with or without $\bot$), we form $P_\bot$ (called P 'lifted') as follows. Take an element $\mathbf{0} \notin P$ and define $\leqslant$ on $P_\bot := P \cup \{\mathbf{0}\}$

by

$$x \leqslant y \text{ if and only if } x = \mathbf{0} \text{ or } x \leqslant y \text{ in } P.$$

Any set S gives rise to an ordered set with $\perp$, as follows. Order S by making it an antichain, $\overline{S}$, and then form $\overline{S}_\perp$. Ordered sets obtained in this way are called **flat**. In applications it is likely that $S \subseteq \mathbb{R}$. In this context we shall, for simplicity, write $S_\perp$ instead of the more correct $\overline{S}_\perp$. Since we shall not have occasion to apply lifting to subsets of $\mathbb{R}$ ordered as chains, this should cause no confusion. Figure 1.4 shows $\mathbb{N}_\perp$.

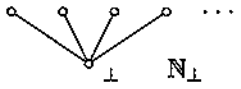


Figure 1.4

1.23 Maximal and minimal elements. Let P be an ordered set and let $Q \subseteq P$. Then $a \in Q$ is a **maximal element of** Q if $a \leqslant x$ and $x \in Q$ imply $a = x$. We denote the set of maximal elements of Q by $\operatorname{Max} Q$. If Q (with the order inherited from P) has a top element, $\top_Q$, then $\operatorname{Max} Q = \{\top_Q\}$; in this case $\top_Q$ is called the **greatest** (or **maximum**) element of Q, and we write $\top_Q = \max Q$. A **minimal** element of $Q \subseteq P$ and $\min Q$, the **least** (or **minimum**) element of Q (when these exist) are defined dually, that is by reversing the order.

Figure 1.5 illustrates the distinction between 'maximal' and 'maximum': P_1 has maximal elements a_1, a_2, a_3, but no greatest element; a is the greatest element of P_2.

Figure 1.5

Let P be a finite ordered set. Then any non-empty subset of P has at least one maximal element and, for each $x \in P$, there exists $y \in \operatorname{Max} P$ with $x \leqslant y$. In general a subset Q of an ordered set P may have many maximal elements, just one, or none. A subset of the chain $\mathbb{N}$ has a maximal element if and only if it is finite and non-empty. In the subset Q of $\wp(\mathbb{N})$ consisting of all subsets of $\mathbb{N}$ other than $\mathbb{N}$ itself, there is no top element, but $\mathbb{N} \setminus \{n\} \in \operatorname{Max} Q$ for each $n \in \mathbb{N}$. The subset of $\wp(\mathbb{N})$ consisting of all finite subsets of $\mathbb{N}$ has no maximal elements. An important set-theorists' tool, known as **Zorn's Lemma**, guarantees the existence of maximal elements, under suitable conditions. Zorn's Lemma is discussed in Chapter 10.

Referring to the examples in 1.9–1.11, we see that the maximal elements in Σ^{**} are the infinite strings, those in $(X \multimap Y)$ are the total maps and the maximal elements in the interval approximations to real numbers are the 1-element intervals. This suggests that when an order relation models information, we might expect a correlation between maximal elements and totally defined elements.

The next two subsections introduce important constructs for building new ordered sets.

1.24 Sums of ordered sets. There are several different ways to join two ordered sets together. In each of these constructions we require that the sets being joined are disjoint (and we shall assume this for the remainder of this subsection). This is no real restriction since we can always find isomorphic copies of the original ordered sets which are disjoint; see Exercise 1.9 for a formal approach to this process.

Suppose that P and Q are (disjoint) ordered sets. The **disjoint union** $P \dot\cup Q$ of P and Q is the ordered set formed by defining $x \leqslant y$ in $P \dot\cup Q$ if and only if either $x, y \in P$ and $x \leqslant y$ in P or $x, y \in Q$ and $x \leqslant y$ in Q. A diagram for $P \dot\cup Q$ is formed by placing side by side diagrams for P and Q.

Again let P and Q be (disjoint) ordered sets. The **linear sum** $P \oplus Q$ is defined by taking the following order relation on $P \cup Q$: $x \leqslant y$ if and only if

$$\begin{aligned} & x, y \in P \text{ and } x \leqslant y \text{ in } P, \\ \text{or} \quad & x, y \in Q \text{ and } x \leqslant y \text{ in } Q, \\ \text{or} \quad & x \in P \text{ and } y \in Q. \end{aligned}$$

A diagram for $P \oplus Q$ (when P and Q are finite) is obtained by placing a diagram for P directly below a diagram for Q and then adding a line segment from *each* maximal element of P to *each* minimal element of Q. The lifting construction is a special case of a linear sum: $P_\perp$ is just $\mathbf{1} \oplus P$. Similarly, $P \oplus \mathbf{1}$ represents P with a (new) top element added.

Each of the operations $\dot\cup$ and $\oplus$ is associative: for (pairwise disjoint) ordered sets P, Q and R,

$$P \dot\cup (Q \dot\cup R) = (P \dot\cup Q) \dot\cup R \text{ and } P \oplus (Q \oplus R) = (P \oplus Q) \oplus R.$$

This allows us to write iterated disjoint unions and linear sums unambiguously without brackets. We denote by $\mathbf{M}_n$ the sum $\mathbf{1} \oplus \overline{\mathbf{n}} \oplus \mathbf{1}$.

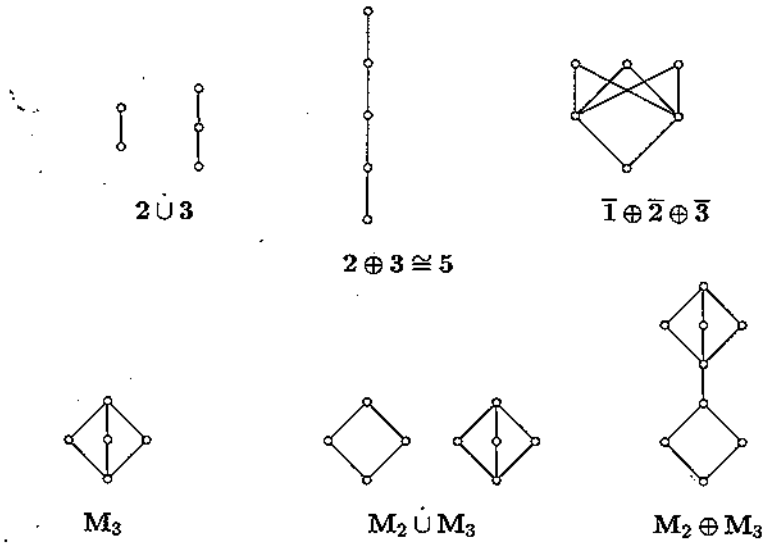


Figure 1.6

Figure 1.6 shows examples of sums. Further notions of sum, tailored to particular applications, are presented in 8.4 and in Exercise 3.12.

1.25 Products. Let $P_1, \dots, P_n$ be ordered sets. The Cartesian product $P_1 \times \cdots \times P_n$ can be made into an ordered set by imposing the coordinatewise order defined by

$$(x_1, \dots, x_n) \leqslant (y_1, \dots, y_n) \iff (\forall i)\, x_i \leqslant y_i \text{ in } P_i.$$

Given an ordered set P, the notation P^n is used as shorthand for the n–fold product $P \times \cdots \times P$.

As an aside we remark that there is another way to order the product of ordered sets P and Q. Define the **lexicographic order** by $(x_1, x_2) \leqslant (y_1, y_2)$ if $x_1 < y_1$ or $(x_1 = y_1$ and $x_2 \leqslant y_2)$. By iteration a lexicographic order can be defined on any finite product of ordered sets. Unless otherwise stated we shall always equip a product with the coordinatewise order.

Informally, a product $P \times Q$ is drawn by replacing each point of a diagram of P by a copy of a diagram for Q, and connecting 'corresponding' points; this assumes that the points are placed in such a way that the rules for diagram-drawing in 1.15 are obeyed. Figure 1.7(i) shows diagrams for some simple products. In Figure 1.7(ii) we depict the four-dimensional hypercube $\mathbf{2}^4$ in various ways. The right-hand representation is obtained by thinking of $\mathbf{2}^4$ as order-isomorphic to $\mathbf{2} \times \mathbf{2}^3$.

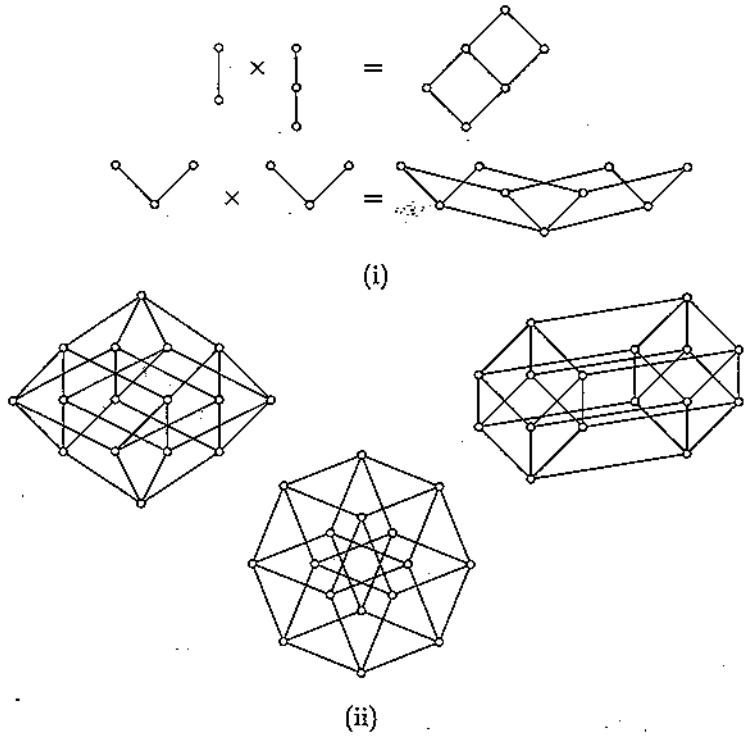


Figure 1.7

The diagram for $\mathbf{2}^3$ is the same as that for $\wp(\{1,2,3\})$ in Figure 1.2(iii). By 1.18, these two ordered sets are isomorphic. We can prove directly a more general result.

1.26 Proposition. *Let $X \doteq \{1, 2, \dots, n\}$ and define $\varphi\colon \wp(X) \to \mathbf{2}^n$ by $\varphi(A) = (\varepsilon_1, \dots, \varepsilon_n)$ where*

$$\varepsilon_i = \begin{cases} 1 & \text{if } i \in A, \\ 0 & \text{if } i \notin A. \end{cases}$$

Then φ is an order-isomorphism.

Proof. Given $A, B \in \wp(X)$, let $\varphi(A) = (\varepsilon_1, \dots, \varepsilon_n)$ and $\varphi(B) = (\delta_1, \dots, \delta_n)$. Then

$$\begin{aligned} A \subseteq B &\iff (\forall i)\, i \in A \text{ implies } i \in B \\ &\iff (\forall i)\, \varepsilon_i = 1 \text{ implies } \delta_i = 1 \\ &\iff (\forall i)\, \varepsilon_i \leqslant \delta_i \\ &\iff \varphi(A) \leqslant \varphi(B) \text{ in } \mathbf{2}^n. \end{aligned}$$

To show φ is onto, take $x = (\varepsilon_1, \dots, \varepsilon_n) \in \mathbf{2}^n$. Then $x = \varphi(A)$, where $A = \{ i \mid \varepsilon_i = 1 \}$, so φ is onto. □

Down-sets and up-sets

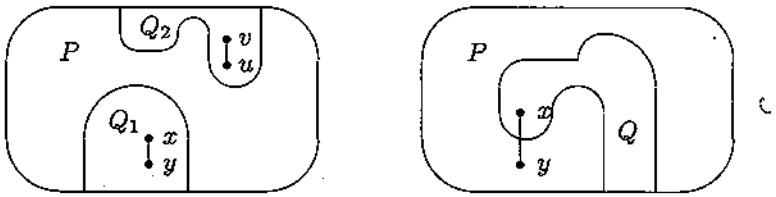
Associated with any ordered set are two important families of sets. They play a central role in the representation theory developed in later chapters.

1.27 Definitions and remarks. Let P be an ordered set and $Q \subseteq P$.

(i) Q is a **down-set** (alternative terms include **decreasing set** and **order ideal**) if, whenever $x \in Q$, $y \in P$ and $y \leqslant x$, we have $y \in Q$.

(ii) Dually, Q is an **up-set** (alternative terms are **increasing set** and **order filter**) if, whenever $x \in Q$, $y \in P$ and $y \geqslant x$, we have $y \in Q$.

It may help to think of a down-set as one which is 'closed under going down'. Down-sets and up-sets may be depicted in a stylized way in a 'directional Venn diagram'; see Figure 1.8. Such drawings do not have the formal status of diagrams, as defined in 1.15.



Q_1 a down-set, Q_2 an up-set — Q not a down-set

Figure 1.8

Given an arbitrary subset Q of P and $x \in P$, we define

$$\downarrow Q := \{ y \in P \mid (\exists x \in Q)\, y \leqslant x \} \text{ and } \uparrow Q := \{ y \in P \mid (\exists x \in Q)\, y \geqslant x \},$$
$$\downarrow x := \{ y \in P \mid y \leqslant x \} \text{ and } \uparrow x := \{ y \in P \mid y \geqslant x \}.$$

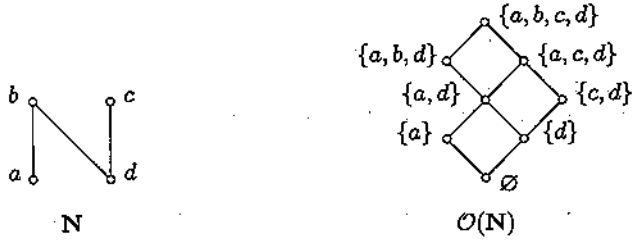
These are read 'down Q', etc. It is easily checked that $\downarrow Q$ is the smallest down-set containing Q and that Q is a down-set if and only if $Q = \downarrow Q$, and dually for $\uparrow Q$. Clearly $\downarrow\{x\} = \downarrow x$, and dually. Down-sets (up-sets) of the form $\downarrow x$ ($\uparrow x$) are called **principal**.

1.28 The ordered set $\mathcal{O}(P)$ of down-sets. The family of all down-sets of P is denoted by $\mathcal{O}(P)$. It is itself an ordered set, under the inclusion order. The letter $\mathcal{O}$ is traditional here; it comes from the term 'order ideal'. When P is finite, every non-empty down-set Q of P is expressible in the form $\bigcup_{i=1}^{k} \downarrow x_i$ (where $\{x_1, \dots, x_k\} = \operatorname{Max} Q$ is an antichain). This provides a recipe for finding $\mathcal{O}(P)$, though one which is practical only when P is small. See also Exercise 1.14, which presents a 'divide and conquer' strategy for calculating $|\mathcal{O}(P)|$ which is more efficient.

1.29 Examples.

(1) Consider the ordered set in Figure 1.1(iii). The sets $\{c\}, \{a,b,c,d,e\}$ and $\{a,b,d,f\}$ are all down-sets. The set $\{b,d,e\}$ is not a down-set; we have $\downarrow\{b,d,e\} = \{a,b,c,d,e\}$. The set $\{e,f,g\}$ is an up-set, but $\{a,b,d,f\}$ is not.

(2) Figure 1.9 shows $\mathcal{O}(P)$ in a simple case.



N — $\mathcal{O}(\mathbf{N})$

Figure 1.9

(3) If P is an antichain, then $\mathcal{O}(P) = \mathcal{P}(P)$.

(4) If P is the chain $\mathbf{n}$, then $\mathcal{O}(P)$ consists of all the sets $\downarrow x$ for $x \in P$, together with the empty set. Hence $\mathcal{O}(P)$ is an $(n+1)$-element chain. If P is the chain $\mathbb{Q}$ of rational numbers, then $\mathcal{O}(P)$ contains the empty set, $\mathbb{Q}$ itself and all sets $\downarrow x$ (for $x \in \mathbb{Q}$). There are other sets in $\mathcal{O}(P)$ too: for example, $\downarrow x \setminus \{x\}$ (for $x \in \mathbb{Q}$) and $\{ y \in \mathbb{Q} \mid y < a \}$ (for $a \in \mathbb{R} \setminus \mathbb{Q}$).

The following handy lemma connects the order relation and down-sets. The proof is an easy but instructive exercise.

1.30 Lemma. *Let P be an ordered set and $x, y \in P$. Then the following are equivalent:*

(i) $x \leqslant y$;

(ii) $\downarrow x \subseteq \downarrow y$;

(iii) $(\forall Q \in \mathcal{O}(P))\, y \in Q \implies x \in Q$.

1.31 $\mathcal{O}(P)$ and duality. Besides being related by duality, down-sets and up-sets are related by complementation: Q is a down-set of P if and only if $P \setminus Q$ is an up-set of P (equivalently, a down-set of P^∂). The proof is left as an exercise. For subsets A, B of P, we have $A \subseteq B$ if and only if $P\setminus A \supseteq P\setminus B$. It follows that

$$\mathcal{O}(P)^\partial \cong \mathcal{O}(P^\partial),$$

the order-isomorphism being the complementation map.

The next proposition shows how $\mathcal{O}(P)$ can be analysed for various compound ordered sets P. Another result of the same sort appears in Exercise 1.18.

1.32 Proposition. *Let P be an ordered set. Then*

(i) $\mathcal{O}(P \oplus \mathbf{1}) \cong \mathcal{O}(P) \oplus \mathbf{1}$ *and* $\mathcal{O}(\mathbf{1} \oplus P) \cong \mathbf{1} \oplus \mathcal{O}(P)$;

(ii) $\mathcal{O}(P_1 \cup P_2) \cong \mathcal{O}(P_1) \times \mathcal{O}(P_2)$.

Proof. (i) The down-sets of $P \oplus \mathbf{1}$ are the down-sets of P together with $P \oplus \mathbf{1}$ itself. The down-sets of $\mathbf{1} \oplus P$ are the empty set and all down-sets of P with the least element of $\mathbf{1} \oplus P$ adjoined. The required isomorphisms are now easily set up.

(ii) It is easily verified that the map $U \mapsto (U \cap P_1, U \cap P_2)$ defines an order-isomorphism from $\mathcal{O}(P_1 \cup P_2)$ to $\mathcal{O}(P_1) \times \mathcal{O}(P_2)$. □

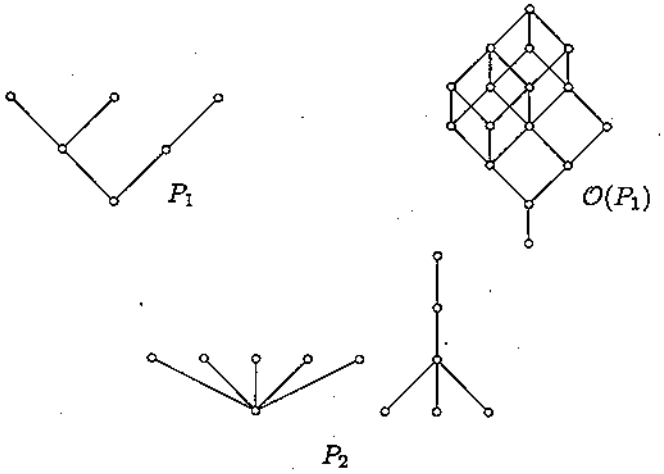


Figure 1.10

1.33 Examples. Consider Figure 1.10.

(1) The ordered set P_1 may be thought of as $\mathbf{1}\oplus((\mathbf{1}\oplus\overline{\mathbf{2}})\cup\mathbf{2})$. By 1.32(i) and (ii), we see that $\mathcal{O}(P_1)$ is isomorphic to $\mathbf{1} \oplus ((\mathbf{1} \oplus \mathbf{2}^2) \times \mathbf{3})$.

(2) Repeated use of Proposition 1.32 gives $\mathcal{O}(P_2) \cong (\mathbf{1}\oplus\mathbf{2}^5) \times (\mathbf{2}^3 \oplus \mathbf{3})$. This ordered set is too complicated to draw effectively, but we know at least that its size is $(1 + 2^5) \times (2^3 + 3) = 363$.

Maps between ordered sets

We have already made use of maps of a very special type between ordered sets, namely order-isomorphisms. In this section, structure-preserving maps are considered more generally.

1.34 Definitions. Let P and Q be ordered sets. A map $\varphi\colon P \to Q$ is said to be

(i) **order-preserving** (or, alternatively, **monotone**) if $x \leqslant y$ in P implies $\varphi(x) \leqslant \varphi(y)$ in Q;

(ii) an **order-embedding** (and we write $\varphi\colon P \hookrightarrow Q$) if $x \leqslant y$ in P if and only if $\varphi(x) \leqslant \varphi(y)$ in Q;

(iii) an **order-isomorphism** if it is an order-embedding which maps P onto Q (recall 1.4).

It is important to appreciate the difference between the notions 'order-preserving map' and 'order-embedding'. The distinction is illustrated by the examples below.

1.35 Examples.

(1) Figure 1.11 shows some maps between ordered sets. The map φ_1 is not order-preserving. Each of φ_2 to φ_5 is order-preserving, but not an order-embedding. The map φ_6 is an order-embedding, but not an order-isomorphism.

(2) Let P be any ordered set. Then, by 1.30, the map $x \mapsto {\downarrow}x$ sets up an order-embedding from P into $\mathcal{O}(P)$.

1.36 Remarks. The following are all easy to prove.

(1) Let $\varphi\colon P \to Q$ and $\psi\colon Q \to R$ be order-preserving maps. Then the composite map $\psi \circ \varphi$, given by $(\psi \circ \varphi)(x) = \psi(\varphi(x))$ for $x \in P$, is order-preserving. More generally the composite of a finite number of order-preserving maps is order-preserving, if it is defined.

(2) Let $\varphi\colon P \hookrightarrow Q$ and let $\varphi(P)$ (defined to be $\{\varphi(x) \mid x \in P\}$) be the image of φ. Then $\varphi(P) \cong P$. This justifies the use of the term embedding.

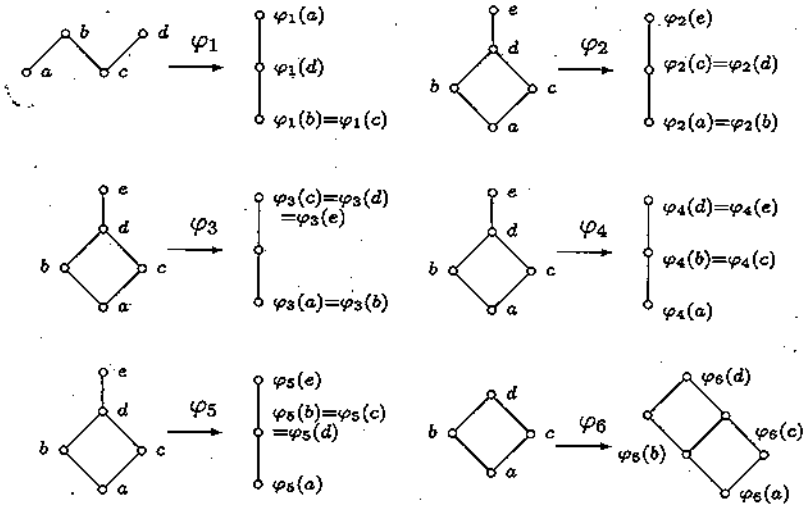


Figure 1.11

(3) An order-embedding is automatically a one-to-one map (by the argument given in 1.4).

(4) Ordered sets P and Q are order-isomorphic if and only if there exist order-preserving maps $\varphi\colon P \to Q$ and $\psi\colon Q \to P$ such that $\varphi \circ \psi = \mathrm{id}_Q$ and $\psi \circ \varphi = \mathrm{id}_P$ (where $\mathrm{id}_S\colon S \to S$ denotes the **identity map** on S given by $\mathrm{id}_S(x) = x$ for all $x \in S$).

1.37 Ordered sets of maps. In elementary analysis an ordering of functions is used and understood often without a formal definition ever being given. Consider, for example, the statement '$\sin x \leqslant |x|$ on $\mathbb{R}$'. The order relation implicit here is the **pointwise order**: for functions $f, g\colon \mathbb{R} \to \mathbb{R}$, the relation $f \leqslant g$ means $f(x) \leqslant g(x)$ for all $x \in \mathbb{R}$.

Pointwise ordering need not be confined to real-valued functions on $\mathbb{R}$. Suppose X is any set and Y an ordered set. We may order the set Y^X of all maps from X to Y as follows. We put $f \leqslant g$ if and only if $f(x) \leqslant g(x)$ in Y, for all $x \in X$. When X is an n-element set, then Y^X is really just Y^n, as defined in 1.25.

Any subset Q of Y^X inherits the pointwise order. When X is itself an ordered set, we may take Q to be the set of all order-preserving maps from X to Y; the resulting ordered set is denoted $Y^{(X)}$. We sometimes write $(X \to Y)$ in place of Y^X and $\langle X \to Y\rangle$ in place of $Y^{(X)}$. This alternative notation is needed because the notation Y^X and $Y^{(X)}$ becomes unwieldy when X or Y is of the form $P_\perp$ or when higher-order functions are involved. (A higher-order function means a function which maps functions to functions; a typical example is the map V in 8.12.) As an example, we note that $\langle X \to 2\rangle \cong O(X)^\partial$; see Exercise 1.25.

1.38 Speaking categorically. In modern pure mathematics it is rare for a class of structures of a given type to be introduced without an associated class of structure-preserving maps following hard on its heels. Ordered sets + order-preserving maps is one example. Others are groups + group homomorphisms, vector spaces over a field + linear maps, topological spaces + continuous maps, and we later meet lattices + lattice homomorphisms, CPOs + continuous maps, etc., etc. The recognition that an appropriate unit for study is a class of objects together with its structure-preserving maps (or **morphisms**) leads to category theory. Informally, a **category** is a class of objects + morphisms, with an operation of composition of morphisms satisfying a set of natural conditions suggested by examples such as those above.

Commuting diagrams of objects and morphisms expressing properties of categories and *their* structure-preserving maps (called **functors**) form the basis of category theory. The representation theory we present in Chapters 5 and 11 is a prime example of a topic which owes its development to the apparatus of category theory. We do not have sufficient need to call on the theory of categories to warrant setting up its formalism here, but it would be wrong not to acknowledge its subliminal influence.

Exercises

Exercises from the text. Verify the unproved claims in 1.27. Prove Lemma 1.30. Prove the unproved assertions in 1.31 and 1.36.

1.1 Let P be a set on which a binary relation $<$ is defined such that, for all $x, y, z \in P$,

(a) $x < x$ is false,

(b) $x < y$ and $y < z$ imply $x < z$.

Prove that if $\leqslant$ is defined by

$$x \leqslant y \iff (x < y \text{ or } x = y),$$

then $\leqslant$ is an order on P, and moreover every order on P arises from a relation $<$ satisfying (a) and (b). (A binary relation satisfying (a) and (b) is called a **strict order**.)

1.2 There is a list of 16 diagrams of four-element ordered sets such that every four-element ordered set can be represented by one of

the diagrams in the list. (That is, up to order-isomorphism, there are just 16 four-element ordered sets.) Find such a list.

1.3 Recall from 1.5 that $\preccurlyeq$ is defined on (any subset of) $\mathbb{N}_0$ by $m \preccurlyeq n$ if and only if m divides n. Draw a diagram for each of the following subsets of $\langle \mathbb{N}_0; \preccurlyeq \rangle$:

(i) $\{1,2,3,5,30\}$, (v) $\{2,3,12,18\}$,
(ii) $\{1,2,3,4,12\}$, (vi) $\{1,2,3,4,6,12\}$,
(iii) $\{1,2,5,10\}$, (vii) $\{1,2,3,12,18,0\}$,
(iv) $\{1,2,4,8,16\}$, (viii) $\{1,2,3,5,6,10,15,30\}$.

1.4 Let $P = \{a,b,c,d,e,f,u,v\}$. Draw the diagram of the ordered set $\langle P; \leqslant \rangle$ where

$$v<a,\ v<b,\ v<c,\ v<d,\ v<e,\ v<f,\ v<u,$$
$$a<c,\ a<d,\ a<e,\ a<f,\ a<u,$$
$$b<c,\ b<d,\ b<e,\ b<f,\ b<u,$$
$$c<d,\ c<e,\ c<f,\ c<u,$$
$$d<e,\ d<f,\ d<u,$$
$$e<u,\ f<u.$$

1.5 Prove that the ordered set Σ^{**} of all binary strings is a tree (that is, an ordered set P with $\bot$ such that $\downarrow x$ is a chain for each $x \in P$). For each $u \in \Sigma^{**}$ describe the set of elements covering u.

1.6 Let P be the set of all finite binary strings (including the empty string). Define $\leqslant$ on P by $u \leqslant v$ if and only if v is a prefix of u or there exist (possibly empty) strings x,y,z such that $v = x0y$ and $u = x1z$. Show that $\leqslant$ is an order on P and that $\langle P; \leqslant \rangle$ is a chain with a $\top$ but no $\bot$. Draw a diagram of the induced order on the seven strings of length less than three and another for the fifteen strings of length less than four.

Let Q be the set of all finite or infinite binary strings with $\leqslant$ defined as for P; then again $\langle Q; \leqslant \rangle$ is a chain. Does Q have a $\bot$? Show that if u is an infinite string then (i) there is no string v such that $u \prec v$, and (ii) there is a string w such that $w \prec u$ if and only if u contains only a finite number of zeros. [Hint. First consider the particular cases $u_1 = 1010101\overline{1}$ and $u_2 = \overline{10}$, where $\overline{x} = xxx\ldots$ for any finite string x.]

1.7 Let P and Q be ordered sets. Prove that $(a_1,b_1) \prec (a_2,b_2)$ in $P \times Q$ if and only if

$$(a_1 = a_2 \ \&\ b_1 \prec b_2) \quad \text{or} \quad (a_1 \prec a_2 \ \&\ b_1 = b_2).$$

Figure 1.12

1.8 Draw the diagrams of the products shown in Figure 1.12.

1.9 Let P and Q be ordered sets with $P \cap Q \neq \varnothing$. Give formal definitions of the ordered sets $P \cup Q$ and $P \oplus Q$. [Hint. Define appropriate orders on $(\{0\} \times P) \cup (\{1\} \times Q)$.]

1.10 Let P and Q be chains. Prove that $P \times Q$ is a chain in the lexicographic order. Prove that $P \times Q$ is a chain in the coordinatewise order if and only if at most one of P and Q has more than one element.

1.11 Let P and Q be ordered sets and equip $P \times Q$ with the lexicographic order. Describe the down-sets of $P \times Q$.

1.12 Let A and B be down-sets of P. Prove that $A \prec B$ in $\langle \mathcal{O}(P); \subseteq \rangle$ if and only if $B = A \cup \{b\}$ for some minimal element b of $P \setminus A$.

1.13 Draw and label a diagram for $\mathcal{O}(P)$ for each of the ordered sets P of Figure 1.13.

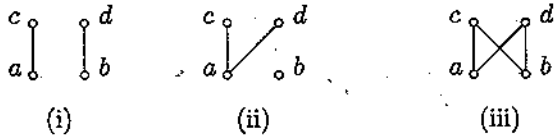


Figure 1.13

1.14 Let P be a finite ordered set.

(i) Show that $Q = \downarrow \mathrm{Max}\, Q$, for all $Q \in \mathcal{O}(P)$.

(ii) Establish a one-to-one correspondence between the elements of $\mathcal{O}(P)$ and antichains in P.

(iii) Hence show that, for all $x \in P$,

$$|\mathcal{O}(P)| = |\mathcal{O}(P \setminus \{x\})| + |\mathcal{O}(P \setminus (\downarrow x \cup \uparrow x))|.$$

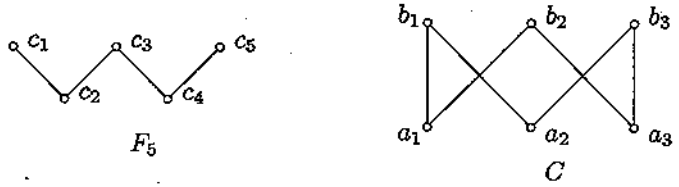


Figure 1.14

1.15 (i) Let F_5 be the 5-element ordered set (a **fence**) shown in Figure 1.14. By applying Exercise 1.14(iii) with $x = c_3$, or otherwise, find the number of down-sets of F_5.

(ii) Let C be the 6-element ordered set (a **crown**) shown in Figure 1.14. Calculate $|\mathcal{O}(C)|$ and deduce the number of elements in $\mathcal{O}((\wp(\{1,2,3\}); \subseteq))$.

1.16 Let $F_n = \{x_1, \ldots, x_n\}$ with $x_1 > x_2 < x_3 > \ldots x_n$ (and no other comparabilities). An ordered set isomorphic to F_n or F_n^∂ is called an n-element **fence**. The Fibonacci sequence $1, 1, 2, 3, 5, 8, \ldots$ is defined by $f_1 = f_2 = 1$ and $f_k = f_{k-2} + f_{k-1}$ for all $k \geqslant 3$. By applying 1.31 and Exercise 1.14 with $x = x_n$ and using induction, or otherwise, show that $|\mathcal{O}(F_n)| = |\mathcal{O}(F_n^\partial)| = f_{n+2}$ for all $n \geqslant 1$.

1.17 Let P be a non-empty ordered set.

(i) Let Q be a subset of P which is both an up-set and a down-set and let $x \neq y$ in P, with $x \in Q$. Assume that there exists a fence in P joining x and y (that is, a fence as defined in Exercise 1.16 with $x = x_1$ and $y = x_n$). Prove that $y \in Q$.

(ii) Prove that the following conditions are equivalent:

(a) the only up-sets of P belonging to $\mathcal{O}(P)$ are $\varnothing$ and P;

(b) P is not the disjoint union of non-empty ordered sets P_1 and P_2;

(c) P is such that, for any two points x, y in P, there is a fence joining x and y (such an ordered set P is called **connected**).

1.18 Let P and Q be ordered sets with $\top$ and $\bot$, respectively. The **vertical sum** $P \overline{\oplus} Q$ of P and Q is obtained from the linear sum $P \oplus Q$ by identifying the top of P with the bottom of Q.

(i) Let P and Q be finite ordered sets. Show that

$$\mathcal{O}(P \oplus Q) \cong \mathcal{O}(P) \overline{\oplus} \mathcal{O}(Q).$$

(ii) Derive Proposition 1.32(i) from (i) above.

1.19 Use the method illustrated in Examples 1.33, along with the result of the previous exercise, to describe $\mathcal{O}(P)$ for each of the ordered sets P in Figure 1.15. Give the cardinality of $\mathcal{O}(P)$ in each case.

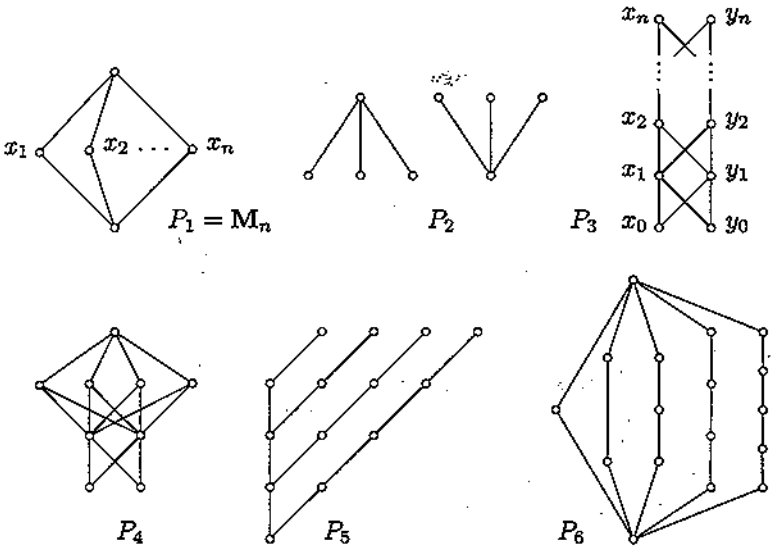


Figure 1.15

1.20 Let P be the 8-element ordered set shown in Figure 1.16.

(i) Explain carefully why $|\mathcal{O}(P)| \neq 31$.

(ii) Find the correct value for $|\mathcal{O}(P)|$.

[Do not find all the down-sets!]

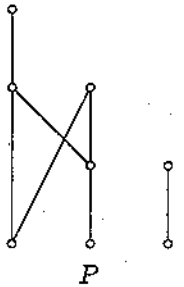


Figure 1.16

1.21 Let X be a topological space satisfying

(T_0) given $x \neq y$ in X there exists either an open set U such that $x \in U$, $y \notin U$ or an open set V such that $x \notin V$, $y \in V$.

Show that $\leqslant$, defined by $x \leqslant y$ if and only if $x \in \overline{\{y\}}$, is an order on X. Describe the down-sets and the up-sets for this order.

1.22 In which of the following cases is the map $\varphi\colon P \to Q$ order-preserving?

(i) $P = Q = \langle \mathbb{Z}; \leqslant \rangle$, and $\varphi(x) = x + 1$.

(ii) $P = \langle \mathcal{P}(S); \subseteq \rangle$ with $|S| > 1$, $Q = \mathbf{2}$, and $\varphi(U) = 1$ if $U \neq \varnothing$ and $\varphi(\varnothing) = 0$.

(iii) $P = \langle \mathcal{P}(S); \subseteq \rangle$ with $|S| > 1$, $Q = \mathbf{2}$, and $\varphi(U) = 1$ if $U = S$ and $\varphi(U) = 0$ if $U \neq S$.

(iv) $P = Q = \langle \mathbb{N}_0; \preccurlyeq \rangle$, and $\varphi(x) = nx$ (with $n \in \mathbb{N}_0$ fixed).

(v) $P = \langle \mathcal{P}(S); \subseteq \rangle$, $Q = \mathbf{2}$, and $\varphi(U) = 1$ if $x \in U$ and $\varphi(U) = 0$ otherwise (with $x \in S$ fixed).

(vi) $P = Q = \langle \mathcal{P}(\mathbb{N}); \subseteq \rangle$, and φ defined by

$$\varphi(U) = \begin{cases} \{1\} & \text{if } 1 \in U, \\ \{2\} & \text{if } 2 \in U \text{ and } 1 \notin U, \\ \varnothing & \text{otherwise.} \end{cases}$$

1.23 Let $\mathbb{N}_0^* := \{\, n \in \mathbb{N} \mid 2 \text{ does not divide } n \,\} \cup \{0\}$.

(i) Define φ from $\langle \mathbb{N}_0; \preccurlyeq \rangle$ to $\langle \mathbb{N}_0^*; \preccurlyeq \rangle$ by $\varphi(0) = 0$ and otherwise $\varphi(n) = n/2^k$, where 2^k is the highest power of 2 which divides n. Show that φ is order-preserving and maps onto $\mathbb{N}_0^*$ but is not an order-isomorphism.

(ii) Show that $\langle \mathbb{N}_0; \preccurlyeq \rangle$ is order-isomorphic to $\langle \mathbb{N}_0^*; \preccurlyeq \rangle$.

1.24 Let P and Q be ordered sets.

(i) Show that $\varphi\colon P \to Q$ is order-preserving if and only if $\varphi^{-1}(A)$ is a down-set in P whenever A is a down-set in Q.

(ii) Assume $\varphi\colon P \to Q$ is order-preserving. Then, by (i), the map $\varphi^{-1}\colon \mathcal{O}(Q) \to \mathcal{O}(P)$ is well defined.

(a) Show that φ is an order-embedding if and only if φ^{-1} maps $\mathcal{O}(Q)$ onto $\mathcal{O}(P)$.

(b) Show that φ maps onto Q if and only if the map $\varphi^{-1}\colon \mathcal{O}(Q) \to \mathcal{O}(P)$ is one-to-one.

1.25 (i) Prove that a subset U of an ordered set P is an up-set if and only if its characteristic function $\chi_U\colon P \to \mathbf{2}$ is order-preserving. (Here $\chi_U(x) = 1$ if $x \in U$ and $\chi_U(x) = 0$ if $x \notin U$.) Show further that $\mathcal{O}(P) \cong \langle P \to \mathbf{2} \rangle^\partial$.

(ii) Prove that $\langle P \to \mathbf{2} \rangle^\partial \cong \langle P^\partial \to \mathbf{2} \rangle$.

1.26 Prove that, for all ordered sets P, Q and R,

$$\langle P \to \langle Q \to R \rangle \rangle \cong \langle P \times Q \to R \rangle.$$

1.27 Let X be any set and let Y be an ordered set. Show that $f \prec g$ in Y^X if and only if there exists $x_0 \in X$ such that

(a) $f(x) = g(x)$ for all $x \in X \setminus \{x_0\}$,

(b) $f(x_0) \prec g(x_0)$.

Assume that X is a finite ordered set. Show that $f \prec g$ in $Y^{(X)}$ if and only if there exists $x_0 \in X$ such that (a) and (b) hold. [Hint. The 'if' direction is straightforward. To prove the 'only if' direction, argue by contradiction. Assume that $f \prec g$ and suppose that f and g differ at more than one element of X. Choose $y \in X$ minimal with respect to $f(y) < g(y)$ and let $z \in X \setminus \{y\}$ satisfy $f(z) < g(z)$. Define $h\colon X \to Y$ by

$$h(x) = \begin{cases} f(x) & \text{if } x = y, \\ g(x) & \text{if } x \neq y. \end{cases}$$

Use the minimality of y to prove that h is order-preserving and then show that $f < h < g$. Finally, show that if $f \prec g$ and $f(x_0) < g(x_0)$, then $f(x_0) \prec g(x_0)$.]

1.28 Draw diagrams for $P^{(\mathbf{2})}$ and $P^{(\mathbf{3})}$ where P is $\mathbf{2}$, $\mathbf{3}$ or $V := \mathbf{1} \oplus (\mathbf{1} \cup \mathbf{1})$. Label the elements—an element of $P^{(\mathbf{2})}$ may be labelled xy where $x \leqslant y$ in P and similarly an element of $P^{(\mathbf{3})}$ may be labelled xyz where $x \leqslant y \leqslant z$ in P.

1.29 Let $\rho \subseteq P \times P$ be a binary relation on a set P. Then the **transitive closure**, ρ^t, of ρ is defined by $a \,\rho^t\, b$ if and only if

$$(\exists n \in \mathbb{N})(\exists z_0, z_1, \ldots, z_n \in P)\, a = z_0 \,\rho\, z_1 \,\rho\, z_2 \,\rho \cdots \rho\, z_{n-1} \,\rho\, z_n = b.$$

(i) Let $\leqslant$ be an order on P and assume that $a \parallel b$ for some $a, b \in P$. Show that ρ^t is an order on P where

$$\rho := \{\, (x, y) \mid x \leqslant y \,\} \cup \{(a, b)\}.$$

[Hint. While this can be done directly, it is easier to work with the corresponding strict order, $<$, and show that the transitive closure of

$$\{(x,y) \mid x < y\} \cup \{(a,b)\}$$

is also a strict order. See Exercise 1.1.]

(ii) Use (i) to show that if P is finite then every order $\leqslant$ on P has a **linear extension**, that is, there is an order $\leqslant_1$ such that $\langle P; \leqslant_1 \rangle$ is a chain and for all $a, b \in P$ we have $a \leqslant b$ implies $a \leqslant_1 b$. (The infinite case will be proved in Exercise 10.2.)

(iii) Use (i) to show that if $\langle P; \leqslant \rangle$ is a finite ordered set, then there is a finite number of chains $\langle P; \leqslant_1 \rangle, \ldots, \langle P; \leqslant_n \rangle$ such that for all $a, b \in P$ we have

$$a \leqslant b \iff (a \leqslant_1 b \,\&\, \cdots \,\&\, a \leqslant_n b).$$

(iv) Show that $\leqslant_1$ is a linear extension of the order $\leqslant$ on P if and only if $\langle P; \leqslant_1 \rangle$ is a chain and the identity map $\mathrm{id}\colon P \to P$ is an order-preserving map from $\langle P; \leqslant \rangle$ to $\langle P; \leqslant_1 \rangle$.

(v) Draw and label a diagram for every possible linear extension of the ordered set **N** given in Figure 1.9.

1.30 Let P be a finite ordered set. The **width** of P is defined to be the size of the largest antichain in P and is denoted by $w(P)$.

(i) Find $w(P)$ for each of the ordered sets P in Figure 1.13 and show, in each case, that P can be written as a union of $w(P)$ many chains.

(ii) Show that if a finite ordered set P can be written as the union of n chains, then $n \geqslant w(P)$.

(iii) **Dilworth's Theorem** states that the width $w(P)$ of a finite ordered set P equals the least $n \in \mathbb{N}$ such that P can be written as a union of n chains. The more intrepid may try to find their own proof of this important result. Alternatively, a much easier, but still valuable, exercise is to rewrite the snappy 14-line proof given in the paper by H. Tverberg in *J. Combinatorial Theory* **3** (1967), pp. 305–306, explaining every step in detail.

2

Lattices and Complete Lattices

Many important properties of an ordered set P are expressed in terms of the existence of certain upper bounds or lower bounds of subsets of P. Two of the most important classes of ordered sets defined in this way are lattices and complete lattices. Here we present the basic theory of such ordered sets, and also consider lattices as algebraic structures in a way that is reminiscent of the study of, for example, groups or rings.

Lattices as ordered sets

It is a fundamental property of the real numbers, $\mathbb{R}$, that if I is a closed and bounded interval in $\mathbb{R}$, then every subset of I has both a least upper bound (or supremum) and a greatest lower bound (or infimum) in I. These concepts pertain to any ordered set.

2.1 Definitions. Let P be an ordered set and let $S \subseteq P$. An element $x \in P$ is an **upper bound** of S if $s \leqslant x$ for all $s \in S$. A **lower bound** is defined dually. The set of all upper bounds of S is denoted by S^u (read as 'S **upper**') and the set of all lower bounds by S^ℓ (read as 'S **lower**'):

$$S^u := \{x \in P \mid (\forall s \in S)\, s \leqslant x\} \text{ and } S^\ell := \{x \in P \mid (\forall s \in S)\, s \geqslant x\}.$$

Since $\leqslant$ is transitive, S^u is always an up-set and S^ℓ a down-set. If S^u has a least element x, then x is called the **least upper bound** of S. Equivalently, x is the least upper bound of S if

(i) x is an upper bound of S, and

(ii) $x \leqslant y$ for all upper bounds y of S.

The least upper bound of S exists if and only if there exists $x \in P$ such that

$$(\forall y \in P)\,[((\forall s \in S)\, s \leqslant y) \iff x \leqslant y],$$

and this characterizes the least upper bound of S. This way of presenting the definition is slicker, but is less transparent until the two-step version has been fully mastered.

Dually, if S^ℓ has a greatest element, x, then x is called the **greatest lower bound** of S. Since least elements and greatest elements are unique (see 1.23), least upper bounds and greatest lower bounds are unique when they exist. The least upper bound of S is also called the **supremum** of S

and is denoted by $\sup S$; the greatest lower bound of S is also called the **infimum** of S and is denoted by $\inf S$.

2.2 Top and bottom. In the definitions of $\sup S$ and $\inf S$ the two extreme cases, where S is empty or S is P itself, warrant a brief investigation. Recall from 1.21 that, when they exist, the top and bottom elements of P are denoted by $\top$ and $\bot$ respectively. It is easily seen that if P has a top element, then $P^u = \{\top\}$ in which case $\sup P = \top$. When P has no top element, we have $P^u = \varnothing$ and hence $\sup P$ does not exist. By duality, $\inf P = \bot$ whenever P has a bottom element. Now let S be the empty subset of P. Then every element $x \in P$ satisfies (vacuously) $s \leqslant x$ for all $s \in S$. Thus $\varnothing^u = P$ and hence $\sup \varnothing$ exists if and only if P has a bottom element, and in that case $\sup \varnothing = \bot$. Dually, $\inf \varnothing = \top$ whenever P has a top element.

2.3 Notation. Looking ahead, we shall adopt the following neater notation: we write $x \vee y$ (read as 'x **join** y') in place of $\sup\{x, y\}$ when it exists and $x \wedge y$ (read as 'x **meet** y') in place of $\inf\{x, y\}$ when it exists. Similarly we write $\bigvee S$ (the '**join of** S') and $\bigwedge S$ (the '**meet of** S') instead of $\sup S$ and $\inf S$ when these exist. It is sometimes necessary to indicate that the join or meet is being found in a particular ordered set P, in which case we write $\bigvee_P S$ or $\bigwedge_P S$. We shall often encounter the join or meet of a set S of the form $S = \{A_i\}_{i \in I}$ where I is some indexing set. We write, for example, $\bigvee_{i \in I} A_i$, as this is neater than the strictly correct notation $\bigvee\{A_i \mid i \in I\}$ for $\bigvee S$. No confusion with the usage $\bigvee_P S$ is likely to arise.

We shall be particularly interested in ordered sets in which $x \vee y$ and $x \wedge y$ exist for all $x, y \in P$.

2.4 Definitions. Let P be a non-empty ordered set.

(i) If $x \vee y$ and $x \wedge y$ exist for all $x, y \in P$, then P is called a **lattice**.

(ii) If $\bigvee S$ and $\bigwedge S$ exist for all $S \subseteq P$, then P is called a **complete lattice**.

In what follows we focus first on lattices, deferring a systematic study of complete lattices until later.

2.5 Remarks on $\vee$ and $\wedge$.

(1) Let P be any ordered set. If $x, y \in P$ and $x \leqslant y$, then $\{x, y\}^u = \uparrow y$ and $\{x, y\}^\ell = \downarrow x$. Since the least element of $\uparrow y$ is y and the greatest element of $\downarrow x$ is x, we have $x \vee y = y$ and $x \wedge y = x$ whenever $x \leqslant y$. In particular, since $\leqslant$ is reflexive, we have $x \vee x = x$ and $x \wedge x = x$.

(2) In an ordered set P, the least upper bound $x \vee y$ of $\{x, y\}$ may fail to exist for two different reasons:

(a) because x and y have no common upper bound, or

(b) because they have no *least* upper bound.

In Figure 2.1(i) we have $\{a, b\}^u = \varnothing$ and hence $a \vee b$ does not exist. In (ii) we find that $\{a, b\}^u = \{c, d\}$ and thus $a \vee b$ does not exist as $\{a, b\}^u$ has no least element.

(i) (ii)

Figure 2.1

(3) Consider the ordered set drawn in Figure 2.2. It is tempting, at first sight, to think that $b \vee c = i$. On more careful inspection we find that $\{b, c\}^u = \{\top, h, i\}$. Since $\{b, c\}^u$ has distinct minimal elements, namely h and i, it cannot have a least element; hence $b \vee c$ does not exist. On the other hand, $\{a, b\}^u = \{\top, h, i, f\}$ has a least element, namely f, and thus $a \vee b = f$.

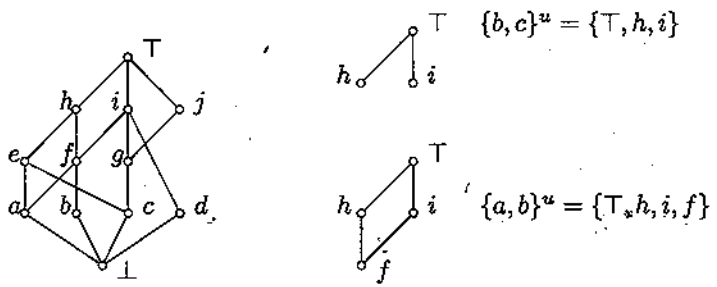


Figure 2.2

(4) Let P be a lattice. Then for all $a, b, c, d \in P$,

(i) $a \leqslant b$ implies $a \vee c \leqslant b \vee c$ and $a \wedge c \leqslant b \wedge c$,

(ii) $a \leqslant b$ and $c \leqslant d$ imply $a \vee c \leqslant b \vee d$ and $a \wedge c \leqslant b \wedge d$.

We leave the proofs as an exercise.

(5) Let P be a lattice. Let $a, b, c \in P$ and assume that $b \leqslant a \leqslant b \vee c$. Since $c \leqslant b \vee c$, we have $(b \vee c) \vee c = b \vee c$, by (1). Thus, by (4)(i),

$$b \vee c \leqslant a \vee c \leqslant (b \vee c) \vee c = b \vee c,$$

whence $a \vee c = b \vee c$; see Figure 2.3. This simple observation and its dual are particularly useful when calculating joins and meets on a diagram – once we know the join of b and c, the join of c with the intermediate element a is forced.

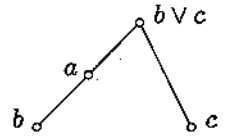


Figure 2.3

2.6 Examples.

(1) Let P be a non-empty ordered set. By Remark 2.5(1), if $x \leqslant y$ then $x \vee y = y$ and $x \wedge y = x$. Hence to show that P is a lattice, it suffices to prove that $x \vee y$ and $x \wedge y$ exist in P for all non-comparable pairs $x, y \in P$. In particular, every chain is a lattice in which $x \vee y = \max\{x, y\}$ and $x \wedge y = \min\{x, y\}$. Thus each of $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{Z}$ and $\mathbb{N}$ is a lattice under its usual order. None of them is complete; every one lacks a top element, and a complete lattice must have top and bottom elements (see 2.2). However, if $x < y$ in $\mathbb{R}$, then the closed interval $[x, y]$ is a complete lattice (by the completeness axiom for $\mathbb{R}$). Failure of completeness in $\mathbb{Q}$ is more fundamental than in $\mathbb{R}$. In $\mathbb{Q}$, it is not merely the lack of top and bottom elements which causes problems; for example, the set $\{s \in \mathbb{Q} \mid s^2 < 2\}$ has upper bounds but no least upper bound in $\mathbb{Q}$.

(2) For any set X, the ordered set $\langle \mathcal{P}(X); \subseteq \rangle$ is a complete lattice in which

$$\bigvee\{A_i \mid i \in I\} = \bigcup\{A_i \mid i \in I\},$$
$$\bigwedge\{A_i \mid i \in I\} = \bigcap\{A_i \mid i \in I\}.$$

As with $\bigvee$ and $\bigwedge$, we shall henceforth indicate the index set by subscripting and, for example, instead of $\bigcup\{A_i \mid i \in I\}$ we shall write $\bigcup_{i \in I} A_i$ or occasionally, when no confusion is likely, simply $\bigcup A_i$. We verify the assertion about meets; that about joins is proved dually. Let $\{A_i\}_{i \in I}$ be a family of elements of $\mathcal{P}(X)$. Since $\bigcap_{i \in I} A_i \subseteq A_j$ for all $j \in I$, it follows that $\bigcap_{i \in I} A_i$ is a lower bound for $\{A_i\}_{i \in I}$. Also, if $B \in \mathcal{P}(X)$ is a lower bound of $\{A_i\}_{i \in I}$, then $B \subseteq A_i$ for all $i \in I$ and hence $B \subseteq \bigcap_{i \in I} A_i$. Thus $\bigcap_{i \in I} A_i$ is indeed the greatest lower bound of $\{A_i\}_{i \in I}$ in $\mathcal{P}(X)$.

(3) Let $\varnothing \neq \mathcal{L} \subseteq \mathcal{P}(X)$. Then $\mathcal{L}$ is known as a **lattice of sets** if it is closed under finite unions and intersections and a **complete lattice** of sets if it is closed under arbitrary unions and intersections. If $\mathcal{L}$ is a lattice of sets, then $\langle \mathcal{L}; \subseteq \rangle$ is a lattice in which $A \vee B = A \cup B$ and $A \wedge B = A \cap B$. Similarly, if $\mathcal{L}$ is a complete lattice of sets, then $\langle \mathcal{L}; \subseteq \rangle$ is a complete lattice with join given by set union and meet given by set intersection. Further details will be given in Lemma 2.28 and Corollary 2.29.

Let P be an ordered set and consider the ordered set $\mathcal{O}(P)$ of all down-sets of P introduced in 1.27. If $\{A_i\}_{i \in I} \subseteq \mathcal{O}(P)$, then $\bigcup_{i \in I} A_i$ and $\bigcap_{i \in I} A_i$ both belong to $\mathcal{O}(P)$. Hence $\mathcal{O}(P)$ is a complete lattice of sets, called the **down-set lattice** of P.

(4) The ordered set $\mathbf{M}_n$ (for $n \geqslant 1$) introduced in Chapter 1 (see Figure 2.4) is easily seen to be a lattice. Let $x, y \in \mathbf{M}_n$ with $x \parallel y$. Then x and y are in the central antichain of $\mathbf{M}_n$ and hence $x \vee y = \top$ and $x \wedge y = \bot$.

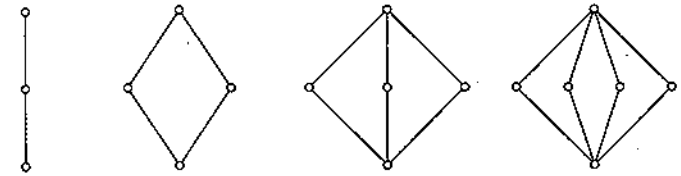


Figure 2.4

(5) Consider the ordered set $\langle \mathbb{N}_0; \preccurlyeq \rangle$ of non-negative integers ordered by division (see 1.5). Recall that k is the greatest common divisor (or highest common factor) of m and n if

(a) k divides both m and n (that is, $k \preccurlyeq m$ and $k \preccurlyeq n$),

(b) if j divides both m and n, then j divides k (that is, $j \preccurlyeq k$ for all lower bounds j of $\{m, n\}$).

Thus the greatest common divisor of m and n is precisely the meet of m and n in $\langle \mathbb{N}_0; \preccurlyeq \rangle$. Dually, the join of m and n in $\langle \mathbb{N}_0; \preccurlyeq \rangle$ is given by their least common multiple. You should convince yourself that these statements remain valid when m or n equals 0. Thus $\langle \mathbb{N}_0; \preccurlyeq \rangle$ is a lattice in which

$$m \vee n = \operatorname{lcm}\{m, n\} \quad \text{and} \quad m \wedge n = \gcd\{m, n\}.$$

Exercise 2.36 indicates two proofs that this lattice is complete.

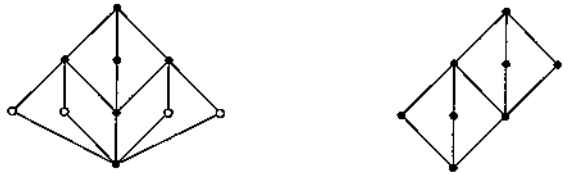
2.7 Lattices of subgroups. Assume that G is a group and $\langle \operatorname{Sub} G; \subseteq \rangle$ is its ordered set of subgroups. Let $H, K \in \operatorname{Sub} G$. It is certainly true that $H \cap K \in \operatorname{Sub} G$, whence $H \wedge K$ exists and equals $H \cap K$. Tiresomely, $H \cup K$ is only exceptionally a subgroup. Nevertheless, $H \vee K$ does exist in $\operatorname{Sub} G$, as (rather tautologically) the subgroup $\langle H \cup K \rangle$ generated by $H \cup K$. Unfortunately there is no convenient general formula for $H \vee K$. This lopsided behaviour is analysed more closely in the next section, where we consider arbitrary joins and meets in $\operatorname{Sub} G$.

Normal subgroups are more amenable. Meet is again given by $\cap$ and join in $\mathcal{N}\text{-}\operatorname{Sub} G$ has a particularly compact description. It is a straightforward exercise in group theory to show that if H, K are normal subgroups of G, then

$$HK := \{ hk \mid h \in H,\ k \in K \}$$

is also a normal subgroup of G. It follows easily that the join in $\mathcal{N}\text{-}\operatorname{Sub} G$ is given by $H \vee K = HK$.

The lattices $\operatorname{Sub} G$ and $\mathcal{N}\text{-}\operatorname{Sub} G$ are given in Figure 2.5 for the group, $\mathbf{D}_4$, of symmetries of a square and for the group $\mathbb{Z}_2 \times \mathbb{Z}_4$. The elements of $\mathcal{N}\text{-}\operatorname{Sub} G$ are shaded. (You should convince yourself, from the diagrams, that these ordered sets are indeed lattices.)



$\operatorname{Sub} \mathbf{D}_4$ with $\mathcal{N}\text{-}\operatorname{Sub} \mathbf{D}_4$ shaded $\qquad$ $\operatorname{Sub} \mathbb{Z}_2 \times \mathbb{Z}_4 = \mathcal{N}\text{-}\operatorname{Sub} \mathbb{Z}_2 \times \mathbb{Z}_4$

Figure 2.5

We may conjecture that properties of a group G are reflected in properties of $\operatorname{Sub} G$ and $\mathcal{N}\text{-}\operatorname{Sub} G$, and vice versa. To take a very simple example, a group G is finite if and only if $\operatorname{Sub} G$ is finite (an easy exercise). More interestingly, many group-theoretic properties of G are equivalent to order- or lattice-theoretic properties of $\operatorname{Sub} G$ or $\mathcal{N}\text{-}\operatorname{Sub} G$. Some results in this spirit can be found in the exercises for Chapters 4 and 5. These results do not appear in the text because their proofs lean much more heavily on group theory than on lattice theory.

Lattices as algebraic structures

We introduced lattices as ordered sets of a special type. However, we may adopt an alternative viewpoint. Given a lattice L, we may define binary operations **join** and **meet** on the non-empty set L by

$$a \vee b := \sup\{a, b\} \quad \text{and} \quad a \wedge b := \inf\{a, b\} \quad (a, b \in L).$$

Note that 2.5(4)(ii) says precisely that the operations $\vee\colon L^2 \to L$ and $\wedge\colon L^2 \to L$ are order-preserving.

In this section we view a lattice as an algebraic structure $\langle L; \vee, \wedge \rangle$ and explore the properties of these binary operations. We first amplify the connection between $\vee, \wedge$ and $\leqslant$. Since we shall often use the following lemma, it deserves a name.

2.8 The Connecting Lemma. *Let L be a lattice and let $a, b \in L$. Then the following are equivalent:*

(i) $a \leqslant b$;

(ii) $a \vee b = b$;

(iii) $a \wedge b = a$.

Proof. It was shown in 2.5(1) that (i) implies both (ii) and (iii). Now assume (ii). Then b is an upper bound for $\{a, b\}$, whence $b \geqslant a$. Thus (i) holds. Similarly, (iii) implies (i). □

2.9 Theorem. *Let L be a lattice. Then $\vee$ and $\wedge$ satisfy, for all $a, b, c \in L$,*

(L1)	$(a \vee b) \vee c = a \vee (b \vee c)$	(*associative laws*)
(L1)$^\partial$	$(a \wedge b) \wedge c = a \wedge (b \wedge c)$	
(L2)	$a \vee b = b \vee a$	(*commutative laws*)
(L2)$^\partial$	$a \wedge b = b \wedge a$	
(L3)	$a \vee a = a$	(*idempotency laws*)
(L3)$^\partial$	$a \wedge a = a$	
(L4)	$a \vee (a \wedge b) = a$	(*absorption laws*)
(L4)$^\partial$	$a \wedge (a \vee b) = a$.	

Proof. Note that the dual of a statement about lattices phrased in terms of $\vee$ and $\wedge$ is obtained simply by interchanging $\vee$ and $\wedge$ (this is the **Duality Principle for lattices**). It is therefore enough to consider (L1)–(L4).

We proved (L3) in 2.5(1) and (L2) is immediate because, for any set S, $\sup S$ is independent of the order in which the elements of S are listed. Also, (L4) follows easily from the Connecting Lemma. To prove (L1) it is enough, thanks to (L2), to show that $(a \vee b) \vee c = \sup\{a, b, c\}$.

This is the case if $\{a \vee b, c\}^u = \{a, b, c\}^u$. But

$$\begin{aligned} d \in \{a,b,c\}^u &\iff d \in \{a,b\}^u \text{ and } d \geqslant c \\ &\iff d \geqslant a \vee b \text{ and } d \geqslant c \\ &\iff d \in \{a \vee b, c\}^u. \end{aligned}$$

□

We now turn things round and start from a set carrying operations $\vee$ and $\wedge$ which satisfy the identities given in the preceding theorem.

2.10 Theorem. *Let $\langle L; \vee, \wedge\rangle$ be a non-empty set equipped with two binary operations which satisfy* (L1)–(L4) *and* $(\text{L1})^\partial$–$(\text{L4})^\partial$ *from* 2.9.

(i) *For all $a, b \in L$, we have $a \vee b = b$ if and only if $a \wedge b = a$.*

(ii) *Define $\leqslant$ on L by $a \leqslant b$ if $a \vee b = b$. Then $\leqslant$ is an order relation.*

(iii) *With $\leqslant$ as in* (ii), *$\langle L; \leqslant\rangle$ is a lattice in which the original operations agree with the induced operations, that is, for all $a, b \in L$,*

$$a \vee b = \sup\{a, b\} \text{ and } a \wedge b = \inf\{a, b\}.$$

Proof. Assume $a \vee b = b$. Then

$$\begin{aligned} a &= a \wedge (a \vee b) && \text{(by } (\text{L4})^\partial\text{)} \\ &= a \wedge b && \text{(by assumption).} \end{aligned}$$

Conversely, assume $a \wedge b = a$. Then

$$\begin{aligned} b &= b \vee (b \wedge a) && \text{(by (L4))} \\ &= b \vee (a \wedge b) && \text{(by } (\text{L2})^\partial\text{)} \\ &= b \vee a && \text{(by assumption)} \\ &= a \vee b && \text{(by (L2)).} \end{aligned}$$

Now define $\leqslant$ as in (ii). Then $\leqslant$ is reflexive by (L3), antisymmetric by (L2) and transitive by (L1). The details are left as an exercise.

To show that $\sup\{a, b\} = a \vee b$ in the ordered set $\langle L; \leqslant\rangle$, it must first be checked that $a \vee b \in \{a, b\}^u$ and second that $d \in \{a, b\}^u$ implies $d \geqslant a \vee b$. To do this, remember that $\leqslant$ is given by $p \leqslant q$ if and only if $p \vee q = q$ and justify each step by appealing to one of the identities. The characterization of inf is obtained by duality (again). □

2.11 Stocktaking. We have shown that lattices can be completely characterized in terms of the join and meet operations. We may henceforth say 'let L be a lattice', replacing L by $\langle L; \leqslant\rangle$ or by $\langle L; \vee, \wedge\rangle$ if we want to emphasize that we are thinking of it as a special kind of ordered set or as an algebraic structure.

In a lattice L, associativity of $\vee$ and $\wedge$ allows us to write iterated joins and meets unambiguously without brackets. An easy induction shows that these correspond to sups and infs in the expected way:

$$\bigvee\{a_1, \ldots, a_n\} = a_1 \vee \cdots \vee a_n,$$

for $a_1, \ldots, a_n \in L$ $(n \geqslant 1)$, and dually; observe that the case $n = 3$ underlies the proof of (L1) in 2.9. Consequently, $\bigvee F$ and $\bigwedge F$ exist for any finite, non-empty subset F of a lattice.

2.12 Definitions. Let L be a lattice. It may happen that $\langle L; \leqslant\rangle$ has top and bottom elements $\top$ and $\bot$ as defined in 1.21. When thinking of L as $\langle L; \vee, \wedge\rangle$, it is appropriate to view these elements from a more algebraic standpoint. We say L has a **one** if there exists $1 \in L$ such that $a = a \wedge 1$ for all $a \in L$. Dually, L is said to have a **zero** if there exists $0 \in L$ such that $a = a \vee 0$ for all $a \in L$. The lattice $\langle L; \vee, \wedge\rangle$ has a one if and only if $\langle L; \leqslant\rangle$ has a top element $\top$ and, in that case, $1 = \top$. A dual statement holds for 0 and $\bot$. A lattice $\langle L; \vee, \wedge\rangle$ possessing 0 and 1 is called **bounded**. A finite lattice is automatically bounded, with $1 = \bigvee L$ and $0 = \bigwedge L$. Recalling 2.6(5), note that $\langle \mathbb{N}_0; \text{lcm}, \text{gcd}\rangle$ is bounded – with $1 = 0$ and $0 = 1$!!

Sublattices, products and homomorphisms

This section presents methods for deriving new lattices.

2.13 Sublattices. Let L be a lattice and $\varnothing \neq M \subseteq L$. Then M is a **sublattice** of L if

$$a, b \in M \text{ implies } a \vee b \in M \text{ and } a \wedge b \in M.$$

We denote the collection of all sublattices of L by $\text{Sub}\, L$ and let $\text{Sub}_0 L = \text{Sub}\, L \cup \{\varnothing\}$; both are ordered by inclusion.

2.14 Examples.

(1) Any one-element subset of a lattice is a sublattice. More generally, any non-empty chain in a lattice is a sublattice. (In fact, when testing that a non-empty subset M is a sublattice, it is sufficient to consider non-comparable elements a, b in 2.13.)

(2) In the diagrams in Figure 2.6 the shaded elements in lattices (i) and (ii) form sublattices, while those in (iii) and (iv) do not.

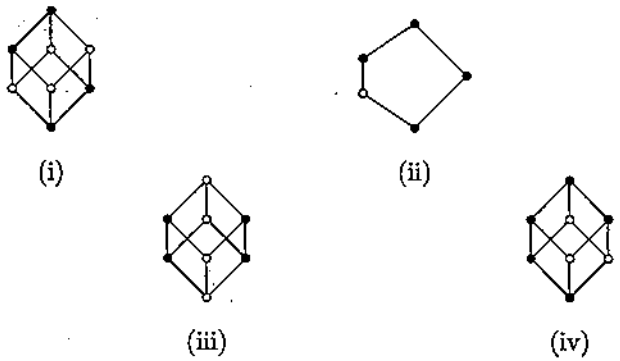


Figure 2.6

(3) A subset M of a lattice $\langle L; \leqslant \rangle$ may be a lattice in its own right without being a sublattice of L; see Figure 2.6(iv) for an example.

2.15 Products. Let L and K be lattices. Define $\vee$ and $\wedge$ coordinatewise on $L \times K$, as follows:

$$(\ell_1, k_1) \vee (\ell_2, k_2) = (\ell_1 \vee \ell_2, k_1 \vee k_2),$$
$$(\ell_1, k_1) \wedge (\ell_2, k_2) = (\ell_1 \wedge \ell_2, k_1 \wedge k_2).$$

It is routine to check that $L \times K$ satisfies the identities (L1)–(L4)$^\partial$ and therefore is a lattice. Also

$$\begin{aligned}(\ell_1, k_1) \vee (\ell_2, k_2) = (\ell_2, k_2) &\iff \ell_1 \vee \ell_2 = \ell_2 \text{ and } k_1 \vee k_2 = k_2 \\ &\iff \ell_1 \leqslant \ell_2 \text{ and } k_1 \leqslant k_2 \\ &\iff (\ell_1, k_1) \leqslant (\ell_2, k_2),\end{aligned}$$

with respect to the order on $L \times K$ defined in 1.25. Hence the lattice formed by taking the ordered set product of lattices L and K is the same as that obtained by defining $\vee$ and $\wedge$ coordinatewise on $L \times K$.

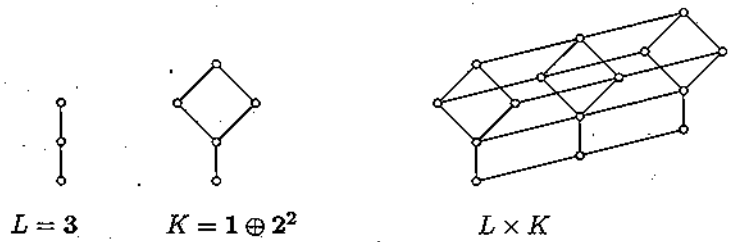


Figure 2.7

Figure 2.7 shows the product of the lattices $L = \mathbf{3}$ and $K = \mathbf{1} \oplus \mathbf{2}^2$. Notice how (isomorphic copies) of L and K sit inside $L \times K$ as the sublattices $L \times \{0\}$ and $\{0\} \times K$. It is routine to show that the product of lattices L and K always contains sublattices isomorphic to L and K.

Iterated products and powers are defined in the obvious way. It is possible to define the product of an infinite family of lattices, but we shall not need this construction.

2.16 Homomorphisms. From the viewpoint of lattices as algebraic structures it is natural to regard as canonical those maps between lattices which preserve the operations join and meet. Since lattices are also ordered sets, order-preserving maps are also available. We need to explore the relationship between these classes of maps. We begin with some definitions.

Let L and K be lattices. A map $f\colon L \to K$ is said to be a **homomorphism** (or, for emphasis, **lattice homomorphism**) if f is **join-preserving** and **meet-preserving**, that is, for all $a, b \in L$,

$$f(a \vee b) = f(a) \vee f(b) \quad \text{and} \quad f(a \wedge b) = f(a) \wedge f(b).$$

A bijective homomorphism is a **(lattice) isomorphism**. If $f\colon L \to K$ is a one-to-one homomorphism, then the sublattice $f(L)$ of K is isomorphic to L and we refer to f as an **embedding (of L into K)**.

2.17 Remarks.

(1) It is straightforward to check that the inverse of an isomorphism is a homomorphism and hence is also an isomorphism.

(2) We write $L \rightarrowtail K$ to indicate that the lattice K has a sublattice isomorphic to the lattice L. It follows from (ii) of the proposition below that $M \rightarrowtail L$ implies $M \hookrightarrow L$ (recall 1.34).

(3) For bounded lattices L and K it is often appropriate to consider homomorphisms $f\colon L \to K$ such that $f(0) = 0$ and $f(1) = 1$. Such maps are called **$\{0,1\}$-homomorphisms**. They arise in Chapters 5 and 11.

2.18 Examples. Referring back to Figure 1.11, we see that each of φ_2–φ_6 is an order-preserving map from one lattice to another. The maps φ_2 and φ_3 are homomorphisms, the remainder are not. Neither join nor meet is preserved by φ_4. The map φ_5 preserves joins but does not preserve all meets; φ_6 is meet-preserving but does not preserve all joins. Thus in general an order-preserving map may not be a homomorphism.

The possible demarcation dispute between order-isomorphism and lattice isomorphism does not arise, as 2.19(ii) below shows.

2.19 Proposition. *Let L and K be lattices and $f: L \to K$ a map.*

(i) *The following are equivalent:*

(a) *f is order-preserving;*

(b) $(\forall a, b \in L)\ f(a \vee b) \geqslant f(a) \vee f(b)$;

(c) $(\forall a, b \in L)\ f(a \wedge b) \leqslant f(a) \wedge f(b)$.

In particular, if f is a homomorphism, then f is order-preserving.

(ii) *f is a lattice isomorphism if and only if it is an order-isomorphism.*

Proof. Part (i) is an easy consequence of the Connecting Lemma. Consider (ii). Assume that f is a lattice isomorphism. Then, by the Connecting Lemma,

$$a \leqslant b \Leftrightarrow a \vee b = b \Leftrightarrow f(a \vee b) = f(b) \Leftrightarrow f(a) \vee f(b) = f(b) \Leftrightarrow f(a) \leqslant f(b),$$

whence f is an order-embedding and so is an order-isomorphism. Conversely, assume that f is an order-isomorphism. Then f is bijective (see 1.4). By (i) and duality, to show that f is a lattice isomorphism it suffices to show that $f(a) \vee f(b) \geqslant f(a \vee b)$ for all $a, b \in L$. Since f is surjective, there exists $c \in L$ such that $f(a) \vee f(b) = f(c)$. Then $f(a) \leqslant f(c)$ and $f(b) \leqslant f(c)$. Since f is an order-embedding, it follows that $a \leqslant c$ and $b \leqslant c$, whence $a \vee b \leqslant c$. Because f is order-preserving, $f(a \vee b) \leqslant f(c) = f(a) \vee f(b)$, as required. □

Ideals and filters

Ideals are of fundamental importance in algebra. Filters, the order duals of lattice ideals, have a variety of applications in logic and topology. Ideals, specifically prime ideals, which we consider in Chapter 10, form the basis for our representation theory in Chapter 11.

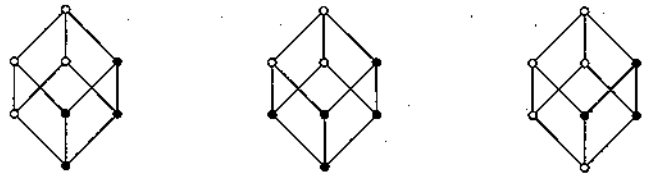
2.20 Definitions. Let L be a lattice. A non-empty subset J of L is called an **ideal** if

(i) $a, b \in J$ implies $a \vee b \in J$,

(ii) $a \in L$, $b \in J$ and $a \leqslant b$ imply $a \in J$.

See Figure 2.8 for illustrations.

The definition can be more compactly stated by declaring an ideal to be a non-empty down-set closed under join. We spelt out the definition as we did to draw a parallel between a lattice ideal and an ideal in a ring. Later, Exercise 4.29 shows the connection to be stronger than just



shaded elements an ideal | shaded elements not an ideal | shaded elements not an ideal

Figure 2.8

an analogy. Clearly, every ideal J of a lattice L is a sublattice, since $a \wedge b \leqslant a$ for any $a, b \in L$.

A dual ideal is called a filter. Specifically, a non-empty subset G of L is called a **filter** if

(i) $a, b \in G$ implies $a \wedge b \in G$,

(ii) $a \in L$, $b \in G$ and $a \geqslant b$ imply $a \in G$.

The set of all ideals (filters) of L is denoted by $\mathcal{I}(L)$ (by $\mathcal{F}(L)$), and carries the usual inclusion order.

An ideal or filter is called **proper** if it does not coincide with L. It is a very easy exercise to show that an ideal J of a lattice with 1 is proper if and only if $1 \notin J$, and dually, a filter G of a lattice with 0 is proper if and only if $0 \notin G$. For each $a \in L$, the set $\downarrow a$ is an ideal; it is known as the **principal ideal** generated by a. Dually, $\uparrow a$ is a **principal filter**.

2.21 Examples.

(1) In a finite lattice, every ideal or filter is principal: the ideal J equals $\downarrow \bigvee J$, and dually for a filter. (See also Exercise 2.37(ii).)

(2) Let L and K be bounded lattices and $f: L \to K$ a $\{0,1\}$-homomorphism. Then $f^{-1}(0)$ is an ideal and $f^{-1}(1)$ is a filter in L.

(3) The following are ideals in $\mathcal{P}(X)$:

(a) all subsets not containing a fixed element of X;

(b) all finite subsets (this ideal is non-principal if X is infinite).

(4) Let $(X; T)$ be a topological space and let $x \in X$. Then the set $\{V \subseteq X \mid (\exists U \in T)\, x \in U \subseteq V\}$ is a filter in $\mathcal{P}(X)$. Convergence in a topological space can be elegantly formulated in terms of such neighbourhood filters; see for example [5].

Complete lattices and $\bigcap$-structures

We now return to complete lattices, which were briefly introduced at the start of this chapter. Recall from 2.4 that a complete lattice is defined to be a non-empty, ordered set P such that the join (supremum), $\bigvee S$, and the meet (infimum), $\bigwedge S$, exist for every subset S of P.

We first collect together in a sequence of elementary lemmas useful information for computing with arbitrary joins and meets, extending the results for binary joins and meets presented earlier. The first lists some immediate consequences of the definitions of least upper bound and greatest lower bound.

2.22 Lemma. *Let P be an ordered set, let $S,T \subseteq P$ and assume that $\bigvee S$, $\bigvee T$, $\bigwedge S$ and $\bigwedge T$ exist in P.*

(i) *$s \leqslant \bigvee S$ and $s \geqslant \bigwedge S$ for all $s \in S$.*

(ii) *Let $x \in P$; then $x \leqslant \bigwedge S$ if and only if $x \leqslant s$ for all $s \in S$.*

(iii) *Let $x \in P$; then $x \geqslant \bigvee S$ if and only if $x \geqslant s$ for all $s \in S$.*

(iv) *$\bigvee S \leqslant \bigwedge T$ if and only if $s \leqslant t$ for all $s \in S$ and all $t \in T$.*

(v) *If $S \subseteq T$, then $\bigvee S \leqslant \bigvee T$ and $\bigwedge S \geqslant \bigwedge T$.*

A straightforward application of this lemma yields the next one, which shows that join and meet behave well with respect to set union. We leave the proof as an exercise.

2.23 Lemma. *Let P be a lattice, let $S,T \subseteq P$ and assume that $\bigvee S$, $\bigvee T$, $\bigwedge S$ and $\bigwedge T$ exist in P. Then*

$$\bigvee(S \cup T) = (\bigvee S) \vee (\bigvee T) \quad \text{and} \quad \bigwedge(S \cup T) = (\bigwedge S) \wedge (\bigwedge T).$$

An easy induction now yields the following result, previously noted in 2.11, but worth re-iterating. The corollary follows from 2.2.

2.24 Lemma. *Let P be a lattice. Then $\bigvee F$ and $\bigwedge F$ exist for every finite, non-empty subset F of P.*

2.25 Corollary. *Every finite lattice is complete.*

We now describe how joins and meets interact with order-preserving maps and order-isomorphisms. First we need a definition.

2.26 Definition. Let P and Q be ordered sets and $\varphi\colon P \to Q$ a map. Then we say that φ **preserves existing joins** if whenever $\bigvee S$ exists in P then $\bigvee \varphi(S)$ exists in Q and $\varphi(\bigvee S) = \bigvee \varphi(S)$. Preservation of existing meets is defined dually.

2.27 Lemma. *Let P and Q be ordered sets and $\varphi\colon P \to Q$ be an order-preserving map.*

(i) *Assume that $S \subseteq P$ is such that $\bigvee S$ exists in P and $\bigvee \varphi(S)$ exists in Q. Then $\varphi(\bigvee S) \geqslant \bigvee \varphi(S)$. Dually, $\varphi(\bigwedge S) \leqslant \bigwedge \varphi(S)$ if both meets exist.*

(ii) *Assume now that $\varphi\colon P \to Q$ is an order-isomorphism. Then φ preserves all existing joins and meets.*

The next lemma is useful for showing that certain subsets of complete lattices are themselves complete lattices. Its corollary, obtained by taking P to be a powerset with the inclusion order, is used sufficiently frequently to deserve an explicit statement.

2.28 Lemma. *Let Q be a subset, with the induced order, of some ordered set P and let $S \subseteq Q$. If $\bigvee_P S$ exists and belongs to Q, then $\bigvee_Q S$ exists and equals $\bigvee_P S$ (and dually for $\bigwedge_Q S$).*

Proof. For any $x \in S$ we have $x \leqslant \bigvee_P S$; since $\bigvee_P S \in Q$ by hypothesis, it acts as an upper bound for S in Q. Further, if y is any upper bound for S in Q, it is also an upper bound for S in P and so $y \geqslant \bigvee_P S$. □

2.29 Corollary. *Let $\mathfrak{L}$ be a family of subsets of a set X and let $\{A_i\}_{i\in I}$ be a subset of $\mathfrak{L}$.*

(i) *If $\bigcup_{i\in I} A_i \in \mathfrak{L}$, then $\bigvee_{\mathfrak{L}}\{A_i \mid i \in I\}$ exists and equals $\bigcup_{i\in I} A_i$.*

(ii) *If $\bigcap_{i\in I} A_i \in \mathfrak{L}$, then $\bigwedge_{\mathfrak{L}}\{A_i \mid i \in I\}$ exists and equals $\bigcap_{i\in I} A_i$.*

Consequently, any (complete) lattice of sets is a (complete) lattice with joins and meets given by union and intersection.

To show that an ordered set is a complete lattice requires only half as much work as the definition would have us believe.

2.30 Lemma. *Let P be an ordered set such that $\bigwedge S$ exists in P for every non-empty subset S of P. Then $\bigvee S$ exists in P for every subset S of P which has an upper bound in P; indeed, $\bigvee S = \bigwedge S^u$.*

Proof. Let $S \subseteq P$ and assume that S has an upper bound in P; thus $S^u \neq \varnothing$. Hence, by assumption, $a := \bigwedge S^u$ exists in P. We claim that $\bigvee S = a$. The details are left as an exercise. □

2.31 Theorem. *Let P be a non-empty ordered set. Then the following are equivalent:*

(i) *P is a complete lattice;*

(ii) *$\bigwedge S$ exists in P for every subset S of P;*

(iii) *P has a top element, $\top$, and $\bigwedge S$ exists in P for every non-empty subset S of P.*

Proof. It is trivial that (i) implies (ii), and (ii) implies (iii) since the meet of the empty subset of P exists only if P has a top element (by 2.2). It follows easily from the previous lemma that (iii) implies (i); the details are left to the reader. □

This theorem has a simple corollary which, nevertheless, yields many examples of complete lattices.

2.32 Corollary. *Let X be a set and let $\mathfrak{L}$ be a family of subsets of X, ordered by inclusion, such that*

(a) $\bigcap_{i\in I} A_i \in \mathfrak{L}$ *for every non-empty family* $\{A_i\}_{i\in I} \subseteq \mathfrak{L}$, *and*

(b) $X \in \mathfrak{L}$.

Then $\mathfrak{L}$ is a complete lattice in which

$$\bigwedge_{i\in I} A_i = \bigcap_{i\in I} A_i,$$
$$\bigvee_{i\in I} A_i = \bigcap \{ B \in \mathfrak{L} \mid \bigcup_{i\in I} A_i \subseteq B \}.$$

Proof. By Theorem 2.31, to show that $\langle \mathfrak{L}; \subseteq \rangle$ is a complete lattice it suffices to show that $\mathfrak{L}$ has a top element and that the meet of every non-empty subset of $\mathfrak{L}$ exists in $\mathfrak{L}$. By (b), $\mathfrak{L}$ has a top element, namely X. Let $\{A_i\}_{i\in I}$ be a non-empty subset of $\mathfrak{L}$; then (a) gives $\bigcap_{i\in I} A_i \in \mathfrak{L}$. Therefore Corollary 2.29 implies that $\bigwedge_{i\in I} A_i$ exists and is given by $\bigcap_{i\in I} A_i$. Thus $\langle \mathfrak{L}; \subseteq \rangle$ is a complete lattice. Since X is an upper bound of $\{A_i\}_{i\in I}$ in $\mathfrak{L}$, Lemma 2.30 gives

$$\begin{aligned} \bigvee_{i\in I} A_i &= \bigwedge \{ A_i \mid i \in I \}^u \\ &= \bigcap \{ B \in \mathfrak{L} \mid (\forall i \in I)\, A_i \subseteq B \} \\ &= \bigcap \{ B \in \mathfrak{L} \mid \bigcup_{i\in I} A_i \subseteq B \}. \end{aligned}$$ □

2.33 Definitions. If $\mathfrak{L}$ is a non-empty family of subsets of X which satisfies condition (a) of Corollary 2.32, then $\mathfrak{L}$ is called an **intersection structure** (or **$\bigcap$-structure**) on X. If $\mathfrak{L}$ also satisfies (b), we refer to it as a **topped intersection structure** on X. An alternative term is **closure system**; see 7.4.

Intersection structures which occur in computer science are usually topless while those in algebra are almost invariably topped. In a complete lattice $\mathfrak{L}$ of this type, the meet is just set intersection, but in general the join is not set union. This is illustrated in the examples which follow.

2.34 Examples.

(1) Consider $(X \multimap Y)$, where X and Y are any non-empty sets. From the observations in 1.10 we saw that the map $\pi \mapsto \operatorname{graph}\pi$ is an order-embedding of $(X \multimap Y)$ into $\mathcal{P}(X \times Y)$. Let $\mathfrak{L}$ be the family of subsets of $X \times Y$ which are graphs of partial maps. To prove that $\mathfrak{L}$ is closed under intersections, use the characterization given in 1.10: if $S \subseteq X \times Y$, then $S \in \mathfrak{L}$ if and only if $(s,y) \in S$ and $(s,y') \in S$ imply $y = y'$. Thus $\mathfrak{L}$ is an $\bigcap$-structure. It is not topped unless $|Y| = 1$.

(2) Each of the following is a topped $\bigcap$-structure and so forms a complete lattice under inclusion:

- the subgroups, $\operatorname{Sub} G$, of a group G;
- the normal subgroups, $\mathcal{N}\text{-}\operatorname{Sub} G$, of a group G;
- the equivalence relations on a set X;
- the subspaces, $\operatorname{Sub} V$ of a vector space V;
- the convex subsets of a real vector space;
- the subrings of a ring;
- the ideals of a ring;
- $\operatorname{Sub}_0 L$, the sublattices of a lattice L, with the empty set adjoined (note that $\operatorname{Sub} L$ is not closed under intersections, except when $|L| = 1$);
- the ideals of a lattice L with 0 (or, if L has no zero element, the ideals of L with the empty set added), and dually for filters.

These families all belong to a class of $\bigcap$-structures – called **algebraic $\bigcap$-structures** because of their provenance – which we shall consider further in Chapter 7.

(3) The closed subsets of a topological space are closed under finite unions and finite intersections and hence form a lattice of sets in which $A \vee B = A \cup B$ and $A \wedge B = A \cap B$. In fact, the closed sets form a topped $\bigcap$-structure and consequently the lattice of closed sets is complete. The formulae for arbitrary (rather than finite) joins and meets given in 2.32 show that, in general, meet is given by intersection while the join of a family of closed sets is not their union but is obtained by forming the closure of their union.

(4) Since the open subsets of a topological space are closed under arbitrary union and include the empty set, the dual of 2.31 shows that they form a complete lattice under inclusion. The dual version

of 2.32 shows that join and meet are given by

$$\bigvee_{i\in I} A_i = \bigcup_{i\in I} A_i \text{ and } \bigwedge_{i\in I} A_i = \mathrm{Int}\big(\bigcap_{i\in I} A_i\big),$$

where $\mathrm{Int}(A)$ denotes the interior of A.

We conclude this section with a famous theorem concerning complete lattices, with an appealingly simple proof. Given an ordered set P and a map $F\colon P \to P$, an element $x \in P$ is called a **fixpoint** of F if $F(x) = x$. A full discussion of fixpoints is presented in Chapter 8 and some applications of them are discussed in Chapter 9.

2.35 The Knaster–Tarski Fixpoint Theorem. *Let L be a complete lattice and $F\colon L \to L$ an order-preserving map. Then*

$$\alpha := \bigvee \{\, x \in L \mid x \leqslant F(x) \,\}$$

is a fixpoint of F. Further, α is the greatest fixpoint of F. Dually, F has a least fixpoint, given by $\bigwedge\{\, x \in L \mid F(x) \leqslant x \,\}$.

Proof. Let $H = \{\, x \in L \mid x \leqslant F(x) \,\}$. For all $x \in H$ we have $x \leqslant \alpha$, so $x \leqslant F(x) \leqslant F(\alpha)$. Thus $F(\alpha) \in H^u$, whence $\alpha \leqslant F(\alpha)$. We now use this inequality to prove the reverse one (!) and thereby complete the proof that α is a fixpoint. Since F is order-preserving, $F(\alpha) \leqslant F(F(\alpha))$. This says $F(\alpha) \in H$, so $F(\alpha) \leqslant \alpha$. If β is any fixpoint of F then $\beta \in H$, so $\beta \leqslant \alpha$. □

2.36 Remark. Theorems which produce fixpoints of certain maps have been extensively exploited in computer science. The Knaster–Tarski Theorem is no exception, but it also has an important application of a different kind: it yields as a by-product the famous Schröder–Bernstein Theorem stating that there is a bijection between sets A and B if there exist one-to-one maps from A to B and from B to A. For the proof, see Exercise 2.32.

Chain conditions and completeness

By Corollary 2.25, every finite lattice is complete. There are various finiteness conditions, of which 'P is finite' is the strongest, which will guarantee that a lattice P is complete.

2.37 Definitions. Let P be an ordered set.

(i) If $C = \{c_0, c_1, \ldots, c_n\}$ is a finite chain in P with $|C| = n+1$, then we say that the **length** of C is n.

(ii) P is said to have **length** n, written $\ell(P) = n$, if the length of the longest chain in P is n.

(iii) P is of **finite length** if it has length n for some $n \in \mathbb{N}_0$.

(iv) P has **no infinite chains** if every chain in P is finite.

(v) P satisfies the **ascending chain condition**, (ACC), if given any sequence $x_1 \leqslant x_2 \leqslant \cdots \leqslant x_n \leqslant \cdots$ of elements of P, there exists $k \in \mathbb{N}$ such that $x_k = x_{k+1} = \ldots$. The dual of the ascending chain condition is the **descending chain condition**, (DCC).

2.38 Examples.

(1) The lattices $\mathbf{M}_n$ of Figure 2.4 are of length 2. A lattice of finite length has no infinite chains and so satisfies both (ACC) and (DCC).

(2) The lattice $\langle \mathbb{N}_0; \preccurlyeq \rangle$ satisfies (DCC) but not (ACC).

(3) Consider the lattices in Figure 2.9. The chain $\mathbb{N}$ satisfies (DCC) but not (ACC), and, dually, $\mathbb{N}^\partial$ satisfies (ACC) but not (DCC). The lattice $\mathbf{1} \oplus (\bigcup_{n\in\mathbb{N}} \mathbf{n}) \oplus \mathbf{1}$ is the simplest example of a lattice which has no infinite chains but is not of finite length.

(4) The finiteness conditions in 2.37 first arose in 'classical' algebra. For example, it can be shown that a vector space V is finite dimensional if and only if $\mathrm{Sub}\, V$ is of finite length, in which case $\dim V = \ell(\mathrm{Sub}\, V)$.

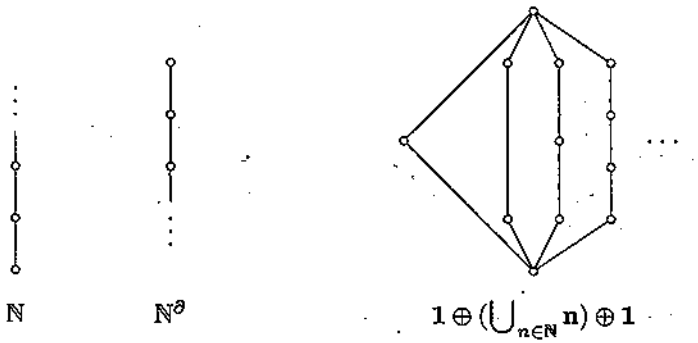


Figure 2.9

A formal proof of the following lemma requires some form of the axiom of set theory known as the **Axiom of Choice**. We give an informal derivation here and reveal in Chapter 10 how this can be converted into a formal proof (see 10.1).

2.39 Lemma. *An ordered set P satisfies* (ACC) *if and only if every non-empty subset A of P has a maximal element.*

Informal proof: We shall prove the contrapositive in both directions, that is, we prove that P has an infinite ascending chain if and only if there is a non-empty subset A of P which has no maximal element.

Assume that $x_1 < x_2 < \cdots < x_n < \cdots$ is an infinite ascending chain in P; then clearly $A := \{ x_n \mid n \in \mathbb{N} \}$ has no maximal element. Conversely, assume that A is a non-empty subset of P which has no maximal element. Let $x_1 \in A$. Since x_1 is not maximal in A, there exists $x_2 \in A$ with $x_1 < x_2$. Similarly, there exists $x_3 \in A$ with $x_2 < x_3$. Continuing in this way (and this is where the Axiom of Choice comes in) we obtain an infinite ascending chain in P. □

2.40 Theorem. *An ordered set P has no infinite chains if and only if it satisfies both* (ACC) *and* (DCC).

Proof. Clearly if P has no infinite chains, then it satisfies both (ACC) and (DCC). Suppose that P satisfies both (ACC) and (DCC) and contains an infinite chain C. Note that if A is a non-empty subset of C, then A has a maximal element m, by 2.39. If $a \in A$, then since C is a chain we have $a \leqslant m$ or $m \leqslant a$. But $m \leqslant a$ implies $m = a$ by the maximality of m. Hence $a \leqslant m$ for all $a \in A$, and so every non-empty subset of C has a greatest element. Let x_1 be the greatest element of C, let x_2 be the greatest element of $C \setminus \{x_1\}$, and in general let x_{n+1} be the greatest element of $C \setminus \{x_1, x_2, \ldots, x_n\}$. Then $x_1 \succ x_2 \succ \cdots \succ x_n \succ \cdots$ is an infinite, descending, covering chain in P, ↯. □

Lattices with no infinite chains are complete as the following more general result shows.

2.41 Theorem. *Let P be a lattice.*

(i) *If P satisfies* (ACC), *then for every non-empty subset A of P there exists a finite subset F of A such that $\bigvee A = \bigvee F$ (which exists in P by 2.24).*

(ii) *If P has a bottom element and satisfies* (ACC), *then P is complete.*

(iii) *If P has no infinite chains, then P is complete.*

Proof. Assume that P satisfies (ACC) and let A be a non-empty subset of P. Then, by 2.24,

$$B := \{ \textstyle\bigvee F \mid F \text{ is a finite non-empty subset of } A \}$$

is a well-defined subset of P. Since B is non-empty, 2.39 guarantees that B has a maximal element $m = \bigvee F$ for some finite subset F of A.

Let $a \in A$. Then $\bigvee(F \cup \{a\}) \in B$ and $m = \bigvee F \leqslant \bigvee(F \cup \{a\})$ by 2.22(v). Thus $m = \bigvee F = \bigvee(F \cup \{a\})$ since m is maximal in B. As $m = \bigvee(F \cup \{a\})$ we have $a \leqslant m$, whence m is an upper bound of A. Let $x \in P$ be an upper bound of A. Then x is an upper bound of F since $F \subseteq A$ and hence $m = \bigvee F \leqslant x$. Thus m is the least upper bound of A; that is, $\bigvee A = m = \bigvee F$. Hence (i) holds.

Combining (i) with the dual of 2.31 yields (ii), and since a lattice with no infinite chains has a bottom element and satisfies (ACC), (iii) follows from (ii). □

Join-irreducible elements

The Fundamental Theorem of Arithmetic says that every natural number is a product of prime numbers. Since prime numbers are just the product-irreducible natural numbers (other than 1) an analogous result for lattices would state that every element is a meet of meet-irreducible elements or, dually, a join of join-irreducible elements. This will not be true in general but will hold provided we impose an appropriate finiteness condition. We prefer to build from the bottom up rather than the top down and consequently focus on joins rather than meets.

2.42 Definitions. Let L be a lattice. An element $x \in L$ is **join-irreducible** if

(i) $x \neq 0$ (in case L has a zero);

(ii) $x = a \vee b$ implies $x = a$ or $x = b$ for all $a, b \in L$.

Condition (ii) is equivalent to the more pictorial

(ii)′ $a < x$ and $b < x$ imply $a \vee b < x$ for all $a, b \in L$.

A **meet-irreducible** element is defined dually. We denote the set of join-irreducible elements of L by $\mathcal{J}(L)$ and the set of meet-irreducible elements by $\mathcal{M}(L)$. Each of these sets inherits L's order relation, and will be regarded as an ordered set.

Let P be an ordered set and let $Q \subseteq P$. Then Q is called **join-dense** in P if for every element $a \in P$ there is a subset A of Q such that $a = \bigvee_P A$. The dual of join-dense is **meet-dense**.

2.43 Examples of join-irreducible elements.

(1) In a chain, every non-zero element is join-irreducible. Thus if L is an n-element chain, then $\mathcal{J}(L)$ is an $(n-1)$-element chain.

(2) In a finite lattice L, an element is join-irreducible if and only if it has exactly one lower cover. This makes $\mathcal{J}(L)$ extremely easy to

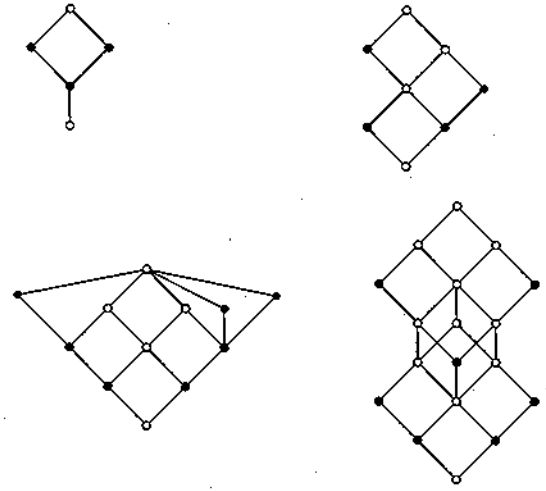


Figure 2.10

identify from a diagram of L. Figure 2.10 gives some examples. The join-irreducible elements are shaded.

(3) Consider the lattice $\langle \mathbb{N}_0; \text{lcm}, \text{gcd}\rangle$. A non-zero element $m \in \mathbb{N}_0$ is join-irreducible if and only if m is of the form p^r, where p is a prime and $r \in \mathbb{N}$.

(4) In a lattice $\wp(X)$ the join-irreducible elements are exactly the singleton sets, $\{x\}$, for $x \in X$.

(5) It is easily seen that the lattice of open subsets of $\mathbb{R}$ (that is, subsets which are unions of open intervals) has no join-irreducible elements.

2.44 Remarks. In the definition in 2.42 we have debarred 0 from being regarded as join-irreducible. This is necessary for the claim on lower covers in 2.43(2) to be valid and leads to tidier statements of results in Chapter 5. Certainly we can never write 0 as a non-empty join, $\bigvee_P A$, unless $0 \in A$. To compensate for this we have not excluded $A = \varnothing$ in the definition of join-density in 2.42. Recall from 2.2 that $\bigvee_P \varnothing = 0$ in a lattice P with zero. Insisting that 0 is not join-irreducible is the lattice-theoretic equivalent of declaring that 1 is not a prime number.

Our examples have shown that join-irreducible elements do not necessarily exist in infinite lattices. On the other hand, it is easy to see that in a finite lattice every element is a join of join-irreducible elements. The next proposition proves more.

2.45 Proposition. *Let L be a lattice satisfying* (DCC).

(i) *Suppose $a, b \in L$ and $a \not\leqslant b$. Then there exists $x \in \mathcal{J}(L)$ such that $x \leqslant a$ and $x \not\leqslant b$.*

(ii) *$a = \bigvee \{x \in \mathcal{J}(L) \mid x \leqslant a\}$ for all $a \in L$.*

These conclusions hold in particular if L is finite.

Proof. Let $a \not\leqslant b$ and let $S := \{x \in L \mid x \leqslant a \text{ and } x \not\leqslant b\}$. The set S is non-empty since it contains a. Hence, since L satisfies (DCC), there exists a minimal element x of S (via the dual of Lemma 2.39). We claim that x is join-irreducible. Suppose that $x = c \vee d$ with $c < x$ and $d < x$. By the minimality of x, neither c nor d lies in S. We have $c < x \leqslant a$, so $c \leqslant a$, and similarly $d \leqslant a$. Therefore $c, d \notin S$ implies $c \leqslant b$ and $d \leqslant b$. But then $x = c \vee d \leqslant b$, $\lightning$. Thus $x \in \mathcal{J}(L) \cap S$, which proves (i).

Let $a \in L$ and let $T := \{x \in \mathcal{J}(L) \mid x \leqslant a\}$. Clearly a is an upper bound of T. Let c be an upper bound of T. We claim that $a \leqslant c$. Suppose that $a \not\leqslant c$; then $a \not\leqslant a \wedge c$. By (i) there exists $x \in \mathcal{J}(L)$ with $x \leqslant a$ and $x \not\leqslant a \wedge c$. Hence $x \in T$ and consequently $x \leqslant c$ since c is an upper bound of T. Thus x is a lower bound of $\{a, c\}$ and consequently $x \leqslant a \wedge c$, $\lightning$. Hence $a \leqslant c$, as claimed. This proves that $a = \bigvee T$ in L, whence (ii) holds. □

Part (iii) of our final result for this chapter is the promised analogue of (the existence portion of) the Fundamental Theorem of Arithmetic. (See Exercise 4.20 for a lattice-theoretic analogue of the uniqueness part of the theorem.)

The join-density of the join-irreducibles in a finite lattice plays a vital role in Chapter 5. More immediately, in Chapter 3 we exploit the join-density results in Theorem 2.46 in tandem with their meet-density duals.

2.46 Theorem. *Let L be a lattice.*

(i) *If L satisfies* (DCC)*, then $\mathcal{J}(L)$ and, more generally, any subset Q which contains $\mathcal{J}(L)$ is join-dense in L.*

(ii) *If L satisfies* (ACC) *and Q is join-dense in L, then, for each $a \in L$, there exists a finite subset F of Q such that $a = \bigvee F$.*

(iii) *If L has no infinite chains, then, for each $a \in L$, there exists a finite subset F of $\mathcal{J}(L)$ such that $a = \bigvee F$.*

(iv) *If L has no infinite chains, then Q is join-dense in L if and only if $\mathcal{J}(L) \subseteq Q$.*

Proof. (i) is an immediate consequence of part (ii) of the previous proposition, (ii) follows immediately from Theorem 2.41(i). Since no infinite chains implies both (ACC) and (DCC), (iii) is a consequence of (i)

and (ii). One direction of (iv) follows from (i). In the other direction, assume that Q is join-dense in L and let $x \in \mathcal{J}(L)$. By (ii), there is a finite subset F of Q such that $x = \bigvee F$. Since x is join-irreducible we have $x \in F$ and hence $x \in Q$. Thus, $\mathcal{J}(L) \subseteq Q$. □

Exercises

Exercises from the text. Prove the claims in 2.5(4). Show that the inverse of a lattice isomorphism is also a lattice isomorphism. Complete the proof of Proposition 2.19(i). Prove the unproved assertions in 2.21. Complete the proofs of 2.22, 2.23, 2.24. Prove the assertions in 2.34(2) concerning the families of sublattices and of ideals of a lattice. Prove the assertion that, if G is any group, $\mathrm{Sub}\, G$ is finite if and only if G is finite (see 2.7).

2.1 Consider the diagram in Figure 2.11 of the ordered subset $P = \{1,2,3,4,5,6,7\}$ of $\langle \mathbb{N}_0; \preccurlyeq \rangle$. Find the join and meet, where they exist, of each of the following subsets of P. Either specify the join or meet or indicate why it fails to exist.

(i) $\{3\}$, (ii) $\{4,6\}$, (iii) $\{2,3\}$, (iv) $\{2,3,6\}$, (v) $\{1,5\}$.

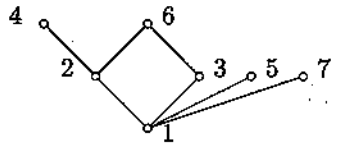


Figure 2.11

2.2 The ordered subset $Q = \{1,2,4,5,6,12,20,30,60\}$ of $\langle \mathbb{N}_0; \preccurlyeq \rangle$ is not a lattice. Draw a diagram of Q (it is 'cube-like') and find elements $a, b, c, d \in Q$ such that $a \vee b$ and $c \wedge d$ do not exist in Q.

2.3 Consider the ordered set drawn in Figure 2.2. Use the technique illustrated in Example 2.5(3) to calculate the following elements or to explain why they fail to exist.

(i) $h \wedge i$ (ii) $a \vee c$ (iii) $h \wedge j$ (iv) $(a \vee b) \vee c$ (v) $a \vee (b \vee c)$ (vi) $c \vee d$ (vii) $a \vee (c \vee d)$ (viii) $(a \vee c) \vee d$ (ix) $\bigvee\{a,c,d\}$.

2.4 Repeat Exercise 2.3 for the following elements of the ordered set shown in Figure 2.12.

(i) $\ell \wedge e$ (ii) $(\ell \wedge e) \wedge k$ (iii) $c \vee e$ (iv) $d \vee e$ (v) $c \vee (d \vee e)$ (vi) $\bigvee\{c,d,e\}$ (vii) $(j \wedge \ell) \wedge k$ (viii) $j \wedge (\ell \wedge k)$ (ix) $j \wedge k$.

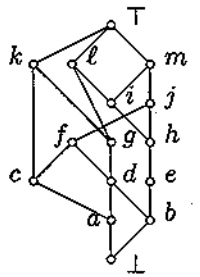


Figure 2.12

2.5 Give an example of an ordered set P in which there are three elements x, y, z such that

(a) $\{x,y,z\}$ is an antichain,

(b) $x \vee y$, $y \vee z$ and $z \vee x$ fail to exist,

(c) $\bigvee\{x,y,z\}$ exists.

(Of course, P will have more than three elements.)

2.6 Let P be an ordered set.

(i) Prove that if $A \subseteq P$ and $\bigwedge A$ exists in P, then

$$\bigcap\{\downarrow a \mid a \in A\} = \downarrow(\textstyle\bigwedge A).$$

(ii) Formulate and prove the dual result.

2.7 In Theorem 2.9 more identities were listed than was necessary. Prove that each of (L3) and $(\mathrm{L3})^\partial$ may be derived from (L4) and $(\mathrm{L4})^\partial$. (Do not use any other identities.)

2.8 Write out the duals of the following statements. (Each is true in all lattices, so its dual is also true in all lattices.)

(i) If z is an upper bound of $\{x,y\}$, then $x \vee y \leqslant z$.

(ii) $a \wedge b \leqslant a \leqslant a \vee b$ and $a \wedge b \leqslant b \leqslant a \vee b$.

(iii) If $a \prec c$ and $b \prec c$, with $a \neq b$, then $a \vee b = c$.

2.9 Let $A = (a_{ij})$ be an $m \times n$ matrix whose entries are elements of a lattice L.

(i) Prove the Mini-Max Theorem, *viz.*

$$\bigvee_{j=1}^{n}\Big(\bigwedge_{i=1}^{m} a_{ij}\Big) \leqslant \bigwedge_{k=1}^{m}\Big(\bigvee_{\ell=1}^{n} a_{k\ell}\Big),$$

that is, (the join of the meets of the columns of A) $\leqslant$ (the meet of the joins of the rows of A). (For the proof, this should be treated as a result about suprema and infima, that is, about a lattice *qua* ordered set. However, its applications, such as those in (ii) and (iii), are primarily to lattices viewed as algebraic structures.)

(ii) By applying (i) to a suitable 2×2 matrix, derive the **distributive inequality**

$$a \wedge (b \vee c) \geqslant (a \wedge b) \vee (a \wedge c).$$

(iii) By applying (i) to a suitable 3×3 matrix, derive the **median inequality**

$$(a \wedge b) \vee (b \wedge c) \vee (c \wedge a) \leqslant (a \vee b) \wedge (b \vee c) \wedge (c \vee a).$$

[Hint. Each row and each column of such a 3×3 matrix must contain exactly two of the elements a, b, c with one repetition.]

2.10 Consider the lattices L_1, L_2 and L_3 in Figure 2.13.

(i) Find L_1 as a sublattice of L_2.

(ii) The shaded elements of L_3 do not form a sublattice. Why?

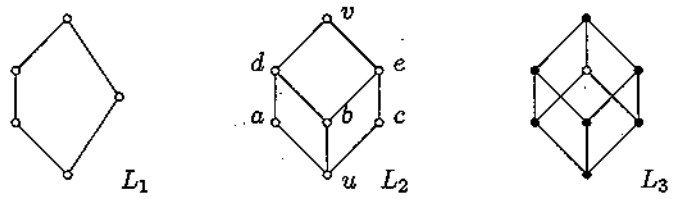


Figure 2.13

2.11 Let L be a lattice. Prove that the following are equivalent:

(i) L is a chain;

(ii) every non-empty subset of L is a sublattice;

(iii) every two-element subset of L is a sublattice.

2.12 (i) Draw a labelled diagram of the lattice $\langle \mathrm{Sub}_0\mathbf{3}; \subseteq \rangle$.

(ii) (a) Find a lattice L such that $\langle \mathrm{Sub}_0 L; \subseteq \rangle$ has the diagram given in Figure 2.14. (Justify your answer by labelling the elements of your guess for L and then drawing a labelled diagram of $\mathrm{Sub}_0 L$.)

(b) Prove that, up to isomorphism, there is only one such lattice L.

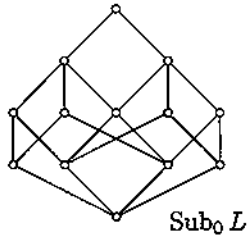


Figure 2.14

2.13 (i) Up to isomorphism there are exactly 10 lattices L for which $1 \leqslant |L| \leqslant 5$. Draw a diagram for each of these lattices.

(ii) Establish the number of sublattices of an n-element chain.

(iii) Prove that, up to isomorphism, there is *at most* one lattice L for which $\langle \mathrm{Sub}_0 L; \subseteq \rangle$ has the diagram given in Figure 2.15.

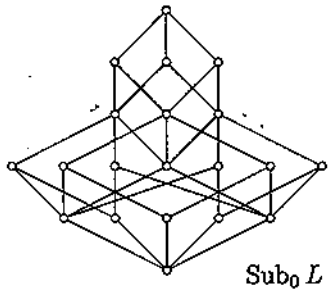


Figure 2.15

2.14 (i) Show that there is no lattice L such that $\mathrm{Sub}_0\, L$ is isomorphic to the lattice $\mathfrak{L}_1$ in Figure 2.16.

(ii) Explain carefully why there are only two possible lattices L such that $\mathrm{Sub}_0\, L$ is given by the lattice $\mathfrak{L}_2$ in Figure 2.16.

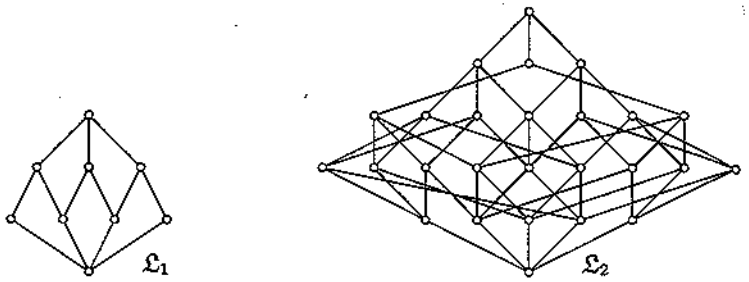


Figure 2.16

2.15 Let L be a lattice and for each $X \subseteq L$ let

$$[X] := \bigcap \{ K \in \mathrm{Sub}_0 L \mid X \subseteq K \}.$$

Show that if X is non-empty, then $[X]$ is the smallest sublattice of L which contains X.

The sublattice $[X]$ is called the **sublattice generated by** X. The definition in terms of set-intersection does not give a viable method for calculating $[X]$ in a finite lattice. Here is an alternative method for obtaining $[X]$ from a diagram of L. Let $\varnothing \neq X \subseteq L$ and define recursively

$$X_0 := X \quad \text{and} \quad X_{k+1} := \{ a \vee b \mid a, b \in X_k \} \cup \{ a \wedge b \mid a, b \in X_k \}.$$

(i) Show that $X_k \subseteq X_{k+1}$ for all $k \in \mathbb{N}_0$.

(ii) Show that $\bigcup \{ X_k \mid k \in \mathbb{N}_0 \}$ is a sublattice of L.

(iii) Show that $[X] = \bigcup \{ X_k \mid k \in \mathbb{N}_0 \}$.

(iv) For each of the lattices in Figure 2.17 take X to be the set of shaded elements. Find $[X]$ in each case. As you proceed, label each element in terms of $\vee, \wedge$ and the generators $a, b, c, \ldots$.

2.16 Draw the product of the lattices $\mathbf{3}$ and $\mathbf{2^2 \oplus 1}$ and shade in elements which form a sublattice isomorphic to $\mathbf{1 \oplus (2 \times 3) \oplus 1}$.

2.17 Let L and K be lattices with 0 and 1 and let $M = L \times K$. Show that there exist $a, b \in M$ such that

(i) $\downarrow a \cong L$ and $\downarrow b \cong K$,

(ii) $a \wedge b = (0,0)$ and $a \vee b = (1,1)$.

Is the lattice given in Figure 2.18 a product of two lattices each with more than one element? (Justify your answer via a careful case-by-case analysis.)

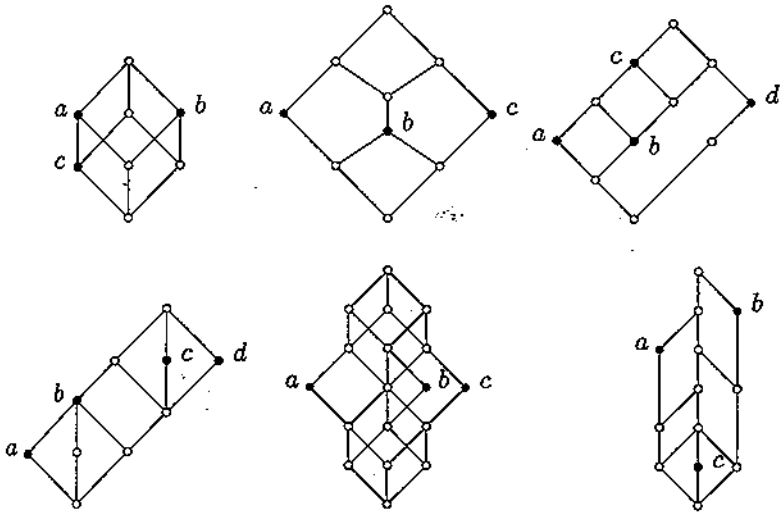


Figure 2.17

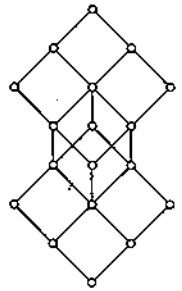


Figure 2.18

2.18 Consider again the maps defined in Exercise 1.22. Which of these maps are (i) join-preserving, (ii) homomorphisms?

2.19 Let $f : L \to K$ be a lattice homomorphism.

(i) Show that if $M \in \mathrm{Sub}\, L$ then $f(M) \in \mathrm{Sub}\, K$.

(ii) Show that if $N \in \mathrm{Sub}\, K$ then $f^{-1}(N) \in \mathrm{Sub}_0 L$.

2.20 For each of the elements $a \in \mathbb{N}_0$ listed below, draw a diagram of the principal ideal $\downarrow a$ of the lattice $\langle \mathbb{N}_0; \preccurlyeq \rangle$:

$$a = 1, 2, 3, 6, 8, 12, 13, 16, 21, 24, 30, 36.$$

2.21 (i) Give an example of a non-principal ideal in the chain $\mathbb{R}$.

(ii) Find all ideals in (a) $\mathbb{Z}$, (b) $\mathbb{Q}$ (with their usual orders). Which of these ideals are principal?

2.22 Let L be a lattice and let $\varnothing \neq A \subseteq L$. Show that

$$(A] := \downarrow\{ a_1 \vee \cdots \vee a_n \mid n \in \mathbb{N},\ a_1, \ldots, a_n \in A \}$$

is an ideal and moreover it is contained in any ideal J of L which contains A. (See also 7.5(4).)

2.23 Let I be an ideal in a lattice L and let $a \in L$. Show that the set $\downarrow\{ a \vee c \mid c \in I \}$ is an ideal and is the smallest ideal in L containing both I and a.

2.24 The set $S = \{ (i,j) \in \mathbb{N} \times \mathbb{N} \mid i < j \}$ is given the order defined by

$$(i,j) \leqslant (i',j') \iff j \leqslant i' \text{ or } (i = i' \ \&\ j \leqslant j').$$

(i) Draw a diagram of S (as a subset of $\mathbb{N} \times \mathbb{N}$).

(ii) Show that S is a lattice having the property that, for any $n \geqslant 3$ and $a_1, \ldots, a_n \in S$, the element $a_1 \wedge a_2 \wedge \cdots \wedge a_n$ can be expressed as the meet of at most two of $a_1, \ldots, a_n$.

(iii) Show that the set of ideals of S, ordered by inclusion, is isomorphic to

$$\{ (i,j) \mid 1 \leqslant i < j \leqslant \infty \} \cup \{(\infty, \infty)\}$$

with the order given by

$$(i,j) \leqslant (i',j') \iff j \leqslant i' \text{ or } (i = i' \ \&\ j \leqslant j').$$

2.25 A subset A of $\mathbb{N}$ is called **cofinite** if $\mathbb{N} \setminus A$ is finite.

(i) Show that the collection $\mathfrak{L}_1$ of cofinite subsets of $\mathbb{N}$ is a lattice of sets.

(ii) Show that the collection $\mathfrak{L}_2$ of subsets of $\mathbb{N}$ which are either finite or cofinite is a lattice of sets.

(iii) Let $A_n := \mathbb{N} \setminus \{2, 4, \ldots, 2n\}$ be obtained from $\mathbb{N}$ by deleting the first n even natural numbers. Show that if $B \subseteq A_n$ for all $n \in \mathbb{N}$ then B is not cofinite and deduce that neither $\mathfrak{L}_1$ nor $\mathfrak{L}_2$ is complete.

2.26 (i) Verify that the linear sum $P \oplus Q$ of (complete) lattices is a (complete) lattice.

(ii) We saw in 2.15 that the product $P \times Q$ of lattices P and Q is again a lattice. Verify that $P \times Q$ is a complete lattice if P and Q are, with joins and meets being formed coordinatewise.

2.27 Assume that P is any set and Q a (complete) lattice.

(i) Show that, under the usual pointwise order, the set Q^P of all maps from P to Q is a (complete) lattice, with the join, φ, of $\{ \varphi_i \mid i \in I \}$ given pointwise by

$$(\forall x \in P)\, \varphi(x) = \bigvee \{ \varphi_i(x) \mid i \in I \},$$

and similarly for meet.

(ii) Assume in addition that P carries an order relation and that all the maps φ_i are order-preserving. Show that the maps $\bigvee\{ \varphi_i \mid i \in I \}$ and $\bigwedge\{ \varphi_i \mid i \in I \}$ are also order-preserving and deduce that $Q^{(P)}$ is a (complete) lattice.

2.28 Let $\langle P; \leqslant \rangle$ be an ordered set. A (possibly empty) subset K of P is called **(order-)convex** if whenever $a, b \in K$ with $a \leqslant b$, every element of P between a and b is also in K; more formally,

$$(\forall a, b \in K)(\forall x \in P)\ (a \leqslant x \leqslant b \implies x \in K).$$

Let $\mathcal{K}(P) := \{ K \subseteq P \mid K \text{ is a convex subset of } P \}$.

(i) Find all the convex subsets of the three-element chain $0 < a < 1$. Draw and label the lattice $\langle \mathcal{K}(\mathbf{3}); \subseteq \rangle$. [Hint. Start with the smallest convex subsets and work up.]

(ii) Prove that $\mathcal{K}(P)$ is a topped intersection structure on P.

2.29 Let P be a complete lattice. Prove that there is a topped $\bigcap$-structure $\mathfrak{L}$ on the set P such that $P \cong \mathfrak{L}$. [Hint. Show that the image of the order-embedding $\varphi\colon P \to \mathcal{O}(P)$ defined by $\varphi(x) = \downarrow x$ is a topped $\bigcap$-structure on P.]

2.30 Draw the subgroup lattice $\operatorname{Sub} G$ and shade in the elements of $\mathcal{N}\text{-}\operatorname{Sub} G$ for each of the following groups: $\mathbf{S}_3$, $\mathbb{Z}_6$, $\mathbb{Z}_{12}$, $\mathbf{A}_4$, $\mathbf{D}_p$ (the symmetries of a regular p-gon) where p is an odd prime, the quaternion group.

2.31 Let H and K be finite groups such that $\gcd(|H|, |K|) = 1$. Show that $\operatorname{Sub}(H \times K) \cong \operatorname{Sub} H \times \operatorname{Sub} K$, where on the left we have the usual coordinatewise product of groups and on the right the coordinatewise product of ordered sets.

2.32 (i) Use the Knaster–Tarski Fixpoint Theorem to prove Banach's Decomposition Theorem:

Let X and Y be sets and let $f\colon X \to Y$ and $g\colon Y \to X$ be maps. Then there exist disjoint subsets X_1 and X_2 of X and disjoint subsets Y_1 and Y_2 of Y such that $f(X_1) = Y_1$, $g(Y_2) = X_2$, $X = X_1 \cup X_2$ and $Y = Y_1 \cup Y_2$.

[Hint. Consider the map $F : \wp(X) \to \wp(X)$ defined by $F(S) = X \setminus g(Y \setminus f(S))$ for $S \subseteq X$.]

(ii) Use (i) to obtain the Schröder–Bernstein Theorem:

Let X and Y be sets and suppose there exist one-to-one maps $f: X \to Y$ and $g: Y \to X$. Then there exists a bijective map h from X onto Y.

2.33 Describe all lattices of length 2, proving that your list is complete.

2.34 Prove that $P \times Q$ satisfies (ACC) if and only if both P and Q do.

2.35 Let P and Q be ordered sets of finite length. Prove that

$$\ell(P \times Q) = \ell(P) + \ell(Q).$$

2.36 (i) Show that $\langle \mathbb{N}_0; \preccurlyeq \rangle$ satisfies (DCC) and deduce that $\langle \mathbb{N}_0; \preccurlyeq \rangle$ is a complete lattice.

(ii) Prove directly that $\bigwedge S$ exists in $\langle \mathbb{N}_0; \preccurlyeq \rangle$ for every non-empty subset S of $\mathbb{N}_0$.

2.37 Let L be a lattice.

(i) Let $J_1 \subseteq J_2 \subseteq \cdots$ be a chain of ideals of L. Show that their union, $\bigcup_{n \in \mathbb{N}} J_n$, is an ideal of L.

(ii) Show that every ideal of L is principal if and only if L satisfies (ACC).

2.38 Let L be a lattice satisfying both (ACC) and (DCC). Let $a, b \in L$. Show that $a \leqslant b$ if and only if, for all join-irreducible elements x and all meet-irreducible elements y, the inequalities $x \leqslant a$ and $b \leqslant y$ together imply $x \leqslant y$.

2.39 Let P be an ordered set and let $Q \subseteq P$. Show that the following are related by (i) $\Leftrightarrow$ (ii) $\Rightarrow$ (iii) in general and are equivalent if P is a complete lattice:

(i) Q is join-dense in P;

(ii) $a = \bigvee_P (\downarrow a \cap Q)$ for all $a \in P$;

(iii) for all $a, b \in P$ with $b < a$ there exists $x \in Q$ with $x \leqslant a$ and $x \not\leqslant b$.

[Hint. To show that (iii) implies (ii) when P is a complete lattice, apply (iii) with $b := \bigvee_P (\downarrow a \cap Q)$.]

3

Formal Concept Analysis

Hierarchies occur often both within mathematics and in the 'real' world and the theory of ordered sets and lattices provides a natural setting in which to discuss and analyse them. In this chapter we take a brief excursion into **formal concept analysis** in order to get a feel for the potential of lattice theory in the analysis of hierarchies of concepts.

Contexts and their concepts

3.1 What is a concept? This would appear to be a question for philosophers rather than for mathematicians. Indeed, traditional philosophy's answer provides us with the basis for our formal definition. A concept is considered to be determined by its extent and its intent: the extent consists of all objects belonging to the concept (as the reader belongs to the concept 'living person') while the intent is the collection of all attributes shared by the objects (as all living persons share the attribute 'can breathe'). As it is often difficult to list all the objects belonging to a concept and usually impossible to list all its attributes, it is natural to work within a specific context in which the objects and attributes are fixed.

3.2 A context for the planets. The information presented in Table 3.1 gives a (somewhat limited) context for the planets of our solar system.

	size			distance from sun		moon	
	small	medium	large	near	far	yes	no
Mercury	×			×			×
Venus	×			×			×
Earth	×			×		×	
Mars	×			×		×	
Jupiter			×		×	×	
Saturn			×		×	×	
Uranus		×			×	×	
Neptune		×			×	×	
Pluto	×				×	×	

Table 3.1

The objects are the planets while the attributes are the seven indicated properties relating to size, distance from the sun and existence of a moon. That the ith object possesses the jth attribute is indicated by a $\times$ in the ij-position of the table. A concept of this context will consist of an ordered pair (A, B), where A (the extent) is a subset of the nine planets and B (the intent) is a subset of the seven properties. To demand that the concept is determined by its extent and by its intent means that B should contain just those properties shared by all the planets in A and, similarly, the planets in A should be precisely those sharing all the properties in B. A simple procedure for finding a concept is as follows: take an object, say the planet Earth, and let B be the set of attributes which it possesses, in this case

$$B = \{\text{size–small, distance–near, moon–yes}\},$$

then let A be the set of all planets possessing all the attributes in B, in this case

$$A = \{\text{Earth, Mars}\}.$$

Then (A, B) is a concept. More generally, we may begin with a set of objects rather than a single object. Concepts may also be obtained via a similar process commencing with a set of attributes.

It is usual to regard a concept (A_1, B_1) as being 'less general' than a concept (A_2, B_2) if the extent A_1 of (A_1, B_1) is contained in the extent A_2 of (A_2, B_2). Thus an order is defined on the set of concepts by

$$(A_1, B_1) \leqslant (A_2, B_2) \iff A_1 \subseteq A_2.$$

The apparent asymmetry in this definition is illusory since $A_1 \subseteq A_2$ is equivalent to $B_1 \supseteq B_2$ (see Lemma 3.5). The resulting ordered set of concepts for our planetary context is the lattice given in Figure 3.1 later in the chapter. With respect to this order, the concept (A, B) constructed above is the smallest concept whose extent contains the planet Earth and is represented by the circle labelled E Ma in Figure 3.1. The significance of the labelling of the lattice will be explained when we return to this example in 3.12.

We now wish to abstract the previous example. The resulting theory has a broad range of applications outside mathematics, for example in the social sciences, as well as having something useful to say within lattice theory itself.

3.3 Contexts and concepts. A **context** is a triple (G, M, I) where G and M are sets and $I \subseteq G \times M$. The elements of G and M are called **objects** and **attributes** respectively. As usual, instead of writing

$(g, m) \in I$ we write gIm and say 'the object g has the attribute m'. (The letters G and M come from the German: Gegenstände and Merkmale.) In finite examples we specify the context in the same manner as in Table 3.1, by means of a **cross-table**.

For $A \subseteq G$ and $B \subseteq M$, define

$$A' = \{m \in M \mid (\forall g \in A)\, gIm\},$$
$$B' = \{g \in G \mid (\forall m \in B)\, gIm\};$$

so A' is the set of attributes common to all the objects in A and B' is the set of objects possessing the attributes in B. Then a **concept** of the context (G, M, I) is defined to be a pair (A, B) where $A \subseteq G$, $B \subseteq M$, $A' = B$ and $B' = A$. The **extent** of the concept (A, B) is A while its **intent** is B. Note that a subset A of G is the extent of some concept if and only if $A'' = A$ in which case the unique concept of which A is an extent is (A, A'). Of course, the corresponding statement applies to those subsets B of M which are the intent of some concept. The maps $': A \mapsto A'$ and $': B \mapsto B'$ are traditionally called the **polars** of the relation $I \subseteq G \times M$. The set of all concepts of the context (G, M, I) is denoted by $\mathfrak{B}(G, M, I)$. (Again the choice of letter comes from the German: $\mathfrak{B}$ for Begriff.)

The framework within which we are working – a pair of sets, G, M, and a binary relation I linking them – is extremely general, and encompasses contexts which might not at first sight be viewed in terms of an object-attribute correspondence. Consider, for example, a computer program modelled by an input-output relation R between a finite set of initial states X and a finite set of final states Y with xRy if and only if the program when started in state x can terminate in state y. Then (X, Y, R) is the context for what is known as a (non-deterministic) transition system. Here A' (for $A \subseteq X$) is to be interpreted as the set of final states in which the program can terminate when started from any one of the states in A.

3.4 The ordering of concepts. Let (G, M, I) be a context. For concepts (A_1, B_1) and (A_2, B_2) in $\mathfrak{B}(G, M, I)$ we write $(A_1, B_1) \leqslant (A_2, B_2)$, if $A_1 \subseteq A_2$. Also, $A_1 \subseteq A_2$ implies that $A_1' \supseteq A_2'$, and the reverse implication is valid too, because $A_1'' = A_1$ and $A_2'' = A_2$. We therefore have

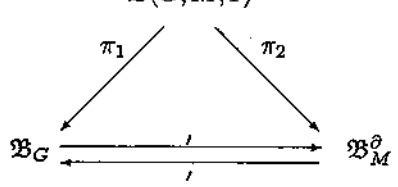
$$(A_1, B_1) \leqslant (A_2, B_2) \iff A_1 \subseteq A_2 \iff B_1 \supseteq B_2.$$

We can then see easily that the relation $\leqslant$ is an order on $\mathfrak{B}(G, M, I)$. As we see in Proposition 3.6, $\langle \mathfrak{B}(G, M, I); \leqslant \rangle$ is a complete lattice; it is known as the **concept lattice** of the context (G, M, I).

Note that each concept (A, B) is uniquely determined by either its first component, A, or by its second component, B, and that the order $\leqslant$ is completely determined by the inclusion order on $\wp(G)$ or equivalently by the reverse inclusion order on $\wp(M)$. This leads us to consider

$$\mathfrak{B}_G := \{ A \subseteq G \mid A'' = A \} \quad \text{and} \quad \mathfrak{B}_M := \{ B \subseteq M \mid B'' = B \},$$

both ordered by inclusion. The map $\pi_1 \colon (A, B) \mapsto A$ gives an order-isomorphism between $\mathfrak{B}(G, M, I)$ and $\mathfrak{B}_G$ while $\pi_2 \colon (A, B) \mapsto B$ gives an order-isomorphism between $\mathfrak{B}(G, M, I)$ and $\mathfrak{B}_M^\partial$. We therefore have a commutative diagram



with the indicated maps setting up order-isomorphisms.

While developing the theory, use of $\mathfrak{B}_G$ and $\mathfrak{B}_M$ allows us some notational simplifications. However in applications, we choose to work with $\mathfrak{B}(G, M, I)$ since we want both the extent and the intent of each concept to be visible.

The claims in the next lemma follow directly from the definitions of A' and B'. The details are left to the reader. The final statement, (P5), tells us that the polar maps $' \colon \wp(G) \to \wp(M)^\partial$ and $' \colon \wp(M)^\partial \to \wp(G)$ set up a Galois connection, as defined in 7.23. We shall see in Chapter 7 that (P1)–(P3) in 3.5 are instances of properties which hold for any Galois connection. The important fact that $\mathfrak{B}(G, M, I)$ is a complete lattice is derived below directly from Lemma 3.5. Alternatively, this result may be seen as coming from the intimate relationship between Galois connections, closure operators and topped $\bigcap$–structures which we reveal in Chapter 7.

3.5 Lemma. *Assume that (G, M, I) is a context and let $A, A_j \subseteq G$ and $B, B_j \subseteq M$, for $j \in J$. Then*

(P1) $A \subseteq A''$ *and* $B \subseteq B''$,

(P2) $A_1 \subseteq A_2 \Longrightarrow A_1' \supseteq A_2'$ *and* $B_1 \subseteq B_2 \Longrightarrow B_1' \supseteq B_2'$,

(P3) $A' = A'''$ *and* $B' = B'''$,

(P4) $\left(\bigcup_{j\in J} A_j\right)' = \bigcap_{j\in J} A_j'$ *and* $\left(\bigcup_{j\in J} B_j\right)' = \bigcap_{j\in J} B_j'$,

(P5) $A \subseteq B' \Longleftrightarrow A' \supseteq B$.

3.6 Proposition. *Let (G, M, I) be a context. Then $(\mathfrak{B}(G, M, I); \leqslant)$ is a complete lattice in which join and meet are given by*

$$\bigvee_{j\in J} (A_j, B_j) = \left(\left(\bigcup_{j\in J} A_j\right)'', \bigcap_{j\in J} B_j \right),$$

$$\bigwedge_{j\in J} (A_j, B_j) = \left(\bigcap_{j\in J} A_j, \left(\bigcup_{j\in J} B_j\right)'' \right).$$

Proof. We shall prove that $\mathfrak{B}_G$ is a topped $\bigcap$-structure. Let $A_j \in \mathfrak{B}_G$ for $j \in J$. Then $A_j'' = A_j$ for each j. By (P1) in Lemma 3.5,

$$\bigcap_{j\in J} A_j \subseteq \left(\bigcap_{j\in J} A_j\right)''.$$

Also $\bigcap_{j\in J} A_j \subseteq A_k$ for all $k \in J$, which, by (P2), implies that

$$\left(\bigcap_{j\in J} A_j\right)'' \subseteq A_k'' = A_k \text{ for all } k \in J,$$

whence

$$\left(\bigcap_{j\in J} A_j\right)'' \subseteq \bigcap_{k\in J} A_k = \bigcap_{j\in J} A_j.$$

Therefore $\left(\bigcap_{j\in J} A_j\right)'' = \bigcap_{j\in J} A_j$ and hence $\bigcap_{j\in J} A_j \in \mathfrak{B}_G$. Also, $G \subseteq G''$ so that necessarily $G = G''$, which shows that $\mathfrak{B}_G$ is topped.

By 2.32, $\mathfrak{B}_G$ is a complete lattice in which meet is given by intersection. A formula for the join is given in 2.32 but we shall proceed more directly. We claim that

$$\bigvee_{j\in J} A_j = \left(\bigcup_{j\in J} A_j\right)''.$$

Let A be the set on the right-hand side. Certainly $A'' = A$, by (P3), and $\bigcup_{j\in J} A_j \subseteq A$, by (P1). Hence A is an upper bound for $\{A_j\}_{j\in J}$ in $\mathfrak{B}_G$. Also, if Y is any upper bound in $\mathfrak{B}_G$ for $\{A_j\}_{j\in J}$, then

$$\bigcup_{j\in J} A_j \subseteq Y \Longrightarrow \left(\bigcup_{j\in J} A_j\right)'' \subseteq Y'' = Y.$$

Therefore A is indeed the required join. We may now appeal to 3.4, 2.27(ii) and 3.5 to deduce that $\mathfrak{B}(G, M, I)$ is a complete lattice in which joins and meets are given by the stated formulae. (Alternatively it can be verified directly that these formulae do indeed provide the least upper bound and greatest lower bound of $\{(A_j, B_j)\}_{j\in J}$ in $\mathfrak{B}(G, M, I)$.) □

The fundamental theorem of concept lattices

Before giving a range of examples of contexts and their concept lattices, including a discussion of the planets example, we present the theory that enables us to analyse concept lattices. We begin by relating the cross-table to the lattice.

3.7 Relating G and M to $\mathfrak{B}(G,M,I)$. Let (G,M,I) be a context and $\mathfrak{B}(G,M,I)$ the associated set of concepts. We define $\gamma\colon G \to \mathfrak{B}(G,M,I)$ and $\mu\colon M \to \mathfrak{B}(G,M,I)$ by

$$\gamma(g) := (\{g\}'', \{g\}') \quad \text{and} \quad \mu(m) := (\{m\}', \{m\}'')$$

for all $g \in G$ and $m \in M$. Note that (P3) implies that $\gamma(g)$ and $\mu(m)$ are indeed concepts. The set $\{m\}'$ is called an **attribute-extent**. To simplify notation we shall henceforth drop the braces, and write g' in place of $\{g\}'$ and m' in place of $\{m\}'$.

The maps γ and μ encode the relation I within the ordered set $\mathfrak{B}(G,M,I)$ of concepts. Indeed, gIm is equivalent to $\gamma(g) \leqslant \mu(m)$ as the following simple calculation shows. Let $g \in G$ and $m \in M$; then

$$\begin{aligned} gIm &\iff g \in m' \\ &\iff g'' \subseteq m''' = m' \\ &\iff \gamma(g) \leqslant \mu(m). \end{aligned}$$

(The forward direction of the second equivalence above uses (P2) and (P3) while the backward implication uses (P1).)

Recall from 2.42 that a subset Q of an ordered set P is called **join-dense** if every element of P is the join of a subset of Q. **Meet-dense** is defined dually.

3.8 Theorem. *Let (G,M,I) be a context and $L = \mathfrak{B}(G,M,I)$ the associated complete lattice of concepts. Then the mappings $\gamma\colon G \to L$ and $\mu\colon M \to L$ are such that the set $\gamma(G)$ is join-dense in L, the set $\mu(M)$ is meet-dense in L, and gIm is equivalent to $\gamma(g) \leqslant \mu(m)$ for each $g \in G$ and $m \in M$.*

Proof. Let $(A,B) \in \mathfrak{B}(G,M,I)$. Then

$$\begin{aligned} \bigvee \gamma(A) &= \bigvee_{g\in A} \gamma(g) \\ &= \bigvee_{g\in A} (g'', g') \\ &= \Big((\bigcup_{g\in A} g'')'', \bigcap_{g\in A} g' \Big). \end{aligned}$$

But, by (P4),

$$\bigcap_{g\in A} g' = \big(\bigcup_{g\in A} \{g\}\big)' = A' = B.$$

Since (A,B) and $\bigvee\gamma(A)$ are elements of $\mathfrak{B}(G,M,I)$ with the same second coordinate, $\bigvee\gamma(A) = (A,B)$. Similarly, we find $\bigwedge\mu(B) = (A,B)$. Consequently $\gamma(G)$ is join-dense and $\mu(M)$ is meet-dense in L. We have already proved in 3.7 that gIm if and only if $\gamma(g) \leqslant \mu(m)$ in $\mathfrak{B}(G,M,I)$, for all $g \in G$ and $m \in M$. □

We now turn things round and prove that every complete lattice arises, in a generally non-unique way, as a concept lattice.

3.9 Theorem. *Let L be a complete lattice, let G and M be sets and assume that there exist mappings $\gamma\colon G \to L$ and $\mu\colon M \to L$ such that $\gamma(G)$ is join-dense in L and $\mu(M)$ is meet-dense in L. Define I by $gIm \Leftrightarrow \gamma(g) \leqslant \mu(m)$, for all $g \in G$ and $m \in M$. Then L is isomorphic to $\mathfrak{B}(G,M,I)$.*

In particular, any complete lattice L is isomorphic to the concept lattice $\mathfrak{B}(L,L,\leqslant)$.

Proof. We first prove that, for all $A \subseteq G$ and $m \in M$,

$$m \in A' \iff \bigvee\gamma(A) \leqslant \mu(m).$$

Indeed,

$$\begin{aligned} m \in A' &\iff (\forall g \in A)\, gIm \\ &\iff (\forall g \in A)\, \gamma(g) \leqslant \mu(m) && \text{(by definition)} \\ &\iff \mu(m) \text{ is an upper bound of } \gamma(A) \\ &\iff \bigvee\gamma(A) \leqslant \mu(m). \end{aligned}$$

Dually, we have $g \in B' \iff \gamma(g) \leqslant \bigwedge \mu(B)$ for $B \subseteq M$ and $g \in G$.

We are now ready to set up an order-isomorphism between L and $\mathfrak{B}_G$. Recalling 1.36(4), we shall do this by defining a pair of mutually inverse order-preserving maps $\varphi\colon \mathfrak{B}_G \to L$ and $\psi\colon L \to \mathfrak{B}_G$. Define φ by $\varphi(A) = \bigvee\gamma(A)$ for all $A \in \mathfrak{B}_G$. Since, for all $A_1, A_2 \in \mathfrak{B}_G$,

$$\begin{aligned} A_1 \subseteq A_2 &\Longrightarrow \gamma(A_1) \subseteq \gamma(A_2) \\ &\Longrightarrow \bigvee\gamma(A_1) \leqslant \bigvee\gamma(A_2) && \text{(by 2.22(v))}, \end{aligned}$$

the map φ is order-preserving. Let $x \in L$. Define

$$A_x := \{\, g \in G \mid \gamma(g) \leqslant x \,\} \text{ and } B_x := \{\, m \in M \mid x \leqslant \mu(m) \,\}.$$

For all $m \in M$ we have

$$\begin{aligned} m \in A'_x &\iff \bigvee \gamma(A_x) \leqslant \mu(m) && \text{(from above)} \\ &\iff x \leqslant \mu(m) && \text{(as } \gamma(G) \text{ is join-dense)} \\ &\iff m \in B_x. \end{aligned}$$

Consequently $A'_x = B_x$ and dually (since $\mu(M)$ is meet-dense in L) we find $B'_x = A_x$. Thus $A_x \in \mathfrak{B}_G$. We may now define $\psi : L \to \mathfrak{B}_G$ by $\psi(x) = A_x$ for all $x \in L$. Since $x \leqslant y$ implies $A_x \subseteq A_y$, the map ψ is order-preserving. Clearly

$$\varphi(\psi(x)) = \varphi(A_x) = \bigvee \gamma(A_x) = x,$$

for all $x \in L$, since $\gamma(G)$ is join-dense in L. Now let $A \in \mathfrak{B}_G$; we shall prove that $\psi(\varphi(A)) = A$. Let $x := \varphi(A) = \bigvee \gamma(A)$; we wish to show that $A_x = A$. For all $g \in A$ we have $\gamma(g) \leqslant \bigvee \gamma(A) = x$ and therefore $g \in A_x$. Hence $A \subseteq A_x$. In the other direction,

$$\begin{aligned} g \in A_x &\Longrightarrow \gamma(g) \leqslant x = \bigvee \gamma(A) \\ &\Longrightarrow (\forall m \in A')\, \gamma(g) \leqslant \mu(m) && \text{(as } (\forall m \in A') \textstyle\bigvee \gamma(A) \leqslant \mu(m)) \\ &\Longrightarrow (\forall m \in A')\, gIm && \text{(by assumption)} \\ &\Longrightarrow g \in A'' = A && \text{(since } A \in \mathfrak{B}_G). \end{aligned}$$

Thus $A_x \subseteq A$ and so $A_x = A$, as required.

Consequently, φ and ψ are order-preserving and mutually inverse, whence $\mathfrak{B}(G, M, I)$ is order-isomorphic to L by 1.36(4).

Finally, given a complete lattice L, we can choose $G = M = L$, and define both γ and μ to be id_L (the identity map on L). The conditions of the first part of the theorem are clearly satisfied. Since I equals $\leqslant$, it follows that $L \cong \mathfrak{B}(L, L, \leqslant)$. □

3.10 Summing up. Proposition 3.6 and Theorem 3.8 tell us that the concepts of a context (G, M, I) form a complete lattice $\mathfrak{B}(G, M, I)$ into which G and M map in the manner described above.. Theorem 3.9 supplies the converse. This bevy of results completely characterizes concept lattices, and is known collectively as the **fundamental theorem of concept lattices**.

We conclude this section by presenting some simple inbred examples of contexts and their concept lattices from lattice theory itself.

3.11 Lattice-theoretic examples.

(1) Our first example amplifies the last part of Theorem 3.9. Let L be a complete lattice. Then

$$\mathfrak{B}(L, L, \leqslant) = \{ (\downarrow x, \uparrow x) \mid x \in L \}.$$

The theorem asserts that $L \cong \mathfrak{B}(L, L, \leqslant)$. The isomorphism is given by $x \mapsto (\downarrow x, \uparrow x)$. This map is an order-isomorphism by Lemma 1.30 and its dual and so a lattice isomorphism by Proposition 2.19(ii). In 7.38 we explore $\mathfrak{B}(P, P, \leqslant)$ for an ordered set P.

(2) For a given complete lattice L, the context $(L, L, \leqslant)$ is not the only one whose concept lattice is isomorphic to L, nor is it necessarily the most natural one. Consider the complete lattice $\mathcal{P}(X)$ for some set X. While the concept lattice of the context $(\mathcal{P}(X), \mathcal{P}(X), \subseteq)$ is isomorphic to $\mathcal{P}(X)$, both of the following concept lattices have the property that $\mathfrak{B}_G$ (as defined in 3.4) actually equals $\mathcal{P}(X)$.

(i) $\mathfrak{B}(X, X, \neq) = \{ (A, X \setminus A) \mid A \subseteq X \} \cong \mathfrak{B}_G = \mathcal{P}(X)$.

(ii) $\mathfrak{B}(X, \mathcal{P}(X), \in) = \{ (A, \{ B \in \mathcal{P}(X) \mid A \subseteq B \}) \mid A \subseteq X \} \cong \mathfrak{B}_G = \mathcal{P}(X)$.

(3) In contrast to (2)(i) above, we have

$$\mathfrak{B}(X, X, =) = \{ (\{x\}, \{x\}) \mid x \in X \} \cup \{ (\varnothing, X), (X, \varnothing) \}.$$

Hence, if $|X| = n$, then $\mathfrak{B}(X, X, =) \cong \mathfrak{B}_G \cong \mathbf{M}_n$.

(4) Let P be an ordered set and consider the context $(G, M, \not\geqslant)$ where $G = M = P$. It is easy to check the following.

(i) $g' = P \setminus \downarrow g$ and $m' = P \setminus \uparrow m$ for $g \in G$ and $m \in M$.

(ii) For $A \subseteq G$ and $B \subseteq M$ we have $A' := P \setminus \downarrow A$ and $B' := P \setminus \uparrow B$.

For $A \subseteq G$ we now have

$$A'' = P \setminus \uparrow(P \setminus \downarrow A) = P \setminus (P \setminus \downarrow A) = \downarrow A$$

because $P \setminus \downarrow A$ is an up-set. So, for $A, B \subseteq P$, we have that (A, B) is a concept if and only if $A \in \mathcal{O}(P)$ and $B = P \setminus A$. It follows that $\mathfrak{B}(P, P, \not\geqslant)$ is isomorphic to the down-set lattice $\mathcal{O}(P)$ of P.

(5) Let L be a lattice with no infinite chains, in particular any finite lattice. Then, by Theorem 2.41 (or just Corollary 2.25 if L is finite), L is complete and, by Theorem 2.46, the subsets $\mathcal{J}(L)$ of join-irreducible elements and $\mathcal{M}(L)$ of meet-irreducible elements are join-dense and meet-dense in L, respectively. Thus Theorem 3.9 yields $L \cong \mathfrak{B}(\mathcal{J}(L), \mathcal{M}(L), \leqslant)$ with the isomorphism given by

$$x \mapsto (\downarrow x \cap \mathcal{J}(L), \uparrow x \cap \mathcal{M}(L)).$$

From theory to practice

In this section we apply our theory to non-mathematical examples. As we have indicated already, concept lattices are a valuable tool in analysing a data set (the objects) in terms of the properties (the attributes) of its members. The examples which follow are chosen to illustrate this. But it should be remembered that contexts may be used as mathematical models in a wide variety of settings, both mathematical and non-mathematical, and may be analysed with the aid of the fundamental theorem in the same manner as the examples we present.

3.12 Returning to the planets. The concept lattice of the planetary context given in Table 3.1 and considered earlier in 3.2 is presented in Figure 3.1. The labels indicate the mappings γ and μ of the fundamental theorem, that is, the circle labelled g represents the concept (g'', g') for all $g \in G$ and similarly the circle labelled m represents the concept (m', m'') for each $m \in M$. The map $\psi\colon L \mapsto \mathfrak{B}_G$ defined in the proof of the fundamental theorem allows us to read off the extent and intent of each concept from the labelling: if x is an element of the lattice, then the proof of Theorem 3.9 tells us that the corresponding concept is

$$(A_x, B_x) := (\{\, g \in G \mid \gamma(g) \leqslant x \,\}, \{\, m \in M \mid x \leqslant \mu(m) \,\}).$$

For example, the element in the middle of the diagram corresponds to the concept

({Earth, Mars, Pluto}, {size–small, moon–yes}).

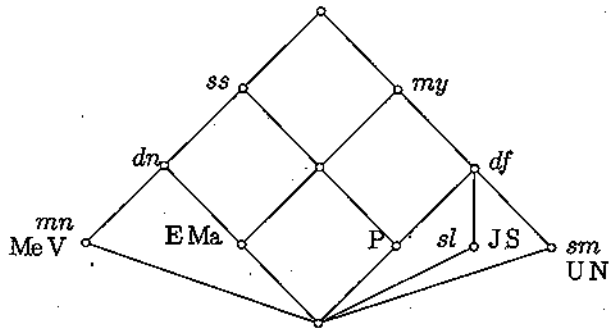


Figure 3.1

The concept lattice provides a basic analysis of a context: it yields an appropriate classification of the objects and at the same time indicates

the implications between the attributes. For the concept lattice of a context to be of practical use, we must be able to *determine* which pairs (A, B), with $A \subseteq G$ and $B \subseteq M$, are concepts of the context, and we must then be able to *describe* the resulting lattice of concepts.

3.13 The determination problem. A simple-minded and extremely inefficient way of determining all the concepts of a context (G, M, I) would be to form (A'', A') for all $A \subseteq G$ (or (B', B'') for all $B \subseteq M$). Here is an efficient alternative: for $A \subseteq G$ and $B \subseteq M$,

$$B' = \bigcap_{m \in B} m' \quad \text{and} \quad A' = \bigcap_{g \in A} g'.$$

In particular, if (A, B) is a concept of the context (G, M, I), then

$$A = \bigcap_{m \in B} m' \quad \text{and} \quad B = \bigcap_{g \in A} g'.$$

For example, in the context for the planets considered in 3.2 and 3.12, we first determine the extents

$$ss', sm', sl', dn', df', my' \text{ and } mn'$$

and then obtain all other extents by forming intersections. The intent corresponding to each extent is then easily calculated. Alternatively, we first determine all the intents and then the corresponding extents.

Once we have found all of the concepts and drawn the lattice, Theorem 3.9 allows us to confirm that our diagram and its labelling are correct. Since L is finite, by Theorem 2.46(iv) the set $\gamma(G)$ is join-dense in L if and only if $\mathcal{J}(L) \subseteq \gamma(G)$. Thus we need to check that each join-irreducible element has a label $g \in G$. Dually, we must check that each meet-irreducible element has a label $m \in M$. In the case of the concept lattice of the planetary context it is easily seen that $\mathcal{J}(L) = \gamma(G)$ and $\mathcal{M}(L) = \mu(M)$; see Figure 3.1. (The observant reader may have noticed that the dual of the lattice in Figure 3.1 occurred back in Figure 2.10 with its join-irreducible elements shaded.) Finally, Theorem 3.9 tells us that we must confirm that gIm holds in the context if and only if $\gamma(g) \leqslant \mu(m)$ holds in the lattice. We do this by checking that, for all $m \in M$, the set of object labels in $\downarrow m$ is precisely the attribute-extent m'. For example, comparing the lattice in Figure 3.1 with the cross-table in Table 3.1 we see that the set of object labels in $\downarrow dn$ is {Me, V, E, Ma} which agrees with the attribute-extent dn', as required.

We now convert the preceding observations into a systematic method for generating all the concepts of a context and for drawing the concept lattice.

3.14 An algorithm for drawing concept lattices. Assume we have the cross-table of a context with the object set G down the side and the attribute set M across the top. The following instructions will generate a list of the extents of all concepts of the context in an order which is convenient for drawing the lattice of all concepts.

Step 1. Find all extents of the concepts of the context (G, M, I).

(1.1) Draw up a table with two columns headed **Attributes** and **Extents**. Leave the first cell of the Attributes column empty and write G in the first cell of the Extents column.

(1.2) Find a maximal attribute-extent, say m'.

(1.2.1) If the set m' is not already in the Extents column, add the row $[m \mid m']$ to the table. Intersect the set m' with all previous extents in the Extents column. Add these intersections to the Extents column (unless they are already in the list) and leave the corresponding cells in the Attribute column empty.

(1.2.2) If the set m' is already in the Extents column, add the label m to the attribute cell of the row where m' previously occurred.

(1.3) Delete the column below m from the table.

(1.4) If the last column has been deleted, stop, otherwise return to (1.2).

Step 2. Draw the diagram with m and m' labels.

Start at the top of the diagram with one point labelled G. Work down the list of Extents in the table from Step 1. For each set S in the list, add an appropriately positioned new point to the diagram. Below the point corresponding to S list the elements in S. If S is an attribute-extent, say $S = m'$, add the label m above the point corresponding to S.

Step 3. Redraw the diagram with g and m labels.

(3.1) Redraw the diagram. Add the m labels as in the first diagram.

(3.2) For each object g in G, add a label g below the point on the diagram which has the smallest extent containing the object g. (This can be found from the first diagram.) Alternatively, the point to be labelled g can be obtained by finding the point $\bigwedge\{m \mid gIm\}$ – find the set from the g-row of the cross-table.

Step 4. Use the fundamental theorem to check the answer.

(4.1) Check that every join-irreducible element has a label $g \in G$.

(4.2) Check that every meet-irreducible element has a label $m \in M$.

(4.3) Check that $(\forall g \in G)(\forall m \in M)\ gIm \iff g \leqslant m$ by checking that, for all $m \in M$, the set of object labels in $\downarrow m$ is exactly the attribute-extent m'.

Hint. When drawing the lattice, try to minimize the number of slopes and to use parallelograms whenever possible.

3.15 Example. The tables and lattices in Figure 3.2 illustrate the algorithm at work. Observe that initially there are three maximal columns: a', b' and d'. Once the column a' has been removed from the table the columns e' and f' also become maximal and the algorithm would allow us to deal with any one of b', d', e' and f' next. Nevertheless, it helps when drawing the lattice if we deal with the 'old' maximal columns first before considering the 'new' maximal columns. We have indicated this via bold horizontal lines in the Attributes/Extents table. Note that, in accordance with Step (1.2.2), the row [b, d | STUW] corresponds to the fact that $b' = d' = \{\text{STUW}\}$. The row [g | U] was struck out after it was noticed that U already occurred in the list and, again in accordance with Step (1.2.2), the label g was added to the Attribute cell of the previous occurrence.

	a	b	c	d	e	f	g
S	×	×		×		×	
T	×	×		×	×		
U	×	×		×	×	×	×
V	×		×		×	×	
W		×		×			
X	×					×	

↓

Attributes	Extents
	G
a	STUVX
b, d	STUW STU
e	TUV TU
f g (added later)	SUVX SU UV U
c	V Ø
~~g~~	~~U~~

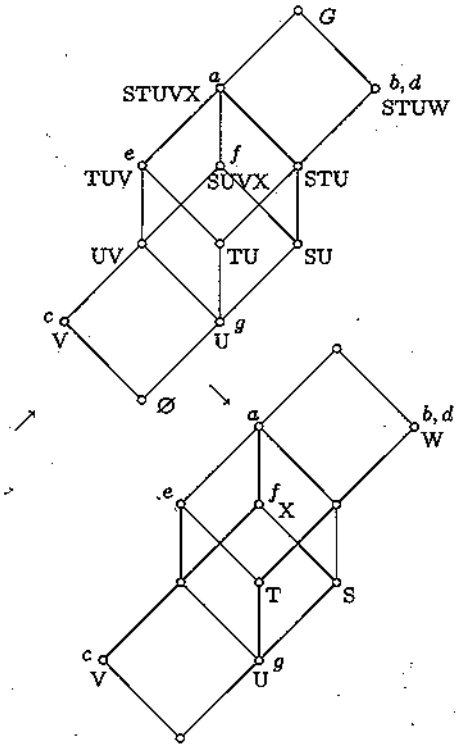


Figure 3.2

3.16 A psychometric test. The context given in Table 3.2 shows the characteristics attributed by a patient to his relatives as well as his own ideal. (For typographical reasons the cross-table has been given with the objects listed horizontally and the attributes vertically.)

	Self	My Ideal	Father	Mother	Sister	Brother-in-Law
vulnerable	×	×	×	×	×	
reserved	×		×	×	×	
self-confident	×	×				×
dutiful		×	×	×	×	×
happy	×	×	×	×	×	×
difficult	×		×	×	×	
attentive	×	×	×		×	×
easily offended			×	×		
not hot-tempered	×	×	×	×	×	
anxious	×		×	×	×	
talkative						×
superficial			×	×		×
sensitive	×	×	×	×	×	
ambitious	×	×	×	×	×	×

Table 3.2

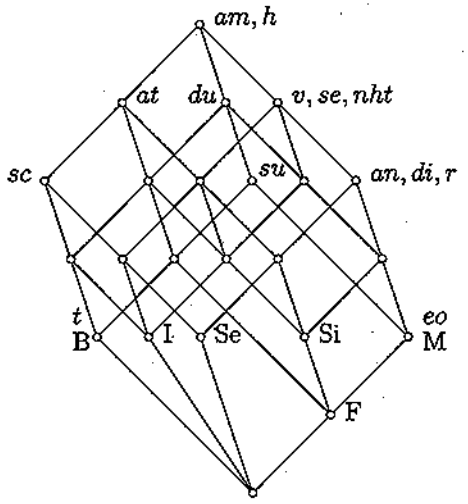


Figure 3.3

The concept lattice (see Figure 3.3) is intended in this case to aid therapists in their analysis of the patient's responses. Consider, for example, the concepts $(\{\text{Se, F, M, Si}\}, \{an, di, r, v, se, nht, am, h\})$ and $(\{\text{I, Se, F, M, Si}\}, \{v, se, nht, am, h\})$. They indicate that the patient regards his immediate family (Self, Father, Mother and Sister) as being

anxious, difficult and reserved and that these are precisely the attributes which distinguish his immediate family from his Ideal Self. We suggest that the reader do some of the simpler exercises for this chapter before working out the Attributes/Extents table for this context and drawing the lattice. Nevertheless, it would be an instructive task to carry out Step 3 right now and so confirm that lattice is as shown in Figure 3.3.

3.17 The description problem. The most natural and informative description of the concept lattice $\mathfrak{B}(G, M, I)$ is a well-drawn diagram labelled by (names for) the elements of G and M. If the number of concepts is small, then the algorithm may be applied and the task may be done by hand. As the size of the context and its lattice of concepts grows, the need for appropriate computer software becomes more apparent. The theory of 'readable diagrams' and computer-based implementations of such a theory, which was in its infancy when the first edition of this book was published, has now grown into a flourishing industry. (See B. Ganter and R. Wille [48].)

Exercises

Exercises from the text. Prove Lemma 3.5. Find all concepts for the context of the psychometric test given in 3.16 and verify that the lattice given in Figure 3.3 is correct. Confirm the claims made in 3.11(4).

3.1 Let (G, M, I) be a context and let $A \subseteq G$ and $B \subseteq M$. Show that the following are equivalent:

(a) $A \subseteq B'$; (b) $B \subseteq A'$; (c) $A \times B \subseteq I$.

3.2 Draw and label the concept lattice for each of the contexts below. In each case the incidence relation I is given in Table 3.3 with G listed vertically and M horizontally.

(1) Large German cities.

$G = \{\text{Hamburg, München, Köln}\}$,

$M = \{\geqslant \textbf{1.5 million}, \geqslant \textbf{1.25 million}, \geqslant \textbf{1.0 million}\}$.

(2) Triangular shapes.

$G = \{\text{Setsquare, Give Way Sign, Delta}\}$,

$M = \{\text{rightangled, isosceles, equilateral}\}$.

(3) Primary colour decomposition.

$G = \{\text{Orange, Green, Violet}\}$,

$M = \{\text{blue, red, yellow}\}$.

(4) Temperatures.

$G = \{\text{Cold, Tepid, Warm}\}$,

$M = \{\leqslant 10°\text{C}, \leqslant 20°\text{C}, \geqslant 10°\text{C}, \geqslant 20°\text{C}\}$.

(5) Watercourses.

$G = \{\text{Channel, Brook, Stream, River}\}$,

$M = \{\text{very small, small, large, very large}\}$.

(6) Family.

$G = \{\text{Father, Mother, Son, Daughter}\}$,

$M = \{\text{old, young, male, female}\}$.

	⩾1.5m	⩾1.25m	⩾1.0m
H	×	×	×
M		×	×
K			×

Context (1)

	r	i	e
S	×		
G		×	×
D		×	

Context (2)

	b	r	y
O		×	×
G	×		×
V	×	×	

Context (3)

	⩽10°	⩽20°	⩾10°	⩾20°
C	×	×		
T		×	×	
W			×	×

Context (4)

	vs	s	l	vl
C	×	×		
B		×		
S			×	
R			×	×

Context (5)

	o	y	m	f
F	×		×	
M	×			×
S		×	×	
D		×		×

Context (6)

Table 3.3

3.3 A student presented the lattices L_1 and L_2 shown in Figure 3.4 as concept lattices with object labels A, B, C, D, E, F, G and attribute labels 1, 2, 3, 4, 5, 6, 7. One is correct and one is wrong.

(i) Decide which is incorrect and explain why.

(ii) For the correctly labelled concept lattice, write down the cross-table of the context which gave rise to this concept lattice.

3.4 Find all concepts of the planetary context discussed in 3.2 and 3.12 and verify that the lattice given in Figure 3.1 is correct.

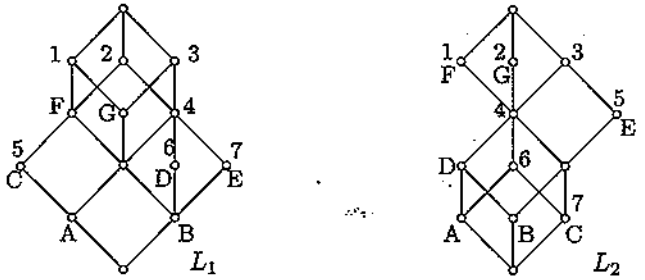


Figure 3.4

3.5 Consider a context having the cross-table shown in Table 3.4.

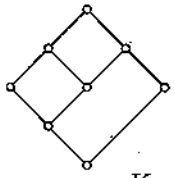
(i) Find all concepts of this context.

(ii) Draw and label the lattice of concepts of this context.

	s	t	u	v	w
A		×	×		
B	×			×	
C		×		×	×
D		×	×	×	×

Table 3.4

3.6 Consider the lattice K in Figure 3.5. Complete the cross-table for a context (G, M, I), with $G = \{\text{A, B, C, D}\}$ and $M = \{\text{m, n, o, p}\}$, such that K is isomorphic to $\mathfrak{B}(G, M, I)$.



K

	m	n	o	p
A				
B				
C				
D				

Figure 3.5

3.7 Consider the context (G, M, I) and lattice L shown in Figure 3.6. Label the lattice to indicate the mappings γ and μ given that L is the concept lattice $\mathfrak{B}(G, M, I)$.

3.8 Consider the context (G, M, I) shown in Table 3.5. Find the 15 concepts and draw a labelled diagram of $\mathfrak{B}(G, M, I)$. Show that no relation strictly smaller than I yields the same concept lattice.

	p	q	r	s	t	u	v
P	×	×	×	×	×	×	
Q	×	×	×	×	×		×
R	×	×	×	×		×	×
S	×	×	×				
T	×	×			×		
U	×		×			×	
V		×	×				×

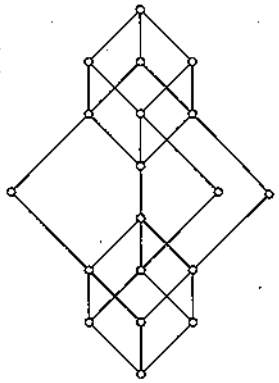


Figure 3.6

	a	b	c	d	e	f	g	h
A	×			×	×			×
B	×	×					×	×
C	×		×			×		×
D	×	×	×	×				
E	×	×			×	×		
F	×		×		×		×	

Table 3.5

3.9 Consider the task-information context given in Table 3.6 which comes from the German national division of health and welfare. The context shows which personal information is required by law for the different tasks of a local medical subdivision. The tasks (objects) and information (attributes) are as follows:

TASKS
1. confirmation of request for preventative care;
2. calculation of insurance benefits;
3. confirmation of incapacity to work;
4. confirmation of correct diagnosis;
5. work preliminary to rehabilitation;
6. verification of sickness;
7. advice to clients;
8. advice on general preventatives;
9. updating the statistics of the medical service.

INFORMATION
a. name and address of client;
b. career history;
c. kind of membership;

d. name of responsible agency;
e. family medical history;
f. vocational education;
g. number of certificates.

	a	b	c	d	e	f	g
1	×	×	×	×	×	×	
2	×	×	×	×	×	×	
3	×	×	×	×	×	×	
4	×	×	×	×			
5	×	×	×	×	×	×	
6	×		×	×		×	
7	×	×	×		×	×	×
8							
9			×				×

Table 3.6

Find the 13 concepts of this context then draw and label the resulting concept lattice.

3.10 The concept lattice of a task-information context (G, M, I) like the one in the previous example provides a natural hierarchical classification of the tasks and indicates the dependencies between the information. In some situations it may be more important to know which information is not required for a particular task; for example, if we wish to restrict access to personal or classified information. Then it is more appropriate to work with the **complementary context** $(G, \overline{M}, \overline{I})$ where

$$\overline{M} := \{\, \overline{m} \mid m \in M \,\} \text{ and } g\overline{I}\overline{m} \iff gIm \text{ is false.}$$

(Here $\overline{m}$ is a new symbol to be thought of as 'not m'.) Consider the information-task context in the previous exercise. Find the 11 concepts of the complementary context, then draw and label the resulting lattice $\mathfrak{B}(G, \overline{M}, \overline{I})$.

3.11 Let G and M be finite sets.

(i) Assume that $\langle G; \leqslant \rangle$ and $\langle M; \leqslant \rangle$ are chains and $I \subseteq G \times M$ is a down-set of the product. Show that $\mathfrak{B}(G, M, I)$ is a chain.

(ii) Assume that (G, M, I) is a context such that $L := \mathfrak{B}(G, M, I)$ is a chain. Show that G and M may be linearly ordered such that I becomes a down-set of $G \times M$. [Hint. Choose orders on G and M so that $\langle G; \leqslant \rangle$ and $\langle M; \leqslant \rangle$ are chains and $\gamma\colon G \to L$ is order-preserving while $\mu\colon M \to L$ is order-reversing.]

3.12 The **horizontal sum** of two bounded ordered sets P and Q is obtained from their disjoint union $P \dot\cup Q$ by identifying the bottoms of the two ordered sets and also identifying the tops. The **vertical sum** of P and Q is obtained from the linear sum $P \oplus Q$ by identifying the top of P with the bottom of Q.

Let (G_1, M_1, I_1) and (G_2, M_2, I_2) be contexts such that $G_1 \cap G_2 = M_1 \cap M_2 = \varnothing$ and $G_i' = M_i' = \varnothing$ for $i = 1, 2$ and let $L_i := \mathfrak{B}(G_i, M_i, I_i)$ for $i = 1, 2$. Prove the following claims.

(i) $\mathfrak{B}(G_1 \dot\cup G_2, M_1 \dot\cup M_2, I_1 \dot\cup I_2)$ is isomorphic to the horizontal sum of L_1 and L_2.

(ii) $\mathfrak{B}(G_1 \dot\cup G_2, M_1 \dot\cup M_2, I_1 \dot\cup I_2 \dot\cup (G_1 \times M_2))$ is isomorphic to the vertical sum of L_1 and L_2.

(iii) $\mathfrak{B}(G_1 \dot\cup G_2, M_1 \dot\cup M_2, I_1 \dot\cup I_2 \dot\cup (G_1 \times M_2) \dot\cup (G_2 \times M_1))$ is isomorphic to the direct product of L_1 and L_2.

3.13 Use the previous exercise to find the concept lattice for each of the contexts given in Table 3.7. In each case draw the lattice.

	a	b	c	d	e	f
A		×	×	×	×	×
B			×	×	×	×
C		×		×	×	×
D	×	×	×		×	×
E	×	×	×			×
F	×	×	×			

Context (1)

	a	b	c	d	e	f
A		×	×	×	×	×
B	×		×	×	×	×
C				×	×	×
D			×		×	×
E				×		×
F					×	

Context (2)

	a	b	c	d	e	f	g	h
A		×	×	×				
B	×		×	×				
C				×				
D	×		×					
E						×	×	×
F							×	×
G						×		×
H								

Context (3)

Table 3.7

4

Modular, Distributive and Boolean Lattices

In Chapter 2 we began an exploration of the algebraic theory of lattices, armed with enough axioms on $\vee$ and $\wedge$ to ensure that each lattice $\langle L; \vee, \wedge \rangle$ arose from a lattice $\langle L; \leqslant \rangle$ and vice versa. Now we introduce identities linking join and meet which are not implied by the laws (L1)–(L4) and their duals $(\text{L1})^\partial$–$(\text{L4})^\partial$ defining lattices (recall 2.9). These hold in many of our examples of lattices, in particular in powersets. In the second part of the chapter we abstract a different feature of powersets, namely the existence of complements.

Lattices satisfying additional identities

Before formally introducing modular and distributive lattices we prove three lemmas which will put the definitions in 4.4 into perspective. The import of these lemmas is discussed in 4.5.

4.1 Lemma. *Let L be a lattice and let $a, b, c \in L$. Then*

(i) $a \wedge (b \vee c) \geqslant (a \wedge b) \vee (a \wedge c)$, *and dually,*

(ii) $a \geqslant c$ *implies* $a \wedge (b \vee c) \geqslant (a \wedge b) \vee c$, *and dually,*

(iii) $(a \wedge b) \vee (b \wedge c) \vee (c \wedge a) \leqslant (a \vee b) \wedge (b \vee c) \wedge (c \vee a)$.

Proof. We leave (i) and (iii) as exercises. (Alternatively, see Exercise 2.9.) By the Connecting Lemma, (ii) is a special case of (i). □

4.2 Lemma. *Let L be a lattice. Then the following are equivalent:*

(i) $(\forall a, b, c \in L)\ a \geqslant c \implies a \wedge (b \vee c) = (a \wedge b) \vee c$;

(ii) $(\forall a, b, c \in L)\ a \geqslant c \implies a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$;

(iii) $(\forall p, q, r \in L)\ p \wedge (q \vee (p \wedge r)) = (p \wedge q) \vee (p \wedge r)$.

Proof. The Connecting Lemma gives the equivalence of (i) and (ii). To prove that (iii) implies (ii), assume that $a \geqslant c$ and apply (iii) with $p = a$, $q = b$ and $r = c$. Conversely, assume (ii) holds and that p, q and r are any elements of L. We may put $a = p$, $b = q$ and $c = p \wedge r$ in (ii), and this gives (iii). □

4.3 Lemma. *Let L be a lattice. Then the following are equivalent:*

(D) $(\forall a, b, c \in L)\ a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$;

$(\text{D})^\partial$ $(\forall p, q, r \in L)\ p \vee (q \wedge r) = (p \vee q) \wedge (p \vee r)$.

Proof. Assume (D) holds. Then, for $p, q, r \in L$,

$$\begin{aligned}(p \vee q) \wedge (p \vee r) &= ((p \vee q) \wedge p) \vee ((p \vee q) \wedge r) && \text{(by (D))}\\ &= p \vee (r \wedge (p \vee q)) && \text{(by (L2)}^\partial \text{ \& (L4)}^\partial\text{)}\\ &= p \vee ((r \wedge p) \vee (r \wedge q)) && \text{(by (D))}\\ &= p \vee (q \wedge r) && \text{(by (L1), (L2)}^\partial \text{ \& (L4))}\end{aligned}$$

so (D) implies $(D)^\partial$. By duality, $(D)^\partial$ implies (D) too. □

4.4 Definitions. Let L be a lattice.

(i) L is said to be **distributive** if it satisfies the **distributive law**,

$$(\forall a, b, c \in L)\ \ a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c).$$

(ii) L is said to be **modular** if it satisfies the **modular law**,

$$(\forall a, b, c \in L)\ \ a \geqslant c \implies a \wedge (b \vee c) = (a \wedge b) \vee c.$$

4.5 Remarks.

(1) Lemma 4.1 shows that any lattice is 'half-way' to being both modular and distributive. To establish distributivity or modularity we only need to check an inequality (see 4.6(5) for an example).

(2) Lemma 4.2 serves two purposes. It shows that any distributive lattice is modular. Also it reveals that the rather mysterious modular law can be reformulated as an identity, a fact we need in 4.7. The modular law may be regarded as licence to rebracket $a \wedge (b \vee c)$ as $(a \wedge b) \vee c$, provided $a \geqslant c$. This observation has no mathematical content, but is useful as an *aide-mémoire*.

(3) Providentially, distributivity can be defined either by (D) or by $(D)^\partial$ (from Lemma 4.3). Thus the apparent asymmetry between join and meet in 4.4(i) is illusory. In other words, L is distributive if and only if L^∂ is. An application of the Duality Principle shows that L is modular if and only if L^∂ is.

(4) The universal quantifiers in Lemmas 4.2 and 4.3 are essential. It is not true, for example, that if particular elements a, b and c in an arbitrary lattice satisfy $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$, then they also satisfy $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$.

4.6 Examples.

(1) Any powerset lattice $\mathcal{P}(X)$ is distributive. More generally, any lattice of sets is distributive. In 10.21 we prove the striking result that *every* distributive lattice is isomorphic to a lattice of sets.

(2) Any chain is distributive (enumerate cases, or (slicker) do Exercise 4.2, or (less elementary) use 4.10).

(3) The lattice $\langle \mathbb{N}_0; \mathrm{lcm}, \gcd \rangle$ is distributive. See Exercise 4.9 for a useful necessary and sufficient condition for distributivity which applies neatly to this example.

(4) Exercise 4.15 asks for a proof that the subgroup lattice of the infinite cyclic group $\langle \mathbb{Z}; + \rangle$ is isomorphic to $\langle \mathbb{N}_0; \mathrm{lcm}, \gcd \rangle^\partial$. Consequently $\mathrm{Sub}\,\mathbb{Z}$ is distributive. Now consider a finite group G. Exercise 5.16 will show that $\mathrm{Sub}\,G$ is distributive if G is cyclic. The converse is also true but is much harder to prove. It requires a more extensive treatment of subgroup lattices than we have space to include.

(5) Our examples of classes of modular lattices come from algebra.

(i) We noted in 2.7 that the set $\mathcal{N}\text{-}\mathrm{Sub}\,G$ of normal subgroups of a group G forms a lattice under the operations

$$H \wedge K = H \cap K \qquad \text{and} \qquad H \vee K = HK,$$

with $\subseteq$ as the underlying order. Let $H, K, N \in \mathcal{N}\text{-}\mathrm{Sub}\,G$, with $H \supseteq N$. Take $g \in H \wedge (K \vee N)$, so $g \in H$ and $g = kn$, for some $k \in K$ and $n \in N$. Then $k = gn^{-1} \in H$, since $H \supseteq N$ and H is a subgroup. This proves that $g \in (H \wedge K) \vee N$. Hence

$$H \wedge (K \vee N) \subseteq (H \wedge K) \vee N.$$

Since the reverse inequality holds in any lattice (by 4.1) the lattice $\mathcal{N}\text{-}\mathrm{Sub}\,G$ is modular, for any group G.

(ii) It can be shown in a similar way that the lattice of subspaces of a vector space (see 2.34(2)) is modular.

(6) Consider the lattices $\mathbf{M}_3$ (the **diamond**) and $\mathbf{N}_5$ (the **pentagon**) shown in Figure 4.1. The lattice $\mathbf{M}_3$ arose in 1.16 as $\mathcal{N}\text{-}\mathrm{Sub}\,\mathbf{V}_4$. Hence, by (5)(i), $\mathbf{M}_3$ is modular. It is, however, not distributive. To see this, note that in the diagram of $\mathbf{M}_3$

$$p \wedge (q \vee r) = p \wedge 1 = p \neq 0 = 0 \vee 0 = (p \wedge q) \vee (p \wedge r).$$

The lattice $\mathbf{N}_5$ is not modular (and so also not distributive): in the diagram we have

$$u \geqslant w \text{ and } u \wedge (v \vee w) = u \wedge 1 = u > w = 0 \vee w = (u \wedge v) \vee w.$$

Figure 4.1

These innocent-looking examples turn out to play a crucial role in the identification of non-modular and non-distributive lattices; see the discussion of the $\mathbf{M_3}$–$\mathbf{N_5}$ Theorem below.

4.7 Sublattices, products and homomorphic images. By 2.13, 2.15 and Exercise 2.19, new lattices can be manufactured by forming sublattices, products and homomorphic images. Modularity and distributivity are preserved by these constructions, as follows.

(i) If L is a modular (distributive) lattice, then every sublattice of L is modular (distributive).

(ii) If L and K are modular (distributive) lattices, then $L \times K$ is modular (distributive).

(iii) If L is modular (distributive) and K is the image of L under a homomorphism, then K is modular (distributive).

Here (i) is immediate and (ii) holds because $\vee$ and $\wedge$ are defined coordinatewise on products. For (iii) we use the fact that a join- and meet-preserving map preserves any lattice identity; for the modular case we then invoke (i) $\Leftrightarrow$ (iii) in 4.2.

A particularly useful consequence of the above results deserves to be singled out as a proposition.

4.8 Proposition. *If a lattice is isomorphic to a sublattice of a product of distributive (modular) lattices, then it is distributive (modular).*

4.9 Examples. Consider Figure 4.2. The lattice L_1 is distributive because it is a sublattice of $4 \times 4 \times 2$, as indicated. (Any product of chains is, of course, distributive.) The lattice L_2 is isomorphic to the shaded sublattice of the modular lattice $\mathbf{M_3} \times \mathbf{2}$ shown alongside and so is itself modular.

The $\mathbf{M_3}$–$\mathbf{N_5}$ Theorem

We have as yet no way of showing that the distributive law or the modular law is *not* satisfied except a random search for elements for which the law fails. The $\mathbf{M_3}$–$\mathbf{N_5}$ Theorem remedies this in a most satisfactory way. It implies that it is possible to determine whether or not a finite lattice is modular or distributive from its diagram. The first part of the theorem is due to R. Dedekind and the second to G. Birkhoff.

We adopt a head-on approach to the proof. This has the disadvantage that it does not reveal why the theorem works. For a more illuminating treatment, beyond the scope of this book, see [40].

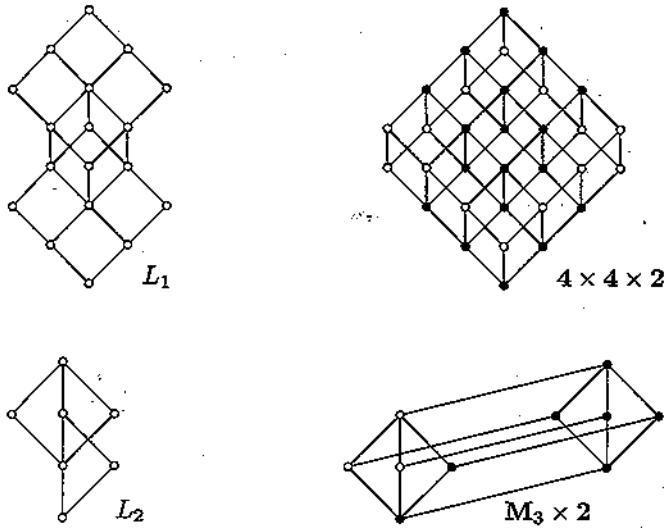


Figure 4.2

Recall from 2.17(2) that we write $M \rightarrowtail L$ to indicate that the lattice L has a sublattice isomorphic to the lattice M.

4.10 The $\mathbf{M_3}$–$\mathbf{N_5}$ Theorem. *Let L be a lattice.*

(i) *L is non-modular if and only if $\mathbf{N_5} \rightarrowtail L$.*

(ii) *L is non-distributive if and only if $\mathbf{N_5} \rightarrowtail L$ or $\mathbf{M_3} \rightarrowtail L$.*

Proof. By 4.6(6) and 4.7, it will be enough to prove that a non-modular lattice has a sublattice isomorphic to $\mathbf{N_5}$ and that a lattice which is modular but not distributive has a sublattice isomorphic to $\mathbf{M_3}$.

Figure 4.3

Assume that L is not modular. Then there exist elements d, e and f such that $d > f$ and $v > u$, where

$$u = (d \wedge e) \vee f \quad \text{and} \quad v = d \wedge (e \vee f).$$

We aim to prove that $e \wedge u = e \wedge v$ $(= p$ say) and $e \vee u = e \vee v$ $(= q$ say). Then our required sublattice has elements u, v, e, p, q (which are clearly distinct). The lattice identities give

$$v \wedge e = (e \wedge (e \vee f)) \wedge d = d \wedge e \text{ and } u \vee e = (e \vee (d \wedge e)) \vee f = e \vee f.$$

Also, by 2.5(4),

$$d \wedge e = (d \wedge e) \wedge e \leqslant u \wedge e \leqslant v \wedge e = d \wedge e$$

and, similarly,

$$e \vee f = u \vee e \leqslant v \vee e \leqslant e \vee f \vee e = e \vee f.$$

This proves our claims and so completes the proof of (i).

Now assume that L is modular but not distributive. We build a sublattice isomorphic to $\mathbf{M}_3$. Take d, e and f such that $(d \wedge e) \vee (d \wedge f) < d \wedge (e \vee f)$. Let

$$\begin{aligned} p &:= (d \wedge e) \vee (e \wedge f) \vee (f \wedge d), \\ q &:= (d \vee e) \wedge (e \vee f) \wedge (f \vee d), \\ u &:= (d \wedge q) \vee p, \\ v &:= (e \wedge q) \vee p, \\ w &:= (f \wedge q) \vee p. \end{aligned}$$

Clearly $u \geqslant p$, $v \geqslant p$ and $w \geqslant p$. Also, by Lemma 4.1(iii), we have $p \leqslant q$. Hence $u \leqslant (d \wedge q) \vee q = q$. Similarly, $v \leqslant q$ and $w \leqslant q$. Our candidate for a copy of $\mathbf{M}_3$ has elements $\{p, q, u, v, w\}$. We need to check that this subset has the correct joins and meets, and that its elements are distinct.

We shall repeatedly appeal to the modular law, *viz.*

(M) $a \geqslant c$ implies $a \wedge (b \vee c) = (a \wedge b) \vee c$.

For each application of (M) we underline the elements a and c involved. In the calculations which follow we use the commutative and associative laws many times without explicit mention. We have $d \wedge q = d \wedge (e \vee f)$, by $(\mathrm{L4})^\partial$. Also

$$\begin{aligned} d \wedge p &= \underline{d} \wedge ((e \wedge f) \vee \underline{((d \wedge e) \vee (d \wedge f))}) \\ &= (d \wedge (e \wedge f)) \vee ((d \wedge e) \vee (d \wedge f)) \\ &= (d \wedge e) \vee (d \wedge f). \end{aligned}$$

Thus $p = q$ is impossible. We conclude that $p < q$.

We next prove that $u \wedge v = p$. We have

$$\begin{aligned} u \wedge v &= ((d \wedge q) \vee \underline{p}) \wedge \underline{((e \wedge q) \vee p)} \\ &= ((e \wedge \underline{q}) \vee \underline{p}) \wedge (d \wedge q)) \vee p && \text{(by (M))} \\ &= ((q \wedge (e \vee p)) \wedge (d \wedge q)) \vee p && \text{(by (M))} \\ &= ((e \vee p) \wedge (d \wedge q)) \vee p \\ &= ((d \wedge (e \vee f)) \wedge (e \vee (f \wedge d))) \vee p && \text{(by (L4) \& } (\mathrm{L4})^\partial\text{)} \\ &= (d \wedge \underline{((e \vee f)} \wedge \underline{(e} \vee (f \wedge d)))) \vee p \\ &= (d \wedge (((e \vee f) \wedge (f \wedge d)) \vee e)) \vee p && \text{(by (M))} \\ &= (\underline{d} \wedge \underline{((f \wedge d)} \vee e)) \vee p && \text{(since } d \wedge f \leqslant f \leqslant e \vee f\text{)} \\ &= ((d \wedge e) \vee (f \wedge d)) \vee p && \text{(by (M))} \\ &= p. \end{aligned}$$

In exactly the same way, $v \wedge w = p$ and $w \wedge u = p$. Similar calculations yield $u \vee v = v \vee w = w \vee u = q$.

Finally, it is easy to see that if any two of the elements u, v, w, p, q are equal, then $p = q$, which is impossible. □

4.11 Applying the $\mathbf{M}_3$–$\mathbf{N}_5$ Theorem. Consider the four lattices in Figure 4.4. The lattices L_1 and L_2 have sublattices isomorphic to $\mathbf{N}_5$, explicitly exhibited, and $\mathbf{M}_3 \rightarrowtail L_3$. The $\mathbf{M}_3$–$\mathbf{N}_5$ Theorem implies, immediately and conclusively, that L_1 and L_2 are non-modular and that L_3 is non-distributive.

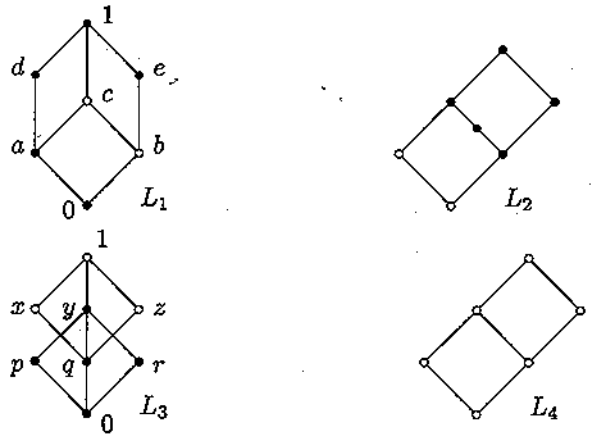


Figure 4.4

It is apparent from the diagrams that $\mathbf{N}_5$ does not embed in L_3 and that neither $\mathbf{N}_5$ nor $\mathbf{M}_3$ embeds in L_4. However, to justify such assertions fully requires a tedious enumeration of cases. For example, suppose $\{u, a, b, c, v\}$, with $u < c < a < v$, $u < b < v$, were a sublattice of L_3 isomorphic to $\mathbf{N}_5$. Since L_3 and $\mathbf{N}_5$ both have length 3, we must have $u = 0$ and $v = 1$. Since $a \wedge b = c \wedge b = 0$ and $a \vee b = c \vee b = 1$, by duality and symmetry we may assume without loss of generality that $a = r$, $c = p$ and $b = x$. But $\{0, r, x, p, 1\}$ is not a sublattice of L_3, $\lightning$.

To decide whether a given lattice L is non-modular, modular but non-distributive, or distributive, we therefore proceed as follows. If a sublattice of L isomorphic to $\mathbf{N}_5$ ($\mathbf{M}_3$) can be exhibited, then L is non-modular (non-distributive), by the $\mathbf{M}_3$–$\mathbf{N}_5$ Theorem. If a search for a copy of $\mathbf{N}_5$ (of either $\mathbf{N}_5$ or $\mathbf{M}_3$) fails, we conjecture that L is modular (distributive). To substantiate this claim we first try to apply the sublattice-of-a-product technique. Our remarks in Example 4.12 show that there are cases where this is doomed to fail. A fall-back method, relying on a rather tricky proof, is given in Exercise 4.18.

It should be emphasized that the statement of the $\mathbf{M}_3$–$\mathbf{N}_5$ Theorem refers to the occurrence of the pentagon or diamond as a *sublattice* of L; this means that the joins and meets in a candidate copy of $\mathbf{N}_5$ or $\mathbf{M}_3$ must be the same as those in L. In Figure 4.4, the pentagon $K = \{0, a, b, d, 1\}$ in L_1 is *not* a sublattice; $a \vee b = c \notin K$. In the other direction, in applying Proposition 4.8 one must be sure to embed the given lattice as a *sublattice*. Thus it would be erroneous to conclude from Figure 2.6(iv) that $\mathbf{N}_5$ is distributive: $\mathbf{N}_5$ sits inside the distributive lattice $\mathbf{2}^3$, but not as a sublattice.

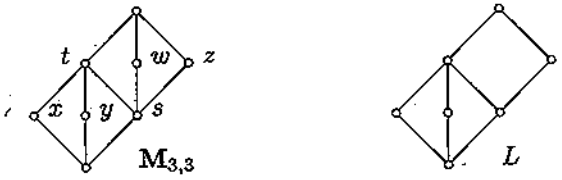


Figure 4.5

4.12 Example. Consider the lattice $\mathbf{M}_{3,3}$ in Figure 4.5. It can be shown (see Exercise 6.15) that if $\mathbf{M}_{3,3} \leqslant A \times B$ for some lattices A and B, then $\mathbf{M}_{3,3} \rightarrowtail A$ or $\mathbf{M}_{3,3} \rightarrowtail B$. Consequently the sublattice-of-a-product technique cannot be applied to $\mathbf{M}_{3,3}$. Nevertheless, $\mathbf{M}_{3,3}$ is modular. To see this, note that for $u \in \{x, y, z\}$, the sublattice $\mathbf{M}_{3,3} \setminus \{u\}$ is isomorphic to L (Figure 4.5) or to its dual, both of which are modular

(see 4.9 and 4.5(3)). Thus any sublattice of $\mathbf{M}_{3,3}$ isomorphic to $\mathbf{N}_5$ would need to contain the antichain $\{x, y, z\}$, which is impossible.

Boolean lattices and Boolean algebras

A Boolean algebra is a distributive lattice with additional structure which mimics the complementation in a powerset. Our first task is to define complements in an arbitrary lattice.

4.13 Complements. Let L be a lattice with 0 and 1. For $a \in L$, we say $b \in L$ is a **complement** of a if $a \wedge b = 0$ and $a \vee b = 1$. If a has a *unique* complement, we denote this complement by a'.

Assume L is distributive and suppose that b_1 and b_2 are both complements of a. Then

$$b_1 = b_1 \wedge 1 = b_1 \wedge (a \vee b_2) = (b_1 \wedge a) \vee (b_1 \wedge b_2) = b_1 \wedge b_2.$$

Hence $b_1 \leqslant b_2$ by the Connecting Lemma. Interchanging b_1 and b_2 gives $b_2 \leqslant b_1$. Therefore in a distributive lattice an element can have at most one complement. It is easy to find examples of non-unique complements in non-distributive lattices: look at $\mathbf{M}_3$ or $\mathbf{N}_5$ (Figure 4.1).

A lattice element may have no complement. The only complemented elements in a bounded chain are 0 and 1. If $\mathfrak{L} \subseteq \mathcal{P}(X)$ is a lattice of sets, then an element $A \in \mathfrak{L}$ has a complement if and only if $X \setminus A$ belongs to $\mathfrak{L}$. Thus the complemented elements of $\mathcal{O}(P)$ are the sets which are simultaneously down-sets and up-sets. (See Exercise 1.17.)

4.14 Definition. A lattice L is called a **Boolean lattice** if

(i) L is distributive,

(ii) L has 0 and 1,

(iii) each $a \in L$ has a (necessarily unique) complement $a' \in L$.

The following lemma collects together properties of the complement in a Boolean lattice.

4.15 Lemma. *Let L be a Boolean lattice. Then*

(i) $0' = 1$ *and* $1' = 0$,

(ii) $a'' = a$ *for all* $a \in L$,

(iii) *de Morgan's laws hold: for all* $a, b \in L$,

$$(a \vee b)' = a' \wedge b' \quad \text{and} \quad (a \wedge b)' = a' \vee b',$$

(iv) $a \wedge b = (a' \vee b')'$ *and* $a \vee b = (a' \wedge b')'$ *for all* $a, b \in L$,

(v) $a \wedge b' = 0$ *if and only if* $a \leqslant b$ *for all* $a, b \in L$.

Proof. To prove $p = q'$ in L it is sufficient to prove that $p \vee q = 1$ and $p \wedge q = 0$, since the complement of q is unique. This observation makes the verification of (i)–(iii) entirely routine. Part (iv) follows from (ii) and (iii), while (v) is an easy exercise. □

4.16 Boolean algebras. A Boolean lattice was defined to be a special kind of distributive lattice. In such a lattice it is often more natural to regard the distinguished elements 0 and 1 and the unary operation $'$ as an integral part of the structure, with their properties embodied in axioms. Accordingly, a **Boolean algebra** is defined to be a structure $\langle B; \vee, \wedge, ', 0, 1\rangle$ such that

(i) $\langle B; \vee, \wedge\rangle$ is a distributive lattice,

(ii) $a \vee 0 = a$ and $a \wedge 1 = a$ for all $a \in B$,

(iii) $a \vee a' = 1$ and $a \wedge a' = 0$ for all $a \in B$.

This viewpoint is extended to other concepts from Chapter 2. We say that a subset A of a Boolean algebra B is a **subalgebra** if A is a sublattice of B which contains 0 and 1 and is such that $a \in A$ implies $a' \in A$. Given Boolean algebras B and C, a map $f\colon B \to C$ is a **Boolean homomorphism** if f is a lattice homomorphism which also preserves $0, 1$ and $'$ (that is, $f(0) = 0, f(1) = 1$ and $f(a') = (f(a))'$ for all $a \in B$). Lemma 4.17 shows that these conditions are not independent.

4.17 Lemma. *Let $f\colon B \to C$, where B and C are Boolean algebras.*

(i) *Assume f is a lattice homomorphism. Then the following are equivalent:*

(a) $f(0) = 0$ *and* $f(1) = 1$;

(b) $f(a') = (f(a))'$ *for all* $a \in B$.

(ii) *If f preserves $'$, then f preserves $\vee$ if and only if f preserves $\wedge$.*

Proof. (i) To confirm that (a) implies (b) use the equations

$$0 = f(0) = f(a \wedge a') = f(a) \wedge f(a'),$$
$$1 = f(1) = f(a \vee a') = f(a) \vee f(a').$$

Conversely, if (b) holds, we have

$$f(0) = f(a \wedge a') = f(a) \wedge (f(a))' = 0,$$
$$f(1) = f(a \vee a') = f(a) \vee (f(a))' = 1.$$

(ii) Assume f preserves $'$ and $\vee$. By Lemma 4.15(iv),

$$f(a \wedge b) = f((a' \vee b')') = (f(a' \vee b'))' = (f(a') \vee f(b'))'$$
$$= ((f(a))' \vee (f(b))')' = f(a) \wedge f(b),$$

for all $a, b \in B$. The converse is proved dually. □

4.18 Examples of Boolean algebras. Here we present some old friends in new clothes and add some examples of infinite Boolean algebras which are important in Chapters 10 and 11.

(1) For any set X, let $A' := X \setminus A$ for all $A \subseteq X$. Then the structure $(\wp(X); \cup, \cap, ', \varnothing, X)$ is a Boolean algebra known as the **powerset algebra** on X. By an **algebra of sets** (also known as a **field of sets**) we mean a subalgebra of some powerset algebra $\wp(X)$, that is, a family of sets which forms a Boolean algebra under the set-theoretic operations.

We shall prove in Chapter 5 that every *finite* Boolean algebra is isomorphic to $\wp(X)$ for some finite set X. Example (2) below shows that there are infinite Boolean algebras which are not powerset algebras. However, we show in Chapter 10 that *every* Boolean algebra is isomorphic to an algebra of sets. Further, in 10.24 we characterize the powerset algebras among Boolean algebras.

(2) The **finite-cofinite algebra** of the set X is defined to be

$$\mathrm{FC}(X) := \{\, A \subseteq X \mid A \text{ is finite or } X \setminus A \text{ is finite} \,\}.$$

It is easily checked that this is an algebra of sets, but we claim that $\mathrm{FC}(\mathbb{N})$ is not isomorphic to $\wp(X)$ for any set X. One way to arrive at this is to consider cardinalities. It is a standard exercise on countable sets to prove that $\mathrm{FC}(\mathbb{N})$ is countable. On the other hand, Cantor's Theorem implies that any powerset is either finite or uncountable. A more lattice-theoretic proof involves completeness. Exercise 2.25 shows that $\mathrm{FC}(\mathbb{N})$ is not complete. But $\wp(X)$ is always complete and an isomorphism must preserve completeness, by 2.27(ii).

(3) The family of all clopen subsets of a topological space $(X; T)$ is an algebra of sets. Clearly this example will not be of much interest unless X has plenty of clopen sets. The significance of Boolean algebras of this sort emerges in Chapter 11 where we show that every Boolean algebra can be concretely represented as such an algebra.

(4) For $n \geqslant 1$ the lattice $\mathbf{2}^n$ is lattice-isomorphic to $\wp(\{1, 2, \ldots, n\})$, which is a Boolean algebra. Hence $\mathbf{2}^n$ is a Boolean algebra, with

$$0 = (0, 0, \ldots, 0) \quad \text{and} \quad 1 = (1, 1, \ldots, 1),$$
$$(\varepsilon_1, \ldots, \varepsilon_n)' = (\eta_1, \ldots, \eta_n), \text{ where } \eta_i = 0 \iff \varepsilon_i = 1.$$

The simplest non-trivial Boolean algebra of all is $\mathbf{2} = \{0, 1\}$. It arises frequently in logic and computer science as an algebra of truth

values. In such contexts the symbols F and T, or alternatively $\bot$ and $\top$, are used in place of 0 and 1. We have

$$\mathsf{F} \vee \mathsf{F} = \mathsf{F} \wedge \mathsf{F} = \mathsf{F} \wedge \mathsf{T} = \mathsf{T}' = \mathsf{F},$$
$$\mathsf{T} \wedge \mathsf{T} = \mathsf{F} \vee \mathsf{T} = \mathsf{T} \vee \mathsf{T} = \mathsf{F}' = \mathsf{T}.$$

In the next section we begin to explore the way Boolean algebras model deductive reasoning involving statements which are assigned value T (true) or F (false).

Boolean terms and disjunctive normal form

Historically, Boolean algebras are inextricably linked to logic, and it is in this context that students in a variety of disciplines encounter them, at levels ranging from primary school to graduate. Many Boolean algebra applications, circuit design for example, are specialist topics which rely on the laws of Boolean algebra but quickly leave lattice theory behind. We only hint at such applications.

4.19 Truth tables: the algebra of propositions. In propositional calculus, propositions are designated by **propositional variables** which take values in $\{\mathsf{F}, \mathsf{T}\}$. Admissible compound statements are formed using **logical connectives**. Connectives include 'and', 'or' and 'not', denoted respectively by our old friends $\wedge, \vee$ and $'$. Another natural connective is 'implies' ($\rightarrow$). Compound statements built from these are assigned the expected truth values according to the truth values of their constituent parts. For example, $p \wedge q$ has value T if and only if both p and q have value T and $p \rightarrow q$ has value T unless p has value T and q has value F. (Any puzzlement resulting from the fact that $p \rightarrow q$ is true whenever p is false should be dispelled by reading $p \rightarrow q$ as 'if p, then q'.)

Formally, we take an infinite set of propositional variables, denoted $p, q, r, \ldots$, and define a **wff** (or **well-formed formula**) by the rules:

(i) any propositional variable standing alone is a wff (optionally, constant symbols T and F may also be included as wffs);

(ii) if φ and ψ are wffs, so are $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, φ', $(\varphi \rightarrow \psi)$ (this clause is suitably adapted if a different set of connectives is used);

(iii) any wff arises from a finite number of applications of (i) and (ii).

Thus $((p \wedge q') \vee r)'$ and $((p' \rightarrow q) \rightarrow ((p' \rightarrow q') \rightarrow p))$ are wffs while $(((p \vee q) \wedge p)$ (invalid bracketing) and $\vee \rightarrow q$ (arrant nonsense) are not. The parentheses may appear to clutter up wffs unnecessarily. Included as dictated by the definition, they guarantee non-ambiguity. In practice we drop parentheses where no ambiguity would result, just as if we were writing a string of joins, meets and complements in a lattice.

A wff φ involving the propositional variables $p_1, \ldots, p_n$ defines a **truth function** F_φ of n variables. For a given assignment of values in $\{\mathsf{F}, \mathsf{T}\}$ to $p_1, \ldots, p_n$, simply substitute these values into φ and compute the resulting expression in the Boolean algebra $\{\mathsf{F}, \mathsf{T}\}$ to obtain the value of F_φ. Conventionally, truth functions are presented via **truth tables**, as illustrated in Table 4.1.

p	q	$p \rightarrow q$
T	T	T
T	F	F
F	T	T
F	F	T

p_1	p_2	p_3	$(p_1 \vee p_2)$	$(p_1' \vee p_3)$	$((p_1 \vee p_2) \wedge (p_1' \vee p_3))'$
T	T	T	T	T	F
T	T	F	T	F	T
T	F	T	T	T	F
T	F	F	T	F	T
F	T	T	T	T	F
F	T	F	T	T	F
F	F	T	F	T	T
F	F	F	F	T	T

Table 4.1

Two wffs φ and ψ are called **logically equivalent** (written $\varphi \equiv \psi$) if they define the same truth function, that is, they give rise to the same truth table. For the purposes of deductive reasoning, logically equivalent wffs are essentially the same. It is an easy exercise on truth tables to prove that, for any wffs φ and ψ,

$$(\varphi \wedge \psi) \equiv (\varphi' \vee \psi')', \quad (\varphi \vee \psi) \equiv (\varphi' \wedge \psi')',$$
$$(\varphi \rightarrow \psi) \equiv (\varphi' \vee \psi), \quad (\varphi \wedge \psi) \equiv (\varphi \rightarrow \psi')'.$$

A proof by induction on the number of connectives then shows that any wff built using $\vee, \wedge$ and $'$ is logically equivalent to one built using $\rightarrow$ and $'$, and vice versa. Therefore, up to logical equivalence, we arrive at the same set of wffs whether we take $\{\vee, \wedge, ', \rightarrow\}$, just $\{\rightarrow, '\}$ or just $\{\vee, \wedge, '\}$ as the basic set of connectives. The choice of $\{\rightarrow, '\}$ is the most natural for studying logic, while $\{\vee, \wedge, '\}$ brings out the connections with Boolean algebras.

The set of wffs, with $\vee, \wedge$ and $'$ as operations, closely resembles a Boolean lattice. The axioms do not hold if $=$ is taken to mean 'is the

same wff as', but it is a routine matter to show that they all hold if $=$ is read as 'is logically equivalent to'. For example, to establish (L4) note that $\varphi \vee (\varphi \wedge \psi)$ takes value **T** if and only if φ does, so $\varphi \vee (\varphi \wedge \psi) \equiv \varphi$. If **F** and **T** are included as wffs, to serve as 0 and 1, we obtain a Boolean algebra. We meet this **algebra of propositions** again, in a more formal guise, in Chapter 11.

4.20 Boolean terms. We used the Boolean symbols, representing logical connectives, to build the formulae of propositional calculus from propositional variables. This construction, freed from the trappings of logic, has a variety of applications. We define the class **BT** of **Boolean terms** (or **Boolean polynomials**) as follows. Let S be a set of variables, whose members will be denoted by letters such as $x, y, z, x_1, x_2, \ldots$, and let $\vee, \wedge, ', 0, 1$ be the symbols used to axiomatize Boolean algebras. Then

(i) $0, 1 \in \mathbf{BT}$ and $x \in \mathbf{BT}$ for all $x \in S$,

(ii) if $p, q \in \mathbf{BT}$ then $(p \vee q), (p \wedge q)$ and p' belong to **BT**,

(iii) every element of **BT** is an expression formed by a finite number of applications of (i) and (ii).

A Boolean term p whose variables are drawn from among $x_1, \ldots, x_n$ will be written $p(x_1, \ldots, x_n)$. Examples of Boolean terms, illustrating the building process, are

$$1,\ x,\ y,\ y',\ (x \vee y'),\ (1 \wedge (x \vee y')),\ (1 \wedge (x \vee y'))'.$$

Just as numbers may be substituted into 'ordinary' polynomials, elements of any Boolean algebra B may be substituted for the variables of a Boolean term to yield an element of B. In particular we may take $B = \mathbf{2}$. Thence every Boolean term $p(x_1, \ldots, x_n)$ defines a map $F_p \colon \mathbf{2}^n \to \mathbf{2}$. The map F_p associated with p can be specified by a 'truth table' in just the same way as a wff determines a truth function; the only difference is that each entry of the table is 0 or 1, instead of **F** or **T**.

It is usual to use p to denote both the term and the function F_p it induces. As it simplifies the notation, we shall do this from time to time when no confusion results.

We say that $p(x_1, \ldots, x_n)$ and $q(x_1, \ldots, x_n)$ are **equivalent**, and write $p \equiv q$, if p and q have the same truth function, that is, $F_p = F_q$. It is easy to see that, for instance, $(x \wedge y')' \equiv (x' \vee y)$; just check that both sides give the same truth table. Note that the right-hand side can be obtained from the left by applying the laws of Boolean algebra, treating the variables as though they were Boolean algebra elements:

$$(x \wedge y')' = (x' \vee y'') = (x' \vee y).$$

In general, whenever $q(x_1, \ldots, x_n)$ can be obtained from $p(x_1, \ldots, x_n)$ by the laws of Boolean algebra, we have $p \equiv q$. We see in 4.24 that the converse is also true. Where removal of parentheses from a Boolean term would, up to equivalence, not result in ambiguity, we shall omit the parentheses. For example, we shall write $x \vee y \vee z$ in place of either $(x \vee (y \vee z))$ or $((x \vee y) \vee z)$.

We prove in 4.23 that *every* map from $\mathbf{2}^n$ to $\mathbf{2}$ coincides with F_p for some Boolean term p in n variables. This is a surprising and important theorem. To motivate both the theorem and its proof, we preface it with a brief discussion of one application.

4.21 Boolean terms and computer architecture. The design of a computer system may be viewed hierarchically: from customer requirement, down in multiple stages through high level programming language, machine code and integrated circuits to semiconductors. The circuits are carried on chips of silicon and consist of interconnected groups of transistors. A crucial feature of transistors is that although subjected to continuously varying voltages, they either allow current to pass to the best of their ability or not at all, and so act as electrical switches. Transistors are linked to create gates. A gate recognizes only two levels of voltage: high (denoted 1) and low (denoted 0). It may be regarded as having n inputs each taking value 0 or 1, and having one or more outputs, each taking value 0 or 1 (depending on the combination of inputs). A basic kit for constructing circuits consists of

AND gate	two inputs; output 1 if and only if both inputs are 1,
OR gate	two inputs; output 1 if and only if either input is 1,
NOT gate	one input; output 0 if and only if the input is 1.

Figure 4.6(i) shows a stylized representation of these gates and 4.6(ii) a gate **diagram** for a circuit. The same input–output behaviour results from ordinary electric switches wired in series (for an AND gate) and in parallel (for an OR gate). Thus gate diagrams are really what in olden, pre-transistor days were known as **series-parallel switching circuits**.

Clearly AND, OR and NOT gates mimic $\wedge, \vee$ and $'$ acting on $\{0, 1\}$. Thus a gate diagram with k outputs corresponds to a k-element set of Boolean terms. The 2–output diagram in Figure 4.6(ii) corresponds to the 2-element set of terms $\{x_1 \wedge x_2,\ (x_1 \vee x_2) \wedge (x_3' \vee (x_4' \wedge x_5))\}$.

The problem of constructing a gate diagram to model a circuit with specified characteristics is just that of finding a Boolean term with a given truth function. Theorem 4.23 solves this problem, but in a way which is in general highly redundant. In circuit design this may be very undesirable. A complicated Boolean term may lead to a circuit which is

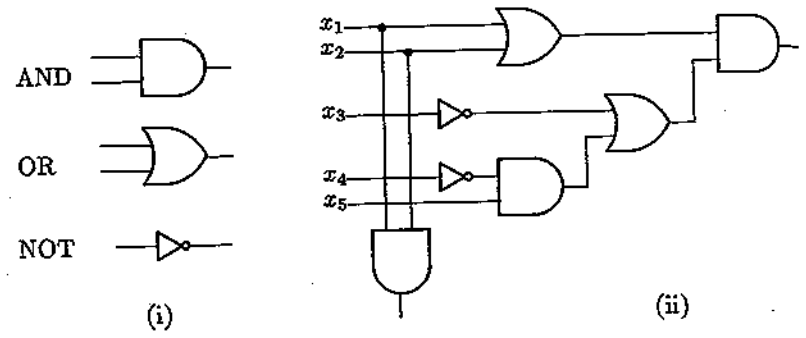


Figure 4.6

costly (many connectives entail many gates), hard to realize compactly on a chip (if the term is 'irregular'), or slow (if the term involves many sub-terms or long strings of joins or meets). This difficulty can in theory be overcome by using the laws of Boolean algebra to replace a complicated Boolean term by a simpler equivalent one; see Example 4.22(2). However, a single chip may have millions of transistors, and direct implementation of complicated circuits would be impractical and wastefully repetitive. Instead a modular approach is adopted. At the level above gate diagrams in the design hierarchy comes microprogramming. This deals with the implementation of relatively simple modular components (adders, memories, etc.) from which, in turn, more complex processing units are constructed. Effective system design depends on good communication between adjacent levels in the hierarchy. Thus gate diagrams cannot be divorced from microprogramming (above) and the wiring of transistors (below). The use of Boolean terms is enmeshed with the methodology of these related topics, so that we can only illustrate it in a limited way. A much fuller account can be found in [31], for example.

4.22 Examples.

(1) Each stage in the binary addition of two numbers involves addition modulo 2, with a 'carry' if two 1s are added. For example, suppose we wish to add 6, represented as 110, to 3, represented as 011. We add the final bits 0 and 1 to give 1 and no carry, then add the penultimate bits to give 0 with a carry 1, and finally add the leftmost bits, remembering to include the carry. This gives the expected result, $1001 = 9$. A circuit (**an adder**) to execute this procedure can be built from components known as **half-adders**, where each half-adder carries out a single 'sum and carry' operation on a pair of bits. Input–output for a half-adder is shown Table 4.2.

x	y	*sum*	*carry*
0	0	0	0
0	1	1	0
1	0	1	0
1	1	0	1

Table 4.2

It is immediate that *carry* is given by the Boolean term $x \wedge y$. For *sum* we require a term $p(x, y)$ which takes value 1 when exactly one of x, y takes value 1. Looking at lines 2 and 3 of the table we see that $x \wedge y'$ and $x' \wedge y$ have this property; no other meets of pairs do. It is then routine to verify that the term $(x \wedge y') \vee (x' \wedge y)$ has exactly the truth table we require for *sum*. The associated gate diagram is shown in Figure 4.7. This can be simplified by having tailor-made gates for other 2-variable truth functions; here an XOR gate is wanted, modelling the exclusive form of OR.

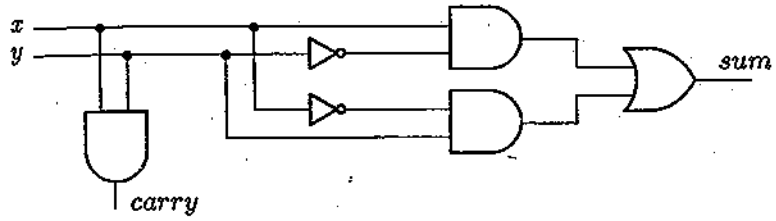


Figure 4.7

(2) Our second example concerns a circuit to execute a logical operation as opposed to an arithmetical one. We seek $p(x, y, z)$ having the truth function given in Table 4.3.

x	y	z	$p(x, y, z)$
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	0
1	1	0	1
1	1	1	1

Table 4.3

Notice that if $x = 1$ then $p(x, y, z)$ must take the same value as y and otherwise must take the value of z. We are therefore modelling `if-then-else`, for which an appropriate term is $(x \wedge y) \vee (x' \wedge z)$. Let us now see how we can arrive at this by the technique we used in the previous example. We want $p(x, y, z) = 1$ for the combinations of truth values in rows 2, 4, 7 and 8. The terms $x' \wedge y' \wedge z$, $x' \wedge y \wedge z$, $x \wedge y \wedge z'$ and $x \wedge y \wedge z$ give value 1 on these rows. Taking the join of these we have a candidate for $p(x, y, z)$, namely

$$(x' \wedge y' \wedge z) \vee (x' \wedge y \wedge z) \vee (x \wedge y \wedge z') \vee (x \wedge y \wedge z)$$
$$\equiv ((x' \wedge z) \wedge (y \vee y')) \vee ((x \wedge y) \wedge (z \vee z')) \equiv (x \wedge y) \vee (x' \wedge z).$$

The construction in the following proof generalizes that used above. Take the truth table associated with a given truth function $F\colon \mathbf{2}^n \to \mathbf{2}$. For each row (element of $\mathbf{2}^n$) on which F has value 1, form the meet of n symbols by selecting for each variable x either x or x' depending on whether x has value 1 or 0 in that row. The join of these terms, p, is such that $F = F_p$.

4.23 Theorem. *Every map $F : \mathbf{2}^n \to \mathbf{2}$ coincides with F_p for some Boolean term $p(x_1, \dots, x_n)$. A suitable term p may be described as follows: For $\mathbf{a} = (a_1, \dots, a_n) \in \mathbf{2}^n$, define $p_{\mathbf{a}}(x_1, \dots, x_n)$ by*

$$p_{\mathbf{a}}(x_1, \dots, x_n) = x_1^{\varepsilon_1} \wedge \cdots \wedge x_n^{\varepsilon_n} \text{ where } x_j^{\varepsilon_j} = \begin{cases} x_j & \text{if } a_j = 1, \\ x_j' & \text{if } a_j = 0. \end{cases}$$

Then define

$$p(x_1, \dots, x_n) = \bigvee \{ p_{\mathbf{a}}(x_1, \dots, x_n) \mid F(\mathbf{a}) = 1 \}.$$

Proof. Let $\mathbf{a} = (a_1, \dots, a_n) \in \mathbf{2}^n$ and $\mathbf{b} = (b_1, \dots, b_n) \in \mathbf{2}^n$. We have carefully chosen the term $p_{\mathbf{a}}(x_1, \dots, x_n)$ so that

$$F_{p_{\mathbf{a}}}(b_1, \dots, b_n) = \begin{cases} 1 & \text{if } \mathbf{b} = \mathbf{a}, \\ 0 & \text{if } \mathbf{b} \neq \mathbf{a}. \end{cases}$$

We claim that $F = F_p$. Assume that $F(\mathbf{b}) = 1$. Then

$$\begin{aligned} F_p(b_1, \dots, b_n) &= \bigvee \{ F_{p_{\mathbf{a}}}(b_1, \dots, b_n) \mid F(\mathbf{a}) = 1 \} \\ &\geqslant F_{p_{\mathbf{b}}}(b_1, \dots, b_n) && \text{(since } F(\mathbf{b}) = 1\text{)} \\ &= 1 && \text{(by the above).} \end{aligned}$$

Thus $F(\mathbf{b}) = 1$ implies $F_p(\mathbf{b}) = 1$. Now assume $F(\mathbf{b}) = 0$. Then $F(\mathbf{a}) = 1$ implies $\mathbf{b} \neq \mathbf{a}$, so $F_{p_{\mathbf{a}}}(b_1, \dots, b_n) = 0$. Therefore

$$F_p(b_1, \dots, b_n) = \bigvee \{ F_{p_{\mathbf{a}}}(b_1, \dots, b_n) \mid F(\mathbf{a}) = 1 \} = 0.$$

Thus $F(\mathbf{b}) = 0$ implies $F_p(\mathbf{b}) = 0$. Hence $F = F_p$, as claimed. □

4.24 Disjunctive normal form. A Boolean term $p(x_1, \dots, x_n)$ is said to be in **full disjunctive normal form**, or DNF, if it is a join of distinct meets of the form $x_1^{\varepsilon_1} \wedge \cdots \wedge x_n^{\varepsilon_n}$. By definition, x^ε equals x if $\varepsilon = 1$, and x' if $\varepsilon = 0$; terms of the form x^ε are known as **literals**.

Theorem 4.23 implies that any Boolean term is equivalent to a term in DNF (in the setting of propositional calculus this is just the classic result that any wff is logically equivalent to a wff in DNF). Note that the Boolean term 0 is already in DNF as it is the join of the empty set. At the other end of the spectrum, the DNF of the Boolean term 1 is the join of all 2^n meets of the form $x_1^{\varepsilon_1} \wedge \cdots \wedge x_n^{\varepsilon_n}$. Each truth function uniquely determines, and is determined by, a DNF term, so $p \equiv q$ in BT if and only if each of p and q is equivalent to the same DNF, in the sense that every meet of literals occurring in the DNF of p occurs in the DNF of q and vice versa. We have already remarked that applying the laws of Boolean algebra to a Boolean term yields an equivalent term. We observe that this process can be used to reduce any term $p(x_1, \dots, x_n)$ to DNF, as outlined below.

(i) Use de Morgan's laws to reduce $p(x_1, \dots, x_n)$ to literals combined by joins and meets.

(ii) Use the distributive laws repeatedly, with the lattice identities, to obtain a join of meets of literals.

(iii) Finally, we require each x_i to occur, either primed or not, once and once only in each meet term. This is achieved by dropping any terms containing both x_i and x_i', for any i. If neither x_j nor x_j' occurs in $\bigwedge_{k \in K} x_k^{\varepsilon_k}$, note that

$$\bigwedge_{k \in K} x_k^{\varepsilon_k} \equiv \Big(\bigwedge_{k \in K} x_k^{\varepsilon_k} \Big) \wedge (x_j \vee x_j') \equiv \Big(\bigwedge_{k \in K} x_k^{\varepsilon_k} \wedge x_j \Big) \vee \Big(\bigwedge_{k \in K} x_k^{\varepsilon_k} \wedge x_j' \Big).$$

Repeating this for each missing variable we arrive at a term in DNF.

This process shows that a term may be converted to DNF by the laws of Boolean algebra. We may therefore assert that $p \equiv q$ if and only if $F_p(a_1, \dots, a_n) = F_q(a_1, \dots, a_n)$ for any elements $a_1, \dots, a_n$ in any Boolean algebra B, and it is possible to test whether this is true by reducing each of p and q to DNF; see Exercise 4.36.

For the term $((p_1 \vee p_2) \wedge (p_1' \vee p_3))'$ the truth table in Table 4.1 and Theorem 4.23 give the DNF

$$(p_1' \wedge p_2' \wedge p_3) \vee (p_1 \wedge p_2 \wedge p_3') \vee (p_1 \wedge p_2' \wedge p_3') \vee (p_1' \wedge p_2' \wedge p_3').$$

For comparison, try obtaining this by the process above.

Finally we give a theorem which is essentially a reformulation of Theorem 4.23.

4.25 Theorem. *Let B be the Boolean algebra $\mathbf{2}^{2^n}$. Then B is generated by n elements, in the sense that there exists an n-element subset X of B such that the smallest Boolean subalgebra of B containing X is B.*

Proof. Identify B with the Boolean algebra $\wp(2^n)$. Define X to be $\{e_1, \ldots, e_n\}$, where $e_i := \{(a_1, \ldots, a_n) \in \mathbf{2}^n \mid a_i = 1\}$ for $i = 1, \ldots, n$. Then for each $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbf{2}^n$ we have

$$\{\mathbf{a}\} = \bigcap \{e_i \mid a_i = 1\} \cap \bigcap \{e_i' \mid a_i = 0\}.$$

Each non-empty element of B is a union of singletons, $\{\mathbf{a}\}$, and hence expressible as a join of meets of elements of the form e_i or e_i'; noting that $\varnothing = e_1 \cap e_1'$ takes care of the empty set. □

Exercises

Exercises from the text. Prove Lemma 4.1(i), (iii). Prove that L^∂ is modular if L is (see 4.5(3)). Complete the proof of Lemma 4.15 (i)–(iv) (see Exercise 4.26 for (v)).

4.1 Show that $\mathbf{M}_3$ and $\mathbf{N}_5$ are the only non-distributive lattices with fewer than 6 elements.

4.2 (i) Find a set X of least cardinality such that the chain $\mathbf{3}$ is isomorphic to a sublattice of $\wp(X)$. Conclude that the chains $\mathbf{1}, \mathbf{2}$ and $\mathbf{3}$ are distributive.

(ii) Let C be any chain and let $x, y, z \in C$. Show that $x \wedge (y \vee z)$ and $(x \wedge y) \vee (x \wedge z)$ both belong to $\{x, y, z\}$.

(iii) By combining (i) and (ii) show that every chain is distributive.

4.3 Which of the lattices of Figure 4.8 are distributive and which are modular? (Use the $\mathbf{M}_3$–$\mathbf{N}_5$ Theorem and Proposition 4.8, as explained in 4.11, to justify your answers.)

4.4 Repeat the previous exercise for the lattices in Figure 2.17.

4.5 Use the $\mathbf{M}_3$–$\mathbf{N}_5$ Theorem to show that if L is a distributive lattice, then L^∂ is also distributive. (This yields a non-computational proof of Lemma 4.3.)

4.6 Use the $\mathbf{M}_3$–$\mathbf{N}_5$ Theorem to show that every lattice L of length 2 is modular (Exercise 2.33 sought a description of all such lattices). (Hence, in particular, $\mathbf{M}_n$ is modular for all n.)

4.7 Find all pairs of lattices (L, K) (up to isomorphism) such that

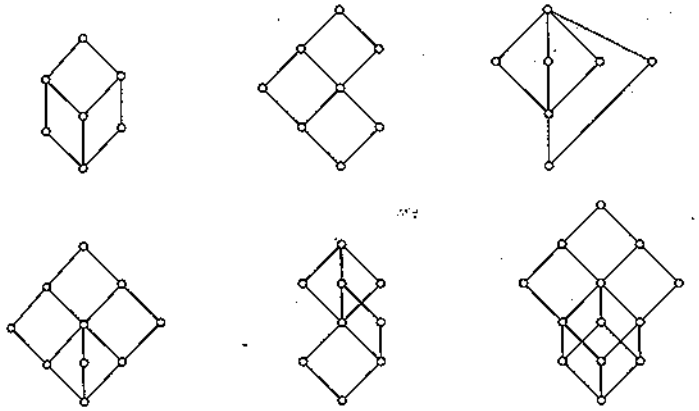


Figure 4.8

(a) $L \times K$ contains 20 elements;
(b) L is non-distributive;
(c) K has at least 3 elements;
(d) the greatest element of $L \times K$ covers precisely 4 elements.

4.8 Determine the lattices L and K to within isomorphism given that
(a) $|L \times K| = 18$,
(b) L is non-distributive,
(c) K has at least 3 elements,
(d) $L \times K$ is modular,
(e) the bottom element of $L \times K$ is covered by exactly 4 elements.
Fully explain your answer. Finish off by drawing and labelling L, K and $L \times K$.

4.9 (i) Let L be a distributive lattice and let $a, b, c \in L$. Prove that

$$(a \vee b = c \vee b \;\&\; a \wedge b = c \wedge b) \Longrightarrow a = c. \qquad (*)$$

(ii) Find elements a, b, c in $\mathbf{M}_3$ violating $(*)$. Do the same for $\mathbf{N}_5$.

(iii) Deduce that a lattice L is distributive if and only if $(*)$ holds for all $a, b, c \in L$.

4.10 Guess and then prove a characterization of modularity similar to the characterization of distributivity given in Exercise 4.9.

4.11 It was seen in Example 2.6(5) and Exercise 2.36 that when $\mathbb{N}_0$ is ordered by division the resulting ordered set is a complete lattice in which finite joins are given by lcm and finite meets by gcd.

(i) Use Exercise 4.9 to show that $\langle \mathbb{N}_0; \mathrm{lcm}, \mathrm{gcd}\rangle$ is distributive.

(ii) Let $A = \{3, 5, 7, \dots\}$ be the set of odd primes. Calculate $\bigvee A$ in $\langle \mathbb{N}_0; \preccurlyeq\rangle$. Hence show that $\langle \mathbb{N}_0; \preccurlyeq\rangle$ fails the **Join-Infinite Distributive law**

$$x \wedge \bigvee \{ y_i \mid i \in I \} = \bigvee \{ x \wedge y_i \mid i \in I \}.$$

(Compare this with Proposition 10.25.)

4.12 (i) Prove that a lattice L is distributive if and only if for each $a \in L$, the map $f_a \colon L \to {\downarrow}a \times {\uparrow}a$ defined by

$$f_a(x) = (x \wedge a, x \vee a) \text{ for all } x \in L$$

is a one-to-one homomorphism. [Use Exercise 4.9.]

(ii) Prove that, if L is distributive and possesses 0 and 1, then f_a (as defined in (i)) is an isomorphism if and only if a has a complement in L.

4.13 Let L be a lattice. Show that L is modular if and only if for all $a, b \in L$ the maps $j_b : x \mapsto x \vee b$ and $m_a : y \mapsto y \wedge a$ are mutually inverse lattice isomorphisms between $[a \wedge b, a]$ and $[b, a \vee b]$. Deduce that, if L is finite, then L is modular if and only if

$$[a \wedge b, a] \cong [b, a \vee b] \quad \text{for all } a, b \in L.$$

(Here $[c, d] := \{ x \in L \mid c \leqslant x \leqslant d \}$.)

4.14 (For those with some group theory behind them.)

(i) Prove that no group G satisfies $\mathrm{Sub}\, G \cong \mathbf{N}_5$.

(ii) Prove that if a group G satisfies $\mathrm{Sub}\, G \cong \mathbf{M}_3$, then $G \cong \mathbf{V}_4 \cong \mathbb{Z}_2 \times \mathbb{Z}_2$.

4.15 Show that the lattice $\mathrm{Sub}\, \mathbb{Z}$ of subgroups of the infinite cyclic group $\langle \mathbb{Z}; + \rangle$ is isomorphic to $\langle \mathbb{N}_0; \preccurlyeq\rangle^\partial$ and hence is distributive.

4.16 (i) Let L and K be lattices and assume that $\mathbf{N}_5 \rightarrowtail L \times K$. Show that $\mathbf{N}_5 \rightarrowtail L$ or $\mathbf{N}_5 \rightarrowtail K$.

(ii) Under the additional assumption that both L and K are modular, repeat (i) with $\mathbf{M}_3$ in place of $\mathbf{N}_5$.

4.17 (i) Use Remark 2.5(5) to show that $\mathbf{N}_5 \rightarrowtail L$ if and only if there exist five elements $u, a, b, c, v \in L$ such that

(a) $u < b < a < v$ and $u < c < v$,

(b) $a \wedge c = u$ and $b \vee c = v$.

(ii) Let $\{u, a, b, c, v\}$ be a sublattice of L isomorphic to $\mathbf{N}_5$ and assume $a_1, b_1 \in L$ with $b \leqslant b_1 < a_1 \leqslant a$. Show that $\{u, a_1, b_1, c, v\}$ is also a sublattice of L isomorphic to $\mathbf{N}_5$.

4.18 Assume that L is a lattice, J is an ideal of L, and G is a filter of L such that $L = J \cup G$ and $J \cap G \neq \varnothing$.

(i) Show that, if L has a sublattice M with $M \cong \mathbf{M}_3$, then either $M \subseteq J$ or $M \subseteq G$. (This does not require $J \cap G \neq \varnothing$.)

(ii) Show that if $x \in J$, $y \in G$ and $x \leqslant y$, then there exists $z \in J \cap G$ such that $x \leqslant z \leqslant y$. [Hint. Consider $x \vee (y \wedge w)$, where $w \in J \cap G$.]

(iii) (a) Prove that if $\mathbf{N}_5 \rightarrowtail L$, then $\mathbf{N}_5 \rightarrowtail J$ or $\mathbf{N}_5 \rightarrowtail G$.

[Hint. Let $N = \{u, a, b, c, v\}$: see Figure 4.9. The non-trivial case occurs when $u \in L \setminus G$ and $v \in L \setminus J$. Show that (up to duality) it suffices to consider the configuration on the left in Figure 4.9. Apply (ii) to obtain z, then add to the diagram the elements $z \wedge b$ and $a \vee (z \wedge b)$. Show that if $c \vee (z \wedge b) = a \vee (z \wedge b)$, then $\mathbf{N}_5 \rightarrowtail J$; otherwise, add $c \vee (z \wedge b)$ to the diagram and show that then $\mathbf{N}_5 \rightarrowtail G$. Exercise 4.17 will come in handy.]

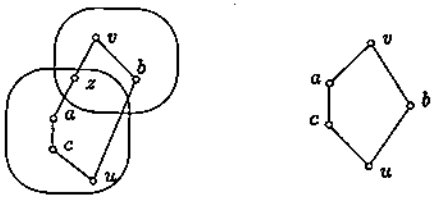


Figure 4.9

(b) Show by example that the assumption $J \cap G \neq \varnothing$ is necessary in (a).

(iv) Prove that L is a modular (distributive) lattice if and only if both J and G are modular (distributive) lattices.

(v) Use (iv) to prove that $\mathbf{M}_{3,3}$ is modular. (See Figure 4.5 and Example 4.12.)

4.19 Prove that the following are equivalent for a distributive lattice L:

(i) L is finite;

(ii) L is of finite length;

(iii) The set $\mathcal{J}(L)$ of join-irreducible elements of L is finite.

[Hint. For (ii) ⇒ (iii) let C be a (necessarily finite) maximal chain in L. Define $\varphi\colon \mathcal{J}(L) \to C$ by $\varphi(x) := \bigwedge(\uparrow x \cap C)$. Show that φ is one-to-one: assume that $\varphi(x) = \varphi(y) = a$, let b be the lower cover of a in C and calculate $x \wedge (b \vee y)$. For (iii) ⇒ (i), appeal to Theorem 2.46(iii).]

4.20 Let L be a lattice, let $S \subseteq L$ be such that $\bigvee S$ exists in L. Then $\bigvee S$ is called an **irredundant join** if $\bigvee(S \setminus \{s\}) \neq \bigvee S$ for all $s \in S$.

(i) Show that if $\bigvee S$ is an irredundant join then S is an antichain.

Let L be a lattice such that $\downarrow a$ has no infinite chains for all $a \in L$.

(ii) Show that every element a of L has a representation as an irredundant join, $a = \bigvee S$, of a finite subset S of $\mathcal{J}(L)$.

(iii) Show that in both $\mathbf{M}_3$ and $\mathbf{N}_5$ the representation guaranteed by (ii) is not always unique.

(iv) Show that if L is distributive, then every element of L has a unique representation as an irredundant join of join-irreducible elements. [Hint. Let $a \in L$ and let $S, T \subseteq \mathcal{J}(L)$ be such that $a = \bigvee S$ and $a = \bigvee T$ are irredundant. By Exercise 4.19, both S and T are finite. Consider $t \wedge \bigvee S$ for $t \in T$ and $s \wedge \bigvee T$ for $s \in S$.]

(v) Use Exercise 4.11(i) and Example 2.43(3) to show that when interpreted in the lattice $\langle \mathbb{N}_0; \operatorname{lcm}, \gcd\rangle$, part (iv) becomes a reformulation of the Fundamental Theorem of Arithmetic.

4.21 For $n \in \mathbb{N}_0$, consider the sublattice $L = \downarrow n$ of $\langle \mathbb{N}_0; \preccurlyeq\rangle$. When does $m \in L$ have a complement? Give a formula for m' in L when it exists. Characterize those $n \in \mathbb{N}_0$ such that $L = \downarrow n$ is a Boolean lattice.

4.22 For a Boolean lattice B and $a, b \in B$ such that $a \leqslant b$, show that the interval sublattice

$$[a, b] := \uparrow a \cap \downarrow b = \{x \in B \mid a \leqslant x \leqslant b\}$$

is a Boolean lattice. When is $[a, b]$ a Boolean subalgebra of B? [Hint. First show that for any distributive lattice L the map $f\colon L \to L$, given by $f(x) := (x \vee a) \wedge b$, is a homomorphism. Then calculate $f(L)$.]

4.23 Use Exercise 4.12 to give a proof by induction on $|B|$ that $|B| = 2^n$, for some $n \in \mathbb{N}_0$, for every finite Boolean lattice B.

4.24 Let S be a set and $f\colon S \to S$ any map. A subset A of S is called **f-invariant** if $x \in A$ implies $f(x) \in A$; note that $\varnothing$ is f-invariant. Denote the set of all f-invariant subsets of S by $\mathfrak{L}(S, f)$.

(i) Show that $\mathfrak{L}(S, f)$ is a lattice of sets.

(ii) Show that if S is finite and f is bijective, then $\mathfrak{L}(S, f)$ is an algebra of sets. Give an example to show that in general the finiteness of S is necessary here.

(iii) Prove that if $\mathfrak{L}(S, f)$ is an algebra of sets, then f is bijective.

4.25 Show that the following hold in all Boolean algebras:

(i) $(a \wedge b) \vee (a' \wedge b) \vee (a \wedge b') \vee (a' \wedge b') = 1$;

(ii) $(a \wedge b) \vee (a' \wedge c) = (a \wedge b) \vee (a' \wedge c) \vee (b \wedge c)$;

(iii) $a = b \iff (a \wedge b') \vee (a' \wedge b) = 0$;

(iv) $a \wedge b \leqslant c \vee d \iff a \wedge c' \leqslant b' \vee d$.

4.26 Let B be an ordered set such that $a \wedge b$ exists in B for all $a, b \in B$. Show that B is a Boolean lattice if and only if for all $x \in B$ there exists $x' \in B$ such that for all $y \in B$

$$x \leqslant y \iff x \wedge y' = 0,$$

where 0 is some fixed element of B. (This rather tricky exercise in axiomatics gives an extremely useful characterization of Boolean lattices as it eliminates the need to deal with joins and the distributive law. It is interesting that O. Frink's original proof of this result is short and non-computational but relies, unnecessarily, on Zorn's Lemma, (ZL) (see 10.2).)

4.27 A ring B with identity is called a **Boolean ring** if $x^2 = x$ for all $x \in B$.

(i) Show that the following identities hold in a Boolean ring:

$$\text{(a) } xy + yx = 0; \quad \text{(b) } x + x = 0; \quad \text{(c) } xy = yx.$$

(ii) Let B be a Boolean algebra and define $+$ and $\cdot$ on B by

$$x + y := (x \wedge y') \vee (x' \wedge y), \quad x \cdot y := x \wedge y.$$

Show that $\langle B; +, \cdot\rangle$ is a Boolean ring.

(iii) Let B be a Boolean ring and define $\vee$ and $\wedge$ on B by

$$x \vee y := x + y + xy, \quad x \wedge y := xy, \quad x' := 1 + x.$$

Show that $\langle B; \vee, \wedge, ', 0, 1\rangle$ is a Boolean algebra. [Hint. Use Exercise 4.26.]

(iv) Show that the correspondence between Boolean algebras and Boolean rings established in (ii) and (iii) is a bijective one.

4.28 Show that under the correspondence set up in the previous exercise, the Boolean algebra **2** corresponds to the Boolean ring $\mathbb{Z}_2$. Show that the additive group of the Boolean ring corresponding to the Boolean algebra $\mathbf{2}^2$ is isomorphic to the Klein 4-group. (In general, the additive group of the Boolean ring corresponding to $\mathbf{2}^n$ is isomorphic to $\mathbb{Z}_2^n$.)

4.29 Let B be a Boolean algebra. Prove that I is a lattice ideal in B if and only if I is a ring ideal in the ring $\langle B; +, \cdot \rangle$ defined in Exercise 4.27.

4.30 Use Theorem 4.23 to find a Boolean term $p(x,y)$ such that F_p coincides with the function *equal*: $\mathbf{2}^2 \to \mathbf{2}$, where

$$equal(x,y) := \begin{cases} 1 & \text{if } x = y, \\ 0 & \text{if } x \neq y. \end{cases}$$

4.31 For any non-empty set S, the **ternary discriminator** on S is a function d: $S^3 \to S$ given by

$$d(x,y,z) := \begin{cases} z & \text{if } x = y, \\ x & \text{if } x \neq y. \end{cases}$$

(This function plays an important role in universal algebra.)

(i) Give a formula for the ternary discriminator on **2** in terms of $\vee, \wedge, '$ and the function *equal*: $\mathbf{2}^2 \to \mathbf{2}$ of the previous exercise. Hence give a Boolean term $p(x,y,z)$ such that F_p coincides with the ternary discriminator on **2**.

(ii) Use Theorem 4.23 to find a Boolean term $q(x,y,z)$ such that F_q equals the ternary discriminator on **2**.

(iii) Show via the laws of Boolean algebra that for every Boolean algebra B and all $a, b, c \in B$ we have $p(a,b,c) = q(a,b,c)$.

(iv) Find a 'simple' term $r(x,y,z)$ such that F_r agrees with the ternary discriminator on **2**.

4.32 Design a gate diagram for a majority voting machine with four inputs which allows a current to flow when three or more of the inputs are 1.

4.33 A safe–unsafe decision circuit is required. There are to be three inputs and two outputs, one to operate a green light and the other to operate a red light. The green light should come on if all inputs indicate safety (1), and the red light should come on if any input indicates danger (0). Design the gate diagram.

4.34 The Duality Principle extends to Boolean algebras. Given a statement Φ about Boolean algebras, involving the symbols $\vee, \wedge, ', 0, 1$ and $\leqslant$, indicate the replacements needed to produce the dual statement Φ^∂. Hence formulate and prove the **Boolean Duality Principle**.

Define what it means for a Boolean term to be in **conjunctive normal form**, or CNF. (This is simply the dual of DNF.)

State the dual of Theorem 4.23 which ensures that every Boolean term is equivalent to a term in CNF.

4.35 For each of the following Boolean terms, p, draw up a truth table for F_p then apply Theorem 4.23 to obtain an equivalent polynomial in DNF. Show that the algorithm given in 4.24 yields the same DNF.

(i) $x \wedge (y \vee z)$.

(ii) $(x \vee y') \wedge z$.

(iii) $(x \vee y') \wedge (y \vee z') \wedge (z \vee x')$.

(iv) $((x \wedge y') \wedge z')' \wedge (x \vee z)$.

4.36 Test each of the following proposed Boolean algebra identities by reducing both sides to DNF:

(i) $(a \wedge b)' \vee (a \wedge c') = (a \wedge (b \vee c')')'$;

(ii) $(a \wedge b) \vee (a \wedge c') \vee (b' \wedge c') = (a \wedge b) \vee (b \vee c)'$;

(iii) $(a' \vee b') \vee (c' \wedge (a \vee b))' = a$.

4.37 The binary operation, $|$, of **inclusive denial** on a Boolean algebra is defined by $x \mid y := x' \vee y'$. Show that $\vee$, $\wedge$, $'$, 0 and 1 can each be defined in terms of $|$ alone. (It is a somewhat more difficult exercise to prove that if $*$ is a binary connective from which each of $\vee$, $\wedge$, $'$, 0 and 1 can be defined, then $*$ is logically equivalent to either inclusive denial or to its dual, $\downarrow$, defined by $x \downarrow y := x' \wedge y'$ and known as **joint denial**.)

4.38 Let B be a Boolean algebra and for $X \subseteq B$ let $[X]$ be the smallest subalgebra of B containing X (*cf.* Exercise 2.15). Show that

$$[X] = \{\, p(a_1, \ldots, a_n) \mid n \in \mathbb{N}_0, p \in \mathbf{BT}, a_1, \ldots, a_n \in B \,\}.$$

(Note that the case $n = 0$ is included to cover the '0–ary' Boolean terms 0 and 1.)

4.39 Find the subalgebra of the powerset algebra $\wp(\{1,2,3,4,5\})$ generated by $X = \{\{1,2\}, \{1,2,3,4\}\}$. Draw a diagram of $[X]$ and label each element with an appropriate Boolean term $p(a,b)$ where $a = \{1,2\}$ and $b = \{1,2,3,4\}$. (Compare with Exercise 2.15.)

5

Representation: the Finite Case

In previous chapters we have introduced various classes of lattices. We have given examples of members of these classes, and described some of their general properties. We now turn our attention to structure theorems. Later (see Chapters 10 and 11) we give a concrete representation, as a lattice of sets, of any (bounded) distributive lattice. This chapter deals, less ambitiously, with the finite case, and reveals a very satisfactory correspondence between finite distributive lattices and finite ordered sets. We show that any finite distributive lattice L can be realized as a lattice $\mathcal{O}(P)$ of down-sets built from a suitable subset P of L. We begin by discussing in general terms the problem of finding a subset of a lattice L which, as an ordered set, uniquely determines L.

Building blocks for lattices

This chapter draws on the final section of Chapter 2, concerned with join-irreducible elements. In the remarks below we also make links with the material on concept lattices from Chapter 3, but familiarity with this is not essential for the representation theory that follows.

5.1 Remarks on lattice-building. In 2.42 we defined a non-zero element x of a lattice L to be **join-irreducible** if $x = a \vee b$ implies $x = a$ or $x = b$ for all $a, b \in L$. Proposition 2.45 showed that if L satisfies (DCC), and hence certainly if L is finite, the set $\mathcal{J}(L)$ of join-irreducible elements of L is join-dense; that is, that every element of L can be obtained as a (possibly empty) join of elements from $\mathcal{J}(L)$.

We exploited join-dense subsets and, dually, meet-dense subsets of a complete lattice (in particular of a finite lattice) in our study of concept lattices. Theorem 3.9 says that we can reconstruct a complete lattice $\langle L; \leqslant \rangle$ from a join-dense subset G and a meet-dense subset M of L by forming all concepts of the context $(G, M, \leqslant)$. However this may be a substantial labour. Ideally, we should like a more direct way of building a lattice L from a suitable 'skeletal' subset P of L. We should like P to have the following properties:

(i) P is 'small' and readily identifiable;

(ii) L is uniquely determined by the ordered set P.

Even more nebulously, we should also like:

(iii) the process for obtaining L from P is simple to carry out.

Conditions (i) and (ii) pull in opposite directions, since (ii) requires P to be, in some sense, large. Good candidates for sets satisfying (ii) are, as we have already seen, those which are join-dense, or (dually) meet-dense, or both. However, the fact that Theorems 3.8 and 3.9 are converses to one another suggests that we cannot expect to improve significantly on the results in Chapter 3 so that in general (iii) is difficult to achieve. We return to these issues in 7.43.

In Chapter 4 we saw that many important lattices are distributive. There are very amenable concrete lattices of this type, namely down-set lattices and, in the Boolean case, powerset lattices. This encourages us to investigate whether the join-irreducible elements of a finite *distributive* lattice form a good skeleton for it. We shall see that the answer – in the finite case – is a resounding 'yes!'

Before studying distributive lattices we look at join-irreducible elements in the more special Boolean case, which is especially simple and easy to motivate.

5.2 Atoms. Our archetypal example of a Boolean algebra is a powerset algebra $\langle \wp(X); \cup, \cap, ', \varnothing, X \rangle$. Any $A \in \wp(X)$ is a union of singleton sets $\{x\}$ for $x \in A$; indeed the singletons are precisely the join-irreducible elements. But note too that the singletons are exactly the elements in $\wp(X)$ which cover 0.

Let L be a lattice with least element 0. Then $a \in L$ is called an **atom** if $0 \prec a$. The set of atoms of L is denoted by $\mathcal{A}(L)$. The lattice L is called **atomic** if, given $a \neq 0$ in L, there exists $x \in \mathcal{A}(L)$ such that $x \leqslant a$. Every finite lattice is atomic. By contrast, it may happen that an infinite lattice has no atoms at all. The chain of non-negative real numbers provides an example. Even a Boolean lattice may have no atoms; see Exercise 10.11.

The following lemma compares atoms and join-irreducible elements. It shows that in any Boolean lattice, $\mathcal{J}(L)$ coincides with $\mathcal{A}(L)$.

5.3 Lemma. *Let L be a lattice with least element 0. Then*

(i) $0 \prec x$ *in* L *implies* $x \in \mathcal{J}(L)$,

(ii) *if L is a Boolean lattice,* $x \in \mathcal{J}(L)$ *implies* $0 \prec x$.

Proof. To prove (i), suppose by way of contradiction that $0 \prec x$ and $x = a \vee b$ with $a < x$ and $b < x$. Since $0 \prec x$, we have $a = b = 0$, whence $x = 0$, $\lightning$.

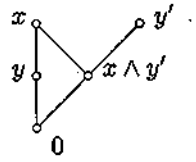


Figure 5.1

Now assume L is a Boolean lattice and that $x \in \mathcal{J}(L)$. Suppose $0 \leqslant y < x$; we want $y = 0$. We have (see Figure 5.1)

$$x = x \vee y = (x \vee y) \wedge (y' \vee y) = (x \wedge y') \vee y.$$

Since x is join-irreducible and $y < x$, we must have $x = x \wedge y'$, whence $x \leqslant y'$. But then $y = x \wedge y \leqslant y' \wedge y = 0$, so $y = 0$. This proves (ii). □

Finite Boolean algebras are powerset algebras

The set of atoms, $\mathcal{A}(L)$, of a finite Boolean lattice L admirably meets the building block criterion (i) in 5.1. For criterion (ii) we have the following lemma. This confirms the join-density of $\mathcal{A}(L)$ in L and is a consequence of Proposition 2.45 and Lemma 5.3. We give a direct proof because this is so simple.

5.4 Lemma. *Let B be a finite Boolean lattice. Then, for each $a \in B$,*

$$a = \bigvee \{ x \in \mathcal{A}(B) \mid x \leqslant a \}.$$

Proof. Fix $a \in B$. Let $S = \{ x \in \mathcal{A}(B) \mid x \leqslant a \}$. Certainly a is an upper bound for S. Let b be any upper bound for S. To complete the proof we require $a \leqslant b$. Suppose not, so $0 < a \wedge b'$, by 4.15(v). Choose $x \in \mathcal{A}(B)$ such that $0 \prec x \leqslant a \wedge b'$. Then $x \in S$, so $x \leqslant b$. Since $x \leqslant b'$ also holds, we have $x \leqslant b \wedge b' = 0$, ↯. □

The preceding lemma tells us how each individual element of L is determined by the atoms, but it doesn't by itself fulfill the aim of building block criterion (ii). For this we want to construct the lattice L as a whole from $\mathcal{A}(L)$. The next theorem tells us how to do this.

5.5 The representation theorem for finite Boolean algebras. *Let B be a finite Boolean algebra. Then the map*

$$\eta \colon a \longmapsto \{ x \in \mathcal{A}(B) \mid x \leqslant a \}$$

is an isomorphism of B onto $\wp(X)$, where $X = \mathcal{A}(B)$, with the inverse of η given by $\eta^{-1}(S) = \bigvee S$ for $S \in \wp(X)$.

Proof. We first show that η maps B onto $\wp(X)$. Clearly $\varnothing = \eta(0)$. Now let $S = \{a_1, \ldots, a_k\}$ be a non-empty set of atoms of B and define $a = \bigvee S$. We claim $S = \eta(a)$. Certainly $S \subseteq \eta(a)$. Now let x be any atom such that $x \leqslant a = a_1 \vee \cdots \vee a_k$. For each i, we have $0 \leqslant x \wedge a_i \leqslant x$. Because x is an atom, either $x \wedge a_i = 0$ for all i or there exists j such $x \wedge a_j = x$. In the former case, $x = x \wedge a = (x \wedge a_1) \vee \cdots \vee (x \wedge a_k) = 0$, ↯. Therefore $x \leqslant a_j$ for some j, which forces $x = a_j$, as a_j and x are atoms. Hence $\eta(a) \subseteq S$, as we wished to show.

Let $a, b \in B$. Then $\eta(a) \subseteq \eta(b)$ implies, by Lemma 5.4, that $a = \bigvee \eta(a) \leqslant \bigvee \eta(b) = b$. It is trivial (by the transitivity of $\leqslant$) that $\eta(a) \subseteq \eta(b)$ whenever $a \leqslant b$. So η is an order-isomorphism. By 2.19(ii) and 4.17, η is an isomorphism of Boolean algebras. □

5.6 Corollary. *Let B be a finite lattice. Then the following statements are equivalent:*

(i) *B is a Boolean lattice;*

(ii) *$B \cong \wp(\mathcal{A}(B))$;*

(iii) *B is isomorphic to $\mathbf{2}^n$, for some $n \geqslant 0$.*

Further, any finite Boolean lattice has 2^n elements, for some $n \geqslant 0$.

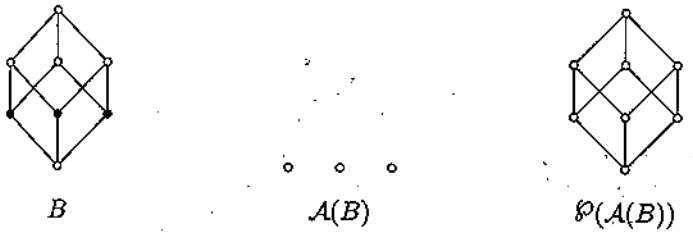


Figure 5.2

5.7 Remarks. To cement the ideas of this section, we illustrate 5.5 in the case that B is a Boolean algebra with 3 atoms; see Figure 5.2.

Corollary 5.6 is a very satisfactory result, but in a sense also a disappointing one. There is little variety among finite Boolean algebras. As we shall see, finite distributive lattices are a much richer class, but with an equally satisfying representation theory.

Finite distributive lattices are down-set lattices

We have seen that abstract finite Boolean algebras mimic finite powerset algebras, with atoms playing the role that singleton sets play in the concrete setting. In just the same way, we shall see that abstract finite distributive lattices mimic finite down-set lattices. Here atoms are replaced by join-irreducible elements and so our first task is to discover how $\mathcal{J}(\mathcal{O}(P))$ is related to P.

5.8 The join-irreducible elements of a down-set lattice. Let P be an ordered set. We claim that each set $\downarrow x$, for $x \in P$, is join-irreducible in $\mathcal{O}(P)$. Suppose that $\downarrow x = U \cup V$, where $U, V \in \mathcal{O}(P)$. Without loss of generality, $x \in U$. But then $\downarrow x \subseteq U$. Since $\downarrow x = U \cup V$ implies $U \subseteq \downarrow x$, we conclude that $\downarrow x = U$. This shows that $\downarrow x \in \mathcal{J}(\mathcal{O}(P))$.

Now assume that P is finite. Any non-empty $U \in \mathcal{O}(P)$ is the union of sets $\downarrow x_i$, $i = 1, \ldots, k$, where $x_i \parallel x_j$ for $i \neq j$ (recall Exercise 1.14). Unless $k = 1$, the set U is not join-irreducible. Hence $\mathcal{J}(\mathcal{O}(P)) = \{\downarrow x \mid x \in P\}$.

In the previous paragraph P must be finite: $\{q \in \mathbb{Q} \mid q < 0\}$ is join-irreducible in $\mathcal{O}(\mathbb{Q})$, but is not of the form $\downarrow x$.

5.9 Theorem. *Let P be a finite ordered set. Then the map $\varepsilon : x \mapsto \downarrow x$ is an order-isomorphism from P onto $\mathcal{J}(\mathcal{O}(P))$.*

Proof. Lemma 1.30 implies that ε is an order-embedding of P into $\mathcal{O}(P)$ and 5.8 gives that the image of ε is $\mathcal{J}(\mathcal{O}(P))$. □

For order-isomorphic ordered sets P and Q we have $\mathcal{O}(P) \cong \mathcal{O}(Q)$. Therefore Theorem 5.9 tells us that, when L is a finite down-set lattice $\mathcal{O}(P)$, we have $L \cong \mathcal{O}(\mathcal{J}(L))$. We now look at $\mathcal{O}(\mathcal{J}(L))$ more generally.

5.10 Examples and remarks. Figure 5.3 shows L, $\mathcal{J}(L)$ and $\mathcal{O}(\mathcal{J}(L))$ for some small lattices L (cf. 1.29). In the right-hand diagrams the join-irreducible elements have been shaded, to illustrate Theorem 5.9: in every case the shaded elements in $\mathcal{O}(\mathcal{J}(L))$ form an ordered set isomorphic to $\mathcal{J}(L)$.

Observe that $L \cong \mathcal{O}(\mathcal{J}(L))$ only in the first two examples, and that, of the four lattices L, just the first two are distributive. Since $\mathcal{O}(\mathcal{J}(L))$ is always distributive, we cannot have $L \cong \mathcal{O}(\mathcal{J}(L))$ unless L is distributive, and any proof establishing this isomorphism must make explicit use of distributivity of L. None of the results on join-irreducible elements proved so far brought in the distributive law. The next result does.

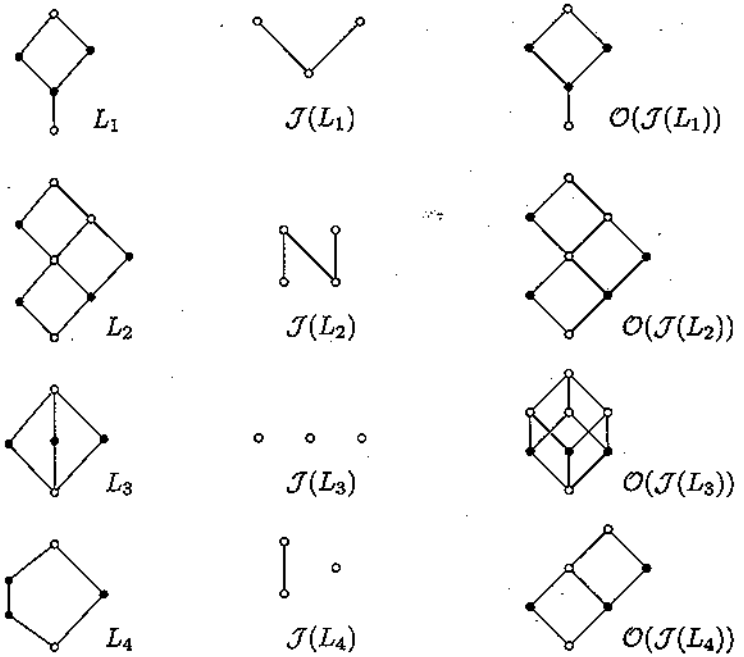


Figure 5.3

5.11 Lemma. *Let L be a distributive lattice and let $x \in L$, with $x \neq 0$ in case L has a zero. Then the following are equivalent:*

(i) *x is join-irreducible;*

(ii) *if $a, b \in L$ and $x \leqslant a \vee b$, then $x \leqslant a$ or $x \leqslant b$;*

(iii) *for any $k \in \mathbb{N}$, if $a_1, \ldots, a_k \in L$ and $x \leqslant a_1 \vee \cdots \vee a_k$, then $x \leqslant a_i$ for some i $(1 \leqslant i \leqslant k)$.*

Proof. To prove that (i) implies (ii), we assume that $x \in \mathcal{J}(L)$ and that $a, b \in L$ are such that $x \leqslant a \vee b$. We have

$$\begin{aligned} x &= x \wedge (a \vee b) && \text{(since } x \leqslant a \vee b\text{)} \\ &= (x \wedge a) \vee (x \wedge b) && \text{(since } L \text{ is distributive).} \end{aligned}$$

Because x is join-irreducible, $x = x \wedge a$ or $x = x \wedge b$. Hence $x \leqslant a$ or $x \leqslant b$, as required.

That (ii) implies (iii) is proved by induction on k; the case $k = 1$ is trivial and (ii) gets the induction started at $k = 2$.

It is trivial that (iii) implies (ii), so it only remains to deduce (i) from (ii). Suppose (ii) holds and that $x = a \vee b$. Then certainly $x \leqslant a \vee b$,

so $x \leqslant a$ or $x \leqslant b$. But $x = a \vee b$ forces $x \geqslant a$ and $x \geqslant b$. Hence $x = a$ or $x = b$. □

We can now justify the claim made in the title of this section by showing that every finite distributive lattice is (isomorphic to) a lattice of down-sets.

5.12 Birkhoff's representation theorem for finite distributive lattices. *Let L be a finite distributive lattice. Then the map $\eta\colon L \to \mathcal{O}(\mathcal{J}(L))$ defined by*

$$\eta(a) = \{x \in \mathcal{J}(L) \mid x \leqslant a\} \quad (= \mathcal{J}(L) \cap {\downarrow}a)$$

is an isomorphism of L onto $\mathcal{O}(\mathcal{J}(L))$.

Proof. It is immediate that $\eta(a) \in \mathcal{O}(\mathcal{J}(L))$ (since $\leqslant$ is transitive). By Proposition 2.19, it remains only to show that η is an order-isomorphism.

To prove that $a \leqslant b$ implies $\eta(a) \subseteq \eta(b)$, use 1.30. To prove that $\eta(a) \subseteq \eta(b)$ implies $a \leqslant b$, use Proposition 2.45 and 2.22(v) to obtain

$$a = \bigvee \eta(a) \leqslant \bigvee \eta(b) = b.$$

Finally, we prove that η is onto. Certainly $\varnothing = \eta(0)$. Now let $\varnothing \neq U \in \mathcal{O}(\mathcal{J}(L))$ and write $U = \{a_1, \ldots, a_k\}$. Define a to be $a_1 \vee \cdots \vee a_k$. We claim $U = \eta(a)$. To prove this, first let $x \in U$, so $x = a_i$ for some i. Then x is join-irreducible and $x \leqslant a$, hence $x \in \eta(a)$. In the reverse direction, suppose $x \in \eta(a)$. Then $x \leqslant a = a_1 \vee \cdots \vee a_k$ and Lemma 5.11 implies $x \leqslant a_i$ for some i. Since U is a down-set and $a_i \in U$, we have $x \in U$. □

5.13 Corollary. *Let L be a finite lattice. Then the following statements are equivalent:*

(i) *L is distributive;*

(ii) *$L \cong \mathcal{O}(\mathcal{J}(L))$;*

(iii) *L is isomorphic to a down-set lattice;*

(iv) *L is isomorphic to a lattice of sets;*

(v) *L is isomorphic to a sublattice of $\mathbf{2}^n$ for some $n \geqslant 0$.*

5.14 Remark. In 5.10 we stressed that no non-distributive lattice could be isomorphic to a down-set lattice. Birkhoff's representation theorem provides an alternative to the $\mathbf{M}_3$–$\mathbf{N}_5$ Theorem for establishing non-distributivity of a finite lattice L; see 4.11. If $L \cong \mathcal{O}(\mathcal{J}(L))$ fails, then L cannot be distributive. This applies, for example, to $L_3 = \mathbf{M}_3$ and $L_4 = \mathbf{N}_5$ in 5.10. Further illustrations are given in 5.16.

Finite distributive lattices and finite ordered sets in partnership

With Birkhoff's representation theorem in hand, our knowledge of finite ordered sets becomes a powerful asset in the study of finite distributive lattices. Questions concerning finite distributive lattices can often be answered by converting them into simpler but equivalent questions concerning ordered sets. The map $\mathcal{J}$ from finite distributive lattices to finite ordered sets is the distributive-lattice theorist's logarithm. In particular, it maps products to sums.

5.15 The join-irreducible elements of a product lattice. Consider the product $L_1 \times L_2$ of lattices L_1 and L_2 each with a least element, but not necessarily distributive. First note that $(x_1, x_2) = (x_1, 0) \vee (0, x_2)$. Thus (x_1, x_2) is not join-irreducible unless either x_1 or x_2 is zero. Further, $x_1 = a_1 \vee b_1$ in L_1 implies $(x_1, 0) = (a_1, 0) \vee (b_1, 0)$. It follows that

$$\mathcal{J}(L_1 \times L_2) \subseteq (\mathcal{J}(L_1) \times \{0\}) \cup (\{0\} \times \mathcal{J}(L_2)).$$

It is readily seen that the reverse inclusion also holds. It follows easily that we have an order-isomorphism

$$\mathcal{J}(L_1 \times L_2) \cong \mathcal{J}(L_1) \dot\cup \mathcal{J}(L_2).$$

For an example, see Figure 5.4. The join-irreducible elements are shaded.

Figure 5.4

Now assume that L_1 and L_2 are finite and distributive. In this case, the result of the previous paragraph can be derived from Theorem 5.9, Birkhoff's representation theorem and the fact that $\mathcal{O}(P_1 \dot\cup P_2)$ is isomorphic to $\mathcal{O}(P_1) \times \mathcal{O}(P_2)$ (see 1.32):

$$\begin{aligned}\mathcal{J}(L_1 \times L_2) &\cong \mathcal{J}(\mathcal{O}(\mathcal{J}(L_1)) \times \mathcal{O}(\mathcal{J}(L_2))) \\ &\cong \mathcal{J}(\mathcal{O}(\mathcal{J}(L_1) \dot\cup \mathcal{J}(L_2))) \cong \mathcal{J}(L_1) \dot\cup \mathcal{J}(L_2).\end{aligned}$$

5.16 Examples. Consider Figure 5.5.

(1) Consider the lattice L_1. The ordered set $\mathcal{J}(L_1)$ is shown alongside. Since $\mathcal{J}(L_1) \cong \mathbf{1} \dot\cup (\mathbf{1} \oplus \overline{\mathbf{2}})$, by 1.32 we have $\mathcal{O}(\mathcal{J}(L_1)) \cong \mathbf{2} \times (\mathbf{1} \oplus \mathbf{2}^2)$, which has 10 elements. We deduce that L_1 is not isomorphic to $\mathcal{O}(\mathcal{J}(L_1))$, so that L_1 is not distributive.

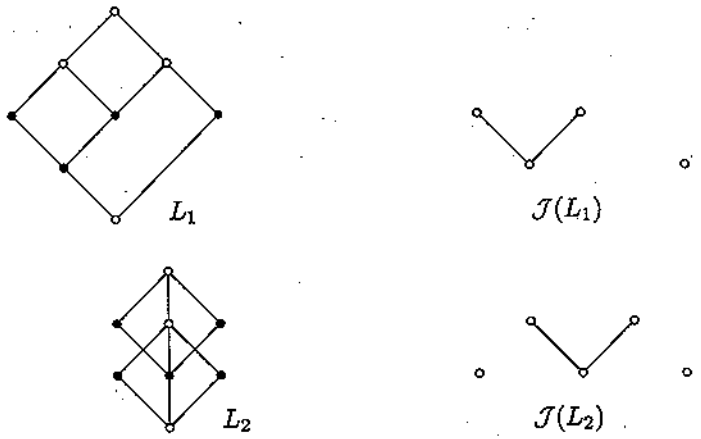


Figure 5.5

(2) Now consider L_2. We could immediately describe $\mathcal{O}(\mathcal{J}(L_2))$. Instead, we note that $L_2 \cong L_2^\partial$ yet $\mathcal{J}(L_2)$ is not isomorphic to its order dual. Hence, by 1.31, L_2 cannot be isomorphic to $\mathcal{O}(\mathcal{J}(L_2))$ and consequently is not distributive.

5.17 A fruitful partnership. We denote by $\mathbf{D}_F$ the class of all finite distributive lattices and by $\mathbf{P}_F$ the class of all finite ordered sets. Theorems 5.12 and 5.9 assert that

$$L \cong \mathcal{O}(\mathcal{J}(L)) \quad \text{and} \quad P \cong \mathcal{J}(\mathcal{O}(P))$$

for all $L \in \mathbf{D}_F$ and $P \in \mathbf{P}_F$. We call $\mathcal{J}(L)$ the **dual** of L and $\mathcal{O}(P)$ the **dual** of P. (The use of the word dual here should of course not be confused with that in 1.19.)

When we identify each finite distributive lattice L with the isomorphic lattice $\mathcal{O}(\mathcal{J}(L))$ of down-sets of $\mathcal{J}(L)$, we may regard $\mathbf{D}_F$ as consisting of the concrete lattices $\mathcal{O}(P)$, for $P \in \mathbf{P}_F$, rather than as abstract objects satisfying certain identities.

Up to isomorphism, we have a one-to-one correspondence

$$\mathcal{O}(P) = L \quad \underset{\longleftarrow}{\longrightarrow} \quad P = \mathcal{J}(L)$$

for $L \in \mathbf{D}_F$ and $P \in \mathbf{P}_F$.

Describing $P = \mathcal{J}(L)$, given L, is entirely straightforward. Those who worked through Exercise 1.13 will appreciate that describing L, given P, can be laborious, even when P is quite small. Computer

programs have been devised for drawing $\mathcal{O}(P)$ for a given finite ordered set P. These are viable only so long as $\mathcal{O}(P)$ remains reasonably small (of the order of hundreds). It is possible for $|\mathcal{O}(P)|$ to grow extremely fast as P increases, so that the problem of determining $\mathcal{O}(P)$ from P becomes intractable, even with computer assistance. The example $P = \wp(X)$, for $|X| = 1, 2, 3, \ldots$, is instructive. Figure 5.6 shows $\wp(X)$ and $\mathcal{O}(\wp(X))$ for $|X| = 3$ and the accompanying table $|\wp(X)|$ and $|\mathcal{O}(\wp(X))|$ for $|X| \leqslant 8$. The size of $\mathcal{O}(\wp(X))$ for $|X| = 9$ remains elusive.

$\lvert X\rvert$	$\lvert\wp(X)\rvert$	$\lvert\mathcal{O}(\wp(X))\rvert$
1	2	3
2	4	6
3	8	20
4	16	168
5	32	7581
6	64	7828354
7	128	2414682040998
8	256	5613043722868755790788

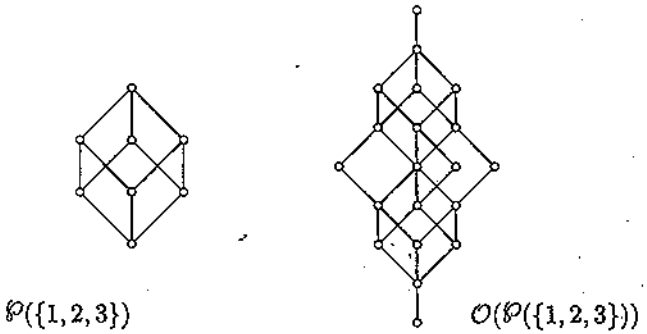


Figure 5.6

The above observations show that the dual of a finite distributive lattice is generally much smaller and less complex than the lattice itself. This means that lattice problems concerning $\mathbf{D}_F$ are likely to become simpler when translated into problems about $\mathbf{P}_F$. We may regard the maps $L \mapsto \mathcal{J}(L)$ and $P \mapsto \mathcal{O}(P)$ as playing a role analogous to that of the logarithm and exponential functions (and this analogy is strengthened by 5.15).

Special properties of a finite distributive lattice are reflected in special properties of its dual. The following lemma provides an elementary example. The descriptions of $\mathcal{O}(\overline{\mathbf{n}})$ and $\mathcal{O}(\mathbf{n})$ come from 1.32. The proofs are left as easy exercises.

5.18 Lemma. *Let $L = \mathcal{O}(P)$ be a finite distributive lattice. Then*

(i) *L is a Boolean lattice if and only if P is an antichain; $\mathcal{O}(\overline{\mathbf{n}}) = 2^n$.*

(ii) *L is a chain if and only if P is a chain; $\mathcal{O}(\mathbf{n}) = \mathbf{n+1}$.*

Our discussion of the partnership between $\mathbf{D}_F$ and $\mathbf{P}_F$ would be seriously incomplete were we not to consider structure-preserving maps (recall 1.38). Theorem 5.19 sets up a one-to-one correspondence between $\{0,1\}$-homomorphisms from $\mathcal{O}(P)$ to $\mathcal{O}(Q)$ and order-preserving maps from Q to P, for $P, Q \in \mathbf{P}_F$ (note the reversal of the direction). This theorem is harder to formulate and to prove than any of our preceding results on duals. Some of the difficulty stems from our hitherto admirable choice of the join-irreducible elements as the basis for our representation theory. In Chapter 11 we remove the finiteness restrictions under which we are currently working. We shall then obtain a more natural version of Theorem 5.19, as a special case of Theorem 11.31. In the meantime we recommend only to the most intrepid readers the exercise of proving Theorem 5.19 directly.

5.19 Theorem. *Let P and Q be finite ordered sets and let $L = \mathcal{O}(P)$ and $K = \mathcal{O}(Q)$.*

Given a $\{0,1\}$-homomorphism $f\colon L \to K$, there is an associated order-preserving map $\varphi_f\colon Q \to P$ defined by

$$\varphi_f(y) = \min\{\, x \in P \mid y \in f(\downarrow x)\,\}$$

for all $y \in Q$.

Given an order-preserving map $\varphi\colon Q \to P$, there is an associated $\{0,1\}$-homomorphism $f_\varphi\colon L \to K$ defined by

$$f_\varphi(a) = \varphi^{-1}(a) \text{ for all } a \in L.$$

Equivalently,

$$\varphi(y) \in a \text{ if and only if } y \in f_\varphi(a) \quad \text{for all } a \in L, \quad y \in Q.$$

The maps $f \mapsto \varphi_f$ and $\varphi \mapsto f_\varphi$ establish a one-to-one correspondence between $\{0,1\}$-homomorphisms from L to K and order-preserving maps from Q to P.

Further,

(i) *f is one-to-one if and only if φ_f is onto,*

(ii) *f is onto if and only if φ_f is an order-embedding.*

5.20 Example. Figure 5.7 shows an order-preserving map $\varphi\colon Q \to P$ and the associated $\{0,1\}$-homomorphism $f\colon \mathcal{O}(P) \to \mathcal{O}(Q)$. The image of f is shaded.

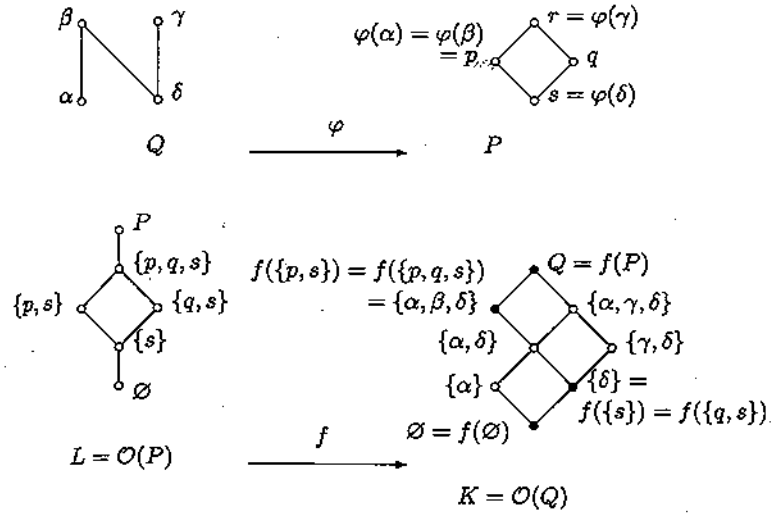


Figure 5.7

5.21 Stocktaking. Combining 5.17 and 5.19, we have a correspondence between $\mathbf{D}_F$ + $\{0,1\}$-homomorphisms and $\mathbf{P}_F$ + order-preserving maps which establishes what in technical parlance is called a **duality** or a **dual equivalence of categories**. In Chapter 11 we specify more explicitly what we mean by a duality, in the context of the representation of all bounded distributive lattices. The import of the duality in the finite case should already be clear: statements about finite distributive lattices can be translated into statements about finite ordered sets, and vice versa. We can now see that our two uses of the word 'dual' have an underlying commonality. If, in an ordered set P, we think of $x \leqslant y$ as representing an 'arrow' from x to y, then P^∂ is obtained by reversing the arrows. Similarly, for $L, K \in \mathbf{D}_F$, a $\{0,1\}$-homomorphism $f\colon L \to K$ provides an 'arrow' from L to K, and Theorem 5.19 shows that when we pass from $\mathbf{D}_F$ to $\mathbf{P}_F$ the arrows again reverse. Category theory is exactly the tool needed to formalize this hand waving. The step up into the wide blue, category-theoretic yonder is not a large one but is beyond the scope of our work here.

Exercises

5.1 Consider the lattices in Figure 5.8.

(i) Draw labelled diagrams of the ordered sets $\mathcal{J}(L)$ and $\mathcal{M}(L)$ for each of the lattices.

(ii) Draw a labelled diagram of $\mathcal{O}(\mathcal{J}(L))$ in each case and comment on your results.

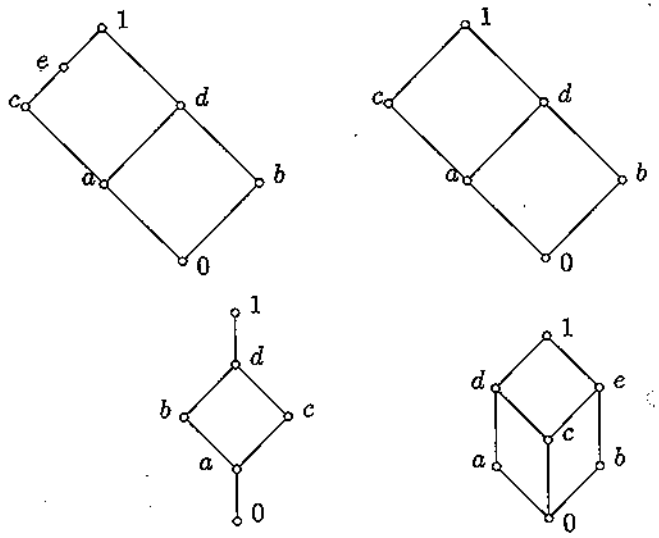


Figure 5.8

5.2 Verify Theorem 5.9 for the ordered sets given in Figure 5.9.

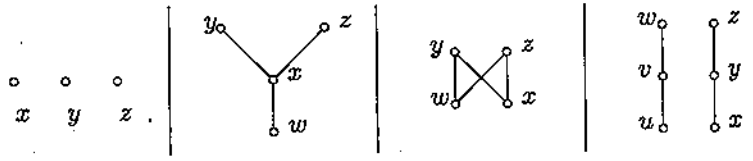


Figure 5.9

5.3 Let L be a distributive lattice and define $\eta\colon L \to \mathcal{O}(\mathcal{J}(L))$ as in Birkhoff's representation theorem. Prove directly that η preserves $\vee$, that is, $\eta(a \vee b) = \eta(a) \cup \eta(b)$ for all $a, b \in L$.

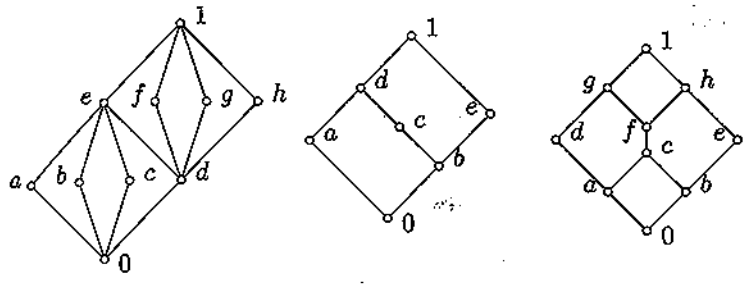


Figure 5.10

5.4 Use Theorem 5.12 to show that the lattices in Figure 5.10 are not distributive.

5.5 Let L be $\langle \mathbb{N}_0; \mathrm{lcm}, \gcd \rangle$.

(i) Draw diagrams of $\downarrow 4$ and $\downarrow 12$.

(ii) Show that for every $k \geqslant 1$ there exists $n_k \in \mathbb{N}$ such that $\downarrow n_k \cong \wp(\{1, 2, \ldots, k\})$.

(iii) Deduce that every finite distributive lattice can be embedded into L.

(iv) Give an example of a countable distributive lattice which cannot be embedded into L.

5.6 Let L be a finite distributive lattice. Prove that there exists a finite Boolean lattice B and an embedding $\eta\colon L \to B$ such that η is a $\{0,1\}$-homomorphism. Show further that, if $|L| = n$, then B can be chosen so that $|B| \leqslant 2^{n-1}$.

5.7 Let L be a finite distributive lattice. Prove by the steps below that $\mathcal{J}(L) \cong \mathcal{M}(L)$. (An alternative proof, by duality, is indicated in Exercise 5.8.)

(i) Let $x \in \mathcal{J}(L)$. Show that there exists $\widehat{x} \in L$ such that $\downarrow \widehat{x} = L \setminus \uparrow x$. [Hint. Let $\widehat{x} := \bigvee(L \setminus \uparrow x)$ and then use Lemma 5.11 to show that $\widehat{x} \not\geqslant x$.]

(ii) Show that for all $x \in \mathcal{J}(L)$ the element $\widehat{x}$ defined in (i) is meet-irreducible.

(iii) Prove that the map $\varphi\colon \mathcal{J}(L) \to \mathcal{M}(L)$, given by $\varphi(x) = \widehat{x}$ for all $x \in \mathcal{J}(L)$, is an order-isomorphism. [Hint. Recall from Lemma 1.30 that $x \leqslant y$ if and only if $\uparrow x \supseteq \uparrow y$. When proving that φ maps onto $\mathcal{M}(L)$, use the dual of (i) and (ii).]

5.8 Let P be a finite ordered set.

(i) Show that a down-set U is meet-irreducible in $\mathcal{O}(P)$ if and only if it is of the form $P \setminus \uparrow x$ for some $x \in P$.

(ii) Use (i) to show that P is order-isomorphic to $\mathcal{M}(\mathcal{O}(P))$.

(iii) Conclude that $\mathcal{J}(\mathcal{O}(P)) \cong \mathcal{M}(\mathcal{O}(P))$. Use Birkhoff's representation theorem to deduce that $\mathcal{J}(L)$ is order-isomorphic to $\mathcal{M}(L)$ for any finite distributive lattice L.

5.9 Give counterexamples to each of the following statements.

(i) If L and K are finite lattices and $\mathcal{J}(L)$ and $\mathcal{J}(K)$ are order-isomorphic, then L and K are isomorphic.

(ii) If L is a distributive lattice, then $L \cong \mathcal{O}(\mathcal{J}(L))$. [Hint. Consider a suitable infinite chain.]

(iii) If L is a finite distributive lattice and $\mathcal{J}(L)$ is a lattice, then $\mathcal{J}(L)$ is a sublattice of L.

(iv) If L is a finite lattice, then $\mathcal{J}(L) \cong \mathcal{M}(L)$. [Hint. Note Exercise 5.7.]

5.10 Use Lemma 5.3(i) and the fact that $\mathcal{J}(L_1 \times L_2) \cong \mathcal{J}(L_1) \dot{\cup} \mathcal{J}(L_2)$, for finite lattices L_1 and L_2, to answer the question posed in Exercise 2.17.

5.11 Let L be a finite distributive lattice.

(i) Use Exercise 1.17 and Birkhoff's representation theorem to prove that the following conditions are equivalent:

(a) the only complemented elements of L are 0 and 1;

(b) L is **indecomposable** (that is, there do not exist lattices L_1, L_2 with $|L_1| > 1$, $|L_2| > 1$ such that $L \cong L_1 \times L_2$).

(ii) Consider the lattice K shown in Figure 5.11, which you may assume to be distributive. Show that K is indecomposable.

5.12 Let L, H_1, H_2, K_1 and K_2 be finite distributive lattices such that

$$H_1 \times H_2 \cong L \cong K_1 \times K_2.$$

Prove that there exist finite distributive lattices M_1, M_2, M_3 and M_4 such that

$$L \cong M_1 \times M_2 \times M_3 \times M_4,$$

with

$$H_1 \cong M_1 \times M_2, \quad H_2 \cong M_3 \times M_4,$$
$$K_1 \cong M_1 \times M_3, \quad K_2 \cong M_2 \times M_4.$$

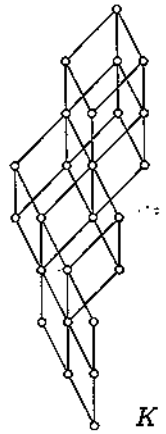


Figure 5.11

5.13 Prove that the length of a finite distributive lattice L equals $|\mathcal{J}(L)|$. [Hint. Use Exercise 1.12.]

5.14 Let L be a finite distributive lattice. Prove that the width of $\mathcal{J}(L)$ equals the least $n \in \mathbb{N}$ such that L can be embedded into a product of n chains. [Hint. Use the duality between $\mathbf{D}_F$ and $\mathbf{P}_F$ to reinterpret Dilworth's Theorem – see Exercise 1.30.]

5.15 Consider the (distributive) lattice $\langle \mathbb{N}_0; \mathrm{lcm}, \gcd \rangle$ and let $n \in \mathbb{N}$.

(i) Describe $\mathcal{J}(\downarrow n)$ (recall 2.43(3)).

(ii) Let $n = p_1^{k_1} \dots p_s^{k_s}$ with the p_i pairwise distinct primes. Show that

$$\downarrow n \cong (\mathbf{k}_1 \oplus \mathbf{1}) \times \cdots \times (\mathbf{k}_s \oplus \mathbf{1}).$$

5.16 Prove that the lattice $\mathrm{Sub}\,\mathbb{Z}_n$ of subgroups of the cyclic group $\mathbb{Z}_n = \{0, 1, \dots, n-1\}$, under addition modulo n, is a finite product of finite chains. [Hint. First prove that $\mathrm{Sub}\,\mathbb{Z}_n$ is isomorphic to the dual of the principal ideal $\downarrow n$ in $\langle \mathbb{N}_0; \mathrm{lcm}, \gcd \rangle$, then prove that $(\downarrow n)^\partial \cong \downarrow n$ and finally apply Exercise 5.15.]

5.17 Let $m_1, m_2, \dots, m_s \in \mathbb{N}$ with $m_i > 1$ for all i. Use the duality between $\mathbf{D}_F$ and $\mathbf{P}_F$ to prove that

$$\mathbf{m}_1 \times \cdots \times \mathbf{m}_s \cong \mathbf{2}^t \Longrightarrow m_1 = \ldots = m_s = 2 \text{ and } s = t.$$

5.18 Use Exercises 5.16 and 5.17 along with 4.6(4) to characterize those groups G such that $\mathrm{Sub}\,G$ is (i) a finite chain, (ii) isomorphic to $\wp(X)$ for some finite set X.

5.19 Recall from 1.37 that, if P and Q are ordered sets, then $Q^{(P)}$, or alternatively $\langle P \to Q \rangle$, denotes the set of order-preserving maps from P to Q with the pointwise order.

(i) Let L be a lattice and P an ordered set. Show that $L^{(P)}$ is a sublattice of L^P and hence is distributive whenever L is.

(ii) Use Exercises 1.25 and 1.26 to show that, for all ordered sets P and Q,
$$\mathcal{O}(Q)^{(P)} \cong \mathcal{O}(P^\partial \times Q).$$

(iii) Conclude that, if $L \in \mathbf{D}_F$ and $P \in \mathbf{P}_F$, then
$$\mathcal{J}(L^{(P)}) \cong P^\partial \times \mathcal{J}(L).$$

(iv) Hence draw diagrams of $\mathbf{2}^{(\mathbf{4})}$ and $\mathbf{4}^{(\mathbf{2})}$.

5.20 Let L be a finite distributive lattice with $|L| > 1$ and let X denote the set of all $\{0,1\}$-homomorphisms from L to $\mathbf{2}$ ordered pointwise.

(i) Let $x \in \mathcal{J}(L)$ and define $f_x : L \to \mathbf{2}$ by
$$f_x(a) = \begin{cases} 1 & \text{if } a \geqslant x, \\ 0 & \text{if } a \ngeq x, \end{cases}$$
that is, f_x is the characteristic function of $\uparrow x$. Show that f_x is a $\{0,1\}$-homomorphism.

(ii) Prove that the map $\varepsilon : \mathcal{J}(L) \to X^\partial$, defined by $\varepsilon(x) = f_x$ for all $x \in \mathcal{J}(L)$, is an order-isomorphism from $\mathcal{J}(L)$ onto X^∂ whose inverse $\eta : X^\partial \to \mathcal{J}(L)$ is given by
$$\eta(f) = \bigwedge \{ a \in L \mid f(a) = 1 \} \text{ for all } f \in X.$$

5.21 Consider the distributive lattices L_1–L_4 shown in Figure 5.12.

(i) Is it possible to find an onto $\{0,1\}$-homomorphism $f : L \to K$ where $L = L_1$ and $K = L_2$? (Use Theorem 5.19(ii) to justify your answer.)

(ii) Repeat (i) with $L = L_1$ and $K = L_3$.

(iii) Repeat (i) with $L = L_3$ and $K = L_4$.

(iv) Repeat (i) with $L = \mathbf{1} \oplus \mathbf{3}^3 \oplus \mathbf{2}$ and $K = \mathbf{1} \oplus (\mathbf{2} \times \mathbf{3}) \oplus \mathbf{1}$.

(v) Repeat (i) with $L = \mathbf{2}^2 \overline{\oplus} \mathbf{2}^2$ and $K = \mathbf{4}$.

5.22 A lattice L with 0 is said to be **pseudocomplemented** if, for each $a \in L$, there exists an element $a^* \in L$ such that, for all $b \in L$,
$$a \wedge b = 0 \iff b \leqslant a^*,$$
that is, $a^* = \max\{ b \in L \mid a \wedge b = 0 \}$.

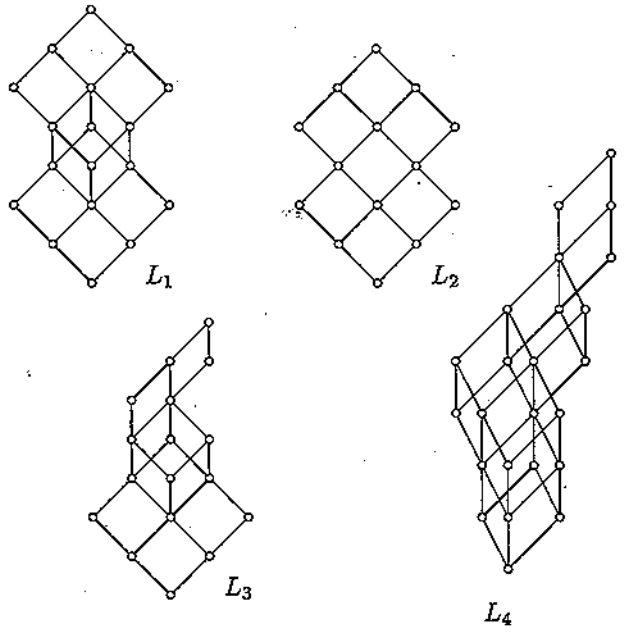


Figure 5.12

(i) Show that any Boolean lattice is pseudocomplemented.

(ii) Show that any bounded chain is pseudocomplemented.

(iii) Show that any finite distributive lattice is pseudocomplemented.

(iv) Prove that, if $P \in \mathbf{P}_F$, then for each $a \in L := \mathcal{O}(P)$ we have $a^* = P \setminus \uparrow a$. (Here $\uparrow a$ is calculated in P not in L.)

(v) Give an example of a bounded distributive lattice which is not pseudocomplemented.

5.23 Let P and Q be finite ordered sets and let $L = \mathcal{O}(P)$ and $K = \mathcal{O}(Q)$ be the corresponding finite distributive lattices. Assume that $f : L \to K$ is a $\{0,1\}$-homomorphism and that $\varphi : Q \to P$ is the dual of f. (See Theorem 5.19.)

Show that the following are equivalent:

(i) $f(a^*) = (f(a))^*$ for all $a \in L$ (where a^* is as given in Exercise 5.22);

(ii) $\operatorname{Min}(\varphi(y)) = \varphi(\operatorname{Min}(y))$ for all $y \in Q$ (where $\operatorname{Min}(z)$ denotes the set of minimal elements in $\downarrow z$).

6

Congruences

Lattices of congruences play a central role in lattice theory and in algebra more widely. This chapter develops the rudiments of a theory which goes way beyond the scope of an introductory text such as this.

Introducing congruences

In group theory courses it is customary, after homomorphisms have been introduced, to go on to define normal subgroups and quotient groups (factor groups) and to reveal the intimate connection between these concepts that is summed up in the fundamental Homomorphism Theorem (also called the First Isomorphism Theorem). We begin with a summary of the basic group theory results, expressed in a form that will make the parallels with the lattice case stand out clearly. This summary is prefaced by a brief refresher on equivalence relations.

6.1 Equivalence relations: a recap. We recall that an equivalence relation on a set A is a binary relation on A which is reflexive, symmetric and transitive. We write $a \equiv b \pmod{\theta}$ or $a\,\theta\,b$ to indicate that a and b are related under the relation θ; we use instead the notation $(a,b) \in \theta$ where it is appropriate to be formally correct and to regard θ as a subset of $A \times A$.

An equivalence relation θ on A gives rise to a partition of A into non-empty disjoint subsets. These subsets are the equivalence classes or **blocks** of θ; we shall usually use the term block, as it is more in keeping with the pictorial approach we shall be adopting. A typical block is of the form $[a]_\theta := \{\, x \in A \mid x \equiv a \pmod{\theta} \,\}$. In the opposite direction, a partition of A into a union of non-empty disjoint subsets gives rise to an equivalence relation whose blocks are the subsets in the partition. This correspondence between equivalence relations and partitions is set out more explicitly in, for example, [9].

6.2 The group case. Let G and H be groups and $f\colon G \to H$ be a group homomorphism. We may define an equivalence relation θ on G by

$$(\forall a, b \in G)\; a \equiv b \pmod{\theta} \iff f(a) = f(b).$$

This relation and the partition of G it induces have the following important properties.

(1) The relation θ is compatible with the group operation in the sense that, for all $a, b, c, d \in G$,

$$a \equiv b \pmod{\theta}\ \&\ c \equiv d \pmod{\theta} \implies ac \equiv bd \pmod{\theta}.$$

(2) The block $N = [1]_\theta := \{\, g \in G \mid g \equiv 1 \pmod{\theta} \,\}$ is a normal subgroup of G.

(3) For each $a \in G$, the block $[a]_\theta := \{\, g \in G \mid g \equiv a \pmod{\theta} \,\}$ equals the (left) coset $aN := \{\, an \mid n \in N \,\}$.

(4) The natural definition,

$$[a]_\theta[b]_\theta := [ab]_\theta \quad \text{for all } a, b \in G,$$

yields a well-defined group operation on $\{\, [a]_\theta \mid a \in G \,\}$; by (2) and (3), the resulting group is precisely the quotient group G/N and hence (by the Homomorphism Theorem for groups) is isomorphic to the subgroup $f(G)$ of H.

6.3 Speaking universally. It should now be apparent that much of the above is not particular to groups and group homomorphisms, but will apply, *mutatis mutandis*, to lattices and lattice homomorphisms. In fact, the natural setting for the Homomorphism Theorem and its consequences is neither group theory nor lattice theory but **universal algebra.** This is the general theory of classes of algebraic structures, of which groups, rings, lattices, bounded lattices, vector spaces, ... are examples. Lattice theory and universal algebra have a close and symbiotic relationship: results from universal algebra (such as the Homomorphism Theorem) specialize to classes of lattices, and lattices arise naturally in the study of abstract algebras, as lattices of congruences, for example. For references, see Appendix B.

We now introduce congruences on lattices, as in the group case using homomorphisms as our starting point. We say that an equivalence relation θ on a lattice L is **compatible with join and meet** if, for all $a, b, c, d \in L$,

$$a \equiv b \pmod{\theta} \text{ and } c \equiv d \pmod{\theta}$$

imply

$$a \vee c \equiv b \vee d \pmod{\theta} \text{ and } a \wedge c \equiv b \wedge d \pmod{\theta}.$$

6.4 Lemma. *Let L and K be lattices and let $f\colon L \to K$ be a lattice homomorphism. Then the equivalence relation θ defined on L by*

$$(\forall a, b \in L)\; a \equiv b \pmod{\theta} \iff f(a) = f(b)$$

is compatible with join and meet.

Proof. It is elementary that θ is indeed an equivalence relation. Now assume $a \equiv b \pmod{\theta}$ and $c \equiv d \pmod{\theta}$, so that $f(a) = f(b)$ and $f(c) = f(d)$. Hence, since f preserves join,

$$f(a \vee c) = f(a) \vee f(c) = f(b) \vee f(d) = f(b \vee d).$$

Therefore $a \vee c \equiv b \vee d \pmod{\theta}$. Dually, θ is compatible with meet. □

6.5 Definitions and examples. An equivalence relation on a lattice L which is compatible with both join and meet is called a **congruence** on L. If L and K are lattices and $f\colon L \to K$ is a lattice homomorphism, then the associated congruence θ on L, defined in 6.4, is known as the kernel of f and is denoted by $\ker f$. The set of all congruences on L is denoted by $\operatorname{Con} L$. Examples of homomorphisms and their kernels are given in Figure 6.1.

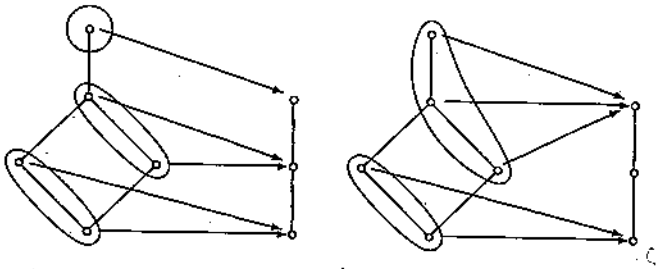


Figure 6.1

Note that a congruence on a lattice L can be indicated on a diagram by placing a loop around the elements in each block of the corresponding partition.

The following lemma is handy when calculating with congruences.

6.6 Lemma.

(i) *An equivalence relation θ on a lattice L is a congruence if and only if, for all $a, b, c \in L$,*

$a \equiv b \pmod{\theta} \implies a \vee c \equiv b \vee c \pmod{\theta}$ *and* $a \wedge c \equiv b \wedge c \pmod{\theta}$.

(ii) *Let θ be a congruence on L and let $a, b, c \in L$.*

(a) *If $a \equiv b \pmod{\theta}$ and $a \leqslant c \leqslant b$, then $a \equiv c \pmod{\theta}$.*

(b) *$a \equiv b \pmod{\theta}$ if and only if $a \wedge b \equiv a \vee b \pmod{\theta}$.*

Proof. (i) Assume that θ is a congruence on L. If $a \equiv b \pmod{\theta}$, then, since $c \equiv c \pmod{\theta}$, we have $a \vee c \equiv b \vee c \pmod{\theta}$ and $a \wedge c \equiv b \wedge c \pmod{\theta}$. The converse is left as an exercise.

(ii) Let θ be a congruence on L. To prove (a), note first that $a \leqslant c \leqslant b$ implies $a = a \wedge c$ and $c = b \wedge c$. Assume $a \equiv b \pmod{\theta}$. Then $a \wedge c \equiv b \wedge c \pmod{\theta}$, so $a \equiv c \pmod{\theta}$.

Finally we consider (b). If $a \equiv b \pmod{\theta}$, then $a \vee a \equiv b \vee a \pmod{\theta}$ and $a \wedge a \equiv b \wedge a \pmod{\theta}$ by the definition of a congruence. The lattice identities imply $a \equiv a \vee b \pmod{\theta}$ and $a \equiv a \wedge b \pmod{\theta}$. Since θ is transitive and symmetric, we deduce $a \wedge b \equiv a \vee b \pmod{\theta}$.

Conversely, assume $a \wedge b \equiv a \vee b \pmod{\theta}$. We have $a \wedge b \leqslant a \leqslant a \vee b$, so $a \wedge b \equiv a \pmod{\theta}$, by (a), and similarly $a \wedge b \equiv b \pmod{\theta}$. Because θ is symmetric and transitive, it follows that $a \equiv b \pmod{\theta}$. □

6.7 Quotient lattices. Given an equivalence relation θ on a lattice L there is a natural way to try to define operations $\vee$ and $\wedge$ on the set

$$L/\theta := \{\, [a]_\theta \mid a \in L \,\}$$

of blocks. Namely, for all $a, b \in L$, we 'define'

$$[a]_\theta \vee [b]_\theta := [a \vee b]_\theta \text{ and } [a]_\theta \wedge [b]_\theta := [a \wedge b]_\theta.$$

These formulæ will produce well-defined operations precisely when they are independent of the elements chosen to represent the equivalence classes, that is, when

$$[a_1]_\theta = [a_2]_\theta \text{ and } [b_1]_\theta = [b_2]_\theta$$

imply

$$[a_1 \vee b_1]_\theta = [a_2 \vee b_2]_\theta \text{ and } [a_1 \wedge b_1]_\theta = [a_2 \wedge b_2]_\theta,$$

for all $a_1, a_2, b_1, b_2 \in L$. Since, for all $a_1, a_2 \in L$,

$$[a_1]_\theta = [a_2]_\theta \Leftrightarrow a_1 \in [a_2]_\theta \Leftrightarrow a_1 \equiv a_2 \pmod{\theta},$$

it follows that $\vee$ and $\wedge$ are well defined on L/θ if and only if θ is a congruence. When θ is a congruence on L, we call $\langle L/\theta; \vee, \wedge \rangle$ the **quotient lattice of L modulo θ**. Our next lemma justifies this terminology.

6.8 Lemma. *Let θ be a congruence on the lattice L. Then $\langle L/\theta; \vee, \wedge \rangle$ is a lattice and the natural quotient map $q\colon L \to L/\theta$, defined by $q(a) := [a]_\theta$, is a homomorphism.*

We can now state the Homomorphism Theorem for lattices. Its proof is a routine verification.

6.9 Theorem. *Let L and K be lattices, let f be a homomorphism of L onto K and define $\theta = \ker f$. Then the map $g\colon L/\theta \to K$, given by $g([a]_\theta) = f(a)$ for all $[a]_\theta \in L/\theta$, is well defined, that is,*

$$(\forall a, b \in L)\, [a]_\theta = [b]_\theta \text{ implies } g([a]_\theta) = g([b]_\theta).$$

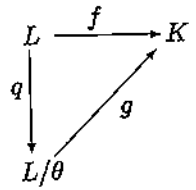


Figure 6.2

Moreover g is an isomorphism between L/θ and K. Furthermore, if q denotes the quotient map, then $\ker q = \theta$ *and the diagram in Figure 6.2 commutes, that is,* $g \circ q = f$.

Theorem 6.10 is, likewise, the Boolean algebra version of the Homomorphism Theorem; as for lattices, the proof is a straightforward verification. An equivalence relation θ on a Boolean algebra B is a **Boolean congruence** if it is a lattice congruence such that $a \equiv b \pmod{\theta}$ implies $a' \equiv b' \pmod{\theta}$, for all $a, b \in B$.

6.10 Theorem. *Let B and C be Boolean algebras, let f be a Boolean homomorphism of B onto C and define $\theta = \ker f$. Then θ is a Boolean congruence and the map $g\colon B/\theta \to C$, given by $g([a]_\theta) = f(a)$ for all $[a]_\theta \in B/\theta$, is a well-defined isomorphism between B/θ and C.*

Congruences and diagrams

In our discussion of congruences we have so far treated lattices as algebraic structures; the underlying order on a lattice has not been mentioned. In this section we redress the balance.

6.11 Blocks of congruences. Some examples of congruences and the resulting quotient lattices are given in Figure 6.3.

When considering the blocks of a congruence θ on L, it is best to think of each block X as an entity in its own right rather than as the block $[a]_\theta$ associated with some $a \in L$, as the latter gives undue emphasis to the element a. Intuitively, the quotient lattice L/θ is obtained by collapsing each block to a point.

Assume we are given a diagram of a finite lattice L and loops are drawn on the diagram representing a partition of L. Two natural geometric questions arise.

(a) How can we tell if the equivalence relation corresponding to the partition is a congruence?

(b) If we know that the loops define the blocks of a congruence θ, how do we go about drawing L/θ?

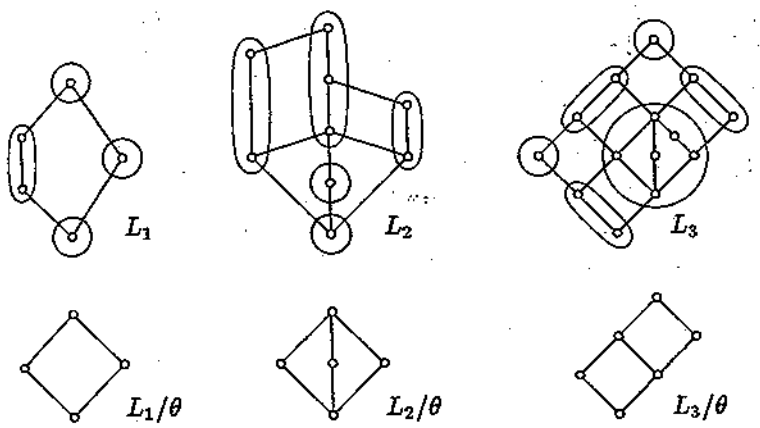


Figure 6.3

There is, of course, no definitive way to draw a lattice diagram. By providing a description of the order and the covering relation on L/θ, Lemma 6.12 provides as good an answer as we can expect to Question (b). The proof of the lemma is left as an easy exercise.

6.12 Lemma. *Let θ be a congruence on a lattice L and let X and Y be blocks of θ.*

(i) *$X \leqslant Y$ in L/θ if and only if there exist $a \in X$ and $b \in Y$ such that $a \leqslant b$.*

(ii) *$X \prec Y$ in L/θ if and only if $X < Y$ in L/θ and $a \leqslant c \leqslant b$ implies $c \in X$ or $c \in Y$, for all $a \in X$, all $b \in Y$ and all $c \in L$.*

(iii) *If $a \in X$ and $b \in Y$, then $a \vee b \in X \vee Y$ and $a \wedge b \in X \wedge Y$.*

We pursue an answer to Question (a) by first looking for properties that the blocks of a congruence must possess. The blocks are certainly sublattices and are convex. (A subset Q of an ordered set P is **convex** if $x \leqslant z \leqslant y$ implies $z \in Q$ whenever $x, y \in Q$ and $z \in P$.) A third property, also with a geometric flavour, relates elements in different blocks. It is described in 6.13. Theorem 6.14 shows that these properties of blocks characterize congruences among equivalence relations.

6.13 The quadrilateral argument. Let L be a lattice and suppose that $\{a, b, c, d\}$ is a 4-element subset of L. Then a, b and c, d are said to be opposite sides of the **quadrilateral** $\langle a, b; c, d\rangle$ if $a < b$, $c < d$ and either

$$(a \vee d = b \text{ and } a \wedge d = c) \quad \text{or} \quad (b \vee c = d \text{ and } b \wedge c = a).$$

We say that the blocks of a partition of L are **quadrilateral-closed** if whenever a, b and c, d are opposite sides of a quadrilateral and $a, b \in A$ for some block A then $c, d \in B$ for some block B (see Figure 6.4). Note that for a covering pair $a \prec b$, we often indicate $a \equiv b \pmod{\theta}$ on a diagram by drawing a wavy line from a to b, to be thought of as a spring which collapses a and b together. See the $\mathbf{N}_5$ in Figure 6.6.

Figure 6.4

6.14 Theorem. *Let L be a lattice and let θ be an equivalence relation on L. Then θ is a congruence if and only if*

(i) *each block of θ is a sublattice of L,*

(ii) *each block of θ is convex,*

(iii) *the blocks of θ are quadrilateral-closed.*

Proof. Although the necessity of (i), (ii) and (iii) follows easily from Lemma 6.6, we elect to give a proof based on Lemma 6.12 as this illustrates the use of blocks of θ rather than the relation θ itself. Assume that θ is a congruence on L and let X and Y be blocks of θ.

(i) If $a, b \in X$, then $a \vee b \in X \vee X = X$ and $a \wedge b \in X \wedge X = X$, by 6.12(iii). Hence X is a sublattice of L.

(ii) Let $a, b \in X$, let $c \in L$ with $a \leqslant c \leqslant b$ and assume that c belongs to the block Z of θ. Then, by 6.12(i), we have $X \leqslant Z \leqslant X$ in L/θ and hence $X = Z$. Thus $c \in Z = X$ and hence X is convex.

(iii) Let a, b and c, d be opposite sides of a quadrilateral, with $a \vee d = b$ and $a \wedge d = c$ (see Figure 6.4). We assume that $a, b \in X$ and $d \in Y$ and seek to prove that $c \in Y$. Since $d \leqslant b$ we have $Y \leqslant X$ (by 6.12(i)) and thus $c = a \wedge d \in X \wedge Y = Y$, as required.

The converse is much harder. Assume that (i), (ii) and (iii) hold. By Lemma 6.6, θ is a congruence provided that, for all $a, b, c \in L$,

$$a \equiv b \pmod{\theta} \text{ implies } a \vee c \equiv b \vee c \pmod{\theta} \text{ and } a \wedge c \equiv b \wedge c \pmod{\theta}.$$

Let $a, b, c \in L$ with $a \equiv b \pmod{\theta}$. By duality it is enough to show that $a \vee c \equiv b \vee c \pmod{\theta}$. Define $X := [a]_\theta = [b]_\theta$. Since X is a sublattice

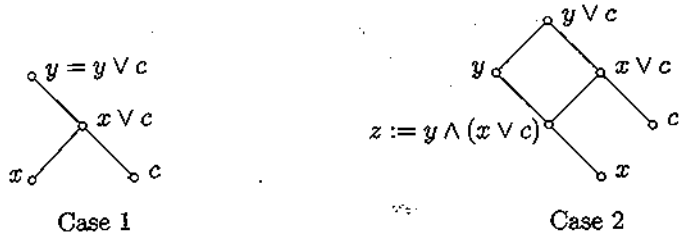


Figure 6.5

of L, we have $x := a \wedge b \in X$ and $y := a \vee b \in X$. Our first claim is that $x \vee c \equiv y \vee c \pmod{\theta}$. We distinguish two cases (see Figure 6.5).

Case 1: $c \leqslant y$. We have $x \leqslant x \vee c \leqslant y \vee c = y$; the second inequality holds because $x \leqslant y$ (see 2.5(4)) and the final equality follows from the Connecting Lemma. Hence $x \vee c \equiv y \vee c \pmod{\theta}$, since the block X contains both x and y and is convex.

Case 2: $c \nleqslant y$. Since $x \leqslant y$, we have $x \vee c \leqslant y \vee c$. If $x \vee c = y \vee c$, then $x \vee c \equiv y \vee c \pmod{\theta}$ as θ is reflexive; thus we may assume that $x \vee c < y \vee c$. Since x is a lower bound of $\{y, x \vee c\}$ we have $x \leqslant z := y \wedge (x \vee c) \leqslant y$. Now $x \leqslant z \leqslant x \vee c$ implies $z \vee c = x \vee c$ (see 2.5(4)) and hence $z \neq y$ as $y \vee c > x \vee c$. Consequently, z, y and $x \vee c, y \vee c$ are opposite sides of a quadrilateral. Since the block X is convex and $x, y \in X$, it follows that $z \in X$. Since $z, y \in X$ and θ is quadrilateral-closed it follows that $x \vee c$ and $y \vee c$ belong to the same block, say Y. Thus $x \vee c \equiv y \vee c \pmod{\theta}$, as claimed.

We show $a \vee c \equiv b \vee c \pmod{\theta}$ by showing that $a \vee c$ and $b \vee c$ both belong to the block Y. Since $a \wedge b \leqslant a \leqslant a \vee b$ and $a \wedge b \leqslant b \leqslant a \vee b$ we have (by 2.5(4) again) $x \vee c = (a \wedge b) \vee c \leqslant a \vee c \leqslant a \vee b \vee c = y \vee c$ and $x \vee c = (a \wedge b) \vee c \leqslant b \vee c \leqslant a \vee b \vee c = y \vee c$. Since $x \vee c, y \vee c \in Y$ and Y is convex, it follows that $a \vee c, b \vee c \in Y$. □

The lattice of congruences of a lattice

So far, we have looked at congruences one at a time. We now look at the congruences on a lattice collectively. Much of the significance of congruences comes from the fact that the congruences of a lattice (or any other algebraic structure) form a lattice.

6.15 The lattice of congruences of a lattice. An equivalence relation θ on a lattice L is a subset of L^2 (see 6.1). We can rewrite the compatibility conditions in the form

$$(a, b) \in \theta \text{ and } (c, d) \in \theta \text{ imply } (a \vee c, b \vee d) \in \theta \text{ and } (a \wedge c, b \wedge d) \in \theta.$$

As this says precisely that θ is a sublattice of L^2, we could define congruences to be those subsets of L^2 which are both equivalence relations and sublattices of L^2. With this viewpoint, the set $\operatorname{Con} L$ of congruences on a lattice L is a family of sets, and is ordered by inclusion. It is easily seen to be a topped $\bigcap$-structure on L^2. Hence $\operatorname{Con} L$, when ordered by inclusion, is a complete lattice, by 2.32. The least element, $\mathbf{0}$, and greatest element, $\mathbf{1}$, are given by $\mathbf{0} = \{(a,a) \mid a \in L\}$ and $\mathbf{1} = L^2$.

6.16 Principal congruences. The characterization in 6.14 allows us to see how a congruence spreads through a lattice L, that is, it permits us to answer the question:

> 'Which pairs of elements c, d in L must be collapsed in order to obtain a congruence θ which collapses the pair a, b?'

The smallest congruence collapsing a given pair of elements a and b is denoted by $\theta(a,b)$; it is called the **principal congruence generated by** (a,b). Since $\operatorname{Con} L$ is a topped $\bigcap$-structure, $\theta(a,b)$ exists for all $(a,b) \in L^2$: indeed,

$$\theta(a,b) = \bigwedge \{\theta \in \operatorname{Con} L \mid (a,b) \in \theta\}.$$

The diagrams of $\mathbf{N}_5$ and $\mathbf{M}_3$ in Figure 6.6 show the partitions corresponding to the principal congruences $\theta(a,1)$ and $\theta(0,c)$ respectively. We use the quadrilateral argument to justify these claims.

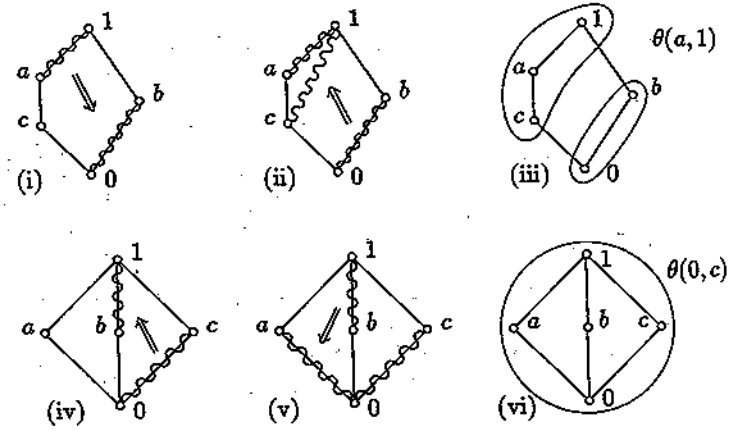


Figure 6.6

(1) To find the blocks of the principal congruence $\theta(a,1)$ on $\mathbf{N}_5$, we first use the quadrilateral $\langle a,1;0,b\rangle$ to show that $a \equiv 1$ implies $0 \equiv b$ (here $\equiv$ denotes equivalence with respect to $\theta(a,1)$. The quadrilateral $\langle 0,b;c,1\rangle$ yields $c \equiv 1 \pmod{\theta}$. Since blocks of $\theta(a,1)$ are convex, we deduce that $a, c, 1$ lie in the same block. It is clear that $\{0,b\}$ and $\{a,c,1\}$ are convex sublattices and together are quadrilateral-closed. Thus they form the blocks of $\theta(a,1)$ on $\mathbf{N}_5$.

(2) The diagrams in Figure 6.6(iv)–(vi) illustrate the application of the quadrilateral argument to $\mathbf{M}_3$, starting with the pair $(0,c)$. After step (ii) we deduce that $a, c, 0$ lie in the same block, say A. Since the blocks of a congruence are sublattices, we have $1 = a \vee c \in A$ and $0 = a \wedge c \in A$. Thus, since blocks are convex, A is the only block. Hence $\theta(0,c) = \mathbf{1}$.

The next lemma indicates why principal congruences are important.

6.17 Lemma. *Let L be a lattice and let $\theta \in \operatorname{Con} L$. Then*

$$\theta = \bigvee \{\theta(a,b) \mid (a,b) \in \theta\}.$$

Consequently the set of principal congruences is join-dense in $\operatorname{Con} L$.

Proof. We verify that θ is the least upper bound in $\langle \operatorname{Con} L; \subseteq\rangle$ of the set $S = \{\theta(a,b) \mid (a,b) \in \theta\}$. First note that the definition of $\theta(a,b)$ implies that $\theta(a,b) \subseteq \theta$ whenever $(a,b) \in \theta$. Therefore θ is an upper bound for S. Now assume that ψ is any upper bound for S. This means that $\theta(a,b) \subseteq \psi$ for any pair $(a,b) \in \theta$. But $(a,b) \in \theta(a,b)$ always, so $(a,b) \in \theta$ implies $(a,b) \in \psi$, as required. □

6.18 The join of two congruences. As frequently occurs with topped $\bigcap$-structures, the join in $\operatorname{Con} L$ is not generally given by set union: the union of two equivalence relations is often not an equivalence relation because transitivity fails. We now describe finite joins in $\operatorname{Con} L$.

Let L be a lattice and let $\alpha, \beta \in \operatorname{Con} L$. We say that a sequence $z_0, z_1, \ldots, z_n$ **witnesses** $a\,(\alpha \vee \beta)\,b$ if $a = z_0$, $z_n = b$ and $z_{k-1}\,\alpha\,z_k$ or $z_{k-1}\,\beta\,z_k$ for $1 \leqslant k \leqslant n$. We claim that $a\,(\alpha \vee \beta)\,b$ if and only if for some $n \in \mathbb{N}$ there exists a sequence $z_0, z_1, \ldots, z_n$ which witnesses $a\,(\alpha \vee \beta)\,b$. To prove the claim, define a relation θ on L by $a\,\theta\,b$ if and only if for some $n \in \mathbb{N}$ there exists a sequence $z_0, z_1, \ldots, z_n$ which witnesses $a\,(\alpha \vee \beta)\,b$. The following are straightforward to check:

(i) $\theta \in \operatorname{Con} L$,

(ii) $\alpha \subseteq \theta$ and $\beta \subseteq \theta$, and

(iii) if $\alpha \subseteq \gamma$ and $\beta \subseteq \gamma$ for some $\gamma \in \operatorname{Con} L$, then $\theta \subseteq \gamma$.

Consequently θ is indeed the least upper bound of α and β in $\operatorname{Con} L$.

A property which distinguishes lattices (and, more generally, lattices with additional operations) from algebras in general is that the congruence lattice of a lattice is distributive. This turns out to have many far-reaching consequences. The study of congruence-distributive algebras, as they are called, is integral to any modern course on universal algebra. Our proof that congruence lattices of lattices are distributive has a universal algebraic flavour: it shows that the crux of the matter is three simple identities satisfied by the **median** term, $m(x,y,z) := (x \wedge y) \vee (y \wedge z) \vee (z \wedge x)$, *viz.*

$$m(x,x,y) = m(x,y,x) = m(y,x,x) = x,$$

rather than anything intrinsically lattice theoretic.

6.19 Theorem. *The lattice* $\operatorname{Con} L$ *is distributive for any lattice* L.

Proof. Let $\alpha, \beta, \gamma \in \operatorname{Con} L$. By Lemma 4.1 it will suffice to show that $\alpha \wedge (\beta \vee \gamma) \leqslant (\alpha \wedge \beta) \vee (\alpha \wedge \gamma)$. Assume that $a\,(\alpha \wedge (\beta \vee \gamma))\,b$. Then $a\,\alpha\,b$ and by 6.18 there is a sequence $a = z_0, z_1, \ldots, z_n = b$ which witnesses $a\,(\beta \vee \gamma)\,b$. We wish to construct a sequence which witnesses $a\,((\alpha \wedge \beta) \vee (\alpha \wedge \gamma))\,b$. The identities satisfied by the median term give in particular $a = m(a,b,z_0)$ and $b = m(a,b,z_n)$. Furthermore, since $a\,\alpha\,b$, for $i = 0, \ldots, n-1$, we have $m(a,b,z_i)\,\alpha\,m(a,b,z_{i+1})$, since

$$m(a,b,z_i)\,\alpha\,m(a,a,z_i) = a = m(a,a,z_{i+1})\,\alpha\,m(a,b,z_{i+1}).$$

Observe also that, if $c\,\theta\,d$, then $m(a,b,c)\,\theta\,m(a,b,d)$ for all $c, d \in L$ and all $\theta \in \operatorname{Con} L$. For $i = 0, \ldots, n-1$ we can apply this with $c = z_i$, $d = z_{i+1}$ and θ as either β or γ. We deduce that the sequence

$$a = m(a,b,z_0), m(a,b,z_1), \ldots, m(a,b,z_n) = b$$

witnesses $a\,((\alpha \wedge \beta) \vee (\alpha \wedge \gamma))\,b$. □

6.20 Groups revisited. Let G be a group. We showed in 6.1 that there is a correspondence between normal subgroups of G and equivalence relations compatible with the group structure, that is, group congruences. Denote the set of all such congruences by $\operatorname{Con} G$. Each congruence is regarded as a subset of $G \times G$ and $\operatorname{Con} G$ is given the inclusion order inherited from $\mathcal{P}(G \times G)$. This makes $\operatorname{Con} G$ into a topped $\bigcap$-structure, and so a complete lattice, in just the same way that $\operatorname{Con} L$ is, for L a lattice. It is then easy to see that $\operatorname{Con} G \cong \mathcal{N}\text{-Sub}\, G$. We have already seen that $\mathcal{N}\text{-Sub}\, G$ is modular (see 4.6(5)). Consequently, $\operatorname{Con} G$ is modular. However, even for very small groups it may not be distributive. For $G = \mathbf{V}_4$, the Klein 4-group, for example, we have $\mathcal{N}\text{-Sub}\, G \cong \mathbf{M}_3$.

A very important problem in group theory, solved around 1980, was the classification of all finite simple groups. A group is defined to be simple if it has no proper non-trivial normal subgroups. In a similar way, it is natural to investigate lattices L which are **simple** in the sense that they have precisely two congruences, namely $\mathbf{0}$ and $\mathbf{1}$. Since it can be proved that every lattice can be embedded into a simple lattice, simple lattices can be very complicated!

Exercises

Exercises from the text. Prove Lemma 6.8, Theorem 6.9, Theorem 6.10 and Lemma 6.12. Prove that $\operatorname{Con} G \cong \mathcal{N}\text{-Sub}\, G$, for every group G.

6.1 Let $\leqslant$ be a quasi-order on a set P, that is, $\leqslant$ is reflexive and transitive. Define θ on P by $a \equiv b \pmod{\theta} \Leftrightarrow a \leqslant b \;\&\; b \leqslant a$. Prove that the relation $\sqsubseteq$ defined on P/θ by

$$[a]_\theta \sqsubseteq [b]_\theta \iff (\exists x \in [a]_\theta)(\exists y \in [b]_\theta)\, x \leqslant y$$

is an order relation and that $[a]_\theta \sqsubseteq [b]_\theta$ if and only if $a \leqslant b$.

6.2 Draw the diagram of L/θ for each θ shown in Figure 6.7.

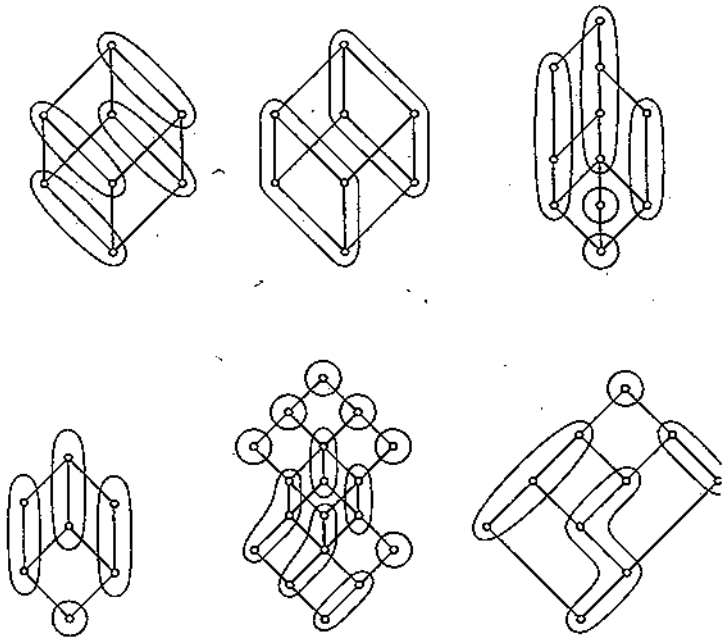


Figure 6.7

6.3 Show that a homomorphism $f: L \to K$ is an embedding if and only if $\ker f = \mathbf{0}$.

6.4 Let J be an ideal of a lattice L and define a relation θ_J on L by

$$\theta_J := \{(a,b) \in L^2 \mid (\exists c \in J)\, a \vee c = b \vee c\}.$$

Prove that L is distributive if and only if for every ideal J of L, the relation θ_J is a congruence on L and J is a block of the corresponding partition of L.

6.5 Show that an equivalence relation on a Boolean algebra B is a Boolean congruence if and only if it is a lattice congruence. Prove that every congruence on a Boolean algebra B is of the form θ_J for some ideal J in B (see Exercise 6.4) and moreover $(a,b) \in \theta_J$ if and only if $(a \wedge b') \vee (a' \wedge b) \in J$.

6.6 Let L be a lattice and let $a, b, c, d \in L$.

(i) Show that $\theta(a,b) \subseteq \theta(c,d)$ if and only if $a \equiv b \,(\mathrm{mod}\, \theta(c,d))$.

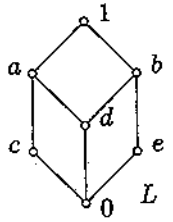
(ii) Show that $\theta(a,b) = \theta(a \wedge b, a \vee b)$.

6.7 Let L be a distributive lattice and assume that $c \leqslant d$ in L. Prove that $(a,b) \in \theta(c,d) \iff a \wedge c = b \wedge c \;\&\; a \vee d = b \vee d$.

6.8 Consider the lattices L and K in Figure 6.8.

(i) Assume that $\theta \in \mathrm{Con}\, L$ and that $a \equiv c \,(\mathrm{mod}\, \theta)$. Show that $b \equiv e \,(\mathrm{mod}\, \theta)$. Hence find $\theta(a,c)$ and then draw $L/\theta(a,c)$.

(ii) Assume that $\theta \in \mathrm{Con}\, K$ and that $a \equiv b \,(\mathrm{mod}\, \theta)$. Show that $b \equiv 0 \,(\mathrm{mod}\, \theta)$. Hence find $\theta(a,b)$ and then draw $L/\theta(a,b)$.



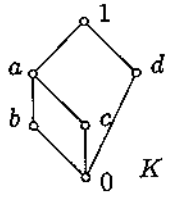


Figure 6.8

6.9 Let L be the lattice in Figure 6.9.

(i) Let θ be a congruence on L with $0 \equiv a \,(\mathrm{mod}\, \theta)$. Show carefully that $d \equiv 1 \,(\mathrm{mod}\, \theta)$.

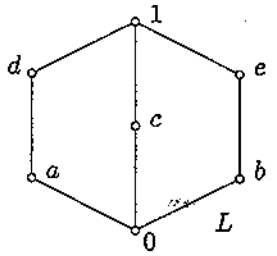


Figure 6.9

(ii) Let $\theta \in \mathrm{Con}\, L$ with $0 \equiv a \,(\mathrm{mod}\, \theta)$. Show that $0 \equiv 1 \,(\mathrm{mod}\, \theta)$ and hence explain why $\theta = \mathbf{1}$ in $\mathrm{Con}\, L$.

6.10 Give an example of a congruence $\theta \in \mathrm{Con}\, \mathbf{4}$ such that θ is not principal.

6.11 Find all congruences on $\mathbf{N}_5$, then draw the lattice $\mathrm{Con}\, \mathbf{N}_5$. [Hint. First find all principal congruences on $\mathbf{N}_5$, then show that they are closed under $\vee$ in $\mathrm{Con}\, \mathbf{N}_5$. It then follows by Lemma 6.17 that every congruence on $\mathbf{N}_5$ is principal.]

6.12 For each of the following lattices L, find all congruences and then draw $\mathrm{Con}\, L$:

$$(a)\ L = \mathbf{2} \times \mathbf{3}; \quad (b)\ L = \mathbf{2}^2 \oplus \mathbf{1}; \quad (c)\ L = \mathbf{4}.$$

6.13 Recall from 6.20 that a lattice L is called **simple** if it has precisely two congruences, namely $\mathbf{0}$ and $\mathbf{1}$.

(i) Show that L is simple if and only if for all $a, b \in L$ with $a \neq b$ we have $\theta(a,b) = \mathbf{1}$.

(ii) Show that each of the following lattices is simple:

(a) $\mathbf{M}_3$; (b) $\mathbf{M}_n$ (for $n \geqslant 3$); (c) $\mathbf{M}_{3,3}$.

See Figures 2.4 and 4.5.

6.14 Consider again the second half of the proof of the $\mathbf{M}_3$–$\mathbf{N}_5$ Theorem in 4.10. This actually shows that, if L is modular and $a, b, c \in L$, then $f: \mathbf{M}_3 \to L$ defined by

$$\begin{aligned} f(0) = p &:= (a \wedge b) \vee (b \wedge c) \vee (c \wedge a), \\ f(1) = q &:= (a \vee b) \wedge (b \vee c) \wedge (c \vee a), \\ f(x) = u &:= (a \wedge q) \vee p, \\ f(y) = v &:= (b \wedge q) \vee p, \\ f(z) = w &:= (c \wedge q) \vee p, \end{aligned}$$

is a homomorphism. Use the fact that $\mathbf{M}_3$ is simple (Exercise 6.13) to show that f is one-to-one, and therefore an embedding, provided $p \neq q$. (This gives a more sophisticated proof of the last line of the proof of the $\mathbf{M}_3$–$\mathbf{N}_5$ Theorem.)

6.15 (i) Let L, K_1 and K_2 be lattices and assume that the map $\varphi\colon L \to K_1 \times K_2$ is an embedding. Define congruences $\theta_1, \theta_2 \in \operatorname{Con} L$ by $\theta_i = \ker(\pi_i \circ \varphi)$ where $\pi_i : K_1 \times K_2 \to K_i$ is the natural projection $(i = 1, 2)$. Show that $\theta_1 \wedge \theta_2 = 0$ in $\operatorname{Con} L$.

(ii) Let $\theta_1, \theta_2 \in \operatorname{Con} L$ with $\theta_1 \wedge \theta_2 = \mathbf{0}$. Show that the map $\psi : L \to L/\theta_1 \times L/\theta_2$ defined by $\psi(a) := ([a]_{\theta_1}, [a]_{\theta_2})$ is an embedding.

(iii) Use (i) to show that each of the following lattices L has the property that if $L \rightarrowtail K_1 \times K_2$, then $L \rightarrowtail K_1$ or $L \rightarrowtail K_2$.

(*a*) $L = \mathbf{M}_3$; (*b*) $L = \mathbf{N}_5$; (*c*) $L = \mathbf{M}_{3,3}$ (see Figure 4.5).

[Hint. If you haven't already done so, you'll first need to calculate $\operatorname{Con} L$ in each case (Exercises 6.11 and 6.13). Then use Exercise 6.3.]

6.16 (i) Let P be an ordered set and let Q be any subset of P. Define $\theta_{[Q]}$ by

$$(A, B) \in \theta_{[Q]} \iff A \cap (P \setminus Q) = B \cap (P \setminus Q) \quad (A, B \in \mathcal{O}(P)).$$

Prove that $\theta_{[Q]}$ is a congruence on $\mathcal{O}(P)$.

(ii) Let $C, D \in \mathcal{O}(P)$ with $C \subseteq D$. Use Exercise 6.7 to prove that $\theta(C, D) = \theta_{[D \setminus C]}$, and hence prove that the atoms in $\operatorname{Con}(\mathcal{O}(P))$ are precisely the congruences $\theta_{[\{x\}]}$ $(x \in P)$.

6.17 Prove that if L is a finite distributive lattice then $\operatorname{Con} L$ is isomorphic to 2^n where $n = |\mathcal{J}(L)|$. [Hint. Here are three possible (but related) approaches.

(a) Prove the result indirectly via Theorem 5.19 and the fact that there is a bijection between congruences on L and the kernels of homomorphisms with domain L (see Lemma 6.8 and Theorem 6.9).

(b) Define $f\colon \wp(\mathcal{J}(L)) \to \operatorname{Con} L$ by $f(Y) = \theta_{[Y]}$ for each subset Y of $\mathcal{J}(L)$, where, for $a, b \in L$,

$$a \equiv b \,(\mathrm{mod}\, \theta_{[Y]}) \iff {\downarrow}a \cap (\mathcal{J}(L) \setminus Y) = {\downarrow}b \cap (\mathcal{J}(L) \setminus Y),$$

and show that f is an order-isomorphism.

(c) Combine 6.17 and the result in Exercise 6.16(ii).]

7

Complete Lattices and Galois Connections

We return now to the study of complete lattices. We introduced such lattices back in Chapter 2 where we saw many examples arising as topped $\bigcap$–structures. In Chapter 3 we gave a glimpse of the role of complete lattices in concept analysis. Here we pursue these ideas further by presenting the circle of ideas that link complete lattices and topped $\bigcap$–structures with closure operators and Galois connections. In 7.30 we summarize for reference the various correspondences.

En route, we pursue the theory of complete lattices in a different direction. Very many lattices arising in algebra are topped $\bigcap$–structures of a special type. We explore the features of these lattices and of the associated closure operators. The topped structures studied here have as topless counterparts the computer scientists' domains which we study in Chapter 9.

Finally, as an application of results here and in Chapter 3, we investigate the Dedekind–MacNeille completion of an ordered set.

Closure operators

Assume $(X, \mathcal{T})$ is a topological space. There is an intimate connection between the topped $\bigcap$–structure $\Gamma(X)$ consisting of the closed subsets of X and the closure operator $^- : \wp(X) \to \wp(X)$ which maps a subset A of X to its closure $\overline{A}$. Namely,

$$\Gamma(X) = \{\, A \subseteq X \mid \overline{A} = A \,\}$$

and

$$\overline{A} = \bigcap \{\, B \in \Gamma(X) \mid A \subseteq B \,\},$$

for all $A \subseteq X$. In fact, this connection has nothing to do with topology and exists for any topped $\bigcap$–structure.

To meet our needs when discussing Galois connections, we work initially in an arbitrary ordered set rather than in a powerset.

7.1 Closure operators. Let P be an ordered set. A map $c\colon P \to P$ is called a **closure operator** (on P) if, for all $x, y \in P$,

(clo1) $x \leqslant c(x)$,

(clo2) $x \leqslant y \Longrightarrow c(x) \leqslant c(y)$,

(clo3) $c(c(x)) = c(x)$.

An element $x \in P$ is called **closed** if $c(x) = x$. The set of all closed elements of P is denoted by P_c.

If $P = \langle \wp(X); \subseteq \rangle$ for some set X, we customarily use the symbol C rather than c and shall refer to a closure operator $C\colon \wp(X) \to \wp(X)$ **on** X. Of course, if $\langle X, T \rangle$ is a topological space, then the topological closure map $^-$ is a closure operator C on X, and $\wp(X)_C = \Gamma(X)$.

7.2 Proposition. *Let c be a closure operator on an ordered set P.*

(i) *$P_c = \{ c(x) \mid x \in P \}$ and P_c contains the top element of P when it exists.*

(ii) *Assume P is a complete lattice.*

(a) *For any $x \in P$,*

$$c(x) = \bigwedge\nolimits_P \{ y \in P_c \mid x \leqslant y \}.$$

(b) *P_c is a complete lattice, under the order inherited from P, such that, for every subset S of P_c,*

$$\bigwedge\nolimits_{P_c} S = \bigwedge\nolimits_P S \text{ and } \bigvee\nolimits_{P_c} S = c\big(\bigvee\nolimits_P S\big).$$

Proof. (i) Let $y \in P$. If $y \in P_c$, then $y = c(y)$. If $y = c(x)$ for some $x \in P$, then $c(y) = c(c(x)) = c(x) = y$, by (clo3), and so $y \in P_c$. If $\top$ exists in P, then $\top = c(\top)$ by (clo1).

We now prove (ii)(a). By (clo2), $c(x) \in \{ y \in P_c \mid x \leqslant y \}^\ell$. Since $c(x)$ belongs to $\{ y \in P_c \mid x \leqslant y \}$ by part (i) and (clo1), it is the *greatest* lower bound. Now consider (ii)(b). To show P_c is a complete lattice it suffices by 2.31 to show that $\bigwedge_{P_c} S$ exists for every $S \subseteq P_c$. By 2.28, this happens provided $\bigwedge_P S \in P_c$, and in that case $\bigwedge_P S$ serves as $\bigwedge_{P_c} S$. But $c\big(\bigwedge_P S\big) \leqslant c(s) = s$ for all $s \in S$, so $c\big(\bigwedge_P S\big) \leqslant \bigwedge_P S$; by (clo1) this inequality must be an equality. Finally, note that

$$\begin{aligned}
\bigvee\nolimits_{P_c} S &= \bigwedge\nolimits_{P_c} \{ y \in P_c \mid (\forall s \in S)\, s \leqslant y \} && \text{(by 2.30)} \\
&= \bigwedge\nolimits_P \{ y \in P_c \mid (\forall s \in S)\, s \leqslant y \} && \text{(from above)} \\
&= \bigwedge\nolimits_P \{ y \in P_c \mid \textstyle\bigvee_P S \leqslant y \} \\
&= c\big(\bigvee\nolimits_P S\big) && \text{(by (ii)(a)).} \qquad \square
\end{aligned}$$

The next result says that every topped $\bigcap$-structure gives rise to a closure operator and conversely. The first part is a specialization of Proposition 7.2(ii)(b). The proofs of the remaining assertions are left as exercises.

7.3 Theorem. *Let C be a closure operator on a set X. Then the family*

$$\mathfrak{L}_C := \{ A \subseteq X \mid C(A) = A \}$$

of closed subsets of X is a topped $\bigcap$-structure and so forms a complete lattice, when ordered by inclusion, in which

$$\bigwedge_{i \in I} A_i = \bigcap_{i \in I} A_i,$$
$$\bigvee_{i \in I} A_i = C\Big(\bigcup_{i \in I} A_i\Big).$$

Conversely, given a topped $\bigcap$-structure $\mathfrak{L}$ on X, the formula

$$C_{\mathfrak{L}}(A) := \bigcap \{ B \in \mathfrak{L} \mid A \subseteq B \}.$$

defines a closure operator $C_{\mathfrak{L}}$ on X.

7.4 Remarks. The relationship between closure operators and topped $\bigcap$-structures on a given set is a bijective one: the closure operator induced by the topped $\bigcap$-structure $\mathfrak{L}_C$ is C itself, and, similarly, the topped $\bigcap$-structure induced by the closure operator $C_{\mathfrak{L}}$ is $\mathfrak{L}$; in symbols,

$$C_{(\mathfrak{L}_C)} = C \quad \text{and} \quad \mathfrak{L}_{(C_{\mathfrak{L}})} = \mathfrak{L}.$$

Thus, whether we work with a topped $\bigcap$-structure or the corresponding closure operator is a matter of convenience.

It is worth recalling that every complete lattice arises (up to order-isomorphism) as a topped $\bigcap$-structure on some set (see Exercise 2.29), so that equivalently, every complete lattice is isomorphic to the lattice of closed sets with respect to some closure operator.

We have already observed that the topological closure map is a closure operator on X, for any topological space $(X, \mathcal{T})$. It is also natural to consider the map sending a subset A of X to its interior, A°. This map is not a closure operator because $A^\circ \subseteq A$ rather than $A \subseteq A^\circ$. However by reversing the order on $\wp(X)$ (that is, by considering $\wp(X)^\partial$) we do get a closure operator. This device allows topological interior operators and their abstract counterparts to be brought within the scope of our theory. An order-reversal of this sort appears in 7.27.

It is easy to identify the closure operators associated with the topped $\bigcap$-structures introduced earlier. They are all very natural.

7.5 Examples.

(1) Let G be a group. Then the closure operator corresponding to the topped $\bigcap$-structure $\operatorname{Sub} G$ maps a subset A of G to the subgroup $\langle A \rangle$ generated by A.

(2) Let V be a vector space over a field F and let $\mathrm{Sub}\,V$ be the complete lattice of linear subspaces of V. The corresponding closure operator on V maps a subset A of V to its linear span.

(3) Let L be a lattice and for each $X \subseteq L$ let

$$[X] := \bigcap \{ K \in \mathrm{Sub}_0 L \mid X \subseteq K \}.$$

Then $[-]\colon \wp(L) \to \wp(L)$ is the closure operator corresponding to the topped $\bigcap$-structure $\mathrm{Sub}_0 L$. (Recall Exercise 2.15.)

(4) Let L be a lattice with 0. Then the closure operator corresponding to the topped $\bigcap$-structure $\mathcal{I}(L)$ consisting of all ideals of L is $(-]\colon \wp(L) \to \wp(L)$, as defined in Exercise 2.22.

(5) Let P be an ordered set. The map $\downarrow\colon \wp(P) \to \wp(P)$ introduced in 1.27 is easily seen to be a closure operator. The corresponding topped $\bigcap$-structure is the down-set lattice $\mathcal{O}(P)$.

Complete lattices coming from algebra: algebraic lattices

Any topped $\bigcap$-structure is a complete lattice in which meet is given by set intersection, but, regrettably, join is usually not given by union. We now explore the circumstances under which joins *are* given by union, and discover that this occurs in a natural way for certain joins in many of the lattices arising in algebra.

7.6 Example. Let G be a group and $\mathcal{H} := \{H_i\}_{i \in I}$ be a non-empty family of subgroups of G with the property that, for each $i_1, i_2 \in I$, there exists $k \in I$ such that $H_{i_1} \cup H_{i_2} \subseteq H_k$. We claim that $H := \bigcup_{i \in I} H_i$ is a subgroup. Choose $g_1, g_2 \in H$. It suffices to show that $g_1 g_2^{-1} \in H$. For $j = 1, 2$, there exists $i_j \in I$ such that $g_j \in H_{i_j}$. By hypothesis we can find $H_k \in \mathcal{H}$ so that $H_{i_1} \subseteq H_k$ and $H_{i_2} \subseteq H_k$. Then g_1, g_2 both belong to a *common* subgroup H_k, so $g_1 g_2^{-1} \in H_k$. Hence $g_1 g_2^{-1} \in H$, as required.

As a special case, note that if $H_1 \subseteq H_2 \subseteq \cdots$ is a non-empty chain of subgroups, then $\bigcup_{n \geqslant 1} H_n$ is a subgroup.

It should be clear that what is crucial to the argument above is not that it concerns groups, but the existence, for a given pair H_1, H_2 of members of $\mathcal{H}$, of a member H of $\mathcal{H}$ which contains both H_1 and H_2, so that we can exploit the closure properties of the group operations in H. This leads us to the next definition.

7.7 Definition. Let S be a non-empty subset of an ordered set P. Then S is said to be **directed** if, for every pair of elements $x, y \in S$, there exists $z \in S$ such that $z \in \{x, y\}^u$. An easy induction shows that S is directed if and only if, for every finite subset F of S, there exists $z \in S$ such that $z \in F^u$.

When D is a directed set for which $\bigvee D$ exists then we often write $\bigsqcup D$ in place of $\bigvee D$ as a reminder that D is directed. Directed joins arise very naturally in the context of computer science. Example 7.8(5) hints at the way directed sets may arise in approximation. This theme is picked up in Chapter 8, which is concerned with CPOs: a **CPO** is an ordered set P with $\bot$ in which $\bigsqcup D$ exists for every directed subset D of P.

7.8 Examples.

(1) In any ordered set P, any non-empty chain is directed and any subset of P with a greatest element is directed.

(2) The only directed subsets of an antichain are the singletons. More generally, in an ordered set with (ACC) (recall 2.37) a set is directed if and only if it has a greatest element.

(3) Let X be a set. Then any non-empty family $\mathcal{D}$ of subsets of X which is closed under finite unions is directed: for $A, B \in \mathcal{D}$, we have $A \cup B \in \{A, B\}^u$ in $\mathcal{D}$. Hence, for example, the family of finite subsets of $\mathbb{N}$ is directed.

(4) The finitely generated subgroups of a group G form a directed subset $\mathcal{L}$ of $\mathrm{Sub}\,G$. To check this claim, let H and K be subgroups of G generated, respectively, by $\{a_1, \ldots, a_m\}$ and $\{b_1, \ldots, b_n\}$. Let M be the subgroup generated by $\{a_1, \ldots, a_m, b_1, \ldots, b_n\}$. Then $M \in \{H, K\}^u$ in $\mathcal{L}$. Notice that, by contrast with the preceding example, the exhibited upper bound is *not* given by set union: in general $H \cup K$ is, of course, *not* a subgroup.

(5) The graph of a map $f\colon \mathbb{N} \to \mathbb{N}$ is the union of the set of all graphs of partial maps $\sigma \leqslant f$ with $\mathrm{dom}\,\sigma$ finite (more informally, f can be built up from partial maps each specifying a finite amount of information about f). This family of approximating partial maps σ is a directed subset of $(\mathbb{N} \multimap \mathbb{N})$.

7.9 Directed families of sets. We can restate the result in 7.6 as asserting that the union of a directed family of subgroups of a group is again a subgroup. In an exactly analogous way, the union of a directed family of subspaces of a vector space is a subspace, the union of a directed family of ideals in a lattice is an ideal (Exercise 2.37(i) is a special case), and so on. The union of a directed family of sets will be called a **directed union**.

Now recall 2.29: if $\{A_i\}_{i\in I}$ is a subset of a family $\mathfrak{L}$ of subsets of a set X, then

$$\bigcup_{i\in I} A_i \in \mathfrak{L} \implies \bigvee_{\mathfrak{L}} \{A_i \mid i \in I\} \text{ exists and equals } \bigcup_{i\in I} A_i.$$

We deduce that if the family $\mathfrak{L}$ is closed under directed unions, we have $\bigcup_{i\in I} A_i = \bigvee_{i\in I} A_i = \bigcup_{i\in I} A_i$ whenever $\{A_i\}_{i\in I} \subseteq \mathfrak{L}$ is directed.

The following simple observation is very useful. A subset $\mathcal{D} = \{A_i\}_{i\in I}$ of $\mathcal{P}(X)$ is directed if and only if, given $A_{i_1}, \ldots, A_{i_n}$ in $\mathcal{D}$, there exists $k \in I$ such that $A_{i_j} \subseteq A_k$ for $i = 1, \ldots, n$ (equivalently, $\bigcup\{A_{i_j} \mid j = 1, \ldots, n\} \subseteq A_k$). It follows that if $\mathcal{D}$ is directed and $Y = \{y_1, \ldots, y_n\}$ is a finite subset of $\bigcup A_i$ then there exists $A_k \in \mathcal{D}$ such that $Y \subseteq A_k$.

7.10 Definitions. A non-empty family $\mathfrak{L}$ in $\mathcal{P}(X)$ is said to be **closed under directed unions** if $\bigcup_{i\in I} A_i \in \mathfrak{L}$ for any directed family $\mathcal{D} = \{A_i\}_{i\in I}$ in $\mathfrak{L}$.

A non-empty family $\mathfrak{L}$ of subsets of a set X is said to be an **algebraic $\bigcap$-structure** if

(i) $\bigcap_{i\in I} A_i \in \mathfrak{L}$ for any non-empty family $\{A_i\}_{i\in I}$ in $\mathfrak{L}$,

(ii) $\bigcup_{i\in I} A_i \in \mathfrak{L}$ for any directed family $\{A_i\}_{i\in I}$ in $\mathfrak{L}$.

Thus an algebraic $\bigcap$-structure is an $\bigcap$-structure which is closed under directed unions. In such a structure the join of any directed family is given by set union.

7.11 Examples. We see from 7.6 that the $\bigcap$-structure $\operatorname{Sub} G$ is algebraic. Similarly, each of the $\bigcap$-structures presented in 7.5 can be shown to be algebraic. We can also add to the list given there the congruence lattice $\operatorname{Con} L$, for any lattice L.

Theorem 7.3 set up a correspondence between topped $\bigcap$-structures and closure operators on a set. This specializes in a very satisfactory way to the algebraic case.

7.12 Definition. A closure operator C on a set X is called **algebraic** if, for all $A \subseteq X$,

$$C(A) = \bigcup \{C(B) \mid B \subseteq A \text{ and } B \text{ is finite}\}.$$

It is easy to show that, for any closure operator C,

$$C(A) \supseteq \bigcup \{C(B) \mid B \subseteq A \text{ and } B \text{ is finite}\}$$

(Exercise 7.1), so that to prove that a closure operator C is algebraic it is only necessary to prove the reverse inclusion.

7.13 Example. Recall from 7.5(1) that the closure operator corresponding to the $\bigcap$-structure $\operatorname{Sub} G$ maps a subset A of G to the subgroup $\langle A\rangle$ generated by A. We claim that this closure operator is algebraic. This follows, via 7.14, from the fact that $\operatorname{Sub} G$ is an algebraic $\bigcap$-structure, but the direct proof below is also instructive. By the remark in 7.12, it is sufficient to show that

$$\langle A\rangle \subseteq \bigcup \{\langle B\rangle \mid B \subseteq A \text{ and } B \text{ is finite}\}.$$

Let $g \in \langle A\rangle$; then there exist $a_1, a_2, \ldots, a_n \in A$ such that $g = a_1' a_2' \ldots a_n'$, where $a_i' \in \{a_i, a_i^{-1}\}$ for each i. Thus $g \in \langle\{a_1, \ldots, a_n\}\rangle$, and this gives the required containment.

7.14 Theorem. *Let C be a closure operator on a set X and let $\mathfrak{L}_C$ be the associated topped $\bigcap$-structure. Then the following are equivalent:*

(i) *C is an algebraic closure operator;*

(ii) *for every directed family $\{A_i\}_{i\in I}$ of subsets of X,*

$$C\Big(\bigcup_{i\in I} A_i\Big) = \bigcup_{i\in I} C(A_i);$$

(iii) *$\mathfrak{L}_C$ is an algebraic $\bigcap$-structure.*

Proof. Assume (i) holds and let $\{A_i\}_{i\in I}$ be a directed family of subsets of X. First observe that if B is finite and $B \subseteq \bigcup_{i\in I} A_i$, then $B \subseteq A_k$ for some $k \in I$ (see 7.9). Consequently,

$$\begin{aligned} C\Big(\bigcup_{i\in I} A_i\Big) &= \bigcup\{C(B) \mid B \subseteq \textstyle\bigcup_{i\in I} A_i \text{ and } B \text{ is finite}\} \qquad \text{(by (i))} \\ &= \bigcup\{C(B) \mid B \subseteq A_k \text{ for some } k \in I \text{ and } B \text{ is finite}\} \\ &\qquad\qquad \text{(by the observation above)} \\ &\subseteq \bigcup_{i\in I} C(A_i). \end{aligned}$$

The reverse inclusion is always valid (Exercise 7.1). Hence (i) $\Rightarrow$ (ii).

Since $\mathfrak{L}_C = \{C(A) \mid A \subseteq X\}$, it is trivial that (ii) implies (iii). Now assume (iii). Let $A \subseteq X$. The family

$$\mathcal{D} := \{C(B) \mid B \subseteq A \text{ and } B \text{ is finite}\}$$

is directed. Hence $\bigcup \mathcal{D} \in \mathfrak{L}_C$. Also $A \subseteq \bigcup \mathcal{D}$ since, for each $x \in A$, we have $x \in \{x\} \subseteq C(\{x\}) \subseteq \bigcup \mathcal{D}$. Hence

$$\begin{aligned} C(A) &\subseteq C\big(\bigcup \mathcal{D}\big) \\ &= \bigcup \mathcal{D} \qquad\qquad (\text{since } \bigcup \mathcal{D} \in \mathfrak{L}_C) \\ &= \bigcup \big\{ C(B) \mid B \subseteq A \text{ and } B \text{ is finite} \big\}. \end{aligned}$$

As previously noted, the reverse inclusion always holds. It follows that C is algebraic. □

Our next objective is to characterize in a lattice-theoretic way the closures, with respect to an algebraic closure operator on X, of the finite subsets of X. Before we can do this we need further definitions.

7.15 Definitions. Let L be a complete lattice and let $k \in L$.

(i) k is called **finite** (in L) if, for every directed set D in L,

$$k \leqslant \bigsqcup D \Longrightarrow k \leqslant d \text{ for some } d \in D.$$

The set of finite elements of L is denoted $F(L)$.

(ii) k is said to be **compact** if, for every subset S of L,

$$k \leqslant \bigvee S \Longrightarrow k \leqslant \bigvee T \text{ for some finite subset } T \text{ of } S.$$

The set of compact elements of L is denoted $K(L)$.

Lemma 7.16 reconciles these definitions. The term 'compact' (cribbed from topology) is commonly used in the context of algebra; the notion of a 'finite' element comes from computer science. Observe that, unlike compactness, finiteness makes sense in ordered sets in which joins exist for directed subsets, but not necessarily for all subsets.

7.16 Lemma. *Let L be a complete lattice. Then $F(L) = K(L)$. Further, $k_1 \vee k_2 \in F(L)$ whenever $k_1, k_2 \in F(L)$.*

Proof. Assume first that $k \in K(L)$ and that $k \leqslant \bigsqcup D$, where D is directed. Then there exists a finite subset F of D such that $k \leqslant \bigvee F$. Because D is directed, we can find $d \in D$ with $d \in F^{\mathrm{u}}$. Then $k \leqslant d$, so $k \in F(L)$.

Conversely, assume that $k \in F(L)$ and that $k \leqslant \bigvee S$. The set

$$D = \big\{ \bigvee T \mid T \subseteq S \text{ and } T \text{ is finite} \big\}$$

is directed and it is easy to see that $\bigsqcup D = \bigvee S$ (see Exercise 7.5). Applying the finiteness condition, we find a finite subset T of S with $k \leqslant \bigvee T$.

The second part is left as an exercise. □

7.17 Examples. Table 7.1 lists the finite (alias compact) elements in various complete lattices. The assertions can easily be verified directly. The first four lattices are topped $\bigcap$-structures and for these Lemma 7.19 below can alternatively be used.

Note that $\bot$ in a complete lattice is always finite. As a simple example of a non-finite element we have the top element of $\mathbb{N} \oplus \mathbf{1}$. Now consider $\langle \mathbb{N}_0; \preccurlyeq \rangle$. We claim that no element other than 1 $(= \bot)$ is compact. Since 0 $(= \top)$ is the join of the set of all primes but is not the join of any finite set of primes, 0 is not compact. Now let $n \in \mathbb{N}_0$ with $n \notin \{0, 1\}$ and let S be the set of primes which do not divide n. Then S is infinite and so $\bigvee S = \top$ $(= 0)$, because any non-zero element of $\mathbb{N}_0$ has only finitely many prime divisors. Hence $n \preccurlyeq \bigvee S$ but $n \not\preccurlyeq \bigvee T$ for any finite subset T of S, whence n is not compact.

L	$F(L)$
$\wp(X)$ $(X$ a set$)$	Finite subsets
$\mathcal{O}(P)$ $(P$ an ordered set$)$	Sets $\downarrow F$ $(F$ finite$)$
$\mathrm{Sub}\, G$ $(G$ a group$)$	Finitely generated subgroups
$\mathrm{Sub}\, V$ $(V$ a vector space$)$	Finite-dimensional subspaces
Complete lattice with (ACC)	All elements of L
$[0, 1]$	0 only

Table 7.1

We now work towards a characterization of the ordered sets which can be concretely represented as topped algebraic $\bigcap$-structures.

7.18 Definition. A complete lattice L is said to be **algebraic** if, for each $a \in L$,

$$a = \bigvee \big\{ k \in K(L) \mid k \leqslant a \big\}.$$

The terminology will be justified by the next lemma which implies that many lattices arising in algebra are algebraic; see 7.21. In the following proofs, finiteness is more convenient to work with than compactness. Lemma 7.16 then allows results about algebraic lattices to be stated, as is traditional, in terms of compact elements.

7.19 Lemma. *Let C be an algebraic closure operator on X and $\mathfrak{L}_C$ the associated topped algebraic $\bigcap$-structure. Then $\mathfrak{L}_C$ is an algebraic lattice in which an element A is finite (equivalently, compact) if and only if $A = C(Y)$ for some finite set $Y \subseteq X$.*

Proof. We show that the finite elements are the closures of the finite sets. Then Definition 7.12 implies that $\mathfrak{L}_C$ is an algebraic lattice.

Let Y be a finite subset of X and let $A = C(Y)$. Take a directed set $\mathcal{D}$ in $\mathfrak{L}_C$ with $A \subseteq \bigsqcup \mathcal{D}$. Then, since $\bigsqcup$ coincides with $\bigcup$ in $\mathfrak{L}_C$,

$$Y \subseteq C(Y) = A \subseteq \bigsqcup \mathcal{D} = \bigcup \mathcal{D}.$$

As Y is finite and $\mathcal{D}$ directed, there exists $B \in \mathcal{D}$ such that $Y \subseteq B$. Then

$$A = C(Y) \subseteq C(B) = B,$$

so A is finite in $\mathfrak{L}_C$.

Conversely, assume that $A \in \mathfrak{L}_C$ is a finite element. Certainly

$$A = \bigsqcup \{ C(Y) \mid Y \subseteq A \text{ and } Y \text{ is finite} \}$$

(see 7.12). Invoke the finiteness of A in $\mathfrak{L}_C$ to find a finite set $Y \subseteq A$ such that $A \subseteq C(Y)$. The reverse inclusion holds since $Y \subseteq A$ implies $C(Y) \subseteq C(A) = A$. □

7.20 Theorem.

(i) *Let $\mathfrak{L}$ be a topped algebraic $\bigcap$–structure. Then $\mathfrak{L}$ is an algebraic lattice.*

(ii) *Let L be an algebraic lattice and define $D_a := \{ k \in K(L) \mid k \leqslant a \}$ for each $a \in L$. Then $\mathfrak{L} := \{ D_a \mid a \in L \}$ is a topped algebraic $\bigcap$–structure isomorphic to L.*

Proof. Part (i) follows from the preceding lemma and 7.14. For the converse, (ii), we leave the proof that $\mathfrak{L}$ is a topped $\bigcap$–structure as an exercise and prove here that the map $\varphi\colon a \mapsto D_a$ is an isomorphism of L onto $\mathfrak{L}$ and that $\mathfrak{L}$ is algebraic. Because L is algebraic, $D_a \subseteq D_b$ in $\mathfrak{L}$ implies $a = \bigvee D_a \leqslant \bigvee D_b = b$ in L. The reverse implication holds always. Therefore φ is an order-isomorphism.

Take a directed subset $\mathcal{D} = \{ D_c \mid c \in C \}$ of $\mathfrak{L}$. As φ is an order-isomorphism, the indexing set C is a directed subset of L. Define $a = \bigsqcup C$. We claim that $\bigcup \mathcal{D} = D_a$ and so belongs to $\mathfrak{L}$. Indeed,

$$\begin{aligned} k \in D_a &\iff k \in K(L) = F(L) \text{ and } k \leqslant a = \textstyle\bigsqcup C \\ &\iff k \in F(L) \text{ and } k \leqslant c \text{ for some } c \in C \\ &\iff k \in D_c \text{ for some } c \in C \\ &\iff k \in \textstyle\bigcup \mathcal{D}. \end{aligned}$$

Hence $\mathfrak{L}$ is closed under directed unions and so is algebraic. □

7.21 Examples. The following are topped $\bigcap$–structures arising from algebraic closure operators and so are algebraic lattices (see 7.13 for a typical proof):

- $\wp(X)$, for any set X;
- any complete lattice of sets and, in particular, the down-set lattice $\mathcal{O}(P)$, for any ordered set P (in Theorem 10.29 below we characterize down-set lattices amongst algebraic lattices);
- $\operatorname{Sub} G$, for any group G (see also 7.8(4));
- $\operatorname{Sub} V$, for any vector space V;
- $\mathcal{I}(L)$, the ideal lattice of any lattice L with 0;
- $\operatorname{Con} L$, for any lattice L.

In addition, the chains $\mathbf{n}$, for $n \geqslant 1$, and $\mathbb{N} \oplus \mathbf{1}$ are algebraic lattices. Further, any lattice L with a bottom element and satisfying (ACC) is an algebraic lattice: by 2.41 L is a complete lattice and, as noted in 7.17, every element $x \in L$ is compact, and so is the join of $\downarrow x \cap K(L)$. As an example of an infinite algebraic lattice of this type we have $\langle \mathbb{N}_0; \preccurlyeq \rangle^\partial$. On the other hand, $\langle \mathbb{N}_0; \preccurlyeq \rangle$ is *not* algebraic, since its only compact element is $\perp$ (that is, 1 – see 7.17).

Galois connections

We now introduce a class of examples which will later prove to be very significant and which we shall relate to closure operators.

7.22 Contexts and their polar maps. Let (G, M, I) be a context and consider the polar maps ${}^{\triangleright}\colon \wp(G) \to \wp(M)^\partial$ and ${}^{\triangleleft}\colon \wp(M)^\partial \to \wp(G)$ given by

$$\begin{aligned} A^{\triangleright} &:= A' = \{ m \in M \mid (\forall g \in A)\, gIm \}, \\ B^{\triangleleft} &:= B' = \{ g \in G \mid (\forall m \in B)\, gIm \}; \end{aligned}$$

(see 3.3). These polar maps satisfy, for $A \subseteq G$ and $B \subseteq M$,

$$A^{\triangleright} \leqslant B \iff A \leqslant B^{\triangleleft}$$

(see (P5) in Lemma 3.5); here $\leqslant$ on the left-hand side is the order $\subseteq$ of $\wp(G)$ and $\leqslant$ on the right-hand side is the order $\supseteq$ of $\wp(M)^\partial$.

We are now ready to reveal that the pair of maps $(', ')$ provides one example of a notion of widespread occurrence and considerable importance: a Galois connection between two ordered sets.

7.23 Galois connections. Let P and Q be ordered sets. A pair $({}^{\triangleright}, {}^{\triangleleft})$ of maps ${}^{\triangleright}\colon P \to Q$ and ${}^{\triangleleft}\colon Q \to P$ (called **right** and **left** respectively) is a **Galois connection** between P and Q if, for all $p \in P$ and $q \in Q$,

(Gal) $p^{\triangleright} \leqslant q \iff p \leqslant q^{\triangleleft}$.

The map $^{\triangleright}$ is called the **lower adjoint of** $^{\triangleleft}$ and the map $^{\triangleleft}$ the **upper adjoint of** $^{\triangleright}$; the terms 'lower' and 'upper' here refer to the side of $\leqslant$ on which the map appears.

Galois connections manifest themselves in many settings, from algebra to computer science. It is common for ubiquitous concepts to adopt different guises in different contexts. Galois connections illustrate this all too well. There are two versions of the definitions: the one we adopt here, in which the paired maps are order-preserving (see 7.26) and the other in which they are order-reversing, as would occur in our context example if we worked with $\wp(M)$ rather than $\wp(M)^\partial$. Historically, and in algebra, there are arguments for order-reversal: the most famous Galois connection of all, that discovered by Galois, between the subgroups of a Galois group $G(K,F)$ and the fields intermediate between F and K, is order-reversing. In computer science the maps are usually taken to be order-preserving. At a theoretical level, the difference is not significant: we can swap backwards and forwards between the two versions by swapping between Q and Q^∂. There is also, regrettably, no uniformly adopted notation. We have used the symbols $^{\triangleright}$ and $^{\triangleleft}$ as they make it easy to keep track of which map is which.

7.24 Order-theoretic examples.

(1) Suppose that sets P and Q are ordered by the discrete order, $=$. Then $^{\triangleright}: P \to Q$ and $^{\triangleleft}: Q \to P$ set up a Galois connection between P and Q if and only if these maps are set-theoretic inverses of each other.

(2) Consider again 7.22. The $'$ maps associated with a context (G, M, I) give maps $^{\triangleright}: \wp(G) \to \wp(M)^\partial$ and $^{\triangleleft}: \wp(M)^\partial \to \wp(G)$ such that $(^{\triangleright}, ^{\triangleleft})$ is a Galois connection between $\wp(G)$ and $\wp(M)^\partial$.

(See Exercise 7.18 for an alternative way to construct a Galois connection out of a relation $R \subseteq A \times B$ between sets A and B.)

(3) Let P be an ordered set. For $A \subseteq P$ we have previously defined the sets of upper and lower bounds of A as

$$A^u := \{y \in P \mid (\forall x \in A)\, x \leqslant y\},$$
$$A^\ell := \{y \in P \mid (\forall x \in A)\, y \leqslant x\}.$$

It is easy to see that $(^u, ^\ell)$ is a Galois connection between $\wp(P)$ and $\wp(P)^\partial$:

$$\begin{aligned} A^u \supseteq B &\iff (\forall y \in B)((\forall x \in A)\, x \leqslant y) \\ &\iff (\forall x \in A)((\forall y \in B)\, y \geqslant x) \\ &\iff A \subseteq B^\ell. \end{aligned}$$

In fact this is a special case of (2) in which $G = M = P$ and I is the relation $\leqslant$ (regarded as a subset of $P \times P$). We shall return to this important example in the next section, where we discuss the Dedekind–MacNeille completion.

(4) Let P be an ordered set. For $A \subseteq P$, define

$$A^{\triangleright} := P \setminus {\downarrow}A \quad \text{and} \quad A^{\triangleleft} := P \setminus {\uparrow}A.$$

Then $(^{\triangleright}, ^{\triangleleft})$ establishes a Galois connection between $\wp(P)$ and $\wp(P)^\partial$. Again this is a specialization of (2).

7.25 A context for programs. We now give a computational interpretation to Example 7.24(2).

Following up our introductory remarks in 1.8, we shall view a computer program as operating on a (finite or infinite) state space X, whose elements are vectors of which each component is of an appropriate datatype, these components being the values which the program's variables can take. It will be helpful to distinguish between a set G of initial states and a set M of final states.

Consider a deterministic program P which terminates when started from a state in a subset S of the set G of initial states. Then P may be viewed as a partial function σ_P from G to the set M of final states, with $\operatorname{dom}\sigma_P = S$: given initial state $x \in \operatorname{dom}\sigma_P$ the program terminates in the final state $\sigma_P(x) \in M$. A non-deterministic (and possibly non-terminating) program is modelled not by a partial function but by a relation $R \subseteq G \times M$. To illustrate, we consider $G = \{A, B, C, D, E\}$, $M = \{a, b, c, d, e\}$ and suppose that possible behaviours of the program are given by the relation R shown in Table 7.2. Thus the program, when started from initial state x, must terminate in a final state y for which xRy. So, when started from state B, the final state is one of c, d or e, and if the program is to terminate in state c, then it must begin from either A or B. The model allows for the possibility that the program fails to terminate, by having an initial state (that is, E) with no associated final states.

	a	b	c	d	e
A		×	×	×	×
B			×	×	×
C	×	×		×	×
D	×	×			
E					

Table 7.2

We now look at a different computational model. Consider again a set G of initial states and set M of final states. We may, as in 1.6, think of the predicates $\mathbb{P}(G)$ on G and $\mathbb{P}(M)$ on M as being identified with $\wp(G)$ and $\wp(M)$, respectively. As usual, powersets are ordered by $\subseteq$ and predicates by $\Rrightarrow$. We think of a subset Y of M as specifying a **postcondition**: a property of final states true precisely in Y. Similarly, a predicate on initial states is a **precondition**. We may stipulate what a (non-deterministic) program should do by giving conditions on the state of the system before and after the program's execution. For example, with $G = \mathbb{N}$ and $M = \mathbb{R}$, to find a square root of a natural number x in the range 1 to 10, we could consider a variable y, take the precondition '$1 \leqslant x \leqslant 10$' and the postcondition '$y^2 = x$', with only the value of y being changed in the transition of the system from its initial state to its final state. Our aim would still be met if we required '$y > 0$ & $y^2 = x$' (a more restrictive postcondition than before) and '$1 \leqslant x \leqslant 100$' (a less restrictive precondition). In general a program will still do what is demanded of it if a postcondition is strengthened or a precondition is weakened. Such changes result in a refinement of the given program (recall the notion of refinement described in 1.8). Now fix G and M, and consider a relation $R \in \mathcal{R} := \wp(G \times M)$ modelling a program P, possibly non-deterministic but which we assume, for simplicity, to be terminating (the case of non-termination requires more care and involves the liftings of G and M and an appropriate subset of $\wp(G_\perp \times M_\perp)$). For a given postcondition Y, the **weakest precondition** $wp_R(Y)$, is the set of input states x such that P is guaranteed to terminate in a state in Y when it is started from x. We may regard wp_R as a map from $\mathbb{P}(M)$ to $\mathbb{P}(G)$; it preserves $\Rrightarrow$. Such maps are called **predicate transformers**. Given a predicate transformer $T \in \mathcal{T} := (\mathbb{P}(M) \to \mathbb{P}(G))$ we may associate a relation R_T by

$$xR_Ty \iff (\forall Y \in \wp(M))\, x \in T(Y) \implies y \in Y \quad (\text{for } x \in G,\, y \in M).$$

It can be verified that, for $R \in \mathcal{R}$ and $T \in \mathcal{T}$,

$$R \subseteq R_T \iff wp_R \Lleftarrow T.$$

Thus there is a Galois connection between $\langle \mathcal{R}; \subseteq \rangle$ and $\langle \mathcal{T}; \Rrightarrow \rangle^\partial$. Order reversal here is essential; known preservation properties of $wp\colon \mathcal{R} \to \mathcal{T}$ require that wp should be the *upper* adjoint (see 7.31). Taking these ideas one step further, the program itself may be modelled by a predicate transformer. The Galois connection above allows movement backwards and forwards between relational and predicate transformer semantics.

The preceding discussion has been used to introduce, within the framework of contexts, some terminology which is well established in computer science and which we need later. Meanwhile, we return to theory. The following lemma establishes the basic calculational properties

of Galois connections. Compare with Lemma 3.5! Note that $({}^\triangleright, {}^\triangleleft)$ is a Galois connection between P and Q if and only if $({}^\triangleleft, {}^\triangleright)$ is a Galois connection between Q^∂ and P^∂. Consequently we have a 'buy one, get one free' situation, and to prove (Gal1)–(Gal3) below we only need to verify the first from each pair of mutually dual assertions. Sometimes, (Gal1) and (Gal3) are referred to as the **Cancellation Rule** and the **Semi-inverse Rule**.

7.26 Lemma. *Assume $({}^\triangleright, {}^\triangleleft)$ is a Galois connection between ordered sets P and Q. Let $p, p_1, p_2 \in P$ and $q, q_1, q_2 \in Q$. Then*

(Gal1) $p \leqslant p^{\triangleright\triangleleft}$ *and* $q^{\triangleleft\triangleright} \leqslant q$,

(Gal2) $p_1 \leqslant p_2 \implies p_1^\triangleright \leqslant p_2^\triangleright$ *and* $q_1 \leqslant q_2 \implies q_1^\triangleleft \leqslant q_2^\triangleleft$,

(Gal3) $p^\triangleright = p^{\triangleright\triangleleft\triangleright}$ *and* $q^\triangleleft = q^{\triangleleft\triangleright\triangleleft}$.

Conversely, a pair of maps ${}^\triangleright : P \to Q$ and ${}^\triangleleft : Q \to P$ satisfying (Gal1) *and* (Gal2) *for all $p, p_1, p_2 \in P$ and for all $q, q_1, q_2 \in Q$ sets up a Galois connection between P and Q.*

Proof. For $p \in P$, we have $p^\triangleright \leqslant p^\triangleright$ from which we obtain $p \leqslant p^{\triangleright\triangleleft}$ by putting $q = p^\triangleright$ in (Gal). Hence (Gal) implies (Gal1).

Consider (Gal2). We have

$$\begin{aligned} p_1 \leqslant p_2 &\implies p_1 \leqslant p_2^{\triangleright\triangleleft} && \text{(by (Gal1) and transitivity)} \\ &\iff p_1^\triangleright \leqslant p_2^\triangleright && \text{(from (Gal)).} \end{aligned}$$

We now prove (Gal3). Applying ${}^\triangleright$ to the inequality $p \leqslant p^{\triangleright\triangleleft}$ in (Gal1) we have, by (Gal2), $p^\triangleright \leqslant p^{\triangleright\triangleleft\triangleright}$. Also, by (Gal) with $p^{\triangleright\triangleleft}$ in place of p and $p^\triangleright$ in place of q,

$$p^{\triangleright\triangleleft} \leqslant p^{\triangleright\triangleleft} \implies p^{\triangleright\triangleleft\triangleright} \leqslant p^\triangleright.$$

Lastly, assume that (Gal1) and (Gal2) hold universally. Let $p^\triangleright \leqslant q$. By (Gal2), $p^{\triangleright\triangleleft} \leqslant q^\triangleleft$. Also, (Gal1) gives $p \leqslant p^{\triangleright\triangleleft}$. Hence $p \leqslant q^\triangleleft$ by transitivity. The reverse implication follows in the same way. □

The exercises give a sample of further triangle-juggling games that can be played, yielding in particular alternative characterizations of Galois connections.

We turn now to the relationship between Galois connections and the other structures we have been studying. We first reveal a connection between Galois connections and closure operators. This is important because it shows that Galois connections give rise to complete lattices.

7.27 From a Galois connection to a closure operator. Let $({}^{\triangleright}, {}^{\triangleleft})$ be a Galois connection between ordered sets P and Q^{∂} (note that we have Q^{∂} not Q here). Then

(i) $c := {}^{\triangleright\triangleleft} : P \to P$ and $k := {}^{\triangleleft\triangleright} : Q \to Q$ are closure operators. (Because we write the triangles to the right of their arguments, the left-hand map in each composition is performed first.)

(ii) Let

$$P_c := \{ p \in P \mid p^{\triangleright\triangleleft} = p \} \quad \text{and} \quad Q_k := \{ q \in Q \mid q^{\triangleleft\triangleright} = q \}.$$

Then ${}^{\triangleright} : P_c \to Q_k^{\partial}$ and ${}^{\triangleleft} : Q_k^{\partial} \to P_c$ are mutually inverse order-isomorphisms.

We leave the routine verification of (i) as an exercise in using the properties in 7.26. Now consider (ii). Using (Gal3) we see that ${}^{\triangleright}$ maps P_c onto Q_k^{∂} and that ${}^{\triangleleft}$ maps Q_k^{∂} onto P_c and that these maps are inverse to each other. Since they are also order-preserving (by (Gal2)), they are order-isomorphisms (by 1.36(4)).

7.28 From a closure operator to a Galois connection. In a somewhat contrived way, we can recognize that every closure operator arises as the composite of the left and right maps of a Galois connection. To see this, let $c : P \to P$ be a closure operator. Define $Q := P_c$, let ${}^{\triangleright} : P \to P_c$ be given by $p^{\triangleright} := c(p)$, and ${}^{\triangleleft} : P_c \to P$ be the inclusion map. Then $c = {}^{\triangleright\triangleleft}$.

7.29 Concept lattices reviewed. We have already noted in 7.24(2) that with every context (G, M, I) there is an associated Galois connection $(', ')$ between $\mathcal{P}(G)$ and $\mathcal{P}(M)^{\partial}$. According to 7.27 there is an associated closure operator c on $\mathcal{P}(G)$ which maps each $A \subseteq G$ to A''. Similarly, we have a closure operator k on $\mathcal{P}(M)$ which maps $B \subseteq M$ to B''; note that there is no order-reversal on $\mathcal{P}(M)$ here. Under the correspondence between closure operators and topped $\bigcap$-structures (7.3), c corresponds to $\mathfrak{B}_G$ and k to $\mathfrak{B}_M$; recall from 3.4 that both $\mathfrak{B}_G$ and $\mathfrak{B}_M^{\partial}$ are isomorphic to $\mathfrak{B}(G, M, I)$. The fact that $\mathfrak{B}(G, M, I)$ is a complete lattice (Proposition 3.6) can now be seen in context.

7.30 The provenance of complete lattices: summing up. We bring together here the correspondences which provide alternative ways in which complete lattices manifest themselves.

- Every topped $\bigcap$-structure is a complete lattice (2.32) and, up to isomorphism, every complete lattice arises this way (Exercise 2.29).
- The set of concepts $\mathfrak{B}(G, M, I)$ associated with a context (G, M, I) forms a complete lattice, and, up to isomorphism, every complete lattice arises this way (the fundamental theorem of concept lattices, 3.6–3.9).
- There is a bijective correspondence between closure operators on a set X and topped $\bigcap$-structures on X (7.3).
- Every Galois connection $({}^{\triangleright}, {}^{\triangleleft})$ gives rise to a pair of closure operators, ${}^{\triangleright\triangleleft}$ and ${}^{\triangleleft\triangleright}$, and thence to an isomorphic pair of complete lattices (7.27).

We have seen that Galois connections occur widely and have a natural place in the theory of complete lattices. But much of their usefulness stems from the remaining results in this section, which reveal the interaction between the maps in a Galois connection and joins and meets.

7.31 Proposition. *Let $({}^{\triangleright}, {}^{\triangleleft})$ be a Galois connection between ordered sets P and Q. Then ${}^{\triangleright}$ preserves existing joins in the sense defined in 2.26. Likewise, ${}^{\triangleleft}$ preserves existing meets.*

Proof. We first define $z := \bigvee_P S$ and show that $z^{\triangleright}$ is *an* upper bound for $S^{\triangleright}$. By (Gal2),

$$(\forall s \in S)\, s \leqslant z \Longrightarrow (\forall s \in S)\, s^{\triangleright} \leqslant z^{\triangleright}.$$

Now let q be any upper bound for $S^{\triangleright}$. Then

$$\begin{aligned} (\forall s \in S)\, s^{\triangleright} \leqslant q &\iff (\forall s \in S)\, s \leqslant q^{\triangleleft} && \text{(by (Gal))} \\ &\Longrightarrow \textstyle\bigvee_P S \leqslant q^{\triangleleft} && \text{(by definition of } \textstyle\bigvee_P S\text{)} \\ &\iff \left(\textstyle\bigvee_P S\right)^{\triangleright} \leqslant q && \text{(by (Gal)).} \end{aligned}$$

We conclude that $z^{\triangleright}$ is the *least* upper bound of $S^{\triangleright}$. □

The following lemma could have been presented in Chapter 1 along with results about order-preserving maps but its significance would have been far from obvious at that stage. Observe that Exercise 1.24 implies that the inverse image under an order-preserving map of a principal down-set is a down-set; the import of condition (ii) is that it is a *principal* down-set.

7.32 Lemma. *Let P and Q be ordered sets and $\varphi : P \to Q$ an order-preserving map. Then the following are equivalent:*

(i) *there exists an order-preserving map $\varphi^{\sharp} : Q \to P$ such that both $\varphi^{\sharp} \circ \varphi \geqslant \mathrm{id}_P$ and $\varphi \circ \varphi^{\sharp} \leqslant \mathrm{id}_Q$;*

(ii) *for each $q \in Q$ there exists a (necessarily unique) $s \in P$ such that $\varphi^{-1}(\downarrow q) = \downarrow s$.*

Proof. Assume (i). We claim that $\varphi^{-1}(\downarrow q) = \downarrow\varphi^\sharp(q)$. We have

$$\begin{aligned} p \in \varphi^{-1}(\downarrow q) &\iff \varphi(p) \leqslant q \\ &\implies (\varphi^\sharp \circ \varphi)(p) \leqslant \varphi^\sharp(q) \text{ (since } \varphi^\sharp \text{ is order-preserving)} \\ &\implies p \leqslant \varphi^\sharp(q) \qquad (\text{from } \varphi^\sharp \circ \varphi \geqslant \mathrm{id}_P \ \&\ \text{transitivity}) \\ &\iff p \in \downarrow\varphi^\sharp(q). \end{aligned}$$

For the other direction, let $p \in \downarrow\varphi^\sharp(q)$. This yields $\varphi(p) \leqslant (\varphi \circ \varphi^\sharp)(q)$ from which we can deduce that $\varphi(p) \leqslant q$, so that $p \in \varphi^{-1}(\downarrow q)$. Therefore (ii) holds.

Now assume (ii). For each $q \in Q$ we have a unique element $s \in P$, depending on q, such that $\varphi^{-1}(\downarrow q) = \downarrow s$. Define $\varphi^\sharp(q) := s$. Restated, this means that

$$(\forall q \in Q)(\forall p \in P)\ \varphi(p) \leqslant q \iff p \leqslant \varphi^\sharp(q).$$

We now see that the pair $(\varphi, \varphi^\sharp)$ is a Galois connection between P and Q, so that the properties in (i) follow from Lemma 7.26. □

7.33 $^\triangleright$ from $^\triangleleft$ and $^\triangleleft$ from $^\triangleright$. We can now interpret the proof of (i) ⇒ (ii) in Lemma 7.32, and the dual proof, in the notation we customarily use for Galois connections. We obtain the important fact that in a Galois connection $(^\triangleright, {}^\triangleleft)$ each of $^\triangleright$ and $^\triangleleft$ uniquely determines the other via the formulae

$$\begin{aligned} p^\triangleright &= \min\{\, q \in Q \mid p \leqslant q^\triangleleft \,\}, \\ q^\triangleleft &= \max\{\, p \in P \mid p^\triangleright \leqslant q \,\}. \end{aligned}$$

(For a subset S of an ordered set, $\min S$ and $\max S$ denote respectively the least and greatest elements of S, when these exist.)

7.34 Proposition. *Let P and Q be ordered sets and $\varphi\colon P \to Q$ be a map.*

(i) *Assume P is a complete lattice. Then φ preserves arbitrary joins if and only if φ possesses an upper adjoint $\varphi^\sharp$ (that is, $(\varphi, \varphi^\sharp)$ is a Galois connection).*

(ii) *Assume Q is a complete lattice. Then φ preserves arbitrary meets if and only if φ possesses a lower adjoint $\varphi^\flat$ (that is, $(\varphi^\flat, \varphi)$ is a Galois connection).*

Proof. We prove (i). The backward implication comes from Proposition 7.31. For the forward implication, assume that φ preserves arbitrary joins. Note first that φ is order-preserving, by Proposition 2.19. It

therefore suffices to show that condition (ii) in Lemma 7.32 is satisfied. Let $q \in Q$. We claim that

$$s := \bigvee_P \{\, p \in P \mid \varphi(p) \leqslant q \,\} \ (= \bigvee_P \varphi^{-1}(\downarrow q))$$

is such that $\varphi^{-1}(\downarrow q) = \downarrow s$. It follows immediately that $\varphi^{-1}(\downarrow q) \subseteq \downarrow s$. Since φ preserves arbitrary joins,

$$\varphi(s) = \bigvee_Q \{\, \varphi(p) \mid p \in P \text{ with } \varphi(p) \leqslant q \,\}$$

and hence $\varphi(s) \leqslant q$. For any $p \in \downarrow s$, we have $\varphi(p) \leqslant q$, because φ is order-preserving and $\leqslant$ is transitive. Therefore $\downarrow s \subseteq \varphi^{-1}(\downarrow q)$. □

In 7.25, we indicated one way in which Galois connections arise in the study of computational models, but we gave no intimation of what makes this concept valuable to computer scientists. We conclude this section by giving a very brief overview of the role Galois connections can play in the development of programs from specifications. We seek here to build a bridge between pure order theory and one important area of application. Our primary aim is to inform a mathematical readership for whom navigation through the vast literature can be daunting. References to much fuller accounts can be found in Appendix B.

7.35 Refinement. The development of a computer program often starts from a specification of the task the program is to perform and finishes with program code in a selected programming language so that the final program correctly meets the initial specification. At the initial (abstract) level, the specification must be precise and intelligible; it will not necessarily indicate an algorithm, or even a strategy, for performing the desired task: for example, 'find a solution x of the equation $x^2 = 2$ correct to five decimal places' is a valid specification. In the final (concrete) implementation, the program code may not be easily intelligible. At the outset, even when an algorithm is available, it may perform the desired task in an inefficient way. On the other hand, the efficiency of the final executable program is likely to be a major issue. In practice the transformation from specification to executable program will be carried out in stages. The objective is to do this by applying fixed rules which guarantee correctness at each step. This process is known as stepwise refinement.

By way of illustration, let us consider an imperative style of programming which is free of the distracting detail that intrudes in a final implementation. Programs are described operationally, that is, by the actions they perform: assignment, sequential composition, `abort` and so forth. A suitable setting in which to work is provided by a **specification space** $\langle \mathcal{S}; \sqsubseteq \rangle$ of **commands**, relative to some fixed state space X.

This space consists of a chosen imperative programming language ('real' program code) augmented by specifications (descriptions of computations, not cast in an executable form). The specification space is thus designed to be a universe within which program development can be carried out. It is worthwhile to include in $\mathcal{S}$ commands such as `magic` (which miraculously meets every specification) that are far removed from code for feasible specifications; `magic` provides a top for $\mathcal{S}$. The admittance of commands for arbitrary non-deterministic choice corresponds to the existence of arbitrary meets in $\mathcal{S}$; given a family of commands, non-deterministic choice will select one of these commands, but without there being any control on which one. So (recall 2.31) $\langle \mathcal{S}; \sqsubseteq \rangle$ is a complete lattice. Mathematically this is advantageous. It means that the full power of the theory of Galois connections is available. In particular maps from $\mathcal{S}$ to $\mathcal{S}$ preserving arbitrary meets (joins) possess lower (upper) adjoints. The existence of such adjoints guarantees the existence of commands that may assist in program development. Further, the calculational rules obeyed by Galois connections will supply laws governing these commands.

The notion of refinement is more versatile than the discussion above may suggest. In general, one mechanism is said to be refined by another of like type when every specification satisfied by the first is satisfied also by the second. Here the term mechanism can refer to an individual program, either in a formal guise or semantically modelled by relations, predicate transformers or whatever, or to a command. Refinement may also involve enriching a language with new constructs to increase its expressivity (algorithmic refinement). Alternatively it might involve moving from abstract datatypes such as sets and bags to more concrete versions such as strings and arrays (data refinement). Extending our earlier usage, we denote by $\sqsubseteq$ the relation 'is refined by'. It is clearly reflexive and transitive, and so a quasi-order. Its transitivity validates a stepwise strategy in moving from abstract to concrete. It is entirely reasonable for many aspects of program development not to distinguish between algorithmic specifications with the same end result or between syntactically distinct programs which achieve the required objective. With this perspective we may regard $\sqsubseteq$ as an order. (The process of obtaining an order from a quasi-order was explained formally in Exercise 6.1.)

The instances in which refinement is most productive are those in which one ordered structure $\langle A; \leqslant \rangle$ is refined by another $\langle C; \leqslant \rangle$, and where there exists a Galois connection $({}^{\triangleright}, {}^{\triangleleft})$ between A (abstract) and C (concrete). As we have done earlier, we think of the orders on A and C as having the interpretation 'is less informative than', as in 1.8, so that $x \leqslant y$ means that y is at least as good as x or that y serves every purpose that x does. In case A and C are semantic models for programming at different levels of abstraction it may help to think of the Galois maps ${}^{\triangleright} : A \to C$ (concretization) and ${}^{\triangleleft} : C \to A$ (abstraction) as, respectively, compilation and verification. We think of $c^{\triangleleft}$ as the best available approximation, in the more abstract structure A, to the element c in the more concrete structure C. Assume that a programming command is described by c in C and suppose that we wish to show that c implements some specification a in the more abstract model A. We can either prove that $a^{\triangleright} \leqslant c$ in the more concrete model or (equivalently, since $({}^{\triangleright}, {}^{\triangleleft})$ is a Galois connection) prove that $a \leqslant c^{\triangleleft}$ in the more abstract model. We have $c^{\triangleleft\triangleright} \leqslant c$, for any $c \in C$. This expresses the fact that, in general, abstraction results in a loss of information: if we have concrete knowledge given by $c \in C$ and we abstract this via ${}^{\triangleleft}$ and then concretize via ${}^{\triangleright}$, the result, $c^{\triangleleft\triangleright}$, may contain less information than the original c.

In summary, there is a trade-off between different semantic models of programming: crude models may be easy to understand but cannot fully capture intended behaviours, whereas more refined models may be unacceptably unwieldy. There is clear merit in being able to move backwards and forwards between two such models via a Galois connection, faithfully lifting up the semantic features of the simpler one to the more sophisticated one. In the context of relational and predicate transformer models for imperative programs (see 7.25), such Galois maps are provided by the map taking a relation to the associated weakest precondition transformer and its upper adjoint (here the natural order on predicate transformers has to be reversed to make the maps order-preserving). This Galois connection can be extended from models of programs to models of specifications to yield a powerful refinement calculus.

Completions

This section provides important applications of our earlier theory, in particular of the fundamental theorem of concept lattices, but it is not used later in the book. In general, there are many ways in which an ordered set can be embedded into a complete lattice. In this section we shall concentrate on one such embedding. This is associated with the Galois connection $({}^{u}, {}^{\ell})$ introduced in 7.24(3); it generalizes Dedekind's construction of $\mathbb{R}$ as the completion by cuts of $\mathbb{Q}$.

7.36 Definition and remarks. Let P be an ordered set. If $\varphi\colon P \hookrightarrow L$ and L is a complete lattice, then we say that L is a **completion** of P (via

the order-embedding φ). It follows from Lemma 1.30 (see Exercise 2.29) that the map $\varphi : x \mapsto {\downarrow}x$ is an order-embedding of P into $\mathcal{O}(P)$. We saw in 2.6(3) that $\mathcal{O}(P)$ is a complete lattice; hence $\mathcal{O}(P)$ is a completion of P. This completion is unnecessarily large. For example, if P is a complete lattice then P is a completion of itself (via the identity map) while $\mathcal{O}(P)$ is much larger. Another completion of an ordered set is the ideal completion (see Exercise 9.6).

7.37 Resumé. For ease of reference we repeat the definition and basic properties of the maps ^{u} and $^\ell$ on an ordered set P. The properties come from 7.26, or can easily be verified directly.

Let $A \subseteq P$. Then A 'upper' and A 'lower' are defined by

$$A^u := \{ x \in P \mid (\forall a \in A)\, a \leqslant x \} \quad \text{and} \quad A^\ell := \{ x \in P \mid (\forall a \in A)\, a \geqslant x \}.$$

For subsets A and B of P, we have

(i) $A \subseteq A^{u\ell}$ and $A \subseteq A^{\ell u}$,

(ii) if $A \subseteq B$, then $A^u \supseteq B^u$ and $A^\ell \supseteq B^\ell$,

(iii) $A^u = A^{u\ell u}$ and $A^\ell = A^{\ell u \ell}$.

Further, A^u is an up-set and A^ℓ is a down-set.

7.38 The Dedekind–MacNeille completion. Let P be an ordered set. We define

$$\mathbf{DM}(P) := \{ A \subseteq P \mid A^{u\ell} = A \}.$$

This is the topped $\bigcap$–structure on P corresponding to the closure operator $C(A) := A^{u\ell}$ on P. Therefore the ordered set $\langle \mathbf{DM}(P); \subseteq \rangle$ is a complete lattice. It is known as the **Dedekind–MacNeille completion** of P. (It is also referred to as the **completion by cuts** or the **normal** completion of P.)

As we noted in 7.24, the Galois connection associated with the context $(P, P, \leqslant)$ is $(^u, ^\ell)$. Therefore we may view $\mathbf{DM}(P)$ as the lattice $\mathfrak{B}_P$ associated with $(P, P, \leqslant)$, meaning that $\mathfrak{B}_P$ is the isomorphic copy of $\mathfrak{B}(P, P, \leqslant)$ obtained via the projection $\pi_1 : (A, B) \mapsto A$.

7.39 Lemma. *Let P be an ordered set.*

(i) *For all $x \in P$, we have $({\downarrow}x)^{u\ell} = {\downarrow}x$ and hence ${\downarrow}x \in \mathbf{DM}(P)$.*

(ii) *If $A \subseteq P$ and $\bigvee A$ exists in P, then $A^{u\ell} = {\downarrow}(\bigvee A)$.*

Proof. (i) Let $y \in ({\downarrow}x)^u$; then $z \leqslant y$ for all $z \in {\downarrow}x$ so, in particular, $x \leqslant y$ (as $x \in {\downarrow}x$) and hence $y \in {\uparrow}x$. Thus $({\downarrow}x)^u \subseteq {\uparrow}x$. If $y \in {\uparrow}x$, then $y \geqslant x$ and so, by transitivity, $y \geqslant z$ for all $z \in {\downarrow}x$, that is, $y \in ({\downarrow}x)^u$. Thus ${\uparrow}x \subseteq ({\downarrow}x)^u$. Therefore $({\downarrow}x)^u = {\uparrow}x$ and, by duality, $({\uparrow}x)^\ell = {\downarrow}x$. Thus $({\downarrow}x)^{u\ell} = ({\uparrow}x)^\ell = {\downarrow}x$.

(ii) Let $A \subseteq P$ and assume that $\bigvee A$ exists in P. Of course $\bigvee A \in A^u$. Thus $x \in A^{u\ell}$ implies that $x \leqslant \bigvee A$ and hence $x \in {\downarrow}(\bigvee A)$. Consequently $A^{u\ell} \subseteq {\downarrow}(\bigvee A)$. Since $\bigvee A$ is the *least* upper bound of A we have $\bigvee A \leqslant y$ for all $y \in A^u$ and hence $\bigvee A \in A^{u\ell}$. Since $A^{u\ell}$ is a down-set this gives ${\downarrow}(\bigvee A) \subseteq A^{u\ell}$. Hence $A^{u\ell} = {\downarrow}(\bigvee A)$, as required. □

7.40 Theorem. *Let P be an ordered set and define $\varphi : P \to \mathbf{DM}(P)$ by $\varphi(x) = {\downarrow}x$ for all $x \in P$.*

(i) *$\mathbf{DM}(P)$ is a completion of P via the map φ.*

(ii) *φ preserves all joins and meets which exist in P.*

Proof. (i) As we saw above, $\langle \mathbf{DM}(P); \subseteq \rangle$ is a complete lattice and the order-embedding φ maps P into $\mathbf{DM}(P)$.

(ii) Let $A \subseteq P$ and assume that $\bigvee A$ exists in P. We must show that $\varphi(\bigvee A) = \bigvee \varphi(A)$, that is, ${\downarrow}(\bigvee A) = \bigvee \{ {\downarrow}a \mid a \in A \}$ in $\mathbf{DM}(P)$. Clearly, ${\downarrow}(\bigvee A)$ is an upper bound for $\{ {\downarrow}a \mid a \in A \}$. Now choose any $B \in \mathbf{DM}(P)$ which is an upper bound for $\{ {\downarrow}a \mid a \in A \}$. Since $a \in {\downarrow}a \subseteq B$, for all $a \in A$, we have $A \subseteq B$. Hence,

$${\downarrow}(\bigvee A) = A^{u\ell} \subseteq B^{u\ell} = B,$$

by 7.39(ii).

Now assume that $\bigwedge A$ exists in P. We must show that

$${\downarrow}(\bigwedge A) = \bigwedge \{ {\downarrow}a \mid a \in A \}.$$

Since $\mathbf{DM}(P)$ is a topped $\bigcap$–structure, we have

$$\bigwedge \{ {\downarrow}a \mid a \in A \} = \bigcap \{ {\downarrow}a \mid a \in A \}, \text{ in } \mathbf{DM}(P),$$

and hence Exercise 2.6 yields the result. □

We can now use the fundamental theorem of concept lattices to characterize the Dedekind–MacNeille completion.

7.41 Theorem. *Let P be an ordered set and let $\varphi : P \to \mathbf{DM}(P)$ be the order-embedding of P into its Dedekind–MacNeille completion given by $\varphi(x) = {\downarrow}x$.*

(i) *$\varphi(P)$ is both join-dense and meet-dense in $\mathbf{DM}(P)$.*

(ii) *Let L be a complete lattice and assume that P is a subset of L which is both join-dense and meet-dense in L. Then $L \cong \mathbf{DM}(P)$ via an order-isomorphism which agrees with φ on P.*

Proof. Consider (i). The order-isomorphism π_1 between $\mathfrak{B}(P, P, \leqslant)$ and $\mathbf{DM}(P)$ preserves join- and meet-density. Hence join- and meet-density of $\varphi(P)$ follow from Theorem 3.8. Part (ii) follows from Theorem 3.9, with the isomorphism ψ defined by $\psi(x) = \{ g \in P \mid g \leqslant x \}$. (Each of γ and μ is simply the natural embedding of P into L.) □

7.42 Theorem. *Let L be a lattice with no infinite chains. Then*

$$L \cong \mathbf{DM}(\mathcal{J}(L) \cup \mathcal{M}(L)).$$

Moreover, $\mathcal{J}(L) \cup \mathcal{M}(L)$ is the smallest subset of L which is both join-dense and meet-dense in L.

Proof. For the first part, recall Theorems 3.9, 2.41 and 2.46.

It remains to show that, if P is both join-dense and meet-dense in L, then $\mathcal{J}(L) \cup \mathcal{M}(L) \subseteq P$. Let $x \in \mathcal{J}(L)$. Since P is join-dense there exists a subset A of P with $x = \bigvee A$. Hence, by 2.41(i), there is a finite subset F of A with $x = \bigvee F$. Since x is join-irreducible we have $x \in F$ and so $x \in A$. Hence $\mathcal{J}(L) \subseteq P$. By duality, $\mathcal{M}(L) \subseteq P$ too. □

7.43 Remarks. In 5.1 we put forward properties we would wish a representing subset P for a lattice L to possess: L should be determined by P, in a simple manner, and P should be 'small' and easily identifiable. We showed in Chapter 5 that when L is a finite distributive lattice then $P = \mathcal{J}(L)$ (or, equivalently, $P = \mathcal{M}(L)$) serves the purpose admirably. We now see that, so long as the lattice L has no infinite chains, then the ordered set $P = \mathcal{J}(L) \cup \mathcal{M}(L)$ is a possible substitute 'skeleton'. However, in general we must construct a Dedekind–MacNeille completion to recapture L from P and it is debatable whether this process could be described as simple. Nonetheless a representation for arbitrary finite lattices can be developed from this starting point, and results about finite lattices derived from corresponding results on finite ordered sets. We do not pursue this further.

7.44 Examples. The characterization of the Dedekind–MacNeille completion provided by Theorem 7.41 may be used in several ways. For example, given an ordered set P, it may allow us to guess the structure of $\mathbf{DM}(P)$ without actually constructing the family of subsets of P which satisfy $A^{u\ell} = A$. Alternatively, given a complete lattice L it may enable us to find subsets P of L such that $L \cong \mathbf{DM}(P)$. These ideas are illustrated in the examples which follow. These examples also illustrate Theorem 7.42. In the diagram of each lattice L the set of shaded elements is $\mathcal{J}(L) \cup \mathcal{M}(L)$. Indeed it was via Theorem 7.42 that the examples were constructed in the first place.

(1) It is not difficult to see that every real number $x \in \mathbb{R}$ satisfies $\bigvee_{\mathbf{R}}(\downarrow x \cap \mathbb{Q}) = x = \bigwedge_{\mathbf{R}}(\uparrow x \cap \mathbb{Q})$ and hence $\mathbb{Q}$ is both join-dense and meet-dense in $\mathbb{R} \cup \{-\infty, \infty\}$. Consequently $\mathbb{R} \cup \{-\infty, \infty\}$ is (order-isomorphic to) the Dedekind–MacNeille completion of $\mathbb{Q}$.

(2) $\mathbf{DM}(\mathbb{N}) \cong \mathbb{N} \oplus \mathbf{1}$.

(3) For any set X, the complete lattice $\wp(X) \cong \mathbf{DM}(P)$ where

$$P = \{\{x\} \mid x \in X\} \cup \{X \setminus \{x\} \mid x \in X\}.$$

(4) The Dedekind–MacNeille completion of an n-element antichain (for $n \geqslant 1$) is order-isomorphic to the lattice $\mathbf{M}_n$ (see Figure 2.4); recall 3.11(3).

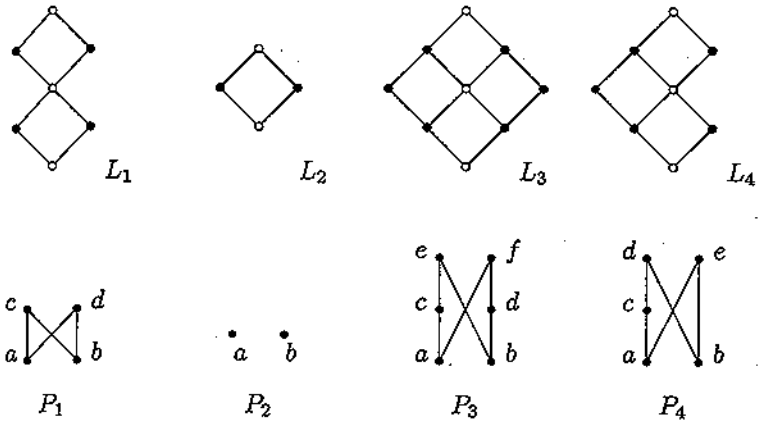


Figure 7.1

(5) Each pair of diagrams in Figure 7.1 may be interpreted either as an ordered set P_i along with its Dedekind–MacNeille completion $L_i \cong \mathbf{DM}(P_i)$ or as a lattice L_i with a distinguished subset P_i such that $L_i \cong \mathbf{DM}(P_i)$. In each case the elements of P_i are shaded.

Exercises

Exercises from the text. Complete the proof of Theorem 7.3. Verify the assertions in 7.5. Prove the second part of Lemma 7.16 and confirm the results given in Table 7.1. Verify the assertions in 7.27.

7.1 Suppose that C is a closure operator on X and let $A \subseteq X$ and $A_i \subseteq X$ for each $i \in I$. Show that

(i) $C\left(\bigcup_{i\in I} A_i\right) \supseteq \bigcup_{i\in I} C(A_i)$,

(ii) $C(A) \supseteq \bigcup\{C(B) \mid B \subseteq A \text{ and } B \text{ is finite}\}$.

7.2 Let $L = \{0\} \cup \{2^r 3^s \mid r, s = 0, 1, 2, 3, \dots\}$ be ordered by $\preccurlyeq$ where $m \preccurlyeq n \Leftrightarrow (\exists k \in \mathbb{N}_0)\, n = km$.

(i) Prove that L is a complete, distributive lattice.

(ii) Identify the compact (equivalently, finite) elements in L and in the dual lattice L^∂. Is L an algebraic lattice? Is L^∂ an algebraic lattice?

7.3 In which of the following cases is D a directed subset of P?

(i) $P = \wp(\mathbb{N})$, $D = \{ X \subseteq \mathbb{N} \mid (\forall m \in \mathbb{N})\, 2m \in X \}$.

(ii) $P = \wp(\mathbb{N})$, $D = \{ X \subseteq \mathbb{N} \mid \mathbb{N} \setminus X \text{ is finite} \}$.

(iii) $P = (\mathbb{N} \multimap \mathbb{N})$, $D = \{ \pi \in P \mid \pi(n) \leqslant 2 \text{ for all } n \in \operatorname{dom} \pi \}$.

(iv) $P = \mathbb{N}_0^2$, $D = \{ (0,n) \mid n \in \mathbb{N} \} \cup \{ (n,0) \mid n \in \mathbb{N} \}$.

(v) $P = (\mathbb{N}_0 \to \mathbb{N}_0)$, $D = \{f_i\}_{i \geqslant 1}$, where $f_i \colon \mathbb{N}_0 \to \mathbb{N}_0$ is defined by $f_i(i) = 1$ and $f_i(j) = 0$ for $j \neq i$.

(vi) $P = (\mathbb{N}; \preccurlyeq)$, $D = \{ 2^a 3^b \mid a, b \in \mathbb{N} \text{ and } a + b \text{ is prime} \}$.

7.4 Let R be a commutative ring with identity, 1, and let $\mathfrak{L}$ be the $\bigcap$–structure of all ideals of R. Prove that $R \in K(\mathfrak{L})$.

7.5 Let P be an ordered set and let $S \subseteq P$. Prove that if the joins on the right-hand side exist then $\bigvee S$ exists and is given by

$$\bigvee S = \bigsqcup \{ \bigvee F \mid \varnothing \neq F \subseteq S,\ F \text{ finite} \}.$$

7.6 A non-empty ordered set S is called a **join semilattice** if $x \vee y$ exists in S for all $x, y \in S$. As in the lattice case, a non-empty subset J of S is an **ideal** of S if J is closed under finite joins and going down. (See 2.20.)

(i) Show that the set $\mathcal{I}(S)$ of ideals of a join semilattice with $\bot$ is a topped algebraic $\bigcap$–structure on S and that $K(\mathcal{I}(S)) \cong S$.

(ii) By Lemma 7.16, $K(L)$ is a join semilattice with $\bot$ for any complete lattice L. Prove that the following are equivalent:

(a) L is an algebraic lattice;

(b) $L \cong \mathcal{I}(K(L))$;

(c) $L \cong \mathcal{I}(S)$ where S is some join semilattice with $\bot$.

[Hint. Show that J is an ideal of $K(L)$ if and only if J is of the form D_a for some $a \in L$, then appeal to Theorem 7.20.]

7.7 Let L be an algebraic lattice and K a non-empty subset of L such that $\bigvee_L S$ and $\bigwedge_L S$ belong to K for every non-empty subset S of K. Show that K is an algebraic lattice. [Hint. Represent L as a topped algebraic $\bigcap$–structure on some set X and show that K is also a topped algebraic $\bigcap$–structure on some subset Y of X.]

Deduce that the interval $[x, y] := \{ z \in L \mid x \leqslant z \leqslant y \}$ is an algebraic lattice for all $x, y \in L$ with $x \leqslant y$.

7.8 Let L be a complete lattice. Prove that the following statements are equivalent:

(i) $K(L) = L$;

(ii) every directed subset of L has a greatest element;

(iii) L satisfies (ACC).

7.9 Let P be an arbitrary ordered set and let $\langle \mathcal{O}(P); \subseteq \rangle$ be its lattice of down-sets. Assume that $\mathcal{O}(P)$ satisfies (ACC). Prove that P satisfies (ACC). Hence or otherwise prove that if P is infinite then $\mathcal{O}(P)$ has an infinite descending chain. Deduce that $\mathcal{O}(P)$ satisfies both (ACC) and (DCC) if and only if P is finite.

7.10 Let P be an ordered set. Define $X = \{ {\downarrow}x \mid x \in P \}$, ordered by inclusion, and $Y = \{ {\uparrow}x \mid x \in P \}$, ordered by reverse inclusion. Define ${}^\triangleright \colon X \to Y$ by ${\downarrow}x \mapsto {\uparrow}x$ and ${}^\triangleleft \colon Y \to X$ by ${\uparrow}x \mapsto {\downarrow}x$. Check that $({}^\triangleright, {}^\triangleleft)$ sets up a Galois connection between X and Y.

7.11 Assume that $({}^\triangleright, {}^\triangleleft)$ is a Galois connection between ordered sets P and Q. Prove the following are equivalent for $p_1, p_2 \in P$:

(i) $p_1{}^\triangleright \leqslant p_2{}^\triangleright$;

(ii) $p_1{}^{\triangleright\triangleleft} \leqslant p_2{}^{\triangleright\triangleleft}$;

(iii) $p_1 \leqslant p_2{}^{\triangleright\triangleleft}$.

7.12 Let P and Q be ordered sets and let ${}^\triangleright \colon P \to Q$ and ${}^\triangleleft \colon Q \to P$ be maps. Prove that the following are equivalent:

(i) $({}^\triangleright, {}^\triangleleft)$ is a Galois connection;

(ii) the following hold for all $p \in P$, $q \in Q$:

(a) ${}^\triangleright$ is order-preserving,

(b) $q^{\triangleleft\triangleright} \leqslant q$,

(c) $p^\triangleright \leqslant q \implies p \leqslant q^\triangleleft$;

(ii)$^\partial$ the following hold for all $p \in P$, $q \in Q$:

(a)$^\partial$ ${}^\triangleleft$ is order-preserving,

(b)$^\partial$ $p \leqslant p^{\triangleright\triangleleft}$,

(c)$^\partial$ $p^\triangleright \leqslant q \impliedby p \leqslant q^\triangleleft$.

7.13 Let $({}^\triangleright, {}^\triangleleft)$ be a Galois connection.

(i) Prove that the following are equivalent:

(a) ${}^\triangleright$ is a surjective map;

(b) $^{\triangleleft}$ is an injective map;
(c) $q^{\triangleleft} = \max\{ s \in P \mid s^{\triangleright} = q \}$;
(d) $^{\triangleright\triangleleft} = \mathrm{id}_Q$.

(ii) Formulate the dual statement.

7.14 In which of the following cases does $\varphi\colon P \to Q$ possess (a) an upper adjoint $\varphi^\sharp\colon Q \to P$, (b) a lower adjoint $\varphi^\flat\colon Q \to P$? Describe the maps $\varphi^\sharp$ and $\varphi^\flat$ when they exist.

(i) $P = Q = \mathbb{N}$ and $\varphi(n) = mn$, for fixed m in $\mathbb{N}$.

(ii) $P = Q = \mathbb{R}$ and φ is the function **floor**, so that $\varphi(x)$, usually denoted $\lfloor x \rfloor$, is the greatest integer $\leqslant x$.

(iii) $P = Q = \wp(S)$ for some set S and $\varphi(Y) = A \cap Y$, where A is a fixed subset of S.

7.15 Let P, Q and R be ordered sets and let $\varphi\colon P \to Q$ and $\psi\colon Q \to R$ be order-preserving maps. Prove that if φ and ψ have upper adjoints $\varphi^\sharp$ and $\psi^\sharp$, respectively, then the composite $\psi \circ \varphi$ has upper adjoint $\varphi^\sharp \circ \psi^\sharp$ (so that $(\psi\circ\varphi, \varphi^\sharp \circ \psi^\sharp)$ is a Galois connection).

7.16 Let P and Q be ordered sets and let $\varphi_1, \varphi_2\colon P \to Q$ be order-preserving maps which have upper adjoints $\varphi_1^\sharp, \varphi_2^\sharp\colon Q \to P$. Show that, in the pointwise order, $\varphi_1 \leqslant \varphi_2$ if and only if $\varphi_1^\sharp \geqslant \varphi_2^\sharp$.

7.17 Let P and Q be complete lattices and, for $i \in I$, let $\varphi_i\colon P \to Q$ be an order-preserving map with an upper adjoint $\varphi_i^\sharp\colon Q \to P$. Define $\varphi\colon P \to Q$ by $\varphi(x) := \bigvee_{i\in I} \varphi_i(x)$, for all $x \in P$. Show that φ has an upper adjoint given by $\varphi^\sharp(y) = \bigwedge_{i\in I} \varphi_i^\sharp(y)$, for all $y \in Q$.

7.18 Let $R \subseteq A \times B$ be a relation between sets A and B and define maps $F_R\colon \wp(A) \to \wp(B)$ and $G_R\colon \wp(B) \to \wp(A)$ by

$$F_R(S) := \{ b \in B \mid (\exists a \in S)\, (a,b) \in R \}, \text{ and}$$
$$G_R(T) := \{ a \in A \mid (\forall b \in B)\, ((a,b) \in R \Rightarrow b \in T) \},$$

for all $S \subseteq A$ and $T \subseteq B$. Prove that (F_R, G_R) is a Galois connection between $\wp(A)$ and $\wp(B)$.

Conversely, let (F,G) be a Galois connection between $\wp(A)$ and $\wp(B)$ and define $R \subseteq A \times B$ by

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

Prove that $(F,G) = (F_R, G_R)$.

7.19 Let L and M be bounded lattices, with associated ideal lattices $\mathcal{I}(L)$ and $\mathcal{I}(M)$, and let (f,g) be a Galois connection between L and M.

(i) Show that there is a well-defined map $G\colon \mathcal{I}(M) \to \mathcal{I}(L)$ given for $J \in \mathcal{I}(M)$ by $G(J) := f^{-1}(J)$, and show also that $G(J) = \downarrow g(J)$.

(ii) Show that G has a lower adjoint F, given by $F(I) = \downarrow f(I)$ for $I \in \mathcal{I}(L)$.

7.20 Let L be a bounded lattice, with ideal lattice $\mathcal{I}(L)$ and filter lattice $\mathcal{F}(L)$. Consider the relation $R \subseteq \mathcal{I}(L) \times \mathcal{F}(L)$ given by

$$(I,F) \in R \iff I \cap F \neq \varnothing.$$

Consider the context $(\mathcal{I}(L), \mathcal{F}(L), R)$, the associated Galois connection between $\wp(\mathcal{I}(L))$ and $\wp(\mathcal{F}(L))^\partial$ and the closure operator c on $\wp(\mathcal{I}(L))$ as defined in 7.27.

(i) Prove that, for any $a \in L$,

$$c(\{\downarrow a\}) = \{ I \in \mathcal{I}(L) \mid a \in I \}.$$

(ii) Deduce that α given by $\alpha\colon a \mapsto \{ I \in \mathcal{I}(L) \mid a \in I \}$, is an order-embedding of L into the complete lattice $c(\wp(\mathcal{I}(L)))$ (that is, the image of the map c).

(iii) Prove that, calculated in $c(\wp(\mathcal{I}(L)))$, we have, for all $S \subseteq L$,

$$\textstyle\bigwedge \alpha(S) = \{ I \in \mathcal{I}(L) \mid I \supseteq S \},$$
$$\textstyle\bigvee \alpha(S) = \{ I \in \mathcal{I}(L) \mid (\forall F \in \mathcal{F}(L))\, (F \supseteq S \Rightarrow F \cap I \neq \varnothing) \}.$$

(iv) Prove that $c(\mathcal{S}) = \bigvee \{ \bigwedge \alpha(I) \mid I \in \mathcal{S} \}$, for all $\mathcal{S} \subseteq \mathcal{I}(L)$.

7.21 Recall that a quasi-order on a set P is a binary relation on P which is reflexive and transitive. Let P be a set and assume that $\leqslant_1$ and $\leqslant_2$ are quasi-orders on P such that $x \leqslant_1 y$ and $x \leqslant_2 y$ imply $x = y$. Define, for a subset Y of P,

$$Y^{\triangleright} := \{ x \in P \mid (\forall y)(y \leqslant_1 x \Rightarrow y \notin Y) \},$$
$$Y^{\triangleleft} := \{ x \in P \mid (\forall y)(y \leqslant_2 x \Rightarrow y \notin Y) \}.$$

A subset Y of P is called **stable** if $Y^{\triangleright\triangleleft} = Y$; denote the set of all stable subsets of P by $S(P)$.

(i) Verify that the pair $({}^{\triangleright}, {}^{\triangleleft})$ sets up a Galois connection between the $\leqslant_1$-down-sets of P ordered by $\subseteq$ and the $\leqslant_2$-down-sets of P ordered by $\supseteq$. (Down-sets of a quasi-ordered set are defined in just the same way as down-sets of an ordered set.)

(ii) Verify that $\langle \mathcal{S}(P); \subseteq \rangle$ is a lattice in which meet and join are given by $Y \wedge Z = Y \cap Z$ and $Y \vee Z = (Y^{\triangleright} \cap Z^{\triangleright})^{\triangleleft}$.

(iii) Let P be a 3-element set $\{a, b, c\}$ with $\leqslant_1$ defined by $x \leqslant_1 y$ if and only if $x = y$ or $x = a$ and $y = b$ and $\leqslant_2$ defined by $x \leqslant_2 y$ if and only if $x = y$ or $x = a$ and $y = c$. Identify the stable subsets of P and prove that $\langle \mathcal{S}(P); \subseteq \rangle \cong \mathbf{N}_5$.

(iv) Let P be a 6-element set with elements $a_1, a_2, \ldots, a_6$ and define $\leqslant_1$ and $\leqslant_2$ by

$$x \leqslant_1 y \iff x = y \text{ or } \{x, y\} \in \{\{a_1, a_2\}, \{a_3, a_4\}, \{a_5, a_6\}\},$$
$$x \leqslant_2 y \iff x = y \text{ or } \{x, y\} \in \{\{a_2, a_3\}, \{a_4, a_5\}, \{a_6, a_1\}\}.$$

Find the stable subsets of P and prove that $\langle \mathcal{S}(P); \subseteq \rangle \cong \mathbf{M}_3$.

(This exercise hints at the concrete representation for arbitrary bounded lattices, due to A. Urquhart, which in the finite case generalizes Birkhoff's representation. In the distributive case, $\leqslant_1$ and $\leqslant_2$ are respectively $\leqslant$ and $\geqslant$, for an order $\leqslant$.)

7.22 Let L be a complete lattice, let P be a subset of L which is both join-dense and meet-dense in L and let $A \subseteq P$. Prove that $A^{u\ell} = {\downarrow}(\bigvee_L A) \cap P$. (Here $A^{u\ell}$ is calculated completely within P.)

7.23 Find the Dedekind–MacNeille completion of each of the ordered sets P_1 to P_4 in Figure 7.1 and thereby verify that $L_i \cong \mathbf{DM}(P_i)$ in each case. [Use the definition of $\mathbf{DM}(P)$; do not appeal to Theorem 7.41.]

7.24 Let Q be an ordered set, let P be a subset of Q with the inherited order and assume that P is both join-dense and meet-dense in Q. Let $\varphi\colon P \to \mathbf{DM}(P)$ be the embedding given in 7.41. Prove that there is an order-embedding ψ of Q into $\mathbf{DM}(P)$ such that ψ agrees with φ on P, that is, $\psi(x) = \varphi(x)$ for all $x \in P$.

8

CPOs and Fixpoint Theorems

This chapter pursues themes hinted at in earlier chapters. Its main focus is on order-theoretic fixpoint theorems, of which the Knaster–Tarski Theorem in Chapter 2 was a first example, within the setting of complete lattices. We mostly work with CPOs, which generalize complete lattices and which we introduced in passing in the previous chapter. For a comparative summary of the various theorems see 8.24.

CPOs

Consider a map $f\colon \mathbb{N} \to \mathbb{N}$. We may regard each partial function σ in $(\mathbb{N} \multimap \mathbb{N})$, with $\sigma \leqslant f$ and $\operatorname{dom}\sigma$ finite, as supplying an approximation to f containing a finite amount of information. Such approximations often arise as the output after a finite number of steps in a computation of the values of f. These partial maps form a directed set in $(\mathbb{N} \multimap \mathbb{N})$ with join f. In the context of information orderings more generally, directed joins provide a natural framework in which to model partial approximations to total objects. These observations lead us to study CPOs systematically.

8.1 Definition. We say that an ordered set P is a **CPO** (an abbreviation for **complete partially ordered set**) if

(i) P has a bottom element, $\bot$,

(ii) $\bigsqcup D$ $(:= \bigvee D)$ exists for each directed subset D of P.

Some authors do not require a CPO to have a bottom element and so omit (i), using the term **pointed CPO** when both (i) and (ii) hold. When we do not wish to assume $\bot$ is present in an ordered set P satisfying (ii) or do not want to regard $\bot$, even if it exists, as an integral part of the structure, we say P is a **pre-CPO**. Note that $P_\bot$ is a CPO whenever P is a pre-CPO. In the literature a pre-CPO is often called a **dcpo** (for directedly complete partial order) and a CPO is called a **dcppo** (the extra 'p' standing for 'pointed').

8.2 Examples.

(1) Recall that in an ordered set P satisfying (ACC) a non-empty subset is directed if and only if it has a greatest element. Therefore P is a pre-CPO and its lifting, $P_\bot$, is a CPO. In particular, any flat

ordered set ($S_\perp$, for some set S) is a CPO. Indeed, this example is the primary motivation for introducing the lifting construction.

(2) Any complete lattice is a CPO.

(3) An algebraic $\bigcap$-structure is a CPO, with $\bigsqcup$ coinciding with $\bigcup$.

(4) We claim that the set Σ^{**} of all binary strings is a CPO under the prefix ordering. A directed subset D of Σ^{**} is necessarily a chain. Its join is the well-defined string having nth element defined and equal to δ ($\delta = 0$ or 1) if and only if some $u \in D$ has nth element defined and equal to δ.

Many additional CPOs can be constructed by taking suitable subsets of CPOs or by combining CPOs in various ways.

8.3 Sub-CPOs. Given a CPO P, we say that a subset Q of P is a **sub-CPO** of P if

(i) the bottom element of P belongs to Q,

(ii) whenever D is a directed subset of Q, the join $\bigsqcup_P D$ belongs to Q.

By 2.28, condition (ii) is equivalent to

(ii)′ whenever D is a directed subset of Q, the join $\bigsqcup_Q D$ exists and coincides with $\bigsqcup_P D$.

In a pre-CPO, (i) is dropped. It is possible for a subset of a CPO P to be a CPO in its own right, without being a sub-CPO of P. For example, under the inclusion order,

$$\{\, S \subseteq \mathbb{N} \mid S \text{ is finite or } S = \mathbb{N} \,\}$$

is a CPO but is not a sub-CPO of $\wp(\mathbb{N})$: consider the join of the directed set $\{\, S \subseteq \mathbb{N} \setminus \{1\} \mid S \text{ finite} \,\}$.

8.4 Sums and products of CPOs. Let P and Q be ordered sets each with a bottom element. Then their disjoint union $P \cup Q$ fails to have a bottom element. There are two ways to modify this construction and stay within the class of ordered sets with $\perp$. The first is to form $(P \cup Q)_\perp$; this is the **separated sum** of P and Q, which we write as $P \oplus_\perp Q$.

Figure 8.1

Alternatively, we may form the **coalesced sum**, $P \oplus_\vee Q$, by taking $P \cup Q$ and identifying the two bottom elements. Figure 8.1 shows examples of such sums.

8.5 Lemma. *Let P and Q be CPOs and let S be a set.*

(i) *Each of $P \oplus_\perp Q$, $P \oplus_\vee Q$ and $P \times Q$ is a CPO.*

(ii) *The power $(S \to P)$ of P is a CPO in which directed joins are calculated pointwise. More precisely, if $\{\varphi_i\}_{i \in I}$ is a directed subset of $(S \to P)$, then, for each $s \in S$, the set $\{\varphi_i(s)\}_{i \in I}$ is a directed subset of P and $(\bigsqcup_{i \in I} \varphi_i)(s) = \bigsqcup_{i \in I} \varphi_i(s)$.*

Proof. It is elementary to show that each of $P \oplus_\perp Q$ and $P \oplus_\vee Q$ is a CPO. Exercise 8.5 provides guidance on how to show that $P \times Q$ is also a CPO. If $\{\varphi_i\}_{i \in I}$ is a directed subset of $(S \to P)$, then, for each $s \in S$, the set $\{\varphi_i(s)\}_{i \in I}$ is a directed subset of P and consequently, since P is a CPO, we may define a map $\varphi\colon S \to P$ by $\varphi(s) := \bigsqcup_{i \in I} \varphi_i(s)$. We leave it as an easy exercise to show that φ is the least upper bound in $(S \to P)$ of $\{\varphi_i\}_{i \in I}$. □

8.6 Continuous maps. In analysis, a function is continuous if it preserves limits. In a context in which a computation is modelled as the join (= limit) of a directed set, it is natural to consider a map as being continuous if it is compatible with the formation of directed joins; see also Exercise 8.8. Formally, we say that $\varphi\colon P \to Q$ (where P and Q are pre-CPOs) is **continuous** if, for every directed set D in P, the subset $\varphi(D)$ of Q is directed and

$$\varphi\Big(\bigsqcup D\Big) = \bigsqcup \varphi(D) \quad \Big(:= \bigsqcup \{\, \varphi(x) \mid x \in D \,\}\Big).$$

Note that since the empty set is not directed (by definition), a continuous map need not preserve bottoms. A map $\varphi\colon P \to Q$ such that $\varphi(\perp) = \perp$ is called **strict**. The natural structure-preserving maps for pre-CPOs are the continuous maps and for CPOs the strict continuous maps. The next lemma shows that every continuous map is order-preserving. Where the order represents 'is less defined than' or 'is a worse approximation than', an order-preserving map φ is one which is such that, the better the input x, the better the output $\varphi(x)$. Thus only maps which are order-preserving are likely to be of computational significance. For many applications, continuity, which is generally a strictly stronger property than order-preservation, is the appropriate one.

8.7 Lemma. *Let P and Q be CPOs and φ be a map from P to Q.*

(i) *Suppose D is a directed subset of P and φ is order-preserving. Then $\varphi(D)$ is a directed subset of Q and $\bigsqcup \varphi(D) \leqslant \varphi(\bigsqcup D)$. In*

particular, $\bigsqcup_{n\geqslant 0}\varphi(x_n) \leqslant \varphi(\bigsqcup_{n\geqslant 0} x_n)$, *for any ascending chain* $x_0 \leqslant x_1 \leqslant x_2 \leqslant \cdots$ *in* P.

(ii) *If* $\varphi(D)$ *is directed and* $\bigsqcup\varphi(D) \leqslant \varphi(\bigsqcup D)$, *for every directed set* D *in* P, *then* φ *is order-preserving.*

Proof. The first part is a consequence of 2.27(i) and of 7.8(1). For the second, take x, y in P such that $x \leqslant y$. Then $D := \{x, y\}$ is directed, whence $\varphi(x) \leqslant \bigsqcup\varphi(D) \leqslant \varphi(\bigsqcup D) = \varphi(y)$. □

8.8 Order-preserving versus continuous.

(1) Not every order-preserving map between CPOs is continuous. Consider $\varphi\colon \wp(\mathbb{N}) \to \wp(\mathbb{N})$ defined by

$$\varphi(S) = \begin{cases} \varnothing & \text{if } S \text{ is finite,} \\ \mathbb{N} & \text{otherwise.} \end{cases}$$

The collection $\mathcal{D}$ of finite subsets of $\mathbb{N}$ is directed, with $\bigsqcup\varphi(\mathcal{D}) = \varnothing$ and $\varphi(\bigsqcup\mathcal{D}) = \mathbb{N}$.

(2) The continuity condition can be awkward to check. It is therefore useful to know when it is satisfied automatically. Using 7.8(2) and 8.7(i), we see that an order-preserving map $\varphi\colon P \to Q$ is continuous whenever P satisfies (ACC), and in particular whenever P is the flat CPO $S_\perp$, for some set S.

Let P and Q be pre-CPOs. The set of all continuous maps from P to Q, with the pointwise order, is denoted $[P \to Q]$. This function space construction provides an important way of building new CPOs.

8.9 Theorem. *Let* P *and* Q *be pre-CPOs. Then* $[P \to Q]$ *is a pre-CPO, and is a CPO whenever* Q *is a CPO. Directed joins in* $[P \to Q]$ *are calculated pointwise.*

Proof. Note that the constant map onto $\perp$ acts as $\perp$ in $[P \to Q]$ whenever $\perp$ exists in Q. Let $E = \{\varphi_i\}_{i\in I}$ be a directed subset of $[P \to Q]$. For all $x \in P$, the subset $\{\varphi_i(x)\}_{i\in I}$ of Q is directed, since E is directed, and so $\bigsqcup_{i\in I}\varphi_i(x)$ exists in the pre-CPO Q. Thus the pointwise join, $\varphi := \bigsqcup_{i\in I}\varphi_i$, of E is well defined and is order-preserving by Exercise 2.27. We establish that φ is continuous by showing that $\varphi(\bigsqcup D) \leqslant \bigsqcup\varphi(D)$ for an arbitrary directed subset D of P. Bracket-

pushing, we have

$$\begin{aligned}\varphi(\textstyle\bigsqcup D) &= (\bigsqcup_{i\in I}\varphi_i)(\bigsqcup D) \\ &= \bigsqcup_{i\in I}(\varphi_i(\bigsqcup D)) \qquad \text{(by the definition of } \textstyle\bigsqcup_{i\in I}\varphi_i\text{)} \\ &= \bigsqcup_{i\in I}(\bigsqcup_{x\in D}\varphi_i(x)),\end{aligned}$$

because each φ_i is continuous. But for each $y \in D$ and each $j \in I$,

$$\varphi_j(y) \leqslant \bigsqcup_{x\in D}(\bigsqcup_{i\in I}\varphi_i(x)).$$

Hence, for each $j \in I$, we have $\bigsqcup_{x\in D}\varphi_j(x) \leqslant \bigsqcup_{x\in D}(\bigsqcup_{i\in I}\varphi_i(x))$. So

$$\varphi(\textstyle\bigsqcup D) = \bigsqcup_{i\in I}(\bigsqcup_{x\in D}\varphi_i(x)) \leqslant \bigsqcup_{x\in D}(\bigsqcup_{i\in I}\varphi_i(x)) = \bigsqcup_{x\in D}\varphi(x) = \bigsqcup\varphi(D).$$ □

Note that if P is a pre-CPO and Q is a CPO, then $[P \to Q]$ is a sub-CPO of the power $(P \to Q)$: see Lemma 8.5.

8.10 Directed sets versus chains. The directed sets in terms of which CPOs have been defined link CPOs with the algebraic lattices in Chapter 7 and with the domains which we shall meet in Chapter 9. Arguably, in the context of approximations, it would be adequate and less complicated to work with ascending chains rather than with directed sets. The proof of Theorem 8.15 below seems to bear this out.

It turns out that an ordered set is a CPO provided that each chain has a least upper bound in P. (Note that the join of the empty chain guarantees the existence of $\perp$.) We omit the proof of this highly non-trivial result, which we record below as a theorem. Exercise 8.9 seeks a proof in the countable case; the general case requires the machinery of ordinals. The theorem would make it legitimate for directed sets to be replaced by non-empty chains in the remainder of this chapter, but no major simplification would result from this.

8.11 Theorem. *Let* P *be an ordered set. Then* P *is a CPO if and only if each chain has a least upper bound in* P.

In the literature, an ordered set in which each chain has a least upper bound is often called **completely inductive** (the reason for this term emerges in 10.4). Thus, CPOs and completely inductive ordered sets are one and the same thing.

CPOs of partial maps

Within computer science the fixpoint theorems, which will be the focus of the remainder of this chapter, are frequently applied to CPOs of partial maps. These seminal examples deserve further attention before we continue with our general theory.

Let S be any non-empty set. We can use the lifting process, introduced in 1.22, to give simple but important insights into the ordered set of partial maps on S.

Define $S_\perp$ as in 1.22; as before, we denote the bottom element of $S_\perp$ by $\mathbf{0}$. We can use the extra element $\mathbf{0}$ to convert any partial map $\pi \in (S \multimap S)$ into an 'ordinary' map defined on the whole of S: we let $\pi_\perp : S \to S_\perp$ be given by

$$\pi_\perp(x) = \begin{cases} \pi(x) & \text{if } x \in \operatorname{dom}\pi, \\ \mathbf{0} & \text{otherwise.} \end{cases}$$

We can then define a map V from $(S \multimap S)$ to $(S \to S_\perp)$ by $V(\pi) = \pi_\perp$.

8.12 Lemma. *Let S be a non-empty set. Then the map $V : \pi \mapsto \pi_\perp$ is an order-isomorphism between $(S \multimap S)$ and $(S \to S_\perp)$, and both these ordered sets are CPOs.*

Proof. For $\pi, \sigma \in (S \multimap S)$, we have

$$\begin{aligned} V(\pi) \leqslant V(\sigma) \text{ in } (S \to S_\perp) &\iff (\forall x \in S)(\pi_\perp(x) \leqslant \sigma_\perp(x)) \\ &\iff (\forall x \in S)(\pi_\perp(x) = \mathbf{0} \text{ or } \pi_\perp(x) = \sigma_\perp(x)) \\ &\iff (\forall x \in S)(x \in \operatorname{dom}\pi \Rightarrow (x \in \operatorname{dom}\sigma \text{ and } \pi(x) = \sigma(x))) \\ &\iff \pi \leqslant \sigma \text{ in } (S \multimap S). \end{aligned}$$

(The first equivalence uses the definition of $\leqslant$ in $(S \to S_\perp)$ and the second uses the definition of $\leqslant$ in $S_\perp$.) We have proved that V is an order-embedding. To show that V is onto, take $f : S \to S_\perp$. Let $T := \{ x \in S \mid f(x) \neq \mathbf{0} \}$ and define $\pi : T \to S$ by $\pi(x) = f(x)$ for $x \in T$. Then $V(\pi) = f$. Hence V is an order-isomorphism.

Since $S_\perp$ is a (flat) CPO, $(S \to S_\perp)$ is also a CPO, by 8.5. Since V is an order-isomorphism, $(S \multimap S)$ is a CPO too. Alternatively, it is easy to see this directly once we identify each partial map on S with its graph (see 1.10). □

Lemma 8.12 lets us identify the CPOs $(S \multimap S)$ and $(S \to S_\perp)$ and, for a partial map $\pi \in (S \multimap S)$, to write $\pi(x) = \perp$ to mean that $x \in S \setminus \operatorname{dom}\pi$. The bottom element of $(S \multimap S)$ is (the map whose graph is) $\varnothing$; this corresponds to the map in $(S \to S_\perp)$ which sends every element of S to $\perp$ in $S_\perp$. We denote this map by $\perp$ in either case.

To round out this section, we present two examples which illustrate why, in the context of computations with natural numbers,

- it is natural to work with CPOs of the form $P = (S \multimap S)$, and
- why fixpoints and, in particular, least fixpoints of maps $F : P \to P$ play an important role.

We begin with an example which is a favourite with computer scientists.

8.13 The factorial function. Consider the set $\mathbb{N}_0 = \{0, 1, \dots\}$. The factorial function, **fact**, is the map on $\mathbb{N}_0$ given by $\mathbf{fact}(k) = k!$. It satisfies

$$\mathbf{fact}(k) = \begin{cases} 1 & \text{if } k = 0, \\ k\,\mathbf{fact}(k-1) & \text{if } k \geqslant 1. \end{cases}$$

To each map $f : \mathbb{N}_0 \to \mathbb{N}_0$ we may associate a new map $\overline{f}$ given by

$$\overline{f}(k) = \begin{cases} 1 & \text{if } k = 0, \\ kf(k-1) & \text{if } k \geqslant 1. \end{cases}$$

Define a map F on $(\mathbb{N}_0 \to \mathbb{N}_0)$ by $F(f) = \overline{f}$. Then we must have $F(\mathbf{fact}) = \mathbf{fact}$.

To determine $\mathbf{fact}(k)$ for a given $k \geqslant 1$, we need to know $\mathbf{fact}(k-1)$, and unless $k = 1$ this requires knowledge of $\mathbf{fact}(k-2)$, and so on. What we have here is a **recursive equation** $F(f) = f$, satisfied by **fact** (recursive, from its latin roots, meaning running backwards). The entire factorial function cannot be unwound from its recursive specification in a finite number of steps. However, we can, for each $n \in \mathbb{N}_0$, determine in a finite number of steps the partial map f_n which is the restriction of **fact** to $\{0, 1, \dots, n\}$; the graph of f_n is $\{(0,1), (1,1), \dots, (n, n!)\}$. To accommodate approximations to **fact**, it is therefore natural to work not with maps from $\mathbb{N}_0$ to $\mathbb{N}_0$ but with all partial maps on $\mathbb{N}_0$. When this is done, we regard $\overline{f}$ as having $\{0\} \cup \{ k \mid k-1 \in \operatorname{dom} f \}$ as its domain, for each $f \in (\mathbb{N}_0 \multimap \mathbb{N}_0)$. In summary, we may extend F to a map from $(\mathbb{N}_0 \multimap \mathbb{N}_0)$ to itself and regard the factorial function as a solution of the recursive equation $F(f) = f$.

In a similar way, we may consider the solutions $f \in (\mathbb{Z} \multimap \mathbb{Z})$ of $G(f) = f$, where $G(f)(0) = 1$, $G(f)(k) = kf(k-1)$ if $k \geqslant 1$ and $f(k-1) \neq \perp$, and $G(f)(k) = f(k)$ if $k < 0$. A fixpoint f of G is uniquely determined on non-negative integers. However, $f(-1)$ may be assigned any value $\alpha \in \mathbb{Z}$, so that f contains extraneous unforced information. Thus G has many different fixpoints. Nevertheless, G does

have a least fixpoint: the partial map coinciding with **fact** on $\mathbb{N}_0$ and otherwise undefined. Typically in applications it is the *least* fixpoint which is of most significance.

Fixpoint theorems

Much mathematical effort is expended on solving equations. These may be of very diverse types, but many important ones can be expressed in the form $F(x) = x$, where $F\colon X \to X$ is a map; the set X might be a set of real numbers, maps, or sets, or might be of some other type. A solution of such a **fixpoint equation**, when one exists, often has to be obtained by a process of successive approximation. Order theory plays a role when X carries an order and when the solution can be realized as the join of elements which approximate it. CPOs provide a natural class of ordered sets within which to develop such a fixpoint theory.

As indicated above, a compelling reason for investigating fixpoints comes from computer science. Our objective here is to present the rudiments of the theory without discussing specialized applications. However, in the previous section we gave an intimation of a connection between fixpoints and recursive programs. In Chapter 9, which deals with domains, we hint at the use of fixpoints in the solution of domain equations.

We now turn to generalities.

8.14 Definitions and notation. We shall be concerned with maps $F\colon P \to P$, where P is some ordered set. We call such a map F a **self-map on** P (another commonly used term is **endofunction**). We say that $x \in P$ is a **fixpoint** of F if $F(x) = x$, a **pre-fixpoint** if $F(x) \leqslant x$, and a **post-fixpoint** if $x \leqslant F(x)$. The sets of such points are denoted, respectively, by $\mathrm{fix}(F)$, $\mathrm{pre}(F)$, and $\mathrm{post}(F)$; each carries the induced order from P. The least element of $\mathrm{fix}(F)$, when it exists, is denoted $\mu(F)$, and the greatest by $\nu(F)$, if this exists.

We have already proved one fixpoint theorem. The Knaster–Tarski Theorem, 2.35, tells us that every order-preserving self-map on a complete lattice has both least and greatest fixpoints. But this theorem suffers from two disadvantages from the standpoint of computer science. First, it concerns a map on a complete lattice, which is necessarily topped. Second, it does not give an algorithmic procedure for finding a fixpoint. We overcome the first difficulty by moving from a complete lattice to a CPO. For a continuous map on a CPO we can also overcome the second defect very easily.

We shall use the n-fold composite, F^n, of a map $F : P \to P$. This is defined as follows: F^0 is the identity map and $F^n = F \circ F^{n-1}$ for $n \geqslant 1$. If F is order-preserving, so is F^n.

8.15 CPO Fixpoint Theorem I. *Let P be a CPO, let F be an order-preserving self-map on P and define $\alpha := \bigsqcup_{n\geqslant 0} F^n(\bot)$.*

(i) *If $\alpha \in \mathrm{fix}(F)$, then $\alpha = \mu(F)$.*

(ii) *If F is continuous, then the least fixpoint $\mu(F)$ exists and equals α.*

Proof. (i) Certainly $\bot \leqslant F(\bot)$. Applying the map F^n, we have $F^n(\bot) \leqslant F^{n+1}(\bot)$, for all n. Hence we have a chain

$$\bot \leqslant F(\bot) \leqslant \cdots \leqslant F^n(\bot) \leqslant F^{n+1}(\bot) \leqslant \cdots$$

in P. Since P is a CPO, $\alpha := \bigsqcup_{n\geqslant 0} F^n(\bot)$ exists. Let β be any fixpoint of F. By induction, $F^n(\beta) = \beta$ for all n. We have $\bot \leqslant \beta$, whence we obtain $F^n(\bot) \leqslant F^n(\beta) = \beta$ by applying F^n. The definition of α forces $\alpha \leqslant \beta$. Hence if α is a fixpoint then it is the least fixpoint.

(ii) It will be enough to show that $\alpha \in \mathrm{fix}(F)$. We have

$$\begin{aligned} F\Big(\bigsqcup_{n\geqslant 0} F^n(\bot)\Big) &= \bigsqcup_{n\geqslant 0} F(F^n(\bot)) \quad \text{(since } F \text{ is continuous)} \\ &= \bigsqcup_{n\geqslant 1} F^n(\bot) \\ &= \bigsqcup_{n\geqslant 0} F^n(\bot) \qquad \text{(since } \bot \leqslant F^n(\bot) \text{ for all } n\text{).} \end{aligned}$$

□

8.16 Iteration and convergence. An instructive parallel may be drawn between CPO Fixpoint Theorem I and another well-known fixpoint theorem. Banach's Contraction Mapping Theorem asserts that a contraction map F on a complete metric space has a fixpoint, and this can be obtained as the limit of the sequence defined iteratively by $x_{n+1} := F(x_n)$, starting from a first approximation, x_0. The approximating sequence converges because the metric space is complete, and its limit is a fixpoint thanks to the continuity of the map. In 8.15 we simply have order-theoretic notions of completeness and continuity replacing the topological ones in Banach's Theorem.

Consider the chain $P = \mathbb{N} \oplus \mathbf{2}$ and let F be the (order-preserving) map fixing $\top$ and taking every other element to its upper cover. This map has $\top$ as its unique fixpoint, but $\top \neq \bigsqcup_{n\geqslant 0} F^n(\bot)$. This example shows that for a non-continuous map F on a CPO we cannot expect $\bigsqcup_{n\geqslant 0} F^n(\bot)$ necessarily to provide a fixpoint even when one exists.

Referring back to 8.10, we note that the proof of CPO Fixpoint Theorem I uses joins of chains, rather than of directed sets. This suggests

that a definition of CPOs in terms of chains would be quite justifiable for fixpoint theory. Note too that in the proof we explicitly used the existence of $\perp$ in P but we did not require our map F to preserve $\perp$. This suggests that we should work in CPOs, rather than pre-CPOs, but that we should not demand that maps be strict.

8.17 Applying CPO Fixpoint Theorem I to CPOs of partial maps. In many practical applications of CPO Fixpoint Theorem I, the map F is defined on a CPO of the form $(S \multimap S)$. Observe that a map $F\colon (S \multimap S) \to (S \multimap S)$ is order-preserving if and only if

$$\operatorname{graph} f \subseteq \operatorname{graph} g \implies \operatorname{graph} F(f) \subseteq \operatorname{graph} F(g),$$

for all $f, g \in (S \multimap S)$. When this condition is satisfied, $\{F^n(\perp)\}_{n\geqslant 0}$ forms a directed set of partial maps with successively bigger domain sets. Further, $\operatorname{graph}(\bigsqcup_{n\geqslant 0} F^n(\perp)) = \bigcup_{n\geqslant 0} \operatorname{graph} F^n(\perp)$. In applications it is usually helpful to work with graphs.

The graph perspective makes it transparent that the domain of definition of an element σ of $(S \multimap S)$ critically affects its relationships, in the order of $(S \multimap S)$, to other elements. This can potentially cause difficulty when we wish to define a map $F\colon (S \multimap S) \to (S \multimap S)$, so that the image elements $F(\pi)$ are partial maps (as occurs in 8.13, for example): we need to take care to prescribe their domains appropriately, especially where we have a specification 'by cases'. We always take such maps $F(\pi)$ to have the maximum possible domain. In other words, $\operatorname{dom} F(\pi)$ consists of all points $x \in S$ for which the definition of $(F(\pi))(x)$ makes sense. In simple cases, this is just what common sense would lead us to do.

8.18 The factorial function revisited. The recursive specification for **fact** comes within the scope of CPO Fixpoint Theorem I. In 8.13, we recognized **fact** as the solution in $(\mathbb{N}_0 \multimap \mathbb{N}_0)$ of a fixpoint equation $F(f) = f$, where

$$(F(f))(k) = \begin{cases} 1 & \text{if } k = 0, \\ kf(k-1) & \text{if } k \geqslant 1. \end{cases}$$

It is obvious that F is order-preserving. It is true, but far less obvious, that F is continuous (Exercise 8.17 outlines a proof). We have $\operatorname{graph} F(\perp) = \{(0,1)\}$, $\operatorname{graph} F^2(\perp) = \{(0,1),(1,1)\}$, etc. An easy induction confirms that $f_n = F^{n+1}(\perp)$ for all n, where $\{f_n\}_{n\geqslant 0}$ is the sequence of partial maps defined in 8.13 as approximations to **fact**. The directed join of $\{f_n\}_{n\geqslant 0}$ is obtained by taking the map (it is a well-defined map!) whose graph is $\bigcup_{n\geqslant 0} \operatorname{graph} f_n$, so that $\bigsqcup_{n\geqslant 0} F^n(\perp)$ is the map **fact**: $k \mapsto k!$ on $\mathbb{N}_0$. It is the least fixpoint of F in $(\mathbb{N}_0 \multimap \mathbb{N}_0)$.

8.19 Further examples. Only for a map which is continuous does CPO Fixpoint Theorem I guarantee the existence of a fixpoint. In some simple cases continuity can be side-stepped by appealing to the first part of the theorem. This is convenient since verifying continuity can be a non-trivial undertaking, especially when the underlying CPO is a set of maps.

(1) Let Σ^{**} be the CPO of all binary strings. Given a finite string u and an arbitrary string v, we denote by uv the string obtained by concatenating u and v. Let $F(u) = 01u$ for $u \in \Sigma^{**}$. It is intuitively clear that the fixpoint equation $F(u) = u$ has a unique solution, namely α, the infinite string of alternating zeros and ones. This is exactly the solution we obtain by taking the empty string, $\varnothing$, and forming $\bigsqcup_{n\geqslant 0} F^n(\varnothing)$. Clearly, $F^n(\varnothing)$ is the $2n$-element string $0101\ldots01$. The join in Σ^{**} of these strings is α, which is certainly a fixpoint of F. Trivially F is order-preserving so, by 8.15(i), α is indeed the least fixpoint.

(2) The fixpoint formula in CPO Fixpoint Theorem I may be used to determine certain recursively defined maps of more than one variable. For each $\pi \in (\mathbb{N}^2 \multimap \mathbb{N})$ define $\overline{\pi}$ by

$$\overline{\pi}(j,k) = \begin{cases} 1 & \text{if } j = k, \\ (k+1)\pi(j,k+1) & \text{otherwise;} \end{cases}$$

the domain of $\overline{\pi}$ is $\{(j,k) \in \mathbb{N}^2 \mid j = k \text{ or } (j,k+1) \in \operatorname{dom}\pi\}$. We now define $F\colon (\mathbb{N}^2 \multimap \mathbb{N}) \to (\mathbb{N}^2 \multimap \mathbb{N})$ to be the map taking π to $\overline{\pi}$. It is easy to see that F is order-preserving. A rather complicated induction on n shows that, for all $n \geqslant 1$, we have

$$F^n(\perp)(j,k) = \begin{cases} j!/k! & \text{if } 0 \leqslant j-k \leqslant n-1, \\ \perp & \text{otherwise.} \end{cases}$$

The join of $\{F^n(\perp)\}_{n\geqslant 0}$ is the element σ of $(\mathbb{N}^2 \multimap \mathbb{N})$ that is undefined at (j,k) if $k > j$ and takes value $j!/k!$ there if $k \leqslant j$. It is easy to check that $\overline{\sigma} = \sigma$ and hence, by 8.15(i), $\mu(F) = \sigma$. It is not difficult to show that σ is the unique fixpoint of F in $(\mathbb{N}^2 \multimap \mathbb{N})$. Therefore, the recursive equation

$$\pi(j,k) = \begin{cases} 1 & \text{if } j = k, \\ (k+1)\pi(j,k+1) & \text{otherwise} \end{cases}$$

has a unique solution in $(\mathbb{N}^2 \multimap \mathbb{N})$, given by the partial map σ.

Although we have shown that it is sometimes possible to get useful information from Theorem 8.15 for maps which are not continuous or not known to be continuous, it is clearly worthwhile to ask whether every

order-preserving self-map on a CPO has a fixpoint or, better, a least fixpoint. This proves to be the case, but it is much less straightforward to establish than either Theorem 2.35 or Theorem 8.15.

We begin with a simple but very useful observation which characterizes a least fixpoint. A special case of this was implicit in the proof of the Knaster–Tarski Theorem, 2.35.

8.20 Least fixpoints and least pre-fixpoints. *Let P be an ordered set and let F be an order-preserving self-map on P.*

(i) *Assume that F possesses a least pre-fixpoint $\mu_*(F)$. Then F has a least fixpoint, $\mu(F)$, which satisfies*

$$F(x) \leqslant x \Rightarrow \mu(F) \leqslant x \qquad \text{(the \textbf{Induction Rule})}.$$

Indeed, $\mu(F) = \mu_*(F)$.

(ii) *Assume that P is a complete lattice. Then $\mu_*(F)$ exists and hence* (i) *is applicable.*

Proof. (i) Assume that $\mu_*(F)$ exists. The Induction Rule says simply that $\mu(F)$ is a lower bound of $\mathrm{pre}(F)$. Since $\mu_*(F)$ is certainly a lower bound of $\mathrm{pre}(F)$, it suffices to prove that $\mu(F) = \mu_*(F)$. As $\mu_*(F) \in \mathrm{pre}(F)$, we have $F(\mu_*(F)) \leqslant \mu_*(F)$. Applying the order-preserving map F we find that $F(F(\mu_*(F))) \leqslant F(\mu_*(F))$, that is, $F(\mu_*(F)) \in \mathrm{pre}(F)$. Since $\mu_*(F)$ is the least element of $\mathrm{pre}(F)$ we therefore have $\mu_*(F) \leqslant F(\mu_*(F))$. Hence, $\mu_*(F)$ is a fixpoint of F. Since $\mathrm{fix}(F) \subseteq \mathrm{pre}(F)$ we must then have $\mu(F) = \mu_*(F)$.

For (ii), we have an obvious candidate for $\mu_*(F)$, namely $\bigwedge \mathrm{pre}(F)$. We need to check that $\bigwedge \mathrm{pre}(F) \in \mathrm{pre}(F)$. But this follows easily from the assumption that F is order-preserving. □

Before stating and proving our second CPO fixpoint theorem, we require a further concept and a preliminary lemma which is of interest in its own right. Let P be a CPO and let $F\colon P \to P$. The map F is said to be **increasing** if it satisfies $x \leqslant F(x)$ for all $x \in P$, that is, $\mathrm{post}(F) = P$. While increasing maps do not occur in the statement of CPO Fixpoint Theorem II, they will play a crucial role in its proof.

8.21 Lemma. *Let P be a CPO. Then the increasing order-preserving self-maps on P have a common fixpoint.*

Proof. Denote by $I(P)$ the set of increasing order-preserving self-maps on the CPO P. The set $I(P)$ is non-empty since $\mathrm{id}_P \in I(P)$. Let $F, G \in I(P)$ and $x \in P$. Then $F(x) \leqslant F(G(x))$ since G is increasing and F is order-preserving and $G(x) \leqslant F(G(x))$ since F is increasing. Therefore the map $F \circ G$ is an upper bound in $I(P)$ for $\{F, G\}$. Thus

$I(P)$ is a directed subset of the CPO $\langle P \to P\rangle$ of all order-preserving self-maps on P.

Let $H := \bigsqcup I(P)$ be the join of $I(P)$ in $\langle P \to P\rangle$. Since joins in $\langle P \to P\rangle$ are calculated pointwise, it is trivial that $H \in I(P)$. Let $G \in I(P)$. Since $G \circ H \in I(P)$ and H is the join of the set $I(P)$, we have $G \circ H \leqslant H$. On the other hand, $H \leqslant G \circ H$ since G is increasing. Hence $G \circ H = H$. It follows immediately that $H(x)$ is a fixpoint of G for all $x \in P$. □

8.22 CPO Fixpoint Theorem II. *Let P be a CPO and let $F\colon P \to P$ be order-preserving. Then F has a least fixpoint.*

Proof. We say that a subset Y of P is **F-invariant** if $F(Y) \subseteq Y$. Since the set of all F-invariant sub-CPOs of P is closed under intersection, P has a smallest F-invariant sub-CPO P_0. We shall show that P_0 has a top, $\top_{P_0}$, and that $\mu(F) = \top_{P_0}$.

Define a map $\Phi\colon \wp(P) \to \wp(P)$ by

$$\Phi(X) := \{\bot\} \cup F(X) \cup \{\textstyle\bigsqcup D \mid D \subseteq X \ \&\ D \text{ is directed}\}$$

for all $X \subseteq P$. It is easily checked that Φ is order-preserving. By the Knaster–Tarski Fixpoint Theorem, 2.35, the map Φ has a least fixpoint, given by $\bigcap\{X \in \wp(P) \mid \Phi(X) \subseteq X\}$. By the definition of Φ, this least fixpoint is the smallest F-invariant sub-CPO P_0 of P.

Claim 1. $P_0 \subseteq \mathrm{post}(F)$.

Proof. Since $\wp(P)$ is a complete lattice, the Induction Rule applies (see 8.20). Thus, in order to show that $P_0 := \mu(\Phi) \subseteq \mathrm{post}(F)$, it suffices to show that $\Phi(\mathrm{post}(F)) \subseteq \mathrm{post}(F)$. Let $Q := \mathrm{post}(F)$. Certainly, (a) $\bot \in Q$. Since F is order-preserving, $\mathrm{post}(F)$ is F-invariant, that is, (b) $F(Q) \subseteq Q$. Let D be a directed subset of Q. As F is order-preserving, $F(D)$ is directed and $\bigsqcup F(D) \leqslant F(\bigsqcup D)$. Since $x \leqslant F(x)$ for all $x \in D$, we also have $\bigsqcup D \leqslant \bigsqcup F(D)$, whence $\bigsqcup D \leqslant F(\bigsqcup D)$. Consequently, (c) $\bigsqcup D \in \mathrm{post}(F) = Q$. By (a), (b) and (c), we have $\Phi(Q) \subseteq Q$, as required.

Claim 2. If $x \in P$ is a fixpoint of F, then $P_0 \subseteq {\downarrow}x$.

Proof. Assume that $F(x) = x$. Since P_0 is the smallest F-invariant sub-CPO of P, it suffices to show that ${\downarrow}x$ is an F-invariant sub-CPO of P. The set ${\downarrow}x$ is a sub-CPO of P for all $x \in P$. Since $F(x) = x$, the fact that ${\downarrow}x$ is F-invariant is an immediate consequence of the fact that F is order-preserving.

As $F(P_0) \subseteq P_0$ we may define $G := F{\restriction}_{P_0}\colon P_0 \to P_0$. By Claim 1, G is increasing and hence, by Lemma 8.21 applied to the CPO P_0, we have $G(a) = a$, and therefore $F(a) = a$, for some $a \in P_0$.

Claim 3. The element a is both the top of P_0 and the least fixpoint of F.

Proof. Assume that $F(b) = b$ for some $b \in P$. It is clear from the definition of Φ that $\Phi(\downarrow b) \subseteq \downarrow b$. Again, by the Induction Rule (8.20), we obtain $P_0 = \mu(\Phi) \subseteq \downarrow b$. Since $a \in P_0$ we have $a \leqslant b$. □

CPO Fixpoint Theorem II is central to the study of fixpoints in computer science. The elegant proof of it presented above is due to D. Pataraia. Previous proofs came in two flavours. Some were quite straightforward but relied on the Axiom of Choice via Zorn's Lemma. (We present such a proof in Chapter 10: see 10.5.) Others avoided the Axiom of Choice but relied upon a further fixpoint theorem, stated below as CPO Fixpoint Theorem III. This result is of independent interest. However, especially when compared with Pataraia's proof of 8.22, its known proofs appear convoluted; Exercise 8.21 gives a step-by-step guide to one of these.

8.23 CPO Fixpoint Theorem III. *Let P be a CPO and let F be an increasing self-map on P. Then F has a minimal fixpoint.*

In fact, if P_0 is the smallest F-invariant sub-CPO of P, then P_0 has a top, $\top_{P_0}$, and $\top_{P_0}$ is a minimal element of $\mathrm{fix}(F)$. Note that we are not asserting the existence of a *least* fixpoint; the reader should construct a counterexample to show that F may not possess a least fixpoint.

8.24 Stocktaking. We have now presented three theorems guaranteeing the existence of a least fixpoint for a self-map F on an ordered set P and one asserting the existence of a minimal fixpoint.

- The **Knaster–Tarski Theorem** (2.35) asserts that when F is order-preserving and P is a complete lattice then F has least and greatest fixpoints given respectively by $\bigwedge \mathrm{pre}(F)$ and $\bigvee \mathrm{post}(F)$.
- **CPO Fixpoint Theorem I** (8.15) asserts that when F is continuous and P a CPO then F has a least fixpoint given by $\bigsqcup_{n \geqslant 0} F^n(\bot)$.
- **CPO Fixpoint Theorem II** (8.22) asserts that F has a least fixpoint if F is order-preserving and P is a CPO.
- **CPO Fixpoint Theorem III** (8.23) asserts that F has a minimal fixpoint if P is a CPO and F is increasing.

Readers who know about ordinals may have surmised that, in the case that F is order-preserving but not necessarily continuous, it is possible to extend the formula in 8.15 by taking a limit over a chain indexed by ordinals. This is indeed the case; an outline can be found in Exercise 8.19.

The Knaster–Tarski Theorem and CPO Fixpoint Theorem II are in a sense optimal. The following converses exist. Both are difficult to prove; see [71].

8.25 Theorem. *Let P be an ordered set.*

(i) *If P is a lattice and every order-preserving map $F\colon P \to P$ has a fixpoint, then P is a complete lattice.*

(ii) *If every order-preserving map $F\colon P \to P$ has a least fixpoint, then P is a CPO.*

Proposition 8.26 gives information about the set of all fixpoints of an order-preserving map on a complete lattice and on a CPO.

8.26 Proposition. *Let F be an order-preserving self-map on an ordered set P.*

(i) *If P is a complete lattice, then so is* $\mathrm{fix}(F)$.

(ii) *If P is a CPO, then so is* $\mathrm{fix}(F)$.

Proof. We prove only the second claim. For the proof of the first see Exercise 8.23.

Let P be a CPO and assume that F is order-preserving. By 8.22, $\mu(F)$ exists, so $\mathrm{fix}(F)$ has a bottom element. Now let D be a directed set in $\mathrm{fix}(F)$ and let $\alpha := \bigsqcup D$. Take $x \geqslant \alpha$. Then, for all $d \in D$, we have $x \geqslant d$, so that $F(x) \geqslant F(d) = d$. Hence $F(x) \geqslant \alpha$. We have shown that F maps $\uparrow\alpha$ into itself and may therefore consider the restriction G of F to $\uparrow\alpha$. Certainly $\uparrow\alpha$ is a CPO in the induced order. By 8.22, G has a least fixpoint, and this necessarily acts as $\bigsqcup_{\mathrm{fix}(F)} D$. We deduce that $\mathrm{fix}(F)$ is a CPO. □

Calculating with fixpoints

The preceding section has provided a clutch of theorems guaranteeing that certain maps on ordered sets possess fixpoints or, better still, (necessarily unique) least fixpoints. In this section we change our perspective and show that much information can be gleaned about least fixpoints from some simple calculational rules.

8.27 The need for a calculus for fixpoints. Often, computational procedures are modelled as least fixpoints. It may be important to verify that alternative formulations of a given procedure – for example, different implementations of a 'while' command – really do result in the same output. When the alternatives are specified as the least fixpoints

of two different self-maps F and G on some ordered set P, then we require $\mu(F) = \mu(G)$. Suppose, to take a very simple case, that we can show that $\mu(F) \in \mathrm{fix}(G)$ and $\mu(G) \in \mathrm{fix}(F)$. Then $\mu(F) \leqslant \mu(G)$ and $\mu(G) \leqslant \mu(F)$, so we do indeed have $\mu(F) = \mu(G)$. In more complicated situations, we may seek to show that $\mu(F)$ satisfies properties characterizing $\mu(G)$, or vice versa. We may also want to be able to recast expressions involving one or more least fixpoints in more amenable forms.

We have already given one rule that is frequently used in fixpoint calculus, namely the Induction Rule given in 8.20. Some of the other most frequently used calculational rules for fixpoints are presented below in 8.28–8.31 and in Exercises 8.28–8.31. As we shall see, the Induction Rule is often employed in the proofs. Since we are more interested in least fixpoints than in least pre-fixpoints, our focus in this section is on complete lattices, so that the Induction Rule tells us that $F(x) \leqslant x$ implies $\mu(F) \leqslant x$. We start with a very simple application.

8.28 The Monotonicity Rule. *Let F and G be order-preserving self-maps on a complete lattice P. Then $F \leqslant G$ implies $\mu(F) \leqslant \mu(G)$. (As usual, maps are ordered pointwise.)*

Proof. By the Induction Rule applied to F, the required conclusion follows if we can prove that $F(\mu(G)) \leqslant \mu(G)$. But since $F \leqslant G$ we have $F(\mu(G)) \leqslant G(\mu(G)) = \mu(G)$. □

8.29 The Rolling Rule. *Let P and Q be complete lattices and let $F\colon P \to Q$ and $G\colon Q \to P$ be order-preserving. Then*

$$G(\mu(F \circ G)) = \mu(G \circ F).$$

Proof. By 8.20, it suffices to prove that $G(\mu(F \circ G))$ is the least pre-fixpoint of $G \circ F$. For this, note first that

$$(G \circ F)(G(\mu(F \circ G))) = G((F \circ G)(\mu(F \circ G))) = G(\mu(F \circ G)).$$

Hence $G(\mu(F \circ G))$ is a fixpoint and therefore a pre-fixpoint of $G \circ F$. Now assume that $x \in P$ is such that $(G \circ F)(x) \leqslant x$. We require $G(\mu(F \circ G)) \leqslant x$. We have

$$\begin{aligned}
(G \circ F)(x) \leqslant x &\Longrightarrow (F \circ G)(F(x)) \leqslant F(x) \\
&\qquad \text{(since } F \text{ is order-preserving)} \\
&\Longrightarrow \mu(F \circ G) \leqslant F(x) \\
&\qquad \text{(by the characterization of } \mu(F \circ G)) \\
&\Longrightarrow G(\mu(F \circ G)) \leqslant (G \circ F)(x) \\
&\qquad \text{(since } G \text{ is order-preserving)} \\
&\Longrightarrow G(\mu(F \circ G)) \leqslant x \\
&\qquad \text{(by the assumption } (G \circ F)(x) \leqslant x). \qquad \square
\end{aligned}$$

8.30 The Fusion Rule. *Assume that P and Q are complete lattices and that $F\colon P \to Q$ possesses an upper adjoint $F^\sharp\colon Q \to P$. Assume further that $G\colon P \to P$ and $H\colon Q \to Q$ are order-preserving. Then, if $\mu(G)$ and $\mu(H)$ exist,*

(i) $F \circ G \leqslant H \circ F \Longrightarrow F(\mu(G)) \leqslant \mu(H)$,

(ii) $F \circ G = H \circ F \Longrightarrow F(\mu(G)) = \mu(H)$.

(So the two-stage process of forming the fixpoint $\mu(G)$ and then applying F can be 'fused' into the single operation of forming the fixpoint $\mu(H)$. Contrariwise, the Fusion Rule may be regarded as decomposing $\mu(H)$. It is also known as the **Transfer Rule** since F can be seen as taking the least fixpoint $\mu(G)$ in P to the least fixpoint $\mu(H)$ in Q.)

Proof. For (i), assume $F \circ G \leqslant H \circ F$. We use the defining property, (Gal), of Galois connections from 7.23 and also (Gal1) from 7.26. We have

$$\begin{aligned}
&F(F^\sharp(\mu(H))) \leqslant \mu(H) && \text{(by (Gal1))} \\
&\Longrightarrow (H \circ F)(F^\sharp(\mu(H))) \leqslant \mu(H) && \text{(as } H \text{ is order-preserving)} \\
&\Longrightarrow (F \circ G)(F^\sharp(\mu(H))) \leqslant \mu(H) && \text{(by hypothesis)} \\
&\Longleftrightarrow G(F^\sharp(\mu(H))) \leqslant F^\sharp(\mu(H)) && \text{(by (Gal))} \\
&\Longrightarrow \mu(G) \leqslant F^\sharp(\mu(H)) && \text{(by 8.20 for } G) \\
&\Longleftrightarrow F(\mu(G)) \leqslant \mu(H) && \text{(by (Gal)).}
\end{aligned}$$

Now assume that $H \circ F = F \circ G$. To prove that $F(\mu(G)) \geqslant \mu(H)$ it suffices to show that $F(\mu(G))$ is a fixpoint of H. But $H(F(\mu(G))) = (H \circ F)(\mu(G)) = (F \circ G)(\mu(G)) = F(\mu(G))$, which, in combination with (i), proves (ii). □

8.31 The Exchange Rule. *Let P and Q be complete lattices and let $F, G\colon P \to Q$ and $H\colon Q \to P$ be order-preserving maps.*

(i) *Assume that F is the lower adjoint in a Galois connection $(F, F^\sharp)$ between P and Q. Then*

$$F \circ H \circ G \leqslant G \circ H \circ F \Longrightarrow \begin{cases} \mu(F \circ H) \leqslant \mu(G \circ H), \\ \mu(H \circ F) \leqslant \mu(H \circ G). \end{cases}$$

(ii) *Assume that both F and G are lower adjoints. Then*

$$F \circ H \circ G = G \circ H \circ F \implies \begin{cases} \mu(F \circ H) = \mu(G \circ H), \\ \mu(H \circ F) = \mu(H \circ G). \end{cases}$$

Proof. For (i) we have

$$\begin{aligned} F \circ H \circ G \leqslant G \circ H \circ F &\implies F(\mu(H \circ G)) \leqslant \mu(G \circ H) \\ &\qquad \text{(by the Fusion Rule)} \\ &\implies (F \circ H)(\mu(G \circ H)) \leqslant \mu(G \circ H) \\ &\qquad \text{(by the Rolling Rule)} \\ &\implies \mu(F \circ H) \leqslant \mu(G \circ H) \\ &\qquad \text{(by the Induction Rule).} \end{aligned}$$

Also

$$\begin{aligned} F \circ H \circ G \leqslant G \circ H \circ F &\implies \mu(F \circ H) \leqslant \mu(G \circ H) \\ &\qquad \text{(from above)} \\ &\implies H(\mu(F \circ H)) \leqslant H(\mu(G \circ H)) \\ &\qquad \text{(since } H \text{ is order-preserving)} \\ &\implies \mu(H \circ F) \leqslant \mu(H \circ G) \\ &\qquad \text{(by the Rolling Rule).} \end{aligned}$$

Part (ii) follows from part (i), as stated and also with F and G interchanged, and antisymmetry of $\leqslant$. □

We conclude this section by returning to continuous maps on CPOs. We have seen in 8.18 and 8.19 that, in the context of programs, mathematical induction can sometimes be used to establish what a least fixpoint outputs for a given input. There is, however, a formulation of induction which is often more convenient to apply. Like the Induction Rule given in 8.20, the Principle of Fixpoint Induction uses information about a map to glean information about the least fixpoint.

8.32 Principle of Fixpoint Induction for continuous maps. Let F be a continuous self-map on a CPO P and let $S \subseteq P$ satisfy:

(FI1) $\bot \in S$;

(FI2) $x \in S$ implies $F(x) \in S$;

(FI3) for any chain $x_0 \leqslant x_1 \leqslant x_2 \leqslant \cdots$ in S, we have $\bigsqcup_{n \geqslant 0} x_n \in S$.

Then $\mu(F) \in S$. This is derived by taking the formula for $\mu(F)$ given in Theorem 8.15 and using mathematical induction. To illustrate in a simple case, let F be a continuous map on a CPO P and let G and H be continuous maps on P such that $F \circ G = G \circ F$, $F \circ H = H \circ F$ and $G(\bot) = H(\bot)$. We can prove by induction that $G(F^n(\bot)) = H(F^n(\bot))$ for all n. Since G and H are continuous, and $\mu(F) = \bigsqcup_{n \geqslant 0} F^n(\bot)$ by Theorem 8.15, we deduce that $G(\mu(F)) = H(\mu(F))$. Alternatively we can argue as follows. Let $S := \{ x \in P \mid G(x) = H(x) \}$. Then (FI1) holds by hypothesis, (FI2) because F commutes with each of G and H, and (FI3) because G and H are continuous.

In general the subset S of P above is taken to be a set on which some property desired of a least fixpoint holds. It will be of the form $S = \{ x \in P \mid \mathbf{P}(x) \}$, where $\mathbf{P}$ is some predicate. When P is a CPO of partial maps, we might, for example, have $\mathbf{P}(f)$ as the assertion that f is not defined at some given point c (or, alternatively, that a computation to determine $f(c)$ does not terminate). As an example, consider $P = (\mathbb{N}^2 \multimap \mathbb{N})$ and

$$\overline{f}(j,k) = (F(f))(j,k) := \begin{cases} 1 & \text{if } j = k, \\ (k+1)(k+2)f(j,k+2) & \text{otherwise.} \end{cases}$$

We claim that the associated least fixpoint, $\mu(F)$, is undefined at (j,k) if $j < k$. To check this, take $S = \{ g \in P \mid j < k \implies g(j,k) = \bot \}$ and show that (FI1), (FI2) and (FI3) are satisfied. This technique of fixpoint induction can be adapted to a wide range of circumstances, but does have its limitations, since not all predicates yield a set S on which the Principle of Fixpoint Induction works.

Exercises

Exercises from the text. Complete the proof of Lemma 8.5. Prove directly that $(S \multimap S)$ is a CPO, for each set S (see the proof of 8.12). Prove all the claims made in 8.19(2). Give an example of an increasing map F on a CPO which has no least fixpoint.

8.1 Let P, Q and R be ordered sets. Is it true that $P \odot (Q \odot R) \cong (P \odot Q) \odot R$ when $\odot$ denotes the operation of forming (i) the separated sum, (ii) the coalesced sum?

8.2 Let P be a pre-CPO and let $A_i \subseteq P$ for all $i \in I$. Suppose that A_i is a directed subset of P for all $i \in I$ and that $\{A_i\}_{i \in I}$ is a directed family of subsets of P. Show that both $\bigcup_{i \in I} A_i$ and $\{\bigsqcup A_i\}_{i \in I}$ are directed subsets of P and that

$$\bigsqcup \left(\bigcup_{i \in I} A_i \right) = \bigsqcup_{i \in I} \left(\bigsqcup A_i \right).$$

8.3 Which of the following are (pre)-CPOs?

(i) $\{0,1,2\}^*$ (the set of finite strings of zeros, ones and twos with order defined as in 1.9).

(ii) The set of all finite strings of zeros and ones with order defined as in Exercise 1.6.

(iii) The family of all countable (including finite) subsets of $\mathbb{R}$, ordered by inclusion.

(iv) $\{1/n \mid n \in \mathbb{N}\} \cup \{0\}$.

(v) The chain $\mathbb{Q} \cap [0,1]$.

(vi) $\{1 - 1/2^n - 1/5^m \mid n, m \in \mathbb{N}\}$.

(In (iv)–(vi) the order is that induced from the chain $\mathbb{R}$.)

8.4 Let C be a closure operator on a set X. Show that C is algebraic if and only if it is continuous as a self-map on $\mathcal{P}(X)$.

8.5 Let P_1 and P_2 be CPOs. Let $D \subseteq P_1 \times P_2$ be a directed set. Define D_1 and D_2 by

$$D_1 := \{x_1 \in P_1 \mid (\exists x_2 \in P_2)\,(x_1, x_2) \in D\},$$
$$D_2 := \{x_2 \in P_2 \mid (\exists x_1 \in P_1)\,(x_1, x_2) \in D\}.$$

Show that D_1 and D_2 are directed and that $\bigsqcup D = (\bigsqcup D_1, \bigsqcup D_2)$. Deduce that $P_1 \times P_2$ is a CPO.

8.6 Let P_1 and P_2 be CPOs. Define the projections $\pi_1 \colon P_1 \times P_2 \to P_1$ and $\pi_2 \colon P_1 \times P_2 \to P_2$ by $\pi_i(x_1, x_2) = x_i \quad (i = 1, 2)$.

(i) Prove that π_1 and π_2 are continuous.

(ii) Prove that a map $\varphi \colon Q \to P_1 \times P_2$, where Q is another CPO, is continuous if and only if both $\pi_1 \circ \varphi$ and $\pi_2 \circ \varphi$ are continuous.

8.7 Let P_1, P_2 and Q be CPOs and let $\varphi \colon P_1 \times P_2 \to Q$ be a map.

(i) Define $\varphi^x \colon P_2 \to Q$ and $\varphi_y \colon P_1 \to Q$ by

$$\varphi^x(v) = \varphi(x, v)\,(v \in P_2) \quad \text{and} \quad \varphi_y(u) = \varphi(u, y)\,(u \in P_1).$$

Prove that φ is continuous if and only if φ^x and φ_y are continuous for all x, y (that is, φ is continuous if and only if it is continuous in each variable separately).

(ii) Let φ be continuous and define $F \colon P_1 \to [P_2 \to Q]$ by

$$(F(x))(y) = \varphi(x, y) \quad (x \in P_1,\ y \in P_2).$$

Prove that F is well defined and continuous and deduce that $[(P_1 \times P_2) \to Q] \cong [P_1 \to [P_2 \to Q]]$.

8.8 Let P be a CPO. Let $\mathcal{F}$ be the family of sets $U \in \mathcal{O}(P)$ such that $\bigsqcup D \in U$ whenever D is a directed subset of U.

(i) Identify $\mathcal{F}$ when P is

(a) $\mathbb{N} \oplus \mathbf{1}$, (b) $\overrightarrow{\mathbf{2}}_\perp$, (c) $(\mathbb{N} \times \mathbb{N}) \oplus \mathbf{1}$.

(ii) Show that $\mathcal{F}$ contains $\varnothing$ and P and is closed under arbitrary intersections and finite unions (and so is the family of closed sets for a topology $\mathcal{T}$ on P (the **Scott topology**)).

(iii) Let P and Q be CPOs and topologize each as above. Prove that a map $\varphi \colon P \to Q$ is topologically continuous if and only if it is continuous in the CPO sense. [Hint. Recall that φ is topologically continuous if and only if $\varphi^{-1}(V)$ is closed in P whenever V is closed in Q.]

8.9 Let P be a countable ordered set such that $\bigsqcup C$ $(= \bigvee C)$ exists for every non-empty chain C in P and let $D = \{x_0, x_1, x_2, \ldots\}$ be a directed subset of P. For each finite subset F of D let u_F be an upper bound for F in D. Define sets D_i, for $i \in \mathbb{N}_0$, as follows:

$$D_0 = \{x_0\}, \qquad D_{i+1} = D_i \cup \{y_{i+1}, u_{D_i \cup \{y_{i+1}\}}\},$$

where y_{i+1} is the element x_n in $D \setminus D_i$ with subscript n chosen as small as possible. Prove that

(i) for each i, the set D_i is directed and has at least i elements;

(ii) the sets D_i form a chain;

(iii) $\{\bigvee D_i\}_{i \geqslant 1}$ is a chain in P and its join is $\bigsqcup D$.

8.10 Let S be a set. Prove that the ordered set $(S \multimap S)$ (or equivalently $(S \to S_\perp)$) is order-isomorphic to the sub-CPO of $[S_\perp \to S_\perp]$ consisting of the strict maps, that is, those which map $\perp$ to $\perp$. (This result, combined with that in the next exercise, gives a simple way of showing that certain maps defined on CPOs of partial maps are order-preserving.)

8.11 Let S be any set and let $\varphi, \psi \colon S_\perp \to S_\perp$ be order-preserving. Show that $F \colon [S_\perp \to S_\perp] \to [S_\perp \to S_\perp]$, given by $F(f) := \psi \circ f \circ \varphi$ for all $f \in [S_\perp \to S_\perp]$, is well defined and order-preserving. [Hint. Use 8.8(2).]

8.12 For each partial map π on $\mathbb{N}$ define another such map by $F(\pi) = \overline{\pi}$, where

$$\overline{\pi}(k) = \begin{cases} 1 & \text{if } k = 1, \\ \pi(k-1) + k & \text{if } k > 1. \end{cases}$$

(Here $\overline{\pi}$ is to have the maximum domain allowed by its specification (see 8.17), so that $\operatorname{dom}\overline{\pi} = \{1\} \cup \{k+1 \mid k \in \operatorname{dom}\pi\}$.) Let π be a solution in $(\mathbb{N} \multimap \mathbb{N})$ to the fixpoint equation $F(\pi) = \pi$. Verify that π is total and that $\pi(k+1) = \pi(k) + k + 1$, for all $k \in \mathbb{N}$. Deduce, by induction, that $\pi(k) = \sum_{i=1}^{k} i$, for all $k \in \mathbb{N}$.

8.13 For each of the following recursive specifications, construct a map $F\colon P \to P$ whose fixpoints satisfy the specification, show that F is order-preserving and describe $F^0(\bot)$, $F^1(\bot)$, $F^2(\bot)$, $F^3(\bot)$, $F^n(\bot)$ and $\bigsqcup_{n \geqslant 0} F^n(\bot)$ via the graphs of the corresponding partial maps or, if you can, via non-recursively defined (partial) maps. Base your answer on our treatment of **fact**—see 8.13 and 8.18.

(i) $P = (\mathbb{N} \multimap \mathbb{N})$,

$$f(k) = \begin{cases} 1 & \text{if } k = 1, \\ (2k-1) + f(k-1) & \text{otherwise.} \end{cases}$$

(ii) $P = (\mathbb{N}_0 \multimap \mathbb{N}_0)$,

$$f(k) = \begin{cases} 1 & \text{if } k = 0, \\ f(k+1) & \text{otherwise.} \end{cases}$$

(iii) $P = (\mathbb{N}_0 \times \mathbb{N}_0 \multimap \mathbb{N}_0)$,

$$f(j,k) = \begin{cases} k & \text{if } j = 0, \\ 1 + f(j-1,k) & \text{otherwise.} \end{cases}$$

(iv) $P = (\mathbb{N}_0 \multimap \mathbb{N}_0)$,

$$f(k) = \begin{cases} k - 10 & \text{if } k > 100, \\ f(f(k+11)) & \text{otherwise.} \end{cases}$$

8.14 Find all fixpoints of $F\colon (\mathbb{N}_0 \multimap \mathbb{N}_0) \to (\mathbb{N}_0 \multimap \mathbb{N}_0)$, where

$$(F(f))(k) = \begin{cases} 1 & \text{if } k = 0, \\ f(k+1) & \text{otherwise.} \end{cases}$$

8.15 Consider maps in $P = (\mathbb{Z} \times \mathbb{Z} \multimap \mathbb{Z})$. Show that each of

$$(j,k) \mapsto j+1 \text{ and } (j,k) \mapsto \begin{cases} j+1 & \text{if } j \geqslant k, \\ k-1 & \text{otherwise,} \end{cases}$$

is a solution of

$$f(j,k) = \begin{cases} k+1 & \text{if } j = k, \\ f(j, f(j-1, k+1)) & \text{otherwise.} \end{cases}$$

What is the least solution in P?

8.16 Construct the least solutions to the following equations in the CPO $P = \wp(X)$:

(i) $S = S \cup T$ (T fixed);

(ii) (with $X = \mathbb{N}$) $S = S \cup \{1\} \cup \{n+2 \mid n \in S\}$;

(iii) (with $X = \mathbb{N}^2$)

$$S = \{(n,n) \mid n \in \mathbb{N}\} \cup \{(n,m+1) \mid (n,m) \in S\};$$

(iv) (with X a finite group) $S = T \cup S \cup S \cdot S$ (T fixed).

[Hint. Find a suitable order-preserving map F of which the required set is to be the least fixpoint, guess a formula for $F^n(\bot)$ and prove by induction that it works, and finally verify that $\bigsqcup F^n(\bot)$ is a fixpoint.] (See Theorem 8.15(i).)

8.17 Given $f\colon \mathbb{N}_0 \to (\mathbb{N}_0)_\bot$, let $\overline{f}$ be defined by

$$\overline{f}(k) = \begin{cases} 1 & \text{if } k = 0, \\ k f(k-1) & \text{if } k \geqslant 1 \text{ and } f(k-1) \neq \bot, \\ \bot & \text{otherwise.} \end{cases}$$

Define $F\colon (\mathbb{N}_0 \to (\mathbb{N}_0)_\bot) \to (\mathbb{N}_0 \to (\mathbb{N}_0)_\bot)$ by $F(f) = \overline{f}$.

(i) Show that F is order-preserving.

(ii) Prove that F is continuous. [Hint. Let $D = \{g_i\}_{i \in I}$ be a directed set in $(\mathbb{N}_0 \to (\mathbb{N}_0)_\bot)$, and let $g := \bigsqcup_{i \in I} g_i$. Check that $\overline{g}(k) = \bigsqcup_{i \in I} \overline{g}_i(k)$ for all $k \in \mathbb{N}_0$ by considering the cases

(a) $k = 0$,

(b) $g_i(k-1) = \bot$ for all $i \in I$,

(c) $g_i(k-1) \neq \bot$ for some $i \in I$.]

(This exercise establishes the continuity of the map of which the factorial function is the least fixpoint.)

8.18 Let P be a CPO and F an order-preserving self-map on P. Let

$$Q := \mathrm{post}(F) \cap \mathrm{fix}(F)^\ell$$
$$= \{ x \in P \mid x \leqslant F(x) \ \ \& \ \ (\forall y \in \mathrm{fix}(F))\, x \leqslant y \}.$$

(i) Show that Q is a sub-CPO of P.

(ii) Show that if Q has a maximal element, β, then the least fixpoint $\mu(F)$ of F exists and equals β.

(It is important to notice that in the definition of Q we are not claiming that F has a fixpoint. The second clause in the definition of Q is satisfied vacuously if $\mathrm{fix}(F) = \varnothing$. It serves to ensure that a fixpoint of F in Q is the least fixpoint of F.)

8.19 (For those who know about ordinals.) Let P be a CPO and $F\colon P \to P$ an order-preserving map. Let

$$F^0(\bot) = \bot,$$
$$F^{\beta+1}(\bot) = F(F^\beta(\bot)),$$
$$F^\alpha(\bot) = \bigsqcup \{ F^\beta(\bot) \mid \beta < \alpha \} \text{ if } \alpha \text{ is a limit ordinal.}$$

Prove that $F^\alpha(\bot)$ is well defined for all ordinals α by showing (by transfinite induction) that, if $\alpha \leqslant \beta$ and $F^\beta(\bot)$ is defined, then $F^\alpha(\bot) \leqslant F^\beta(\bot)$. Argue by contradiction to show (with the aid of Hartogs' Theorem) that, for some ordinal, α, the point $F^\alpha(\bot)$ is a fixpoint of F.

8.20 Prove CPO Fixpoint Theorem III (8.23) by following the steps below. (Readers used to working with ordinals will find that the verifications required are very reminiscent of arguments involving ordinals directly; cf. Exercise 8.19.)

Let P_0 be the smallest F-invariant sub-CPO of P. We shall work entirely in P_0 and shall show that F has a fixpoint in P_0. An element $x \in P_0$ is called a **roof** (of F) if $F(y) \leqslant x$ for all $y < x$.

(i) Show that if $x \in P_0$ is a roof and $y \in P_0$, then either $y \leqslant x$ or $F(x) \leqslant y$. [Hint. Show that Z_x is an F-invariant sub-CPO of P, where $Z_x := \{ y \in P_0 \mid y \leqslant x \text{ or } F(x) \leqslant y \}$. Conclude that $Z_x = P_0$.]

(ii) Show that each element of P_0 is a roof. [Hint. Show that $Z := \{ x \in P_0 \mid x \text{ is a roof} \}$ is an F-invariant sub-CPO of P_0.]

(iii) Show that P_0 is a chain.

(iv) Explain why P_0 has a top element, $\top_{P_0}$, and show that $\top_{P_0}$ is a minimal fixpoint of F.

8.21 Deduce CPO Fixpoint Theorem II from CPO Fixpoint Theorem III by applying the latter to the set Q defined in Exercise 8.18.

8.22 Let P be an ordered set and Q a complete lattice. Given a map $f\colon P \to Q$, define $\overline{f}\colon P \to Q$ by

$$\overline{f}(x) = \bigvee \{ f(y) \mid y \leqslant x \}.$$

Show that $\overline{f}$ is order-preserving and that $f = \overline{f}$ if and only if f is order-preserving. Show further that F defined by $F\colon f \mapsto \overline{f}$ is an order-preserving map from Q^P to $Q^{(P)} \subseteq Q^P$ whose set of fixpoints is exactly $Q^{(P)}$.

8.23 Let P be a complete lattice and $F\colon P \to P$ be order-preserving. Let X be a subset of $\mathrm{fix}(F)$. Define

$$Y = \{ y \in P \mid (\forall x \in X)\, x \leqslant F(y) \leqslant y \}$$

and let $\alpha = \bigwedge_P Y$. Prove that $\alpha = \bigvee_{\mathrm{fix}(F)} X$. Deduce that $\mathrm{fix}(F)$ is a complete lattice.

8.24 Let L be a complete lattice and define $F : \wp(L) \to \wp(L)$ by $A \mapsto {\downarrow}\bigvee A$. Show that F is order-preserving and that $\mathrm{fix}(F) \cong L$. (This is the long proof that every complete lattice is, up to isomorphism, a lattice of fixpoints. Find the short proof.)

8.25 Let P be a CPO and let $F, G\colon P \to P$ be continuous self-maps on P. Assume that $F \leqslant G$ (pointwise).

(i) Use CPO Fixpoint Theorem I to prove that $\mu(F) \leqslant \mu(G)$. (Compare with 8.28.)

(ii) Now assume that $F(x) < G(x)$ for all $x \in P$. Prove that $\mu(F) < \mu(G)$.

(iii) Prove that $\mu(F) \lneqq \mu(F \circ G) < \mu(G)$ and that the first inequality is strict if F is an injective map.

8.26 Let P be a CPO. Let $\mu\colon [P \to P] \to P$ be the fixpoint operator which assigns to each continuous map $f\colon P \to P$ its least fixpoint $\mu(f)$ and, for each $n \in \mathbb{N}$, let μ_n be defined by $\mu_n(f) = f^n(\bot)$.

(i) Show, by induction, that μ_n is continuous for each n.

(ii) Let $Q = ([P \to P] \to P)$, the CPO of all maps from $[P \to P]$ to P. Show that $\{\mu_n\}_{n \geqslant 1}$ is a chain in Q.

(iii) Show that $\bigsqcup_{n \geqslant 1} \mu_n = \mu$ in Q. Deduce that μ is continuous. (This strengthens the monotonicity assertion in Exercise 8.25(i).)

8.27 Let P be a CPO and let F and G be continuous self-maps on P such that $F \circ G = G \circ F$.

(i) Show that $\{ F^m(G^n(\bot)) \mid n, m \geqslant 0 \}$ forms a directed set.

(ii) Show that $\mu(F \circ G)$ is a fixpoint of both F and G.

(iii) Show that $\mu(F \circ G) = \mu(F)$ if and only if $\mu(G) \leqslant \mu(F)$.

8.28 Let F be an order-preserving self-map on a complete lattice P. Prove that $\mu(F \circ F) = \mu(F)$ (the **Square Rule**).

8.29 Let L be a complete lattice and F, G be order-preserving self-maps on L. Prove that if F and G have a common fixpoint then they have a least common fixpoint, given by

$$\bigwedge \{ x \in L \mid F(x) \leqslant x \ \& \ G(x) \leqslant x \}.$$

Prove further that, if $F \circ G = G \circ F$, then the least common fixpoint of F and G is $\mu(F \circ G)$.

8.30 Let P be a complete lattice and assume that $F\colon P \times P \to P$ is order-preserving. Consider the induced 'diagonal' self-map on P given by $\Delta(x) := F(x,x)$. Define maps $F_y\colon u \mapsto F(u,y)$ and $F^x\colon v \mapsto F(x,v)$, then define $\ell\colon x \mapsto \mu(F^x)$ and $r\colon y \mapsto \mu(F_y)$. Show that all five maps are order-preserving and then prove that $\mu(\Delta) = \mu(\ell) = \mu(r)$.

(This is the **Diagonal Rule**. If $F(x,x)$ is a complex expression involving several occurrences of the variable x, the Diagonal Rule may be applied to eliminate the occurrences of x one at a time.)

8.31 Let P be a complete lattice and assume that F and G are order-preserving maps from $P \times P$ to P. We seek the least fixpoint of $H\colon P \times P \to P \times P$ defined by $H(x,y) = (F(x,y), G(x,y))$. Define maps F_y and G^x as in the previous exercise and then define $M\colon x \mapsto F(x, \mu(G^x))$ and $N\colon y \mapsto G(\mu(F_y), y)$. Show that all five maps are order-preserving, then prove that $\mu(H) = (\mu(M), \mu(N))$. [Hint. Use Exercise 8.30.]

(This exercise illustrates the process known as **mutual recursion**. A definition by mutual recursion can be viewed as a single fixpoint equation in which the variable is a vector, (x,y) in our case. On the other hand, it may be viewed as a collection of equations which may be solved on an individual basis. The exercise asserts that the two approaches are compatible.)

9

Domains and Information Systems

This chapter contains material at the interface between order theory and computer science. It discusses domains – those CPOs in which each element is the supremum of its 'finite' approximations – paralleling the discussion given in Chapter 7 of the class of algebraic lattices lying within the class of complete lattices. These domains provide a setting for denotational semantics, in a way we outline in 9.33. An alternative approach to domains, via information systems, is also presented. The final section of the chapter returns to fixpoint theory, and shows how the theorems in Chapter 8 can be applied in the solution of recursive equations, and in particular domain equations.

Domains for computing

This section brings together the notions of directed joins and of finite approximations.

9.1 Definitions. Let S be a non-empty subset of an ordered set P. Then S is said to be **consistent** if, for every finite subset F of S, there exists $z \in P$ such that $z \in F^u$.

9.2 Remarks. Non-consistency arises only in ordered sets without $\top$. A directed set is, of course, consistent. The difference between the two notions is in the location of upper bounds: for D to be directed we require every finite subset F of D to have some upper bound which is a member of D, but for consistency an upper bound in P suffices.

9.3 Complete semilattices. We introduced CPOs as a class of ordered sets in which suitable joins exist, but meets play no role, yet Lemma 2.30 shows that meets may sneak in by the back door. The situation is clarified by the following lemma, whose proof is left as an exercise. A CPO satisfying the equivalent conditions of the lemma is called a **complete semilattice**. Adjoining a top to such a CPO creates a complete lattice. It is not true that removal of the top from an arbitrary complete lattice leaves a complete semilattice; consider $\mathbb{N} \oplus \mathbf{1}$.

9.4 Lemma. *Let P be a CPO. Then the following are equivalent:*

(i) *P is* **consistently complete**, *that is, $\bigvee S$ exists for every consistent set S in P;*

(ii) *$\bigvee S$ exists whenever $S^u \neq \varnothing$;*

(iii) *$\bigwedge S$ exists whenever $S \neq \varnothing$;*

(iv) *$P \oplus \mathbf{1}$ is a complete lattice.*

We referred earlier to elements with a connotation of 'finiteness'. We now give the promised discussion of this concept in the context of CPOs.

9.5 Definition. Let P be a CPO and let $k \in P$. Then k is called **finite** (in P) if, for every directed set D in P,

$$k \leqslant \bigsqcup D \Longrightarrow k \leqslant d \text{ for some } d \in D.$$

The set of finite elements of P is denoted $F(P)$. In case P is a complete lattice this is just the definition we gave in 7.15.

9.6 Examples. In addition to the finite elements in complete lattices given in 7.17, we note the following examples in CPOs which are not complete lattices. In the ordered set Σ^{**} of all binary strings the set $F(\Sigma^{**})$ is Σ^* (finite strings). In $(X \multimap X)$, the set of partial maps on X, the finite elements are the maps with finite domain. In an ordered set P with (ACC) we have $F(P) = P$.

9.7 Domains. Our information-content examples mostly lack a top but are in all other respects very like the algebraic lattices we studied in Chapter 7. If L is a complete semilattice (as defined in 9.3) then L is called an **algebraic semilattice** or a **domain** if, for all $a \in L$,

$$a = \bigsqcup \{ k \in F(L) \mid k \leqslant a \}.$$

(The join here is taken over a directed set, by 9.4 and 7.16.) The term domain is used widely in computer science, but not consistently (and certainly not directedly!); its meaning ranges from CPO through algebraic semilattice to algebraic semilattice with at most countably many finite elements. This last property is important in connection with computability questions which we shall not consider.

With the aid of Lemma 9.4 it is easily seen that adjoining a top to a domain gives an algebraic lattice and adjoining a top to an algebraic $\bigcap$-structure gives a topped algebraic $\bigcap$-structure. This observation yields a topless counterpart to Theorem 7.20, which in one direction provides a concrete realization of domains and in the other an order-theoretic characterization of algebraic $\bigcap$-structures.

9.8 Theorem.

(i) *Let $\mathfrak{L}$ be an algebraic $\bigcap$-structure. Then $\mathfrak{L}$ is a domain.*

(ii) *Let L be a domain and define $D_a := \{ k \in F(L) \mid k \leqslant a \}$ for each $a \in L$. Then $\mathfrak{L} := \{ D_a \mid a \in L \}$ is an algebraic $\bigcap$-structure isomorphic to L.*

9.9 Examples. Any algebraic lattice is a domain. In addition the following are domains:

(i) any flat CPO, and in particular $\mathbb{N}_\perp$;

(ii) $(S \multimap S)$, for any $S \subseteq \mathbb{R}$;

(iii) Σ^{**} (all binary strings).

9.10 Continuous maps between domains. Let P and Q be domains. A map $\varphi\colon P \to Q$ is continuous if and only if it is determined by finite approximations, that is, by its effect on the finite elements of P. Where domains are used as computational models, a finite element may be interpreted as an object conveying a finite amount of information. In this case the continuity condition asserts that to obtain a finite amount of information about $\varphi(x)$ it is only necessary to input a finite amount of information about x; this is exactly the import of (iii) in 9.11. Compare this with the requirement that φ be order-preserving, which may be informally stated as 'more information in implies more information out'.

9.11 Proposition. *Let P and Q be domains and $\varphi\colon P \to Q$ be order-preserving. Then the following are equivalent:*

(i) *φ is continuous;*

(ii) *$\varphi(x) = \bigsqcup \{ \varphi(k) \mid k \in F(P) \text{ and } k \leqslant x \}$ for each $x \in P$;*

(iii) *$D_{\varphi(x)} \subseteq {\downarrow}\varphi(D_x)$ for all $x \in P$.*

Further, $[P \to Q]$ (the continuous maps from P to Q) is isomorphic to $\langle F(P) \to Q \rangle$ (the order-preserving maps from $F(P)$ to Q).

Proof. We have (i) $\Rightarrow$ (ii) as $D_x := \{k \in F(P) \mid k \leqslant x\}$ is directed. Assume (ii) holds and let $k' \in D_{\varphi(x)}$. Then $k' \leqslant \bigsqcup \{ \varphi(k) \mid k \in D_x \}$. Directedness implies $k' \leqslant \varphi(k)$ for some $k \in D_x$. Thus (iii) holds.

To prove (iii) $\Rightarrow$ (i), let $D \subseteq P$ be directed and take $x := \bigsqcup D$. Let $k' \in D_{\varphi(x)}$. Thus, by (iii), $k' \leqslant \varphi(k)$ for some $k \in D_x$. Since $k \in F(P)$ and $k \leqslant x = \bigsqcup D$, we have $k \leqslant d$ for some $d \in D$. Thus $k' \leqslant \varphi(k) \leqslant \varphi(d) \leqslant \bigsqcup \varphi(D)$. Hence $\bigsqcup \varphi(D)$ is an upper bound of $D_{\varphi(x)}$ and consequently $\varphi(x) = \bigsqcup D_{\varphi(x)} \leqslant \bigsqcup \varphi(D)$.

Finally we note that the restriction map $\varphi \mapsto \varphi{\restriction}_{F(P)}$ sets up an order-isomorphism from $[P \to Q]$ to $\langle F(P) \to Q \rangle$; it is onto because,

for any $\psi \in (F(P) \to Q)$, the map $x \mapsto \bigsqcup\{\psi(k) \mid k \in D_x\}$ is in $[P \to Q]$ and extends ψ. The details are left to the reader. □

9.12 Drawing parallels. We conclude this section by giving a diagram to show the interrelation between the various structures introduced in this chapter and in Chapter 7. A double arrow indicates a bijective correspondence and a single arrow an inclusion. The numbers refer to the subsections in which proofs of equivalence are given. The boxes indicate the main areas of application. The structures with 'algebraic' appended to their names have 'plenty of finite elements' (finite here having its technical meaning) and the various types of $\bigcap$–structures provide concrete examples or representations of the more abstract structures in the left-hand part of the diagram. Vertically we have a hierarchy of progressively weaker conditions on join and meet.

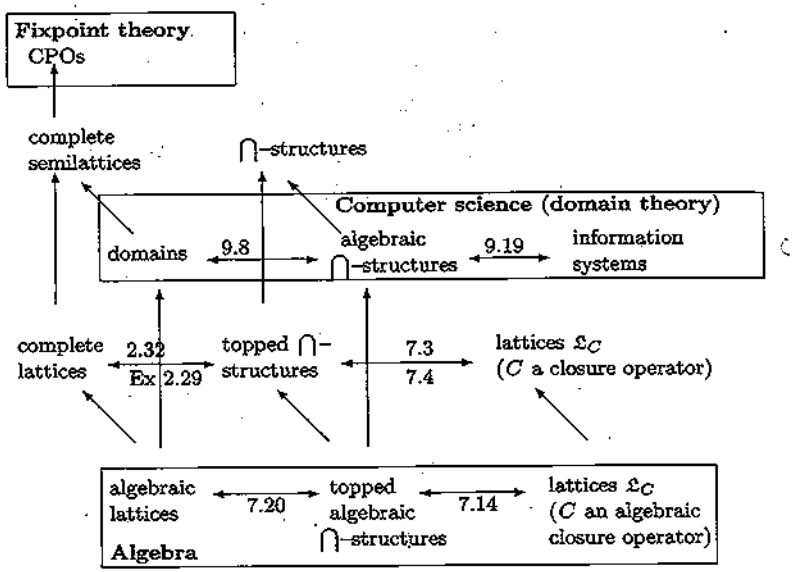


Figure 9.1

Domains re-modelled: information systems

We already have two ways of presenting domains: abstractly, as algebraic semilattices, and, more concretely, as algebraic $\bigcap$–structures. A third way to arrive at domains is through information systems. This approach has a particularly intuitive appeal, because it capitalizes on the finiteness and information-content ideas which pervade the theory of computation.

In this section we give a brief introduction to information systems and relate them to domains and algebraic $\bigcap$–structures. We contend that a judicious combination of the techniques and constructions derived from these alternative viewpoints gives an economical and intuitive approach to domain theory.

The starting point for the notion of an information system is the idea of identifying an object with a set of propositions true of it and adequate to define it. These propositions are to be thought of as tokens, each bearing a finite amount of information. Thus, if the objects to be described are the maps from $\mathbb{N}$ to $\mathbb{N}$, a suitable set of tokens would be $\mathbb{N} \times \mathbb{N}$, with the single token (m, n) true of f if and only if $f(m) = n$.

In the theory developed below, an information system has three constituents: a set A of tokens, a family Con of finite subsets of A (representing gobbets of consistent information) and a relation $\vdash$ of entailment (identifying implied, or superfluous, information). From an information system we build an $\bigcap$–structure $\mathfrak{L}$ on A so that each member of $\mathfrak{L}$ is a set of tokens whose finite subsets lie in Con and which contains all entailed tokens. The members of $\mathfrak{L}$, which are known as the elements of the information system, thus serve to represent the objects determined by consistent information. Further, $\mathfrak{L}$ is a domain and, conversely, every domain is associated with an information system in a canonical way.

9.13 Definition. An **information system** is a triple $\mathbf{A} = \langle A, Con, \vdash \rangle$ consisting of

(i) a set A of **tokens**;

(ii) a non-empty set Con of finite subsets of A which satisfy

(IS1) $Y \in Con$ and $Z \subseteq Y$ implies $Z \in Con$,

(IS2) $a \in A$ implies $\{a\} \in Con$;

(iii) $\vdash$ is a relation (**entailment**) between members of Con and members of A (formally $\vdash$ is a subset of $Con \times A$) satisfying

(IS3) $Y \cup \{a\} \in Con$ whenever $Y \in Con, a \in A$ and $Y \vdash a$,

(IS4) $Y \in Con$ and $a \in Y$ implies $Y \vdash a$,

(IS5) if $Y, Z \in Con$ and $a \in A$ satisfy $Y \vdash b$, for all $b \in Z$, and $Z \vdash a$, then $Y \vdash a$.

Read $Y \vdash a$ as 'Y entails a' or 'a is deducible from Y'. An arbitrary subset $X \subseteq A$ is said to be **consistent** if every finite subset of X is in Con. We adopt the notation $Y \Subset A$ to mean that Y is a finite (possibly empty) subset of A.

9.14 Remarks. Axioms (IS1), (IS3) and (IS4) formalize commonsense features of consistency and entailment and (IS2) just says that every token contributes some information. The most mysterious axiom, (IS5), is a transitivity condition; it may be interpreted as saying that if from Y we can deduce enough information, Z, to deduce a, then we can deduce a from Y.

From (IS2) and (IS1) we deduce that $\varnothing$ (representing 'no information') is in Con. Further, (IS1) implies that Con is precisely the set of all finite consistent subsets of A.

The following rules are frequently used in proofs:

(a) if $Y \in Con$, $Z \subseteq Y$ and $Z \vdash a$, then $Y \vdash a$;

(b) if $Y \in Con$, Z is finite and $Y \vdash a$ for every $a \in Z$, then $Z \in Con$.

The first is entirely elementary. To prove the second, use induction on $|Z|$ to show $Y \cup Z \in Con$ and then appeal to (IS1).

9.15 Examples. Each of the triples $\mathbf{A} = \langle A, Con, \vdash \rangle$ in (1)–(5) below defines an information system, with the specified consistent sets. These examples show that the structure of an information system may be borne mainly by Con, by $\vdash$, or by the interaction between Con and $\vdash$.

(1) Take $A = \mathbb{N} \times \mathbb{N}$ and let

$\varnothing \in Con$ and $\{(m_1, n_1), \ldots, (m_k, n_k)\} \in Con$ if and only if $(m_i = m_j \Rightarrow n_i = n_j)$,

$Y \vdash (n, m)$ if and only if $(n, m) \in Y$.

The consistent sets are just the (graphs of) partial maps on $\mathbb{N}$.

(2) Let V be a vector space. Take $A = V$, Con to be all finite subsets of A and put $Y \vdash v$ if and only if v belongs to the subspace spanned by Y. All subsets of V are consistent.

(3) Take $A = \mathbb{N}$, let Con be all finite subsets of $\mathbb{N}$ and define

$\{n_1, \ldots, n_k\} \vdash n$ if and only if $n \leqslant n_i$ for some i.

All subsets of $\mathbb{N}$ are consistent.

(4) Take $A = \Sigma^*$ and let $\leqslant$ be the order on Σ^* defined in 1.9. Take

$Con = \{ Y \Subset \Sigma^* \mid \sigma, \tau \in Y \Rightarrow \sigma \leqslant \tau \text{ or } \tau \leqslant \sigma \}$,

$Y \vdash \sigma$ if and only if $\sigma \leqslant \tau$ for some $\tau \in Y$.

A subset of A is consistent if and only if it is a subset of $\downarrow\sigma$ for some $\sigma \in \Sigma^{**}$.

(5) Recalling 1.11, take $A = \{ [\underline{x}, \overline{x}] \mid -\infty \leqslant \underline{x} \leqslant \overline{x} \leqslant \infty \}$ and define

$\varnothing \in Con$ and $\{I_1, \ldots, I_k\} \in Con$ if and only if $I_1 \cap \cdots \cap I_k \neq \varnothing$.

$\{I_1, \ldots, I_k\} \vdash [\underline{x}, \overline{x}]$ if and only if $I_1 \cap \cdots \cap I_k \subseteq [\underline{x}, \overline{x}]$.

The non-empty consistent sets are the subsets of A with non-empty intersection.

(6) We may define a family of information systems $\perp_n$, for $n \geqslant 0$, as follows. Let $\perp_0$ be the unique information system with $\varnothing$ as its set of tokens. This has $\varnothing$ as its only consistent set and $\vdash$ the empty relation. Now define inductively

$$a_0 = (0,0), a_1 = \{a_0, (1, a_0)\}, \ldots, a_{n+1} = \{a_n, (1, a_n)\}.$$

This creates a chain $a_0, a_1, \ldots$ of sets satisfying $a_0 \in a_1$, $a_1 \in a_2$ and so on; the process will be familiar to those who have seen the formal construction of the natural numbers. (This rather artificial notation is chosen to fit in with that used in 9.26(i). The pairing with 1 ensures that a new element is added at each stage.) For $n \geqslant 1$, take $\perp_n = \langle A_n, Con_n, \vdash_n \rangle$, where

$A_n = \{a_0, \ldots, a_{n-1}\}$,

Con_n consists of all subsets of A_n,

$Y \vdash a_i$ if and only if there exists $j \geqslant i$ such that $a_j \in Y$.

9.16 Elements. Each of the information systems above has a strong connection with some ordered set we have previously encountered. To make this precise, we need more definitions. Let $\mathbf{A} = \langle A, Con, \vdash \rangle$ be an information system. A set E of tokens is called an **element** of $\mathbf{A}$ if E is consistent and $\vdash$–closed, in the sense that $Y \in Con$, $Y \Subset E$ and $Y \vdash a$ imply $a \in E$. The set of elements of $\mathbf{A}$ is denoted $|\mathbf{A}|$; in 9.18 we show that $|\mathbf{A}|$ is an algebraic $\bigcap$–structure.

For any consistent set X we define

$$\overline{X} := \{ a \in A \mid (\exists Y \Subset X)\, Y \vdash a \};$$

this may be interpreted as the set of tokens deducible from X. When $X \in Con$, we have $\overline{X} = \{ a \in A \mid X \vdash a \}$. Lemma 9.17 shows that $\overline{X}$ is an element whenever X is consistent, and that every element is of this form. This lemma also characterizes elements in a way which reveals that, on the set of consistent subsets of A, the map $X \mapsto \overline{X}$ behaves very like an algebraic closure operator. Before stating the lemma we find the elements of the information systems in 9.15.

(1) $|\mathbf{A}| = (\mathbb{N} \multimap \mathbb{N})$, since all consistent sets are $\vdash$–closed.

(2) $|\mathbf{A}| = \operatorname{Sub} V$. This example typifies the way algebraic lattices come from information systems.

(3) $|\mathbf{A}| = \{ \downarrow m \mid m \in \mathbb{N} \} \cup \{\mathbb{N}\}$.

(4) There is a one-to-one correspondence between the set Σ^{**} (finite or infinite binary strings) and the elements, under which a string is associated with its set of finite initial substrings. Under the inclusion order on $|\mathbf{A}|$, the elements finite in the CPO sense correspond to the finite strings and the maximal elements to the infinite strings.

(5) This example is rather similar to the last, with each element associated to an interval in $\mathbb{R} \cup \{-\infty, \infty\}$. Here the maximal elements correspond to the real numbers together with $-\infty$ and ∞.

(6) The elements of $\bot_n$ are $\varnothing$ and the sets $\{a_0, \ldots, a_k\}$ for $0 \leqslant k < n$. Ordered by inclusion, the elements form an n-element chain.

We prove one implication in the proof of the following lemma to illustrate how the rules in 9.14 and the axioms are employed and set the rest of the proof as an exercise.

9.17 Lemma. *Let $\mathbf{A} = \langle A, Con, \vdash \rangle$ be an information system and let $E \subseteq A$. Then the following are equivalent:*

(i) *E is consistent and $\vdash$–closed (that is, $E \in |\mathbf{A}|$);*

(ii) *$\{\overline{Y} \mid Y \in Con$ and $Y \Subset E\}$ is directed and*

$$E = \bigcup \{\overline{Y} \mid Y \in Con \text{ and } Y \Subset E\};$$

(iii) *$E = \overline{X}$ for some consistent set X.*

Proof of (iii) $\Rightarrow$ (i). Let $Z := \{x_1, \ldots, x_k\} \Subset E$. For each i, there exists $Y_i \Subset X$ with $Y_i \in Con$ and $Y_i \vdash x_i$. Then $Y := Y_1 \cup \cdots \cup Y_k \Subset X$ and so $Y \in Con$, since X is consistent. By Rule (a) in 9.14, $Y \vdash x_i$ for each i. By Rule (b) in 9.14, $Z \in Con$. Hence E is consistent. To show E is $\vdash$–closed, assume $Z \vdash a$. The set Y above is such that $Y \vdash b$ for each $b \in Z$. By (IS5), $Y \vdash a$, so $a \in \overline{X} = E$. $\square$

9.18 Information systems and algebraic $\bigcap$–structures. Take an information system $\mathbf{A} = \langle A, Con, \vdash \rangle$. We now prove our earlier claim that $|\mathbf{A}|$ is an algebraic $\bigcap$–structure, in other words a non-empty family of sets closed under intersections and directed unions of non-empty subfamilies. Since $|\mathbf{A}|$ contains $\overline{\varnothing}$, it is non-empty. It is routine to show that if $\{E_i\}_{i \in I}$ is a non-empty subfamily of $|\mathbf{A}|$ then $\bigcap_{i \in I} E_i$ is consistent and $\vdash$–closed, and so is in $|\mathbf{A}|$. Finally, assume $\mathcal{D} = \{E_i\}_{i \in I}$ is a directed set in $|\mathbf{A}|$ and let $E = \bigcup_{i \in I} E_i$. Take $Y \Subset E$. Because $\mathcal{D}$ is directed, $Y \Subset E_i$ for some i. Since E_i is consistent, we have $Y \in Con$. Assume also $Y \vdash a$. Then $a \in E_i$ since E_i is $\vdash$–closed, so $a \in E$. Therefore E is consistent and $\vdash$–closed. This completes the proof of the claim. It is an easy exercise to prove in addition that the finite elements of $|\mathbf{A}|$ are exactly the sets $\overline{Y}$ where $Y \in Con$.

In the other direction, take an algebraic $\bigcap$–structure $\mathfrak{L}$ and let $\mathbf{IS}(\mathfrak{L})$ be the information system $\langle A, Con, \vdash \rangle$ defined as follows:

(i) $A := \bigcup \mathfrak{L}$;

(ii) $Con := \{Y \mid (\exists U \in \mathfrak{L})\, Y \Subset U\}$;

(iii) $Y \vdash a$ if and only if $a \in \bigcap \{U \in \mathfrak{L} \mid Y \Subset U\}$.

We outline the proof of the following theorem, leaving the reader to supply the details.

9.19 Theorem. *The maps $\mathbf{A} \mapsto |\mathbf{A}|$ and $\mathfrak{L} \mapsto \mathbf{IS}(\mathfrak{L})$ are mutually inverse and set up a bijective correspondence between the class of information systems and the class of algebraic $\bigcap$–structures.*

Proof. Given $\mathbf{A} = \langle A, Con, \vdash \rangle$, we claim that $\mathbf{A} = \mathbf{IS}(|\mathbf{A}|)$. We have:

(i) $A = \bigcup |\mathbf{A}|$ (by (IS2));

(ii) if $Y \Subset A$, then $Y \in Con \Leftrightarrow (\exists E \in |\mathbf{A}|)\, Y \Subset E$ (for the forward implication note $\overline{Y} \in |\mathbf{A}|$ and for the reverse recall that any $E \in |\mathbf{A}|$ is consistent);

(iii) if $Y \in Con$ and $a \in A$, then $Y \vdash a \Leftrightarrow a \in \bigcap \{E \in |\mathbf{A}| \mid Y \Subset E\}$ (for the forward implication recall that any $E \in |\mathbf{A}|$ is $\vdash$–closed and for the reverse use the fact that $Y \subseteq \overline{Y} \in |\mathbf{A}|$).

Let $\mathfrak{L}$ be an algebraic $\bigcap$–structure. The formulae in 9.18 and 9.17 imply that, for an element E of $\mathbf{IS}(\mathfrak{L})$,

$$E = \bigcup \{\bigcap \{U \in \mathfrak{L} \mid U \supseteq Y\} \mid Y \Subset E\},$$

with the union taken over a directed set. Since $\mathfrak{L}$ is algebraic, we have $|\mathbf{IS}(\mathfrak{L})| \subseteq \mathfrak{L}$. Conversely, the definitions of consistency and entailment in $\mathbf{IS}(\mathfrak{L})$ imply that $\mathfrak{L} \subseteq |\mathbf{IS}(\mathfrak{L})|$. $\square$

9.20 Information systems and domains. By combining 9.8 and 9.19 we obtain a bijective correspondence between information systems and domains. Given a domain D, the associated information system $\mathbf{IS}(D)$ has $F(D)$ (the finite elements of D) as its tokens, the finite sets consistent in the sense of 7.7 as the members of Con and $Y \vdash k$ if and only if $k \leqslant \bigvee Y$. Further $|\mathbf{IS}(D)|$ is order-isomorphic to D. The isomorphism is given by $U \mapsto \bigvee U$ and its inverse by $x \mapsto \{k \in F(D) \mid k \leqslant x\}$.

One cautionary remark needs to be made. Let D be the domain of elements of an information system $\mathbf{A}$. Then $|\mathbf{IS}(D)|$ is order-isomorphic to D but in general the set of tokens of $\mathbf{A}$ is quite different from the set of tokens of $\mathbf{IS}(D)$; indeed these sets of tokens may be of different cardinalities. Example 9.15(1) illustrates this well. The original set of

tokens is $\mathbb{N} \times \mathbb{N}$, the domain D of elements is $(\mathbb{N} \multimap \mathbb{N})$ and the token set of $\mathbf{IS}(D)$ is $\{\sigma \in (\mathbb{N} \multimap \mathbb{N}) \mid \operatorname{dom}\sigma \text{ is finite}\}$.

9.21 Technical remark. Take an algebraic $\bigcap$–structure $\mathfrak{L}$ on a set X. As Theorem 9.19 tells us, $|\mathbf{IS}(\mathfrak{L})|$ is the same family of sets as $\mathfrak{L}$. It is an $\bigcap$–structure on $\bigcup\mathfrak{L}$, which may be a proper subset of X. We define $\mathfrak{L}$ to be **full** if $\bigcup\mathfrak{L} = X$ or, in other words, if every point of the base set X belongs to a member of $\mathfrak{L}$. By construction $|\mathbf{A}|$ is full for any information system $\mathbf{A}$. Henceforth we work always with full $\bigcap$–structures and adopt the notation $(\mathfrak{L}, A)$ to indicate that $\mathfrak{L}$ is an $\bigcap$–structure with A as its base set.

9.22 Substructures and subsystems. Let $(\mathfrak{L}, A)$ and $(\mathfrak{K}, B)$ be algebraic $\bigcap$–structures. Then $\mathfrak{L}$ is called a **substructure** of $\mathfrak{K}$, and we write $\mathfrak{L} \sqsubseteq \mathfrak{K}$, if

(i) $A \subseteq B$,

(ii) $\mathfrak{L} = \{U \cap A \mid U \in \mathfrak{K}\}$.

Note that we use the term substructure only in relation to $\bigcap$–structures which are algebraic. If $\mathfrak{L} \sqsubseteq \mathfrak{K}$ and $A = B$, then $\mathfrak{L} = \mathfrak{K}$. This useful property implies in particular that $\sqsubseteq$ is antisymmetric. It is obviously also reflexive and transitive. Therefore $\sqsubseteq$ defines a partial order on the set of substructures of any algebraic $\bigcap$–structure $\mathfrak{J}$. (Those who know a little set theory will appreciate that staying within some fixed $\mathfrak{J}$ is a device to ensure that we work with a set and not a proper class. This restriction causes no difficulties in practice, since it is usually possible to assume that all the $\bigcap$–structures involved in a given problem have their base sets lying in some fixed set X, so we may take $\mathfrak{J} = \wp(X)$.)

The ordering $\sqsubseteq$ of $\bigcap$–structures (or its information system equivalent, given below) is the key to solving domain equations (see 9.32 and 9.36). Exercise 9.15 elucidates how $\sqsubseteq$ works. It shows in particular that, if $\mathfrak{L} \sqsubseteq \mathfrak{K}$, then the map $S \mapsto \bigcap\{T \in \mathfrak{K} \mid S \subseteq T\}$ is a continuous order-embedding of $\mathfrak{L}$ into $\mathfrak{K}$.

Let $\mathbf{A} = \langle A, Con_A, \vdash_A\rangle$ and $\mathbf{B} = \langle B, Con_B, \vdash_B\rangle$ be information systems. Then (an exercise) $|\mathbf{A}| \sqsubseteq |\mathbf{B}|$ if and only if

(i) $A \subseteq B$,

(ii) $Y \in Con_A \iff (Y \subseteq A \;\&\; Y \in Con_B)$,

(iii) $Y \vdash_A a \iff (Y \subseteq A \;\&\; a \in A \;\&\; Y \vdash_B a)$.

When these three conditions hold for $\mathbf{A}$ and $\mathbf{B}$, we say $\mathbf{A}$ is a **subsystem** of $\mathbf{B}$ and write $\mathbf{A} \trianglelefteq \mathbf{B}$. The relation $\mathbf{A} \trianglelefteq \mathbf{B}$ essentially says that $\mathbf{A}$ is less rich in information than $\mathbf{B}$, with each token of information from $\mathbf{A}$ providing a token in $\mathbf{B}$ with the same message. As a simple example we note that, for the information systems $\perp_n$ given in 9.15(6), we have

$$\perp_0 \trianglelefteq \perp_1 \trianglelefteq \ldots \trianglelefteq \perp_n \trianglelefteq \ldots .$$

We next look at the order properties of $\langle \operatorname{Sub}\mathfrak{J}; \sqsubseteq\rangle$, the family of substructures of $\mathfrak{J}$, and re-interpret these in terms of information systems. The proof of Theorem 9.23 is routine and we omit it.

9.23 Theorem. *Let $\mathfrak{J}$ be an algebraic $\bigcap$–structure. Then $\langle \operatorname{Sub}\mathfrak{J}, \sqsubseteq\rangle$ forms a domain.*

(i) *The bottom element is $\mathfrak{N} := \{\varnothing\}$, the unique algebraic $\bigcap$–structure based on $\varnothing$.*

(ii) *Let $\{(\mathfrak{L}_i, A_i)\}_{i\in I}$ be a non-empty family of substructures of $\mathfrak{J}$. Then*

$$\Big\{ X \subseteq \bigcap\nolimits_{i\in I} A_i \mid (\forall j \in I)(\exists Y_j \in \mathfrak{L}_j)\, X = Y_j \cap \bigcap\nolimits_{i\in I} A_i \Big\}$$

is an algebraic $\bigcap$–structure based on $\bigcap_{i\in I} A_i$ and equals $\bigwedge_{i\in I}\mathfrak{L}_i$.

(iii) *Let $\{(\mathfrak{L}_i, A_i)\}_{i\in I}$ be a family of substructures of $\mathfrak{J}$ directed with respect to $\sqsubseteq$. Then*

$$\Big\{ X \subseteq \bigcup\nolimits_{i\in I} A_i \mid (\forall j \in I)\, X \cap A_j \in \mathfrak{L}_j \Big\}$$

is an algebraic $\bigcap$–structure on $\bigcup_{i\in I} A_i$ and equals $\bigsqcup_{i\in I} \mathfrak{L}_i$.

9.24 Proposition. *Let $\{\mathfrak{L}_i\}_{i\in I}$ be a non-empty family of substructures of the algebraic $\bigcap$–structure $\mathfrak{J}$ and let $\mathbf{IS}(\mathfrak{L}_i) = \langle A_i, Con_i, \vdash_i\rangle$.*

(i) $\mathbf{IS}(\bigwedge_{i\in I}\mathfrak{L}_i) = \langle \bigcap A_i, \bigcap Con_i, \bigcap \vdash_i\rangle$.

(ii) *If $\{(\mathfrak{L}_i, A_i)\}_{i\in I}$ is directed,* $\mathbf{IS}(\bigsqcup_{i\in I}\mathfrak{L}_i) = \langle \bigcup A_i, \bigcup Con_i, \bigcup \vdash_i\rangle$.

(Here the entailment relation $\vdash$ of $\langle A, Con, \vdash\rangle$ is taken to be the subset $\{(X, a) \mid X \vdash a\}$ of $Con \times A$.)

Proof. Apply Theorem 9.19. □

9.25 Constructions. We have already noted that CPOs can be combined by forming sums, products and so on. It is far from clear that when we combine domains in such ways we remain within the class of domains. The three-way correspondence

$$\text{information system} \longleftrightarrow \text{algebraic } \bigcap\text{–structure} \longleftrightarrow \text{domain}$$

allows us to carry out constructions in whichever of these settings seems expedient and then use 9.8 and 9.19 to translate them into the formulation required. For the simpler constructions it is very convenient to work with algebraic $\bigcap$–structures, since it is quick and routine to check

that the resulting structures are algebraic, so that the constructions do yield domains from domains. Proposition 9.26 collects a group of such results together and Example 9.27 illustrates the definitions in simple cases. We follow this with an example to show how the corresponding constructions for information systems can be read off, using Proposition 9.26. The other constructions are handled similarly.

At first sight, the definitions we give in 9.26 may look somewhat awkward. In (ii) the device of forming the product of the base sets A and B with $\{0\}$ and $\{1\}$ is necessary to force disjointness (see Exercise 1.9); from an information system viewpoint this is like colouring the tokens of two systems red and blue to keep track of which system they refer to. Similar tricks are employed in (i) and (iii). Also in (iv), the tokens in $\bot_{\mathfrak{L}}$ correspond to 'no information', so there is no harm in deleting them, so that $\bot_{\mathfrak{L}} = \varnothing$, and similarly for $\mathfrak{K}$. This is done in forming the coalesced sum which was introduced, along with separated sum, in 8.4.

9.26 Proposition. *Let $(\mathfrak{L}, A)$ and $(\mathfrak{K}, B)$ be $\bigcap$–structures and let I be a set.*

(i) *For all $T \subseteq A$, define $T_0 := \{(0,0)\} \cup (\{1\} \times T)$. Define*

$$\mathfrak{L}_\bot := \{\, S \subseteq A_0 \mid S = T_0 \text{ for some } T \in \mathfrak{L} \,\} \cup \{\varnothing\}.$$

Then $\mathfrak{L}_\bot$ is an $\bigcap$–structure on A_0 and is order-isomorphic to $\mathbf{1} \oplus \mathfrak{L}$.

(ii) *For all $S \subseteq A$ and $T \subseteq B$, define $S \dot\cup T = (\{0\} \times S) \cup (\{1\} \times T)$. Define*

$$\mathfrak{L} \boxtimes \mathfrak{K} := \{\, S \dot\cup T \mid S \in \mathfrak{L}, T \in \mathfrak{K} \,\}.$$

Then $\mathfrak{L} \boxtimes \mathfrak{K}$ is an $\bigcap$–structure on $A \dot\cup B$ which is order-isomorphic to the ordered-set product $\mathfrak{L} \times \mathfrak{K}$.

(iii) *Define $(I \boxminus \mathfrak{L}) := \{\, \{\, (i,a) \in I \times A \mid a \in f(i) \,\} \mid f \in (I \to \mathfrak{L}) \,\}$. Then $(I \boxminus \mathfrak{L})$ is an $\bigcap$–structure on $I \times A$ which is order-isomorphic to the power $(I \to \mathfrak{L}) = \mathfrak{L}^I$ of $\mathfrak{L}$.*

(iv) *Let $A \dot\cup_\vee B$ be $(\{0\} \times (A \setminus \bot_{\mathfrak{L}})) \cup (\{1\} \times (B \setminus \bot_{\mathfrak{K}}))$ and define*

$$\mathfrak{L} \boxplus_\vee \mathfrak{K} := \{\, \{0\} \times (S \setminus \bot_{\mathfrak{L}}) \mid S \in \mathfrak{L} \,\} \cup \{\, \{1\} \times (T \setminus \bot_{\mathfrak{K}}) \mid T \in \mathfrak{K} \,\} \cup \{\varnothing\}.$$

Then $\mathfrak{L} \boxplus_\vee \mathfrak{K}$ is an $\bigcap$–structure on $A \dot\cup_\vee B$ which is order-isomorphic to the coalesced sum $\mathfrak{L} \oplus_\vee \mathfrak{K}$.

(v) *$\mathfrak{L} \boxplus_\bot \mathfrak{K} := \mathfrak{L}_\bot \boxplus_\vee \mathfrak{K}_\bot$ is an $\bigcap$–structure on $A_0 \dot\cup_\vee B_0$ which is order-isomorphic to the separated sum $\mathfrak{L} \oplus_\bot \mathfrak{K}$.*

Further, if $\mathfrak{L}$ and $\mathfrak{K}$ are algebraic, then so are $\mathfrak{L}_\bot$, $\mathfrak{L} \boxtimes \mathfrak{K}$, $(I \boxminus \mathfrak{L})$, $\mathfrak{L} \boxplus_\vee \mathfrak{K}$ and $\mathfrak{L} \boxplus_\bot \mathfrak{K}$.

9.27 Example. We first illustrate 9.26(i). Take $\mathfrak{N} = \{\varnothing\}$. Then $\mathfrak{N}_\bot$ has base set $\{(0,0)\}$ and consists of the sets $\varnothing$ and $\{(0,0)\}$. Now repeat the process: $(\mathfrak{N}_\bot)_\bot$ has base set $\{(0,0),(1,(0,0))\}$ and consists of the sets $\varnothing$, $\{(0,0)\}$ and $\{(0,0),(1,(0,0))\}$. We have $\mathfrak{N} \sqsubseteq \mathfrak{N}_\bot \sqsubseteq (\mathfrak{N}_\bot)_\bot$, and so on. This chain of $\bigcap$–structures corresponds to $|\bot_0|$, $|\bot_1|$, $|\bot_2|$, ...

Now consider $\mathfrak{N}_\bot \boxplus_\vee (\mathfrak{N}_\bot)_\bot$. This has base set

$$\{(0,(0,0)),(1,(0,0)),(1,(1,(0,0)))\}.$$

Its members are

$$\varnothing,\ \{(0,(0,0))\},\ \{(1,(0,0))\},\ \{(1,(0,0)),(1,(1,(0,0)))\};$$

as anticipated, $\mathfrak{N}_\bot \boxplus_\vee (\mathfrak{N}_\bot)_\bot$ with its inclusion order is isomorphic to $2 \oplus_\vee 3 \cong \mathbf{1} \oplus (\mathbf{1} \dot\cup \mathbf{2})$.

9.28 Example. Take information systems $\mathbf{A} = \langle A, Con_A, \vdash_A \rangle$ and $\mathbf{B} = \langle B, Con_B, \vdash_B \rangle$. Their product is defined to be $\mathbf{IS}(|\mathbf{A}| \boxtimes |\mathbf{B}|)$, equal to $\langle C, Con, \vdash \rangle$, say. Proposition 9.26 tells us immediately that $C = A \dot\cup B$, as defined in 9.26(ii). To describe Con and $\vdash$, we write

$$X_0 := \{\, a \in A \mid (0,a) \in X \,\} \quad \text{and} \quad X_1 := \{\, b \in B \mid (1,b) \in X \,\},$$

for any $X \subseteq A \dot\cup B$. Then, again from 9.26,

$$Con = \{\, X \Subset C \mid X_0 \in Con_A \text{ and } X_1 \in Con_B \,\}$$

and $(X \vdash (0,a) \Leftrightarrow X_0 \vdash_A a)$ and $(X \vdash (1,b) \Leftrightarrow X_1 \vdash_B b)$.

9.29 Approximable mappings. In the applications of domain theory it is very important to know that the CPO of continuous maps from one domain to another is itself a domain. The search for an elementary route to this result has probably been largely responsible for the proliferation of alternative approaches to domains. We go via approximable mappings, which are to information systems what continuous maps are to domains (recall 1.38). An approximable mapping is not a map in the usual sense. It is an 'information-respecting' relation, to be thought of as a machine which from any finite set of information in one information system produces output in another.

Specifically, given information systems $\mathbf{A} = \langle A, Con_A, \vdash_A \rangle$ and $\mathbf{B} = \langle B, Con_B, \vdash_B \rangle$, an **approximable mapping** is a subset r of $Con_A \times B$, satisfying (AM1) and (AM2) below. We read $(Y,b) \in r$ as 'under r (the input) Y produces (the output) b'. Wherever possible, it is best to replace $(Y,b) \in r$ by the more suggestive notation $r\colon Y \rightsquigarrow b$, or simply $Y \rightsquigarrow b$ when the name of the approximable mapping is not required. The axioms that must be satisfied are:

(AM1) $Y \rightsquigarrow b$ for all $b \in Z \Subset B$ implies

(a) $Z \in Con_B$ and

(b) $Z \vdash_B c \Rightarrow Y \rightsquigarrow c$;

(AM2) $Y \vdash_A a$ for all $a \in Y'$ and $Y' \rightsquigarrow b$ imply $Y \rightsquigarrow b$.

Here $\{ X \mid X \rightsquigarrow b \text{ for some } b \in B \}$ should be thought as the inputs and $\{ b \mid X \rightsquigarrow b \text{ for some } X \in Con_A \}$ as the information which is output.

The requirements on r are sensible. Part (a) of (AM1) says that putting consistent information into r gives consistent output, while (b) says that if the total output resulting from Y is enough to deduce c in B, then we get c itself as part of the output. Condition (AM2) ensures that adding extra consistent information to the input doesn't alter the output. Also, the conditions are exactly what is needed to make the next proof work.

The family of all approximable mappings is a family of subsets of $Con_A \times B$ and is readily seen to be an algebraic $\bigcap$–structure. Proposition 9.30 shows that this family gives the function space we require. Note the mix of $\bigcap$–structures and information systems.

9.30 Proposition. *Let $(\mathfrak{L}, A)$ and $(\mathfrak{K}, B)$ be algebraic $\bigcap$–structures. Then the family of approximable mappings from $\mathbf{IS}(\mathfrak{L})$ to $\mathbf{IS}(\mathfrak{K})$ is an algebraic $\bigcap$–structure which is order-isomorphic to $[\mathfrak{L} \to \mathfrak{K}]$. The isomorphism associates to $\varphi \in [\mathfrak{L} \to \mathfrak{K}]$ the approximable mapping s_φ given by $s_\varphi : Y \rightsquigarrow b$ if and only if $b \in \varphi(\overline{Y})$.*

Proof. A routine check confirms that s_φ is an approximable mapping whenever φ is a well-defined order-preserving map.

In the other direction, assume that r is an approximable mapping and define $|r|$ on $\mathfrak{L}$ by

$$|r|(U) := \{ b \in B \mid (\exists Y \Subset U)\, r : Y \rightsquigarrow b \}.$$

To show that $|r| : \mathfrak{L} \to \mathfrak{K}$, it is enough by 9.19 to check that $|r|(U)$ is consistent and $\vdash$–closed for each $U \in \mathfrak{L}$. Let $Z := \{b_1, \ldots, b_n\} \Subset |r|(U)$. For each $i = 1, \ldots, n$, there exists $Y_i \Subset U$ such that $r : Y_i \rightsquigarrow b_i$. Then, by (IS4) and (AM2), we have $r : Y \rightsquigarrow b_i$ for each i, where $Y := Y_1 \cup \cdots \cup Y_n$. By (AM1)(a), Z is consistent in $\mathbf{IS}(\mathfrak{K})$, so $|r|(U)$ is consistent. To check $\vdash$–closure, assume also that $Z \vdash c$. By (AM1)(b), $r : Y \rightsquigarrow c$, whence $c \in |r|(U)$.

We claim that $|r| \in [\mathfrak{L} \to \mathfrak{K}]$. Let $\{U_i\}_{i \in I}$ be a directed family in $\mathfrak{L}$. The directedness implies that $Y \Subset \bigcup U_i$ if and only if $Y \Subset U_j$ for some j. It follows immediately that $|r|(\bigcup U_i) = \bigcup |r|(U_i)$, as required.

We now have maps $\Phi : \varphi \mapsto s_\varphi$ and $\Psi : r \mapsto |r|$. To prove that Φ and Ψ are mutually inverse bijections between $[\mathfrak{L} \to \mathfrak{K}]$ and the family

of approximable mappings, we need $\varphi = |s_\varphi|$ and $r = s_{|r|}$. For $U \in \mathfrak{L}$,

$$\begin{aligned}
\varphi(U) &= \varphi\Big(\bigsqcup \{ \overline{Y} \mid Y \Subset U \}\Big) && \text{(by 9.17)} \\
&= \bigsqcup \{ \varphi(\overline{Y}) \mid Y \Subset U \} && \text{(since } \varphi \text{ is continuous)} \\
&= \{ b \in B \mid (\exists Y \Subset U)\, s_\varphi : Y \rightsquigarrow b \} && \text{(by the definition of } s_\varphi) \\
&= |s_\varphi|(U) && \text{(by the definition of } |\text{–}|).
\end{aligned}$$

Let r be an approximable mapping. Then, for all $Y \in Con_A$ and $b \in B$,

$$\begin{aligned}
s_{|r|} : Y \rightsquigarrow b &\iff b \in |r|(\overline{Y}) \\
&\iff (\exists Z \Subset \overline{Y})\, r : Z \rightsquigarrow b \\
&\iff r : Y \rightsquigarrow b \qquad \text{(by (AM2))}.
\end{aligned}$$

The correspondence between $[\mathfrak{L} \to \mathfrak{K}]$ and the approximable mappings is an order-isomorphism provided Φ and Ψ are order-preserving. It follows from the definition that Ψ is order-preserving. Note also that $\varphi_1 \leqslant \varphi_2$ in $[\mathfrak{L} \to \mathfrak{K}]$ implies that $\varphi_1(\overline{Y}) \subseteq \varphi_2(\overline{Y})$ for all Y such that $Y \Subset U$ for some $U \in \mathfrak{L}$. But $s_{\varphi_i} : Y \rightsquigarrow b$ if and only if $b \in \varphi_i(\overline{Y})$ for $i = 1, 2$. It follows that $\varphi_1 \leqslant \varphi_2$ implies $s_{\varphi_1} \subseteq s_{\varphi_2}$, so Φ is order-preserving. □

Thus the interplay between domains, $\bigcap$–structures and information systems culminates in the following important theorem.

9.31 Theorem. *Let P and Q be domains and let I be a set. Then each of the following is also a domain:*

$$P_\perp,\ P \oplus_\perp Q,\ P \oplus_\vee Q,\ P \times Q,\ (I \to P),\ [P \to Q].$$

9.32 Domain constructors. The various constructions often need to be combined, in order to produce more complex domains. An n **to** m **domain constructor** maps an n–tuple of domains to an m–tuple of domains. Lifting, that is $P \mapsto P_\perp$, and $P \mapsto [[P \to P] \to P]$ are examples of **unary constructors** ($n = m = 1$). Taking $*$ equal to $\oplus_\perp$, $\oplus_\vee$, $\times$ or $\to$ in $(P, Q) \mapsto P * Q$, we obtain **binary constructors** ($n = 2$, $m = 1$).

We indicate in 9.33 how the search for mathematical models for programming languages leads to the problem of solving for P 'domain equations' $P \cong \mathbf{F}(P)$, where $\mathbf{F}$ is some domain constructor. Because of the way P is 'defined' in terms of itself, it is far from clear that a solution to this equation exists. Such apparent circularity is not

an insuperable problem in simpler cases. It is easy to see that the equation $P \cong P_\perp$ is solved by taking $P = \mathbb{N} \oplus \mathbf{1}$ and it is almost as obvious that $P = \Sigma^{**}$ is a solution to $P \cong P \oplus_\perp P$. More complicated equations, for example $P \cong [P \to P]_\perp$, $P \cong Q \oplus_\perp [P \to P_\perp]$ or $P \cong Q \oplus_\perp (P \times P) \oplus_\perp [[P \to P] \to P]$ (where Q is fixed), look far less tractable.

A domain equation $P \cong \mathbf{F}(P)$ can be recast in the language of algebraic $\bigcap$–structures or of information systems. Then $\mathbf{F}$ may be treated as an operator on a CPO of algebraic $\bigcap$–structures ordered by $\sqsubseteq$ or on a CPO of information systems ordered by $\trianglelefteq$ (see 9.22). The fixpoint theorems from Chapter 8 show that the above equation is soluble if $\mathbf{F}$ is continuous, or, if we are willing to appeal to harder theory, just order-preserving.

Let $\mathfrak{X}$ be the family of substructures of some algebraic $\bigcap$–structure, ordered by $\sqsubseteq$. It is elementary that the lifting constructor preserves $\sqsubseteq$. To show that the binary constructors associated with sums, product and function space are order-preserving (as maps defined on $\mathfrak{X} \times \mathfrak{X}$) it is convenient to use the fact that the binary constructor $(\mathfrak{L}, \mathfrak{K}) \mapsto \mathfrak{L} * \mathfrak{K}$, where $\mathfrak{L}$ and $\mathfrak{K}$ range over $\mathfrak{X}$, is order-preserving if and only if each of the unary constructors $\mathfrak{L} \mapsto \mathfrak{L} * \mathfrak{K}$ (for fixed $\mathfrak{K}$) and $\mathfrak{K} \mapsto \mathfrak{L} * \mathfrak{K}$ (for fixed $\mathfrak{L}$) is order-preserving. To prove sufficiency, note that

$$\begin{aligned}(\mathfrak{L}_1, \mathfrak{K}_1) \sqsubseteq (\mathfrak{L}_2, \mathfrak{K}_2) &\Longrightarrow \mathfrak{L}_1 \sqsubseteq \mathfrak{L}_2 \;\&\; \mathfrak{K}_1 \sqsubseteq \mathfrak{K}_2 \\ &\Longrightarrow \mathfrak{L}_1 * \mathfrak{K}_1 \sqsubseteq \mathfrak{L}_2 * \mathfrak{K}_1 \;\&\; \mathfrak{L}_2 * \mathfrak{K}_1 \sqsubseteq \mathfrak{L}_2 * \mathfrak{K}_2 \\ &\Longrightarrow \mathfrak{L}_1 * \mathfrak{K}_1 \sqsubseteq \mathfrak{L}_2 * \mathfrak{K}_2.\end{aligned}$$

The proof of necessity is equally easy. Checking this condition when $*$ is any of $\boxtimes$, $\boxplus_\vee$ or $\boxplus_\perp$ (as in 9.26) or $\to$ (as in 9.30) requires a clear head, but no ingenuity. It is recommended as an exercise to those wishing to become familiar with the definitions.

In fact, each of our constructors is continuous. To prove this we may first invoke Exercise 8.7, which reduces the problem to that of checking continuity of unary operators. By definition, a unary constructor $\mathbf{F}$ on $\mathfrak{X}$ is continuous if and only if

$$\bigsqcup_{i \in I} \mathbf{F}((\mathfrak{L}_i, A_i)) = \mathbf{F}\Big(\bigsqcup_{i \in I} (\mathfrak{L}_i, A_i)\Big)$$

for any directed family $\{(\mathfrak{L}_i, A_i)\}_{i \in I}$ in $\mathfrak{X}$. The remark following the definition of $\sqsubseteq$ in 9.22, together with Exercise 8.7, imply that this holds provided $\mathbf{F}$ is order-preserving and the $\bigcap$–structures $\bigsqcup \mathbf{F}((\mathfrak{L}_i, A_i))$ and $\mathbf{F}(\bigsqcup (\mathfrak{L}_i, A_i))$ have the same base set. We say that $\mathbf{F}$ is **continuous on base sets** if the latter condition is satisfied. In terms of information systems, this just means that every token of $\mathbf{IS}(\bigsqcup \mathfrak{L}_i)$ is a token drawn from $\mathbf{IS}(\mathfrak{L}_i)$ for some i. Checking the continuity of each of our domain constructors is now quite straightforward.

We conclude this chapter with an informal account of the rudiments of denotational semantics, to indicate how domains and fixpoint theory underpin this approach to programming languages. Our discussion is perforce very brief and is aimed principally at readers with some experience of programming. References to full treatments of the subject can be found in Appendix B.

9.33 Denotational semantics and semantic domains. Suppose $\mathbf{L}$ is a programming language specified by formal syntactic rules. One way to analyze $\mathbf{L}$ is to construct a concrete mathematical model $\mathbf{M}$ and a map $\mathcal{V}\colon \mathbf{L} \to \mathbf{M}$ (the valuation map), which to each object P in $\mathbf{L}$ (that is, a program constituent or a complete program) assigns $\mathcal{V}[\![P]\!]$ in $\mathbf{M}$, denoting the 'meaning' of P. For example, P might be the multiplication operator $*$ and $\mathcal{V}[\![P]\!]$ the operation $\times$ of multiplication on $\mathbb{Z}$ (assuming that $\mathbf{M}$ is such that $\mathbb{Z} \subseteq \mathbf{M}$).

The abstract language $\mathbf{L}$ may be regarded as being made up of various 'syntactic categories' (variables, commands, the expressions on which commands act, etc.), with complex program constructs defined in terms of simpler components. We seek to associate a 'semantic domain' to each syntactic category, to act as its concrete realization within $\mathbf{M}$. Starting from the most primitive of $\mathbf{L}$'s syntactic categories, we build up by stages the model $\mathbf{M}$, and the corresponding valuation map. This approach, which we illustrate below, is known as 'denotational semantics'. The model $\mathbf{M}$ enhances our understanding of $\mathbf{L}$ and enables us to reason about it. For example, we might seek to confirm that two programs in $\mathbf{L}$ have the same effect by showing they always have the same meaning in $\mathbf{M}$. In the other direction, the formal language may help us talk about $\mathbf{M}$. The study of propositional logic reveals a similar interplay between syntactic and semantic philosophies (see Chapter 10): the syntactic approach uses a formal language and the semantic approach uses truth values. Predicate logic provides an even closer parallel. Denotational semantics is to programming languages what model theory is to predicate logic.

Just as five year olds are taught to read from Ladybird books rather than *War and Peace*, a study of semantics begins not with a fully-fledged programming language but with 'baby' languages, designed to exhibit particular programming features. High-level languages are built up from language fragments in a modular fashion. We therefore start by looking at a very simple imperative language, adequate only for doing arithmetic

on $\mathbb{N}$. This language has two primitive syntactic categories: *Val* (**basic values**) and *Id* (**identifiers**, which are our variables). The members of *Val* are 'abstract versions' of the values our programs take. They fall into two sets: the **numerals**, $Num = \{\underline{1}, \underline{2}, \dots\}$, and the **Boolean values**, $Bool = \{\mathtt{tt}, \mathtt{ff}\}$. The associated semantic domain is $V := \mathbb{N} \cup \{\mathsf{T}, \mathsf{F}\}$; our notation for 'concrete' truth values is as in Chapter 4. The valuation map takes each member of *Val* to the natural number or truth value it 'means'. Thus, $\mathcal{V}[\![\mathtt{tt}]\!] = \mathsf{T}$ and $\mathcal{V}[\![\underline{42}]\!] = 42$, etc.

Out of *Val* we build the **expressions**. These are of two types, **numeric** and **Boolean**:

$$N ::= N \mid (N_1 + N_2) \mid (N_1 - N_2) \mid I$$
$$B ::= B \mid \neg B \mid (N_1 = N_2).$$

Here we have used what is called BNF notation, a shorthand way of stating which operations are admitted. In this case, each numeric expression is either a numeral, the sum or difference of two numerals, or an identifier. A Boolean expression is either a Boolean value, the negation of a Boolean expression, or the assertion that two numeric expressions are equal. We could, of course, make our language more complex by adding additional clauses to the prescription of the syntax, to represent further operations. We use generic letters (subscripted or unsubscripted) to label objects in the language: N for numeric expressions, E for expressions, C for commands, and so on. We adopt the conventional notation and use $\mathcal{E}$ and $\mathcal{C}$ to denote the restriction of $\mathcal{V}$ to *Exp* and *Cmd*.

Language **L**	Valuation map $\xrightarrow{\mathcal{V}}$	Model **M**
Syntactic categories		Semantic domains
Basic values, *Val*		$V := \mathbb{N} \cup \{\mathsf{T}, \mathsf{F}\}$
Numerals, *Num*		$\mathbb{N}$
Boolean values, *Bool*		$\{\mathsf{T}, \mathsf{F}\}$
Identifiers, *Id*		*Id*
Expressions, *Exp*	$\mathcal{E}$	$(St \to V_\perp)$
Commands, *Cmd*	$\mathcal{C}$	$(St_\perp \to St_\perp)$

Table 9.1

The values assigned to identifiers are not constants. They change as a program runs. A state σ records, as $\sigma(I)$, the current value of each identifier I. Formally σ is a map from *Id* to V. We then define $\mathcal{E}[\![I]\!]\sigma = \sigma(I)$, for $\sigma \in St$, the set of states. This reflects our intention

that the meaning of an identifier at a given moment is the value currently assigned to it. We take the semantic domain for *Exp* to be $(St \to V_\perp)$. This choice allows us to define $\mathcal{E}$ on compound expressions as follows:

$$\mathcal{E}[\![(N_1 + N_2)]\!]\sigma = \mathcal{E}[\![N_1]\!]\sigma + \mathcal{E}[\![N_2]\!]\sigma;$$

$$\mathcal{E}[\![(N_1 - N_2)]\!]\sigma = \begin{cases} \mathcal{E}[\![N_1]\!]\sigma - \mathcal{E}[\![N_2]\!]\sigma & \text{if this makes sense in } \mathbb{N}, \\ \perp & \text{otherwise;} \end{cases}$$

$$\mathcal{E}[\![\neg B]\!]\sigma = \mathsf{T} \iff \mathcal{E}[\![B]\!]\sigma = \mathsf{F};$$

$$\mathcal{E}[\![E_1 = E_2]\!]\sigma = \begin{cases} \mathsf{T} & \text{if } \mathcal{E}[\![E_1]\!]\sigma = \mathcal{E}[\![E_2]\!]\sigma, \\ \mathsf{F} & \text{otherwise.} \end{cases}$$

On the left-hand sides, $+$, $-$ and $=$ are drawn from the formal language, while on the right they have their usual meanings in V. The set V is lifted so that $\perp$ can be used to model an 'error value'.

Finally, we need commands. These are 'state transformers': a command acts on a state and returns a new state. Their semantic domain is $(St_\perp \to St_\perp)$. The lifting on the right allows for an error value. That on the left accommodates non-termination (looping): $\perp$ fed to a command is required to yield $\perp$. Our syntax for commands is

$$C ::= I := E \mid C_1; C_2 \mid \texttt{if } B \texttt{ then } C_1 \texttt{ else } C_2$$

This says that a command either assigns an expression to an identifier, or is two simpler commands in sequence (C_1 followed by C_2), or is an abstract version of `if-then-else`. The meanings of commands are determined as follows:

$\mathcal{C}[\![I := E]\!]\sigma$ is the state which at $I' \neq I$ takes value $\sigma(I')$, and at I takes value $\mathcal{E}[\![E]\!]\sigma$,

$$\mathcal{C}[\![C_1; C_2]\!]\sigma = (\mathcal{C}[\![C_2]\!] \circ \mathcal{C}[\![C_1]\!])\sigma,$$

$$\mathcal{C}[\![\texttt{if } B \texttt{ then } C_1 \texttt{ else } C_2]\!]\sigma = cond(\mathcal{E}[\![B]\!], \mathcal{C}[\![C_1]\!], \mathcal{C}[\![C_2]\!])\sigma.$$

In the last definition, *cond* maps $(St \to \{\mathsf{T}, \mathsf{F}\}) \times (St \to St) \times (St \to St)$ to St and is given by

$$cond(P, f, g)\sigma = \begin{cases} f(\sigma) & \text{if } P(\sigma) = \mathsf{T}, \\ g(\sigma) & \text{otherwise.} \end{cases}$$

Putting together the semantic domains for *Exp* and *Cmd* by forming the separated sum, we get our model **M** for **L**:

$$\mathbf{M} = (St_\perp \to V_\perp) \oplus_\perp (St_\perp \to St_\perp).$$

By Theorem 9.31, **M** is indeed a domain. The construction of **M** has the following features:

(i) the linguistic constructs on the primitive syntactic categories (+, =, etc.) are given global (that is, once-and-for-all) meanings,

and, fundamental to the philosophy of denotational semantics,

(ii) the meaning of a compound syntactic object is determined by the meanings of its syntactic subcomponents.

To add more elaborate commands, such as '`while` B `do` C', which involve recursion, we need the machinery of fixpoints.

9.34 The while-loop. In the examples in Chapter 8 we started from a fixpoint equation and sought its solution(s). More commonly in applications, the starting point is a recursive definition of some map or procedure, with the object to be defined recognizable as a solution of some fixpoint equation. This is the situation here.

We ask how commands of the form '`while` B `do` C' should be assigned a meaning. Intuitively, the interpretation should be: 'so long as B is true, do C repeatedly; once B is false, stop in current state'. We want, for any state σ, $\mathcal{C}[\![\texttt{while } B \texttt{ do } C]\!]\sigma$ to be $\mathcal{C}[\![\texttt{while } B \texttt{ do } C]\!]\mathcal{C}[\![C]\!]\sigma$ if $\mathcal{E}[\![B]\!]$ is true and σ otherwise. More succinctly, we are demanding

$$\mathcal{C}[\![\texttt{while } B \texttt{ do } C]\!] = \mathit{cond}(\mathcal{E}[\![B]\!], \mathcal{C}[\![\texttt{while } B \texttt{ do } C]\!] \circ \mathcal{C}[\![C]\!], \mathrm{id}).$$

This appears to require $\mathcal{C}[\![\texttt{while } B \texttt{ do } C]\!]$ to be defined in terms of itself, indicating that we are dealing with recursion. The position is clarified by writing A for $\mathcal{C}[\![\texttt{while } B \texttt{ do } C]\!]$ (a member of the domain $D := (St_\perp \to St_\perp)$) and F for the map from D to D which sends f to $\mathit{cond}(\mathcal{E}[\![B]\!], f \circ \mathcal{C}[\![C]\!], \mathrm{id})$. Then the while-loop equation becomes $F(A) = A$. Thus the problem of showing that the recursive procedure '`while` B `do` C' does have a proper definition is reduced to that of showing that a certain fixpoint equation has a solution.

The while-loop shows striking similarities to the factorial function which we analysed in 8.13 and 8.18. The successive approximations to $\mathcal{C}[\![\texttt{while } B \texttt{ do } C]\!]$ are given, for $n \geqslant 1$, by $\{W_n(B,C)\}_{n\geqslant 1}$, where $W_n(B,C)$ coincides with '`while` B `do` C' for computations involving fewer than n iterations of the loop, and is undefined otherwise; $W_0(B,C)$ is 'do nothing'. The transition from $W_n(B,C)$ to $W_{n+1}(B,C)$ accomplishes one stage in unwinding the loop. Fixpoint theory allows us to realize a recursively defined object as the limit of partially defined objects which can be specified without recursion.

In programming, there are often alternative ways of modelling intended behaviour, and a consequent need to verify the denotational equivalence of program constructs defined in different ways, for example

different versions of 'while'. Here the advantage of realizing the constructs as least fixpoints emerges. Many of the rules derived in Chapter 8 for calculating with least fixpoints have as their conclusions the equality of least fixpoints of different maps.

Using fixpoint theorems to solve domain equations

Our fixpoint examples so far have concerned maps and procedures. In this section we hint at the potential of CPO Fixpoint Theorem I for creating solutions to domain equations.

9.35 Introducing domain equations. To see how equations between domains may arise, let us consider what might be involved in setting up a semantic model for a functional programming language. This is much more difficult than the imperative programming we considered above. We now suggest, without setting up the syntax and semantics in full, what shape a model $\mathbf{M}$ for a language of this sort might have. We may take *Val* and V as before. We replace the previous categories *Exp* and *Cmd* by a single category of expressions. This encompasses our previous commands and expressions, except that the assignment clause $I := E$ is omitted and a mechanism for defining maps and procedures is substituted. The latter makes use of a new syntactic category of 'declarations'. Values of compound programs are to be determined from the values of their parts, so we must allow $\mathbf{M}$ to contain maps. Maps may well need to act on other maps and we may require procedures to act on maps or even other procedures. We therefore do not want to tie ourselves down by specifying the 'type' of the argument of a map or procedure. This suggests we should take as the semantic domain for expressions the domain $(\mathbf{M} \to \mathbf{M}_\perp)$, so that the meanings of expressions may be maps defined on any part of $\mathbf{M}$; recall 8.12. When all this is fleshed out and made precise (no small undertaking) it leads to the conclusion that a model $\mathbf{M}$ for a functional language might need to satisfy $\mathbf{M} \cong V \oplus_\perp (\mathbf{M} \to \mathbf{M}_\perp)$. This is patently impossible on cardinality grounds, by Cantor's Theorem (see [6]). The cardinality problem disappears if instead of *all* maps from $\mathbf{M}$ to $\mathbf{M}_\perp$ we take $[\mathbf{M} \to \mathbf{M}_\perp]$. Since it is the continuous maps that are the computationally significant ones (recall 9.10), this is a sensible modification in any case. We are now faced with the problem of whether there is a domain $\mathbf{M}$ satisfying

$$\mathbf{M} \cong V \oplus_\perp [\mathbf{M} \to \mathbf{M}_\perp].$$

Fixpoint theory is just what is needed to show that this equation is

soluble, along with many others of a similar kind. The existence of domains to serve as models for functional languages is thus ensured.

9.36 Recursively defined domains. As a simple first example, consider the domain equation $\mathbf{F}(P) \cong P$, where $\mathbf{F}(P) = P_\perp$. Starting from $\mathbf{1}$ and forming $\{\mathbf{F}^n(\mathbf{1})\}_{n\geqslant 1}$, we obtain $\mathbf{1}$, $\mathbf{1}_\perp \cong \mathbf{2}$, $\mathbf{2}_\perp \cong \mathbf{3}$, etc. This makes it highly plausible that the least solution to $\mathbf{F}(P) \cong P$ is the domain $\mathbb{N}\oplus\mathbf{1}$. To confirm this, we must realize $\mathbf{F}$ as an order-preserving map on a CPO of domains and verify that $\mathbb{N}\oplus\mathbf{1}$ is indeed the join in this CPO of the approximations $\{\mathbf{F}^n(\mathbf{1})\}_{n\geqslant 0}$. We do not have an ordering of 'abstract' domains and must therefore realize domains concretely, either as $\bigcap$-structures (ordered by $\sqsubseteq$) or as information systems (ordered by $\trianglelefteq$). Using $\bigcap$-structures, we then have to construct $\mathbf{F}^n(\mathfrak{N})$. The reward for disentangling this notational horror (recall 9.27) is that the correct embeddings are then given at once via Exercise 9.15. This confirms that the 'natural' nesting of the chains $\mathbf{1}, \mathbf{2}, \ldots$ (in which each chain sits, in $\mathbb{N}\oplus\mathbf{1}$, as a down-set in the next) corresponds to the $\sqsubseteq$-ordering of their realizations as $\bigcap$-structures. See Figure 9.2 (in which $\mathbf{0}$ is used as an abbreviation for $(0,0)$).

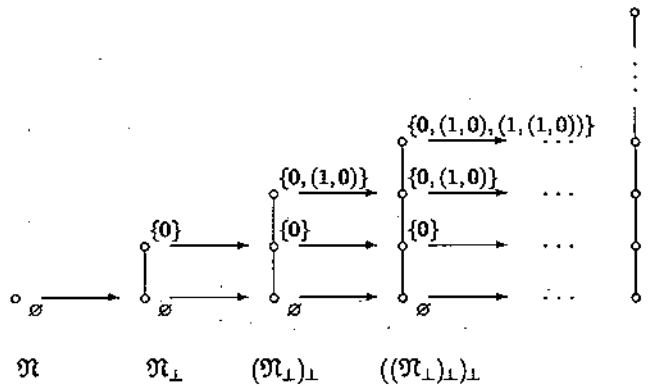


Figure 9.2

As a second example, consider the equation $P \cong P \oplus_\perp P$. Let $\mathbf{F}(P) = P \oplus_\perp P$. Then the least fixpoint of $\mathbf{F}$ is isomorphic to Σ^{**}, the domain of all binary strings. The nth approximation, $\mathbf{F}^n(\varnothing)$, to the solution is the set of strings of length at most n; again the ordering $\sqsubseteq$ is the 'natural' nesting. Finally, we note that we cannot exhibit explicit solutions to domain equations, such as $P \cong [P \to P]_\perp$, which involve function spaces. As those who did Exercise 1.28 will realize, these spaces get very unwieldy very quickly and the limit domain cannot be visualized

easily. However, so long as the domain constructor used is continuous, or more generally order-preserving, we know that there is a solution to the associated fixpoint equation, and this is what really matters, since it ensures that the models required for denotational semantics do indeed exist.

Exercises

Exercises from the text. Prove Rules (a) and (b) in 9.14. Verify that the examples in 9.15 are indeed information systems, with elements as specified in 9.16. Complete the proof of Lemma 9.17 by proving the implications (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii). Prove the claim in 9.18 that $F(|\mathbf{A}|) = \{\overline{X} \mid X \in Con\}$ for any information system $\mathbf{A}$. Complete the proof of Theorem 9.19. Prove the claim in 9.22 that for information systems $\mathbf{A}$ and $\mathbf{B}$ we have $|\mathbf{A}| \sqsubseteq |\mathbf{B}|$ if and only if conditions (i)–(iii) given there hold. Prove Theorem 9.23.

9.1 Consider the examples given in Exercise 7.3. In which cases is D a consistent subset of P?

9.2 Let $P = \{(a,b) \mid a,b \subseteq \mathbb{N}, a \cap b = \varnothing\}$, ordered as a subset of $\wp(\mathbb{N}) \times \wp(\mathbb{N})$. When are two elements (a,b) and (c,d) consistent? Which elements of P are maximal? Which elements are finite? Prove that P is a domain.

Prove that P is isomorphic to the set of all maps from $\mathbb{N}$ to $\overline{\mathbf{2}}_\perp$ with the pointwise order.

9.3 Give an example of a CPO P and elements $x \notin F(P), k \in F(P)$ such that $x < k$.

9.4 Show that if D is a domain then $D\oplus\mathbf{1}$ is an algebraic lattice. [Hint. First consider the particular case where D itself is an algebraic lattice.]

9.5 Let Q be a non-empty ordered set. A non-empty subset J of Q is called an **ideal** of Q if it is a directed down-set. The set of all ideals of Q is denoted by $\mathcal{I}(Q)$.

(i) Show that if Q is a join semilattice then this concept of ideal agrees with the one introduced in Exercise 7.6.

(ii) Give an example of an ordered set Q such that $\mathcal{I}(Q) \cup \{\varnothing\}$ is not an $\bigcap$-structure on Q.

(iii) Show that, if Q is an ordered set, then $\langle \mathcal{I}(Q); \subseteq\rangle$ is a pre-CPO in which $\bigsqcup$ is given by set union, and $\alpha\colon x \mapsto \downarrow x$ is a (well-defined) order-embedding of Q into $\mathcal{I}(Q)$.

(iv) Prove that, if D is a domain, then the map $\varphi: a \mapsto D_a$, where $D_a := \{ f \in F(D) \mid f \leqslant a \}$, is an order-isomorphism of D onto the set $\mathcal{I}(F(D))$ of ideals of $F(D)$.

9.6 Let Q be an ordered set and R a pre-CPO. Then R is called a **free pre-CPO generated by** Q if there is an order-preserving map $\eta: Q \to R$ such that for each order-preserving map $\varphi: Q \to P$, where P is a pre-CPO, there exists a unique continuous map $\varphi': R \to P$ such that $\varphi' \circ \eta = \varphi$.

The pre-CPO $\langle \mathcal{I}(Q); \subseteq \rangle$ is called the **ideal completion** of Q and is denoted by $\mathbf{IC}(Q)$. By Exercise 9.5(iii), $\alpha: x \mapsto {\downarrow}x$ is an order-embedding of Q into the pre-CPO $\mathbf{IC}(Q)$.

(i) Show that, if both R and R' are free pre-CPOs generated by Q, then $R \cong R'$. [Hint. Use 1.36(4).]

(ii) By (i) we may, up to order-isomorphism, refer to *the* free pre-CPO R generated by Q. Show that $\eta: Q \hookrightarrow R$. [Hint. Try $\varphi = \alpha$ in the definition of R.]

(iii) Show that, if J is an ideal of Q, then $J = \bigsqcup_{x \in J} {\downarrow}x$ in $\mathbf{IC}(Q)$.

(iv) Show that $\mathbf{IC}(Q)$ is generated as a pre-CPO by $\alpha(Q)$, that is, the smallest sub-pre-CPO of $\mathbf{IC}(Q)$ containing $\alpha(Q)$ is $\mathbf{IC}(Q)$ itself.

(v) Prove that $\mathbf{IC}(Q)$ is the free pre-CPO generated by Q. [Hint. Take $\eta = \alpha$ and, given $\varphi: Q \to P$, define $\varphi': \mathbf{IC}(Q) \to P$ by $\varphi': J \mapsto \bigsqcup \varphi(J)$. Use Exercise 8.2 to show that φ' is continuous and (iii) above to show that φ' is unique.]

(vi) Show, directly and without reference to the ideal completion, that the uniqueness assumption on the continuous map φ' in the definition of the free pre-CPO generated by Q may be replaced by the assumption that R is generated as a pre-CPO by $\eta(Q)$.

9.7 Let Q be an ordered set and P a pre-CPO. The first part of the proof of Theorem 8.9 shows that $\langle Q \to P \rangle$ is a pre-CPO and the theorem itself implies that $[\mathbf{IC}(Q) \to P]$ is a pre-CPO. Show that $\varphi \mapsto \varphi'$ is an order-isomorphism from $\langle Q \to P \rangle$ onto $[\mathbf{IC}(Q) \to P]$.

9.8 Given a group G, describe an information system $\langle G, Con, \vdash \rangle$ whose set of elements is order-isomorphic to $\operatorname{Sub} G$ (cf. 9.15).

9.9 Let X be a set and let $F := \{ Y \mid Y \Subset X \}$. Describe information systems $\mathbf{X}$ and $\mathbf{F}$, with X and F as their sets of tokens respectively, such that $|\mathbf{X}| = \wp(X)$ while $|\mathbf{F}| \cong \wp(X)$.

Define approximable mappings r from $\mathbf{X}$ to $\mathbf{F}$ and s from $\mathbf{F}$ to $\mathbf{X}$ such that the induced continuous maps $|r|: |\mathbf{X}| \to |\mathbf{F}|$ and $|s|: |\mathbf{F}| \to |\mathbf{X}|$, as defined in the proof of Proposition 9.30, are mutually inverse order-isomorphisms.

9.10 (i) Given an ordered set P, define an information system $\mathbf{P}$, with P as its token set, such that $|\mathbf{P}| = \mathcal{O}(P)$.

(ii) Let Q be an ordered set. Show that $\mathbf{Q} \trianglelefteq \mathbf{P}$ if and only if $Q \subseteq P$ and Q has the order induced from P.

(iii) Assume that $Q \subseteq P$ has the order induced from P. For $Y \Subset P$ and $b \in Q$ define

$$r: Y \rightsquigarrow b \iff b \leqslant a \text{ for some } a \in Y.$$

Show that r is an approximable mapping from $\mathbf{P}$ to $\mathbf{Q}$ and describe the corresponding continuous map from $\mathcal{O}(P)$ to $\mathcal{O}(Q)$.

9.11 Let $A = \mathbb{Q} \cap (0,1)$. Define $\mathbf{A} = \langle A, Con, \vdash \rangle$ by setting $Y \in Con$ if and only if $Y \Subset A$ and $Y \vdash a$ if and only if $a \leqslant y$ for some $y \in Y$. Describe $F(|\mathbf{A}|)$ and show that the non-finite elements are in order-preserving correspondence with the half-open interval $(0,1]$.

9.12 Let $\mathbf{M} = \langle \mathbb{N} \times \mathbb{N}, Con, \vdash \rangle$, where

$Y \in Con \iff Y \Subset \mathbb{N} \times \mathbb{N}$ such that

$$(m_1, n_1), (m_2, n_2) \in Y \text{ and } m_1 \leqslant m_2 \text{ imply } n_1 \leqslant n_2,$$

$Y \vdash (m,n) \iff$

(a) $n = 1$ and $(\exists m_1 \in \mathbb{N})\, m \leqslant m_1$ and $(m_1, 1) \in Y$, or

(b) $(\exists m_1, m_2 \in \mathbb{N})\, m_1 \leqslant m \leqslant m_2$ and $(m_1, n), (m_2, n) \in Y$.

Show that $\mathbf{M}$ is an information system and that

$$|\mathbf{M}| = \{ \varphi \in (\mathbb{N} \multimap \mathbb{N}) \mid \varphi \text{ is order-preserving} \}.$$

9.13 The triple $\mathbf{A} = \langle A, Con, \vdash \rangle$ is defined as follows:

(a) $A = \{ (x_1, x_2) \times (y_1, y_2) \subseteq \mathbb{R} \times \mathbb{R} \mid x_1 < x_2 \,\&\, y_1 < y_2 \}$ (open rectangles in the plane),

(b) $Y \in Con$ if and only if $Y = \varnothing$ or $Y \Subset A \,\&\, \bigcap Y \neq \varnothing$,

(c) $Y \vdash a$ if and only if $\bigcap Y \subseteq a$.

Verify that $\mathbf{A}$ is an information system. Describe the finite elements and the maximal elements of $|\mathbf{A}|$.

9.14 Let $\mathbf{F}$ be the information system given in Example 9.15(1) and let $\mathbf{M}$ be the information system from the previous example. Show that

$$\Sigma\colon Y \rightsquigarrow (m,n) \iff (1,k_1),\ldots,(m,k_m) \in Y \text{ and } \sum_{i=1}^{m} k_i = n$$

defines an approximable mapping from $\mathbf{F}$ to $\mathbf{M}$. Show that the corresponding continuous map σ from $|\mathbf{F}|$ to $|\mathbf{M}|$ satisfies $\sigma(f)(m) = \sum_{i=1}^{m} f(i)$ for each total map $f\colon \mathbb{N} \to \mathbb{N}$. What is the value of $\sigma(f)$ if f is a partial map and $1 \notin \operatorname{dom}\sigma$?

9.15 Let $(\mathfrak{L}, A)$ and $(\mathfrak{K}, B)$ be algebraic $\bigcap$-structures with $\mathfrak{L} \sqsubseteq \mathfrak{K}$. Define $\iota\colon \mathfrak{L} \to \mathfrak{K}$ by $\iota(S) := \bigcap\{T \in \mathfrak{K} \mid S \subseteq T\}$ and $\pi\colon \mathfrak{K} \to \mathfrak{L}$ by $\pi(T) := T \cap A$ for all $S \in \mathfrak{L}$ and $T \in \mathfrak{K}$.

(i) Show that ι is well defined and that $\pi(\iota(S)) = S$ for all $S \in \mathfrak{L}$ and $\iota(\pi(T)) \subseteq T$ for all $T \in \mathfrak{K}$.

(ii) Show that π is continuous and ι is a continuous order-embedding.

9.16 Let $(\mathfrak{L}, A)$ be an $\bigcap$-structure and let $A^1 := (\{0\} \times A) \cup \{(1,1)\}$. Construct an $\bigcap$-structure $\mathfrak{L} \boxplus \mathbf{1}$ on A^1 which is isomorphic as an ordered set to $\mathfrak{L} \oplus \mathbf{1}$.

Show that the domain constructor $\mathfrak{L} \mapsto \mathfrak{L} \boxplus \mathbf{1}$ is (a) $\sqsubseteq$-preserving, (b) continuous on base sets (and hence continuous).

Let $\mathfrak{L}^0 = \mathfrak{N}$, the unique $\bigcap$-structure based on $\varnothing$, and for all $n \in \mathbb{N}_0$ let $\mathfrak{L}^{n+1} := \mathfrak{L}^n \boxplus \mathbf{1}$. Construct $\mathfrak{L}^n$ for $n = 0,1,2,3,4$ and draw a labelled diagram of each. (For notational convenience let $a_0 = (1,1)$ and $a_{n+1} = (0, a_n)$ for $n \geqslant 0$.) Indicate via arrows the embeddings $\iota_0\colon \mathfrak{L}^0 \to \mathfrak{L}^1$, $\iota_1\colon \mathfrak{L}^1 \to \mathfrak{L}^2$, etc. (see Exercise 9.15).

9.17 Consider the following domain equations. In each case, let $\mathbf{F}$ be the corresponding domain constructor. Draw diagrams for $\mathbf{F}^n(\mathbf{1})$ for $n = 1,2,3,4$ and for the least solution P of the equation. Indicate the order-embedding of $\mathbf{F}^n(\mathbf{1})$ into $\mathbf{F}^{n+1}(\mathbf{1})$ for $n = 1,2,3$ and of $\mathbf{F}^4(\mathbf{1})$ into P. [Either (a) work abstractly, treating $\mathbf{1}$ as an abstract ordered set, or (b) work concretely, replacing $\mathbf{1}$ by $\mathfrak{N}$ and the operators $\oplus_\vee$, $\oplus_\perp$ and $\oplus$ by $\boxplus_\vee$, $\boxplus_\perp$ and $\boxplus$, respectively. While the abstract approach has the advantage that $\mathbf{F}^n(\mathbf{1})$ is easily found, the embeddings can be more elusive. Using the concrete approach, constructing $\mathbf{F}^n(\mathfrak{N})$ can be a notational nightmare but the embeddings are given at once via Exercise 9.15. In order to solve (v) and (vi) concretely, the operator $\boxplus$ must first be defined – see Exercise 9.16 for a particular case.]

(i) $P \cong \mathbf{1} \oplus_\vee P$.

(ii) $P \cong \mathbf{1} \oplus_\perp P$.

(iii) $P \cong P \oplus \mathbf{1}$. (See Exercise 9.16.)

(iv) $P \cong \mathbf{1} \oplus P \oplus \mathbf{1} = (P \oplus \mathbf{1})_\perp$.

(v) $P \cong P \oplus (\mathbf{1} \oplus_\perp \mathbf{1})$.

(vi) $P \cong (\mathbf{1} \oplus_\perp \mathbf{1}) \oplus P$.

(vii) $P \cong (\mathbf{1} \oplus_\perp P)_\perp$.

9.18 Show that $(\mathbb{N} \to \mathbf{2})$ is the least solution to the domain equation $P \cong P \times \mathbf{2}$.

10

Maximality Principles

There are many examples in mathematics of statements which, overtly or covertly, assert the existence of an element maximal in some ordered set (commonly, a family of sets under inclusion). The first section of this chapter addresses the question of the existence of maximal elements. This question cannot be answered without a discussion of Zorn's Lemma and the Axiom of Choice, and this necessitates an excursion into the foundations of set theory. It would be inappropriate to include here a full discussion of the role and status in mathematics of Zorn's Lemma and its equivalents. Rather we seek to complement the treatment in set theory texts of this important topic and, although our account is self-contained, it is principally directed at readers who have previously encountered the Axiom of Choice. *En route*, we provide belated justification for the arguments in 2.39, prove some intrinsically interesting results about ordered sets, and derive the results on prime and maximal ideals on which the representation theory in Chapter 11 rests. Those who do not wish to explore this foundational material but who do wish to study Chapter 11, may without detriment, skip over the first section of this chapter; see 10.15.

Do maximal elements exist? – Zorn's Lemma and the Axiom of Choice

Aside from the treatment of ordinals, ordered sets have traditionally played a peripheral role in introductory set theory courses. It is not unusual for ordered sets only to be formally introduced immediately before Zorn's Lemma is presented. By setting Zorn's Lemma in its order-theoretic context we hope to refute the view prevalent amongst some mathematicians that Zorn's Lemma is psychologically unappealing.

10.1 The Axiom of Choice. A non-empty ordered set may, but need not, possess maximal elements (see 1.23 for examples). In 2.39 we gave an informal argument that a non-empty subset A of an ordered set with (ACC) has a maximal element by picking, one element at a time, a strictly increasing chain of elements of A. By a similar argument we may deduce, more generally, that every CPO has a maximal element. The argument, which appeals to CPO Fixpoint Theorem III (8.23), goes as follows. Let P be a CPO and suppose that every element of P fails to be maximal. This means that, for each $x \in P$, the set $\{y \in P \mid y > x\}$ is non-empty. For each x select a point $y > x$. Since y depends on x,

we label it $F(x)$. Then $x \mapsto F(x)$ defines a map F on P to which we may apply CPO Fixpoint Theorem III. Since $x < F(x)$ for every $x \in P$, we have a contradiction, so P does indeed have a maximal element. We must now come out in the open and make it clear how in both the above arguments we invoke the Axiom of Choice. This asserts that it is possible to find a map which picks one element from each member of a family of non-empty sets. (Only for finitely many sets can this be done without recourse to an axiom additional to those used in standard ZF set theory.) Formally, the **Axiom of Choice** may be stated as follows.

(AC) Given a non-empty family $\mathcal{A} = \{A_i\}_{i \in I}$ of non-empty sets, there exists a **choice function for** $\mathcal{A}$, that is, a map

$$f\colon I \to \bigcup_{i \in I} A_i \text{ such that } (\forall i \in I)\, f(i) \in A_i.$$

To apply (AC) to obtain the map Φ required above we take $I = P$ and $A_x = \{y \in P \mid y > x\}$ for each $x \in P$. Likewise, to formalize the proof of 2.39, we take $I = A$ and $A_x = \{y \in A \mid y > x\}$ for each $x \in A$.

10.2 Maximality axioms. In the same way that (AC) may be regarded as an (optional) axiom of set theory, and deductions made from it, we may take the following statement as a postulate.

(ZL) Let P be a non-empty ordered set in which every non-empty chain has an upper bound. Then P has a maximal element.

We shall also need the following three axioms concerning the existence of maximal elements.

(ZL)′ Let $\mathcal{E}$ be a non-empty family of sets such that $\bigcup_{i \in I} A_i \in \mathcal{E}$ whenever $\{A_i\}_{i \in I}$ is a non-empty chain in $\langle \mathcal{E}; \subseteq \rangle$. Then $\mathcal{E}$ has a maximal element.

(ZL)″ Let P be a CPO. Then P has a maximal element.

(KL) Let P be an ordered set. Then every chain in P is contained in a maximal chain.

Clearly (ZL)′ is just the restriction of (ZL) to families of sets. Our next theorem shows that the five assertions (AC), (ZL), (ZL)′, (ZL)″ and (KL) are all equivalent. It is the implication (AC) ⇒ (ZL) that we refer to as **Zorn's Lemma**. (Some authors use Zorn's Lemma to mean the statement (ZL) instead.) Similarly, the implication (AC) ⇒ (KL) is **Kuratowski's Lemma**.

10.3 Theorem. *The conditions* (AC), (ZL), (ZL)′, (ZL)″ *and* (KL) *are equivalent.*

Proof. We prove (AC) ⇒ (ZL)″ ⇒ (KL) ⇒ (ZL) ⇒ (ZL)′ ⇒ (AC). The first implication has been obtained in 10.1 and the fourth is trivial since (ZL)′ is a restricted form of (ZL).

We next prove that (ZL)″ implies (KL). Take an ordered set P and let $\mathcal{P}$ denote the family of all chains in P which contain a fixed chain C^0, and order this family of sets by inclusion. We claim that $\mathcal{P}$ is a CPO. By 8.11, it suffices to show that every chain in $\mathcal{P}$ has a least upper bound in $\mathcal{P}$. Let $\mathcal{C} = \{C_i\}_{i\in I}$ be a chain in $\mathcal{P}$. If I is empty, then $\bigvee_{\mathcal{P}} \mathcal{C} = C^0$ since C^0 is the bottom of $\mathcal{P}$. Now assume that I is non-empty. Let $C = \bigcup_{i\in I} C_i$. We claim that $C \in \mathcal{P}$, that is, C is a chain. Then, $\bigvee_{\mathcal{P}} \mathcal{C} = C$ by 2.29. Let $x, y \in C$. We are required to show that x and y are comparable. There exist $i, j \in I$ such that $x \in C_i$ and $y \in C_j$. Since $\mathcal{C}$ is a chain, we have $C_i \subseteq C_j$ or $C_j \subseteq C_i$. Assume, without loss of generality, that $C_i \subseteq C_j$. Then x, y both belong to the chain C_j, and hence x and y are comparable, whence C is a chain, as required. We may therefore apply (ZL)″ to $\mathcal{P}$ to obtain a maximal element C^* in $\mathcal{P}$.

The next step is to show that (KL) implies (ZL). Let P be a non-empty ordered set in which every non-empty chain has an upper bound. By (KL), an arbitrarily chosen chain C in P is contained in a maximal chain C^*. By hypothesis, C^* has an upper bound u in P. If u were not a maximal element of P, we could find $v > u$. Clearly $v \notin C^*$, since $u \geqslant c$ for all $c \in C^*$. Thus $C^* \cup \{v\}$ would be a chain strictly containing the maximal chain C^*, ↯.

Finally, we prove that (ZL)′ implies (AC). Consider the ordered set P of partial maps from I to $\bigcup_{i\in I} A_i$ (see 1.10). By identifying maps with their graphs we may regard P as a family of sets ordered by inclusion. Let $\mathcal{E} = \{\pi \in P \mid (\forall i \in \operatorname{dom}\pi)\, \pi(i) \in A_i\}$. Certainly $\mathcal{E} \neq \varnothing$ since the partial map with empty domain vacuously belongs to $\mathcal{E}$. Now let $\mathcal{C} = \{\pi_j\}_{j\in J}$ be a non-empty chain in $\mathcal{E}$. Because $\mathcal{C}$ is a chain, the partial maps π_j are consistent and the union of their graphs is the graph of a partial map, which necessarily belongs to $\mathcal{E}$. By (ZL)′, $\mathcal{E}$ has a maximal element, $f\colon \operatorname{dom} f \to \bigcup A_i$, say. Provided f is a total map, it serves as the required choice function. Suppose f is not total, so that there exists $k \in I \setminus \operatorname{dom} f$. Because $A_k \neq \varnothing$, there exists $a_k \in A_k$. Define g by

$$g(j) = \begin{cases} a_k & \text{if } j = k, \\ f(j) & \text{if } j \in \operatorname{dom} f. \end{cases}$$

Then $g \in \mathcal{E}$ and $g > f$, which contradicts the maximality of f. □

10.4 Inductive ordered sets. An ordered set P in which every non-empty chain has an upper bound is often referred to as **inductive**. We may contrast this with our earlier definition of P being **completely inductive**: every chain in P has a least upper bound. In the definition of 'inductive' it is convenient to exclude the empty chain (which, of course, has every element of P as an upper bound). Thus, (ZL) and, via 8.11, (ZL)″ can be restated as

(ZL) every non-empty inductive ordered set has a maximal element,

(ZL)″ every completely inductive ordered set has a maximal element.

Indeed, it was this version of (ZL)″ that was used in the proof of (ZL)″ ⇒ (KL) above.

In Chapter 8 we gave an (AC)-free proof of CPO Fixpoint Theorem II (8.22). If we are willing to use (AC), or one of its equivalents, then we can give a much simpler proof which relies only on an easy exercise from Chapter 8.

10.5 CPO Fixpoint Theorem II, with (AC). *Assume that P is a CPO and that F is an order-preserving self-map on P. Then F has a least fixpoint.*

Proof. Exercise 8.18 tells us that F has a least fixpoint provided a certain CPO Q has a maximal element. If we assume (AC), then, by (ZL)″, Q does have a maximal element and so F has a least fixpoint. □

This proof is unusual in that it relies upon the axiom (ZL)″. Most proofs which appeal to some form of the Axiom of Choice rely on (ZL) or more commonly the particular case (ZL)′ of (ZL).

10.6 (ZL) in action. The axiom (ZL) (or more usually (ZL)′) is used to assert the existence of an object which cannot be directly constructed, such as

(a) a maximal linearly independent subset in a vector space $V \neq \{0\}$,

(b) a maximal ideal in a ring R,

(c) the choice function sought in the last part of the proof of Theorem 10.3.

Proofs involving (ZL)′ have a distinct sameness. Let X be an object whose existence we wish to establish. We proceed as follows:

(i) take a non-empty family $\mathcal{E}$ of sets ordered by inclusion, in which X is a (hypothetical) maximal element;

(ii) check that (ZL)′ is applicable;

(iii) verify that the maximal element supplied by (ZL)′ has all the properties demanded of X.

Let us review these steps in turn. Choosing $\mathcal{E}$ is usually straightforward. For example, we take $\mathcal{E}$ to be all linearly independent subsets of V in (a) and all proper ideals of R in (b). We then have to exhibit an element of $\mathcal{E}$ to ensure $\mathcal{E} \neq \varnothing$. Again this is easy in our examples: take $\{v\}$, where $0 \neq v \in V$ in (a), and the ideal $\{0\}$ in (b). Notice that $\mathcal{E}$ may be thought of as a family of partial objects, having some of the features X should have.

Now consider (ii). To confirm that (ZL)′ applies, we need to show that the union of a non-empty chain of sets in $\mathcal{E}$ is itself in $\mathcal{E}$. Observe the similarity between the arguments in the proof of (ZL)″ ⇒ (KL) and of Example 7.6. A chain is a special case of a directed set. In an algebraic $\bigcap$–structure $\mathfrak{L}$ we have $\bigcup \mathcal{D} \in \mathfrak{L}$ whenever $\mathcal{D}$ is a directed subset of $\mathfrak{L}$. In each of our examples above, and in many other (ZL) applications, $\mathcal{E}$ is an algebraic $\bigcap$–structure, and it is this fact that ensures success in (ii).

Step (iii) is immediate in examples (a) and (b), but not in (c). In cases where (iii) is non-trivial, argument by contradiction is invariably used. Exercise 10.1 and the proof of Theorem 10.18 below provide illustrations.

Prime and maximal ideals

We introduced lattice ideals in Chapter 2, as part of the development of the algebraic theory of lattices. But we did not take the theory far enough to reveal the importance of ideals, or of their order duals, filters.

Join-irreducible elements served very well as building blocks for finite distributive lattices, but we need an alternative if we are to remove the finiteness restriction. Example 2.43(5) shows that an infinite distributive lattice may have no join-irreducible elements at all. We next introduce a class of ideals which will substitute for join-irreducible elements to yield an extension of Birkhoff's representation theorem for the infinite case. But we shall need (ZL) to show that such ideals exist.

10.7 Definitions. Let L be a lattice. Recall from 2.20 that a non-empty subset J of L is called an **ideal** if

(i) $a, b \in J$ implies $a \vee b \in J$,

(ii) $a \in L$, $b \in J$ and $a \leqslant b$ imply $a \in J$;

it is **proper** if $J \neq L$.

A proper ideal J of L is said to be **prime** if $a, b \in L$ and $a \wedge b \in J$ imply $a \in J$ or $b \in J$. The set of prime ideals of L is denoted $\mathcal{I}_p(L)$. It is ordered by set inclusion. A **filter** and **prime filter** are defined dually and the set of prime filters is denoted by $\mathcal{F}_p(L)$.

A subset J of a lattice L is a prime ideal if and only if $L \setminus J$ is a prime filter – a simple exercise. Thus it is easy to switch between $\mathcal{I}_p(L)$ and $\mathcal{F}_p(L)$. In the sequel we work predominantly with prime ideals. The next two results provide evidence that these may act as a substitute for join-irreducible elements in the representation of distributive lattices.

10.8 Lemma.

(i) *Let L be a finite distributive lattice and let $a \in L$. Then the map $x \mapsto L \setminus {\uparrow}x$ is an order-isomorphism of $\mathcal{J}(L)$ onto $\mathcal{I}_p(L)$ that maps $\{ x \in \mathcal{J}(L) \mid x \leqslant a \}$ onto $\{ I \in \mathcal{I}_p(L) \mid a \notin I \}$.*

(ii) *Let B be a finite Boolean algebra and let $a \in B$. Then the map $x \mapsto B \setminus {\uparrow}x$ is a bijection of $\mathcal{A}(L)$ onto $\mathcal{I}_p(B)$ that maps $\{ x \in \mathcal{A}(L) \mid x \leqslant a \}$ onto $\{ I \in \mathcal{I}_p(B) \mid a \notin I \}$.*

Proof. We prove (i); part (ii) then follows from Lemma 5.3. Lemma 5.11 asserts that ${\uparrow}x$ is a prime filter if and only if $x \in \mathcal{J}(L)$. Hence, taking complements, we have

$$\mathcal{I}_p(L) = \{ L \setminus {\uparrow}x \mid x \in \mathcal{J}(L) \}.$$

We now know that φ maps $\mathcal{J}(L)$ onto $\mathcal{I}_p(L)$. By the dual of 1.30, $x \leqslant y$ if and only if ${\uparrow}x \supseteq {\uparrow}y$, so φ is an order-embedding. □

The following corollary reformulates Proposition 2.45(i) – a critical step in the proof of Birkhoff's representation theorem.

10.9 Corollary. *Let L be a finite distributive lattice and let $a \not\leqslant b$ in L. Then there exists $I \in \mathcal{I}_p(L)$ such that $a \notin I$ and $b \in I$.*

Prime ideals are related to ideals of another important type.

10.10 Definitions. Let L be a lattice and I a proper ideal of L. Then I is said to be a **maximal ideal** if the only ideal properly containing I is L. In other words, I is a maximal ideal if and only if it is a maximal element in $\langle \mathcal{I}(L) \setminus \{L\}; \subseteq \rangle$. A **maximal filter**, more usually known as an **ultrafilter**, is defined dually.

10.11 Theorem. *Let L be a distributive lattice with 1. Then every maximal ideal in L is prime. Dually, in a distributive lattice with 0, every ultrafilter is a prime filter.*

Proof. Let I be a maximal ideal in L and let $a, b \in L$. Assume $a \wedge b \in I$ and $a \notin I$; we require $b \in I$. Define $I_a = \downarrow\{ a \vee c \mid c \in I \}$. Then I_a is an ideal containing I and a (Exercise 2.23). Because I is maximal, we have $I_a = L$. In particular $1 \in I_a$, so $1 = a \vee d$ for some $d \in I$. Then

$$I \ni (a \wedge b) \vee d = (a \vee d) \wedge (b \vee d) = b \vee d.$$

Since $b \leqslant b \vee d$, we have $b \in I$. □

In fact, this theorem is true whether or not L has any bounds – a good exercise for the reader. In a Boolean lattice we can do better.

10.12 Theorem. *Let B be a Boolean lattice and let I be a proper ideal in B. Then the following are equivalent:*

(i) *I is a maximal ideal;*

(ii) *I is a prime ideal;*

(iii) *for all $a \in B$, it is the case that $a \in I$ if and only if $a' \notin I$.*

Proof. Theorem 10.11 gives (i) ⇒ (ii). To prove (ii) ⇒ (iii), note that, for any $a \in B$, we have $a \wedge a' = 0$. Because I is prime, $a \in I$ or $a' \in I$. If both a and a' belong to I then $1 = a \vee a' \in I$, ↯.

Finally we prove that (iii) ⇒ (i). Let J be an ideal properly containing I. Fix $a \in J \setminus I$. Then $a' \in I \subseteq J$, so $1 = a \vee a' \in J$. Therefore $J = B$, which shows that I is maximal. □

10.13 Ultrafilters on a set. Let S be a non-empty set. An ultrafilter of the Boolean lattice $\wp(S)$ is called an **ultrafilter on** S. Such ultrafilters are important in logic.

An ultrafilter on S is said to be **principal** if it is a principal filter, and **non-principal** otherwise. For each $s \in S$, the set $\{ A \in \wp(S) \mid s \in A \}$ is a principal ultrafilter on S, and every principal ultrafilter is of this form. Non-principal ultrafilters prove much more elusive; see Exercise 10.8. All ultrafilters on a finite set are, of course, principal.

Theorem 10.14 characterizes the ultrafilters on a set. It takes the dual of Theorem 10.12 and adds two further useful equivalences. For the proof, do (ii) ⇒ (v) ⇒ (iii) ⇒ (iv) ⇒ (i) ⇒ (ii).

10.14 Theorem. *Let $\mathcal{F}$ be a proper filter in $\wp(S)$. Then the following are equivalent:*

(i) *$\mathcal{F}$ is an ultrafilter;*

(ii) *$\mathcal{F}$ is a prime filter;*

(iii) *for each $A \subseteq S$, either $A \in \mathcal{F}$ or $S \setminus A \in \mathcal{F}$;*

(iv) *for each $B \subseteq S$, if $A \cap B \neq \varnothing$ for all $A \in \mathcal{F}$, then $B \in \mathcal{F}$;*

(v) *given pairwise disjoint sets $A_1, \ldots, A_n$ such that $A_1 \cup \cdots \cup A_n = S$, there exists a unique j such that $A_j \in \mathcal{F}$.*

10.15 Do prime ideals exist? We have so far given singularly few non-trivial examples of prime or maximal ideals. This is unfortunate since in order to remove the finiteness condition in Corollary 10.9 we need a plentiful supply of prime ideals.

To help see what is at stake, consider a Boolean lattice B. Theorem 10.12 implies that a prime ideal in B is just a maximal element of $\langle \mathcal{I}(B) \setminus \{B\}; \subseteq \rangle$. The question of the existence of maximal elements was addressed earlier in this chapter. The discussion there showed that this topic has closer affinities with set theory than with lattice theory, and we should not wish knowledge of it to be a prerequisite for proceeding to Chapter 11. The solution to this apparent dilemma is to present a treatment that operates on two levels. The statements (BPI) and (DPI) introduced below assert the existence of certain prime ideals. On one level, (BPI) and (DPI) may be taken as axioms, whose lattice-theoretic implications we pursue. At a deeper level, we show (in a self-contained account which may be omitted) how (BPI) and (DPI) may be derived from (ZL) (see 10.2). Our remarks in 10.19 will reveal that the difference between these two philosophies is less than might appear.

10.16 (DPI) and (BPI). We consider the following assertions, which embody the existence statements we shall require.

(DPI) Given a distributive lattice L and an ideal J and a filter G of L such that $J \cap G = \varnothing$, there exist $I \in \mathcal{I}_p(L)$ and $F = L \setminus I \in \mathcal{F}_p(L)$ such that $J \subseteq I$ and $G \subseteq F$.

(BPI) Given a proper ideal J of a Boolean lattice B, there exists $I \in \mathcal{I}_p(B)$ such that $J \subseteq I$.

The remainder of this section employs (ZL). The proof of the first result is typical of Zorn's Lemma arguments (see 10.6).

10.17 Theorem. (ZL) *implies* (BPI).

Proof. Let B be a Boolean lattice and J be a proper ideal of B. We apply the special case (ZL)′ of (ZL) stated in 10.2 to the set

$$\mathcal{E} := \{ K \in \mathcal{I}(B) \mid B \neq K \supseteq J \},$$

ordered by inclusion. The set $\mathcal{E}$ contains J, and so is non-empty. Let $\mathcal{C} = \{ K_\lambda \mid \lambda \in \Lambda \}$ be a chain in $\mathcal{E}$. We require $K := \bigcup_{\lambda \in \Lambda} K_\lambda \in \mathcal{E}$. Certainly $K \neq B$ (why?), $K \supseteq J$ and K is a down-set. It remains

to prove that $a, b \in K$ implies $a \vee b \in K$. For some $\lambda, \mu \in \Lambda$, we have $a \in K_\lambda$ and $b \in K_\mu$. Since $\mathcal{C}$ is a chain, we may assume without loss of generality that $K_\lambda \subseteq K_\mu$. But then a, b both belong to K_μ, so $a \vee b \in K_\mu \subseteq K$. The maximal element of $\mathcal{E}$ provided by (ZL)′ is just the maximal ideal we require (by 10.12). □

The corresponding result for distributive lattices, often referred to as the **Prime Ideal Theorem**, is slightly more complicated to prove, but essentially just combines the techniques of 10.11 and 10.17.

10.18 Theorem. (ZL) *implies* (DPI).

Proof. We take L, G and J as in the statement (DPI) and define

$$\mathcal{E} = \{ K \in \mathcal{I}(L) \mid K \supseteq J \text{ and } K \cap G = \varnothing \}.$$

An argument mildly more complicated than the one in 10.17 shows that $\langle \mathcal{E}; \subseteq \rangle$ has a maximal element I. It remains to prove that I is prime. To do this we adapt the proof of 10.11, which is the case $G = \{1\}$. Suppose $a, b \in L \setminus I$ but $a \wedge b \in I$. Because I is maximal, any ideal properly containing I is not in $\mathcal{E}$. Consequently $I_a = {\downarrow}\{ a \vee c \mid c \in I \}$ (the smallest ideal containing I and a) intersects G. Therefore there exists $c_a \in I$ such that $a \vee c_a$ is above an element of G and hence is itself in G, because G is an up-set. Similarly we can find $c_b \in I$ such that $b \vee c_b \in G$. Now consider

$$(a \wedge b) \vee (c_a \vee c_b) = ((a \vee c_a) \vee c_b) \wedge ((b \vee c_b) \vee c_a).$$

The right-hand side is in G, since G is a filter, while the left is in I, since I is an ideal. This gives $I \cap G \neq \varnothing$, ↯. □

10.19 A choice of axioms. Some further comments on the relationship between (ZL), (BPI) and (DPI) are appropriate, although we cannot attempt to justify all the assertions we make. Consult [18] and [25] for more details, related results and references.

When L is a distributive lattice with 1, we may take $G = \{1\}$ in (DPI). Then (DPI) implies the existence of a maximal ideal of L containing a given proper ideal J. So (DPI), restricted to Boolean lattices, yields (BPI) as a special case. Much less obviously, (BPI) ⇒ (DPI). This is proved by constructing an embedding of a given distributive lattice into a Boolean lattice, to which (BPI) is applied. Hence (BPI) and (DPI) are equivalent.

We proved in 10.3 that (ZL) is equivalent to the Axiom of Choice, (AC), and remarked that many other equivalents of (AC) (set-theoretic and otherwise) were known. It turns out that one such statement is:

(DMI) every distributive lattice with 1, which has more than one element, contains a maximal ideal.

It is easy to derive (DMI) from (ZL) (see Exercise 10.1). Conversely it can be proved that (AC) can be derived from (DMI), applied to a suitable lattice of sets. A proof that (DMI) ⇒ (DPI) is indicated in Exercise 10.10.

By contrast, (BPI) and (DPI) belong to a family of conditions known to be equivalent to the choice principle $(\mathrm{AC})_F$ (asserting that every family of non-empty *finite* sets has a choice function). It is known that $(\mathrm{AC})_F$ is strictly weaker than (AC), so that it is not true that (DPI) implies (DMI). However, $(\mathrm{AC})_F$ is not derivable within traditional Zermelo–Fraenkel set theory. To obtain results such as (DPI) and (BPI) some additional axiom must be added, and whether this is (AC), (ZL), or even (DPI) itself, is a matter of choice. Thus our suggestion that readers ignorant of (ZL) should take (DPI) as a hypothesis had a sound logical basis.

There are many instances in logic and topology of the construction of ultrafilters in Boolean lattices (especially powerset lattices). The proof via filters of Tychonoff's Theorem is an example. Accordingly we introduce

(BUF) Given a proper filter G of a Boolean lattice B, there exists $F \in \mathcal{F}_p(B)$ such that $G \subseteq F$.

A proper filter (an ultrafilter) of a Boolean lattice B is a proper ideal (a maximal ideal) of B^∂ (which is also a Boolean lattice). Thus the statements (BPI) and (BUF) are equivalent.

We sum up the connections between the various conditions below.

$$\begin{array}{ccccccc} (\mathrm{AC}) & \iff & (\mathrm{ZL}) & \implies & (\mathrm{BPI}) & \iff & (\mathrm{BUF}) \\ \Updownarrow & & & & \Updownarrow & & \Updownarrow \\ (\mathrm{DMI}) & & \implies & & (\mathrm{DPI}) & \iff & (\mathrm{AC})_F \end{array}$$

Powerset algebras and down-set lattices revisited

Our representation theorems in the finite case show that any finite Boolean algebra is isomorphic to a powerset and any finite distributive lattice is isomorphic to the lattice of down-sets of an ordered set. We cannot expect these statements to remain universally true when we delete the word 'finite': we gave in 4.18 an example of a Boolean algebra which is not isomorphic to a powerset algebra.

However, we can use the results of the preceding section to show, easily, that every distributive lattice does have a concrete representation as a lattice of sets, or, in a Boolean case, an algebra of sets. At the end of the section we characterize among Boolean algebras and bounded distributive lattices those which are, respectively, powerset algebras and down-set lattices.

10.20 Lemma. *Let L be a lattice and let $X = \mathcal{I}_p(L)$. Then the map $\eta : L \to \wp(X)$ defined by*

$$\eta\colon a \longmapsto X_a := \{\, I \in \mathcal{I}_p(L) \mid a \notin I \,\}$$

is a lattice homomorphism.

Proof. We have to show that $X_{a\vee b} = X_a \cup X_b$ and $X_{a\wedge b} = X_a \cap X_b$, for all $a, b \in L$. Take $I \in \mathcal{I}_p(L)$. Since I is an ideal,

$$a \vee b \in I \text{ if and only if } a \in I \text{ and } b \in I$$

and, since I is prime,

$$a \wedge b \in I \text{ if and only if } a \in I \text{ or } b \in I.$$

Thus we have

$$\begin{aligned} X_{a\vee b} &= \{\, I \in \mathcal{I}_p(L) \mid a \vee b \notin I \,\} \\ &= \{\, I \in \mathcal{I}_p(L) \mid a \notin I \text{ or } b \notin I \,\} \\ &= X_a \cup X_b. \end{aligned}$$

Similarly, $X_{a\wedge b} = X_a \cap X_b$. □

We would like the map η above to give a faithful copy of L in the lattice $\wp(I_p(L))$. We certainly cannot prove this without the additional hypothesis of distributivity, because a lattice of sets must be distributive. Theorem 10.21 shows that (DPI) is exactly what is needed to ensure that a distributive lattice L has enough prime ideals for $\eta\colon L \to \wp(\mathcal{I}_p(L))$ to be an embedding.

10.21 Theorem. *Let L be a lattice. Then the following are equivalent:*

(i) *L is distributive;*

(ii) *given an ideal J of L and a filter G of L with $J \cap G = \varnothing$, there exists a prime ideal I such that $J \subseteq I$ and $I \cap G = \varnothing$;*

(iii) *given $a, b \in L$ with $a \not\leq b$, there exists a prime ideal I such that $a \notin I$ and $b \in I$;*

(iv) *the map $\eta\colon a \mapsto X_a := \{\, I \in \mathcal{I}_p(L) \mid a \notin I \,\}$ is an embedding of L into $\wp(\mathcal{I}_p(L))$;*

(v) *L is isomorphic to a lattice of sets.*

Proof. The implications (iv) ⇒ (v) ⇒ (i) are trivial and (i) ⇒ (ii) is the statement that (DPI) holds for L. To prove (ii) ⇒ (iii) just take $J = {\downarrow}b$ and $G = {\uparrow}a$ in (DPI).

Since η is a homomorphism, it is order-preserving. To prove that (iii) ⇒ (iv) it is enough to show that $a \not\leq b$ implies $X_a \not\subseteq X_b$. This is true since the prime ideal I supplied by (iii) belongs to $X_a \setminus X_b$. □

For Boolean algebras we have the following result.

10.22 Theorem. *Let B be a Boolean algebra. Then*

(i) *given a proper ideal J of B, there exists a maximal ideal $I \in \mathcal{I}_p(B)$ with $J \subseteq I$,*

(ii) *given $a \neq b$ in B, there exists a maximal ideal $I \in \mathcal{I}_p(B)$ such that I contains one and only one of a and b,*

(iii) *the map $\eta\colon a \mapsto X_a := \{\, I \in \mathcal{I}_p(B) \mid a \notin I \,\}$ is a Boolean algebra embedding of B into the powerset algebra $\wp(\mathcal{I}_p(L))$.*

Proof. We first show that (i), which is the statement (BPI) for B, implies (ii). Take $a, b \in B$ with $a \neq b$. Without loss of generality we may assume $a \not\leq b$ and this gives $1 \neq a' \vee b$. Apply (i) with $J = {\downarrow}(a' \vee b)$. Any prime ideal I containing J contains b, but, by Theorem 10.12, not a.

The map $\eta\colon a \mapsto X_a := \{\, I \in \mathcal{I}_p(B) \mid a \notin I \,\}$ is a lattice homomorphism. Also $X_0 = \varnothing$ because each prime ideal contains 0 and $X_1 = X$ since each prime ideal is proper. So, by 4.17, η is a Boolean algebra homomorphism. Since (ii) holds, η is also one-to-one. □

Our next objective is the characterization of those Boolean algebras which are powerset algebras. Our starting point is the observation that, just as lattices of sets are distributive, complete lattices of sets, and in particular powersets, must satisfy a very strong distributive law.

10.23 Infinite distributive laws. A complete lattice L is said to be **completely distributive** if, for any doubly indexed subset $\{x_{ij}\}_{i\in I, j\in J}$ of L, we have

$$\text{(CD)} \qquad \bigwedge_{i\in I}\Big(\bigvee_{j\in J} x_{ij}\Big) = \bigvee_{\alpha\colon I\to J}\Big(\bigwedge_{i\in I} x_{i\alpha(i)}\Big).$$

The formulation of (CD) is simply a formal way of saying that any meet of joins is converted into the join of all possible elements obtained by taking the meet over $i \in I$ of elements x_{ik} where k depends on i; the functions $\alpha\colon I \to J$ do the job of picking out the indices k. The law (CD) can be shown to be self-dual, as distributivity is (recall 4.3); that is, L satisfies (CD) if and only if L^∂ does.

Certainly any powerset $\langle \wp(X); \subseteq \rangle$ satisfies (CD). So does any complete lattice of sets, and in particular any lattice $\langle \mathcal{O}(P); \subseteq \rangle$, where P is an ordered set.

As an instance of (CD), obtained by taking $I = \{1, 2\}$, $x_{1j} = x$ and $x_{2j} = y_j$ for all $j \in J$, we have the **Join-Infinite Distributive Law**: for any subset $\{y_j\}_{j\in J}$ of L and any $x \in L$,

$$\text{(JID)} \qquad x \wedge \bigvee_{j\in J} y_j = \bigvee_{j\in J} x \wedge y_j.$$

The dual condition is the **Meet-Infinite Distributive Law**, (MID), and it too holds in any completely distributive lattice.

10.24 Theorem. *Let B be a Boolean algebra. Then the following are equivalent:*

(i) *$B \cong \wp(X)$ for some set X;*

(ii) *B is complete and atomic;*

(iii) *B is complete and completely distributive.*

Proof. Certainly (i) implies both (ii) and (iii). Now assume (ii). By exactly the same argument as in the proof of Theorem 5.5, the map

$$\eta\colon a \longmapsto \{x \in \mathcal{A}(B) \mid x \leqslant a\}$$

is a Boolean algebra isomorphism mapping B onto $\wp(\mathcal{A}(B))$. Thus (ii) implies (i).

To complete the proof we show that (iii) implies (ii). We apply (CD) with $I = B$ and $J = \{\pm 1\}$, with

$$x_{ij} = \begin{cases} i & \text{if } j = 1, \\ i' & \text{if } j = -1. \end{cases}$$

Note that, for any i, we have $\bigvee_{j\in J} x_{ij} = i \vee i' = 1$. Therefore, by (CD),

$$\bigvee_{\alpha\colon I\to J} \left(\bigwedge_{i\in I} x_{i\alpha(i)} \right) = 1.$$

Let $y \in B$. Then by (JID) we have

$$\bigvee_{\alpha\colon I\to J} \left(y \wedge \bigwedge_{i\in I} x_{i\alpha(i)} \right) = y.$$

We shall show that $z_\alpha := y \wedge \bigwedge_{i\in I} x_{i\alpha(i)}$ is an atom whenever it is non-zero. Suppose $0 < u \leqslant z_\alpha$. Then $u \leqslant x_{u\alpha(u)}$. This forces $\alpha(u) = 1$ since otherwise $u \leqslant u'$ in contradiction to $u \neq 0$. But $\alpha(u) = 1$ gives $x_{u\alpha(u)} = u$, so that $u \geqslant z_\alpha$. Therefore $u = z_\alpha$, so that $z_\alpha \in \mathcal{A}(B)$, as claimed. □

Characterizing down-set lattices requires substantially more work. We note that any lattice $\mathcal{O}(P)$ is algebraic, and begin by extending our knowledge of algebraic lattices.

Exercise 4.11 seeks a proof that $\langle \mathbb{N}_0; \preccurlyeq \rangle$ fails (JID). By contrast, any bounded distributive lattice which satisfies (ACC) (respectively (DCC)) does satisfy (JID) (respectively (MID)); for the proof use Theorem 2.41(i). The next proposition generalizes this.

10.25 Proposition. *Let L be an algebraic lattice.*

(i) *Meet distributes over directed joins in L, that is,*

$$x \wedge \bigsqcup \{y_i \mid i \in I\} = \bigsqcup \{x \wedge y_i \mid i \in I\}.$$

(ii) *If L is distributive, then it satisfies* (JID).

Proof. (i) Let $D = \{y_i\}_{i\in I}$ be directed. It is easy to see that $\{x \wedge y_i\}_{i\in I}$ is also directed. Note that

$$x \wedge \bigsqcup \{y_i \mid i \in I\} \geqslant \bigsqcup \{x \wedge y_i \mid i \in I\},$$

since the left-hand side is an upper bound for $\{x \wedge y_i\}_{i\in I}$. Suppose for a contradiction that the inequality is strict. Because L is algebraic, this implies that there exists $k \in F(L)$ such that

$$k \leqslant x \wedge \bigsqcup \{y_i \mid i \in I\} \quad \text{but} \quad k \not\leqslant \bigsqcup \{x \wedge y_i \mid i \in I\}.$$

Then $k \leqslant x$ and $k \leqslant \bigsqcup D$, from which we get $k \leqslant y_j$ for some j. But then $k \leqslant x \wedge y_j \leqslant \bigsqcup\{x \wedge y_i\}$, $\lightning$.

For (ii) we draw on Exercise 7.5: for any non-empty set S,

$$\bigvee S = \bigsqcup \{\bigvee F \mid \varnothing \neq F \Subset S\}.$$

Since meet distributes over directed joins and over finite joins it is easy to show that meet distributes over arbitrary joins. We leave the checking of the details as an exercise. □

We have seen various analogues, (CD), (JID) and (MID), of the distributive laws (D) and (D)$^\partial$ that a complete lattice may satisfy. We now turn to analogues of join- and meet-irreducible elements, and introduce a new atomicity condition.

10.26 Definitions. An element a of a complete lattice is called **completely join-irreducible** if $a = \bigvee S$ implies that $a \in S$, for every subset S of L; in particular, $a \neq 0$ (take $S = \varnothing$). The element a is called **completely join-prime** if $a \leqslant \bigvee S$ implies $a \leqslant s$ for some $s \in S$. **Completely meet-irreducible** and **completely meet-prime** are defined dually.

It is easy to see that every completely join-prime element is completely join-irreducible. By a parallel argument to that used to prove (i) $\Rightarrow$ (ii) in Lemma 5.11, in the presence of (JID) every completely join-irreducible element is completely join-prime. We denote the set of completely join-prime elements in L by $\mathcal{J}_p(L)$.

We say that a lattice L is **weakly atomic** if, given $x < y$ in L, there exist $a, b \in L$ such that $x \leqslant b \prec a \leqslant y$. Note that this condition is satisfied in any down-set lattice (use Exercise 1.12) and note too that it is self-dual.

Compare the proof of (ii) below with that of (the dual version of) Proposition 2.45. Note how (ZL) is used to guarantee the existence of a maximal element in the absence of (ACC). We leave the proof of (i) as an exercise (see Exercise 10.13).

10.27 Proposition. *Let L be a complete lattice.*

(i) *Assume that L is algebraic. Then the completely meet-irreducible elements are meet-dense in L.*

(ii) *Assume that L satisfies* (JID) *and is weakly atomic. Then the completely meet-irreducible elements are meet-dense in L.*

Proof. (ii) By the dual of Exercise 2.39, to prove meet-density of a set Q it suffices to show that if $s, t \in L$ with $t > s$ then there exists $m \in Q$ with $m \geqslant s$ and $m \ngeq t$.

Assume L satisfies (JID) and is weakly atomic. Take $t > s$. Then there exist $p, q \in L$ such that $t \geqslant q \succ p \geqslant s$. Define

$$P = \{\, x \in L \mid x \geqslant p \text{ and } x \ngeq q \,\}.$$

The set P contains p and so $P \neq \varnothing$. Let C be a non-empty chain in P, and suppose for a contradiction that $\bigvee C \notin P$. This means that $\bigvee C \geqslant q$. Invoking (JID) we have $\bigvee_{x \in C} (x \wedge q) = q$. If we had $x \wedge q \leqslant p$ for all $x \in C$ then $\bigvee_{x \in C} (x \wedge q) \leqslant p$, $\lightning$. Pick $x \in C$ such that $x \wedge q \nleqslant p$. Then, using the contrapositive of the Connecting Lemma, $p < (x \wedge q) \vee p$. By distributivity, which is implied by (JID), $(x \wedge q) \vee p = (x \vee p) \wedge (q \vee p) = x \wedge q \leqslant q$. Hence, because $q \succ p$, we have $x \wedge q = q$, $\lightning$. By (ZL), P has a maximal element, m say, and this satisfies $m \geqslant p$ and $m \ngeq q$. By transitivity, $m \geqslant s$ and $m \ngeq t$.

Finally suppose for a contradiction that $m = \bigwedge S$ but that $m \neq y$ for every $y \in S$. Because m is maximal in P, every $y \in S$ lies outside P. But $y \geqslant m \geqslant p$, so we must have $y \geqslant q$ for all $y \in S$. But then $m = \bigwedge S \geqslant q$, $\lightning$. Hence m is completely meet-irreducible. □

Our next result about algebraic lattices also requires (ZL). This result does not imply that (ii) in Proposition 10.27 follows from (i) because L is not required to be distributive.

10.28 Proposition. *Every algebraic lattice L is weakly atomic.*

Proof. Let $x < y$ in L. Recall from Exercise 7.7 that $K := [x, y]$ is an algebraic lattice. We claim that if $a \in K$ is finite and $x < a$, then there exists $b \in K$ such that $x \leqslant b \prec a$. To do this we apply (ZL) to the set

$$P = \{\, b \in K \mid x \leqslant b < a \,\}$$

and show that a maximal element b of P is a lower cover of a. We leave the verification as an exercise. □

At last we have the promised characterization of down-set lattices.

10.29 Theorem. *Let L be a lattice. Then the following are equivalent:*

(i) *L is isomorphic to $\mathcal{O}(P)$ for some ordered set P;*

(ii) *L is isomorphic to a complete lattice of sets;*

(iii) *L is distributive and both L and L^∂ are algebraic;*

(iv) *L is complete, L satisfies* (JID) *and the completely join-irreducible elements are join-dense;*

(v) *the map $\eta\colon x \mapsto \{x \in \mathcal{J}_p(L) \mid x \leqslant a\}$ is an isomorphism from L onto $\mathcal{O}(\mathcal{J}_p(L))$;*

(vi) *L is completely distributive and L is algebraic;*

(vii) *L is complete, satisfies* (JID) *and* (MID) *and is weakly atomic.*

Proof. We have the following implications:

(i)	$\Longrightarrow$ (ii)	(trivially),
(ii)	$\Longrightarrow$ (iii)	(trivially),
(iii)	$\Longrightarrow$ (iv)	(by the duals of 10.25 and 10.27),
(iv)	$\Longrightarrow$ (v)	(cf. 5.12 and see 10.26),
(v)	$\Longrightarrow$ (i)	(trivially),
(ii)	$\Longrightarrow$ (vi)	(using 7.21),
(vi)	$\Longrightarrow$ (vii)	(by 10.28),
(vii)	$\Longrightarrow$ (iv)	(by the dual of 10.27(ii)).

These implications show that all seven conditions are equivalent. □

Exercises

Exercises from the text: Complete the proof of Lemma 10.8. Show that 10.11 holds without the assumption that L has any bounds. Fill in the details of the proofs of 10.25(ii) and 10.28.

10.1 Let L be a lattice. Recall that $\mathcal{I}(L)$ denotes the set of all ideals of L. Deduce from (ZL)$'$ that if L has a top then any ideal J in L with $J \neq L$ is contained in an ideal I which is maximal in $\langle \mathcal{I}(L) \setminus \{L\}; \subseteq \rangle$.

10.2 Let $\langle P; \leqslant \rangle$ be an ordered set. By applying (ZL)$''$ to an appropriate family $\mathcal{E}$ of partial orders (regarding an order relation on P as a subset of $P \times P$), show that $\leqslant$ has a linear extension. (This is Szpilrajn's Theorem; you will need its finite version, given in Exercise 1.29(ii), to prove that the maximal element of $\mathcal{E}$ supplied by (ZL)$''$ is a chain.)

10.3 Let L be a lattice. Prove that $\mathcal{I}_p(L) \cong \mathcal{F}_p(L)^\partial$.

10.4 Let L and K be bounded lattices and $f\colon L \to K$ a $\{0,1\}$-homomorphism.

(i) Show that $f^{-1}(0)$ is an ideal in L.

(ii) Show that, if $K = \mathbf{2}$, then $f^{-1}(0)$ is a prime ideal in L.

(iii) Let I be a prime ideal in L. Define $f_I\colon L \to \mathbf{2}$ by

$$f_I(a) = \begin{cases} 1 & \text{if } a \notin I, \\ 0 & \text{if } a \in I. \end{cases}$$

Prove that f_I is a $\{0,1\}$-homomorphism.

(iv) Let X denote the set of all $\{0,1\}$-homomorphisms from L to $\mathbf{2}$, ordered pointwise. Show that there is an order-isomorphism between $\mathcal{I}_p(L)$, ordered by inclusion, and X^∂. [cf. Exercise 5.20 and Corollary 10.9.]

10.5 Let L and K be bounded lattices.

(i) Prove that every ideal of $L \times K$ is of the form $I \times J$ where I is an ideal of L and J is an ideal of K.

(ii) Let I be a prime ideal of L and let J be a prime ideal of K. Show that $I \times K$ and $L \times J$ are prime ideals in $L \times K$ and that every prime ideal of $L \times K$ is of this form.

10.6 Find all prime ideals (prime filters) in $\langle \mathbb{N}_0; \mathrm{lcm}, \gcd \rangle$. (See Lemma 10.8 and Exercise 5.15.) Describe the order on the set of prime filters of $\langle \mathbb{N}_0; \mathrm{lcm}, \gcd \rangle$.

10.7 Let $B = \{b_1, b_2, \ldots, b_n, \ldots\}$ be a countable Boolean lattice. Without using (ZL) or an equivalent, prove that B satisfies (BUF). [Hint. Consider $\bigcup_{n \geqslant 0} G_n$ where $G_0 = G$ and, for $n \geqslant 0$, $G_n = G_{n-1}$ if $b_n' \in G_{n-1}$ and G_n is the smallest filter containing G_{n-1} and b_n otherwise.]

10.8 Let S be an infinite set.

(i) Let $\mathcal{G}$ be the set of cofinite subsets of S.

(a) Show that $\mathcal{G}$ is a filter in $\wp(S)$.

(b) Show that if $\mathcal{F}$ is a proper filter in $\wp(S)$ and $\mathcal{G} \subseteq \mathcal{F}$, then $\mathcal{F}$ is not principal.

(ii) Assume that (BUF) holds. Show that there is a non-principal ultrafilter on S.

(iii) Assume that (ZL) holds. Prove directly from (ZL), or (ZL)$'$, that there is a non-principal ultrafilter on S.

10.9 A filter G of a lattice L is called **distributive** if it satisfies

$$(\forall a, b, c \in L)\, a \vee b, a \vee c \in G \Longrightarrow a \vee (b \wedge c) \in G.$$

(i) Find all distributive filters of $\mathbf{N}_5$ and $\mathbf{M}_5$.

(ii) Prove that L is distributive if and only if every filter of L is distributive.

(iii) Let L be a lattice and G a filter in L. Prove that the following are equivalent:

(a) G is a distributive filter;

(b) every ideal I which is a maximal element of the set $\{K \in \mathcal{I}(L) \mid K \cap G = \varnothing\}$ is a prime ideal;

(c) G is an intersection of prime filters, that is, $G = \bigcap_{i \in I} F_i$ for some family $\{F_i\}_{i \in I}$ of prime filters.

[Hint. The implication (a) $\Rightarrow$ (b) is a refinement of the second portion of the proof of Theorem 10.18, while the implication (b) $\Rightarrow$ (c) is an easy consequence of Exercise 10.9(i). (Note that (ZL) is required.)]

10.10 Let L be a distributive lattice, J an ideal and G a filter of L such that $J \cap G = \varnothing$.

(i) Suppose that there is an onto homomorphism $f\colon L \to K$ such that:

(a) $|K| \geqslant 2$ and K has a 0 and a 1, and

(b) $J \subseteq f^{-1}(\{0\})$ and $G \subseteq f^{-1}(\{1\})$.

Show that (DMI) applied to K yields (DPI) for L.

(ii) Let θ_J be the congruence on L defined in Exercise 6.4 and let θ^G be defined dually. Let $j\colon L \to L/\theta_J$ be the natural map. Show that $j(G)$ is a filter of L/θ_J. Consider the natural map $g\colon L/\theta_J \to (L/\theta_J)/\theta^{j(G)}$. (Note: L/θ_J is distributive.) Show that $f := g \circ j\colon L \to (L/\theta_J)/\theta^{j(G)}$ satisfies (a) and (b) of (i).

Thus (DMI) implies (DPI).

10.11 Let B be the family of all finite unions of subintervals of $\mathbb{R}$ of the form: $(-\infty, a)$, $[a, b)$, and $[b, \infty)$, where $-\infty < a < b < \infty$, together with $\varnothing$. Show that B is a Boolean subalgebra of the powerset algebra $\wp(\mathbb{R})$ and that B has no atoms.

10.12 Let B be an atomic Boolean lattice. Prove that B is weakly atomic.

10.13 Show that in an algebraic lattice L the completely meet-irreducible elements are meet-dense. [Hint. Use the dual of Exercise 2.39. Take $t > s$ and apply (ZL) to the set

$$P = \{\, a \in L \mid a \geqslant s \text{ and } a \ngeq k \,\}$$

where $k \in K(L)$ with $t \geqslant k$ and $s \ngeq k$.]

10.14 Let L be a complete lattice. If $a, b \in L$ satisfy $L = {\uparrow}a \cup {\downarrow}b$, then (a, b) is called a **completely prime pair**. This exercise justifies the terminology.

(i) Prove that if (a, b) is a completely prime pair, then a is completely join-prime and b is completely meet-prime.

(ii) Prove that a is completely join-prime in L if and only if there exist $b \in L$ such that (a, b) is a completely prime pair.

10.15 Let C be a closure operator on a set X and let $\mathfrak{L}_C$ be the corresponding topped $\bigcap$-intersection structure on X.

(i) Show that $A \in \mathfrak{L}_C$ is completely join-irreducible if and only if there exists $a \in X$ such that

(a) $A = C(\{a\})$, and

(b) the set $\{\, x \in X \mid x \in C(\{a\}) \text{ and } a \notin C(\{x\}) \,\}$ belongs to $\mathfrak{L}_C$.

(ii) Let $A, B \in \mathfrak{L}_C$. Show that (A, B) is a completely prime pair in $\mathfrak{L}_C$ if and only if there exists $a \in X$ such that $A = C(\{a\})$ and $B = \{\, x \in X \mid a \notin C(\{x\}) \,\}$.

11

Representation: the General Case

This chapter continues the study of Boolean algebras and distributive lattices. It improves on the representations by sets given in Theorems 10.21 and 10.22, and presents a duality theory as powerful as that we obtained in Chapter 5 in the finite case.

Stone's representation theorem for Boolean algebras

We showed in Chapter 5 that every finite Boolean algebra is isomorphic to some powerset algebra. Finiteness is essential here: we saw in Example 4.18(2) that the finite-cofinite algebra $\mathrm{FC}(\mathbb{N})$ is not isomorphic to a powerset algebra and Theorem 10.24 showed how special the powerset algebras are. However, it is true that any Boolean algebra B is isomorphic to a *subalgebra* of a powerset algebra (Theorem 10.22). We now refine this result, by describing precisely which subalgebra this is.

11.1 The prime ideal space of a Boolean algebra. Let B be a Boolean algebra. Theorem 10.22 tells us that the map

$$\eta\colon a \longmapsto X_a := \{\, I \in \mathcal{I}_p(B) \mid a \notin I \,\}$$

is a Boolean algebra embedding of B into $\wp(\mathcal{I}_p(B))$. What we seek is a characterization of the image of the embedding η. The description of $\operatorname{im}\eta$ has to be in terms of additional structure on the set of prime ideals. A topological structure on $\mathcal{I}_p(B)$ is exactly what we need. A **topology** on a set X is a family of subsets of X containing X and $\varnothing$ and closed under arbitrary unions and finite intersections. Readers whose knowledge of topology is rusty or non-existent will find an outline of the concepts and results we need in Appendix A. References such as A.1 are to this appendix.

The family of clopen subsets of a topological space $\langle X; \mathcal{T} \rangle$ forms a Boolean algebra. This suggests that we might try to impose a topology $\mathcal{T}$ on $\mathcal{I}_p(B)$ so that $\operatorname{im}\eta$ is characterized as the family of clopen subsets of the topological space $\langle \mathcal{I}_p(B); \mathcal{T} \rangle$. It is certainly necessary that

$$X_a := \{\, I \in \mathcal{I}_p(B) \mid a \notin I \,\}$$

be in $\mathcal{T}$ for each $a \in B$. The family $\mathcal{B} := \{\, X_a \mid a \in B \,\}$ is not a topology because it is not closed under the formation of arbitrary unions. We have to define $\mathcal{T}$ on $\mathcal{I}_p(B)$ as follows:

$$\mathcal{T} := \{\, U \subseteq \mathcal{I}_p(B) \mid U \text{ is a union of members of } \mathcal{B} \,\}.$$

In the terminology of A.3, the family $\mathcal{B}$ is a basis for $\mathcal{T}$ (which is indeed a topology). The topological space $\langle \mathcal{I}_p(B); \mathcal{T}\rangle$ is called the **prime ideal space** or **dual space** of B. Let $X := \mathcal{I}_p(B)$. Each element of $\mathcal{B}$ is clopen in X, because $X \setminus X_a = X_{a'}$ and so $X \setminus X_a$ is open. To prove that *every* clopen subset of $\langle X; \mathcal{T}\rangle$ is of the form X_a, we need further information about the prime ideal space.

11.2 Proposition. *Let B be a Boolean algebra. Then the prime ideal space $\langle \mathcal{I}_p(B); \mathcal{T}\rangle$ is compact.*

Proof. Let $\mathcal{U}$ be an open cover of $X := \mathcal{I}_p(B)$. We have to show that there exist finitely many members of $\mathcal{U}$ whose union is X. Every open set is a union of sets X_a and we may therefore assume without loss of generality that $\mathcal{U} \subseteq \mathcal{B}$. Write $\mathcal{U} = \{ X_a \mid a \in A\}$, where $A \subseteq B$. Let J be the smallest ideal containing A, that is (by Exercise 2.22),

$$J = \{ b \in B \mid b \leqslant a_1 \vee \cdots \vee a_n \text{ for some } a_1, \ldots, a_n \in A\}.$$

If J is not proper, then $1 \in J$ and we have $a_1 \vee \cdots \vee a_n = 1$ for some finite subset $\{a_1, \ldots, a_n\}$ of A. Then $X = X_1 = X_{a_1 \vee \ldots \vee a_n} = X_{a_1} \cup \cdots \cup X_{a_n}$ and $\{X_{a_1}, \ldots, X_{a_n}\}$ provides the required finite subcover of $\mathcal{U}$. If J is proper we can use (BPI) to obtain a prime ideal I containing J. But then I belongs to X but to no member of $\mathcal{U}$, $\lightning$. $\square$

11.3 Proposition. *Let $X := \mathcal{I}_p(B)$ and let $\langle X; \mathcal{T}\rangle$ be the prime ideal space of the Boolean algebra B. Then the clopen subsets of X are exactly the sets X_a for $a \in B$. Further, given distinct points $x, y \in X$, there exists a clopen subset V of X such that $x \in V$ and $y \notin V$.*

Proof. As noted above, each set X_a is clopen. Also, given distinct I_1 and I_2 in $\mathcal{I}_p(B)$, there exists, without loss of generality, $a \in I_1 \setminus I_2$. Then X_a contains I_2 but not I_1. This proves the final assertion.

It remains to prove that an arbitrary clopen subset U of X is of the form X_a for some $a \in B$. Because U is open, $U = \bigcup_{a \in A} X_a$ for some subset A of B. But U is also a closed subset of X and so compact (by A.7). Hence there exists a finite subset A_1 of A such that $U = \bigcup_{a \in A_1} X_a$. Then $U = X_a$, where $a = \bigvee A_1$ (see 10.20). $\square$

By combining Theorem 10.22 and the first part of the preceding proposition we obtain Stone's famous representation theorem.

11.4 Stone's representation theorem for Boolean algebras. *Let B be a Boolean algebra. Then the map*

$$\eta\colon a \longmapsto X_a := \{ I \in \mathcal{I}_p(B) \mid a \notin I\}$$

is a Boolean algebra isomorphism of B onto the Boolean algebra of clopen subsets of the dual space $\langle \mathcal{I}_p(B); \mathcal{T}\rangle$ of B.

To exploit this representation to the full we need to know more about topological spaces with the properties possessed by $\mathcal{I}_p(B)$. The last part of Proposition 11.3 asserts that the prime ideal space of a Boolean algebra satisfies a separation condition guaranteeing that the space has 'plenty' of clopen subsets. We next pursue the topological ramifications of this condition.

11.5 Totally disconnected spaces and Boolean spaces. We say that a topological space $\langle X; \mathcal{T}\rangle$ is **totally disconnected** if, given distinct points $x, y \in X$, there exists a clopen subset V of X such that $x \in V$ and $y \notin V$. Topologists usually give a different definition of total disconnectedness. For compact spaces their definition agrees with ours. If $\langle X; \mathcal{T}\rangle$ is both compact and totally disconnected, it is said to be a **Boolean space**. Propositions 11.2 and 11.3 assert that $\langle \mathcal{I}_p(B); \mathcal{T}\rangle$ is a Boolean space for every Boolean algebra B. We denote by $\wp^{\mathcal{T}}(X)$ the family of clopen subsets of a Boolean space $\langle X; \mathcal{T}\rangle$.

Given distinct points x, y in a totally disconnected space X, there exist disjoint clopen sets V and $W := X \setminus V$ such that $x \in V$ and $y \in W$. This implies that a totally disconnected space is Hausdorff; see A.5. In particular, a Boolean space is compact and Hausdorff. Compact Hausdorff spaces are well known to have many nice properties, some of which are stated in A.7 and A.8. Working with Boolean spaces is like working with compact Hausdorff spaces, but with the bonus of having a basis of clopen sets. This is illustrated by the proof of Lemma 11.6 which is analogous to that of Lemma A.8, except that use of the total disconnectedness condition replaces use of the Hausdorff condition.

11.6 Lemma. *Let $\langle X; \mathcal{T}\rangle$ be a Boolean space.*

(i) *Let Y be a closed subset of X and $x \notin Y$. Then there exists a clopen set V such that $Y \subseteq V$ and $x \notin V$.*

(ii) *Let Y and Z be disjoint closed subsets of X. Then there exists a clopen set U such that $Y \subseteq U$ and $Z \cap U = \varnothing$.*

Proof. (i) For each $y \in Y$ there exists a clopen set V_y with $y \in V_y$ and $x \notin V_y$. The open sets $\{ V_y \mid y \in Y\}$ form an open cover of Y. Since Y is compact (by A.7), there exist $y_1, \ldots, y_n \in Y$ such that $Y \subseteq V := V_{y_1} \cup \cdots \cup V_{y_n}$. As a finite union of clopen sets, V is clopen; by construction it does not contain x.

(ii) This is left as an exercise; compare Lemma A.8. $\square$

This section has so far been totally bereft of examples. The simplest type of Boolean space is a finite set with the discrete topology, but infinite examples are more elusive. We shall discuss the prime ideal space

of the finite-cofinite algebra, FC(ℕ), in 11.8. To aid us in identifying this we need an important companion to Theorem 11.4, which provides an indirect way to obtain dual spaces. The proof of Theorem 11.7 uses several topological lemmas. However, the benefits we shall reap from the result make worthwhile the work involved in the proof.

11.7 Theorem.

(i) *Let Y be a Boolean space, let B be the algebra $\wp^T(Y)$ of clopen subsets of Y and let X be the dual space of B. Then Y and X are homeomorphic.*

(ii) *Let C be a Boolean algebra and Y a Boolean space such that $C \cong \wp^T(Y)$. Then the dual space of C is (homeomorphic to) Y.*

Proof. We define $\varepsilon\colon Y \to X$ by $\varepsilon(y) := \{\, a \in B \mid y \notin a \,\}$. Certainly $\varepsilon(y)$ is a prime ideal in B. We shall show that ε is a continuous bijection from Y onto X. It then follows by A.7 that ε is a homeomorphism.

Because Y is totally disconnected, if $y \neq z$ in Y then there exists a clopen subset a of Y such that $y \in a$ and $z \notin a$. Hence ε is injective.

To establish continuity of ε we need A.4: it suffices to show that $\varepsilon^{-1}(X_a)$ is clopen for each $a \in B$. But this is so: by the definition of X_a and the definition of ε we have

$$\varepsilon^{-1}(X_a) = \{\, y \in Y \mid \varepsilon(y) \in X_a \,\} = \{\, y \in Y \mid a \notin \varepsilon(y) \,\} = a.$$

Finally, we prove that ε is surjective. By Lemma A.7, $\varepsilon(Y)$ is a closed subset of X. Suppose by way of contradiction that there exists $x \in X \setminus \varepsilon(Y)$. Then 11.6 and 11.3 imply that there is a subset X_a of X such that $\varepsilon(Y) \cap X_a = \varnothing$ and $x \in X_a$. We have $\varnothing = \varepsilon^{-1}(X_a) = a$. But this contradicts $x \in X_a$, $\lightning$.

This proves (i), and (ii) follows from it. □

11.8 Example: the finite-cofinite algebra FC(ℕ). Denote by $\mathbb{N}_\infty$ the set of natural numbers with an additional point, ∞, adjoined. We define $\mathcal{T}$ as follows: a subset U of $\mathbb{N}_\infty$ belongs to $\mathcal{T}$ if

$$\text{either (a) } \infty \notin U,$$
$$\text{or (b) } \infty \in U \text{ and } \mathbb{N}_\infty \setminus U \text{ is finite.}$$

We leave as an exercise the proof that $\mathcal{T}$ is a topology.

A subset V of $\mathbb{N}_\infty$ is clopen if and only if both V and $\mathbb{N}_\infty \setminus V$ are open. It follows that the clopen subsets of $\mathbb{N}_\infty$ are the finite sets not containing ∞ and their complements.

It is now easy to show that $\mathbb{N}_\infty$ is totally disconnected. Given distinct points $x, y \in \mathbb{N}_\infty$, we may assume without loss of generality that $x \neq \infty$. Then $\{x\}$ is clopen and contains x but not y.

We next prove that $\mathbb{N}_\infty$ is compact. Take an open cover $\mathcal{U}$ of $\mathbb{N}_\infty$. Some member of $\mathcal{U}$ must contain ∞; say U is such a set. Then $\mathbb{N}_\infty \setminus U$ is finite, by (b). Hence only finitely many members of $\mathcal{U}$ are needed to cover $\mathbb{N}_\infty \setminus U$ and these, together with U, provide the required finite subcover of $\mathcal{U}$. (The *cognoscenti* will recognize the space $\mathbb{N}_\infty$ as the 1-point compactification of a countable discrete space.)

The algebra B of clopen sets of the Boolean space $\mathbb{N}_\infty$ consists of the finite sets not containing ∞ and their complements. Define $f\colon \mathrm{FC}(\mathbb{N}) \to B$ by

$$f(a) = \begin{cases} a & \text{if } a \text{ is finite,} \\ a \cup \{\infty\} & \text{if } a \text{ is cofinite.} \end{cases}$$

This map is easily seen to be an isomorphism. Therefore, by Theorem 11.7(ii), the dual space of FC(ℕ) can be identified with $\mathbb{N}_\infty$. We can now recognize the elements of $\mathcal{I}_p(B)$. The points of ℕ are in one-to-one correspondence with the principal prime ideals of FC(ℕ), via the map $n \mapsto \downarrow(\mathbb{N} \setminus \{n\})$. There is a single non-principal prime ideal, associated with ∞: it consists of all finite subsets of ℕ.

The next example extends our repertoire of Boolean spaces. It is included for those with the topological knowledge and sophistication to appreciate it. We shall not use it later.

11.9 Example: a countable atomless Boolean algebra. The example of a Boolean algebra with no atoms given in Exercise 10.11 is uncountable. It might be surmised that all atomless algebras are uncountable, but this is not so.

Let C be the Cantor 'middle third' set, regarded as a subset of $[0, 1]$. Then C is compact, since C is obtained from $[0, 1]$ by removing open intervals. Also, if $x < y$ in C, there exists u such that $x < u < y$ and $u \notin C$. Then $C \cap [0, u]$ is clopen, contains x and does not contain y. It follows that C is a Boolean space. Those reasonably adept at topology can now prove that $\wp^T(C)$ is an example of a countable Boolean algebra with no atoms.

11.10 Duality for Boolean algebras. We have by no means yet fully exploited the power of the Stone representation for Boolean algebras. However, everything we have done so far in this chapter, and much more, extends to the more general setting of bounded distributive lattices. We shall therefore suspend our treatment of the Stone representation at this point, but note that the results of 11.27 and 11.29–11.32 specialize to Boolean algebras.

Meet LINDA: the Lindenbaum algebra

This optional section deals with an important fragment of mathematical logic and the part Boolean algebras play in it. We do not claim to be presenting a primer on formal logic, and those unfamiliar with the subject are referred to standard texts for motivation and background.

There are two quite different approaches to propositional calculus. One is the **semantic** one, based on assignments of truth values, which we discussed in 4.19. In this approach a wff is said to be true if its truth function always takes value **T**. Such a wff is called a **tautology**.

The alternative is a **syntactic** approach, based on a formal deduction system in which a wff is declared to be true if it can be derived from a set of axioms via given deduction rules. We outline one such system.

The reconciliation of these two approaches is discussed in 11.12.

11.11 The formal system L. A **deduction system** consists of

(i) a set of formulae,

(ii) a subset of the formulae designated as axioms,

(iii) a finite set of deduction rules.

The system **L** of propositional calculus is defined as follows.

(i) The formulae are the wffs of propositional calculus, with $\rightarrow$ and $\neg$ as connectives. (In this section we use the notation $\neg\varphi$ for the negation of φ, rather than φ', and introduce $(\varphi \vee \psi)$ as shorthand for $(\neg\varphi \rightarrow \psi)$ and $(\varphi \wedge \psi)$ as shorthand for $\neg(\varphi \rightarrow \neg\psi)$.)

(ii) The axioms of **L** are all wffs of the form

(A1) $(\varphi \rightarrow (\psi \rightarrow \varphi))$,

(A2) $((\varphi \rightarrow (\psi \rightarrow \chi)) \rightarrow ((\varphi \rightarrow \psi) \rightarrow (\varphi \rightarrow \chi)))$,

(A3) $((\neg\varphi \rightarrow \neg\psi) \rightarrow (\psi \rightarrow \varphi))$,

where φ, ψ and χ are any wffs.

(iii) There is a single deduction rule, **modus ponens**:

(MP) from φ and $(\varphi \rightarrow \psi)$ deduce ψ.

A **proof** in **L** is a finite sequence of formulae of which each is either an axiom or is obtained from two previous ones by (MP). A **theorem** of **L** is the last formula in a proof (in other words, it is the culmination of a proof). If φ is a theorem of **L** we write $\vdash_{\mathbf{L}} \varphi$. Theorems *of* **L** must not be confused with theorems *about* the system **L**, which are usually called **metatheorems**.

11.12 Semantics versus syntax. We now have two classes of wffs with a claim to be called true:

(i) the tautologies (semantically true);

(ii) the theorems of the formal system **L** (syntactically true).

We can say that **L** successfully models deductive reasoning with propositions if it is

- **sound**, that is, every theorem is a tautology, and
- **adequate**, that is, every tautology is a theorem.

The major metatheorems of propositional calculus are the Soundness Theorem and the Adequacy Theorem, asserting respectively that **L** is sound and **L** is adequate. Of these, the Soundness Theorem is by far the more elementary. It works because (a) each axiom is a tautology (a routine verification) and (b) from tautologies φ and $\varphi \rightarrow \psi$, the rule (MP) yields another tautology, ψ. The Adequacy Theorem is far more subtle, and it is only the judicious choice of axioms that makes it work. Indeed, it is remarkable that it does. As anyone who has tried will appreciate, establishing that a particular wff of **L** is a theorem can be a tricky business. The Adequacy Theorem says that a given wff is a theorem of **L** if it is a tautology, and this may be confirmed or refuted by writing down the truth table, a purely mechanical process which can be carried out in a finite number of steps.

Below we derive the Adequacy Theorem as a consequence of (BPI), or equivalently of (BUF) (see 10.19). In preparation we need to associate a Boolean algebra with the formal system **L**, which is a syntactic counterpart of the Boolean algebra to which we alluded, rather informally, at the end of 4.19. Before discussing these algebras we recall some facts from propositional calculus which we have not so far needed.

11.13 Valuations. A map v from wffs to $\{\mathbf{F}, \mathbf{T}\}$ is a **valuation** if

(i) $v(\neg\varphi) = \mathbf{T}$ if and only if $v(\varphi) = \mathbf{F}$,

(ii) $v(\varphi \rightarrow \psi) = \mathbf{T}$ unless $v(\varphi) = \mathbf{T}$ and $v(\psi) = \mathbf{F}$.

An equivalent definition requires v to preserve $\vee, \wedge$ and $'$ (interpreted on the set of wffs as connectives and on $\{\mathbf{F}, \mathbf{T}\}$ as in 4.18). Elementary lemmas assert that any assignment of truth values to the propositional variables extends in a unique way to a valuation and that a wff φ is a tautology if and only if $v(\varphi) = \mathbf{T}$ for all valuations v.

Given wffs φ and ψ, we say φ **logically implies** ψ, and write $\varphi \models \psi$, if whenever $v(\varphi) = \mathbf{T}$ for a valuation v, then $v(\psi) = \mathbf{T}$. Note that $\varphi \equiv \psi$ if and only if $\varphi \models \psi$ and $\psi \models \varphi$.

11.14 The Lindenbaum algebra, LINDA. We define equivalence relations, called **semantic equivalence**, $\sim_{\models}$, and **syntactic equivalence**, $\sim_{\vdash}$, on wffs by

$$\varphi \sim_{\models} \psi \text{ if and only if } \varphi \equiv \psi,$$
$$\varphi \sim_{\vdash} \psi \text{ if and only if } \vdash_{\mathbf{L}} (\varphi \to \psi) \text{ and } \vdash_{\mathbf{L}} (\psi \to \varphi).$$

Given the Soundness and Adequacy Theorems, it is an easy exercise to show that $\sim_{\models}$ and $\sim_{\vdash}$ are actually the same relation. However, *en route* to proving the Adequacy Theorem we must treat these relations independently. Let $\sim$ denote either $\sim_{\models}$ or $\sim_{\vdash}$, let $[\varphi]$ be the equivalence class of φ under $\sim$ and denote the set of $\sim$-equivalence classes by LA or, where we need to specify which relation is being used, $\mathrm{LA}_{\models}$ or $\mathrm{LA}_{\vdash}$.

We show that, for either choice of $\sim$, there are natural operations making LA into a Boolean algebra. The most economical route is to define an order relation on LA, to show this makes LA a lattice and finally to show that this lattice is Boolean. All the verifications required are much easier for $\sim_{\models}$ than for $\sim_{\vdash}$. This is only to be expected. In the former case only logical equivalence and implication are involved. In the latter, it is necessary to show that many wffs are theorems of **L**. We only give an indication of the steps, but energetic readers familiar with **L** should be able to complete the proofs. The Deduction Theorem, a standard metatheorem of propositional calculus not relying on the Adequacy Theorem, is an extremely useful tool.

Define $\leqslant$ on $\mathrm{LA}_{\models}$ by

$$[\varphi] \leqslant [\psi] \text{ if and only if } \varphi \models \psi$$

and on $\mathrm{LA}_{\vdash}$ by

$$[\varphi] \leqslant [\psi] \text{ if and only if } \vdash_{\mathbf{L}} (\varphi \to \psi).$$

It can be checked in either case that $\leqslant$ is well defined, that is, $[\varphi] = [\varphi_1]$, $[\psi] = [\psi_1]$ and $[\varphi] \leqslant [\psi]$ together imply $[\varphi_1] \leqslant [\psi_1]$. Further, $\leqslant$ is an order relation. In $\langle \mathrm{LA}; \leqslant \rangle$ there are greatest and least elements,

$$1 = \begin{cases} [\varphi], \text{ where } \varphi \text{ is any tautology} & (\text{for } \sim_{\models}), \\ [\varphi], \text{ where } \varphi \text{ is such that } \vdash_{\mathbf{L}} \varphi & (\text{for } \sim_{\vdash}), \end{cases}$$

and 0, obtained similarly, with $[\varphi]$ replaced by $[\neg\varphi]$.

The next step is to define join, meet and complement on LA. Let

$$[\varphi] \vee [\psi] := [\varphi \vee \psi], \quad [\varphi] \wedge [\psi] := [\varphi \wedge \psi], \quad [\varphi]' := [\neg\varphi].$$

We claim that

(i) $\langle \mathrm{LA}; \leqslant \rangle$ is a lattice with join and meet given by $\vee$ and $\wedge$;

(ii) $\langle \mathrm{LA}; \vee, \wedge \rangle$ is distributive;

(iii) $[\varphi] \vee [\varphi]' = 1$ and $[\varphi] \wedge [\varphi]' = 0$.

Some guidance on checking these claims for $\mathrm{LA}_{\vdash}$ is called for. To show, for example, that $[\varphi \vee \psi]$ is the least upper bound of $[\varphi]$ and $[\psi]$ with respect to $\leqslant$, we need

$$\vdash_{\mathbf{L}} (\varphi \to (\neg\varphi \to \psi)),$$
$$\vdash_{\mathbf{L}} (\psi \to (\neg\varphi \to \psi)),$$
$$\vdash_{\mathbf{L}} ((\varphi \to \chi) \to ((\psi \to \chi) \to ((\neg\varphi \to \psi) \to \chi))) \quad \text{for any wff } \chi.$$

The first of these is a well-known theorem and the second is an instance of an axiom. The third is easily obtained using the Deduction Theorem and the following theorems of **L**:

$$\vdash_{\mathbf{L}} ((\alpha \to \beta) \to (\neg\beta \to \neg\alpha)), \quad \vdash_{\mathbf{L}} ((\neg\alpha \to \neg\beta) \to ((\neg\alpha \to \beta) \to \alpha)).$$

We conclude that each of $\langle \mathrm{LA}_{\models}; \vee, \wedge, ', 0, 1 \rangle$ and $\langle \mathrm{LA}_{\vdash}; \vee, \wedge, ', 0, 1 \rangle$ is a Boolean algebra. The former algebra was introduced, without the formality of equivalence classes, in 4.19.

Once the Adequacy Theorem is established we know that these are actually the same Boolean algebra. It is known as the **Lindenbaum algebra**, or, to its friends, as LINDA.

11.15 Valuations and homomorphisms. There is a connection between valuations and Boolean homomorphisms from $\mathrm{LA}_{\models}$ or $\mathrm{LA}_{\vdash}$ to **2**. Since valuation is a semantic concept, it might seem more natural to consider $\mathrm{LA}_{\models}$. However working instead with $\mathrm{LA}_{\vdash}$ provides the key to the Boolean algebra proof of the Adequacy Theorem in 11.16. Assume that v is a valuation. The Soundness Theorem implies that f_v, given by

$$f_v([\varphi]) = \begin{cases} 1 & \text{if } v(\varphi) = \mathbf{T}, \\ 0 & \text{if } v(\varphi) = \mathbf{F}, \end{cases}$$

is a well-defined map from $\mathrm{LA}_{\vdash}$ to **2**. It is routine to show that f_v is a Boolean homomorphism and that every Boolean homomorphism from $\mathrm{LA}_{\vdash}$ to **2** arises in this way from some valuation.

11.16 The Adequacy Theorem. We promised that we would prove the Adequacy Theorem for the system **L** of propositional calculus by using the Boolean algebra $\mathrm{LA}_{\vdash}$ (recall 11.14).

We wish to prove that $\vdash_{\mathbf{L}} \varphi$ for every tautology φ. We prove the contrapositive. Suppose $\nvdash_{\mathbf{L}} \varphi$. Then $[\varphi] \neq 1$ in $\mathrm{LA}_{\vdash}$. The ideal $\downarrow[\varphi]$ is proper. We appeal to (BPI) to find a prime ideal I such that $\downarrow[\varphi] \subseteq I$. Define a map from the wffs to $\{\mathbf{F}, \mathbf{T}\}$ by

$$v(\psi) = \begin{cases} \mathbf{T} & \text{if } [\psi] \notin I, \\ \mathbf{F} & \text{if } [\psi] \in I. \end{cases}$$

It can be proved directly that this map is a valuation (see 11.13); alternatively this may be verified by introducing as a stepping stone the Boolean homomorphism f on $\mathrm{LA}_{\vdash}$ specified by $f^{-1}(0) = I$. By construction, $v(\varphi) = \mathsf{F}$, so φ is not a tautology. This completes the proof of the Adequacy Theorem.

It is usual in elementary treatments of propositional calculus to take the set of propositional variables to be countable. This assumption is made to avoid having to invoke (ZL). When the set of propositional variables is countable, so is $\mathrm{LA}_{\vdash}$. Exercise 10.7 shows that (BPI) for countable Boolean algebras can be proved without (ZL) (by a process remarkably reminiscent of the technique customarily employed in the proof of the Adequacy Theorem).

We remark also that there is a strong connection between the Compactness Theorem for propositional calculus and the compactness of the dual space of a Boolean algebra. (cf. [60], 1.2.3, which makes this explicit (in a negation-free setting).)

Priestley's representation theorem for distributive lattices

We now move from Boolean algebras to distributive lattices.

11.17 Stocktaking. Let L be a distributive lattice and let $X = \mathcal{I}_p(L)$ be its set of prime ideals ordered, as usual, by inclusion. We already have representations for L in two special cases.

When L is Boolean and X is topologized in the way described above, L is isomorphic to the algebra $\wp^{\mathcal{T}}(X)$ of clopen subsets of X. Every prime ideal of a Boolean algebra is maximal (by 10.12), so the order on X is discrete (that is, $x \leqslant y$ in X if and only if $x = y$). Thus the order has no active role in this case.

When L is finite, L is isomorphic to the lattice $\mathcal{O}(X)$ of down-sets of X (by 5.12 and 10.8). Suppose X has a topology $\mathcal{T}$ making it a Boolean space. Then $\mathcal{T}$ is the discrete topology, in which every subset is clopen (see A.9). In this case the topology contributes nothing.

These observations suggest that to represent L in general we should equip X with the inclusion order *and* a suitable Boolean space topology. A prime candidate for a lattice isomorphic to L would then be the lattice of all clopen down-sets of X. Our remarks above imply that this lattice coincides with $\mathcal{O}(X)$ when L is finite, with $\wp^{\mathcal{T}}(X)$ when L is Boolean (and with $\wp(X)$ when L is both finite and Boolean). We prove in 11.23 that a bounded distributive lattice L is indeed isomorphic to the lattice of clopen down-sets of $\mathcal{I}_p(L)$, ordered by inclusion and appropriately

topologized, and thereby obtain a natural common generalization of Birkhoff's and Stone's theorems. The boundedness restriction is forced upon us because the lattice of clopen down-sets is bounded. Extensions of the theorem to lattices lacking bounds do exist, but we do not consider them here.

11.18 The prime ideal space of a bounded distributive lattice. Let L be a distributive lattice with 0 and 1 and for each $a \in L$ let

$$X_a := \{ I \in \mathcal{I}_p(L) \mid a \notin I \},$$

as before. Let $X := \mathcal{I}_p(L)$. We want a topology $\mathcal{T}$ on X so that each X_a is clopen. Accordingly, we want every element of

$$\mathcal{S} := \{ X_b \mid b \in L \} \cup \{ X \setminus X_c \mid c \in L \}$$

to be in $\mathcal{T}$. Compared with the Boolean case we have a double complication to contend with. The family $\mathcal{S}$ contains sets of two types and it is also not closed under finite intersections. We let

$$\mathcal{B} := \{ X_b \cap (X \setminus X_c) \mid b, c \in L \}.$$

Since L has 0 and 1, the set $\mathcal{B}$ contains $\mathcal{S}$. Also $\mathcal{B}$ is closed under finite intersections. Finally, we define $\mathcal{T}$ as follows: $U \in \mathcal{T}$ if U is a union of members of $\mathcal{B}$. Then $\mathcal{T}$ is the smallest topology containing $\mathcal{S}$; in the terminology of A.3, $\mathcal{S}$ is a subbasis for $\mathcal{T}$ and $\mathcal{B}$ a basis.

11.19 Theorem. *Let L be a bounded distributive lattice. Then the prime ideal space $\langle \mathcal{I}_p(L); \mathcal{T} \rangle$ is compact.*

Proof. Alexander's Subbasis Lemma proved (with the aid of (BPI)) in A.10 tells us that it is sufficient to prove that any open cover $\mathcal{U}$ of $X = \mathcal{I}_p(L)$ by sets in the subbasis $\mathcal{S}$ has a finite subcover. Doing this is only slightly more complicated than proving 11.2. Let

$$\mathcal{U} = \{ X_b \mid b \in A_0 \} \cup \{ X \setminus X_c \mid c \in A_1 \}.$$

Let J be the ideal generated by A_0 (this is $\{0\}$ if A_0 is empty) and let G be the filter generated by A_1 (this is $\{1\}$ if A_1 is empty). Assume first that $J \cap G = \varnothing$ and invoke (DPI) to find a prime ideal I such that $J \subseteq I$ and $G \cap I = \varnothing$. Then $I \notin X_b$ for any $b \in A_0$ and $I \notin X \setminus X_c$ for any $c \in A_1$ and this means that $\mathcal{U}$ does not cover X, $\lightning$.

Hence $J \cap G \neq \varnothing$. Take $a \in J \cap G$. If A_0 and A_1 are both non-empty, there exist $b_1, \ldots, b_j \in A_0$ and $c_1, \ldots, c_k \in A_1$ such that

$$c_1 \wedge \cdots \wedge c_k \leqslant a \leqslant b_1 \vee \cdots \vee b_j,$$

whence

$$X = X_1 = X_{b_1} \cup \cdots \cup X_{b_j} \cup (X \setminus X_{c_1}) \cup \cdots \cup (X \setminus X_{c_k}).$$

In this case, therefore, $\mathcal{U}$ has a finite subcover. The case $A_1 = \varnothing$ is treated as in 11.2 and the case $A_0 = \varnothing$ is similar. □

11.20 Totally order-disconnected spaces. A set X carrying an order relation $\leqslant$ and a topology $\mathcal{T}$ is called an **ordered (topological) space** and denoted $\langle X; \leqslant, \mathcal{T}\rangle$ (or by X where no ambiguity would result). It is said to be **totally order-disconnected** if, given $x, y \in X$ with $x \not\geqslant y$, there exists a clopen down-set U such that $x \in U$ and $y \notin U$. This separation condition is illustrated in Figure 11.1. We call a compact totally order-disconnected space a **Priestley space**. These spaces are also known as ordered Stone spaces or CTOD spaces. We shall denote by $\mathcal{O}^\tau(X)$ the family of clopen down-sets of a Priestley space X. As noted above, $\mathcal{O}^\tau(X)$ coincides with $\wp^\tau(X)$ when the order on X is discrete and with $\mathcal{O}(X)$ when X is finite.

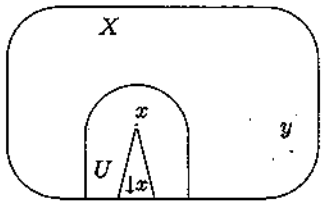


Figure 11.1

Clearly total order-disconnectedness implies total disconnectedness and the two notions coincide when the order is discrete. In many ways Priestley spaces behave like a cross between Boolean spaces and ordered sets. Lemma 11.21 illustrates this. The first part, which is an easy exercise, should be compared with Lemma 1.30. The remainder of the lemma is an analogue for Priestley spaces of Lemma 11.6 and is proved in the same way, but with clopen down-sets replacing clopen sets.

11.21 Lemma. *Let $\langle X; \leqslant, \mathcal{T}\rangle$ be a Priestley space.*

(i) *$x \leqslant y$ in X if and only if $y \in U$ implies $x \in U$ for every $U \in \mathcal{O}^\tau(X)$.*

(ii) (a) *Let Y be a closed down-set in X and let $x \notin Y$. Then there exists a clopen down-set U such that $Y \subseteq U$ and $x \notin U$.*

(b) *Let Y and Z be disjoint closed subsets of X such that Y is a down-set and Z is an up-set. Then there exists a clopen down-set U such that $Y \subseteq U$ and $Z \cap U = \varnothing$.*

We can now characterize clopen sets and clopen down-sets in the dual space $\langle \mathcal{I}_p(L); \subseteq, \mathcal{T}\rangle$ of a bounded distributive lattice L. The proof follows the same lines as that of Proposition 11.3.

11.22 Lemma. *Let L be a bounded distributive lattice with dual space $\langle X; \subseteq, \mathcal{T}\rangle$, where $X = \mathcal{I}_p(L)$. Then*

(i) *the clopen subsets of X are the finite unions of sets of the form $X_b \cap (X \setminus X_c)$ for $b, c \in L$,*

(ii) *the clopen down-sets of X are exactly the sets X_a for $a \in L$.*

11.23 Priestley's representation theorem for distributive lattices. *Let L be a bounded distributive lattice. Then the map*

$$\eta\colon a \longmapsto X_a := \{ I \in \mathcal{I}_p(L) \mid a \notin I \}$$

is an isomorphism of L onto the lattice of clopen down-sets of the dual space $\langle \mathcal{I}_p(L); \subseteq, \mathcal{T}\rangle$ of L.

Proof. Combine Theorem 10.21 and Lemma 11.22(ii). □

Our next task is to give a simultaneous generalization of Theorem 5.9 and Theorem 11.7. Ordered spaces X and Y are 'essentially the same' if there exists a map φ from X onto Y which is simultaneously an order-isomorphism and a homeomorphism. We call such a map an **order-homeomorphism** and say X and Y are **order-homeomorphic**.

11.24 Theorem.

(i) *Let Y be a Priestley space, let L be the lattice $\mathcal{O}^\tau(Y)$ of clopen down-sets of Y and let X be the dual space of L. Then Y and X are order-homeomorphic.*

(ii) *Let L be a bounded distributive lattice and Y a Priestley space such that $\mathcal{O}^\tau(Y) \cong L$. Then the dual space of L is (order-homeomorphic to) Y.*

Proof. The second part follows from the first. The proof of (i) is similar to the proof given for the Boolean case in 11.7, but somewhat more complicated.

We define $\varepsilon\colon Y \to X$ by $\varepsilon(y) := \{ a \in L \mid y \notin a \}$. Certainly $\varepsilon(y)$ is a prime ideal in L. We must show that

(a) ε is an order-embedding (and hence, by 1.36(3), one-to-one);

(b) ε is continuous;

(c) ε maps Y onto X.

Combined with A.7 this will establish (i).

For (a) note that

$$\begin{aligned} y \leqslant z \text{ in } Y &\iff (\forall a \in L)\,(z \in a \Rightarrow y \in a) \quad \text{(by Lemma 11.21(i))} \\ &\iff \varepsilon(y) \subseteq \varepsilon(z). \end{aligned}$$

To prove (b) we need A.4. It implies that (b) holds so long as $\varepsilon^{-1}(X_a)$ and $\varepsilon^{-1}(X \setminus X_a)$ are open for each $a \in L$. We find that

$$\varepsilon^{-1}(X \setminus X_a) = \{\, y \in Y \mid \varepsilon(y) \notin X_a \,\} = Y \setminus \varepsilon^{-1}(X_a).$$

Thus (b) holds provided $\varepsilon^{-1}(X_a)$ is clopen in Y for each $a \in L$. But, by the definition of X_a and the definition of ε, we have

$$\varepsilon^{-1}(X_a) = \{\, y \in Y \mid \varepsilon(y) \in X_a \,\} = \{\, y \in Y \mid a \notin \varepsilon(y) \,\} = a,$$

and this is clopen, by the definition of L.

Finally, we prove (c). By Lemma A.7, $\varepsilon(Y)$ is a closed subset of X. Suppose by way of contradiction that there exists $x \in X \setminus \varepsilon(Y)$. Lemma 11.6(i) implies that there is a clopen subset V of X such that $\varepsilon(Y) \cap V = \varnothing$ and $x \in V$. By 11.22, we may assume that $V = X_b \cap (X \setminus X_c)$ for some $b, c \in L$. We have $\varnothing = \varepsilon^{-1}(V) = b \cap (Y \setminus c)$. Thus $b \subseteq c$, which is impossible since $x \in X_b \cap (X \setminus X_c)$, $\lightning$. □

11.25 Examples. Our first dual space examples involve a minimum of topology. A variety of Priestley spaces can be obtained by equipping the Boolean space $\mathbb{N}_\infty$, introduced in 11.8, with an order. For a very simple example, order $\mathbb{N}_\infty$ as the chain $\mathbb{N}$ with ∞ adjoined as top element, as in Figure 11.2(i). Take $x \not\geqslant y$. Then $y > x$ and $\downarrow x$, which is clopen because it is finite and does not contain ∞, contains x but not y. Hence we have a Priestley space; its lattice of clopen down-sets is isomorphic to the chain $\mathbb{N} \oplus \mathbf{1}$.

Alternatively, consider the ordered space Y obtained by equipping $\mathbb{N}_\infty$ with the order depicted in Figure 11.2(ii). We have $n \succ n-1$ and $n \succ n+1$ for each even n.

For each $n \in \mathbb{N}$, the down-set $\downarrow n$ is finite and does not contain ∞ and so is clopen. Given $x \not\geqslant y$ in Y, we claim that there exists $U \in \mathcal{O}^\tau(Y)$ such that $x \in U$ and $y \notin U$. Either $x \neq \infty$, in which case $y \notin \downarrow x$ and we may take $U = \downarrow x$, or $x = \infty$, in which case we may take $U = Y \setminus \{1, 2, \ldots, 2y\}$. Hence Y is a Priestley space. The associated lattice $\mathcal{O}^\tau(Y)$ – a sublattice of $\mathrm{FC}(\mathbb{N})$ – is easily described.

These examples illustrate how Priestley spaces can be constructed by imposing suitable order relations on a given Boolean space. Alternatively, we might start from an ordered set and try to make it into a

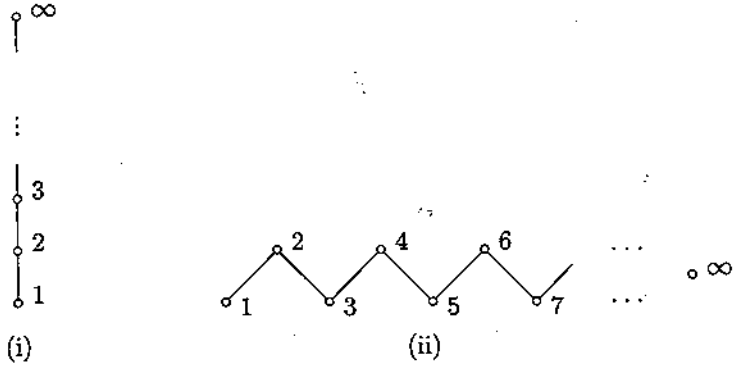


Figure 11.2

Priestley space by topologizing it. This raises the interesting, but difficult, problem of **representability**: when is a given ordered set isomorphic to $\langle \mathcal{I}_p(L); \subseteq \rangle$ for some bounded distributive lattice L, or, more generally, an arbitrary distributive lattice L? For more information on this, see the bibliography of [56].

11.26 Further examples of Priestley spaces.

(1) We indicated in Example 11.9 that the Cantor 'middle third' set, C, regarded as a subset of $[0,1]$, is a Boolean space. In fact with the order inherited from $[0,1]$ it is a Priestley space.

(2) Let S be a set. As a subbasis for a topology on $\langle \wp(S); \subseteq \rangle$ we take the collection of up-sets of the form $\uparrow\{s\}$ (for $s \in S$), and their complements. Alexander's Subbasis Lemma can be used to prove, much as in the proof of 11.19, that $\wp(S)$ then becomes a Priestley space. When S is countable, the underlying Boolean space is homeomorphic to the Cantor space described in 11.9. For further information on this example see [51].

Distributive lattices and Priestley spaces in partnership

The remainder of the chapter parallels the final section of Chapter 5. So far we have presented Boolean algebra results first, and then their distributive lattice counterparts. While we were setting up the necessary machinery this had advantages, since the Boolean case was somewhat simpler. Also this approach gave an easily identified shortest-path route to Stone's Theorem for those interested primarily in Boolean algebras. We contend that, once obtained, the distributive lattice representation is just as easy to work with as that for Boolean algebras and also provides

a richer source of examples. Therefore, we henceforth place the emphasis on distributive lattices, deriving Boolean algebra results as corollaries by specializing to the discrete order.

11.27 Duality. Denote the class of bounded distributive lattices by $\mathbf{D}$, and the class of Priestley spaces (compact totally order-disconnected spaces) by $\mathbf{P}$. Define maps $D\colon \mathbf{D} \to \mathbf{P}$ and $E\colon \mathbf{P} \to \mathbf{D}$ by

$$D\colon L \mapsto \mathcal{I}_p(L) \quad (L \in \mathbf{D}) \quad \text{and} \quad E\colon X \mapsto \mathcal{O}^\tau(X) \quad (X \in \mathbf{P}).$$

Theorems 11.23 and 11.24 assert that, for all $L \in \mathbf{D}$ and $X \in \mathbf{P}$,

$$ED(L) \cong L \quad \text{and} \quad DE(X) \cong X;$$

the latter $\cong$ means 'is order-homeomorphic to'.

We may use the isomorphism between L and $ED(L)$ to represent the members of $\mathbf{D}$ concretely as lattices of the form $\mathcal{O}^\tau(X)$ for $X \in \mathbf{P}$. As an immediate application we note that the representation allows us to construct a 'smallest' Boolean algebra B containing (an isomorphic copy of) a given lattice $L \in \mathbf{D}$: identify L with $\mathcal{O}^\tau(X)$ and take $B = \wp^\tau(X)$. Lemma 11.22 shows how $\mathcal{O}^\tau(X)$ and $\wp^\tau(X)$ are related.

As in the finite case, X is generally very much simpler than $\mathcal{O}^\tau(X)$. We can relate properties of $\mathcal{O}^\tau(X)$ to properties of X, as in 5.18 and 1.32. To complement Chapter 5, we relegate the generalization of these results to the exercises, and use as illustrations in the text a discussion of pseudocomplements (the subject of an exercise in the finite case) and of ideals (of no interest in the finite case since every ideal is then principal).

Before we proceed, a comment on a conflict of notation is called for. We have customarily used lower case letters $a, b, c, \ldots$ for lattice elements. On the other hand, when $L \in \mathbf{D}$ is concretely represented as the lattice $\mathcal{O}^\tau(X)$ of clopen down-sets of its dual space X, it is natural to denote subsets of X, including elements of L, by $U, V, W, \ldots$, and the points of the space X by $x, y, z, \ldots$. The problem is made worse by the fact that the points of a prime ideal space are ideals and, as sets, ideals are usually denoted by upper case letters. We perforce display a kind of notational schizophrenia and switch between these conflicting notational styles as the context indicates.

11.28 Pseudocomplements. We have encountered many distributive lattices which are not Boolean: the requirement that every element a of a bounded lattice L have a complement is stringent. There are many ways to weaken the condition. One possibility is to define the **pseudocomplement** of an element a in a lattice L with 0 to be

$$a^* = \max\{\, b \in L \mid b \wedge a = 0 \,\},$$

if this exists. Pseudocomplements in finite lattices were the subject of Exercise 5.22. Now consider $L = \mathcal{O}^\tau(X)$, where X is a Priestley space. When does $U \in L$ have a pseudocomplement? We claim that this is so if and only if $\uparrow U$ is clopen, and that then $U^* = X \setminus \uparrow U$. To prove the claim, first observe that a down-set W in X does not intersect U if and only if $W \subseteq X \setminus \uparrow U$. Hence $X \setminus \uparrow U$ is the largest down-set disjoint from U and, if it is also clopen, it must be U^*. Conversely, assume U^* exists. Take $x \notin \uparrow U$. We show $x \in U^*$, from which it follows that $U^* = X \setminus \uparrow U$. Exercise 11.14 implies $\uparrow U$ is closed. By the dual of Lemma 11.21(ii) we can find a clopen up-set V such that $x \notin V$ and $\uparrow U \subseteq V$. Then $(X \setminus V) \cap U = \varnothing$, so $X \setminus V \subseteq U^*$ by definition of U^*. This implies that $x \in U^*$. See Figure 11.3.

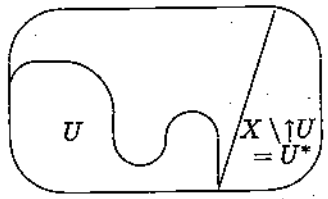


Figure 11.3

11.29 Duality for ideals. Let $L = \mathcal{O}^\tau(X)$ where X is a Priestley space whose family of open down-sets we denote by $\mathcal{L}$. How can we describe the ideals and filters of L in terms of X? An ideal J is determined by its members, which are clopen down-sets of X. Define

$$\Phi(J) = \bigcup \{U \mid U \in J\} \quad (\text{for } J \in \mathcal{I}(L));$$

as a union of clopen sets, $\Phi(J)$ is an open set (but not in general clopen). In the other direction, define

$$\Psi(W) = \{\, U \in \mathcal{O}^\tau(X) \mid U \subseteq W \,\} \quad (\text{for } W \in \mathcal{L});$$

it is easily checked that $\Psi(W)$ is an ideal of L. Further, we claim that

$$\Phi(\Psi(W)) = W \text{ for all } W \in \mathcal{L} \quad \text{and} \quad \Psi(\Phi(J)) = J \text{ for all } J \in \mathcal{I}(L).$$

The first equation asserts that an open down-set W is the union of the clopen down-sets contained in it. To prove this, take any $x \in W$ and use the dual of Lemma 11.21(ii) to find a clopen up-set V containing the closed up-set $X \setminus W$ but with $x \notin V$; then $x \in X \setminus V$, a clopen down-set inside W (see Figure 11.4).

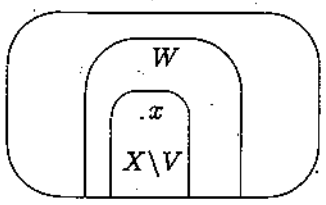


Figure 11.4

Certainly $J \subseteq \Psi(\Phi(J))$ for each ideal J. Take $V \in \Psi(\Phi(J))$. This means that J, regarded as a family of open subsets of X, is an open cover of the clopen set V. By A.7, only finitely many elements of J are needed to cover V, say $U_1, \ldots, U_n$. But $V \subseteq U_1 \cup \cdots \cup U_n$ implies $V \in J$, since J is an ideal. This establishes the second equation.

The bijective correspondence we have set up between $\mathcal{I}(L)$ and $\mathcal{L}$ is in fact a lattice isomorphism (an easy exercise). In addition, special types of ideal correspond to special types of open set (see Exercise 11.17).

Filters may be treated similarly: $\mathcal{F}(L) \cong \mathcal{F}$, the lattice of open up-sets of X.

11.30 Duality. We have amassed a lot of evidence that lattice concepts in $\mathbf{D}$ can be translated into ordered set concepts in $\mathbf{P}$ and vice versa. This '$\mathbf{D}$-$\mathbf{P}$ dictionary' results from there being what is known as a **(full) duality** between $\mathbf{D}$ (bounded distributive lattices + $\{0,1\}$-homomorphisms) and $\mathbf{P}$ (Priestley spaces + continuous order-preserving maps). In line with the philosophy in 1.38, we have here extended the scope of the symbols $\mathbf{D}$ and $\mathbf{P}$ to encompass structure-preserving maps as well as objects. For $L, K \in \mathbf{D}$ and $X, Y \in \mathbf{P}$, we denote the set of $\{0,1\}$-homomorphisms from L to K by $\mathbf{D}(L,K)$ and the set of continuous order-preserving maps from Y to X by $\mathbf{P}(Y,X)$.

The way the duality is required to work is formally laid out in (O) and (M) below. Note the reversal of the directions in (M). We do not digress to discuss the categorical background to these conditions, but do point out that those familiar with dual vector spaces should have a sense of *déjà vu*.

(O) There exist maps $D\colon \mathbf{D} \to \mathbf{P}$ and $E\colon \mathbf{P} \to \mathbf{D}$ such that

(i) for each $L \in \mathbf{D}$, there exists $\eta_L\colon L \to ED(L)$ such that η_L is an isomorphism,

(ii) for each $X \in \mathbf{P}$, there exists $\varepsilon_X\colon X \to DE(X)$ such that ε_X is an order-homeomorphism.

(M) For any $L, K \in \mathbf{D}$, there exists, for each $f \in \mathbf{D}(L,K)$, a map $D(f) \in \mathbf{P}(D(K), D(L))$. For each $X, Y \in \mathbf{P}$, there exists, for each $\varphi \in \mathbf{P}(Y,X)$, a map $E(\varphi) \in \mathbf{D}(E(X), E(Y))$. The maps $D\colon \mathbf{D}(L,K) \to \mathbf{P}(D(K), D(L))$ and $E\colon \mathbf{P}(Y,X) \to \mathbf{D}(E(X), E(Y))$ are bijections and the diagrams below commute.

$$\begin{array}{ccc} L & \xrightarrow{f} & K \\ {\scriptstyle \eta_L}\downarrow & & \downarrow{\scriptstyle \eta_K} \\ ED(L) & \xrightarrow{ED(f)} & ED(K) \end{array} \qquad \begin{array}{ccc} Y & \xrightarrow{\varphi} & X \\ {\scriptstyle \varepsilon_Y}\downarrow & & \downarrow{\scriptstyle \varepsilon_X} \\ DE(Y) & \xrightarrow{DE(\varphi)} & DE(X) \end{array}$$

By Exercise 11.7, $\mathbf{P}(X,2)$ is a $\{0,1\}$-sublattice of 2^X and is isomorphic to $E(X)^\partial$. By Exercise 10.4, $\mathbf{D}(L,2)$ is order-isomorphic to $D(L)^\partial$. Exercise 11.22 indicates that these 'homsets' provide a viable alternative for setting up the duality between $\mathbf{D}$ and $\mathbf{P}$. In fact, the homset approach is technically superior but is inappropriate for a first pass at this topic as it lacks the geometric appeal inherent in the approach via prime ideals and clopen down-sets.

We already have (O), the object part of the duality, with η_L and ε_X the maps supplied by Theorems 11.23 and 11.24 appropriately labelled. Theorem 11.31 provides (M), the correspondence for maps.

11.31 Theorem. *Condition* (M) *holds. Moreover,*

$$a \in (D(f))(y) \iff f(a) \in y \quad \text{for all } f\colon L \to K,\ a \in L,\ y \in D(K).$$

Further,

(i) *f is one-to-one if and only if $D(f)$ is onto,*

(ii) *f is onto if and only if $D(f)$ is an order-embedding.*

Proof. It is elementary to show that, given $\varphi \in \mathbf{P}(Y,X)$, the formula $(E(\varphi))(U) := \varphi^{-1}(U)$ for $U \in E(X)$ defines a $\{0,1\}$-homomorphism $E(\varphi)\colon E(X) \to E(Y)$.

Now assume $f \in \mathbf{D}(L,K)$. Take $y \in D(K)$ and define $(D(f))(y) := f^{-1}(y)$. It is routine to check that $(D(f))(y)$ is a prime ideal in L and that $D(f)$ is order-preserving. To prove $D(f)$ is continuous it is enough (see the proof of 11.24) to prove that $D(f)^{-1}(\eta_L(a))$ is clopen for each $a \in L$. (Here $\eta_L(a)$ is just the set we previously wrote as X_a; we have used the alternative notation because we now have to keep track of subbasic clopen sets in two different dual spaces.) We have

$$\begin{aligned} y \in D(f)^{-1}(\eta_L(a)) &\iff (D(f))(y) \in \eta_L(a) \iff a \notin (D(f))(y) \\ &\iff f(a) \notin y \iff y \in \eta_K(f(a)), \end{aligned}$$

and $\eta_K(f(a))$ is clopen in $D(K)$. The maps D and E are therefore well defined. To verify that diagrams in (M) indeed commute requires a fairly energetic definition-chase, which we leave to the reader, along with the proofs that D and E are bijections on homsets and the proofs of (i) and (ii) (Exercises 11.11 and 11.12). □

Congruences play a very central role in the more advanced theory of lattices. In particular, they are important in the study of distributive lattices with additional operations. It is therefore gratifying that the lattice of congruences of $L \in \mathbf{D}$ turns out to have a very nice concrete representation under our duality.

11.32 Duals of congruences. Let $L \in \mathbf{D}$. As usual, our work is simplified by assuming that $L = \mathcal{O}^\tau(X)$ for some $X \in \mathbf{P}$. We have correspondences (the first by 6.9, the second by 11.31, and the last by A.7):

$$\begin{aligned}\text{congruences on } L &\leftrightarrow \text{surjective } \{0,1\}\text{-homomorphisms with domain } L\\ &\leftrightarrow \text{continuous order-embeddings into } X\\ &\leftrightarrow \text{closed subsets of } X.\end{aligned}$$

Further, a larger congruence on L gives a smaller homomorphic image and hence a smaller closed subset of X. We would therefore expect the lattice $\operatorname{Con} L$ to be isomorphic to the dual of the lattice $\Gamma(X)$ of closed subsets of X. We now elucidate how this isomorphism works, leaving the reader to check the details (not difficult, but a clear head is recommended).

The first step is easy. For each closed subset Y of X, we can define a congruence θ_Y on $\mathcal{O}^\tau(X)$ by

$$(U,V) \in \theta_Y \iff U \cap Y = V \cap Y$$

for $U, V \in \mathcal{O}^\tau(X)$. The congruence θ_Y is the kernel of the $\{0,1\}$-homomorphism f from $\mathcal{O}^\tau(X)$ onto $\mathcal{O}^\tau(Y)$ given by $f(U) := U \cap Y$.

Now let θ be any congruence on L and let $q\colon \mathcal{O}^\tau(X) \to \mathcal{O}^\tau(X)/\theta$ be the natural quotient map. Theorems 6.9 and 11.23 guarantee that there exists an isomorphism $h\colon \mathcal{O}^\tau(X)/\theta \to \mathcal{O}^\tau(Z)$ for some $Z \in \mathbf{P}$. Then $h \circ q$ maps $\mathcal{O}^\tau(X)$ onto $\mathcal{O}^\tau(Z)$. By 11.31 the dual of this map is an order-embedding $\varphi\colon Z \hookrightarrow X$. The set $Y := \varphi(Z)$ is closed in X and $\theta = \theta_Y$.

This shows that $Y \mapsto \theta_{[Y]}$ is a map from $\Gamma(X)$ onto $\operatorname{Con} L$. It is an order-isomorphism between $\Gamma(X)^\partial$ and $\operatorname{Con} L$ provided $\theta_{Y_1} \subseteq \theta_{Y_2}$ implies $Y_1 \supseteq Y_2$ in $\Gamma(X)$ (the reverse implication is trivial). It is enough

to prove that whenever Z is a closed subset of X and $y \in X \setminus Z$, there exist clopen down-sets U, V in X such that $U \cap Z = V \cap Z$ and $y \in V \setminus U$. To do this, apply Exercise 11.14 and Lemma 11.21(ii) twice: first to $\downarrow(Z \cap \downarrow y)$ and $\uparrow y$ to yield U, then to $U \cup \downarrow y$ and $\uparrow(Z \setminus U)$ to yield V.

Hence $\Gamma(X)^\partial \cong \operatorname{Con} \mathcal{O}^\tau(X)$, via the map $Y \mapsto \theta_Y$. We deduce that $\operatorname{Con} \mathcal{O}(X)$ is isomorphic to the lattice of open subsets of X.

Exercises

Exercises from the text. Prove Lemma 11.6(ii). Show that the family $\mathcal{T}$ of subsets of $\mathbb{N}_\infty$, defined in Example 11.8, is a topology on $\mathbb{N}_\infty$ – but beware special cases when checking the topology conditions (T1)–(T3)! (See A.1 for conditions (T1)–(T3).) Prove (or complete the proof of) Lemmas 11.21 and 11.22. Prove the assertion in the penultimate paragraph of 11.29 that $\mathcal{I}(L) \cong \mathcal{L}$. Fill in the details in 11.32.

11.1 A topological space X is called **zero-dimensional** if the clopen subsets of X form a basis for the topology. Prove that the following conditions are equivalent: (One implication is tough!)

(i) X is a Boolean space;

(ii) X is compact, Hausdorff and the only connected subsets of X are the singletons $\{x\}$ for $x \in X$;

(iii) X is compact, T_0 (see Exercise 1.21) and zero-dimensional.

11.2 Let X and Y be topological spaces with $X \cap Y = \varnothing$. There is a natural topology on $X \dot\cup Y$ whose open sets are of the form $A \dot\cup B$ where A is open in X and B is open in Y. This space is called the **disjoint union** of X and Y. (If X and Y are not disjoint, we must first replace them by disjoint copies – see Exercise 1.9 where the corresponding construction was considered for ordered sets.)

(i) Show that if X and Y are (disjoint) Boolean spaces then $X \dot\cup Y$ is also a Boolean space.

(ii) Describe the Boolean algebra of clopen subsets of $X \dot\cup Y$.

11.3 A topological space X is called **extremally disconnected** if the closure of every open set in X is open. Prove that a Boolean lattice B is complete if and only if the Boolean space $D(B)$ is extremally disconnected.

[Hint. Apply the duality. Work with a Boolean space X and the Boolean lattice $\wp^\tau(X)$.]

11.4 Given a chain C define the **interval topology** on C to be the topology which has $\varnothing$, C and the sets of the form $\{x \in C \mid x < c\}$ and $\{x \in C \mid x > c\}$, for $c \in C$, as a subbasis.

(i) Prove that the interval topology on a chain C gives a Boolean space if and only if C is an algebraic lattice.

(ii) (For those who know about ordinals.) Show that the interval topology on an ordinal λ gives a Boolean space if and only if λ is not a limit ordinal.

11.5 (For those who want to get to know LINDA better.)

(i) Check that each axiom of the formal system **L**, as given in 11.11, is a tautology.

(ii) Show that, as claimed in 11.14, the orders defined on $\mathrm{LA}_{\vdash}$ and $\mathrm{LA}_{\dashv}$ are well defined.

(iii) By following the hints given in 11.14, prove that $[\varphi \vee \psi]$ is the least upper bound of $[\varphi]$ and $[\psi]$ in $\mathrm{LA}_{\vdash}$.

(iv) Let v be a valuation. Prove that the map $f_v\colon \mathrm{LA}_{\vdash} \to \mathbf{2}$, as defined in 11.15, is a well-defined Boolean homomorphism and that every Boolean homomorphism from $\mathrm{LA}_{\vdash}$ to $\mathbf{2}$ arises in this way from some valuation.

11.6 Characterize the atoms

(i) in $\wp^{\tau}(X)$ for a Boolean space X,

(ii) in $\mathcal{O}^{\tau}(X)$ for a Priestley space X.

11.7 Let X be a Priestley space.

(i) Assume that $\varphi \in \mathbf{P}(X, \mathbf{2})$. Show that $\varphi^{-1}(0)$ is a clopen down-set in X.

(ii) Let U be a clopen down-set in X. Define $\varphi\colon X \to \mathbf{2}$ by

$$\varphi(x) = \begin{cases} 1 & \text{if } x \notin U, \\ 0 & \text{if } x \in U. \end{cases}$$

Show that $\varphi \in \mathbf{P}(X, \mathbf{2})$.

(iii) Let L denote the set $\mathbf{P}(X, \mathbf{2})$. Show that L is a $\{0,1\}$-sublattice of $\mathbf{2}^X$. Set up an order-isomorphism between $\mathcal{O}^{\tau}(X)$ and L^{∂}.

11.8 (i) Show that if X and Y are (disjoint) Priestley spaces then $X \dot{\cup} Y$ is also a Priestley space (for the definition of the topology on $X \dot{\cup} Y$ see Exercise 11.2).

(ii) (a) Show that, if $X, Y \in \mathbf{P}$, then $\mathcal{O}^{\tau}(X \dot{\cup} Y)$ is isomorphic to $\mathcal{O}^{\tau}(X) \times \mathcal{O}^{\tau}(Y)$.

(b) Let $L, K \in \mathbf{D}$. Use (a) and the duality between **D** and **P** to show there is an order-homeomorphism from the space $D(L \times K)$ of prime ideals of $L \times K$ to $D(L) \dot{\cup} D(K)$. (Compare this with Exercise 10.5.)

11.9 Given $X, Y \in \mathbf{P}$, we denote by $X \oplus Y$ the usual linear sum of the underlying ordered sets endowed with the disjoint union topology defined in Exercise 11.2.

(i) Show that $X \oplus Y \in \mathbf{P}$.

(ii) Show that $\mathcal{O}^{\tau}(X \oplus Y) \cong \mathcal{O}^{\tau}(X) \overline{\oplus} \mathcal{O}^{\tau}(Y)$, where $\overline{\oplus}$ denotes the vertical sum as defined in Exercise 1.18 (the finite case).

(iii) Let $L, K \in \mathbf{D}$. Find all prime ideals in $L \oplus K$. Hence prove, without use of the duality between **D** and **P**, that $D(L \oplus K)$ is order-homeomorphic to $D(L) \oplus \mathbf{1} \oplus D(K)$.

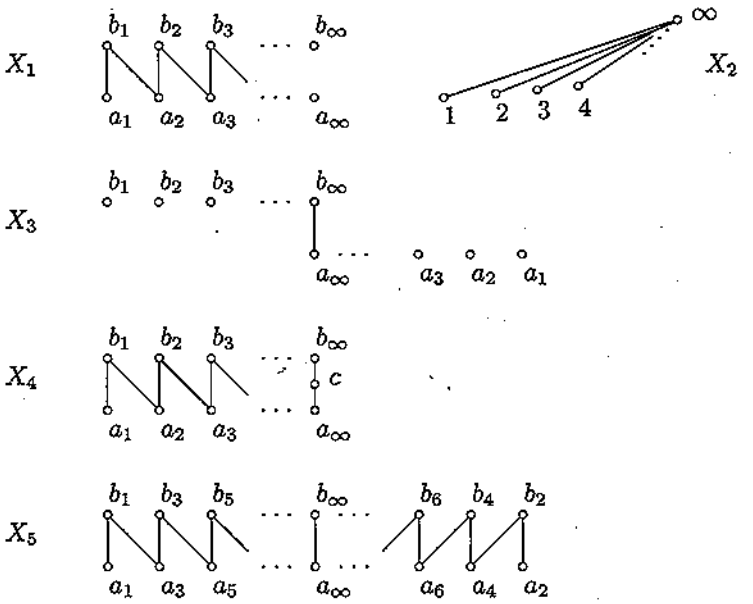


Figure 11.5

11.10 Consider the ordered spaces shown in Figure 11.5. In each case the order and the topology should be apparent. For example, in X_1 the only comparabilities are $a_1 \prec b_1$ and $a_n \prec b_{n-1}$, $a_n \prec b_n$

for $n \geqslant 2$; note, in particular, that $a_\infty \parallel b_\infty$. As a topological space, X_1 is the disjoint union of two copies on $\mathbb{N}_\infty$, namely $\{a_1, a_2, \ldots\} \cup \{a_\infty\}$ and $\{b_1, b_2, \ldots\} \cup \{b_\infty\}$. The other examples are built similarly from one or two copies of $\mathbb{N}_\infty$.

(i) Show that $X_1 \notin \mathbf{P}$.

(ii) Show that $X_2 \in \mathbf{P}$ and describe all the elements of $\mathcal{O}^\tau(X_2)$.

(iii) Consider the Priestley space Y given in Figure 11.2. Show that $\mathcal{O}^\tau(Y)$ is a sublattice of $\mathrm{FC}(\mathbb{N})$ and describe the elements of $\mathcal{O}^\tau(Y)$ (in terms of odd and even numbers).

(iv) Show that $X_3 \in \mathbf{P}$. Show that $\mathcal{O}^\tau(X_3)$ is isomorphic to a sublattice of $\mathrm{FC}(\mathbb{N}) \times \mathrm{FC}(\mathbb{N})$. [Hint. Find a continuous order-preserving map from $\mathbb{N}_\infty \cup \mathbb{N}_\infty$ onto X_3 then use the duality.] Give an explicit description of the elements of this sublattice of $\mathrm{FC}(\mathbb{N}) \times \mathrm{FC}(\mathbb{N})$.

(v) Show that $X_4 \in \mathbf{P}$ and $X_5 \in \mathbf{P}$.

11.11 This exercise completes the proof of Theorem 11.31. Let $L, K \in \mathbf{D}$ and let $f \in \mathbf{D}(L, K)$.

(i) (a) Show that, if I is a prime ideal of K, then $f^{-1}(I)$ is a prime ideal of L.

(b) Show that, if K_1 is a sublattice of K and J is an ideal of K_1, then $\downarrow J$ is an ideal of K.

(ii) (a) Let K_1 be a sublattice of K and let I_1 be a prime ideal of K_1. Show that there exists a prime ideal I of K satisfying $K_1 \cap I = I_1$ [Hint. Apply (DPI) to the ideal $\downarrow I$ and filter $\uparrow(K_1 \setminus I_1)$ of K.]

(b) Prove that, if f is one-to-one, then $D(f)$ is onto.

(c) Prove that, if $D(f)$ is onto, then f is one-to-one. [Hint. Prove the contrapositive. Assume that $a, b \in L$ satisfy $a \neq b$ and $f(a) = f(b)$. Use Theorem 10.21 to obtain a prime ideal $I \in D(L)$ which is not in the image of $D(f)$.]

(iii) (a) Show that, if f is onto, then $D(f)$ is an order-embedding.

(b) Prove that, if $D(f)$ is an order-embedding, then f is onto. [Hint. Again prove the contrapositive. Let $K_1 = f(L)$ and assume that $a \in K \setminus K_1$. Apply (DPI) first to the pair $\downarrow a$, $\uparrow(K_1 \cap \uparrow a)$ to yield a prime ideal I, and then to the pair $\downarrow(K_1 \cap I)$, $\uparrow a$ to yield a second prime ideal J. Show that $D(f)(I) \subseteq D(f)(J)$ while $I \not\subseteq J$.]

11.12 Let $X, Y \in \mathbf{P}$, let X_1 be a closed subset of X and let $\varphi \in \mathbf{P}(Y, X)$.

(i) Let U_1 be a clopen down-set in X_1. Prove that there exists a clopen down-set U in X satisfying $X_1 \cap U = U_1$. [Hint. Apply Lemma 11.21(ii) to the pair $\downarrow U_1$, $\uparrow(X_1 \setminus U_1)$.]

(ii) Prove that, if φ is an order-embedding, then $E(\varphi)$ is onto.

(iii) Use (ii) and the duality between $\mathbf{D}$ and $\mathbf{P}$ to give an alternative proof of Exercise 11.11(iii). [Hint. Use the left-hand commutative diagram in 11.30.]

11.13 A lattice $M \in \mathbf{D}$ is said to be **injective in $\mathbf{D}$** if, for each $K \in \mathbf{D}$ and each $\{0,1\}$-embedding $f\colon L \to K$, every $\{0,1\}$-homomorphism $g\colon L \to M$ can be extended to a homomorphism $\overline{g}\colon K \to M$, that is $\overline{g}(f(a)) = g(a)$ for all $a \in L$. Similarly, a Priestley space Z is **injective in $\mathbf{P}$** if, for every $X \in \mathbf{P}$ and every continuous order-embedding $\varphi\colon Y \to X$, each continuous order-preserving map $\psi\colon Y \to Z$ extends to a continuous order-preserving map $\overline{\psi}\colon X \to Z$.

(i) Use Exercises 11.11(ii)(a) and Exercise 10.4 to show that $\mathbf{2}$ (regarded as a bounded distributive lattice) is injective in $\mathbf{D}$.

(ii) Use Exercises 11.7 and 11.12(i) to show that $\mathbf{2}$ (regarded as a Priestley space) is injective in $\mathbf{P}$.

11.14 Let X be a Priestley space.

(i) Show that $\downarrow y$ and $\uparrow y$ are closed for each $y \in X$.

(ii) Show that, if $Y \subseteq X$ is closed in X, then $\uparrow Y$ and $\downarrow Y$ are closed in X.

(iii) Show that Y is a closed down-set in X if and only if Y is an intersection of clopen down-sets.

(iv) Prove that every directed subset D of X has a join in X. [Hint. Use the duality.] (Hence X is a pre-CPO.)

11.15 For any ordered set P, let $\operatorname{Min} P$ and $\operatorname{Max} P$ denote respectively the set of minimal elements and the set of maximal elements of P.

(i) Prove that if X is a Priestley space then both $\operatorname{Min} X$ and $\operatorname{Max} X$ are non-empty provided X is. [Hint. Use the duality between $\mathbf{D}$ and $\mathbf{P}$.]

(ii) (For those at home with (ZL).) Give a direct proof of the claim in (i). [Hint. When proving that any chain C in X has an upper bound in X, use Exercise 11.14(ii) and the compactness of X to show that $\bigcap_{x \in C} \uparrow x \neq \varnothing$.]

11.16 (Compare with Exercise 11.3.) An ordered space X is called **extremally order-disconnected** if, for every open down-set U in X, the smallest closed down-set containing U is open. Prove that a bounded distributive lattice L is complete if and only if the space $D(L)$ is extremally order-disconnected. [Hint. Apply the duality. Apply Exercise 11.14(iii) to a Priestley space X and the distributive lattice $\mathcal{O}^\tau(X)$.]

11.17 Let X be a Priestley space and let $L = \mathcal{O}^\tau(X)$. (See 11.29.)

(i) Prove that J is a principal ideal in L if and only if $\Phi(J)$ is a clopen down-set.

(ii) Prove that I is a prime ideal in L if and only if $\Phi(I) = X \setminus \uparrow x$ for some $x \in X$.

(iii) Prove that I is a maximal ideal in L if and only if $\Phi(I) = X \setminus \{x\}$ for some maximal element x of X.

11.18 Let $L \in \mathbf{B}$. Use Exercise 11.17 to prove that every prime ideal in L is principal if and only if L is finite.

11.19 Let X be a Priestley space and let $L = \mathcal{O}^\tau(X)$. Describe the mutually inverse maps which establish a bijection between the lattice $\mathcal{F}(L)$ of filters of L and the lattice $\mathcal{F}$ of open up-sets in X.

11.20 Let B be a Boolean algebra. Show that the lattices $\operatorname{Con} B$ of (Boolean) congruences, $\mathcal{I}(B)$ of ideals and $\mathcal{F}(B)$ of filters are isomorphic. [Hint. Use 11.32 and Exercise 6.5.]

Conclude that every (Boolean) congruence θ on a Boolean algebra is of the form θ_J for some ideal J in B (see Exercise 6.4).

11.21 A Priestley space X is called a **p-space** if $\uparrow U$ is open (and therefore clopen, by Exercise 11.14(ii)) for every clopen down-set U in X. Thus, by 11.28, $L = \mathcal{O}^\tau(X)$ is pseudocomplemented if and only if X is a p-space. Let X and Y be p-spaces; then a continuous order-preserving map $\varphi\colon Y \to X$ is called a **p-morphism** if $\varphi(\operatorname{Min} Y \cap \downarrow y) = \operatorname{Min} X \cap \downarrow \varphi(y)$ for all $y \in Y$. Given pseudocomplemented lattices L and K, a map $f\colon L \to K$ is said to **preserve pseudocomplements** if $f(a^*) = f(a)^*$ for all $a \in L$.

(i) Prove that, if X and Y are p-spaces and $\varphi \in \mathbf{P}(Y, X)$, then $E(\varphi)\colon \mathcal{O}^\tau(X) \to \mathcal{O}^\tau(Y)$ preserves pseudocomplements if and only if φ is a p-morphism.

(ii) Show that if X is a p-space then $\operatorname{Min} X$ is closed in X.

(iii) Show that Examples X_2 and X_3 from Exercise 11.10 are not p-spaces while Examples X_4 and X_5 are.

(iv) Let X be a Priestley space such that for every $x \in X$ there is a unique element $m(x) \in \operatorname{Min} X$ such that $m(x) \leqslant x$. Show that X is a p-space if and only if the map $m\colon X \to \operatorname{Min} X$ is continuous.

11.22 This exercise and the next introduce the homset approach to duality referred to in 11.30. Let $L \in \mathbf{D}$ and $X \in \mathbf{P}$. The first task is to topologize $\mathbf{D}(L, \mathbf{2})$. Let $a \in L$ and $\varepsilon \in \{0, 1\}$ and let

$$U(a, \varepsilon) = \{ f \in \mathbf{2}^L \mid f(a) = \varepsilon \}.$$

Let $\mathcal{T}$ be the topology on $\mathbf{2}^L$ which has the sets $U(a, \varepsilon)$, for $a \in L$ and $\varepsilon \in \{0, 1\}$, as a subbasis; thus U is open in $\mathbf{2}^L$ if and only if it is a union of sets each of which is a finite intersection of sets of the form $U(a, \varepsilon)$. We endow $\mathbf{D}(L, \mathbf{2})$ with the subspace topology; thus sets of the form

$$U(a, \varepsilon) \cap \mathbf{D}(L, \mathbf{2}) = \{ f \in \mathbf{D}(L, \mathbf{2}) \mid f(a) = \varepsilon \}$$

constitute a subbasis for the topology on $\mathbf{D}(L, \mathbf{2})$. Show that the map $\varphi\colon \mathcal{I}_p(L) \to \mathbf{2}^L$, given by $\varphi(I) := f_I$ (see Exercise 10.4), is a homeomorphism onto $\mathbf{D}(L, \mathbf{2})$. [Hint. By Exercise 10.4, φ is one-to-one and maps onto $\mathbf{D}(L, \mathbf{2})$, so, by A.7, it remains to show that φ is continuous.]

11.23 Define maps $D'\colon \mathbf{D} \to \mathbf{P}$ and $E'\colon \mathbf{P} \to \mathbf{D}$ by

$$D'\colon L \mapsto \mathbf{D}(L, \mathbf{2}) \qquad (L \in \mathbf{D}),$$
$$E'\colon X \mapsto \mathbf{P}(X, \mathbf{2}) \qquad (X \in \mathbf{P}).$$

Given $L, K \in \mathbf{D}$ and $g \in \mathbf{D}(L, K)$, define $D'(g)\colon D'(K) \to D'(L)$ by $(D'(g))(f) = f \circ g$ for all $f \in D'(K) = \mathbf{D}(K, \mathbf{2})$ and, given $X, Y \in \mathbf{P}$ and $\psi \in \mathbf{P}(Y, X)$, define $E'(\psi)\colon E'(X) \to E'(Y)$ by $(E'(\psi))(\varphi) = \varphi \circ \psi$ for all $\varphi \in E'(X) = \mathbf{P}(X, \mathbf{2})$.

(i) Show that $D'(g)$ is continuous and order-preserving and that $E'(\psi)$ is a $\{0, 1\}$-homomorphism.

(ii) For all $L \in \mathbf{D}$ and $X \in \mathbf{P}$ define $\eta_L\colon L \to E'D'(L)$ and $\varepsilon_X\colon X \to D'E'(X)$ to be the natural 'evaluation maps', that is, $(\eta_L(a))(f) := f(a)$ for all $a \in L$ and $f \in \mathbf{D}'(L) = D(L, \mathbf{2})$ and, similarly, $(\varepsilon_X(x))(\varphi) := \varphi(x)$ for all $x \in X$ and all $\varphi \in E'(X) = \mathbf{P}(X, \mathbf{2})$. Show that η_L is a $\{0, 1\}$-homomorphism, that ε_X is continuous and order-preserving

and that the diagrams below commute.

$$\begin{array}{ccccc} L & \xrightarrow{f} & K & Y & \xrightarrow{\varphi} & X \\ {\scriptstyle\eta_L}\downarrow & & {\scriptstyle\eta_K}\downarrow & {\scriptstyle\varepsilon_Y}\downarrow & & {\scriptstyle\varepsilon_X}\downarrow \\ E'D'(L) & \xrightarrow{E'D'(f)} & E'D'(K) & D'E'(Y) & \xrightarrow{D'E'(\varphi)} & D'E'(X) \end{array}$$

11.24 Let S be a set and equip $\mathbf{2}^S$ with the topology defined in Exercise 11.22. You may assume without proof that this topology is compact. (This is a consequence of a famous result known as Tychonoff's Theorem. Alternatively it can be obtained from Alexander's Subbasis Lemma.)

(i) Show that for all S the ordered space $\mathbf{2}^S$ belongs to $\mathbf{P}$.

(ii) Prove directly from the definition of the topology on $\mathbf{2}^L$ that $\mathbf{D}(L, \mathbf{2})$ is a closed subset of $\mathbf{2}^L$.

(iii) Let X be an ordered space. Show that $X \in \mathbf{P}$ if and only if there exists for some set S an order-homeomorphism φ from X onto a closed subspace of $\mathbf{2}^S$.

Appendix A: a Topological Toolkit

This appendix provides a very concise summary of the results from topology needed in Chapter 11 and its exercises. Our account aims solely to pinpoint those topological ideas we need. Any standard text may be consulted for proofs and further motivation. Our references are to W.A. Sutherland, *An Introduction to Metric and Topological Spaces* [15].

Topology is usually introduced as an abstraction of concepts first met in elementary analysis, such as open neighbourhood and continuous function. In a topological space, a family of **open sets** generalizes the open neighbourhoods of the euclidean spaces $\mathbb{R}^n$. The axioms for a topological space bring under topology's umbrella many structures which are very unlike euclidean spaces. It is certain of such spaces that concern us. The metric spaces which utilize the idea of a distance function analogous to the modulus function on $\mathbb{R}$ and which are frequently used as a stepping stone to topological spaces play no role here.

Proving results in topology demands a certain facility in manipulating sets and maps. The formulae set out in [15], pp. xi–xiii, are a necessary stock in trade.

A.1 Topological spaces. A **topological space** $(X; \mathcal{T})$ consists of a set X and a family $\mathcal{T}$ of subsets of X such that

(T1) $\varnothing \in \mathcal{T}$ and $X \in \mathcal{T}$,

(T2) a finite intersection of members of $\mathcal{T}$ is in $\mathcal{T}$,

(T3) an arbitrary union of members of $\mathcal{T}$ is in $\mathcal{T}$.

The family $\mathcal{T}$ is called a **topology** on X and the members of $\mathcal{T}$ are called **open sets**. We write X in place of $(X; \mathcal{T})$ where $\mathcal{T}$ is the only topology under consideration. The justification for the lopsidedness of conditions (T2) and (T3) is 'because it works'. The standard topology $\mathcal{T}_{\mathbf{R}}$ on $\mathbb{R}$ consists of

$$\{\, U \subseteq \mathbb{R} \mid (\forall x \in U)(\exists \delta > 0)\ (x - \delta, x + \delta) \subseteq U \,\};$$

here δ may depend on x. Equivalently, $\mathcal{T}_{\mathbf{R}}$ consists of those sets which can be expressed as unions of open intervals, together with $\varnothing$. Certainly this family does satisfy (T1), (T2) and (T3). The equation $\bigcap_{n \geqslant 1}(-1/n, 1/n) = \{0\}$ exhibits an intersection of open sets which is not open. Thus (T2) could not be strengthened to require arbitrary intersections of open sets to be open without sacrificing the motivating example of the 'usual' open sets of $\mathbb{R}$.

Given a topological space $(X; \mathcal{T})$ we define a subset of X to be **closed** if it belongs to $\Gamma(X) := \{\, X \setminus U \mid U \in \mathcal{T} \,\}$. The family $\Gamma(X)$ is

closed under arbitrary intersections and finite unions. For every $A \subseteq X$ there exists a smallest closed set $\overline{A}$ containing A; see 7.1 and [15], 3.7.

Sets which are both open and closed are called **clopen**. Very many of the topological spaces encountered in elementary analysis and geometry are **connected**, in the sense that their only clopen subsets are the whole space and the empty set. The spaces used in our representation theory, by contrast, have an ample supply of clopen sets. For a discussion of connectedness, see [15], Chapter 6.

A.2 Subspaces ([15], 3.4). Let $(X;\mathcal{T})$ be a topological space. Any subset Y of X inherits a topology in a natural way. It is given by

$$\mathcal{T}_Y := \{\, V \subseteq Y \mid V = U \cap Y \text{ for some } U \in \mathcal{T} \,\}.$$

A.3 Bases and subbases ([15], 3.2, 3.3). We need to be able to create a topology on a set X in which a specified family $\mathcal{S}$ of subsets of X are open sets. We shall always assume that $\mathcal{S}$ contains $\varnothing$ and X. If $\mathcal{S}$ is already closed under finite intersections, then we define $\mathcal{T}$ to be those sets which are unions of sets in $\mathcal{S}$. Then $\mathcal{T}$ satisfies (T1), (T2) and (T3) and $\mathcal{S}$ is said to be a basis for $\mathcal{T}$. In general, to obtain a topology containing $\mathcal{S}$ we first form $\mathcal{B}$, the family of sets which are finite intersections of members of $\mathcal{S}$, and then define $\mathcal{T}$ to be all arbitrary unions of members of $\mathcal{B}$. In this case $\mathcal{S}$ is called a **subbasis** for $\mathcal{T}$.

A.4 Continuity. Let $(X;\mathcal{T})$ and $(X';\mathcal{T}')$ be topological spaces and $f\colon X \to X'$ a map. Then ([15], 3.7.8, 3.2.5 and p. xii) the following conditions are equivalent:

(i) $f^{-1}(U)$ is open in X whenever U is open in X';

(i)′ $f^{-1}(V)$ is closed in X whenever V is closed in X';

(ii) $f^{-1}(U)$ is open in X for every $U \in \mathcal{S}$, where $\mathcal{S}$ is a given basis or subbasis for $\mathcal{T}'$.

When f satisfies any of these conditions it is said to be **continuous**. In the special case that $(X;\mathcal{T}) = (X';\mathcal{T}') = (\mathbb{R};\mathcal{T}_{\mathbb{R}})$ and $\mathcal{S}$ is the family of subintervals (a,b) (for $-\infty < a < b < \infty$), plus $\mathbb{R}$ and $\varnothing$, (ii) is just a restatement of the ε-δ definition of continuity that so plagues students.

The map $f\colon X \to X'$ is said to be a **homeomorphism** if f is bijective and both f and f^{-1} are continuous. Homeomorphisms are topology's isomorphisms.

A.5 Hausdorff spaces ([15], Chapter 4). Of a hierarchy of possible separation conditions augmenting the topological space axioms, the most important to us is the Hausdorff condition. The topological space $(X;\mathcal{T})$ is said to be **Hausdorff** if, given $x, y \in X$ with $x \neq y$, there exist open sets U_1, U_2 such that $x \in U_1, y \in U_2$ and $U_1 \cap U_2 = \varnothing$. Mnemonically, X is Hausdorff if distinct points can be 'housed off' in disjoint open sets. It is easy to prove the useful result that singleton sets in a Hausdorff space are closed.

A.6 Compactness ([15], Chapter 5). A prime objective of elementary topology is to set in their wider topological context the various results concerning closed bounded subintervals of $\mathbb{R}$ and the continuous real-valued functions on them. The famous Heine–Borel Theorem states that a subset of $\mathbb{R}$ is closed and bounded if and only if it is compact, in the sense we shortly define. Compactness is a fundamental topological concept and may be regarded as a substitute for finiteness. It frequently compensates for the restriction to finite intersections in axiom (T2) by allowing arbitrary families of open sets to be reduced to finite families. All the spaces we use in our representation theory are compact.

Let $(X;\mathcal{T})$ be a topological space and let $\mathcal{U} := \{U_i\}_{i\in I} \subseteq \mathcal{T}$. The family $\mathcal{U}$ is called an **open cover** of $Y \subseteq X$ if $Y \subseteq \bigcup_{i\in I} U_i$. A finite subset of $\mathcal{U}$ whose union still contains Y is a **finite subcover**. We say Y is **compact** if every open cover of Y has a finite subcover.

The lemmas below contain basic results about compact spaces which are also Hausdorff. The first relates compactness and closedness and shows that continuous maps behave well.

A.7 Lemma. *Let $(X;\mathcal{T})$ be a compact Hausdorff space.*

(i) ([15], 5.4.2, 5.6.1) *A subset Y of X is compact if and only if it is closed.*

(ii) ([15], 5.5.1, 5.9.1) *Let $f\colon X \to X'$ be a continuous map, where $(X';\mathcal{T}')$ is any topological space.*

  (a) *$f(X)$ is a compact subset of X'.*

  (b) *If $(X';\mathcal{T}')$ is Hausdorff and $f\colon X \to X'$ is bijective, then f is a homeomorphism.*

Since the next lemma is given in [15] only as an exercise we outline its proof. The lemma strengthens the Hausdorff condition, which is recaptured by taking the closed sets to be singletons.

A.8 Lemma. *Let $(X;\mathcal{T})$ be a compact Hausdorff space.*

(i) *Let V be a closed subset of X and $x \notin V$. Then there exist disjoint open sets W_1 and W_2 such that $x \in W_1$ and $V \subseteq W_2$.*

(ii) *Let V_1 and V_2 be disjoint closed subsets of X. Then there exist disjoint open sets U_1 and U_2 such that $V_i \subseteq U_i$ for $i = 1, 2$.*

Proof. (i) For $y \in V$, use the Hausdorff condition to construct disjoint open sets $U_1^{x,y}$ and $U_2^{x,y}$ containing x and y respectively. Then $\mathcal{U}_2 := \{ U_2^{x,y} \mid y \in V \}$ is an open cover of V, which is compact by A.7. Take a finite subcover $\{ U_2^{x,y_j} \mid j = 1, \ldots, n \}$. Let $U_1^x := \bigcap_{1 \leqslant j \leqslant n} U_1^{x,y_j}$ and $U_2^x := \bigcup_{1 \leqslant j \leqslant n} U_2^{x,y_j}$. Then U_1^x and U_2^x are disjoint, since each U_2^{x,y_j} does not intersect the corresponding U_1^{x,y_j} and so is disjoint from U_1^x. Also U_1^x and U_2^x are open. (Note how compactness reduces the intersection we need to a finite one, so (T2) applies.) These sets contain x and V respectively. Take $W_1 := U_1^x$ and $W_2 := U_2^x$ to obtain (i).

For (ii) we repeat the process, taking $V := V_2$ and letting x vary over V_1. The family $\mathcal{U}_1 := \{ U_1^x \mid x \in V_1 \}$ is an open cover of the compact set V_1. Take a finite subcover $\{ U_1^{x_i} \mid i = 1, \ldots, m \}$ and define $U_1 := \bigcup_{1 \leqslant i \leqslant m} U_1^{x_i}$ and $U_2 := \bigcap_{1 \leqslant i \leqslant m} U_2^{x_i}$. □

The last of this group of lemmas enables us to fit our finite representation theory into the general theory in Chapter 10.

A.9 Lemma. *Let $(X; \mathcal{T})$ be a compact Hausdorff space. Then the following conditions are equivalent:*

(i) *X is finite;*

(ii) *every subset of X is open (that is, $\mathcal{T}$ is discrete);*

(iii) *every subset of X is clopen.*

Proof. Trivially (ii) ⇔ (iii). To prove (iii) ⇒ (i) consider the open cover $\{ \{x\} \mid x \in X \}$. Finally assume (i). For $\varnothing \neq Y \subseteq X$, the set $X \setminus Y$ is a finite union of singleton sets, which are closed because X is Hausdorff. So $X \setminus Y$ is closed, whence Y is open. □

The deepest result about compact spaces we need is Alexander's Subbasis Lemma. We prove this using (BPI); recall 10.17. We have elected to avoid the machinery of product spaces in the main text. Those who already know about product spaces will realize that Alexander's Lemma is closely related to Tychonoff's Theorem. The latter is employed directly in the alternative approach to duality outlined in Exercise 11.22.

A.10 Alexander's Subbasis Lemma. *Let $(X; \mathcal{T})$ be a topological space and $\mathcal{S}$ a subbasis for $\mathcal{T}$. Then X is compact if every open cover of X by members of $\mathcal{S}$ has a finite subcover.*

Proof. Let $\mathcal{B}$ be the basis formed from all finite intersections of members of $\mathcal{S}$. To prove X is compact it is enough to show that every open cover $\mathcal{U}$ of X by sets in $\mathcal{B}$ has a finite subcover. Suppose this is false and let $\mathcal{U}$ be an open cover of X by sets in $\mathcal{B}$ which does not have a finite subcover. Define J to be the ideal in $\wp(X)$ generated by $\mathcal{U}$, so a typical

element of J is a subset of $U_1 \cup \cdots \cup U_k$ for some $U_1, \ldots, U_k \in \mathcal{U}$ (see Exercise 2.22); J is proper, by our hypothesis. Use (BPI) to construct a prime ideal I of $\wp(X)$ containing J. For each $x \in X$, there exists $U(x) \in \mathcal{U}$ with $x \in U(x)$. Each $U(x)$ is a finite intersection of members of $\mathcal{S}$ and belongs to I since $\mathcal{U} \subseteq I$. As I is prime we may assume that $U(x)$ itself lies in $\mathcal{S}$. Let $\mathcal{V} := \{ U(x) \mid x \in X \}$. Then $\mathcal{V}$ is an open cover of X by members of $\mathcal{S}$ and so by assumption has a finite subcover. But then $X = U(x_1) \cup \cdots \cup U(x_n)$ for some finite subset $\{x_1, \ldots, x_n\}$ of X, so that $X \in I$, $\lightning$. □

Appendix B: Further Reading

Background references for related areas of mathematics

[1] S. Abramsky, D. M. Gabbay and T. S. E. Maibaum (eds.), *Handbook of Logic in Computer Science*, Vol. I, Background: mathematical structures, Oxford University Press, 1992. [This includes accounts of basic universal algebra, category theory, topology and logic.]

[2] S. N. Burris, *Logic for Mathematics and Computer Science*, Prentice-Hall, 1998.

[3] S. Burris and H. P. Sankappanavar, *A Course in Universal Algebra*, Springer-Verlag, 1981. (Millennium edition may be freely downloaded from `http://thoralf2.uwaterloo.ca`.)

[4] P. J. Cameron, *Introduction to Algebra*, Oxford University Press, 1998.

[5] J. Dugundji, *Topology*, Allyn and Bacon, 1966.

[6] H. B. Enderton, *Elements of Set Theory*, Academic Press, 1977.

[7] J. B. Fraleigh, *A First Course in Abstract Algebra*, 6th edition, Addison-Wesley, 1999.

[8] D. C. Goldrei, *Classic Set Theory: a Guided Independent Study*, Chapman & Hall/CRC, 1996.

[9] G. Grätzer, *Universal Algebra*, 2nd edition, Springer-Verlag, 1979.

[10] A. G. Hamilton, *Logic for Mathematicians*, 2nd edition, Cambridge University Press, 1988.

[11] W. Hodges, *A Shorter Model Theory*, Cambridge University Press, 1997.

[12] J. Kelley, *General Topology*, Van Nostrand, 1955.

[13] S. Maclane, *Categories for the Working Mathematician*, 2nd edition, Springer-Verlag, 1998.

[14] J. J. Rotman, *An introduction to the theory of groups*, 4th edition, Springer-Verlag, 1995.

[15] W. A. Sutherland, *An Introduction to Metric and Topological Spaces*, Oxford University Press, 1975.

References from the computer science literature

In recent years much of the stimulus for the development of aspects of order theory has come from computer science. Many of the references cited below were not available when the first edition of this book was published. They give a sample of books and papers, written from a computer science perspective, which deal with topics related to lattices and order; further references can be found in their bibliographies.

We draw attention in particular to [25], our primary source for the discussion of refinement in Chapter 7 and to the forthcoming multi-author volume [17], which treats in a unified way a spectrum of material

from pure order theory to practical program construction, and has Galois connections and fixpoint calculus as central themes.

[16] S. Abramsky, D. M. Gabbay and T. S. E. Maibaum (eds.), *Handbook of Logic in Computer Science*, Vol. 3, Semantic structures, Oxford University Press, 1994.

[17] R. Backhouse, R. Crole and J. R. Gibbons, *Algebraic and Coalgebraic Methods in the Mathematics of Program Construction*, Springer Lecture Notes in Computer Science, due for publication in 2002.

[18] A. Edelat, Domains for computation, physics and exact real arithmetic, *Bull. Symbolic Logic* **3** (1997), 401–452.

[19] Eindhoven University of Technology Mathematics of Program Construction Group, Fixed-point calculus, *Inf. Proc. Letters* **53** (1995), 131–136.

[20] C. A. Gunter, *Semantics of Programming Languages*, MIT Press, 1992.

[21] C. A. Gunter and D. S. Scott, Semantic domains, in *Handbook of Theoretical Computer Science*, Vol. B, J. van Leeuween (ed.), Elsevier, 1990, 633–674.

[22] C. A. R. Hoare and H. Jifeng, *Unifying Theories of Programming*, Prentice-Hall, 1998.

[23] J. D. Lawson, The versatile continuous order, *Lecture Notes in Computer Science* **298**, M. Main *et al.* (eds.), Springer-Verlag, 1987, 134–160.

[24] J. Loeckx and K. Sieber, *The Foundations of Program Verification*, 2nd edition, Wiley–Teubner, 1987.

[25] A. K. McIver, C. C. Morgan and J. W. Sanders, *Programs and Abstraction*. Obtainable electronically via: `ftp://ftp.comlab.ox.ac.uk/tmp/Jeff.Sanders/w2.ps`.

[26] A. Melton, D. A. Schmidt and G. E. Strecker, Galois connections and computer science applications, *Lecture Notes in Computer Science* **240**, G. Goos and J. Hartmanis (eds.), Springer-Verlag, 1986, 299–312.

[27] A. W. Roscoe, *The Theory and Practice of Concurrency*, Prentice-Hall 1997.

[28] D. A. Schmidt, *Denotational Semantics*, Allyn & Bacon, 1986.

[29] D. S. Scott, Domains for denotational semantics, *Lecture Notes in Computer Science* **140**, M. Nielsen and E. T. Schmidt (eds.), Springer-Verlag, 1982, 577–613.

[30] V. Stoltenberg-Hansen, I. Lindström and E. R. Griffor, Mathematical theory of domains, *Cambridge Tracts in Computer Science* **22**, Cambridge University Press, 1994.

[31] A. S. Tanenbaum, *Structured Computer Organization*, 4th edition, Prentice-Hall, 1999.

[32] G. Winskel and K. G. Larsen, Using information systems to solve domain equations recursively, *Lecture Notes in Computer Science* **173**, G. Kahn and G. D. Plotkin (eds.), Springer-Verlag, 1984, 109–130.

General references on the theory of ordered sets and lattices

Rather few books have been written on lattices and ordered sets. The successive editions of G. Birkhoff's *Lattice Theory* (1940, 1948 and 1967) are pioneering classics. In the 1960s George Grätzer put forward a proposal for a survey of the whole of lattice theory in depth, of which his textbook [39] on distributive lattices was originally intended as the first part. The rapid development of lattice theory in the following decade quickly made Grätzer's original objective quite impossible to attempt and his *General Lattice Theory*, which appeared in 1978, had more restricted aims. The second edition, [40], published twenty years later, includes appendices by various authors on later developments and a comprehensive bibliography. A four-volume work on the theory of algebras, in which lattices play a central role, was planned in the 1980s by Ralph McKenzie, George McNulty and Walter Taylor. Only the first volume, [42], of this advanced monograph has appeared in print.

Marcel Erné's enchantingly illustrated text [36] covers much of the same elementary material on ordered sets as we do and provides links with set theory. The forthcoming monograph [47] takes the theory of ordered sets further. The Compendium, [38], and Johnstone's treatise, [41], are essential reading, at an advanced level, for those wishing to pursue further the order-theoretic background to domain theory and the connections between order, lattices and topology.

Proceedings of conferences supplement the textbook literature. In particular we draw attention to [43]. This contains a range of interesting articles which give access through their bibliographies to applications of order theory (including those in the social sciences) which we do not have space to reference individually. In addition [43] contains a comprehensive bibliography for ordered sets, up to 1981. Lattices and ordered sets are also served by two specialist journals, *Algebra Universalis* and *Order*, which were launched in 1971 and 1984, respectively.

[33] R. Balbes and Ph. Dwinger, *Distributive Lattices*, University of Missouri Press, 1974.

[34] G. Birkhoff, *Lattice Theory*, 3rd edition, Coll. Publ., XXV, American Mathematical Society, 1967.

[35] P. Crawley and R. P. Dilworth, *Algebraic Theory of Lattices*, Prentice-Hall, 1973.

[36] M. Erné, *Einführung in die Ordnungstheorie*, Bibliographisches Institut, Mannheim, 1982.

[37] R. Freese, J. Ježek and J. B. Nation, Free lattices, *Mathematical Surveys and Monographs* **42**, Amer. Math. Soc., 1995.

[38] G. Gierz, K. H. Hofmann, K. Keimel, J. D. Lawson, M. Mislove and D. S. Scott, *A Compendium of Continuous Lattices*, Springer-Verlag, 1980.

[39] G. Grätzer, *Lattice Theory: First Concepts and Distributive Lattices*, W. H. Freeman, 1971.

[40] G. Grätzer, *General Lattice Theory*, 2nd edition, Birkhäuser Verlag, 1998.

[41] P. T. Johnstone, *Stone Spaces*, Cambridge University Press, 1982.

[42] R. N. McKenzie, G. F. McNulty and W. F. Taylor, *Algebras, Lattices and Varieties*, Vol. I, Wadsworth & Brooks/Cole, 1987.

[43] I. Rival (ed.), Ordered sets, *NATO ASI Series* **83**, Reidel, 1982.

[44] I. Rival (ed.), Graphs and order, *NATO ASI Series* **147**, Reidel, 1985.

[45] I. Rival (ed.), Algorithms and order, *NATO ASI Series* **255**, Reidel, 1989.

[46] I. G. Rosenberg and G. Sabidussi (eds.), Algebras and orders, *NATO ASI Series* C **389**, Kluwer Academic Publishers, 1993.

[47] B. Schröder, *Ordered Sets – Underlying Structure and Open Problems*, Birkhäuser Verlag (forthcoming).

Background and further reading on specialized topics

We do not attempt to attribute all the theorems in the book, but merely give references to the original sources for a few of the major landmarks as well as selected suggestions on specialized topics.

Concept analysis (Chapter 3) is still a relatively new field. The topic was introduced by R. Wille in [43], pp. 445–470. Entry to the subject and to its extensive literature is most easily made through the textbook [48].

[48] B. Ganter and R. Wille, *Formal Concept Analysis*, Springer-Verlag, 1999 (German original published in 1996 by Springer-Verlag).

The theory of **modular** and **distributive lattices** (Chapter 4 and parts of Chapters 10 and 11) is well treated in the books on lattice theory listed above. We also draw attention to two classic papers.

[49] G. Birkhoff, Applications of lattice algebra, *Proc. Camb. Phil. Soc.* **30** (1934), 115–122. [Theorem 4.10(ii).]

[50] R. Dedekind, Über die von drei Moduln erzeugte Dualgruppe, *Math. Ann.* **53** (1900), 371–403. [Modularity of $\mathcal{N}$–Sub$G$ (4.6(5)) and Theorem 4.10(i), implicitly.]

Boolean algebras have an extensive literature separate from their coverage in lattice theory sources.

[51] J. L. Bell and A. B. Slomson, *Models and Ultraproducts*, North-Holland, 1971.

[52] G. Boole, *An Investigation into the Laws of Thought*, London, 1854. (Reprinted by Dover Publications, 1951.)

[53] P. Halmos, *Lectures on Boolean Algebras*, Van Nostrand, 1963.

[54] J. D. Monk (ed.), *Handbook of Boolean Algebras*, 3 vols., North-Holland, 1989.

[55] J. E. Whitesitt, *Boolean Algebra and its Applications*, Addison-Wesley, Reading, Mass., 1961. (Reprinted by Dover Publications, 1995.)

The use of Boolean algebras in circuit design is also covered in [31].

The **duality for distributive lattices** (Chapters 5 in the finite case and Chapter 11 in general) is surveyed in [61], and [40] and a full bibliography up to 1996 can be found in the special issue of the journal *Studia Logica*, [56], devoted to Priestley duality. The historical comments made in these sources are supplemented by the introduction to [41] which puts Marshall Stone's pioneering contributions in perspective. A discussion of Stone's original, purely topological, representation of distributive lattices (equivalent to the order-topological theory we give in Chapter 11) can be found in [33] and [39]. Priestley duality is the stepping-off point for the general theory of natural dualities, see [60]. Our treatment of **Zorn's Lemma** (Chapter 10) is complemented by that in [6]. The characterization of **powerset algebras** and **down-set lattices** (Chapter 10) has a long history: the former is due to A. Tarski [65] from 1930 while the latter was discovered and rediscovered by various authors starting with [63], [57] and [59] in the 1950s.

[56] M. Adams and W. Dziobiak (eds.), *Studia Logica* Special Issue on Priestley duality, *Studia Logica* **56**, Nos. 1–2, 1996.

[57] V. K. Balachandran, A characterization of $\Sigma\Delta$-rings of subsets, *Fund. Math.* **41** (1954), 38–41.

[58] G. Birkhoff, On the combination of subalgebras, *Proc. Camb. Phil. Soc.* **29** (1933), 441–464. [Theorems 5.12 and 10.21.]

[59] G. Bruns, Verbandstheoretische Kennzeichnung vollständiger Mengenringe, *Arch. Math.* **10** (1959), 109–112.

[60] D. M. Clark and B. A. Davey, *Natural Dualities for the Working Algebraist*, Cambridge University Press, 1998.

[61] B. A. Davey and D. Duffus, Exponentiation and duality, in [28], 43–95.

[62] L. Nachbin, *Topology and Order*, Krieger Publishing Co., 1976.

[63] G. R. Raney, Completely distributive complete lattices, *Proc. Amer. Math. Soc.* **3** (1952), 677–680.

[64] M. H. Stone, The theory of representations for Boolean algebras, *Trans. Amer. Math. Soc.* **40** (1936), 37–111. [Theorem 11.4.]

[65] A. Tarski, Une contribution à la théorie de la mesure, *Fund. Math.* **15** (1930), 42–50.

Our discussion of **congruences** (Chapter 6) is rather elementary and is influenced as much by universal algebra as it is by lattice theory. For further information on the role of congruences in universal algebra we refer the reader to [3] and for a thorough treatment of congruences on lattices we refer to [39].

The material on **complete** and **algebraic lattices** (Chapter 7) is classical: they are discussed, in particular, in [33], [34], [37], [40] and [41]. The following original papers on completions are noteworthy.

[66] B. Banaschewski and G. Bruns, Injective hulls in the category of distributive lattices, *J. Reine Angew. Math.* **232** (1968), 102–109. [The source for Theorem 7.41.]

[67] H. M. MacNeille, Partially ordered sets, *Trans. Amer. Math. Soc.* **42** (1937), 416–460.

The theory of **Galois connections** (Chapter 7) is documented in many places and from many perspectives. They are discussed from an algebraic point of view in the first and subsequent editions of G. Birkhoff's *Lattice theory*. The article [68] gives further references for the topic. Nowadays, much of the incentive to study Galois connections comes from computer science. See, in particular, [17], [25] and [26], and the references therein.

[68] M. Erné, J. Koslowski, A. Melton and G. E. Strecker, A primer on Galois connections, *Annals N. Y. Acad. Sci.*, **704** (1993), 103–125.

Our treatment of **CPOs** and **fixpoint theory** (Chapters 8 and 10) has been much influenced by the work of many computer scientists, communicated orally, through unpublished notes and through texts and conference proceedings. Regarding fixpoint calculus, we should make special mention of the work of R. Backhouse and his collaborators.

The proof of CPO Fixpoint Theorem II in 8.22 is due to D. Pataraia and was unpublished as this text went to press; we are grateful for his permission to include this proof. The proof of CPO Fixpoint Theorem III, 8.23, indicated in Exercise 8.20 is due to K. H. Hoffman and is also unpublished and that of CPO Fixpoint Theorem II via (ZL) for CPOs in 10.5 is due to A. W. Roscoe [27]. The obscure history of the various fixpoint theorems and their correct attribution is discussed in [69].

[69] N. Bourbaki, Sur la théoreme de Zorn, *Arch. Math.* (*Basel*) **2** (1949/50), 434–437. [CPO Fixpoint Theorems II and III.]

[70] J.-L. Lassez, V. L. Nguyen, E. A. Sonenberg, Fixedpoint theorems and semantics: a folk tale, *Inf. Proc. Letters* **14** (1982), 112–116.

[71] G. Markowsky, Chain-complete posets and directed sets with applications, *Algebra Universalis* **6** (1976), 53–68.

[72] A. Tarski, A lattice-theoretical fixpoint theorem and its applications, *Pacific J. Math.* **5** (1955), 285–309.

As with the material on Galois connections, CPOs and fixpoint theory, our presentation of the theory of **domains** and **information systems** (Chapter 9) draws heavily on the work of theoretical computer scientists, originally through unpublished notes of D. S. Scott and A. W. Roscoe. Of the more recent literature we recommend in particular the account of domain theory given by S. Abramsky and A. Jung in [16] and the papers of A. Edelat (see [18]).

Notation Index

The symbol := denotes 'equals by definition'. The end of a proof is marked by □. The beginning of an argument by contradiction is flagged by 'suppose [by way of contradiction]' and the lightning symbol, $\lightning$, signals that the required contradiction has been reached. See for example the proof of Lemma '1.17'.

Index

Page numbers given in boldface refer to definitions and those in italic to exercises, with the latter overriding the former where definitions are given in exercises.