Posets

This is an abbreviated version of the Combinatorics Study Group notes by Thomas Britz and Peter Cameron.

1 What is a poset?

A *binary relation* R on a set X is a set of ordered pairs of elements of X, that is, a subset of $X \times X$. We can represent R by a matrix with rows and columns indexed by X, with (x,y) entry 1 if $(x,y) \in R$, 0 otherwise.

The term "poset" is short for "partially ordered set", that is, a set whose elements are ordered but not all pairs of elements are required to be comparable in the order. Just as an order in the usual sense may be strict (as $<$) or non-strict (as $\leq$), there are two versions of the definition of a partial order:

A *strict partial order* is a binary relation S on a set X satisfying the conditions

(R$-$) for no $x \in X$ does $(x,x) \in S$ hold;

(A$-$) if $(x,y) \in S$, then $(y,x) \notin S$;

(T) if $(x,y) \in S$ and $(y,z) \in S$, then $(x,z) \in S$.

A *non-strict partial order* is a binary relation R on a set X satisfying the conditions

(R$+$) for all $x \in X$ we have $(x,x) \in R$;

(A) if $(x,y) \in R$ and $(y,x) \in R$ then $x = y$;

(T) if $(x,y) \in R$ and $(y,z) \in R$ then $(x,z) \in R$.

Condition (A$-$) appears stronger than (A), but in fact (R$-$) and (A) imply (A$-$). So we can (as is usually done) replace (A$-$) by (A) in the definition of a strict partial order. Conditions (R$-$), (R$+$), (A), (T) are called *irreflexivity*, *reflexivity*, *antisymmetry* and *transitivity* respectively. We often write $x < y$ if $(x,y) \in S$, and $x \leq y$ if $(x,y) \in R$. We usually prefer the non-strict version.

The two definitions are essentially the same: we get from one to the other in the obvious way, setting $x \leq y$ if $x < y$ or $x = y$, and setting $x < y$ if $x \leq y$ but $x \neq y$. Thus, a *poset* is a set X carrying a partial order (either strict or non-strict).

If there is ambiguity about R, we simply write $x \leq_R y$.

A *total order* is a partial order in which every pair of elements is comparable, that is, the following condition (known as *trichotomy*) holds:

- for all $x, y \in X$, exactly one of $x <_R y$, $x = y$, and $y <_R x$ holds.

In a poset (X,R), we define the *interval* $[x,y]_R$ to be the set

$$[x,y]_R = \{z \in X : x \leq_R z \leq_R y\}.$$

By transitivity, the interval $[x,y]_R$ is empty if $x \not\leq_R y$. We say that the poset is *locally finite* if all intervals are finite.

The set of positive integers ordered by divisibility (that is, $x \leq_R y$ if x divides y) is a locally finite poset.

2 Properties of posets

An element x of a poset (X,R) is called *maximal* if there is no element $y \in X$ satisfying $x <_R y$. Dually, x is *minimal* if no element satisfies $y <_R x$.

In a general poset there may be no maximal element, or there may be more than one. But in a finite poset there is always at least one maximal element, which can be found as follows: choose any element x; if it is not maximal, replace it by an element y satisfying $x <_R y$; repeat until a maximal element is found. The process must terminate, since by the irreflexive and transitive laws the chain can never revisit any element. Dually, a finite poset must contain minimal elements.

An element x is an *upper bound* for a subset Y of X if $y \leq_R x$ for all $y \in Y$. *Lower bounds* are defined similarly. We say that x is a *least upper bound* or *l.u.b.* of Y if it is an upper bound and satisfies $x \leq_R x'$ for any upper bound x'. The concept of a *greatest lower bound* or *g.l.b.* is defined similarly.

A *chain* in a poset (X,R) is a subset C of X which is totally ordered by the restriction of R (that is, a totally ordered subset of X). An *antichain* is a set A of pairwise incomparable elements.

Infinite posets (such as $\mathbb{Z}$), as we remarked, need not contain maximal elements. *Zorn's Lemma* gives a sufficient condition for maximal elements to exist:

> Let (X,R) be a poset in which every chain has an upper bound. Then X contains a maximal element.

As well known, there is no "proof" of Zorn's Lemma, since it is equivalent to the Axiom of Choice (and so there are models of set theory in which it is true, and models in which it is false). Our proof of the existence of maximal elements in finite posets indicates why this should be so: the construction requires (in general infinitely many) choices of upper bounds for the elements previously chosen (which form a chain by construction).

The *height* of a poset is the largest cardinality of a chain, and its *width* is the largest cardinality of an antichain. We denote the height and width of (X,R) by $h(X)$ and $w(X)$ respectively (suppressing as usual the relation R in the notation).

In a finite poset (X,R), a chain C and an antichain A have at most one element in common. Hence the least number of antichains whose union is X is not less than the size $h(X)$ of the largest chain in X. In fact there is a partition of X into $h(X)$ antichains. To see this, let A_1 be the set of maximal elements; by definition this is an antichain, and it meets every maximal chain. Then let A_2 be the set of maximal elements in $X \setminus A_1$, and iterate this procedure to find the other antichains.

There is a kind of dual statement, harder to prove, known as *Dilworth's Theorem*:

Theorem 1 *Let (X,R) be a finite poset. Then there is a partition of X into $w(X)$ chains.*

An *up-set* in a poset (X,R) is a subset Y of X such that, if $y \in Y$ and $y \leq_R z$, then $z \in Y$. The set of minimal elements in an up-set is an antichain. Conversely, if A is an antichain, then

$$\uparrow(A) = \{x \in X : a \leq_R x \text{ for some } a \in A\}$$

is an up-set. These two correspondences between up-sets and antichains are mutually inverse; so the numbers of up-sets and antichains in a poset are equal.

Down-sets are, of course, defined dually. The complement of an up-set is a down-set; so there are equally many up-sets and down-sets.

3 Hasse diagrams

Let x and y be distinct elements of a poset (X,R). We say that y *covers* x if $[x,y]_R = \{x,y\}$; that is, $x <_R y$ but no element z satisfies $x <_R z <_R y$. In general, there may be no pairs x and y such that y covers x (this is the case in the rational numbers, for example). However, locally finite posets are determined by their covering pairs:

Proposition 2 *Let (X,R) be a locally finite poset, and $x,y \in X$. Then $x \leq_R y$ if and only if there exist elements $z_0,\ldots,z_n$ (for some non-negative integer n) such that $z_0 = x$, $z_n = y$, and z_{i+1} covers z_i for $i = 0,\ldots,n-1$.*

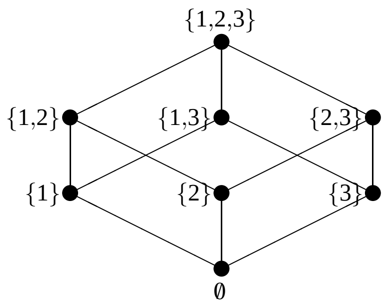


Figure 1: A Hasse diagram

The *Hasse diagram* of a poset (X,R) is the directed graph whose vertex set is X and whose arcs are the covering pairs (x,y) in the poset. We usually draw the Hasse diagram of a finite poset in the plane in such a way that, if y covers x, then the point representing y is higher than the point representing x. Then no arrows are required in the drawing, since the directions of the arrows are implicit.

For example, the Hasse diagram of the poset of subsets of $\{1,2,3\}$ is shown in Figure 1.

4 Linear extensions and dimension

One view of a partial order is that it contains partial information about a total order on the underlying set. This view is borne out by the following theorem. We say that one relation *extends* another if the second relation (as a set of ordered pairs) is a subset of the first.

Theorem 3 *Any partial order on a finite set X can be extended to a total order on X.*

This theorem follows by a finite number of applications of the next result.

Proposition 4 *Let R be a partial order on a set X, and let a,b be incomparable elements of X. Then there is a partial order R' extending R such that $(a,b) \in R'$ (that is, $a < b$ in the order R').*

A total order extending R in this sense is referred to as a *linear extension* of R. (The term “linear order” is an alternative for “total order”.)

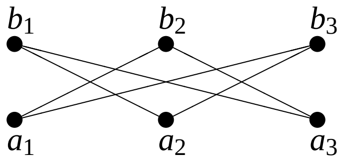


Figure 2: A crown

This proof does not immediately show that every infinite partial order can be extended to a total order. If we assume Zorn's Lemma, the conclusion follows. It cannot be proved from the Zermelo–Fraenkel axioms alone (assuming their consistency), but it is strictly weaker than the Axiom of Choice, that is, the Axiom of Choice (or Zorn's Lemma) cannot be proved from the Zermelo–Fraenkel axioms and this assumption. In other words, assuming the axioms consistent, there is a model in which Theorem 3 is false for some infinite poset, and another model in which Theorem 3 is true for all posets but Zorn's Lemma is false.

The theorem gives us another measure of the size of a partially ordered set. To motivate this, we use another model of a partial order. Suppose that a number of products are being compared using several different attributes. We regard object a as below object b if b beats a on every attribute. If each beats the other on some attributes, we regard the objects as being incomparable. This defines a partial order (assuming that each attribute gives a total order). More precisely, given a set S of total orders on X, we define a partial order R on X by $x <_R y$ if and only if $x <_s y$ for every $s \in S$. In other words, R is the intersection of the total orders in S.

Theorem 5 *Every partial order on a finite set X is the intersection of some set of total orders on X.*

Now we define the *dimension* of a partial order R to be the smallest number of total orders whose intersection is R. In our motivating example, it is the smallest number of attributes which could give rise to the observed total order R.

The *crown* on $2n$ elements $a_1,\ldots,a_n,b_1,\ldots,b_n$ is the partial order defined as follows: for all indices $i \neq j$, the elements a_i and a_j are incomparable, the elements b_i and b_j are incomparable, but $a_i < b_j$; and for each i, the elements a_i and b_i are incomparable. Figure 2 shows the Hasse diagram of the 6-element crown.

Now we have the following result:

Proposition 6 *The crown on $2n$ elements has dimension n.*

5 The Möbius function

Let R be a partial order on the finite set X. We take any linear order extending R, and write $X = \{x_1, \ldots, x_n\}$, where $x_1 < \ldots < x_n$ (in the linear order S): this is not essential but is convenient later.

The *incidence algebra* $\mathcal{A}(R)$ of R is the set of all functions $f : X \times X \to \mathbb{R}$ which satisfy $f(x,y) = 0$ unless $x \leq_R y$ holds. We could regard it as a function on R, regarded as a set of ordered pairs. Addition and scalar multiplication are defined pointwise; multiplication is given by the rule

$$(fg)(x,y) = \sum_z f(x,z)g(z,y).$$

If we represent f by the $n \times n$ matrix A_f with (i,j) entry $f(x_i,x_j)$, then this is precisely the rule for matrix multiplication. Also, if $x \not\leq_R y$, then there is no point z such that $x \leq_R z$ and $z \leq_R y$, and so $(fg)(x,y) = 0$. Thus, $\mathcal{A}(R)$ is closed under multiplication and does indeed form an algebra, a subset of the matrix algebra $M_n(\mathbb{R})$. Also, since f and g vanish on pairs not in R, the sum can be restricted to the interval $[x,y]_R = \{z : x \leq_R z \leq_R y\}$:

$$(fg)(x,y) = \sum_{z \in [x,y]_R} f(x,z)g(z,y).$$

Incidentally, we see that the (i,j) entry of A_f is zero if $i > j$, and so $\mathcal{A}(R)$ consists of upper triangular matrices. Thus, an element $f \in \mathcal{A}(R)$ is invertible if and only if $f(x,x) \neq 0$ for all $x \in X$.

The *zeta-function* ζ_R is the matrix representing the relation R as defined earlier; that is, the element of $\mathcal{A}(R)$ defined by

$$\zeta_R(x,y) = \begin{cases} 1 & \text{if } x \leq_R y, \\ 0 & \text{otherwise.} \end{cases}$$

Its inverse (which also lies in $\mathcal{A}(R)$) is the *Möbius function* μ_R of R. Thus, we have, for all $(x,y) \in R$,

$$\sum_{z \in [x,y]_R} \mu(x,z) = \begin{cases} 1 & \text{if } x = y, \\ 0 & \text{otherwise.} \end{cases}$$

This relation allows the Möbius function of a poset to be calculated recursively. We begin with $\mu_R(x,x) = 1$ for all $x \in X$. Now, if $x <_R y$ and we know the values of $\mu(x,z)$ for all $z \in [x,y]_R \setminus \{y\}$, then we have

$$\mu_R(x,y) = -\sum_{z \in [x,y]_R \setminus \{y\}} \mu_R(x,z).$$

In particular, $\mu_R(x,y) = -1$ if y covers x.

The definition of the incidence algebra and the Möbius function extend immediately to locally finite posets, since the sums involved are over intervals $[x,y]_R$.

The following are examples of Möbius functions.

- The subsets of a set:
  $\mu(A,B) = (-1)^{|B\setminus A|}$ for $A \subseteq B$;

- The subspaces of a vector space $V \subseteq GF(q)^n$:
  $\mu(U,W) = (-1)^k q^{\binom{k}{2}}$ for $U \subseteq W$, where $k = \dim U - \dim W$.

- The (positive) divisors of an integer n:
  $\mu(a,b) = \begin{cases} (-1)^r & \text{if } \frac{b}{a} \text{ is the product of } r \text{ distinct primes;} \\ 0 & \text{otherwise.} \end{cases}$

In number theory, the classical Möbius function is the function of one variable given by $\mu(n) = \mu(1,n)$ (in the notation of the third example above).

The following result is the *Möbius inversion* for locally finite posets. From the present point of view, it is obvious.

Theorem 7 *$f = g\zeta \Leftrightarrow g = f\mu$. Similarly, $f = \zeta g \Leftrightarrow g = \mu f$.*

Example: Suppose that f and g are functions on the natural numbers which are related by the identity $f(n) = \sum_{d|n} g(d)$. We may express this identity as $f = g\zeta$ where we consider f and g as vectors and where ζ is the zeta function for the lattice of positive integer divisors of n. Theorem 7 implies that $g = f\mu$, or

$$g(n) = \sum_{d|n} \mu(d,n) f(d) = \sum_{d|n} \mu\left(\frac{d}{n}\right) f(d),$$

which is precisely the classical Möbius inversion.

Example: Suppose that f and g are functions on the subsets of some fixed (countable) set X which are related by the identity $f(A) = \sum_{B \supseteq A} g(B)$. We may express this identity as $f = \zeta g$ where ζ is the zeta function for the lattice of subsets of X. Theorem 7 implies that $g = \mu f$, or

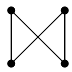
$$g(A) = \sum_{B \supseteq A} \mu(A,B) f(B) = \sum_{B \supseteq A} (-1)^{|B\setminus A|} f(B)$$

which is a rather general form of the inclusion/exclusion principle.

6 Lattices

A *lattice* is a poset (X,R) with the properties

- X has an upper bound 1 and a lower bound 0;
- for any two elements $x,y \in X$, there is a least upper bound and a greatest lower bound of the set $\{x,y\}$.

A simple example of a poset which is not a lattice is the poset .

In a lattice, we denote the l.u.b. of $\{x,y\}$ by $x \vee y$, and the g.l.b. by $x \wedge y$. We commonly regard a lattice as being a set with two distinguished elements and two binary operations, instead of as a special kind of poset.

Lattices can be axiomatised in terms of the two constants 0 and 1 and the two operations $\vee$ and $\wedge$. The result is as follows, though the details are not so important for us. The axioms given below are not all independent. In particular, for finite lattices we don't need to specify 0 and 1 separately, since 0 is just the meet of all elements in the lattice and 1 is their join.

Proposition 8 *Let X be a set, $\wedge$ and $\vee$ two binary operations defined on X, and* 0 *and* 1 *two elements of X. Then $(X,\vee,\wedge,0,1)$ is a lattice if and only if the following axioms are satisfied:*

- *Associative laws: $x \wedge (y \wedge z) = (x \wedge y) \wedge z$ and $x \vee (y \vee z) = (x \vee y) \vee z$;*
- *Commutative laws: $x \wedge y = y \wedge x$ and $x \vee y = y \vee x$;*
- *Idempotent laws: $x \wedge x = x \vee x = x$;*
- *$x \wedge (x \vee y) = x = x \vee (x \wedge y)$;*
- *$x \wedge 0 = 0$, $x \vee 1 = 1$.*

A *sublattice* of a lattice is a subset of the elements containing 0 and 1 and closed under the operations $\vee$ and $\wedge$. It is a lattice in its own right.

The following are a few examples of lattices.

- The subsets of a (fixed) set:
  $$\begin{aligned} A \wedge B &= A \cap B \\ A \vee B &= A \cup B \end{aligned}$$

- The subspaces of a vector space:

$$\begin{aligned} U \wedge V &= U \cap V \\ U \vee V &= \operatorname{span}(U \cup V) \end{aligned}$$

- The partitions of a set:

$$\begin{aligned} R \wedge T &= R \cap T \\ R \vee T &= \overline{R \cup T} \end{aligned}$$

Here $\overline{R \cup T}$ is the partition whose classes are the connected components of the graph in which two points are adjacent if they lie in the same class of either R or T.

7 Distributive and modular lattices

A lattice is *distributive* if it satisfies the *distributive laws*

(D) $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$ and $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$ for all x, y, z.

A lattice is *modular* if it satisfies the *modular law*

(M) $x \vee (y \wedge z) = (x \vee y) \wedge z$ for all x, y, z such that $x \leq z$.

Figure 3 presents a lattice, N_5, which is not modular, as well as a modular lattice, M_3, which is not distributive.

Not only are N_5 and M_3 the smallest lattices with these properties, they are, in a certain sense, the only lattices with these properties. The following theorem states this more precisely.

Theorem 9 *A lattice is modular if and only if it does not contain the lattice N_5 as a sublattice. A lattice is distributive if and only if it contains neither the lattice N_5 nor the lattice M_3 as a sublattice.*

The poset of all subsets of a set S (ordered by inclusion) is a distributive lattice: we have $0 = \emptyset$, $1 = S$, and l.u.b. and g.l.b. are union and intersection respectively. Hence every sublattice of this lattice is a distributive lattice.

Conversely, every finite distributive lattice is a sublattice of the lattice of subsets of a set. We describe how this representation works. This is important in that it gives us another way to look at posets.

N_5 M_3

Figure 3: Two lattices

Let (X,R) be a poset. Recall that an *down-set* in X is a subset Y with the property that, if $y \in Y$ and $z \leq_R y$, then $z \in Y$.

Let L be a lattice. A non-zero element $x \in L$ is called *join-irreducible* if, whenever $x = y \vee z$, we have $x = y$ or $x = z$.

Theorem 10 *(a) Let (X,R) be a finite poset. Then the set of down-sets in X, with the operations of union and intersection and the distinguished elements $0 = \emptyset$ and $1 = X$, is a distributive lattice.*

(b) Let L be a finite distributive lattice. Then the set X of non-zero join-irreducible elements of L is a sub-poset of L.

(c) These two operations are mutually inverse.

Meet-irreducible elements are defined dually, and there is of course a dual form of Theorem 10.

References

[1] B. A. Davey and H. A. Priestley, *Introduction to Lattices and Order*, Cambridge University Press, Cambridge, 1990.

[2] G.-C. Rota, On the foundations of combinatorial theory, I: Theory of Möbius functions, *Z. Wahrscheinlichkeitstheorie* **2** (1964), 340–368.

Peter J. Cameron
May 30, 2003