Chapter 1

Lattice theory

1.1 Partial orders

1.1.1 Binary Relations

A ***binary relation*** R on a set X is a set of pairs of elements of X. That is, $R \subseteq X^2$. We write xRy as a synonym for $(x, y) \in R$ and say that R *holds* at (x, y). We may also view R as a square matrix of 0's and 1's, with rows and columns each indexed by elements of X. Then $R_{xy} = 1$ just when xRy.

The following attributes of a binary relation R in the left column satisfy the corresponding conditions on the right for all x, y, and z. We abbreviate "xRy and yRz" to "$xRyRz$".

empty	$\neg(xRy)$
reflexive	xRx
irreflexive	$\neg(xRx)$
identity	$xRy \leftrightarrow x = y$
transitive	$xRyRz \rightarrow xRz$
symmetric	$xRy \rightarrow yRx$
antisymmetric	$xRyRx \rightarrow x = y$
clique	xRy

For any given X, three of these attributes are each satisfied by exactly one binary relation on X, namely empty, identity, and clique, written respectively $\emptyset$, 1_X, and K_X. As sets of pairs these are respectively the empty set, the set of all pairs (x, x), and the set of all pairs (x, y), for $x, y \in X$. As square $X \times X$ matrices these are respectively the matrix of all 0's, the matrix with 1's on the leading diagonal and 0's off the diagonal, and the matrix of all 1's.

Each of these attributes holds of R if and only if it holds of its converse $\breve{R}$, defined by $xRy \leftrightarrow y\breve{R}x$. This extends to Boolean combinations of these attributes, those formed using "and," "or," and "not," such as "reflexive and either not transitive or antisymmetric". For example we will be defining certain concepts such as partial order and equivalence relation, via such Boolean combinations, mainly using "and." Thus if R is a partial order so is $\breve{R}$, and similarly for the other concepts we define. This is the ***Duality Principle***. Its importance will emerge gradually as we make use of it.

When X is empty so is X^2, whence there is only one binary relation on X, namely the empty relation. Now all the attributes listed above are *universal* in the sense that the variables in their definitions are universally quantified: "for all x, y, z;" none are existentially quantified, as in "there exists w." By convention, when X

is empty a universal statement about X is true; we say it holds *vacuously.* Hence the single binary relation on the empty set enjoys all of these attributes.

When X has n elements, X^2 has n^2, whence there are 2^{n^2} binary relations on X.

The more general notion of binary relation is as a subset of $X \times Y$; instead of saying "on X" we then say "from X to Y." Thus we have been treating the case $X = Y$. This special case suffices for the treatment of partial orders. As we will see later the special case is to the general as monoids are to categories. Binary relations generalize further to n-ary relations as a set of n-tuples indexed from 1 to n, and yet further to I-ary relations where I is an arbitrary index set.

1.1.2 Preorders

A ***preorder*** or ***ordered set*** is a pair $(X, \leq)$ where X is a set and $\leq$ is a reflexive transitive binary relation on X. A ***partial order*** is an antisymmetric preorder. An ***equivalence relation*** is a symmetric preorder.

A familiar example of a preorder is given by the set of points in the line together with the binary relation *left-of.* This relation is certainly not symmetric. However it is antisymmetric, and hence a partial order, since two distinct points cannot both be left of each other.

We may make this example more general by taking instead the set of points in the plane, still ordered by the binary relation *left-of.* We consider vertically aligned points to be left of each other. This relation is still not symmetric, but now it is not antisymmetric either.

The relation of being vertically aligned is an equivalence relation.

The relation of being *strictly* to the left of is an irreflexive transitive relation.

The identity relation 1_X and the clique K_X are both preorders. Other names for 1_X are equality, $=$, and the ***discrete*** order, while K_X is called the ***codiscrete*** or ***chaotic*** order. The empty relation $\emptyset$ however is not a preorder because it is not reflexive, unless X is empty.

There are four reflexive binary relations on a doubleton, corresponding to the four subsets of the two pairs $(0, 1)$ and $(1, 0)$, since reflexivity requires the other two pairs. All four of these reflexive relations are transitive, giving four ways to preorder a doubleton.

In order to find a reflexive relation that is not a preorder we require 3 elements. The relation R on $\{0, 1, 2\}$ satisfying only $0R0R1R1R2R2$ (5 pairs) is reflexive but not transitive.

Each of the set of integers, the set of rationals, and the set of reals forms a partial order under their usual ordering.

Any set of subsets of a set, ordered by set inclusion, forms a partial order. For if $X \subseteq Y \subseteq X$ then for every element x, $x \in X$ iff $x \in Y$, whence $X = Y$.

The natural numbers ordered by the relation $x|y$, defined as $y = zx$ for some natural number z, is a partial order with least element 1 and greatest element 0.

Finite partial orders are conveniently depicted as ***Hasse diagrams.*** This is a two-dimensional representation of a directed acyclic graph all of whose edges are drawn without arrowheads but which are assumed to be directed upwards.

1.1.3 Suborders

$(X, \leq)$ is a ***suborder*** of $(Y, \leq')$ when $X \subseteq Y$ and $\leq$ is the restriction of $\leq'$ to X. That is, for all $x, x' \in X$, $x \leq x'$ if and only if $x \leq' x'$.

The set $(\mathbb{Z}, \leq)$ of integers with their usual order is a suborder of the set $(\mathbb{R}, \leq)$ of reals with their usual order. Any set of subsets of X ordered by inclusion is a suborder of the power set of X ordered by inclusion.

The properties of reflexivity, irreflexivity, transitivity, symmetry, antisymmetry, and cliquehood are all ***preserved by*** suborders. That is, if a preorder is reflexive so are its suborders, and so on. Hence any suborder of a preorder is a preorder, any suborder of a partial order is a partial order, and similarly for equivalence relations.

1.1.4 Direct Products

The ***direct product*** $(X, \leq)$ of a pair $((Y_1, \leq_1), (Y_2, \leq_2))$ consists of the set X of all pairs (y_1, y_2) such that $y_1 \in Y_1$ and $y_2 \in Y_2$, and ordered by $\leq$ such that $(y_1, y_2,) \leq (y_1' y_2')$ just when $y_1 \leq_1 y_1'$ and $y_2 \leq_2 y_2'$.

For example the direct product $(\mathbb{R}, \leq)^2$ of the ordered set of reals with itself is the real plane ordered so that any two points p and q satisfy $p \leq q$ just when p is below and to the left of q. The direct product of the unit interval $[0, 1]$ with the set of all reals is a strip in the plane of unit width or height (depending on which axis we associate each of the two sets with), ordered as for $(\mathbb{R}, \leq)^2$.

For an arbitrary index set I, $(X, \leq)$ is the *direct product* of the family $(Y_i, \leq_i)_{i \in I}$ when X is the cartesian product of $(Y_i)_{i \in I}$ and $x \leq x'$ just when $x_i \leq_i x_i'$ for all $i \in I$.

1.1.5 Monotone functions

Given two preorders $(X, \leq_X)$ and $(Y, \leq_Y)$, a function $f : X \to Y$ is called ***monotone*** when $x \leq_X x'$ implies $f(x) \leq_Y f(x')$.

The squaring function on the natural numbers is monotone with respect to their usual order, as can be inferred from the nonnegative slope of its graph. However the squaring function on the integers is not monotone, since for example we have $-2 \leq -1$ but not $(-2)^2 \leq (-1)^2$. This is clear from the negative slope of x^2 for $x < 0$.

The identity function on any preorder, defined as $f(x) = x$, is monotone. So is any constant function.

Given three sets X, Y, and Z and functions $f : X \to Y$, $g : Y \to Z$, the ***composition*** $g \circ f : X \to Z$ of g with f satisfies $(g \circ f)(x) = g(f(x))$. Sometimes $g \circ f$ is written as $f; g$.

It can be seen that the composition of monotone functions is a monotone function.

Two monotone functions $f : X \to Y$ and $g : Y \to X$ form an ***isomorphism pair*** when $g \circ f = 1_X$ and $f \circ g = 1_Y$. When there exists such an isomorphism pair between X and Y we say that X and Y are *isomorphic.* Isomorphism of preorders $(X, \leq)$ and $(Y, \leq)$ formalizes the intuitive idea that the two orders are identical except for the choice of underlying sets.

For example the ordered sets $(\mathbb{Z}, \leq)$ and $(\mathbb{Z}, \geq)$ consisting of the integers ordered in each direction are isomorphic, with the isomorphism given by negation, and more generally by any function $f_n : \mathbb{Z} \to \mathbb{Z}$ defined as $f_n(x) = n - x$.

1.1.6 Cliques

A clique *of* a preorder P is a suborder of that preorder that is a clique. A ***maximal*** clique P is one that is not a proper subset of any clique of P.

The ***skeleton*** of a preorder P is the partial order whose elements are the maximal cliques of P. (This term is borrowed from category theory.)

We shall call an ***endoskeleton*** of a preorder P a suborder of P consisting of one element from each maximal clique of P.

Cliques carry no more information than their underlying sets, and every preorder can be represented as a partially ordered set of cliques. Up to isomorphism therefore, a preorder can be represented as a partially ordered set labeled with numbers, each label giving the cardinality of its associated clique. To be more precise the numbers should be cardinals, since they may be infinite.

1.1.7 Chains

Elements x and y are ***comparable*** in a preorder when either $x \leq y$ or $y \leq x$, and otherwise are ***incomparable***.

The binary relation of comparability may be seen to be reflexive and symmetric but not in general transitive.

A ***chain*** or ***linear order*** or ***total order*** is a partial order in which all pairs of elements are comparable.

A preorder $(Y, \leq')$ ***augments*** $(X, \leq)$ when $Y = X$ and $x \leq y$ implies $x \leq' y$. Hence a chain can be described as a partial order with no proper augment that is a partial order. (But a chain can always be augmented to a clique.)

A ***linearization*** of a partial order P is a chain augmenting P, i.e. a maximal antisymmetric augment of P.

Theorem 1 *Every partial order $(X, \leq)$ in which x and y are incomparable has an augment in which they are comparable.*

Proof: Form $\leq'$ by adding to $\leq$ all pairs (x', y') for which $x' \leq x$ and $y \leq y'$. The result contains (x, y) since $x \leq x$ and $y \leq y$. It remains reflexive since nothing is removed. It is transitive because for any triple $x' \leq' y' \leq z$, where $x' \leq' y'$ is one of the added pairs, we have $y \leq y' \leq z$ whence (x', z) will also have been added, and similarly for $z \leq x' \leq' y'$. It is antisymmetric because if $x' \leq' y' \leq' x'$ then $y \leq y' \leq' x' \leq x$ contradicting incomparability of x and y. Hence it is a partial order extending $(X, \leq)$ and containing (x, y). ∎

Even though in some cases we may have added infinitely many pairs by this construction merely to get $x \leq y$, the construction is actually minimal: it yields the *least* augment of P for which $x \leq y$ in that augment. For if $x' \leq x$ and $y \leq y'$, then in any augment of P containing $x \leq y$ must contain $x' \leq y'$, by transitivity.

A ***maximal*** chain of a partial order is one such that every element not in the chain is incomparable with some element of the chain.

Kuratowski's Lemma. Every partial order contains a maximal chain.

This is one of a number of equivalent forms of the Axiom of Choice. Another form is that for every set X of disjoint nonempty sets there exists a set Y such that for every set $Z \in X$, $Y \cap Z$ is a singleton. (Thus Y "chooses" a representative of each set in X.) A closely related form is that the cartesian product of a family X_i of nonempty sets is nonempty. Hence there exists an I-tuple in the product, which makes a choice of an element from each X_i. Yet another form is that every set can be well-ordered. (An ordered set is ***well-ordered*** when every suborder has a least element.)

The proposition of exercise 8 at the end of the section is also often referred to as Kuratowski's Lemma; the two are close enough that little harm can result from using the same name for both.

These alternatives can all be derived from each other, so the question of which to take as the *the* Axiom of Choice is arbitrary. A reasonable choice is whichever one seems either the most or least obvious to you, depending on whether you are for or against it.

Our next mission is to show that every partial order has a linearization. The following notions will be helpful here.

Given a family $(X, \leq_i)_{i \in I}$ of preorders on a fixed set X, its *union* $(X, \bigcup_i \leq_i)$ is defined so that $x(\bigcup_i \leq_i)y$ just when there exists $i \in I$ for which $x \leq_i y$. This is the ordinary notion of union applied to the $\leq_i$'s regarded as subsets of X^2.

Lemma 2 *If the family $(X, \leq_i)_{i \in I}$ forms a chain under augmentation (that is, for every pair $\leq_i$, $\leq_j$ one is a subset of the other), its union $(X, \bigcup_i \leq_i)$ is a partial order.*

(It is important to realize that there are two levels of partial ordering here: we are partially ordering partial orders, and the elements of the chain in the lemma are themselves partial orders $(X, \leq_i)$.)

Proof: Abbreviate $\bigcup_i \leq_i$ to just $\leq$. Every $\leq_i$ is reflexive whence so is $\leq$. For transitivity suppose $x \leq y \leq z$. Then $x \leq_i y \leq_j z$ for some i and j. Let $\leq'$ denote whichever of $\leq_i$ and $\leq_j$ augments the other. Then $x \leq' y \leq' z$, whence $x \leq' z$, whence $x \leq z$. For antisymmetry, suppose $x \leq y \leq x$. Then by the same reasoning as for transitivity we have $x \leq' y \leq' x$, whence $x = y$. ■

Theorem 3 *(AC) Every partial order $(X, \leq)$ has a linearization.*

(The "(AC)" is shorthand for "Assuming the Axiom of Choice, ".)

Proof: Take the set of all partial orderings of X, ordered by augmentation, and take a maximal chain containing the given partial order, possible by Kuratowski's Lemma. By the preceding lemma the union of this chain is a partial order. By the previous theorem, if any two elements are incomparable in this order then the order may be extended, contradicting its maximality. Hence it is linear. ■

Theorem 4 *(AC) Given any two incomparable elements of a partial order, there exist linearizations of the order containing these elements in either order.*

Proof: Use the above two theorems, first to extend the partial order in each of the two ways in which $x \leq y$ and $y \leq x$, and then to extend the results to two chains. ■

Corollary 5 *(AC) The intersection of the set of linearizations of a partial order is that partial order.*

1.1.8 Exercises

1. For each of the attributes empty, reflexive, irreflexive, identity, symmetric, antisymmetric, and clique, give an expression in n for the number of binary relations on X having that attribute.

2. For each attribute φ listed above determine the size of the smallest X such that (i) for every attribute ψ above there exists a relation on X satisfying φ but not ψ; (ii) there exists a relation R on X such that X satisfies φ but no other attribute on the list.

3. For any set X, show that the partial orders on X are in 1-1 correspondence with the irreflexive transitive binary relations on X.

4. Formalize and prove the statement that a preorder partially orders its maximal cliques.

5. Show that every endoskeleton of a poset P is isomorphic to the skeleton of P.

6. Show that a partial order is a chain just when its complement is transitive.

7. Show that any two finite chains of the same cardinality are isomorphic. (In contrast it can be shown that there are uncountably many nonisomorphic linear orderings of a countably infinite set. Of these, only four are dense (between every two elements lies another), distinguished by whether they have a least and/or a greatest element.)

8. (Kuratowski's Lemma) Show that every chain in a partial order extends to a maximal chain. (That is, every partial order contains a maximal chain having a specified subchain.)

9. Show that union of binary relations preserves reflexivity and symmetry, but not in general transitivity or antisymmetry.

1.2 Lattices

1.2.1 Bounds

Let $(X, \leq)$ be a poset. We say that an element $x \in X$ is an ***upper bound*** of a subset $Y \subseteq X$, or that x ***bounds*** Y from above, when for all $y \in Y$, $y \leq x$. The dual notion is ***lower bound***: x bounds Y from below when $x \leq y$ for all $y \in Y$. By the Duality Principle everything we have to say below about upper bounds applies equally well, with $\geq$ in place of $\leq$, to lower bounds.

The ***least*** upper bound (lub) of Y, also called ***supremum*** (sup), is that upper bound of Y which is less or equal to every upper bound of Y. A set Y need not have a sup, for example the set of integers in the real line, or for that matter the line itself.

However when the sup does exist it is unique. For if x and x' are both upper bounds then in order for them to both be least we must have $x \leq x'$ and $x' \leq x$, whence $x = x'$ by antisymmetry.

The sup of a set may or may not belong to the set. For example the closed interval $[0, 1]$ of reals contains its sup, namely 1, whereas neither the open interval $(0, 1)$ nor the half open interval $[0, 1)$ contain their sup, which again is 1 in each case.

The dual notion to least upper bound is ***greatest lower bound*** (glb), or ***infimum*** (inf).

1.2.2 Semilattices

First definition. A ***semilattice*** is a partial order $(X, \leq)$ in which every doubleton $\{x, y\}$ has a least upper bound, denoted $x \vee y$ and called the ***join*** of x and y. Even though the relation $\leq$ is partial (i.e. not linear as an order), the operation $\vee$ is total ($x \vee y$ is well-defined for all elements x, y of X).

With this definition there is a dual notion, that of ***lower semilattice*** (so "semilattice" in the above means "upper semilattice"), in which every doubleton has a greatest lower bound, denoted $x \wedge y$ and called their ***meet***.

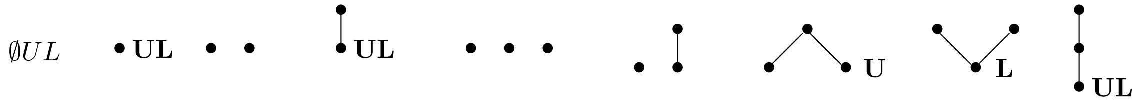
Examples.

1. Any chain forms both an upper and a lower semilattice, with $x \vee y$ being the max (larger) of x and y and $x \wedge y$ being their min (smaller).

2. The power set of a set, ordered by inclusion, forms an upper semilattice with $\vee$ as union, and a lower semilattice with $\wedge$ as intersection.

3. The set of all strings over some alphabet, ordered by the prefix relation ($x \leq xy$ for all strings x, y), forms a lower semilattice, $x \wedge y$ being the longest common prefix of strings x and y. It does not however form an upper semilattice: if x and y do not stand in the prefix relation then they have no upper bound at all, let alone a least such.

4. Here are all the posets (up to isomorphism) with at most 3 elements, marked U if an upper semilattice and L if a lower.



(The empty set $\emptyset$ is vacuously both an upper and lower semilattice, and likewise the singleton, almost as vacuously.)

We may define $\vee$ in terms of $\leq$ via the following equivalence.

$$x \vee y \leq z \quad \text{iff} \quad x \leq z \text{ and } y \leq z \tag{1.1}$$

When this equivalence is satisfied, we deduce that $x \vee y$ is a common upper bound of x and y by setting $z = x \vee y$, thereby satisfying the left hand side and hence the right. We deduce that it is the least upper bound by taking z to be any common upper bound of x and y, thereby satisfying the right hand side and hence the left.

Exercise 1 below asks you to show that the equational theory of a semilattice includes the equations $x \vee (y \vee z) = (x \vee y) \vee z$ (associativity), $x \vee y = y \vee x$ (commutativity), and $x \vee x = x$ (idempotence).

Second Definition. A semilattice $(X, \vee)$ is a pair consisting of a set X and a binary operation $\vee$ which is associative, commutative, and idempotent.

This is our first example of an algebra. An ***algebra*** is defined to consist of a set and operations of various arities on that set.

With this definition there is no classification of semilattices into upper and lower, only a pure notion of semilattice. Changing the notation to $(X, \wedge)$ does not change the concept, only the symbol.

But then we may define $\leq$ in terms of $\vee$ by defining $x \leq y$ as $x \vee y = y$. Exercise 2 asks you to show that $\leq$ so defined is a partial order, and that Equation 1.1 then gives back the binary operation we started with.

Here the opportunity arises for duality to enter: we could just as well have defined $x \leq y$ as $x \vee y = x$, although for sanity we would then have either written the given operation as $\wedge$ or the resulting relation as $x \geq y$.

1.2.3 Lattices

A ***lattice*** is a poset that is simultaneously an upper semilattice and a lower semilattice.

Examples. Every chain is a lattice. The power set $(2^X, \subseteq)$ is a lattice, with union and intersection as join and meet respectively. The natural numbers ordered by $x|y$ is a lattice with lcm and gcd as join and meet.

To axiomatize the theory of lattices requires more than just the axioms of the theory of semilattices together with those of lower semilattices. For one thing there is nothing in the axioms thus far to prevent $\vee$ and $\wedge$ being the same operation.

We add to the axioms for semilattices and lower semilattices the *absorption laws*, $x = x \vee (x \wedge y)$ and its dual $x = x \wedge (x \vee y)$.

Thus a lattice is an algebra $(X, \vee, \wedge)$ satisfying equations expressing associativity, commutativity, and idempotence of $\vee$ and $\wedge$, and satisfying the two absorption equations. The class of lattices is thus a finitely axiomatized equational class.

1.2.4 Distributive Lattices

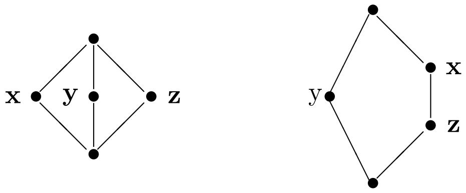
A lattice of fundamental importance is the two-element chain $(2, \vee, \wedge)$. It is the only two-element lattice. This lattice features prominently in logic as the lattice of truth values.

The equational theory of the two-element lattice goes beyond that of lattices, for it includes the ***distributivity law***

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z).$$

A ***distributive lattice*** is a lattice satisfying this law.

Despite the 19th century logician Peirce's conviction otherwise, not every lattice is distributive. The canonical counterexamples are the following two lattices, labeled to indicate the breakdown.



Although $(x \wedge y) \vee (x \wedge z) \leq x \wedge (y \vee z)$ holds in these examples, and indeed in every lattice (Exercise 4), $(x \wedge y) \vee (x \wedge z) \geq x \wedge (y \vee z)$ fails in both.

It is natural to ask whether there are yet more laws to be discovered for $(2, \vee, \wedge)$. The answer is that we now have a complete axiomatization: the theory of $(2, \vee, \wedge)$ is just the theory of distributive lattices. Before stating and proving this theorem we first define some auxiliary notions.

Define a ***lattice term over a set*** V of variables to be a finite binary tree[1] whose leaves are labeled with variables from V and whose remaining nodes are labeled with $\wedge$ or $\vee$.

An ***order filter*** of a poset $(X, \leq)$ is a subset $Y \subseteq X$ such that for all $y \in Y$ and $x \in X$, if $y \leq x$ then $x \in Y$. (The dual notion is ***order ideal***.) An order filter is ***proper*** when it is a proper subset of X, i.e. not X itself, and is ***principal*** when it has a least element.

To each nonempty proper[2] order filter A of the poset $(2^V, \subseteq)$ (the set of all subsets of V ordered by inclusion) associate a term formed as a join of meets, called a ***normal*** term.[3] For each element of A, a nonempty set of variables since A is proper, form the meet of those variables. Then form the join of all those meets, of which there is at least one since A is nonempty.

This specification determines the term only up to associativity and commutativity of its operations.[4] Since these remaining choices do not affect the value of the term we may safely dispense with a more detailed specification and simply regard normal terms as only being defined up to associativity and commutativity.[5]

In any lattice L the normal term associated to A denotes the function $L^V \to L$ obtained from an assignment of elements of L to variables (i.e. an element of L^V) by evaluating the term. On reflection one sees that for

[1] In the sense of a parse tree, i.e. these are trees having a root node, with all edges directed away from the root, and the order of the immediate descendants of each node matters.

[2] When working with lattices with a constant 0 denoting the least element we allow the empty filter as well. Independently, in the presence of a constant 1 denoting the greatest element we also allow the improper filter.

[3] Some readers will recognize this as disjunctive normal form with the additional requirement that the set of disjunctions is closed under conjoining disjuncts with variables.

[4] Idempotence however is catered for, namely by by working with subsets of V and 2^V rather than multisets.

[5] Had we insisted on a specific term we could alphabetize the variables within each meet, and then put the meets in a suitably defined lexicographic order, and associate each meet and the join say to the left, at the expense of pointlessly longer arguments.

the lattice $(2, \vee, \wedge)$ this function is the characteristic function of the given order filter A, i.e. the function from 2^V to 2 which is 1 on the members of A and 0 on the nonmembers.

Another name for "order filter of 2^V" is "monotone V-ary Boolean operation." With $n = |V|$ there are 2^{2^n} n-ary Boolean operations. The fraction of these which are monotone, for n from 1 to 7, is 3/4, 6/16, 20/256, 168/65536, 7581/4294967296, 7828354/18446744073709551616, and 2414682040998/340282366920938463463 374607431768211456. While the denominators are easily enumerated, having a simple closed form, the paradoxically harder task of enumerating the smaller numerators goes by the name of *Dedekind's problem.* There being two constant functions of each arity, we must subtract 2 from each numerator for the case at hand of distributive lattices without constants.

Lemma 6 *The normal terms are in bijection with the nonconstant monotone Boolean operations.*

Proof: This follows directly from the fact that a normal term denotes the monotone function that produced it, and from the definition of normal term as that produced by a nonconstant monotone function, i.e. nonempty proper order ideal, from 2^V to 2. ∎

Lemma 7 *Every lattice-theoretic term φ is equivalent to a normal term, with the equivalence holding in every distributive lattice and hence being an equation of the theory of distributive lattices.*

Proof: We begin by "pushing down the "$\wedge$'s." Choose any subterm of φ of the form $x \wedge (y \vee z)$ such that x does not contain subterms of that form (otherwise we would choose the latter subterm), and rewrite it as $(x \wedge y) \vee (x \wedge z)$. This transformation is justified by the distributivity law in the sense that, for any distributive lattice L and any assignment of elements of L to variables of φ, the transformation leaves unchanged the value of every subterm. (While this should be intuitively clear just by reflecting on how evaluation works, a formal proof would proceed by induction on the height of subterms starting from the variables, whose values the transformation leaves unchanged.)

Now define an ***inversion*** of φ to be a pair consisting of an occurrence of a $\wedge$ above an occurrence of a $\vee$ in the parse tree of φ, not necessarily immediately above. A rewriting step of the above kind eliminates the inversion at the top of the rewritten subterm, and creates no new inversions. (Although x is duplicated it contains no inversions, and although the $\wedge$ is duplicated no inversion associated with that $\wedge$ is duplicated.) Hence this transformation can be repeated only as many times as there were inversions in φ at the outset.

The result of so transforming φ exhaustively is a term of the form $\psi(\mu_1, \ldots, \mu_k)$ where the operations of ψ are all $\vee$ and those of the μ_i's are all $\wedge$.

Duplicate variables within each meet may now be eliminated by associativity and commutativity to bring the duplicates together. and then eliminating them using idempotence of $\wedge$. The same technique permits duplicate meets to be eliminated.

We now have a subset of 2^V which we must make an order filter. If there exists a meet μ_i and a variable p of φ such that $\mu_i \wedge p \neq \mu_j$ for any $j \leq k$, then expand the set of μ's by taking $\mu_{k+1} = \mu_i \wedge p$. This step is justified by $x = x \vee (x \wedge y)$ with $x = \mu_i$ and $y = p$, and we say that μ_i ***subsumes*** $\mu_i \wedge p$. Iterate until the set is closed under conjunction with variables of φ. The result is now a normal term equivalent in every distributive lattice to the one we started with. ∎

We now have enough ammunition to make short work of the theorem promised at the beginning of this section.

Theorem 8 *The equational theory of* **2** *is that of distributive lattices.*

Proof: It suffices to show that every equational property of **2** holds in an arbitrary distributive lattice. Now for $\varphi = \psi$ to hold in **2** is the same thing as saying that they denote the same monotone Boolean formula.

But then by Lemma 7 they are equivalent in every distributive lattice to a common normal term, and hence to each other. ∎

1.2.5 Exercises

1. Show that the operation $\vee$ defined by Equation 1.1 is commutative, associative, and idempotent.

2. (i) Given an associative commutative idempotent binary operation $\vee$, show that the binary relation $x \leq y$ defined by $x \vee y = y$ is a partial order. (ii) Show that plugging this partial order into Equation 1.1 yields the given $\vee$.

3. Give an example of an algebra $(L, \vee, \wedge)$ satisfying associativity, commutativity, and idempotence of both operations and the first absorption equation but not the second.

4. Express the inequality $(x \wedge y) \vee (x \wedge z) \leq x \wedge (y \vee z)$ as an equation. Show that the inequality (and hence the equation) holds in every lattice.

5. Show that a lattice is distributive if and only if it satisfies $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$ (the dual distributivity law).

6. For each of the following subsets of the ordered set $(2^X, \subseteq)$ indicate whether it is an order ideal and whether it is an order filter. In each case if it is such then indicate whether it is principal. If the answer depends on properties of X, say how. (i) All subsets. (ii) The nonempty subsets. (iii) The proper subsets. (iv) Those subsets containing a given element x. (v) Those subsets not containing a given element x.

1.3 Monoids

A ***semigroup*** is a pair $(X, \cdot)$ where X is a set and $\cdot$ is an associative binary operation, one satisfying $x \cdot (y \cdot z) = (x \cdot y) \cdot z$. The operation is generically called *multiplication* and its result the *product* of its arguments.

A ***monoid*** is a triple $(X, \cdot, 1)$ where $(X, \cdot)$ is a semigroup and 1 is an element of X satisfying $1 \cdot x = x = x \cdot 1$. We call 1 the ***identity*** of the monoid.

A ***group*** G is a monoid such that for every $x \in G$ there exists $y \in G$, called the ***inverse*** of x, satisfying $x \cdot y = 1 = y \cdot x$. This condition uniquely determines the inverse of x, for if x has two inverses y and y', we have $y = y \cdot 1 = y \cdot x \cdot y' = 1 \cdot y' = y'$.

A ***commutative*** monoid satisfies $x \cdot y = y \cdot x$. Commutative groups are called ***abelian***.

Examples

1. Any singleton $\{x\}$ with a binary operation and a constant trivially forms a one-element semigroup, monoid, and group, and abelian at that.

2. The triple $(\mathbb{N}, +, 0)$ of natural numbers with the addition operation + and the additive identity 0 forms a commutative monoid.

3. The triple $(\mathbb{Z}, +, 0)$ of integers forms an abelian group.

4. Both $(\mathbb{N}, \times, 1)$ and $(\mathbb{Z}, \times, 1)$ are monoids, but neither is a group. However $(\mathbb{Q} - \{0\}, \times, 1)$, the nonzero rationals under multiplication, is an abelian group.

5. $(\mathbb{Z}_n, +, 0)$ is the abelian group of integers modulo n.

6. The set X^* of all finite strings over a set X (the *alphabet* of X^*) forms a monoid under concatenation of strings with identity element the empty string. It is commutative just when X has at most one element, in which case it is (isomorphic to) $(\mathbb{N}, +, 0)$.

7. The set F of all functions $f : A \to A$ forms a monoid under function composition, with the identity function providing the identity.

8. For $n \geq 1$, the set of all $n \times n$ integer-valued matrices forms a monoid under matrix multiplication, with the $n \times n$ identity matrix I_{nn} as the monoid identity.

9. The set of all functions on the reals, under addition of functions defined by $(f+g)(x) = f(x)+g(x)$, with the constantly zero function as identity, forms a commutative monoid.

Monoids are not the exclusive domain of the mathematician. Here is a basic example from electrical engineering.

Suppose we have inexhaustible supplies of each of several types of electronic black boxes each having a male connector at one end and a matching female connector at the other end, with all male connectors being identical and likewise for the female. Associated with each type of box is its *behavior*, defined in some way. We may plug these boxes together (in a straight line - no loops allowed) to form assemblies which also have certain behaviors. This determines an operation xy on behaviors, such that xy is the behavior of the assembly obtained by plugging the output of an assembly with behavior x into the input of one with behavior y.

The order in which the two connections are made when building a three-box assembly does not matter — to build xyz you can start with either xy or yz. This shows that the operation is associative and hence that the set $\mathcal{M}$ of all behaviors obtainable in this way form a semigroup. Treating the absence of boxes as an assembly, $\mathcal{M}$ is a monoid.

1.3.1 Submonoids

Given a monoid $\mathcal{M} = (X, \cdot, 1)$, we say that a subset Y of X is ***closed*** when it contains 1, and for all y, y' in Y, $y \cdot y'$ is also in Y. Let $\cdot_Y : Y^2 \to Y$ be the binary operation on Y satisfying $y \cdot_Y y' = y \cdot y'$ for all y, y' in Y; call $\cdot_Y$ the ***restriction*** of $\cdot$ to Y. Then $(Y, \cdot_Y, 1)$ must be a monoid. For if not it must violate one of the three monoid equations, but then $\mathcal{M}$ would violate that equation with the same values for the variables of that equation.

We call such a $(Y, \cdot_Y, 1)$ a ***submonoid*** of $(X, \cdot, 1)$.

For example the monoid $(\mathbb{N}, +, 0)$ of natural numbers under addition is a submonoid of the monoid $(\mathbb{Z}, +, 0)$ of integers under addition.

The monoid of all functions $f : A \to A$, denoted A^A, yields many examples of useful submonoids: its injections, its surjections, its bijections (constituting the largest submonoid of A^A that is a group), its continuous functions (defined for an appropriate topology on A), its monotone functions (when A is partially ordered), and so on. In turn A^A is itself a submonoid, namely of the monoid of all binary relations on A under composition of binary relations.

The monoid $(\mathbb{Z}_p, \times, 1)$ of integers modulo a prime p under multiplication has a submonoid forming an Abelian group, obtained simply by omitting 0.

Related notions are subsemigroup and subgroup. A subsemigroup of a semigroup is a subset of the semigroup closed under the operation of the semigroup. A subgroup of a group is a nonempty subset of the group closed under both the unary and binary operations of the group (and hence containing the identity).

The union of two submonoids of $\mathcal{M}$ need not be a monoid. For example consider the submonoid of $(\mathbb{N}, +, 0)$ consisting of all even numbers, and that consisting of all multiples of 3. Their union contains 2 and 3 but not 5 and hence is not a submonoid of $(\mathbb{N}, +, 0)$.

Intersection however is better behaved.

Lemma 9 *The intersection $\bigcap S$ of any set S of submonoids of $\mathcal{M}$ is a submonoid of $\mathcal{M}$.*

Proof: For any k-ary operation f_j and elements $x_1, \ldots, x_k$ in the intersection, $f_j(x_1, \ldots, x_k)$ must be in every submonoid in S, whence it must be in the intersection. ∎

It follows that the submonoids of an algebra form a complete lattice, with $\bigwedge$ provided by intersection. However $\bigvee$ is not union. Let us now see just what $\bigvee$ is.

A subset Y of a monoid $\mathcal{M}$ is said to ***generate*** $\mathcal{M}$ when $\mathcal{M}$ has no proper submonoid which includes Y. For example $(\mathbb{N}, +, 0)$ is generated by $\{1\}$. Furthermore any set of generators of $(\mathbb{N}, +, 0)$ must contain 1, since omitting 1 yields a proper submonoid.

Proposition 1 *Let $\mathcal{M} = (X, \cdot, 1)$ be a monoid, and let Y be a subset of X. Then $\mathcal{M}' = (Z, \cdot', 1)$ is a submonoid of $\mathcal{M}$ generated by Y if and only if $\mathcal{M}'$ is the intersection of all submonoids of $\mathcal{M}$ containing Y.*

Proof: (If) Y is a subset of $\mathcal{M}'$ because the intersection of a set of sets each including Y must itself include Y. $\mathcal{M}'$ is an intersection of submonoids and hence a submonoid. So no proper submonoid of $\mathcal{M}'$ can include Y, whence $\mathcal{M}'$ is generated by Y.

(Only if) If $\mathcal{M}'$ is a submonoid of $\mathcal{M}$ generated by Y, and $\mathcal{M}''$ is a submonoid of $\mathcal{M}$ including Y, then the intersection of $\mathcal{M}'$ and $\mathcal{M}''$ is a submonoid of them both and including Y. But $\mathcal{M}'$ has no such proper submonoids, whence the intersection is M', making $\mathcal{M}'$ a submonoid of $\mathcal{M}''$. Hence $\mathcal{M}'$ is a submonoid of every submonoid of $\mathcal{M}$ including Y, and hence must be the intersection of the set of such. ∎

Another way to say this is that Y generates the least submonoid of $\mathcal{M}$ that includes Y. It can be seen that there is exactly one submonoid of $\mathcal{M}$ generated by any given subset Y of $\mathcal{M}$.

Instead of forming the submonoid generated by Y as an intersection, we may form it as a union as follows.

Given a monoid $\mathcal{M} = (X, \cdot, 1)$, let $F_M : 2^X \to 2^X$ be defined by $F_M(Y) = Y \cup \{1\} \cup \{y \cdot y' \mid y, y' \in Y\}$. Thus $F_M(Y)$ consists of the elements of Y together with 1 and those elements that can be formed from elements of Y via one application of $\cdot$. Define $F^i_M(Y) = F(F(\ldots F(Y)\ldots))$ (F_M repeated i times), and define $F^*_M(Y) = \bigcup_{i \geq 1} F^i(Y)$.

Thus $F^*_M(Y)$ consists of those elements of the monoid representable as a finite product of elements of Y, including Y itself and the empty product 1.

Proposition 2 *$F^*_M(Y)$ is the submonoid of $\mathcal{M}$ generated by Y.*

Proof: Any two elements y, y' in $F^*_M(Y)$ must be in $F^i_M(Y)$ for some i, whence $y \cdot y'$ is in $F^{i+1}_M(Y)$ and hence in $F^*_M(Y)$. Moreover 1 is in $F(Y)$ and hence in $F^*_M(Y)$. Hence $F^*_M(Y)$ is closed. Moreover $F^*_M(Y)$ includes Y since $F_M(Y)$ does. Thus $F^*_M(Y)$ is a monoid including Y. Now for any submonoid $\mathcal{M}' \subseteq \mathcal{M}$ that includes Y we have $F_M(Y) \subseteq F_M(M') = M'$, and by induction $F^i_M(Y) \subseteq \mathcal{M}'$ for all $i \geq 1$. Hence $F^*_M(Y) \subseteq \mathcal{M}'$. Thus $F^*_M(Y)$ is the least monoid including Y. ∎

1.3.2 Direct Products

The ***direct product*** $(X_1, \cdot_1, 1_1) \times (X_2, \cdot_2, 1_2)$ of two monoids is $(X_1 \times X_2, \cdot, (1_1, 1_2))$ where $\cdot : (X_1 \times X_2)^2 \to X_1 \times X_2$ is a binary operation defined as $(x_1, x_2) \cdot (y_1, y_2) = (x_1 \cdot_1 y_1, x_2 \cdot_2 y_2)$.

It can be seen that $(1_1, 1_2)$ is the identity for $\cdot$. Furthermore $\cdot$ is associative. For if it were not then associativity would fail in in either coordinate 1 or 2, whence associativity would fail for the corresponding monoid, a contradiction. Hence the direct product of two monoids is itself a monoid.

For example the direct product of the monoid $(\mathbb{N}, +, 0)$ with itself yields the lattice points (points with integer coordinates) of the upper right quadrant of the plane, under the operation of vector addition $((x, y)+(x', y') = (x + x', y + y'))$, with identity the origin. With $\mathbb{R}$ in place of $\mathbb{N}$ we obtain the whole quadrant.

Direct product generalizes to an arbitrary family straightforwardly. The ***direct product*** of a family $(X_i, \cdot_i, 1_i)_{i\in I}$ is $(\prod_{i\in I} X_i, \cdot, \mathbf{1})$ where $(x \cdot y)_i = x_i \cdot_i y_i$ and $\mathbf{1}_i = 1_i$.

That is, the direct product of a family has for its underlying set the cartesian product of the family of underlying sets, and is equipped with a binary operation $\cdot$ which acts as $\cdot_i$ in the i-th coordinate and a constant $\mathbf{1}$ which acts as 1_i in the i-th coordinate.

The direct product of semigroups and groups is defined similarly. The direct product of a family of semigroups is as for monoids but with the identity $(1_i)_{i\in I}$ omitted. The direct product of a family of groups is as for monoids but with an inverse operation defined as $(x_i)^{-1}_{i\in I} = (x^{-1}_i)_{i\in I}$.

1.3.3 Monoid Homomorphisms

A ***monoid homomorphism*** from $(X, \cdot_X, 1_X)$ to $(Y, \cdot_Y, 1_Y)$ is a function $f : X \to Y$ such that (i) $f(x \cdot_X x') = f(x) \cdot_Y f(x')$ for all $x, x' \in X$, and (ii) $f(1_X) = 1_Y$. The second condition is omitted for semigroup homomorphisms. A group homomorphism is just a monoid homomorphism, in which case it automatically satisfies the condition $h(x^{-1}) = h(x)^{-1}$, since $h(x)h(x^{-1}) = h(xx^{-1}) = h(1) = 1$ whence $h(x^{-1}) = h(x)^{-1}$.

Monoid homomorphisms, or just homomorphisms in this chapter, are traditionally referred to as ***linear*** functions.

The function $h : (N, +, 0) \to (Z_n, +, 0)$ defined as $h(m) = m \mathrm{mod} n$ is a homomorphism, as is $h : (Z, +, 0) \to (Z_n, +, 0)$ defined similarly.

Consider $(\mathbb{R}^{\mathbb{R}}, +, 0)$, the monoid under addition of all functions on the reals. The derivative operator on the submonoid consisting of the differentiable such functions is a homomorphism, since $d(f + g)/dx = df/dx + dg/dx$ and $d0/dx = 0$.

The identity function on a monoid $\mathcal{M}$ is trivially a monoid homomorphism.

The restriction of the identity function to a submonoid $\mathcal{M}'$ of $\mathcal{M}$ is an injective homomorphism from $\mathcal{M}'$ into $\mathcal{M}$, called the ***inclusion*** of $\mathcal{M}'$ into $\mathcal{M}$.

The function $\pi_1 : \mathcal{M} \times \mathbf{N} \to \mathcal{M}$ defined as $\pi_1((x, y)) = x$ is a homomorphism, called the ***first projection*** of $\mathcal{M} \times \mathbf{N}$ onto $\mathcal{M}$. The second projection $\pi_2 : \mathcal{M} \times \mathbf{N} \to \mathbf{N}$ is defined similarly as $\pi_2((x, y)) = y$. More generally the i-th projection $\pi_i : \prod_{i\in I} \mathcal{M}_i \to \mathcal{M}_i$ is defined as $\pi_i(x) = x_i$.

Whereas inclusions are injective, projections are surjective.

An ***isomorphism*** of monoids is a bijective homomorphism. When there exists an isomorphism between two monoids we say that they are ***isomorphic***. An isomorphism of a monoid onto itself is called an ***automorphism***.

1.3.4 Congruences and Quotients

A ***semigroup congruence*** is an equivalence relation $\cong$ on a monoid such that if $x \cong x'$ and $y \cong y'$ then $x \cdot y \cong x' \cdot y'$.

For example the identity relation ($x \cong y$ implies $x = y$) and the complete relation ($x \cong y$ always) are semigroup congruences on any semigroup. The relation "$x = 0 = y$ or $x \neq 0 \neq y$" (an equivalence relation with two classes, $\{0\}$ and the rest) is a congruence on $(\mathbb{N}, +)$ but not on $(\mathbb{Z}, +)$ since $-1 \cong 1$ but $1+1 \not\cong -1+1$. The relation "same parity" (both even or both odd) is a congruence on both $(\mathbb{N}, +)$ and $(\mathbb{Z}, +)$.

An equivalence relation on a set X partitions X into blocks called ***equivalence classes.*** When the equivalence relation is a congruence the blocks are called ***congruence classes.*** We write $[x]_{\cong}$ for the congruence class containing a specified element x.

The congruence classes of a semigroup $\mathcal{S} = (X, \cdot)$ forms a semigroup under the operation $\cdot'$ defined by $[x] \cdot' [y] = [x \cdot y]$, called the ***quotient*** of $\mathcal{S}$ *by* $\cong$, written $\mathcal{S}/\cong$. The definition of a congruence ensures that $\cdot'$ is well-defined in that it does not depend on the choice of representative x of the class $[x]$: any other y in the same class would yield the same result.

A quotient of a monoid is a monoid, with identity [1]. A quotient of a group is a group, with identity [1] and an inverse $[x]^{-1} = [x^{-1}]$ for each class $[x]$.

The ***kernel*** of a homomorphism $h : \mathcal{M} \to \mathbf{N}$, written ker$h$, is the binary relation $\cong$ on X_M defined by $x \cong y$ just when $h(x) = h(y)$.

Proposition 3 *A binary relation on X_M is a congruence on $\mathcal{M}$ just when it is the kernel of a homomorphism $f : \mathcal{M} \to \mathbf{N}$ for some $\mathbf{N}$.*

Proof: It can be seen that the kernel of a homomorphism is a congruence. Conversely any congruence determines a quotient, whose kernel is that congruence. ■

We have defined the quotient $\mathcal{S}/\cong$ to be a semigroup of equivalence classes. The function $h : \mathcal{S} \to \mathcal{S}/\cong$ defined by $h(x) = [x]$ can be seen to be a homomorphism, which we also refer to as a quotient when there is no ambiguity. Note that such a quotient is always surjective.

A congruence $x \cong y$ being the set of pairs (x, y) for which $x \cong y$, the intersection $\cong_1 \cap \cong_2$ of two congruences $\cong_1$ and $\cong_2$ on a semigroup $\mathcal{S}$ then has the property that $x(\cong_1 \cap \cong_2)y$ holds just when $x \cong_1 y$ and $x \cong_2 y$.

Proposition 4 *The intersection of two congruences is itself a congruence.*

Proof: If $x(\cong_1 \cap \cong_2)y$, then $x \cong_1 y$ and $x' \cong_1 y'$, whence $x \cdot x' \cong_1 y \cdot y'$ and similarly $x \cdot x' \cong_2 y \cdot y'$. Hence $x \cdot x'(\cong_1 \cap \cong_2)y \cdot y'$. ■

This generalizes immediately to the intersection of any family $(\cong_i)_{i \in I}$ of congruences on a semigroup $\mathcal{S}$.

1.3.5 Ideals

We have already encountered ideals in the context of ordered sets. Essentially the same notion arises for semigroups and monoids, though the following elementary definition does not make this connection apparent.

An ***ideal*** of a semigroup $\mathcal{S} = (X, \cdot)$ is a nonempty subset $Z \subseteq X$ such that for all $z \in Z$ and $x \in X$, $x+z \in Z$ if and only if $x \in Z$, and $z + x \in Z$ if and only if $x \in Z$.

The "if" direction of the definition ensures that every ideal of $\mathcal{S}$ is a subsemigroup of $\mathcal{S}$. The "only if" direction however rules out some subsemigroups. For example the subsemigroup of $(\mathbb{N}, +)$ omitting just 1 is not an ideal because $2 \in Z$ and $2 + 1 \in Z$ but $1 \notin Z$.

The semigroup X^* of strings on X has two extremal ideals, the singleton $\{\epsilon\}$ and X^* itself. In between there is for each n the ideal I_n consisting of all strings of length a multiple of n. There is also for each subset $Y \subseteq X$ the ideal Y^* of all strings on Y.

Theorem 10 *If Z and Z' are ideals of $\mathcal{S}$ so is $Z \cap Z'$.*

Proof: $x \in \bigcap_i Z_i$ iff $\forall i[x \in Z_i]$ iff $\forall i[(x + z \in Z$ and $x \in Z \cap Z'$ iff $[x \in Z$ and $x \in Z']$ iff $[x + z \in Z$ and $x + z \in Z']$ iff $x + z \in Z \cap Z'$. ■

This generalizes immediately to the intersection $\bigcap_i Z_i$ of a family of ideals $(Z_i)_{i\in I}$ of a semigroup $\mathcal{S}$.

The ideal ***generated by*** a subset X of a semigroup $\mathcal{S}$ is the least ideal of $\mathcal{S}$ containing X. The above theorem ensures that the least such ideal exists, namely the intersection of those ideals containing X.

The following observation relates semigroup ideals to congruences and to ideals in other structures besides semigroups.

Proposition 5 *Z is an ideal of $\mathcal{S} = (X, \cdot)$ if and only if it is the inverse image of the identity 1 of a monoid $\mathcal{M}$ under a semigroup homomorphism $h : \mathcal{S} \to \mathcal{M}$.*

Proof: (If) Let $Z = h^{-1}(1)$. Let $z \in Z$ and $x \in X$. Then $h(x \cdot z) = h(x) \cdot h(z) = h(x)$, so $x \in Z$ iff $x \cdot z \in Z$.

(Only if) Let Z be an ideal of $\mathcal{S}$. Define $x \cong y$ iff there exist $z, z' \in Z$ such that $x \cdot z = y \cdot z'$, making $\cong$ a congruence. Now for any $x \in X$ and $z \in Z$, $[x] + [z] = [x + z] = [x]$, whence the quotient $\mathcal{S}/\cong$ is a monoid with identity $[z] = Z$. Hence Z is the inverse image of $[z]$ under this quotient. ∎

We may now make the connection with order ideals. Recall that an order ideal of an ordered set $(X, \leq)$ is the inverse image of the least element of a partial order $(Y, \leq)$ under a monotone function $f : X \to Y$. The least element plays an analogous role to that of the monoid identity. Note that X need have neither a least element nor be antisymmetric; the notion of order ideal is defined for any ordered set. By the same token, the notion of monoid ideal does not depend on the identity of the monoid, and could have been defined more generally for any semigroup.

A ***principal*** ideal of a monoid $\mathcal{M}$ is an ideal generated by a single element of $\mathcal{M}$.

Theorem 11 *(Principal Ideal Theorem for Z.) Every ideal I in $(\mathbb{Z}, +, 0)$ is principal.*

Proof: If I contains only one number it must be 0, which then generates I. Otherwise, let $x < y$ be two consecutive numbers in I as close together as any pair in I. From $x \cong y$ and $y-x \cong y-x$ we infer $y \cong 2y-x$, whence $2y - x$ is also in I, and more generally I must contain $x + n(y - x)$ for all $n \in Z$, an arithmetic progression infinite in both directions. There is now no room in I for additional numbers without violating the hypothesis that x and y are as close as any pair in I. That I contains 0 determines the alignment of this progression, which can then be seen to consist of all multiples of $y - x$. Hence $y - x$ generates I. ∎

1.3.6 Exercises

1. $(\mathbb{Z}_n, \times, 1)$ for $n \geq 2$ is a commutative monoid but not a group.

2. Suppose the order of the electronic black boxes described near the beginning of the section is immaterial; furthermore two assemblies with the same behavior always connect to yield the identity behavior; and there are only three types of boxes. (i) How many distinct behaviors can be constructed? (ii) Reformulate this electrical engineering problem as a problem in pure mathematics.

3. Show that a monoid has exactly one identity.

4. Show that an Abelian group G may be defined to be a group whose inverse operation is a group homomorphism.

5. Show that up to isomorphism there are just two two-element monoids.

6. For which congruences on $(\mathbb{N}, +)$ is $0 \cong 1$? For which is $0 \cong 2$? (Remember $1 \cong 1$.) What is the least congruence containing $2 \cong 5$? What are the corresponding answers for $(\mathbb{Z}, +)$? (Remember $-1 \cong -1$).

7. Show that for any natural number n, the binary relation "x is congruent to y mod n," defined as $\exists ab[x + an = y + bn]$, is a congruence on any submonoid of $(\mathbb{Z}, +, 0)$, in particular on $(\mathbb{N}, +, 0)$.

8. Show that in X^* each of the following three relations between strings x and y is a congruence: of equal length; agree in their first n symbols; contain the same number of occurrences of each symbol.

9. Given homomorphisms $h : \mathcal{M} \to \mathbf{N}$, $h' : \mathcal{M} \to \mathbf{N}'$, define their *product* $h \times h' : M \to \mathbf{N} \times \mathbf{N}'$ as $(h \times h')(x) = (h(x), h'(x))$. Show $\ker(h \times h') = \ker(h) \cap \ker(h')$. Generalize to an arbitrary family $(h_i)_{i \in I}$ of homomorphisms.

10. Show that a subset Z of a monoid $\mathcal{M}$ is a submonoid of $\mathcal{M}$ if it is an ideal of $\mathcal{M}$, but not necessarily conversely.

11. Show that a subset Z of a group G is a subgroup of G if and only if it is an ideal of G.

1.4 Closure Systems and Galois Connections

1.4.1 Closure Systems

A closure system can be defined in terms of either a closure operator or a closure property. Either one determines the other.

A function $f : X \to X$ on a poset $(X, \leq)$ is a ***closure operator on*** X when it is monotone, idempotent $(f(f(x)) = f(x))$, and increasing $(x \leq f(x))$.

Examples

1. The poset 2^{X^2} of all binary relations on X has a number of associated closure operators: reflexive closure, transitive closure, reflexive transitive closure, symmetric closure, and equivalence closure (reflexive, symmetric, transitive).

2. Existential quantification is a closure operator, in that if $\varphi(x) \to \psi(x)$ then $\exists x\varphi(x) \to \exists x\psi(x)$; $\exists x\exists x\varphi(x) = \exists x\varphi(x)$; and $\varphi(x) \to \exists x\varphi(x)$.

3. Implication $p \to q$ and disjunction $p \vee q$, as a function of q with p held fixed, are both closure operators since they are clearly monotone and idempotent, and furthermore increasing as $q \leq (p \to q)$ and $q \leq (p \vee q)$.

A ***closure property*** on a partial order X is a subset $C \subseteq X$ such that for any subset $D \subseteq C$, the inf of D in X belongs to C, i.e. C is closed under infs in X. Since the inf in X of the empty subset must be the greatest element of X, it follows that any partial order X having a closure property contains a greatest element 1 and moreover 1 must belong to every closure property on X.

Moreover, all X-infs of subsets of C are also C-infs of those subsets since they are in C, whence C is a complete semilattice, being closed under C-infs. However C need not even have X-sups (of its subsets), let alone be closed under them.

Examples The properties reflexive, transitive, and symmetric, as applied to binary relations on a set X, are closure properties on 2^{X^2} ordered as usual, namely by inclusion. This is because given any set of relations satisfying one of these properties, the intersection of that set also satisfies that property. The intersection of the empty set is taken to be X^2 itself, the maximal binary relation, called K_X, which satisfies all of these properties.

Theorem 12 *In a complete lattice, closure operators and closure properties are in 1-1 correspondence.*

Proof: The closure property determined by a closure operator is the property of being a fixpoint of the operator. Conversely the closure operator determined by a closure property is the function mapping x to the infimum of the set of upper bounds on x in the property. ∎

This theorem motivates the following definition. A ***closure system*** is a complete lattice with a closure operator, or equivalently with a closure property.

A ***projection*** is a monotone idempotent function on a poset. Thus a closure operator is an increasing projection. An ***interior operator*** is a decreasing projection, one satisfying $x \geq f(x)$.

Examples 1. The irreflexive interior of a binary relation $R \subseteq X \times X$, obtained by setting the diagonal of R to 0 (i.e. removing all pairs (x, x)) is an interior operation.

2. Universal quantification satisfies $(\forall x.\varphi(x)) \rightarrow \varphi(x)$ and hence is an interior operation.

3. Conjunction with a constant p, as in $p \wedge q$, is an interior operation.

1.4.2 Galois Connections

Given two posets $(X, \leq)$, $(Y, \leq)$, a ***Galois connection*** between them consists of two functions $f : X \rightarrow Y$, $g : Y \rightarrow X$, called ***polarities***, such that for all $x \in X$ and $y \in Y$, $y \leq f(x)$ if and only if $x \leq g(y)$, which we will call the ***Galois condition***.

Examples

1. The canonical example of a Galois connection is that between sets[6] of structures and sets of propositions given a binary relation $s \models p$ of ***satisfaction*** between structures s and propositions p. The ***theory*** $\Theta(C)$ of a set C of structures is the set of those propositions satisfied by every structure in C. Conversely the ***models*** $\mathcal{M}(\Gamma)$ of a set Γ of propositions consists of those structures satisfying every proposition in Γ.

Here the two posets (more correctly partially ordered sets) consist of the set 2^Σ of all sets of structures and the set 2^Π of all sets of propositions, both ordered by inclusion. The two polarities are $\mathcal{M} : 2^\Pi \rightarrow 2^\Sigma$ and $\Theta : 2^\Pi \rightarrow 2^\Sigma$.

To see that these are polarities, suppose that we have sets C of structures and Γ of propositions. Then both $\Gamma \subseteq \Theta(C)$ and $C \subseteq \mathcal{M}(\Gamma)$ are equivalent to the statement, for every structure $s \in C$ and proposition $p \in \Gamma$, $s \models p$. Hence they are equivalent to each other and thus meet the Galois condition.

2. Another example obtained in essentially the same way starts with a poset $(P, \leq)$ and takes both X and Y to be the power set of P, the set of all subsets of P. Given a subset $Q \subseteq P$, define $\mathcal{L}(Q) = \{p \in P | \forall q \in Q.p \leq q\}$ and $\mathcal{U}(Q) = \{p \in P | \forall q \in Q.q \leq p\}$. That is, $\mathcal{L}(Q)$ consists of all lower bounds of the set Q while $\mathcal{U}(Q)$ consists of all upper bounds. Then for all subsets Q, Q' of P, we have $Q \subseteq \mathcal{L}(Q')$ iff $Q' \subseteq \mathcal{U}(Q)$ because both express "$\forall q \in Q \forall q' \in Q'.q \leq q'$."

3. The polarities of the previous two examples were obtained from a binary relation in essentially the same way. Here we give a nontrivially different example. Let $\lfloor x \rfloor$ denote the greatest integer n such that $x \geq \texttt{float}(n)$. (We shall be pedantic here and first convert integers to reals via $\texttt{float} : \mathbf{Z} \rightarrow \mathbb{R}$ when comparing them to reals.) This defines an antimonotone function from $(\mathbb{R}, \geq)$ to $(\mathbf{Z}, \leq)$ (it would be monotone had we chosen to order the reals with $\leq$ instead of $\geq$). With this ordering, $\texttt{float} : \mathbf{Z} \rightarrow \mathbb{R}$ is also antimonotone. Then we have $x \geq \texttt{float}(n)$ iff $n \leq \lfloor x \rfloor$. That is, $\lfloor - \rfloor$ and $\texttt{float}(-)$ are the polarities of a Galois connection between $(\mathbb{R}, \geq)$ and $(\mathbf{Z}, \leq)$.

4. With $(\mathbb{R}, \leq)$ and $(\mathbf{Z}, \geq)$ we obtain essentially the same Galois connection but with $\lceil - \rceil$ (ceiling) in place of $\lfloor - \rfloor$ (floor), where $\lceil x \rceil$ denotes the least integer greater or equal to x.

An equivalent formulation of the Galois condition consists of the following three conditions.

(i) Both f and g are antimonotone.

(ii) For all $x \in X$, $x \leq g(f(x))$.

(iii) For all $y \in Y$, $y \leq f(g(y))$.

[6] These will usually be proper classes. For simplicity of exposition we blur the usual distinction between class and set in this section.

Theorem 13 *The Galois condition holds if and only if conditions (i)-(iii) above hold.*

Proof: (Only if) We prove (i)-(iii) from the Galois condition.

(i) Let $x \leq x'$ in X. Substituting $f(x)$ for y in the Galois condition makes the left hand side true, whence the right hand side, $x \leq g(f(x))$, holds for all $x \in X$. Hence $x' \leq g(f(x'))$, and therefore $x \leq g(f(x'))$. So by the Galois condition $f(x') \leq f(x)$, that is, f is *antimonotone.* By symmetry of the two polarities, g is also antimonotone.

(ii) Substituting $f(x)$ for y in the Galois condition makes it $f(x) \leq f(x)$ if and only if $x \leq g(f(x)))$, whence (i).

(iii) is argued symmetrically.

(If) Given (i)-(iii) we wish to show the Galois condition. So suppose $y \leq f(x)$ for some $x \in X$ and $y \in Y$. By (i) $g(f(x)) \leq g(y)$, which with (ii) yields $x \leq g(y)$ giving one half of the Galois condition. The other half is argued symmetrically. ■

For the moment the main use we make of this theorem is in the proof of the next very important theorem. We remark in passing however that it foreshadows its generalization to two equivalent definitions of adjunction, in the chapter on category theory.

Theorem 14 *The composites $gf : X \to X$ and $fg : Y \to Y$ of the polarities of a Galois connection are closure operations.*

Proof: The previous theorem established the antimonotonicity of the polarities, whence their composites are monotone. It also established that $x \leq g(f(x))$ and $y \leq f(g(y))$, that is, both composites are increasing. It remains to show that the composites are idempotent.

In fact we shall show the stronger result that $fgf = f$, and by symmetry $gfg = g$.

First, since $x \leq gf(x)$, by antimonotonicity of f we have $fgf(x) \leq f(x)$. Second, by symmetry we have $y \leq fg(y)$ for all $y \in Y$, which is made $f(x) \leq fgf(x)$ by substituting $f(x)$ for y. Combining first and second then gives $fgf(x) = f(x)$ as promised. ■

The importance of this theorem is that *any* notion of satisfaction between models and propositions induces a corresponding Galois connection, which in turn induces two closure operators. One operator, theory-of-models, $\Theta\mathcal{M}$, maps any set Γ of propositions to its *deductive closure* $\Theta(\mathcal{M}(\Gamma))$. This is the theory consisting of all propositions true of every model of Γ. The members of $\Theta(\mathcal{M}(\Gamma))$ are called the *consequences* of Γ.

The other operator, $\mathcal{M}\Theta$, maps any set C of structures to what we shall call its *homologue closure*[7] $\mathcal{M}(\Theta(C))$. The members of $\mathcal{M}(\Theta(C))$ are the homologues of C, namely those structures satisfying all the propositions holding of every structure in C.

For example consider structures of the form $(X, \wedge, \vee)$ where $\wedge$ and $\vee$ are arbitrary binary operations on a set X, and sentences of the form $t = u$ where t, u are terms built from variables using the operation symbols $\wedge$ and $\vee$. Say that structure s satisfies $t = u$ when for every assignment of elements of s to variables of t and u, the two terms evaluate to the same element. Then the deductive closure of the axiomatization of lattices, namely the equations for associativity, commutativity, and idempotence of each of the two operations, along with the two absorption laws, is the equational theory of lattices. When we add the distributivity law to Γ the deductive closure becomes the equational theory of distributive lattices.

On the other side, the homologue closure of the (unique) two-element lattice is the set of distributive lattices.

[7] This term was proposed by Bill Rounds in response to the author's request, posted to a logic mailing list, for a suitable name for the concept.

When it is known, from context or otherwise, what the relevant operations are (in this case two binary operations), we may abbreviate the above conditions on the structures, propositions, and the satisfaction relation by talking simply of "equational consequences" and "equational homologues."

Theorem 15 *Polarities map sups to infs.*

Proof: Let X' be a subset of X having a sup $\bigvee X'$. We shall show that $f(\bigvee X') = \bigwedge f(X')$ where $f : X \to Y$ is a polarity of a Galois connection. Now for every $x \in X'$, $x \leq \bigvee X'$ whence by antimonotonicity $f(\bigvee X') \leq f(x)$. Hence $f(\bigvee X') \leq \bigwedge f(X')$, i.e. $f(\bigvee X')$ is a lower bound of $f(X')$.

Now consider any lower bound y of $f(X')$. By the Galois condition, $x \leq g(y)$ for every $x \in X'$, whence $\bigvee X' \leq g(y)$. So again using the Galois condition, $y \leq f(\bigvee X')$. This shows that $f(\bigvee X')$ is the greatest lower bound of $f(X')$, i.e. $f(\bigvee X') = \bigwedge f(X')$. ∎

Corollary 16 *The set of models of the union of two sets of axioms is the intersection of the sets of models of each set of axioms taken separately. Furthermore this generalizes to any number of sets of axioms. On the other side, the theory of the union of two sets of structures is the intersection of the theories of the sets taken separately.*

This corollary is a case where abstraction does not yield any additional insights or simplifications than a more direct proof. After all, it is obvious that the models of the union of two theories Γ_1 and Γ_2 are exactly those structures that are models of Γ_1, and models of Γ_2, and similarly for propositions satisfied by the structures in two sets C_1 and C_2.

In fact it might seem that we were lucky to be able to prove the abstract result at all, since we had no reason to suppose that everything true of our concrete example holds in every Galois connection. We now provide such a reason in the form of the theorem following these definitions.

A Galois connection is called ***concrete*** when it the one determined by two sets A, B and a binary relation $R \subseteq A \times B$ between them. That is, the posets are the respective power sets of A and B, while the polarities $f : 2^A \to 2^B$, $g : 2^B \to 2^A$ are defined by $f(A') = \{b | \forall a \in A'. aRb\}$ and $g(B') = \{a | \forall b \in B'. aRb\}$, for $A' \subseteq A$ and $B' \subseteq B$.

Our canonical example of a Galois connection was concrete: A was a set (or class) Σ of structures, B a set or class Π of propositions, and R the satisfaction relation $\models \subseteq \Sigma \times \Pi$.

Given a Galois connection $G = (X, Y, f, g)$, a ***representation*** of G is another Galois connection $G' = (X', Y', f', g')$ together with functions $F : X \to X'$, $G : Y \to Y'$ such that for all $x \in X$ and $y \in Y$, $f'(F(x)) = F(f(x)$ and $g'(F(y)) = F(g(y))$. We say that $F(x)$ *represents* x and likewise $F(y)$ *represents* y. When F is injective we call the representation an ***embedding*** and say that G embeds in G'.

We call a Galois connection ***representable*** when it embeds in a concrete Galois connection.

Theorem 17 *Every Galois connection (X, Y, f, g) is representable.*

Proof: Take $A = X$ and $B = Y$. Define R such that aRb holds just when either there exists $x \in X$ such that $a \leq x$ and $b \leq f(x)$, or there exists $y \in Y$ such that $a \leq g(y)$ and $b \leq y$. Exercise 5 completes this proof. ∎

1.4.3 Exercises

1. Show that $\mathcal{U}(Q) \cap \mathcal{L}(\mathcal{U}(Q))$ is either empty or a singleton. When empty, show that $\bigvee Q$ is undefined, and when a singleton, that it is $\{\bigvee Q\}$.

2. Show that $f : L \to L$ on a semilattice $(L, \vee)$ is a closure operator if and only if it satisfies

$$x \vee f(f(x)) \leq f(x) \leq f(x \vee y).$$

3. Show that every projection on a lattice L acts as a closure operator on part of L and an interior operator on another part, and as the identity on the intersection of those two parts.

4. Show that the property of being an order filter of a poset is a closure property. Describe the associated closure operator. On which subsets does this operator give the same result as $\mathcal{U}$?

5. Complete the proof of Theorem 17.

6. What is the equational homologue of (i) the first of the two five-element nondistributive lattices depicted earlier? (ii) the second? (iii) both? (In (i) and (ii) the given class of lattices is a singleton, in (iii) a doubleton.)

1.5 Fixpoints

1.5.1 Complete Lattices

We defined a lattice to be a semilattice that is also a lower semilattice, which is to say that it is a partial order all of whose finite nonempty subsets have both least upper and greatest lower bounds. When "finite nonempty" is omitted we obtain the following notions.

An upper ***complete semilattice*** is an upper semilattice each of whose subsets has a sup, and dually for lower complete semilattices. A ***complete lattice*** is a lattice whose two semilattices are both complete, i.e. every subset has both a sup and an inf.

Examples.

1. The power set lattice $(2^X, \cup, \cap)$ is a complete lattice for any choice of X, even empty or uncountable.

2. The closed unit interval $[0, 1]$ of reals is a complete lattice. Every subset of such a bounded set of reals has a sup and an inf. The whole real line however is not, as unbounded subsets lack a sup or inf.

3. The set S of all subalgebras of an algebra is partially ordered by the subalgebra relation (why?), and is closed under arbitrary intersection. Hence every subset T of S has an inf. It is less obvious that it also has a sup, which we now show via the following general theorem.

Theorem 18 *Every complete semilattice L is a complete lattice.*

Proof: We carry out the argument for upper complete semilattices, the other case being simply the dual. For any subset X of L, let $p = \bigvee \mathcal{L}X$, the sup of the set of lower bounds of X. Since every $x \in X$ is an upper bound of $\mathcal{L}X$, p is a simultaneously a lower bound of X and an upper bound of Y. This makes p the greatest lower bound of X, showing that $\bigwedge X$ exists. ∎

1.5.2 Relative Completeness in Sublattices.

We shall shortly have occasion to ask whether a sublattice M of a complete lattice L is itself a complete lattice. This is a more delicate question than it looks, as can be illustrated with the usual Dedekind-cut definition of real numbers as lower sets of rationals. (With this definition we pick up two extra "reals": $-\infty$ as the empty set and ∞ as the set $\mathbb{Q}$ of all rationals; these could be dropped, but keeping them allows us to say that the reals form a complete lattice.) This definition is unambiguous for irrational reals, but to complete the definition for the case of rational reals we must decide (uniformly for all rationals, for simplicity) whether the lower set is closed (i.e. contains the rational it represents). If we decide yes then let us name the set of reals so defined C for closed, and if no then O for open; C and O are distinct in the sense that they are distinct subsets of $2^{\mathbb{Q}}$, though intuitively they should serve equally well as the definition of the set R of reals.

Now the point of having reals as opposed to rationals is to have a complete lattice under the standard arithmetic ordering, one in which every set (every bounded set if we drop $\pm\infty$) of reals has a sup and a inf. So let us check that O and C both have sups and infs for all their subsets.

Our first test case is the subset of positive reals, whose inf had better be 0. The inf in $2^{\mathbb{Q}}$ of the positive C-reals is $\{q \leq 0\}$, which is indeed the C-real 0. However the inf in $2^{\mathbb{Q}}$ of the positive O-reals is also $\{q \leq 0\}$, whereas the O-real for 0 is $\{q < 0\}$. Apparently O is defective!

Our second test case is the subset of negative reals, whose sup had better be 0. In both cases we get the sup in $2^{\mathbb{Q}}$ to be $\{q < 0\}$, which is the O-real for 0. So C is broken too!

This situation is straightened out by taking sups and infs not in $2^{\mathbb{Q}}$ but in C or O itself. The inf in O of the positive O-reals is indeed $\{q < 0\}$, the O-real for 0, and dually for C.

Thus it is important to be clear about which lattice one is taking sups and infs in. We shall sometimes write $\sup_L X$ or $\inf_L X$ to indicate that the sup or inf of X is being taken in L, and sometimes refer to these as L-sups or L-infs. We shall say that a sublattice of a lattice L is ***closed under*** L-sups when it contains all its L-sups, and ***complete in*** L when it is closed under both L-sups and L-infs.

1.5.3 Directed Sets and Continuity

The requirement for complete lattices that *every* subset of a partial order have a sup is very strong. In some situations we may have to settle for sups only of certain kinds of subsets, giving rise to a kind of structure that is not only not a complete lattice but not even a lattice.

A subset Y of a partial order P is called ***directed*** when it is nonempty and any two elements of it have an upper bound in Y. The dual of directed is ***filtered***.

A ***complete partial order***, or ***CPO***, is a partial order such that $\bigvee Y$ is defined for all directed subsets Y.

Often this notion is defined with "nonempty chain" in place of "directed." This gives rise to a slightly different notion from the above, in which case the term ***chain-complete*** partial order, or chain-CPO, is sometimes used to distinguish it from a CPO. Since every nonempty chain is a directed set, every CPO is a chain-CPO.

A ***pointed*** CPO is a CPO with a least element. The presence of a least element is essential for the fixpoint theorems below.

Examples

1. The CPO of finite and infinite strings $(\Sigma^\infty, \sqsupseteq)$ ordered by the prefix relation $u \sqsubseteq w$ defined as $w = uv$ for some $v \in \Sigma^*$. Any directed set in this CPO must be a chain. The sup of any finite nonempty chain is its longest element, while the sup of any infinite chain is the infinite string whose i-th letter is the letter that appears in the i-th position of all strings in the chain of length at least i. That chain need not (but may) include its sup. This is a pointed CPO, having the empty string as its least element.

2. The CPO of all partial functions on the natural numbers, ordered by inclusion. A directed set here is a set of functions all having the same value wherever they are defined. The sup of a directed set is then simply its union (treating functions as their graphs, i.e. sets of pairs). This is a pointed CPO, having the every undefined function as its least element.

Thus far the only kind of function on posets and lattices that we have considered is the monotone function, which preserves order. It can be seen (Exercise 1) that a function between partial orders is monotone just when it preserves sups of finite directed sets. This characterization of monotone functions in terms of which sups they preserve motivates the following definitions.

A function between partial orders is called ***chain-continuous***. ***continuous***, ***finitely additive***, or ***completely additive***, when it preserves sups of nonempty chains, directed sets, nonempty finite sets, or all sets, respectively.

Certain implications between types of functions are made evident by these definitions. Thus every completely additive function is monotone, continuous, finitely additive, and chain-continuous. Every finitely additive function is monotone. And since nonempty chains are directed every continuous function is chain-continuous.

1.5.4 Fixpoint Theorems

We come now to Tarski's fixpoint theorem[8] for monotone functions in complete lattices, adopted by Scott for the foundation of an elegant theory of computation.

A ***fixpoint*** of a function $f\colon L \to L$ on lattices is an element y of L satisfying $f(y) = y$. A ***pre-fixpoint*** of f satisfies $y \leq f(y)$; the dual notion is ***post-fixpoint***.

Theorem 19 *(Tarski, 1942, publ. 1955) Let $f : L \to L$ be a monotone function on a complete lattice L. Let $P \subseteq L$ be the set $\{x \mid x \leq f(x)\}$ of all pre-fixpoints of f, and let $F \subseteq P$ be the set $\{x \mid f(x) = x\}$ of all fixpoints of f. Then F is a suborder of P that by itself is a complete lattice (its sups and infs are not those of L).*

Proof: Let $Y \subseteq F$ be any set of fixpoints. Take $Z = \mathcal{L}_P Y$ and $u = \bigvee_L Z$. Then for any $z \in Z$, $z \leq f(z) \leq f(u)$, whence $f(u)$ is an upper bound of Z, so $u \leq f(u)$, u being the least such upper bound. Hence $u \in P$, so $u = \bigvee_P Z = \bigvee_P \mathcal{L}_P Y = \bigwedge_P Y$. Hence $u \in \mathcal{L}_P Y = Z$. Also $f(u) \leq f(f(u))$ (by $u \leq f(u)$ and monotonicity of f), so $f(u) \in P$. Thus $f(u) \in \mathcal{L}_P f(Y) = \mathcal{L}_P(Y) = Z$. Hence $f(u) \leq u$. Therefore $f(u) = u$ and so $u \in F$. Hence for every subset Y of F, $\bigwedge_P Y$ is in F, making F a complete sublattice of P. ∎

Since a complete lattice has a least element, it follows that every monotone function on a complete lattice has a *least* fixpoint.

The following is the more commonly used fixpoint theorem in computational settings. Its advantage is that it caters for structures with diverging branches that do not reunite higher up, expressing a notion of *conflicting alternatives.*

Theorem 20 *Every continuous function on a pointed CPO has a least fixpoint.*

Proof: Let $f^*(y)$ denote the set $\{y, f(y), f(f(y)), \ldots\}$. If $y \leq f(y)$ then by monotonicity of f, $f(y) \leq f(f(y))$ and so on. Hence $f^*(y)$ is a chain, and therefore a directed set, so $f(\bigvee f^*(y)) = \bigvee f(f^*(y)) = \bigvee f^*(y)$, that is, $\bigvee f^*(y)$ is a fixpoint of f. We may easily meet the prerequisite $y \leq f(y)$ by taking $y = \perp$, the least element of the CPO, giving us $\bigvee f^*(\perp)$ as a fixpoint of f.

To see that $\bigvee f^*(\perp)$ is the least such fixpoint it suffices to show that the upper bounds on $f^*(\perp)$ include all fixpoints of f, since $\bigvee f^*(\perp)$ is the least upper bound. But if z is a fixpoint of f then $\perp \leq z$, so $f(\perp) \leq f(z) = z$, and so on, whence z is an upper bound on $f^*(\perp)$. ∎

A more general theorem unifying both of the above is possible: every monotone function on a CPO has a least fixed point. However the extant proofs are quite a bit longer than the two special cases proved above.

[8] This is often called the Tarski-Knaster theorem. A 1927 paper listing only Knaster as the author but acknowledging Tarski's collaboration showed the theorem for the case of the complete lattice of all subsets of a set. Much later Tarski working alone proved the general version given here.

1.5.5 Exercises.

1. Show that a function between partial orders is monotone just when it preserves sups of finite directed sets.

2. Show that among the subalgebras of an algebra, the least upper bound bound of any set of such subalgebras is the subalgebra generated by the union of that set.

3. Show that if M is a sublattice of L, $\sup_M X \geq \sup_L X$ and $\inf_M X \leq \inf_L X$. Infer that if M is closed under L-sups or L-infs then (i) it is closed under M-sups or M-infs respectively and (ii) it is a complete lattice in itself.

4. Show that C and O are closed under $2^{\mathbb{Q}}$-infs and $2^{\mathbb{Q}}$-sups respectively, and $C \cup O$ under both, where C and O are the sets of order ideals of the chain $\mathbb{Q}$ of rationals described in section 1.5.2. Infer that C, O, and $C \cup O$ as lattices are each complete in themselves. Which of them are complete in $2^{\mathbb{Q}}$? What is the corresponding situation for $C \cap O$?

5. For monotone $f : L \to L$ on a complete lattice L, show that for any pre-fixpoint z of f the set of fixpoints of f above z form a complete lattice. Infer that for any monotone increasing ($f(x) \geq x$) function there exists a least monotone increasing idempotent $f^* : L \to L$ with $f(x) \leq f^*(x)$ for all $x \in L$.

1.6 Quantales

1.6.1 Definitions and examples

A ***quantale*** $(X, \bigvee, \cdot)$ is a structure such that $(X, \bigvee)$ is a complete semilattice, $(X, \cdot)$ is a semigroup, and the equations $x \cdot \bigvee Y = \bigvee(x \cdot Y)$, and $(\bigvee Y) \cdot x = \bigvee(Y \cdot x)$ hold for all $x \in X$ and $Y \subseteq X$. (Notation: $x \cdot Y = \{x \cdot y | y \in Y\}$.)

A quantale ***with unit*** is a quantale $(X, \bigvee, \cdot, 1)$ containing a constant 1 such that $(X, \cdot, 1)$ is a monoid. A ***commutative*** quantale is one satisfying $x \cdot y = y \cdot x$.

For doubleton subsets $Y = \{x, y\}$ we write $\bigvee\{x, y\}$ as $x \vee y$. The sup operation partially orders a quantale as for semilattices: $x \leq y$ holds just when $x \vee y = y$.

Examples

1. Writing Σ^* for the set of all finite strings over an alphabet Σ, the power set 2^{Σ^*} of all formal languages over that alphabet forms a quantale $(2^{\Sigma^*}, \bigcup, \cdot)$ under arbitrary union $\bigcup Y$ of a set Y of languages and concatenation UV of two languages U and V defined as $\{uv | u \in U, v \in W\}$. Expanding this structure with a unit, $(2^{\Sigma^*}, \bigcup, \cdot, \epsilon)$, yields a quantale with unit.

Modifying this example by replacing Σ^* by Σ^+, the set of all finite nonempty strings, yields a quantale with no element available to serve as a unit. Many other modifications are possible, such as taking all strings of length at least 7, all strings of length a multiple (nonzero if you want to eliminate the unit) of some integer, etc.

2. Let A be a set. Then the power set 2^{A^2} of all binary relations R on A forms a quantale $(2^{A^2}, \bigcup, ;)$ under arbitrary union $\bigcup Y$ of a set Y of relations and composition $R; S$ defined as $\{(a, c) | \exists b \in A.(a, b) \in R \wedge (b, c) \in S\}$. This example too can be expanded to a quantale with unit $(2^{A^2}, \bigcup, ;, 1_A)$ where 1_A is the identity relation $\{(a, a) | a \in A\}$.

3. The structure $(\mathbb{N} \cup \{\infty\}, \bigwedge, +)$ is a quantale where $\mathbb{N}$ is the natural numbers, $+$ is numeric addition, $x + \infty = \infty = \infty + x$, and $\bigwedge Y$ is the least element of Y (since every subset of $\mathbb{N}$ has a least element). The same structure expanded to include 0 as a constant, $(\mathbb{N} \cup \{\infty\}, \bigwedge, +, 0)$, is a quantale with unit.

This example has many variants based on the natural numbers $\mathbb{N}$, integers $\mathbb{Z}$, rationals $\mathbb{Q}$, and reals $\mathbb{R}$. In all examples containing at least one positive number, ∞ must be included. Furthermore it must contain a least element (the bottom of the complete semilattice), whence if the quantale is unbounded below (e.g. if it contains infinitely many negative integers) then it must include $-\infty$.

As the first example might suggest, elements of quantales behave very much like binary relations on a fixed set A. The table of kinds of binary relations in Section 1 has its counterpart here, at least for some of the notions.

Let $\mathbf{Q} = (X, \bigvee, \cdot)$ be a quantale. The empty relation corresponds to the least element 0 of $\mathbf{Q}$, the identity relation to the unit 1 when it exists, and the clique to the top element $\top = \bigvee X$.

An element x of $\mathbf{Q}$ is called ***reflexive*** when it satisfies $1 \leq x$, and ***irreflexive*** when it satisfies $x \wedge 1 = 0$. It is called ***transitive*** when it satisfies $xx \leq x$.

1.6.2 Operations on Quantales

All the operations we have seen for semilattices and semigroups generalize straightforwardly to quantales.

Subquantales Given a quantale $\mathbf{Q} = (X, \bigvee, \cdot)$, a quantale $(X', \bigvee', \cdot')$ is a ***subquantale*** of $\mathbf{Q}$ when $X' \subseteq X$ such that $\cdot'$ is the restriction of $\cdot$ to Y', and for all subsets $Y \subseteq X'$, $\bigvee' Y = \bigvee Y$.

The modifications listed above to Example 1 are all subquantales of that example.

Direct Product Given two quantales $(X_1, \bigvee_1, \cdot_1)$, $(X_2, \bigvee_2, \cdot_2)$, their direct product is the quantale $(X_1 \times X_2, \bigvee, \cdot)$. Here $\bigvee Y = (\bigvee_1 Y_1, \bigvee_2 Y_2)$ where $Y \subseteq X_1 \times X_2$ and $Y_1 = \{y \in X_1 | \exists z \in X_2.(y, z) \in Y\}$, and similarly for Y_2, while $(x_1, x_2) \cdot (y_1, y_2) = (x_1 \cdot y_1, x_2 \cdot y_2)$.

Direct product generalizes as usual to any family $\langle (X_i, \bigvee_i, \cdot_i) \rangle_{i \in I}$ of quantales where I is an arbitrary set used to index the family.

Homomorphisms Given two quantales $(X_1, \bigvee_1, \cdot_1)$, $(X_2, \bigvee_2, \cdot_2)$, a ***homomorphism*** is a function $h : X_1 \to X_2$ which is simultaneously a semilattice homomorphism and a semigroup homomorphism. That is, $h(\bigvee_1 Y) = \bigvee_2 h(Y)$ for all $Y \subseteq X$, and $h(x \cdot_1 y) = h(x) \cdot_2 h(y)$ for all $x, y \in X_1$.

A homomorphism $h : (X_1, \bigvee_1, \cdot_1, 1_1) \to (X_2, \bigvee_2, \cdot_2, 1_2)$ between two quantales with unit is a homomorphism of quantales that furthermore preserves the unit, i.e. $h(1_1) = 1_2$.

1.6.3 Residuation

Associated to any quantale $(X, \bigvee, \cdot)$ are two binary operations of ***residuation***. The ***right***[9] ***residual*** of x by y is defined as $y \backslash x = \bigvee \{z | y \cdot z \leq x\}$. The ***left residual*** of x by y is defined dually as $x/y = \bigvee \{z | z \cdot y \leq x\}$. The two residuals coincide ($y \backslash x = x/y$) in a commutative quantale.

Equivalently the two residuals can be defined as follows.

$$\begin{array}{rll} x \cdot y \leq z & \text{iff} & y \leq x \backslash z \\ & \text{iff} & x \leq z/y \end{array}$$

Exercise 4 shows that this definition is equivalent to the first definition. Exercise 5 shows that $x \cdot -$ and $x \backslash -$, as functions of their right hand argument, are the polarities of a Galois connection. From (one direction of)

[9] The name is confusing since the "divisor" y is on the left. The point is that the residue after division is on the right.

the second formulation of a Galois connection we infer that $x;(x\backslash y) \leq y$ and $x\backslash(x;y) \geq y$ (because of the exercise's need to take the order dual of the quantale on one side).

Residuation has a logical interpretation that is made more visible by writing $x\backslash y$ as $x \to y$, i.e. *implication*, $x \leq y$ as $x \vdash y$, *entailment*, and $x;y$ as x,y, *conjunction* (rather than $x \wedge y$, which already has a separate meaning in a quantale, and which besides is commutative, which quantalic conjunction need not be).

In this notation the first inequality of the second formulation of a Galois connection becomes $x, x \to y \vdash y$, recognizable as the logical inference rule *modus ponens*. The second inequality becomes $y \vdash x \to (x,y)$, which is the more boring statement that y is weaker than the implication of x,y by x.

Polarities map sups to infs. However we have taken the order dual of one copy of the quantale, so with respect to the original quantale the polarity $x \cdot -$ preserves sups (which we already had from the definition of quantale), while the polarity $x \to -$ preserves infs, in particular $x \to (y \wedge z) = (x \to y) \wedge (x \to z)$.

1.6.4 Star

Associated with any quantale $(X, \bigvee, \cdot)$ is a unary operation x^+ of ***transitive closure***, defined as the least transitive element greater or equal to x. That is, it satisfies (i) $x \leq x^+$, (ii) $x^+ \cdot x^+ \leq x^+$, and (iii) $x^+ \leq y$ for any y satisfying (i) and (ii) when substituted for x^+.

The reflexive transitive closure or ***star*** x^* of x is defined as $1 \vee x^+$.

Theorem 21 *$x\backslash x$ is reflexive and transitive.*

Define a ***preserver*** of x to be an element y satisfying $x \cdot y \leq x$. Then $x \cdot (x\backslash x) \leq x$ makes $x\backslash x$ a preserver of x, while $y \leq x\backslash x$ for any preserver of x makes $x\backslash x$ the *weakest* preserver of x.

Thus if $x \cdot y \leq x$ we may infer from $(x\backslash x)^* \leq x\backslash x$ and the monotonicity of star that $x \cdot y^* \leq x$.

1.6.5 Exercises

1. Give examples of (a) a quantale of binary relations without unit; (b) a quantale $(X, \bigvee, \cdot)$ such that $(X, \bigwedge, \cdot)$ is not a quantale; (c) a quantale whose semigroup is a group (and show that there are no further examples).

2. Furnish the monoid $(\mathbb{C}, +, 0)$ of complex numbers with a suitable $\bigvee$ making it a quantale.

3. Show that any quantale whose semigroup is a group has exactly one element.

4. Show that the second (iff) definition of the residuals is equivalent to the first definition.

5. Show that for a fixed element x of a quantale $\mathbf{Q}$, the operations $x \cdot y$ and $x\backslash y$ (each as a function of y) form the polarities of a Galois connection between the quantale and its order dual.

6. Derive logic's cut rule $x \to y, y \to z \vdash x \to z$ as a property of a quantale.