Chapter 5

Partial Orders, Lattices, Well Founded Orderings, Equivalence Relations, Distributive Lattices, Boolean Algebras, Heyting Algebras

5.1 Partial Orders

There are two main kinds of relations that play a very important role in mathematics and computer science:

1. Partial orders
2. Equivalence relations.

In this section and the next few ones, we define partial orders and investigate some of their properties.

As we will see, the ability to use induction is intimately related to a very special property of partial orders known as well-foundedness.

Intuitively, the notion of order among elements of a set, X, captures the fact some elements are bigger than others, perhaps more important, or perhaps that they carry more information.

For example, we are all familiar with the natural ordering, $\leq$, of the integers

$$\cdots, -3 \leq -2 \leq -1 \leq 0 \leq 1 \leq 2 \leq 3 \leq \cdots,$$

the ordering of the rationals (where $\frac{p_1}{q_1} \leq \frac{p_2}{q_2}$ iff $\frac{p_2 q_1 - p_1 q_2}{q_1 q_2} \geq 0$, *i.e.*, $p_2 q_1 - p_1 q_2 \geq 0$ if $q_1 q_2 > 0$ else $p_2 q_1 - p_1 q_2 \leq 0$ if $q_1 q_2 < 0$), and the ordering of the real numbers.

In all of the above orderings, note that for any two number a and b, either $a \leq b$ or $b \leq a$.

We say that such orderings are *total* orderings.

A natural example of an ordering which is not total is provided by the subset ordering.

Given a set, X, we can order the subsets of X by the subset relation: $A \subseteq B$, where A, B are any subsets of X.

For example, if $X = \{a, b, c\}$, we have $\{a\} \subseteq \{a, b\}$. However, note that neither $\{a\}$ is a subset of $\{b, c\}$ nor $\{b, c\}$ is a subset of $\{a\}$.

We say that $\{a\}$ and $\{b, c\}$ are *incomparable*.

Now, not all relations are partial orders, so which properties characterize partial orders?

Definition 5.1.1 A binary relation, $\leq$, on a set, X, is a *partial order* (or *partial ordering*) iff it is *reflexive*, *transitive* and *antisymmetric*, that is:

(1) (*Reflexivity*): $a \leq a$, for all $a \in X$;

(2) (*Transitivity*): If $a \leq b$ and $b \leq c$, then $a \leq c$, for all $a, b, c \in X$.

(3) (*Antisymmetry*): If $a \leq b$ and $b \leq a$, then $a = b$, for all $a, b \in X$.

A partial order is a *total order (ordering)* (or *linear order (ordering)*) iff for all $a, b \in X$, either $a \leq b$ or $b \leq a$.

When neither $a \leq b$ nor $b \leq a$, we say that *a and b are incomparable*.

A subset, $C \subseteq X$, is a *chain* iff $\leq$ induces a total order on C (so, for all $a, b \in C$, either $a \leq b$ or $b \leq a$).

The *strict order (ordering), $<$, associated with $\leq$* is the relation defined by: $a < b$ iff $a \leq b$ and $a \neq b$.

If $\leq$ is a partial order on X, we say that the pair $\langle X, \leq \rangle$ is a *partially ordered set* or for short, a *poset*.

Remark: Observe that if $<$ is the strict order associated with a partial order, $\leq$, then $<$ is transitive and *anti-reflexive*, which means that

(4) $a \not< a$, for all $a \in X$.

Conversely, let $<$ be a relation on X and assume that $<$ is transitive and anti-reflexive.

Then, we can define the relation $\leq$ so that $a \leq b$ iff $a = b$ or $a < b$.

It is easy to check that $\leq$ is a partial order and that the strict order associated with $\leq$ is our original relation, $<$.

Given a poset, $\langle X, \leq\rangle$, by abuse of notation, we often refer to $\langle X, \leq\rangle$ as the *poset X*, the partial order $\leq$ being implicit.

If confusion may arise, for example when we are dealing with several posets, we denote the partial order on X by $\leq_X$.

Here are a few examples of partial orders.

1. **The subset ordering**. We leave it to the reader to check that the subset relation, $\subseteq$, on a set, X, is indeed a partial order.

   For example, if $A \subseteq B$ and $B \subseteq A$, where $A, B \subseteq X$, then $A = B$, since these assumptions are exactly those needed by the extensionality axiom.

2. **The natural order on $\mathbb{N}$**. Although we all know what is the ordering of the natural numbers, we should realize that if we stick to our axiomatic presentation where we defined the natural numbers as sets that belong to every inductive set (see Definition 1.10.3), then we haven't yet defined this ordering.

However, this is easy to do since the natural numbers are sets. For any $m, n \in \mathbb{N}$, define $m \leq n$ as $m = n$ or $m \in n$.

Then, it is not hard check that this relation is a total order (Actually, some of the details are a bit tedious and require induction, see Enderton [4], Chapter 4).

3. **Orderings on strings**. Let $\Sigma = \{a_1, \ldots, a_n\}$ be an alphabet. The prefix, suffix and substring relations defined in Section 2.11 are easily seen to be partial orders.

   However, these orderings are not total. It is sometimes desirable to have a total order on strings and, fortunately, the lexicographic order (also called dictionnary order) achieves this goal.

In order to define the *lexicographic order* we assume that the symbols in Σ are totally ordered, $a_1 < a_2 < \cdots < a_n$. Then, given any two strings, $u, v \in \Sigma^*$, we set

$$u \preceq v \quad \left\{ \begin{array}{l} \text{if } v = uy, \text{ for some } y \in \Sigma^*, \text{ or} \\ \text{if } u = xa_iy, \; v = xa_jz, \\ \text{and } a_i < a_j, \text{ for some } x, y, z \in \Sigma^*. \end{array} \right.$$

In other words, either u is a prefix of v or else u and v share a common prefix, x, and then there is a differring symbol, a_i in u and a_j in v, with $a_i < a_j$.

It is fairly tedious to prove that the lexicographic order is a partial order. Moreover, the lexicographic order is a total order.

4. **The divisibility order on $\mathbb{N}$.** Let us begin by defining divisibility in $\mathbb{Z}$.

   Given any two integers, $a, b \in \mathbb{Z}$, with $b \neq 0$, we say that *b divides a* (*a is a multiple of b*) iff $a = bq$ for some $q \in \mathbb{Z}$.

   Such a q is called the *quotient of a and b*. Most number theory books use the notation $b \mid a$ to express that b divides a.

   For example, $4 \mid 12$ since $12 = 4 \cdot 3$ and $7 \mid -21$ since $-21 = 7 \cdot (-3)$ but 3 does not divide 16 since 16 is not an integer multiple of 3.

   We leave the verification that the divisibility relation is reflexive and transitive as an easy exercise. It is also transtive on $\mathbb{N}$ and so, it indeed a partial order on $\mathbb{N}_+$.

Given a poset, $\langle X \leq\rangle$, if X is finite, then there is a convenient way to describe the partial order $\leq$ on X using a graph.

Consider an arbitrary poset, $\langle X \leq\rangle$ (not necessarily finite). Given any element, $a \in X$, the following situations are of interest:

1. For **no** $b \in X$ do we have $b < a$. We say that a is a *minimal element* (of X).
2. There is some $b \in X$ so that $b < a$ and there is **no** $c \in X$ so that $b < c < a$. We say that b is an *immediate predecessor of a*.
3. For **no** $b \in X$ do we have $a < b$. We say that a is a *maximal element* (of X).
4. There is some $b \in X$ so that $a < b$ and there is **no** $c \in X$ so that $a < c < b$. We say that b is an *immediate successor of a*.

Note that an element may have more than one immediate predecessor (or more than one immediate successor).

If X is a finite set, then it is easy to see that every element that is not minimal has an immediate predecessor and any element that is not maximal has an immediate successor (why?).

But if X is infinite, for example, $X = \mathbb{Q}$, this may not be the case. Indeed, given any two distinct rational numbers, $a, b \in \mathbb{Q}$, we have

$$a < \frac{a+b}{2} < b.$$

Let us now use our notion of immediate predecessor to draw a diagram representing a finite poset, $\langle X, \leq \rangle$.

The trick is to draw a picture consisting of nodes and oriented edges, where the nodes are all the elements of X and where we draw an oriented edge from a to b iff a is an immediate predecessor of b.

Such a diagram is called a *Hasse diagram* for $\langle X, \leq \rangle$.

Observe that if $a < c < b$, then the diagram does **not** have an edge corresponding to the relation $a < b$.

A Hasse diagram is an economical representation of a finite poset and it contains the same amount of information as the partial order, $\leq$.

Here is the diagram associated with the partial order on the power set of the two element set, $\{a, b\}$:

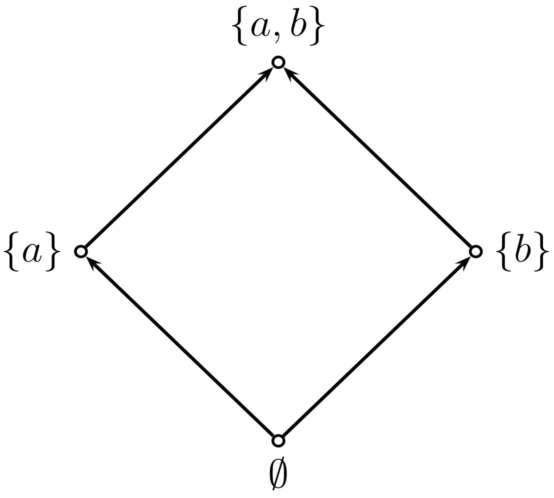

Figure 5.1: The partial order of the power set $2^{\{a,b\}}$

Here is the diagram associated with the partial order on the power set of the three element set, $\{a, b, c\}$:

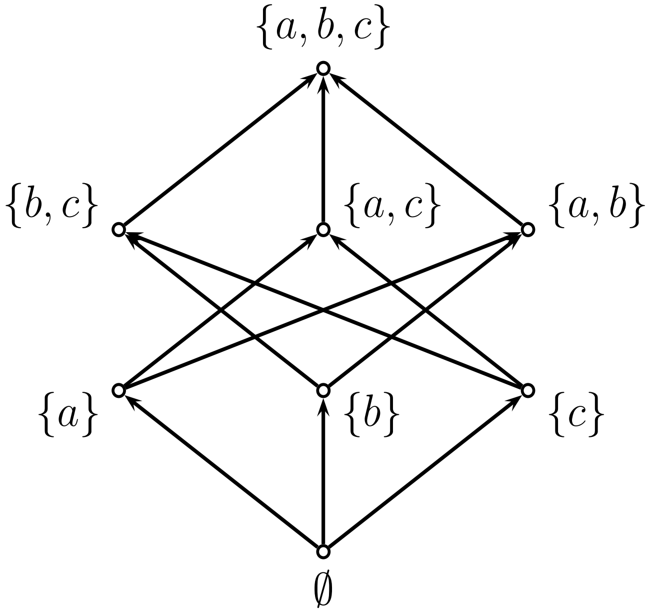

Figure 5.2: The partial order of the power set $2^{\{a,b,c\}}$

Note that $\emptyset$ is a minimal element of the above poset (in fact, the smallest element) and $\{a, b, c\}$ is a maximal element (in fact, the greatest element).

In the above example, there is a unique minimal (resp. maximal) element.

A less trivial example with multiple minimal and maximal elements is obtained by deleting $\emptyset$ and $\{a, b, c\}$:

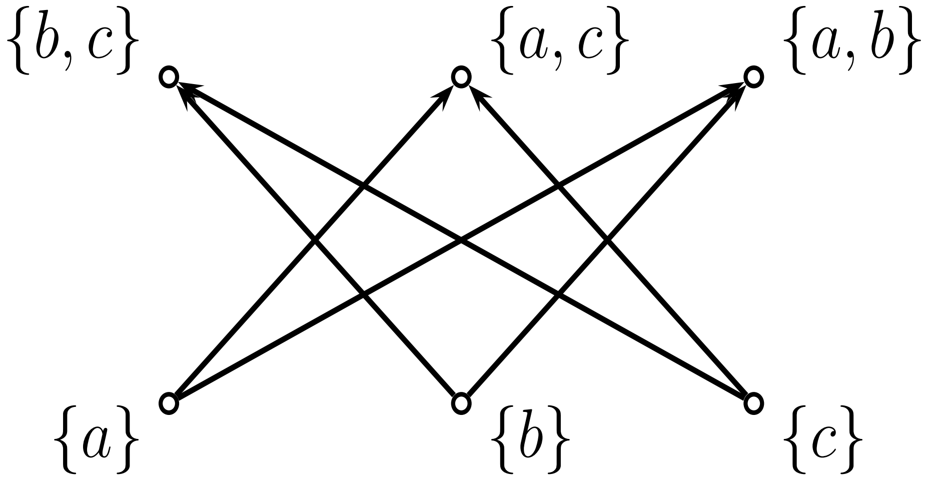

Figure 5.3: Minimal and maximal elements in a poset

Given a poset, $\langle X, \leq \rangle$, observe that if there is some element $m \in X$ so that $m \leq x$ for all $x \in X$, then m is unique.

Such an element, m, is called the *smallest* or the *least element* of X.

Similarly, an element, $b \in X$, so that $x \leq b$ for all $x \in X$ is unique and is called the *greatest element* of X.

We summarize some of our previous definitions and introduce a few more useful concepts in

Definition 5.1.2 Let $\langle X, \leq \rangle$ be a poset and let $A \subseteq X$ be any subset of X. An element, $b \in X$, is a *lower bound of A* iff $b \leq a$ for all $a \in A$.

An element, $m \in X$, is an *upper bound of A* iff $a \leq m$ for all $a \in A$.

An element, $b \in X$, is the *least element of A* iff $b \in A$ and $b \leq a$ for all $a \in A$.

An element, $m \in X$, is the *greatest element of A* iff $m \in A$ and $a \leq m$ for all $a \in A$.

An element, $b \in A$, is *minimal in A* iff $a < b$ for no $a \in A$, or equivalently, if for all $a \in A$, $a \leq b$ implies that $a = b$.

An element, $m \in A$, is *maximal in A* iff $m < a$ for no $a \in A$, or equivalently, if for all $a \in A$, $m \leq a$ implies that $a = m$.

An element, $b \in X$, is the *greatest lower bound of A* iff the set of lower bounds of A is nonempty and if b is the greatest element of this set.

An element, $m \in X$, is the *least upper bound of A* iff the set of upper bounds of A is nonempty and if m is the least element of this set.

Remarks:

1. If b is a lower bound of A (or m is an upper bound of A), then b (or m) may not belong to A.
2. The least element of A is a lower bound of A that also belongs to A and the greatest element of A is an upper bound of A that also belongs to A.

   When $A = X$, the least element is often denoted $\perp$, sometimes 0, and the greatest element is often denoted $\top$, sometimes 1.
3. Minimal or maximal elements of A belong to A but they are not necessarily unique.
4. The greatest lower bound (or the least upper bound) of A may not belong to A. We use the notation $\bigwedge A$ for the greatest lower bound of A and the notation $\bigvee A$ for the least upper bound of A.

In computer science, some people also use $\bigsqcup A$ instead of $\bigvee A$ and the symbol $\bigsqcup$ upside down instead of $\bigwedge$.

When $A = \{a, b\}$, we write $a \wedge b$ for $\bigwedge\{a, b\}$ and $a \vee b$ for $\bigvee\{a, b\}$.

The element $a \wedge b$ is called the *meet of a and b* and $a \vee b$ is the *join of a and b*. (Some computer scientists use $a \sqcap b$ for $a \wedge b$ and $a \sqcup b$ for $a \vee b$.)

5. Observe that if it exists, $\bigwedge \emptyset = \top$, the greatest element of X and if its exists, $\bigvee \emptyset = \perp$, the least element of X.

   Also, if it exists, $\bigwedge X = \perp$ and if it exists, $\bigvee X = \top$.

For the sake of completeness, we state the following fundamental result known as Zorn's Lemma even though it is unlikely that we will use it in this course.

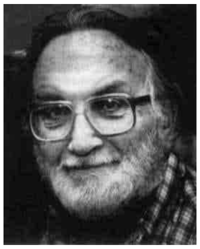

Figure 5.4: Max Zorn, 1906-1993

Zorn's lemma turns out to be equivalent to the axiom of choice.

Theorem 5.1.3 *(Zorn's Lemma) Given a poset, $\langle X, \leq \rangle$, if every nonempty chain in X has an upper-bound, then X has some maximal element.*

When we deal with posets, it is useful to use functions that are order-preserving as defined next.

Definition 5.1.4 Given two posets $\langle X, \leq_X \rangle$ and $\langle Y, \leq_Y \rangle$, a function, $f \colon X \to Y$, is *monotonic* (or *order-preserving*) iff for all $a, b \in X$,

$$\text{if} \quad a \leq_X b \quad \text{then} \quad f(a) \leq_Y f(b).$$

5.2 Lattices and Tarski's Fixed Point Theorem

We now take a closer look at posets having the property that every two elements have a meet and a join (a greatest lower bound and a least upper bound).

Such posets occur a lot more than we think. A typical example is the power set under inclusion, where meet is intersection and join is union.

Definition 5.2.1 A *lattice* is a poset in which any two elements have a meet and a join. A *complete lattice* is a poset in which any subset has a greatest lower bound and a least upper bound.

According to part (5) of the remark just before Zorn's Lemma, observe that a complete lattice must have a least element, $\perp$, and a greatest element, $\top$.

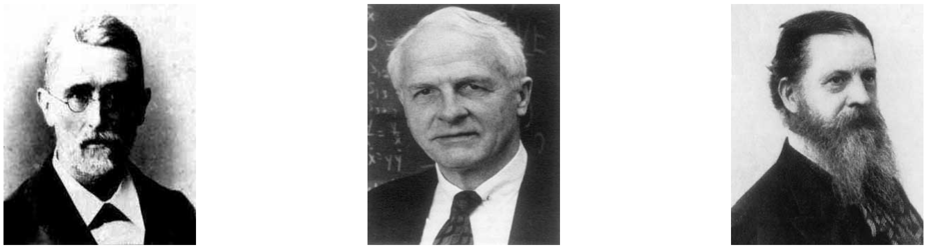

Figure 5.5: J.W. Richard Dedekind, 1831-1916 (left), Garrett Birkhoff, 1911-1996 (middle) and Charles S. Peirce, 1839-1914 (right)

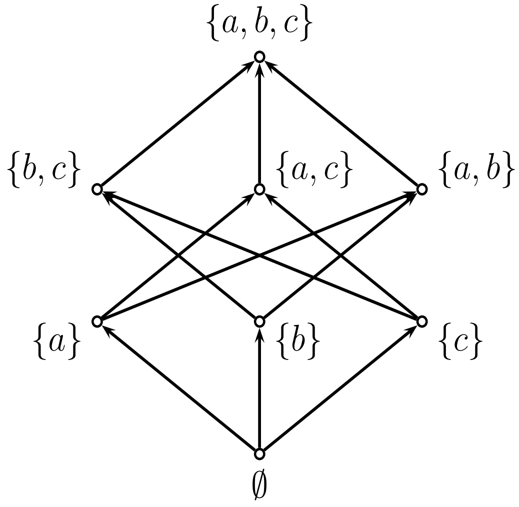

Figure 5.6: The lattice $2^{\{a,b,c\}}$

Remark: The notion of complete lattice is due to G. Birkhoff (1933). The notion of a lattice is due to Dedekind (1897) but his definition used properties (L1)-(L4) listed in Proposition 5.2.2. The use of meet and join in posets was first studied by C. S. Peirce (1880).

Figure 5.6 shows the lattice structure of the power set of $\{a, b, c\}$. It is actually a complete lattice.

It is easy to show that any finite lattice is a complete lattice and that a finite poset is a lattice iff it has a least element and a greatest element.

The poset $\mathbb{N}_+$ under the divisibility ordering is a lattice! Indeed, it turns out that the meet operation corresponds to *greatest common divisor* and the join operation corresponds to *least common multiple*.

However, it is not a complete lattice.

The power set of any set, X, is a complete lattice under the subset ordering.

The following proposition gathers some useful properties of meet and join.

Proposition 5.2.2 *If X is a lattice, then the following identities hold for all $a, b, c \in X$:*

$$\begin{array}{lll}
L1 & a \vee b = b \vee a, & a \wedge b = b \wedge a \\
L2 & (a \vee b) \vee c = a \vee (b \vee c), & (a \wedge b) \wedge c = a \wedge (b \wedge c) \\
L3 & a \vee a = a, & a \wedge a = a \\
L4 & (a \vee b) \wedge a = a, & (a \wedge b) \vee a = a.
\end{array}$$

Properties (L1) correspond to commutativity, properties (L2) to associativity, properties (L3) to idempotence and properties (L4) to absorption. Furthermore, for all $a, b \in X$, we have

$$a \leq b \quad \textit{iff} \quad a \vee b = b \quad \textit{iff} \quad a \wedge b = a,$$

called consistency.

Properties (L1)-(L4) are algebraic properties that were found by Dedekind (1897).

A pretty symmetry reveals itself in these identities: they all come in pairs, one involving $\wedge$, the other involving $\vee$.

A useful consequence of this symmetry is *duality*, namely, that each equation derivable from (L1)-(L4) has a dual statement obtained by exchanging the symbols $\wedge$ and $\vee$.

What is even more interesting is that it is possible to use these properties to define lattices.

Indeed, if X is a set together with two operations, $\wedge$ and $\vee$, satisfying (L1)-(L4), we can define the relation $a \leq b$ by $a \vee b = b$ and then show that $\leq$ is a partial order such that $\wedge$ and $\vee$ are the corresponding meet and join.

Proposition 5.2.3 *Let X be a set together with two operations $\wedge$ and $\vee$ satisfying the axioms (L1)-(L4) of proposition 5.2.2. If we define the relation $\leq$ by $a \leq b$ iff $a \vee b = b$ (equivalently, $a \wedge b = a$), then $\leq$ is a partial order and $(X, \leq)$ is a lattice whose meet and join agree with the original operations $\wedge$ and $\vee$.*

Figure 5.7: Alferd Tarksi, 1902-1983

The following proposition shows that the existence of arbitrary least upper bounds (or arbitrary greatest lower bounds) is already enough ensure that a poset is a complete lattice.

Proposition 5.2.4 *Let $\langle X, \leq \rangle$ be a poset. If X has a greatest element, $\top$, and if every nonempty subset, A, of X has a greatest lower bound, $\bigwedge A$, then X is a complete lattice. Dually, if X has a least element, $\perp$, and if every nonempty subset, A, of X has a least upper bound, $\bigvee A$, then X is a complete lattice*

We are now going to prove a remarkable result due to A. Tarski (discovered in 1942, published in 1955).

A special case (for power sets) was proved by B. Knaster (1928). First, we define fixed points.

Definition 5.2.5 Let $\langle X, \leq\rangle$ be a poset and let $f\colon X \to X$ be a function. An element, $x \in X$, is a *fixed point of f* (sometimes spelled *fixpoint*) iff

$$f(x) = x.$$

An element, $x \in X$, is a *least (resp. greatest) fixed point of f* if it is a fixed point of f and if $x \leq y$ (resp. $y \leq x$) for every fixed point y of f.

Fixed points play an important role in certain areas of mathematics (for example, topology, differential equations) and also in economics because they tend to capture the notion of stability or equilibrium.

We now prove the following pretty theorem due to Tarski and then immediately proceed to use it to give a very short proof of the Schröder-Bernstein Theorem (Theorem 2.9.18).

Theorem 5.2.6 *(Tarski's Fixed Point Theorem) Let $\langle X, \leq\rangle$ be a complete lattice and let $f\colon X \to X$ be any monotonic function. Then, the set, F, of fixed points of f is a complete lattice. In particular, f has a least fixed point,*

$$x_{\min} = \bigwedge\{x \in X \mid f(x) \leq x\}$$

and a greatest fixed point

$$x_{\max} = \bigvee\{x \in X \mid x \leq f(x)\}.$$

It should be noted that the least upper bounds and the greatest lower bounds in F do not necessarily agree with those in X. In technical terms, F is generally not a sublattice of X.

Now, as promised, we use Tarski's Fixed Point Theorem to prove the Schröder-Bernstein Theorem.

Theorem 2.9.18 *Given any two sets, A and B, if there is an injection from A to B and an injection from B to A, then there is a bijection between A and B.*

The proof is probably the shortest known proof of the Schröder-Bernstein Theorem because it uses Tarski's fixed point theorem, a powerful result.

If one looks carefully at the proof, one realizes that there are two crucial ingredients:

1. The set C is closed under $g \circ f$, that is, $g \circ f(C) \subseteq C$.
2. $A - C \subseteq g(B)$.

Using these observations, it is possible to give a proof that circumvents the use of Tarski's theorem. Such a proof is given in Enderton [4], Chapter 6.

We now turn to special properties of partial orders having to do with induction.

5.3 Well-Founded Orderings and Complete Induction

Have you ever wondered why induction on $\mathbb{N}$ actually "works"?

The answer, of course, is that $\mathbb{N}$ was defined in such a way that, by Theorem 1.10.4, it is the "smallest" inductive set!

But this is not a very illuminating answer. *The key point is that every nonempty subset of $\mathbb{N}$ has a least element*.

This fact is intuitively clear since if we had some nonempty subset of $\mathbb{N}$ with no smallest element, then we could construct an infinite strictly decreasing sequence, $k_0 > k_1 > \cdots > k_n > \cdots$. But this is absurd, as such a sequence would eventually run into 0 and stop.

It turns out that the deep reason why induction "works" on a poset is indeed that the poset ordering has a very special property and this leads us to the following definition:

Definition 5.3.1 Given a poset, $\langle X, \leq \rangle$, we say that $\leq$ is a *well-order (well ordering)* and that X is *well-ordered by $\leq$* iff every nonempty subset of X has a least element.

When X is nonempty, if we pick any two-element subset, $\{a, b\}$, of X, since the subset $\{a, b\}$ must have a least element, we see that either $a \leq b$ or $b \leq a$, *i.e.*, *every well-order is a total order*. First, let us confirm that $\mathbb{N}$ is indeed well-ordered.

Theorem 5.3.2 *(Well-Ordering of $\mathbb{N}$) The set of natural numbers, $\mathbb{N}$, is well-ordered.*

Theorem 5.3.2 yields another induction principle which is often more flexible that our original induction principle.

This principle called *complete induction* (or sometimes *strong induction*) was already encountered in Section 2.3.

It turns out that it is a special case of induction on a well-ordered set but it does not hurt to review it in the special case of the natural ordering on $\mathbb{N}$. Recall that $\mathbb{N}_+ = \mathbb{N} - \{0\}$.

Complete Induction Principle on $\mathbb{N}$.

In order to prove that a predicate, $P(n)$, holds for all $n \in \mathbb{N}$ it is enough to prove that

(1) $P(0)$ holds (the base case) and

(2) for every $m \in \mathbb{N}_+$, if $(\forall k \in \mathbb{N})(k < m \Rightarrow P(k))$ then $P(m)$.

As a formula, complete induction is stated as

$$P(0) \wedge (\forall m \in \mathbb{N}_+)[(\forall k \in \mathbb{N})(k < m \Rightarrow P(k)) \Rightarrow P(m)] \Rightarrow (\forall n \in \mathbb{N}) P(n).$$

The difference between ordinary induction and complete induction is that in complete induction, the induction hypothesis, $(\forall k \in \mathbb{N})(k < m \Rightarrow P(k))$, assumes that $P(k)$ holds for all $k < m$ and not just for $m - 1$ (as in ordinary induction), in order to deduce $P(m)$.

This gives us more proving power as we have more knowledge in order to prove $P(m)$.

We will have many occasions to use complete induction but let us first check that it is a valid principle.

Theorem 5.3.3 *The complete induction principle for* $\mathbb{N}$ *is valid.*

Remark: In our statement of the principle of complete induction, we singled out the base case, (1), and consequently, we stated the induction step (2) for every $m \in \mathbb{N}_+$, excluding the case $m = 0$, which is already covered by the base case.

It is also possible to state the principle of complete induction in a more concise fashion as follows:

$$(\forall m \in \mathbb{N})[(\forall k \in \mathbb{N})(k < m \Rightarrow P(k)) \Rightarrow P(m)] \\ \Rightarrow (\forall n \in \mathbb{N})P(n).$$

In the above formula, observe that when $m = 0$, which is now allowed, the premise $(\forall k \in \mathbb{N})(k < m \Rightarrow P(k))$ of the implication within the brackets is trivially true and so, $P(0)$ must still be established.

In the end, exactly the same amount of work is required but some people prefer the second more concise version of the principle of complete induction.

We feel that it would be easier for the reader to make the transition from ordinary induction to complete induction if we make explicit the fact that the base case must be established.

Let us illustrate the use of the complete induction principle by proving that every natural number factors as a product of primes.

Recall that for any two natural numbers, $a, b \in \mathbb{N}$ with $b \neq 0$, we say that *b divides a* iff $a = bq$, for some $q \in \mathbb{N}$.

In this case, we say that *a is divisible by b* and that *b is a factor of a*.

Then, we say that a natural number, $p \in \mathbb{N}$, is a *prime number* (for short, a *prime*) if $p \geq 2$ and if p is only divisible by itself and by 1.

Any prime number but 2 must be odd but the converse is false.

For example, $2, 3, 5, 7, 11, 13, 17$ are prime numbers, but 9 is not.

There are infinitely many prime numbers but to prove this, we need the following Theorem:

Theorem 5.3.4 *Every natural number, $n \geq 2$ can be factored as a product of primes, that is, n can be written as a product, $n = p_1^{m_1} \cdots p_k^{m_k}$, where the p_is are pairwise distinct prime numbers and $m_i \geq 1$ $(1 \leq i \leq k)$.*

For example, $21 = 3^1 \cdot 7^1$, $98 = 2^1 \cdot 7^2$, and $396 = 2^2 \cdot 3^3 \cdot 11$.

Remark: The prime factorization of a natural number is unique up to permutation of the primes $p_1, \ldots, p_k$ but this requires the Euclidean Division Lemma.

However, we can prove right away that there are infinitely primes.

Theorem 5.3.5 *Given any natural number, $n \geq 1$, there is a prime number, p, such that $p > n$. Consequently, there are infinitely many primes.*

Proof. Consider $m = n! + 1$.

As an application of Theorem 5.3.2, we prove the "Euclidean Division Lemma" for the integers.

Theorem 5.3.6 *(Euclidean Division Lemma for $\mathbb{Z}$)* *Given any two integers, $a, b \in \mathbb{Z}$, with $b \neq 0$, there is some unique integer, $q \in \mathbb{Z}$ (the quotient), and some unique natural number, $r \in \mathbb{N}$ (the remainder or residue), so that*

$$a = bq + r \quad \text{with} \quad 0 \leq r < |b|.$$

For example, $12 = 5 \cdot 2 + 2$, $200 = 5 \cdot 40 + 0$, and $42823 = 6409 \times 6 + 4369$.

The remainder, r, in the Euclidean division, $a = bq + r$, of a by b, is usually denoted $a \bmod b$.

We will now show that complete induction holds for a very broad class of partial orders called *well-founded orderings* that subsume well-orderings.

Definition 5.3.7 Given a poset, $\langle X, \leq \rangle$, we say that $\leq$ is a *well-founded ordering (order)* and that X is *well-founded* iff X has **no** infinite strictly decreasing sequence $x_0 > x_1 > x_2 > \cdots > x_n > x_{n+1} > \cdots$.

The following property of well-founded sets is fundamental:

Proposition 5.3.8 *A poset, $\langle X, \leq \rangle$, is well-founded iff every nonempty subset of X has a minimal element.*

So, the seemingly weaker condition that there is **no** infinite strictly decreasing sequence in X is equivalent to the fact that every nonempty subset of X has a minimal element.

If X is a total order, any minimal element is actually a least element and so, we get

Corollary 5.3.9 *A poset, $\langle X, \leq \rangle$, is well-ordered iff $\leq$ is total and X is well-founded.*

Note that the notion of a well-founded set is more general than that of a well-ordered set, since a well-founded set is not necessarily totally ordered.

Remark:

(ordinary) induction on $\mathbb{N}$ is valid
iff
complete induction on $\mathbb{N}$ is valid
iff
$\mathbb{N}$ is well-ordered.

These equivalences justify our earlier claim that the ability to do induction hinges on some key property of the ordering, in this case, that it is a well-ordering.

We finally come to the principle of *complete induction* (also called *transfinite induction* or *structural induction)*, which, as we shall prove, is valid for all well-founded sets.

Since every well-ordered set is also well-founded, complete induction is a very general induction method.

Let $(X, \leq)$ be a well-founded poset and let P be a predicate on X (*i.e.*, a function $P\colon X \to \{\mathbf{true}, \mathbf{false}\}$).

Principle of Complete Induction on a Well-Founded Set.

To prove that a property P holds for all $z \in X$, it suffices to show that, for every $x \in X$,

$(*)$ if x is minimal or $P(y)$ holds for all $y < x$,

$(**)$ then $P(x)$ holds.

The statement $(*)$ is called the *induction hypothesis*, and the implication

for all x, $(*)$ implies $(**)$ is called the *induction step*. Formally, the induction principle can be stated as:

$$(\forall x \in X)[(\forall y \in X)(y < x \Rightarrow P(y)) \Rightarrow P(x)] \\ \Rightarrow (\forall z \in X)P(z) \quad \text{(CI)}$$

Note that if x is minimal, then there is no $y \in X$ such that $y < x$, and $(\forall y \in X)(y < x \Rightarrow P(y))$ is true. Hence, we must show that $P(x)$ holds for every minimal element, x.

These cases are called the *base cases*.

Complete induction is not valid for arbitrary posets (see the problems) but holds for well-founded sets as shown in the following theorem.

Theorem 5.3.10 *The principle of complete induction holds for every well-founded set.*

As an illustration of well-founded sets, we define the *lexicographic ordering* on pairs.

Given a partially ordered set $\langle X, \leq \rangle$, the *lexicographic ordering*, $<<$, on $X \times X$ induced by $\leq$ is defined a follows: For all $x, y, x', y' \in X$,

$$(x, y) << (x', y') \quad \text{iff either}$$

$$\begin{array}{llll} x = x' & \text{and} & y = y' & \text{or} \\ x < x' & \text{or} & & \\ x = x' & \text{and} & y < y'. & \end{array}$$

We leave it as an exercise to check that $<<$ is indeed a partial order on $X \times X$. The following proposition will be useful.

Proposition 5.3.11 *If $\langle X, \leq \rangle$ is a well-founded set, then the lexicographic ordering $<<$ on $X \times X$ is also well founded.*

Example (*Ackermann's function*) The following function, $A\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, known as *Ackermann's function* is well known in recursive function theory for its extraordinary rate of growth. It is defined recursively as follows:

$$\begin{aligned} A(x, y) = \; & \textbf{if } x = 0 \textbf{ then } y + 1 \\ & \textbf{else if } y = 0 \textbf{ then } A(x - 1, 1) \\ & \textbf{else } A(x - 1, A(x, y - 1)). \end{aligned}$$

We wish to prove that A is a total function. We proceed by complete induction over the lexicographic ordering on $\mathbb{N} \times \mathbb{N}$.

1. The base case is $x = 0$, $y = 0$. In this case, since $A(0, y) = y + 1$, $A(0, 0)$ is defined and equal to 1.

2. The induction hypothesis is that for any (m, n), $A(m', n')$ is defined for all $(m', n') << (m, n)$, with $(m, n) \neq (m', n')$.

3. For the induction step, we have three cases:

   (a) If $m = 0$, since $A(0, y) = y + 1$, $A(0, n)$ is defined and equal to $n + 1$.

   (b) If $m \neq 0$ and $n = 0$, since $(m - 1, 1) << (m, 0)$ and $(m-1, 1) \neq (m, 0)$, by the induction hypothesis, $A(m-1, 1)$ is defined, and so $A(m, 0)$ is defined since it is equal to $A(m - 1, 1)$.

   (c) If $m \neq 0$ and $n \neq 0$, since $(m, n - 1) << (m, n)$ and $(m, n-1) \neq (m, n)$, by the induction hypothesis, $A(m, n - 1)$ is defined. Since
   $(m - 1, y) << (m, z)$ and $(m - 1, y) \neq (m, z)$ no matter what y and z are,
   $(m - 1, A(m, n - 1)) << (m, n)$ and
   $(m - 1, A(m, n - 1)) \neq (m, n)$, and by the induction hypothesis, $A(m - 1, A(m, n - 1))$ is defined. But this is precisely $A(m, n)$, and so $A(m, n)$ is defined. This concludes the induction step.

Hence, $A(x, y)$ is defined for all $x, y \geq 0$. □

5.4 Unique Prime Factorization in $\mathbb{Z}$ and GCD's

In the previous section, we proved that every natural number, $n \geq 2$, can be factored as a product of primes numbers.

In this section, we use the Euclidean Division Lemma to prove that such a factorization is unique.

For this, we need to introduce greatest common divisors (gcd's) and prove some of their properties.

In this section, it will be convenient to allow 0 to be a divisor. So, given any two integers, $a, b \in \mathbb{Z}$, we will say that *b divides a and that a is a multiple of b* iff $a = bq$, for some $q \in \mathbb{Z}$.

Contrary to our previous definition, $b = 0$ is allowed as a divisor.

However, this changes very little because if 0 divides a, then $a = 0q = 0$, that is, *the only integer divisible by 0 is 0*.

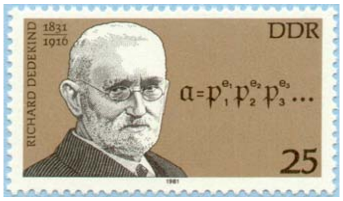

Figure 5.8: Richard Dedekind, 1831-1916

The notation $b \mid a$ is usually used to denote that b divides a. For example, $3 \mid 21$ since $21 = 2 \cdot 7$, $5 \mid -20$ since $-20 = 5 \cdot (-4)$ but 3 does not divide 20.

We begin by introducing a very important notion in algebra, that of an ideal due to Richard Dedekind, and prove a fundamental property of the ideals of $\mathbb{Z}$.

Definition 5.4.1 An *ideal of $\mathbb{Z}$* is any nonempty subset, $\mathfrak{I}$, of $\mathbb{Z}$ satisfying the following two properties:

(ID1) If $a, b \in \mathfrak{I}$, then $b - a \in \mathfrak{I}$.

(ID2) If $a \in \mathfrak{I}$, then $ak \in \mathfrak{I}$ for every $k \in \mathbb{Z}$.

An ideal, $\mathfrak{I}$, is a *principal ideal* if there is some $a \in \mathfrak{I}$, *called a generator*, such that $\mathfrak{I} = \{ak \mid k \in \mathbb{Z}\}$. The equality $\mathfrak{I} = \{ak \mid k \in \mathbb{Z}\}$ is also written as $\mathfrak{I} = a\mathbb{Z}$ or as $\mathfrak{I} = (a)$. The ideal $\mathfrak{I} = (0) = \{0\}$ is called the *null ideal*.

Note that if $\mathfrak{I}$ is an ideal, then $\mathfrak{I} = \mathbb{Z}$ iff $1 \in \mathfrak{I}$.

Since by definition, an ideal $\mathfrak{I}$ is nonempty, there is some $a \in \mathfrak{I}$, and by (ID1) we get $0 = a - a \in \mathfrak{I}$.

Then, for every $a \in \mathfrak{I}$, since $0 \in \mathfrak{I}$, by (ID1) we get $-a \in \mathfrak{I}$.

Theorem 5.4.2 *Every ideal, $\mathfrak{I}$, of $\mathbb{Z}$, is a principal ideal, i.e., $\mathfrak{I} = m\mathbb{Z}$ for some unique $m \in \mathbb{N}$, with $m > 0$ iff $\mathfrak{I} \neq (0)$.*

Theorem 5.4.2 is often phrased: $\mathbb{Z}$ is a *principal ideal domain*, for short, a *PID*.

Note that the natural number m such that $\mathfrak{I} = m\mathbb{Z}$ is a divisor of every element in $\mathfrak{I}$.

Figure 5.9: Étienne Bézout, 1730-1783

Corollary 5.4.3 *For any two integers, $a, b \in \mathbb{Z}$, there is a unique natural number, $d \in \mathbb{N}$, and some integers, $u, v \in \mathbb{Z}$, so that d divides both a and b and*

$$ua + vb = d.$$

(The above is called the Bezout identity.) Furthermore, $d = 0$ iff $a = 0$ and $b = 0$.

Given any nonempty finite set of integers, $S = \{a_1, \ldots, a_n\}$, it is easy to verify that the set

$$\mathfrak{I} = \{k_1 a_1 + \cdots + k_n a_n \mid k_1, \ldots, k_n \in \mathbb{Z}\}$$

is an ideal of $\mathbb{Z}$ and, in fact, the smallest (under inclusion) ideal containing S.

This ideal is called the *ideal generated by S* and it is often denoted $(a_1, \ldots, a_n)$.

Corollary 5.4.3 can be restated by saying that for any two distinct integers, $a, b \in \mathbb{Z}$, there is a unique natural number, $d \in \mathbb{N}$, such that the ideal, (a, b), generated by a and b is equal to the ideal $d\mathbb{Z}$ (also denoted (d)), that is,

$$(a, b) = d\mathbb{Z}.$$

This result still holds when $a = b$; in this case, we consider the ideal $(a) = (b)$.

With a slight (but harmless) abuse of notation, when $a = b$, we will also denote this ideal by (a, b).

The natural number d of corollary 5.4.3 divides both a and b.

Moreover, every divisor of a and b divides $d = ua + vb$. This motivates the definition:

Definition 5.4.4 Given any two integers, $a, b \in \mathbb{Z}$, an integer, $d \in \mathbb{Z}$, is a *greatest common divisor of a and b* (for short, a *gcd of a and b*) if d divides a and b and, for any integer, $h \in \mathbb{Z}$, if h divides a and b, then h divides d. We say that a and b are *relatively prime* if 1 is a gcd of a and b.

Remarks:

1. If $a = b = 0$, then, any integer, $d \in \mathbb{Z}$, is a divisor of 0. In particular, 0 divides 0. According to Definition 5.4.4, this implies $\gcd(0, 0) = 0$.

   The ideal generated by 0 is the trivial ideal, (0), so $\gcd(0, 0) = 0$ is equal to the generator of the zero ideal, (0).

   If $a \neq 0$ or $b \neq 0$, then the ideal, (a, b), generated by a and b is not the zero ideal and there is a unique integer, $d > 0$, such that

$$(a, b) = d\mathbb{Z}.$$

For any gcd, d', of a and b, since d divides a and b, we see that d must divide d'. As d' also divides a and b, the number d' must also divide d. Thus, $d = d'q'$ and $d' = dq$ for some $q, q' \in \mathbb{Z}$ and so, $d = dqq'$ which implies $qq' = 1$ (since $d \neq 0$). Therefore, $d' = \pm d$.

So, according to the above definition, when $(a, b) \neq (0)$, gcd's are not unique. However, exactly one of d' or $-d'$ is positive and equal to the positive generator, d, of the ideal (a, b).

We will refer to this positive gcd as “the” gcd of a and b and write $d = \gcd(a, b)$. Observe that $\gcd(a, b) = \gcd(b, a)$.

For example, $\gcd(20, 8) = 4$, $\gcd(1000, 50) = 50$, $\gcd(42823, 6409) = 17$, and $\gcd(5, 16) = 1$.

2. Another notation commonly found for $\gcd(a, b)$ is (a, b), but this is confusing since (a, b) also denotes the ideal generated by a and b.

3. Observe that if $d = \gcd(a, b) \neq 0$, then d is indeed the largest positive common divisor of a and b since every divisor of a and b must divide d.

   However, we did not use this property as one of the conditions for being a gcd because such a condition does not generalize to other rings where a total order is not available.

   Another minor reason is that if we had used in the definition of a gcd the condition that $\gcd(a, b)$ should be the largest common divisor of a and b, as every integer divides 0, $\gcd(0, 0)$ would be undefined!

4. If $a = 0$ and $b > 0$, then the ideal, $(0, b)$, generated by 0 and b is equal to the ideal, $(b) = b\mathbb{Z}$, which implies $\gcd(0, b) = b$ and similarly, if $a > 0$ and $b = 0$, then $\gcd(a, 0) = a$.

Let $p \in \mathbb{N}$ be a prime number. Then, note that for any other integer, n, if p does not divide n, then $\gcd(p, n) = 1$, as the only divisors of p are 1 and p.

Proposition 5.4.5 *Given any two integers, $a, b \in \mathbb{Z}$, a natural number, $d \in \mathbb{N}$, is the greatest common divisor of a and b iff d divides a and b and if there are some integers, $u, v \in \mathbb{Z}$, so that*

$$ua + vb = d. \qquad \text{(Bezout Identity)}$$

In particular, a and b are relatively prime iff there are some integers, $u, v \in \mathbb{Z}$, so that

$$ua + vb = 1. \qquad \text{(Bezout Identity)}$$

The gcd of two natural numbers can be found using a method involving Euclidean division and so can the numbers u and v.

This method is based on the following simple observation:

Proposition 5.4.6 *If a, b are any two positive integers with $a \geq b$, then for every $k \in \mathbb{Z}$,*

$$\gcd(a, b) = \gcd(b, a - kb).$$

In particular,

$$\gcd(a, b) = \gcd(b, a - b) = \gcd(b, a + b),$$

and if $a = bq + r$ is the result of performing the Euclidean division of a by b, with $0 \leq r < a$, then

$$\gcd(a, b) = \gcd(b, r).$$

Using the fact that $\gcd(a, 0) = a$, we have the following algorithm for finding the gcd of two natural numbers, a, b, with $(a, b) \neq (0, 0)$:

Euclidean Algorithm for Finding the gcd.

The input consists of two natural numbers, m, n, with $(m, n) \neq (0, 0)$.

```
begin
  $a := m$; $b := n$;
  if $a < b$ then
    $t := b$; $b := a$; $a := t$; (swap $a$ and $b$)
  while $b \neq 0$ do
    $r := a \bmod b$; (divide $a$ by $b$ to obtain the remainder $r$)
    $a := b$; $b := r$
  endwhile;
  $\gcd(m, n) := a$
end
```

In order to prove the correctness of the above algorithm, we need to prove two facts:

1. The algorithm always terminates.
2. When the algorithm exits the while loop, the current value of a is indeed $\gcd(m, n)$.

The termination of the algorithm follows by induction on $\min\{m, n\}$.

The correctness of the algorithm is an immediate consequence of Proposition 5.4.6. During any round through the while loop, the invariant $\gcd(a, b) = \gcd(m, n)$ is preserved, and when we exit the while loop, we have

$$a = \gcd(a, 0) = \gcd(m, n),$$

which proves that the current value of a when the algorithm stops is indeed $\gcd(m, n)$.

Let us run the above algorithm for $m = 42823$ and $n = 6409$. There are five division steps:

$$\begin{aligned} 42823 &= 6409 \times 6 + 4369 \\ 6409 &= 4369 \times 1 + 2040 \\ 4369 &= 2040 \times 2 + 289 \\ 2040 &= 289 \times 7 + 17 \\ 289 &= 17 \times 17 + 0, \end{aligned}$$

so we find that

$$\gcd(42823, 6409) = 17.$$

You should also use your computation to find numbers x, y so that

$$42823x + 6409y = 17.$$

Check that $x = -22$ and $y = 147$ work.

The complexity of the Euclidean algorithm to compute the gcd of two natural numbers is quite interesting and has a long history.

It turns out that Gabriel Lamé published a paper in 1844 in which he proved that if $m > n > 0$, then the number of divisions needed by the algorithm is bounded by $5\delta + 1$, where δ is the number of digits in n. For this, Lamé realized that the maximum number of steps is achieved by taking m an n to be two consecutive Fibonacci numbers (see Section 5.7).

Dupré, in a paper published in 1845, improved the upper bound to $4.785\delta + 1$, also making use of the Fibonacci numbers.

Using a variant of Euclidean division allowing negative remainders, in a paper published in 1841, Binet gave an algorithm with an even better bound: $\frac{10}{3}\delta + 1$.

The Euclidean algorithm can be easily adapted to also compute two integers, x and y, such that

$$mx + ny = \gcd(m, n).$$

Such an algorithm is called the *Extended Euclidean Algorithm*.

What can be easily shown is the following proposition:

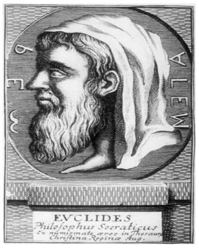

Figure 5.10: Euclid of Alexandria, about 325 BC – about 265 BC

Proposition 5.4.7 *The number of divisions made by the Euclidean Algorithm for gcd applied to two positive integers, m, n, with $m > n$, is at most $\log_2 m + \log_2 n$.*

We now return to Proposition 5.4.5 as it implies a very crucial property of divisibility in any PID.

Proposition 5.4.8 *(Euclid's proposition) Let $a, b, c \in \mathbb{Z}$ be any integers. If a divides bc and a is relatively prime to b, then a divides c.*

In particular, if p is a prime number and if p divides ab, where $a, b \in \mathbb{Z}$ are nonzero, then either p divides a or p divides b.

Proposition 5.4.9 *Let $a, b_1, \ldots, b_m \in \mathbb{Z}$ be any integers. If a and b_i are relatively prime for all i, with $1 \leq i \leq m$, then a and $b_1 \cdots b_m$ are relatively prime.*

One of the main applications of the Euclidean Algorithm is to find the inverse of a number in modular arithmetic, an essential step in the *RSA algorithm*, the first and still widely used algorithm for public-key cryptography.

Given any natural number, $p \geq 1$, we can define a relation on $\mathbb{Z}$, called *congruence*, as follows:

$$n \equiv m \pmod{p}$$

iff $p \mid n - m$, *i.e.*, iff $n = m + pk$, for some $k \in \mathbb{Z}$. We say that *m is a residue of n modulo p*.

The notation for congruence was introduced by Carl Friedrich Gauss (1777-1855), one of the greatest mathematicians of all time.

Figure 5.11: Carl Friedrich Gauss, 1777-1855

Gauss contributed significantly to the theory of congruences and used his results to prove deep and fundamental results in number theory.

If $n \geq 1$ and n and p are relatively prime, an *inverse of n modulo p* is a number, $s \geq 1$, such that

$$ns \equiv 1 \pmod{p}.$$

Using Proposition 5.4.8 (Euclid's proposition), it is easy to see that that if s_1 and s_2 are both inverse of n modulo p, then $s_1 \equiv s_2 \pmod{p}$.

Since finding an inverse of n modulo p means finding some numbers, x, y, so that $nx = 1 + py$, that is, $nx - py = 1$, we can find x and y using the Extended Euclidean Algorithm.

We can now prove the uniqueness of prime factorizations in $\mathbb{N}$. The first rigorous proof of this theorem was given by Gauss.

Theorem 5.4.10 *(Unique Prime Factorization in $\mathbb{N}$)* *For every natural number, $a \geq 2$, there exists a unique set, $\{\langle p_1, k_1\rangle, \ldots, \langle p_m, k_m\rangle\}$, where the p_i's are distinct prime numbers and the k_i's are (not necessarily distinct) integers, with $m \geq 1$, $k_i \geq 1$, so that*

$$a = p_1^{k_1} \cdots p_m^{k_m}.$$

Theorem 5.4.10 is a basic but very important result of number theory and it has many applications.

It also reveals the importance of the primes as the building blocks of all numbers.

Remark: Theorem 5.4.10 also applies to any nonzero integer $a \in \mathbb{Z} - \{-1, +1\}$, by adding a suitable sign in front of the prime factorization.

That is, we have a unique prime factorization of the form

$$a = \pm p_1^{k_1} \cdots p_m^{k_m}.$$

Theorem 5.4.10 shows that $\mathbb{Z}$ is a *unique factorization domain*, for short, a *UFD*.

Such rings play an important role because every nonzero element which is not a unit (*i.e.*, which is not invertible) has a unique factorization (up to some unit factor) into so-called *irreducible elements* which generalize the primes.

Readers who would like to learn more about number theory are strongly advised to read Silverman's delightful and very "friendly" introductory text [13].

5.5 Equivalence Relations and Partitions

Equivalence relations basically generalize the identity relation.

Technically, the definition of an equivalence relation is obtained from the definition of a partial order (Definition 5.1.1) by changing the third condition, antisymmetry, to *symmetry*.

Definition 5.5.1 A binary relation, R, on a set, X, is an *equivalence relation* iff it is *reflexive*, *transitive* and *symmetric*, that is:

(1) (*Reflexivity*): aRa, for all $a \in X$;

(2) (*Transitivity*): If aRb and bRc, then aRc, for all $a, b, c \in X$.

(3) (*symmetry*): If aRb, then bRa, for all $a, b \in X$.

Here are some examples of equivalence relations.

1. The identity relation, id_X, on a set X is an equivalence relation.
2. The relation $X \times X$ is an equivalence relation.
3. Let S be the set of students in CIS160. Define two students to be equivalent iff they were born the same year. It is trivial to check that this relation is indeed an equivalence relation.
4. Given any natural number, $p \geq 1$, recall that we can define a relation on $\mathbb{Z}$ as follows:

   $$n \equiv m \pmod{p}$$

   iff $p \mid n - m$, *i.e.*, $n = m + pk$, for some $k \in \mathbb{Z}$. It is an easy exercise to check that this is indeed an equivalence relation called *congruence modulo p*.

5. Equivalence of propositions is the relation defined so that $P \equiv Q$ iff $P \Rightarrow Q$ and $Q \Rightarrow P$ are both provable (say, classically). It is easy to check that logical equivalence is an equivalence relation.

6. Suppose $f\colon X \to Y$ is a function. Then, we define the relation $\equiv_f$ on X by

   $$x \equiv_f y \quad \text{iff} \quad f(x) = f(y).$$

   It is immediately verified that $\equiv_f$ is an equivalence relation. Actually, we are going to show that every equivalence relation arises in this way, in terms of (surjective) functions.

The crucial property of equivalence relations is that they *partition* their domain, X, into pairwise disjoint nonempty blocks. Intuitively, they carve out X into a bunch of puzzle pieces.

Definition 5.5.2 Given an equivalence relation, R, on a set, X, for any $x \in X$, the set

$$[x]_R = \{y \in X \mid xRy\}$$

is the *equivalence class of x*. Each equivalence class, $[x]_R$, is also denoted $\overline{x}_R$ and the subscript R is often omitted when no confusion arises. The set of equivalence classes of R is denoted by X/R. The set X/R is called the *quotient of X by R* or *quotient of X modulo R*. The function, $\pi\colon X \to X/R$, given by

$$\pi(x) = [x]_R, \quad x \in X,$$

is called the *canonical projection* (or *projection*) of X onto X/R.

Since every equivalence relation is reflexive, *i.e.*, xRx for every $x \in X$, observe that $x \in [x]_R$ for any $x \in R$, that is, every equivalence class is *nonempty*.

It is also clear that the projection, $\pi\colon X \to X/R$, is surjective.

The main properties of equivalence classes are given by

Proposition 5.5.3 *Let R be an equivalence relation on a set, X. For any two elements $x, y \in X$, we have*

$$xRy \quad \textit{iff} \quad [x] = [y].$$

Moreover, the equivalences classes of R satisfy the following properties:

(1) $[x] \neq \emptyset$, for all $x \in X$;

(2) If $[x] \neq [y]$ then $[x] \cap [y] = \emptyset$;

(3) $X = \bigcup_{x \in X} [x]$.

A useful way of interpreting Proposition 5.5.3 is to say that the equivalence classes of an equivalence relation form a partition, as defined next.

Definition 5.5.4 Given a set, X, a *partition of X* is any family, $\Pi = \{X_i\}_{i \in I}$, of subsets of X such that

(1) $X_i \neq \emptyset$, for all $i \in I$ (each X_i is nonempty);

(2) If $i \neq j$ then $X_i \cap X_j = \emptyset$ (the X_i are pairwise disjoint);

(3) $X = \bigcup_{i \in I} X_i$ (the family is exhaustive).

Each set X_i is called a *block* of the partition.

In the example where equivalence is determined by the same year of birth, each equivalence class consists of those students having the same year of birth.

Let us now go back to the example of congruence modulo p (with $p > 0$) and figure out what are the blocks of the corresponding partition. Recall that

$$m \equiv n \pmod{p}$$

iff $m - n = pk$ for some $k \in \mathbb{Z}$.

By the division Theorem (Theorem 5.3.6), we know that there exist some unique q, r, with $m = pq + r$ and $0 \leq r \leq p - 1$. Therefore, for every $m \in \mathbb{Z}$,

$$m \equiv r \pmod{p} \quad \text{with} \quad 0 \leq r \leq p - 1,$$

which shows that there are p equivalence classes,

$$[0], [1], \ldots, [p - 1],$$

where the equivalence class, $[r]$ (with $0 \leq r \leq p - 1$), consists of all integers of the form $pq + r$, where $q \in \mathbb{Z}$, $i.e.$, those integers whose residue modulo p is r.

Proposition 5.5.3 defines a map from the set of equivalence relations on X to the set of partitions on X.

Given any set, X, let $\mathrm{Equiv}(X)$ denote the set of equivalence relations on X and let $\mathrm{Part}(X)$ denote the set of partitions on X.

Then, Proposition 5.5.3 defines the function, $\Pi\colon \mathrm{Equiv}(X) \to \mathrm{Part}(X)$, given by,

$$\Pi(R) = X/R = \{[x]_R \mid x \in X\},$$

where R is any equivalence relation on X. We also write Π_R instead of $\Pi(R)$.

There is also a function, $\mathcal{R}\colon \mathrm{Part}(X) \to \mathrm{Equiv}(X)$, that assigns an equivalence relation to a partition a shown by the next proposition.

Proposition 5.5.5 *For any partition, $\Pi = \{X_i\}_{i \in I}$, on a set, X, the relation, $\mathcal{R}(\Pi)$, defined by*

$$x\mathcal{R}(\Pi)y \quad \textit{iff} \quad (\exists i \in I)(x, y \in X_i),$$

is an equivalence relation whose equivalence classes are exactly the blocks X_i.

Putting Propositions 5.5.3 and 5.5.5 together we obtain the useful fact there is a bijection between $\mathrm{Equiv}(X)$ and $\mathrm{Part}(X)$.

Therefore, in principle, it is a matter of taste whether we prefer to work with equivalence relations or partitions.

In computer science, it is often preferable to work with partitions, but not always.

Proposition 5.5.6 *Given any set, X, the functions* $\Pi\colon \mathrm{Equiv}(X) \to \mathrm{Part}(X)$ *and* $\mathcal{R}\colon \mathrm{Part}(X) \to \mathrm{Equiv}(X)$ *are mutual inverses, that is,*

$$\mathcal{R} \circ \Pi = \mathrm{id} \quad \textit{and} \quad \Pi \circ \mathcal{R} = \mathrm{id}.$$

Consequently, there is a bijection between the set, $\mathrm{Equiv}(X)$*, of equivalence relations on X and the set,* $\mathrm{Part}(X)$*, of partitions on X.*

Now, if $f\colon X \to Y$ is a surjective function, we have the equivalence relation, $\equiv_f$, defined by

$$x \equiv_f y \quad \text{iff} \quad f(x) = f(y).$$

It is clear that the equivalence class of any $x \in X$ is the inverse image, $f^{-1}(f(x))$, of $f(x) \in Y$.

Therefore, there is a bijection between $X/\equiv_f$ and Y. Thus, we can identify f and the projection, π, from X onto $X/\equiv_f$.

If f is not surjective, note that f is surjective onto $f(X)$ and so, we see that f can be written as the composition

$$f = i \circ \pi,$$

where $\pi\colon X \to f(X)$ is the canonical projection and $i\colon f(X) \to Y$ is the *inclusion function* mapping $f(X)$ into Y (*i.e.*, $i(y) = y$, for every $y \in f(X)$).

Given a set, X, the inclusion ordering on $X \times X$ defines an ordering on binary relations on X, namely,

$$R \leq S \quad \text{iff} \quad (\forall x, y \in X)(xRy \Rightarrow xSy).$$

When $R \leq S$, we say that *R refines S*.

If R and S are equivalence relations and $R \leq S$, we observe that every equivalence class of R is contained in some equivalence class of S.

Actually, in view of Proposition 5.5.3, we see that *every equivalence class of S is the union of equivalence classes of R*.

We also note that id_X is the least equivalence relation on X and $X \times X$ is the largest equivalence relation on X.

This suggests the following question: Is $\mathrm{Equiv}(X)$ a lattice under refinement?

The answer is yes. It is easy to see that the meet of two equivalence relations is $R \cap S$, their intersection.

But beware, their join is not $R \cup S$, because in general, $R \cup S$ is not transitive.

However, there is a least equivalence relation containing R and S, and this is the join of R and S. This leads us to look at various closure properties of relations.

5.6 Transitive Closure, Reflexive and Transitive Closure, Smallest Equivalence Relation

Let R be any relation on a set X. Note that R is reflexive iff $\mathrm{id}_X \subseteq R$. Consequently, the smallest reflexive relation containing R is $\mathrm{id}_X \cup R$. This relation is called the *reflexive closure of R*.

Note that R is transitive iff $R \circ R \subseteq R$. This suggests a way of making the smallest transitive relation containing R (if R is not already transitive). Define R^n by induction as follows:

$$\begin{aligned} R^0 &= \mathrm{id}_X \\ R^{n+1} &= R^n \circ R. \end{aligned}$$

Definition 5.6.1 Given any relation, R, on a set, X, the *transitive closure of R* is the relation, R^+, given by

$$R^+ = \bigcup_{n \geq 1} R^n.$$

The *reflexive and transitive closure of R* is the relation, R^*, given by

$$R^* = \bigcup_{n \geq 0} R^n = \mathrm{id}_X \cup R^+.$$

Proposition 5.6.2 *Given any relation, R, on a set, X, the relation R^+ is the smallest transitive relation containing R and R^* is the smallest reflexive and transtive relation containing R.*

If R is reflexive, then it is easy to see that $R \subseteq R^2$ and so, $R^k \subseteq R^{k+1}$ for all $k \geq 0$.

From this, we can show that if X is a finite set, then there is a smallest k so that $R^k = R^{k+1}$.

In this case, R^k is the reflexive and transitive closure of R. If X has n elements it can be shown that $k \leq n-1$.

Note that a relation, R, is symmetric iff $R^{-1} = R$.

As a consequence, $R \cup R^{-1}$ is the smallest symmetric relation containing R.

This relation is called the *symmetric closure of R*.

Finally, given a relation, R, what is the smallest equivalence relation containing R? The answer is given by

Proposition 5.6.3 *For any relation, R, on a set, X, the relation*

$$(R \cup R^{-1})^*$$

is the smalest equivalence relation containing R.

5.7 Fibonacci and Lucas Numbers; Mersenne Primes

We have encountered the Fibonacci numbers (after Leonardo Fibonacci, also known as *Leonardo of Pisa*, 1170-1250) in Section 2.3.

These numbers show up unexpectedly in many places, including algorithm design and analysis, for example, Fibonacci heaps.

The Lucas numbers (after Edouard Lucas, 1842-1891) are closely related to the Fibonacci numbers.

Both arise as special instances of the recurrence relation

$$u_{n+2} = u_{n+1} + u_n, \qquad n \geq 0$$

where u_0 and u_1 are some given initial values.

The *Fibonacci sequence*, (F_n), arises for $u_0 = 0$ and $u_1 = 1$ and the *Lucas sequence*, (L_n), for $u_0 = 2$ and $u_1 = 1$.

Figure 5.12: Leonardo Pisano Fibonacci, 1170-1250 (left) and F Edouard Lucas, 1842-1891 (right)

These two sequences turn out to be intimately related and they satisfy many remarquable identities.

The Lucas numbers play a role in testing for primality of certain kinds of numbers of the form $2^p - 1$, where p is a prime, known as *Mersenne numbers*.

In turns out that the largest known primes so far are Mersenne numbers and large primes play an important role in cryptography.

It is possible to derive a closed formulae for both F_n and L_n using some simple linear algebra.

Observe that the recurrence relation

$$u_{n+2} = u_{n+1} + u_n$$

yields the recurrence

$$\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_n \\ u_{n-1} \end{pmatrix}$$

for all $n \geq 1$, and so,

$$\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} u_1 \\ u_0 \end{pmatrix}$$

for all $n \geq 0$.

Now, the matrix

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

has characteristic polynomial, $\lambda^2 - \lambda - 1$, which has two real roots

$$\lambda = \frac{1 \pm \sqrt{5}}{2}.$$

Observe that the larger root is the famous *golden ratio*, often denoted

$$\varphi = \frac{1+\sqrt{5}}{2} = 1.618033988749\cdots$$

and that

$$\frac{1-\sqrt{5}}{2} = -\varphi^{-1}.$$

Since A has two distinct eigenvalues, it can be diagonalized and it is easy to show that

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} \varphi & -\varphi^{-1} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \varphi & 0 \\ 0 & -\varphi^{-1} \end{pmatrix} \begin{pmatrix} 1 & \varphi^{-1} \\ -1 & \varphi \end{pmatrix}.$$

It follows that

$$\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} \varphi & -\varphi^{-1} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} (\varphi^{-1}u_0 + u_1)\varphi^n \\ (\varphi u_0 - u_1)(-\varphi^{-1})^n \end{pmatrix},$$

and so,

$$u_n = \frac{1}{\sqrt{5}}\big((\varphi^{-1}u_0 + u_1)\varphi^n + (\varphi u_0 - u_1)(-\varphi^{-1})^n\big),$$

for all $n \geq 0$.

For the Fibonacci sequence, $u_0 = 0$ and $u_1 = 1$, so

$$\begin{aligned} F_n &= \frac{1}{\sqrt{5}}\big(\varphi^n - (-\varphi^{-1})^n\big) \\ &= \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right], \end{aligned}$$

a formula established by Jacques Binet (1786-1856) in 1843 and already known to Euler, Daniel Bernoulli and de Moivre.

Since

$$\frac{\varphi^{-1}}{\sqrt{5}} = \frac{\sqrt{5}-1}{2\sqrt{5}} < \frac{1}{2},$$

we see that F_n is the closest integer to $\frac{\varphi^n}{\sqrt{5}}$ and that

$$F_n = \left\lfloor \frac{\varphi^n}{\sqrt{5}} + \frac{1}{2} \right\rfloor .$$

It is also easy to see that

$$F_{n+1} = \varphi F_n + (-\varphi^{-1})^n,$$

which shows that the ratio F_{n+1}/F_n approaches φ as n goes to infinity.

For the Lucas sequence, $u_0 = 2$ and $u_1 = 1$, so

$$\varphi^{-1}u_0 + u_1 = 2\frac{(\sqrt{5}-1)}{2} + 1 = \sqrt{5},$$

$$\varphi u_0 - u_1 = 2\frac{(1+\sqrt{5})}{2} - 1 = \sqrt{5}$$

and we get

$$L_n = \varphi^n + (-\varphi^{-1})^n = \left(\frac{1+\sqrt{5}}{2}\right)^n + \left(\frac{1-\sqrt{5}}{2}\right)^n.$$

Since

$$\varphi^{-1} = \frac{\sqrt{5}-1}{2} < 0.62$$

it follows that L_n is the closest integer to φ^n.

When $u_0 = u_1$, since $\varphi - \varphi^{-1} = 1$, we get

$$u_n = \frac{u_0}{\sqrt{5}}\big(\varphi^{n+1} - (-\varphi^{-1})^{n+1}\big),$$

that is,

$$u_n = u_0 F_{n+1}.$$

Therefore, from now on, we assume that $u_0 \neq u_1$.

It is easy to prove by induction that

Proposition 5.7.1 *The following identities hold:*

$$\begin{aligned}
F_0^2 + F_1^2 + \cdots + F_n^2 &= F_n F_{n+1} \\
F_0 + F_1 + \cdots + F_n &= F_{n+2} - 1 \\
F_2 + F_4 + \cdots + F_{2n} &= F_{2n+1} - 1 \\
F_1 + F_3 + \cdots + F_{2n+1} &= F_{2n+2} \\
\sum_{k=0}^{n} k F_k &= n F_{n+2} - F_{n+3} + 2
\end{aligned}$$

for all $n \geq 0$ (with the third sum interpreted as F_0 for $n = 0$).

Following Knuth (see [8]), the third and fourth identities yield the identity

$$F_{(n \bmod 2)+2} + \cdots + F_{n-2} + F_n = F_{n+1} - 1,$$

for all $n \geq 2$.

The above can be used to prove the *Zeckendorf's representation* of the natural numbers (see Knuth [8], Chapter 6).

Proposition 5.7.2 *(Zeckendorf's representation) Every every natural number, $n \in \mathbb{N}$, with $n > 0$, has a unique representation of the form*

$$n = F_{k_1} + F_{k_2} + \cdots + F_{k_r},$$

with $k_i \geq k_{i+1} + 2$ for $i = 1, \ldots, r - 1$ and $k_r \geq 2$.

For example,

$$\begin{aligned} 30 &= 21 + 8 + 1 \\ &= F_8 + F_6 + F_2 \end{aligned}$$

and

$$\begin{aligned} 1000000 &= 832040 + 121393 + 46368 + 144 + 55 \\ &= F_{30} + F_{26} + F_{24} + F_{12} + F_{10}. \end{aligned}$$

The fact that

$$F_{n+1} = \varphi F_n + (-\varphi^{-1})^n$$

and the Zeckendorf's representation lead to an amusing method for converting between kilometers to miles (see [8], Section 6.6).

Indeed, φ is nearly the number of kilometers in a mile (the exact number is 1.609344 and $\varphi = 1.618033$). It follows that a distance of F_{n+1} kilometers is very nearly a distance of F_n miles!

Thus, to convert a distance, d, expressed in kilometers into a distance expressed in miles, first find the Zeckendorf's representation of d and then shift each F_{k_i} in this representation to F_{k_i-1}.

For example,

$$30 = 21 + 8 + 1 = F_8 + F_6 + F_2$$

so the corresponding distance in miles is

$$F_7 + F_6 + F_1 = 13 + 5 + 1 = 19.$$

The "exact" distance in miles is 18.64 miles.

We can prove two simple formulas for obtaining the Lucas numbers from the Fibonacci numbers and vice-versa:

Proposition 5.7.3 *The following identities hold:*

$$\begin{aligned} L_n &= F_{n-1} + F_{n+1} \\ 5F_n &= L_{n-1} + L_{n+1}, \end{aligned}$$

for all $n \geq 1$.

The Fibonaci sequence begins with

$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610$

and the Lucas sequence begins with

$2, 1, 3, 4, 7, 11, 18, 29, 47, 76, 123, 199, 322, 521, 843, 1364.$

Notice that $L_n = F_{n-1} + F_{n+1}$ is equivalent to

$$2F_{n+1} = F_n + L_n.$$

It can also be shown that

$$F_{2n} = F_n L_n,$$

for all $n \geq 1$.

The proof proceeds by induction but one finds that it is necessary to prove an auxiliary fact:

Proposition 5.7.4 *For any fixed $k \geq 1$ and all $n \geq 0$, we have*

$$F_{n+k} = F_k F_{n+1} + F_{k-1} F_n.$$

The reader can also prove that

$$\begin{aligned}
L_n L_{n+2} &= L_{n+1}^2 + 5(-1)^n \\
L_{2n} &= L_n^2 - 2(-1)^n \\
L_{2n+1} &= L_n L_{n+1} - (-1)^n \\
L_n^2 &= 5F_n^2 + 4(-1)^n.
\end{aligned}$$

Using the matrix representation derived earlier, it can be shown that

Proposition 5.7.5 *The sequence given by the recurrence*

$$u_{n+2} = u_{n+1} + u_n$$

satisfies the following equation:

$$u_{n+1}u_{n-1} - u_n^2 = (-1)^{n-1}(u_0^2 + u_0 u_1 - u_1^2).$$

Figure 5.13: Jean-Dominique Cassini, 1748-1845 (left) and Eugène Charles Catalan, 1814-1984 (right)

For the Fibonacci sequence, where $u_0 = 0$ and $u_1 = 1$, we get the *Cassini identity* (after Jean-Dominique Cassini, also known as Giovanni Domenico Cassini, 1625-1712),

$$F_{n+1}F_{n-1} - F_n^2 = (-1)^n, \qquad n \geq 1.$$

The above identity is a special case of *Catalan's identity*,

$$F_{n+r}F_{n-r} - F_n^2 = (-1)^{n-r+1}F_r^2, \qquad n \geq r,$$

due to Eugène Catalan (1814-1894).

For the Lucas numbers, where $u_0 = 2$ and $u_1 = 1$ we get

$$L_{n+1}L_{n-1} - L_n^2 = 5(-1)^{n-1}, \qquad n \geq 1.$$

In general, we have

$$u_k u_{n+1} + u_{k-1} u_n = u_1 u_{n+k} + u_0 u_{n+k-1},$$

for all $k \geq 1$ and all $n \geq 0$.

For the Fibonacci sequence, where $u_0 = 0$ and $u_1 = 1$, we just reproved the identity

$$F_{n+k} = F_k F_{n+1} + F_{k-1} F_n.$$

For the Lucas sequence, where $u_0 = 2$ and $u_1 = 1$, we get

$$\begin{aligned} L_k L_{n+1} + L_{k-1} L_n &= L_{n+k} + 2L_{n+k-1} \\ &= L_{n+k} + L_{n+k-1} + L_{n+k-1} \\ &= L_{n+k+1} + L_{n+k-1} \\ &= 5F_{n+k}, \end{aligned}$$

that is,

$$L_k L_{n+1} + L_{k-1} L_n = L_{n+k+1} + L_{n+k-1} = 5F_{n+k},$$

for all $k \geq 1$ and all $n \geq 0$.

The identity

$$F_{n+k} = F_k F_{n+1} + F_{k-1} F_n$$

plays a key role in the proof of various divisibility properties of the Fibonacci numbers. Here are two such properties:

Proposition 5.7.6 *The following properties hold:*

1. *F_n divides F_{mn}, for all $m, n \geq 1$.*
2. *$\gcd(F_m, F_n) = F_{\gcd(m,n)}$, for all $m, n \geq 1$.*

An interesting consequence of this divisibility property is that if F_n is a prime and $n > 4$, then n must be a prime.

However, there are prime numbers $n \geq 5$ such that F_n is not prime, for example, $n = 19$, as $F_{19} = 4181 = 37 \times 113$ is not prime.

The gcd identity can also be used to prove that for all m, n with $2 < n < m$, if F_n divides F_m, then n divides m, which provides a converse of our earlier divisibility property.

The formulae

$$\begin{aligned} 2F_{m+n} &= F_m L_n + F_n L_m \\ 2L_{m+n} &= L_m L_n + 5F_m F_n \end{aligned}$$

are also easily established using the explicit formulae for F_n and L_n in terms of φ and φ^{-1}.

The Fibonacci sequence and the Lucas sequence contain primes but it is unknown whether they contain infinitely many primes.

Here are some facts about Fibonacci and Lucas primes taken from *The Little Book of Bigger Primes*, by Paulo Ribenboim [12].

As we proved earlier, if F_n is a prime and $n \neq 4$, then n must be a prime but the converse is false.

For example,

$$F_3, F_4, F_5, F_7, F_{11}, F_{13}, F_{17}, F_{23}$$

are prime but $F_{19} = 4181 = 37 \times 113$ is not a prime.

One of the largest prime Fibonacci numbers if F_{81839}. It has 17103 digits.

Concerning the Lucas numbers, it can also be shown that if L_n is an odd prime and n is not a power of 2, then n is a prime.

Again, the converse is false. For example,

$$L_0, L_2, L_4, L_5, L_7, L_8, L_{11}, L_{13}, L_{16}, L_{17}, L_{19}, L_{31}$$

are prime but $L_{23} = 64079 = 139 \times 461$ is not a prime. Similarly, $L_{32} = 4870847 = 1087 \times 4481$ is not prime! One of the largest Lucas primes is L_{51169}.

Generally, divisibility properties of the Lucas numbers are not easy to prove because there is no simple formula for L_{m+n} in terms of other L_k's.

Nevertheless, we can prove that if $n, k \geq 1$ and k is odd, then L_n divides L_{kn}.

This is not necessarily true if k is even.

For example, $L_4 = 7$ and $L_8 = 47$ are prime.

It should also be noted that not every sequence, (u_n), given by the recurrence

$$u_{n+2} = u_{n+1} + u_n$$

and with $\gcd(u_0, u_1) = 1$ contains a prime number!

According to Ribenboim [12], Graham found an example in 1964 but it turned out to be incorrect. Later, Knuth gave correct sequences (see *Concrete Mathematics* [8], Chapter 6), one of which beginning with

$$\begin{aligned} u_0 &= 62638280004239857 \\ u_1 &= 49463435743205655. \end{aligned}$$

We just studied some properties of the sequences arising from the recurrence relation

$$u_{n+2} = u_{n+1} + u_n.$$

Lucas investigated the properties of the more general recurrence relation

$$u_{n+2} = Pu_{n+1} - Qu_n,$$

where $P, Q \in \mathbb{Z}$ are any integers with $P^2 - 4Q \neq 0$, in two seminal papers published in 1878.

We can prove some of the basic results about these Lucas sequences quite easily using the matrix method that we used before.

The recurrence relation

$$u_{n+2} = Pu_{n+1} - Qu_n$$

yields the recurrence

$$\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = \begin{pmatrix} P & -Q \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_n \\ u_{n-1} \end{pmatrix}$$

for all $n \geq 1$, and so,

$$\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = \begin{pmatrix} P & -Q \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} u_1 \\ u_0 \end{pmatrix}$$

for all $n \geq 0$.

The matrix

$$A = \begin{pmatrix} P & -Q \\ 1 & 0 \end{pmatrix}$$

has characteristic polynomial,
$-(P - \lambda)\lambda + Q = \lambda^2 - P\lambda + Q$, which has discriminant $D = P^2 - 4Q$.

If we assume that $P^2 - 4Q \neq 0$, the polynomial $\lambda^2 - P\lambda + Q$ has two distinct roots:

$$\alpha = \frac{P + \sqrt{D}}{2}, \qquad \beta = \frac{P - \sqrt{D}}{2}.$$

Obviously,

$$\begin{aligned} \alpha + \beta &= P \\ \alpha\beta &= Q \\ \alpha - \beta &= \sqrt{D}. \end{aligned}$$

The matrix A can be diagonalized as

$$A = \begin{pmatrix} P & -Q \\ 1 & 0 \end{pmatrix} = \frac{1}{\alpha - \beta} \begin{pmatrix} \alpha & \beta \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \begin{pmatrix} 1 & -\beta \\ -1 & \alpha \end{pmatrix}.$$

Thus, we get

$$\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = \frac{1}{\alpha - \beta} \begin{pmatrix} \alpha & \beta \\ 1 & 1 \end{pmatrix} \begin{pmatrix} (-\beta u_0 + u_1)\alpha^n \\ (\alpha u_0 - u_1)\beta^n \end{pmatrix}$$

and so,

$$u_n = \frac{1}{\alpha - \beta}\big((-\beta u_0 + u_1)\alpha^n + (\alpha u_0 - u_1)\beta^n\big).$$

Actually, the above formula holds for $n = 0$ only if $\alpha \neq 0$ and $\beta \neq 0$, that is, iff $Q \neq 0$.

If $Q = 0$, then either $\alpha = 0$ or $\beta = 0$, in which case the formula still holds if we assume that $0^0 = 1$.

For $u_0 = 0$ and $u_1 = 1$, we get a generalization of the Fibonacci numbers,

$$U_n = \frac{\alpha^n - \beta^n}{\alpha - \beta}$$

and for $u_0 = 2$ and $u_1 = P$, we get a generalization of the Lucas numbers,

$$V_n = \alpha^n + \beta^n.$$

The orginal Fibonacci and Lucas numbers correspond to $P = 1$ and $Q = -1$.

Since the vectors $\binom{0}{1}$ and $\binom{2}{P}$ are linearly independent, every sequence arising from the recurrence relation

$$u_{n+2} = Pu_{n+1} - Qu_n$$

is a unique linear combination of the sequences (U_n) and (V_n).

It possible to prove the following generalization of the Cassini identity:

Proposition 5.7.7 *The sequence defined by the recurrence*

$$u_{n+2} = Pu_{n+1} - Qu_n$$

(with $P^2 - 4Q \neq 0$*) satisfies the identity:*

$$u_{n+1}u_{n-1} - u_n^2 = Q^{n-1}(-Qu_0^2 + Pu_0u_1 - u_1^2).$$

For the U-sequence, $u_0 = 0$ and $u_1 = 1$, so we get

$$U_{n+1}U_{n-1} - U_n^2 = -Q^{n-1}.$$

For the V-sequence, $u_0 = 2$ and $u_1 = P$, so we get

$$V_{n+1}V_{n-1} - V_n^2 = Q^{n-1}D,$$

where $D = P^2 - 4Q$.

Since $\alpha^2 - Q = \alpha(\alpha - \beta)$ and $\beta^2 - Q = -\beta(\alpha - \beta)$, we easily get formulae expressing U_n in terms of the V's and vice-versa:

Proposition 5.7.8 *We have the following identities relating the U_n and the V_n;*

$$\begin{aligned} V_n &= U_{n+1} - QU_{n-1} \\ DU_n &= V_{n+1} - QV_{n-1}, \end{aligned}$$

for all $n \geq 1$.

Figure 5.14: Marin Mersenne, 1588-1648

The following identities are also easy to derive:

$$\begin{aligned}
U_{2n} &= U_n V_n \\
V_{2n} &= V_n^2 - 2Q^n \\
U_{m+n} &= U_m U_{n+1} - Q U_n U_{m-1} \\
V_{m+n} &= V_m V_n - Q^n V_{m-n}.
\end{aligned}$$

Lucas numbers play a crucial role in testing the primality of certain numbers of the form, $N = 2^p - 1$, called *Mersenne numbers*.

A Mersenne number which is prime is called a *Mersenne prime*.

First, if $N = 2^p - 1$ is prime, then p itself must be a prime.

For $p = 2, 3, 5, 7$ we see that $3 = 2^2 - 1$, $7 = 2^3 - 1$, $31 = 2^5 - 1$, $127 = 2^7 - 1$ are indeed prime.

However, the condition that the exponent, p, be prime is not sufficient for $N = 2^p - 1$ to be prime, since for $p = 11$, we have $2^{11} - 1 = 2047 = 23 \times 89$.

Mersenne (1588-1648) stated in 1644 that $N = 2^p - 1$ is prime when

$$p = 2, 3, 5, 7, 13, 17, 19, 31, 67, 127, 257.$$

Mersenne was wrong about $p = 67$ and $p = 257$, and he missed, $p = 61, 89, 107$.

Euler showed that $2^{31} - 1$ was indeed prime in 1772 and at that time, it was known that $2^p - 1$ is indeed prime for $p = 2, 3, 5, 7, 13, 17, 19, 31$.

Then came Lucas. In 1876, Lucas, proved that $2^{127} - 1$ was prime!

Lucas came up with a method for testing whether a Mersenne number is prime, later rigorously proved correct by Lehmer, and known as the *Lucas-Lehmer test*.

This test does not require the actual computation of $N = 2^p - 1$ but it requires an efficient method for squaring large numbers (less that N) and a way of computing the residue modulo $2^p - 1$ just using p.

Figure 5.15: Derrick Henry Lehmer, 1905-1991

A version of the Lucas-Lehmer test uses the Lucas sequence given by the recurrence

$$V_{n+2} = 2V_{n+1} + 2V_n,$$

starting from $V_0 = V_1 = 2$. This corresponds to $P = 2$ and $Q = -2$.

In this case, $D = 12$ and it is easy to see that $\alpha = 1+\sqrt{3}$, $\beta = 1 - \sqrt{3}$, so

$$V_n = (1+\sqrt{3})^n + (1-\sqrt{3})^n.$$

This sequence starts with

$$2, 2, 8, 20, 56, \cdots$$

Here is the first version of the Lucas-Lehmer test for primality of a Mersenne number:

Theorem 5.7.9 *Lucas-Lehmer test (Version 1) The number, $N = 2^p - 1$, is prime for any odd prime p iff N divides $V_{2^{p-1}}$.*

A proof of the Lucas-Lehmer test can be found in *The Little Book of Bigger Primes* [12]. Shorter proofs exist and are available on the Web but they require some knowledge of algebraic number theory.

The most accessible proof that we are aware of (it only uses the quadratic reciprocity law) is given in Volume 2 of Knuth [9], see Section 4.5.4.

Note that the test does not apply to $p = 2$ because $3 = 2^2 - 1$ does not divide $V_2 = 8$ but that's not a problem.

The numbers $V_{2^{p-1}}$ get large very quickly but if we observe that

$$V_{2n} = V_n^2 - 2(-2)^n,$$

we may want to consider the sequence, S_n, given by

$$S_{n+1} = S_n^2 - 2,$$

starting with $S_0 = 4$.

This sequence starts with

$$4, 14, 194, 37643, 1416317954, \cdots$$

Then, it turns out that

$$V_{2^k} = S_{k-1}2^{2^{k-1}},$$

for all $k \geq 1$. It is also easy to see that

$$S_k = (2+\sqrt{3})^{2^k} + (2-\sqrt{3})^{2^k}.$$

Now, $N = 2^p - 1$ is prime iff N divides $V_{2^{p-1}}$ iff $N = 2^p - 1$ divides $S_{p-2}2^{2^{p-2}}$ iff N divides S_{p-2} (since if N divides $2^{2^{p-2}}$, then N is not prime).

Thus, we obtain an improved version of the Lucas-Lehmer test for primality of a Mersenne number:

Theorem 5.7.10 *Lucas-Lehmer test (Version 2) The number, $N = 2^p - 1$, is prime for any odd prime p iff*

$$S_{p-2} \equiv 0 \pmod{N}.$$

The test does not apply to $p = 2$ because $3 = 2^2 - 1$ does not divide $S_0 = 4$ but that's not a problem.

The above test can be performed by computing a sequence of residues mod N, using the recurrence
$S_{n+1} = S_n^2 - 2$, starting from 4.

As of January 2009, only 46 Mersenne primes are known. The largest one was found in August 2008 by mathematicians at UCLA. This is

$$M_{46} = 2^{43112609} - 1,$$

and it has $12,978,189$ digits!

It is an open problem whether there are infinitely many Mersenne primes.

Going back to the second version of the Lucas-Lehmer test, since we are computing the sequence of S_k's modulo N, the squares being computed never exceed $N^2 = 2^{2p}$.

There is also a clever way of computing $n \bmod 2^p - 1$ without actually performing divisions if we express n in binary.

This is because

$$n \equiv (n \mod 2^p) + \lfloor n/2^p \rfloor \pmod{2^p - 1}.$$

But now, if n is expressed in binary, $(n \mod 2^p)$ consists of the p rightmost (least significant) bits of n and $\lfloor n/2^p \rfloor$ consists of the bits remaining as the head of the string obtained by deleting the rightmost p bits of n.

Thus, we can compute the remainder modulo $2^p - 1$ by repeating this process until at most p bits remain.

Observe that if n is a multiple of $2^p - 1$, the algorithm will produce $2^p - 1$ in binary as opposed to 0 but this exception can be handled easily.

For example

$$\begin{aligned}
916 \mod 2^5 - 1 &= 1110010100_2 \pmod{2^5 - 1} \\
&= 10100_2 + 11100_2 \pmod{2^5 - 1} \\
&= 110000_2 \pmod{2^5 - 1} \\
&= 10000_2 + 1_2 \pmod{2^5 - 1} \\
&= 10001_2 \pmod{2^5 - 1} \\
&= 10001_2 \\
&= 17.
\end{aligned}$$

The Lucas-Lehmer test applied to $N = 127 = 2^7 - 1$ yields the following steps, if we denote $S_k \bmod 2^p - 1$ by r_k:

$r_0 = 4,$

$r_1 = 4^2 - 2 = 14 \pmod{127}$, *i.e.* $r_1 = 14$

$r_2 = 14^2 - 2 = 194 \pmod{127}$, *i.e.* $r_2 = 67$

$r_3 = 67^2 - 2 = 4487 \pmod{127}$, *i.e.* $r_3 = 42$

$r_4 = 42^2 - 2 = 1762 \pmod{127}$, *i.e.* $r_4 = 111$

$r_5 = 111^2 - 2 = 12319 \pmod{127}$, *i.e.* $r_5 = 0.$

As $r_5 = 0$, the Lucas-Lehmer test confirms that
$N = 127 = 2^7 - 1$ is indeed prime.

5.8 Distributive Lattices, Boolean Algebras, Heyting Algebras

If we go back to one of our favorite examples of a lattice, namely, the power set, 2^X, of some set, X, we observe that it is more than a lattice.

For example, if we look at Figure 5.6, we can check that the two identities D1 and D2 stated in the next definition hold.

Definition 5.8.1 We say that a lattice, X, is a *distributive lattice* if (D1) and (D2) hold:

$$\begin{array}{ll} D1 & a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c) \\ D2 & a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c). \end{array}$$

Remark: Not every lattice is distributive but many lattices of interest are distributive.

It is a bit surprising that in a lattice, (D1) and (D2) are actually equivalent.

The reader should prove that every totally ordered poset is a distributive lattice.

The lattice $\mathbb{N}_+$ under the divisibility ordering also turns out to be a distributive lattice.

Another useful fact about distributivity is that in any lattice

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

This is because in any lattice, $a \wedge (b \vee c) \geq a \wedge b$ and $a \wedge (b \vee c) \geq a \wedge c$.

Therefore, in order to establish distributivity in a lattice it suffices to show that

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

Another important property of distributive lattices is the following:

Proposition 5.8.2 *In a distributive lattice, X, if $z \wedge x = z \wedge y$ and $z \vee x = z \vee y$, then $x = y$ (for all $x, y, z \in X$).*

The power set lattice has yet some additional properties having to do with complementation.

First, the power lattice 2^X has a least element $0 = \emptyset$ and a greatest element, $1 = X$.

If a lattice, X, has a least element, 0, and a greatest element, 1, the following properties are clear: For all $a \in X$, we have

$$\begin{aligned} a \wedge 0 &= 0 \qquad & a \vee 0 &= a \\ a \wedge 1 &= a \qquad & a \vee 1 &= 1. \end{aligned}$$

Figure 5.16: Augustus de Morgan, 1806-1871

More importantly, for any subset, $A \subseteq X$, we have the complement, $\overline{A}$, of A in X, which satisfies the identities:

$$A \cup \overline{A} = X, \qquad A \cap \overline{A} = \emptyset.$$

Moreover, we know that the de Morgan identities hold. The generalization of these properties leads to what is called a complemented lattice.

Definition 5.8.3 Let X be a lattice and assume that X has a least element, 0, and a greatest element, 1 (we say that X is a *bounded lattice*). For any $a \in X$, a *complement of a* is any element, $b \in X$, so that

$$a \vee b = 1 \quad \text{and} \quad a \wedge b = 0.$$

If every element of X has a complement, we say that X is a *complemented lattice*.

Remarks:

1. When $0 = 1$, the lattice X collapses to the degenerate lattice consisting of a single element. As this lattice is of little interest, from now on, we will always assume that $0 \neq 1$.
2. In a complemented lattice, complements are generally not unique. However, as the next proposition shows, this is the case for distributive lattices.

Proposition 5.8.4 *Let X be a lattice with least element 0 and greatest element 1. If X is distributive, then complements are unique if they exist. Moreover, if b is the complement of a, then a is the complement of b.*

In view of Proposition 5.8.4, if X is a complemented distributive lattice, we denote the complement of any element, $a \in X$, by $\overline{a}$.

We have the identities

$$\begin{aligned} a \vee \overline{a} &= 1 \\ a \wedge \overline{a} &= 0 \\ \overline{\overline{a}} &= a. \end{aligned}$$

We also have the following proposition about the de Morgan laws.

Proposition 5.8.5 *Let X be a lattice with least element 0 and greatest element 1. If X is distributive and complemented, then the de Morgan laws hold:*

$$\begin{aligned} \overline{a \vee b} &= \overline{a} \wedge \overline{b} \\ \overline{a \wedge b} &= \overline{a} \vee \overline{b}. \end{aligned}$$

All this leads to the definition of a boolean lattice.

Definition 5.8.6 A *Boolean lattice* is a lattice with a least element, 0, a greatest element, 1, and which is distributive and complemented.

Of course, every power set is a boolean lattice, but there are boolean lattices that are not power sets.

Putting together what we have done, we see that a boolean lattice is a set, X, with two special elements, 0, 1, and three operations, $\wedge$, $\vee$ and $a \mapsto \overline{a}$ satisfying the axioms stated in

Proposition 5.8.7 *If X is a boolean lattice, then the following equations hold for all $a, b, c \in X$:*

$$\begin{array}{lll}
L1 & a \vee b = b \vee a, & a \wedge b = b \wedge a \\
L2 & (a \vee b) \vee c = a \vee (b \vee c), & \\
 & (a \wedge b) \wedge c = a \wedge (b \wedge c) & \\
L3 & a \vee a = a, & a \wedge a = a \\
L4 & (a \vee b) \wedge a = a, & (a \wedge b) \vee a = a \\
D1\text{-}D2 & a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c), & \\
 & a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c) & \\
LE & a \vee 0 = a, & a \wedge 0 = 0 \\
GE & a \vee 1 = 1, & a \wedge 1 = a \\
C & a \vee \overline{a} = 1, & a \wedge \overline{a} = 0 \\
I & \overline{\overline{a}} = a & \\
dM & \overline{a \vee b} = \overline{a} \wedge \overline{b}, & \overline{a \wedge b} = \overline{a} \vee \overline{b}.
\end{array}$$

Conversely, if X is a set together with two special elements, 0, 1, *and three operations,* $\wedge$, $\vee$ *and* $a \mapsto \overline{a}$ *satisfying the axioms above, then it is a boolean lattice under the ordering given by* $a \leq b$ *iff* $a \vee b = b$*.*

In view of Proposition 5.8.7, we make the definition:

Definition 5.8.8 A set, X, together with two special elements, 0, 1, and three operations, $\wedge$, $\vee$ and $a \mapsto \overline{a}$ satisfying the axioms of Proposition 5.8.7 is called a *Boolean algebra*.

Proposition 5.8.7 shows that the notions of a Boolean lattice and of a Boolean algebra are equivalent. The first one is order-theoretic and the second one is algebraic.

Remarks:

1. As the name indicates, Boolean algebras were invented by G. Boole (1854). One of the first comprehensive accounts is due to E. Schröder (1890-1895).

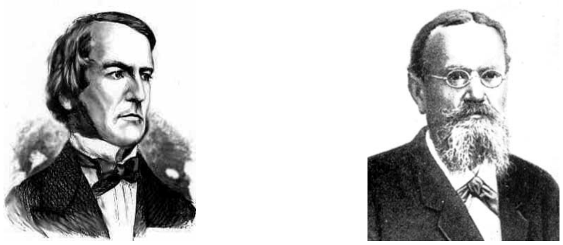

Figure 5.17: George Boole, 1815-1864 (left) and Ernst Schröder 1841-1902 (right)

2. The axioms for Boolean algebras given in Proposition 5.8.7 are not independent. There is a set of independent axioms known as the *Huntington axioms* (1933).

Let p be any integer with $p \geq 2$. Under the division ordering, it turns out that the set, $\mathrm{Div}(p)$, of divisors of p is a distributive lattice.

In general not every integer, $k \in \mathrm{Div}(p)$, has a complement but when it does, $\overline{k} = p/k$.

It can be shown that $\mathrm{Div}(p)$ is a Boolean algebra iff p is not divisible by any square integer (an integer of the form m^2, with $m > 1$).

Classical logic is also a rich source of Boolean algebras.

Indeed, it is easy to show that logical equivalence is an equivalence relation and, as Homework problems, you have shown (with great pain) that all the axioms of Proposition 5.8.7 are provable equivalences (where $\vee$ is disjunction, $\wedge$ is conjunction, $\overline{P} = \neg P$, *i.e.*, negation, $0 = \perp$ and $1 = \top$).

Furthermore, again, as Homework problems you have shown that logical equivalence is compatible with $\vee, \wedge, \neg$ in the following sense: If $P_1 \equiv Q_1$ and $P_2 \equiv Q_2$, then

$$\begin{aligned} (P_1 \vee P_2) &\equiv (Q_1 \vee Q_2) \\ (P_1 \wedge P_2) &\equiv (Q_1 \wedge Q_2) \\ \neg P_1 &\equiv \neg Q_1. \end{aligned}$$

Consequently, for any set, T, of propositions we can define the relation, $\equiv_T$, by

$$P \equiv_T Q \quad \text{iff} \quad T \vdash P \equiv Q,$$

i.e., iff $P \equiv Q$ is provable from T.

Clearly, $\equiv_T$ is an equivalence relation on propositions and so, we can define the operations $\vee$, $\wedge$ and $-$ on the set of equivalence classes, $\mathbf{B}_T$, of propositions as follows:

$$\begin{aligned}
[P] \vee [Q] &= [P \vee Q] \\
[P] \wedge [Q] &= [P \wedge Q] \\
\overline{[P]} &= [\neg P].
\end{aligned}$$

We also let $0 = [\perp]$ and $1 = [\top]$. Then, we get the Boolean algebra, $\mathbf{B}_T$, called the *Lindenbaum algebra* of T.

It also turns out that Boolean algebras are just what's needed to give truth-value semantics to classical logic.

Let B be any Boolean algebra. A *truth assignment* is any function, v, from the set $\mathbf{PS} = \{\mathbf{P}_1, \mathbf{P}_2, \cdots\}$ of propositional symbols to B.

Then, we can evaluate recursively the truth value, $P_B[v]$, in B of any proposition, P, with respect to the truth assignment, v, as follows:

$$\begin{aligned}
(\mathbf{P}_i)_B[v] &= v(P) \\
\perp_B [v] &= 0 \\
\top_B[v] &= 1 \\
(P \vee Q)_B[v] &= P_B[v] \vee P_B[v] \\
(P \wedge Q)_B[v] &= P_B[v] \wedge P_B[v] \\
(\neg P)_B[v] &= \overline{P[v]_B}.
\end{aligned}$$

In the equations above, on the right hand side, $\vee$ and $\wedge$ are the lattice operations of the Boolean algebra, B.

We say that a proposition, P, is *valid in the Boolean algebra B (or B-valid)* if $P_B[v] = 1$ for all truth assignments, v.

We say that P is *(classially) valid* if P is B-valid in all Boolean algebras, B. It can be shown that every provable proposition is valid. This property is called *soundness*.

Conversely, if P is valid, then it is provable. This second property is called *completeness*.

Actually completeness holds in a much stronger sense: If a proposition is valid in the two element Boolean algebra, $\{0, 1\}$, then it is provable!

Figure 5.18: Arend Heyting, 1898-1980

One might wonder if there are certain kinds of algebras similar to Boolean algebras well suited for intuitionistic logic. The answer is yes: Such algebras are called *Heyting algebras*.

In our study of intuitionistic logic, we learned that negation is not a primary connective but instead it is defined in terms of implication by $\neg P = P \Rightarrow \perp$.

This suggests adding to the two lattice operations $\vee$ and $\wedge$ a new operation, $\rightarrow$, that will behave like $\Rightarrow$.

The trick is, what kind of axioms should we require on $\rightarrow$ to "capture" the properties of intuitionistic logic?

Now, if X is a lattice with 0 and 1, given any two elements, $a, b \in X$, experience shows that $a \rightarrow b$ should be the largest element, c, such that $c \wedge a \leq b$. This leads to

Definition 5.8.9 A lattice, X, with 0 and 1 is a *Heyting lattice* iff it has a third binary operation, →, such that

$$c \wedge a \leq b \quad \text{iff} \quad c \leq (a \rightarrow b)$$

for all $a, b, c \in X$. We define the *negation (or pseudo-complement) of a* as $\overline{a} = (a \rightarrow 0)$.

At first glance, it is not clear that a Heyting lattice is distributive but in fact, it is.

The following proposition (stated without proof) gives an algebraic characterization of Heyting lattices which is useful to prove various properties of Heyting lattices.

Proposition 5.8.10 *Let X be a lattice with 0 and 1 and with a binary operation, $\rightarrow$. Then, X is a Heyting lattice iff the following equations hold for all $a, b, c \in X$:*

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

A lattice with 0 and 1 and with a binary operation, $\rightarrow$, satisfying the equations of Proposition 5.8.10 is called a *Heyting algebra*.

So, we see that Proposition 5.8.10 shows that the notions of Heyting lattice and Heyting algebra are equivalent (this is analogous to Boolean lattices and Boolean algebras).

The reader will notice that these axioms are propositions that were shown to be provable intuitionistically in Homework Problems!

The following theorem shows that every Heyting algebra is distributive, as we claimed earlier.

This theorem also shows "how close" to a Boolean algebra a Heyting algebra is.

Theorem 5.8.11 *(a) Every Heyting algebra is distributive.*

(b) A Heyting algebra, X, is a boolean algebra iff $\overline{\overline{a}} = a$ for all $a \in X$.

Remarks:

1. Heyting algebras were invented by A. Heyting in 1930. Heyting algebras are sometimes known as "Brouwerian lattices".

2. Every Boolean algebra is automatically a Heyting algebra: Set $a \rightarrow b = \overline{a} \vee b$.

3. It can be shown that every finite distributive lattice is a Heyting algebra.

We conclude this brief exposition of Heyting algebras by explaining how they provide a truth semantics for intuitionistic logic analogous to the thuth semantics that Boolean algebras provide for classical logic.

As in the classical case, it is easy to show that intuitionistic logical equivalence is an equivalence relation and you have shown (with great pain) that all the axioms of Heyting algebras are intuitionistically provable equivalences (where $\vee$ is disjunction, $\wedge$ is conjunction, and $\rightarrow$ is $\Rightarrow$).

Furthermore, you have also shown that intuitionistic logical equivalence is compatible with $\vee, \wedge, \Rightarrow$ in the following sense: If $P_1 \equiv Q_1$ and $P_2 \equiv Q_2$, then

$$\begin{aligned}
(P_1 \vee P_2) &\equiv (Q_1 \vee Q_2) \\
(P_1 \wedge P_2) &\equiv (Q_1 \wedge Q_2) \\
(P_1 \Rightarrow P_2) &\equiv (Q_1 \Rightarrow Q_2).
\end{aligned}$$

Consequently, for any set, T, of propositions we can define the relation, $\equiv_T$, by

$$P \equiv_T Q \quad \text{iff} \quad T \vdash P \equiv Q,$$

$i.e.$, iff $P \equiv Q$ is provable intuitionistically from T.

Clearly, $\equiv_T$ is an equivalence relation on propositions and we can define the operations $\vee, \wedge$ and $\rightarrow$ on the set of equivalence classes, $\mathbf{H}_T$, of propositions as follows:

$$\begin{aligned}[P] \vee [Q] &= [P \vee Q] \\ [P] \wedge [Q] &= [P \wedge Q] \\ [P] \rightarrow [Q] &= [P \Rightarrow Q].\end{aligned}$$

We also let $0 = [\perp]$ and $1 = [\top]$. Then, we get the Heyting algebra, $\mathbf{H}_T$, called the *Lindenbaum algebra* of T, as in the classical case.

Now, let H be any Heyting algebra. By analogy with the case of Boolean algebras, a *truth assignment* is any function, v, from the set $\mathbf{PS} = \{\mathbf{P}_1, \mathbf{P}_2, \cdots\}$ of propositional symbols to H.

Then, we can evaluate recursively the truth value, $P_H[v]$, in H of any proposition, P, with respect to the truth assignment, v, as follows:

$$\begin{aligned}
(\mathbf{P}_i)_H[v] &= v(P) \\
\perp_H\ [v] &= 0 \\
\top_H[v] &= 1 \\
(P \vee Q)_H[v] &= P_H[v] \vee P_H[v] \\
(P \wedge Q)_H[v] &= P_H[v] \wedge P_H[v] \\
(P \Rightarrow Q)_H[v] &= (P_H[v] \rightarrow P_H[v]) \\
(\neg P)_H[v] &= (P_H[v] \rightarrow 0).
\end{aligned}$$

In the equations above, on the right hand side, $\vee$, $\wedge$ and $\rightarrow$ are the operations of the Heyting algebra, H.

We say that a proposition, P, is *valid in the Heyting algebra H (or H-valid)* if $P_H[v] = 1$ for all truth assignments, v.

We say that P is *HA-valid (or intuitionistically valid)* if P is H-valid in all Heyting algebras, H.

As in the classical case, it can be shown that every intuitionistically provable proposition is HA-valid. This property is called *soundness*.

Conversely, if P is HA-valid, then it is intuitionistically provable. This second property is called *completeness*.

A stronger completeness result actually holds: If a proposition is H-valid in all *finite* Heyting algebras, H, then it is intuitionistically provable.

As a consequence, if a proposition is *not* provable intuitionistically, then it can be falsified in some finite Heyting algebra.

Remark: If X is any set, a *topology on X* is a family, $\mathcal{O}$, of subsets of X satisfying the following conditions:

(1) $\emptyset \in \mathcal{O}$ and $X \in \mathcal{O}$;

(2) For every family (even infinite), $(U_i)_{i\in I}$, of sets $U_i \in \mathcal{O}$, we have $\bigcup_{i\in I} U_i \in \mathcal{O}$.

(3) For every *finite* family, $(U_i)_{1\leq i\leq n}$, of sets $U_i \in \mathcal{O}$, we have $\bigcap_{1\leq i\leq n} U_i \in \mathcal{O}$.

Every subset in $\mathcal{O}$ is called an *open subset* of X (in the topology $\mathcal{O}$) .

The pair, $\langle X, \mathcal{O}\rangle$, is called a *topological space*.

Given any subset, A, of X, the union of all open subsets contained in A is the largest open subset of A and is denoted $\overset{\circ}{A}$.

Given a topological space, $\langle X, \mathcal{O} \rangle$, we claim that $\mathcal{O}$ with the inclusion ordering is a Heyting algebra with $0 = \emptyset$; $1 = X$; $\vee = \cup$ (union); $\wedge = \cap$ (intersection); and with

$$(U \rightarrow V) = \overbrace{(X - U) \cup V}^{\circ}.$$

(Here, $X - U$ is the complement of U in X.)

In this Heyting algebra, we have

$$\overline{U} = \overbrace{X - U}^{\circ}.$$

Since $X - U$ is usually not open, we generally have $\overline{\overline{U}} \neq U$.

Therefore, we see that topology yields another supply of Heyting algebras.

Bibliography

[1] Claude Berge. *Principles of Combinatorics*. Academic Press, first edition, 1971.

[2] J. Cameron, Peter. *Combinatorics: Topics, Techniques, Algorithms*. Cambridge University Press, first edition, 1994.

[3] John H. Conway and K. Guy, Richard. *The Book of Numbers*. Copernicus, Springer-Verlag, first edition, 1996.

[4] Herbert B. Enderton. *Elements of Set Theory*. Academic Press, first edition, 1977.

[5] Jean Gallier. Constructive Logics. Part I: A Tutorial on Proof Systems and Typed λ-Calculi. *Theoretical Computer Science*, 110(2):249–339, 1993.

[6] Jean H. Gallier. *Logic for Computer Science*. Harper and Row, New York, 1986.

[7] Timothy Gowers. *Mathematics: A very Short Introduction*. Oxford University Press, first edition, 2002.

[8] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete Mathematics: A Foundation For Computer Science*. Addison Wesley, second edition, 1994.

[9] Donald E. Knuth. *The Art of Computer Programming, Volume 2: Seminumerical Algorithms*. Addison Wesley, third edition, 1997.

[10] L. Lovász, J. Pelikán, and K. Vesztergombi. *Discrete Mathematics. Elementary and Beyond*. Undergraduate Texts in Mathematics. Springer, first edition, 2003.

[11] Jiri Matousek. *Lectures on Discrete Geometry*. GTM No. 212. Springer Verlag, first edition, 2002.

[12] Paulo Ribenboim. *The Little Book of Bigger Primes*. Springer-Verlag, second edition, 2004.

[13] Joseph H. Silverman. *A Friendly Introduction to Number Theory*. Prentice Hall, first edition, 1997.

[14] Richard P. Stanley. *Enumerative Combinatorics, Vol. I*. Cambridge Studies in Advanced Mathemat-

ics, No. 49. Cambridge University Press, first edition, 1997.

[15] D. van Dalen. *Logic and Structure.* Universitext. Springer Verlag, second edition, 1980.

[16] J.H. van Lint and R.M. Wilson. *A Course in Combinatorics.* Cambridge University Press, second edition, 2001.