CS724: Topics in Algorithms
Set-Theoretical Preliminaries

Prof. Dan A. Simovici

UMASS BOSTON

Set Notations

For a set S we denote by $\mathcal{P}(S)$ the set of its subsets. The set of all finite non-empty subsets of S is denoted by $\mathcal{F}(S)$.

For a finite set S the number of elements of S is denoted by $|S|$. The empty set is denoted by $\emptyset$.

We write $x \in S$ to denote the fact that x is an element of the set S. The usual symbols are used to denote set-theoreical operations: $A \cup B$ is the union of the sets A and B, $A \cap B$ is the intersection of the sets A and B, and $A - B$ is the difference of the sets A and B.

The *symmetric difference* of the sets A and B is denoted by $A \oplus B$. We have

$$A \oplus B = (A - B) \cup (B - A).$$

The Galois Field GF(2)

For set inclusion we write $A \subseteq B$ to denote that each element x of A also belongs to B.
Note that $A = B$ if and only if $A \oplus B = \emptyset$.
Let GF(2) be the 2-element Galois field $\mathrm{GF}(2) = \{0, 1\}$. Addition "$\oplus$" and multiplication "$\cdot$" in this field are defined by the following two tables:

$\oplus$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

The Galois Field GF(2)

The set of subsets $\mathcal{P}(S)$ of a finite set $S = \{x_1, \ldots, x_n\}$ can be organized as a GF(2)-linear space by defining the sum of two subsets U, V as their symmetric difference

$$U \oplus V = (U - V) \cup (V - U).$$

Note that $U \oplus \emptyset = \emptyset \oplus U = U$.
Multiplication with scalars in $\{0, 1\}$ is defined as

$$0\ U = \emptyset \text{ and } 1\ U = U,$$

for every $U \in \mathcal{P}(S)$.

A basis in this linear space is the collection $\{\{x_1\}, \ldots, \{x_n\}\}$. Every subset U of S can be uniquely written as

$$U = a_1\{x_1\} \oplus \cdots \oplus a_n\{x_n\},$$

where

$$a_i = \begin{cases} 1 & \text{if } x_i \in U, \\ 0 & \text{if } x_i \not\in\in U, \end{cases}$$

for $1 \leqslant i \leqslant n$. Thus, the GF(2)-linear space of subsets of S is of dimension n.

The Galois Field GF(2)

For $U = a_1\{x_1\} \oplus \cdots \oplus a_n\{x_n\}$ and $V = b_1\{x_1\} \oplus \cdots \oplus b_n\{x_n\}$ the *inner product* is be defined as

$$(U, V) = a_1 b_1 \oplus \cdots \oplus a_n b_n.$$

Observe that $(U, V) = 0$ if and only if the set $U \cap V$ contains an even number of elements.
A non-empty set U can be orthogonal on itself if and only if it contains an even number of elements. Such a vector is referred to as being *self-orthogonal*.

Example

Let $S = \{x_1, x_2, x_3, x_4, x_5\}$ and let $U = \{x_1, x_2, x_4, x_5\}$. We have $U = 1\ \{x_1\} \oplus 1\ \{x_2\} \oplus 0\ \{x_3\} \oplus 1\ \{x_4\} \oplus 1\ \{x_5\}$, hence $(U, U) = 1 \oplus 1 \oplus 1 \oplus 0 \oplus 1 = 0$. Thus, U is a self-orthogonal vector in $\mathcal{P}(S)$.

A subspace of $\mathcal{P}(S)$ is a collection $\mathcal{T}$ of subsets of S such that $U, V \in \mathcal{T}$ implies $U \oplus V \in \mathcal{T}$.
The collection $\mathcal{T}^\perp$ that consists of subsets that are orthogonal on every set in the subspace $\mathcal{T}$ is a subspace; we refer to $\mathcal{T}^\perp$ as the *orthogonal subspace* of $\mathcal{T}$. Clearly, $\mathcal{T}^\perp$ consists of those subsets W of S whose intersection with every set of $\mathcal{T}$ contains an even number of elements.
Suppose that $\mathcal{T}$ is a subspace of $\mathcal{P}(S)$ of dimension k. There are k subsets of S, $U_1, \ldots, U_k$ such that every set $T \in \mathcal{T}$ can be written as

$$T = a_1 U_1 \oplus \cdots \oplus a_k U_k,$$

UMASS BOSTON

Definition

Let T be a subset of the *characteristic vector* of T is the vector $1_T \in \{0,1\}^n$ whose components $(1_T)_1, \ldots, (1_T)_n$ are defined by:

$$(1_T)_i = \begin{cases} 1 & \text{if } x_i \in T, \\ 0 & \text{if } x_i \not\in T, \end{cases}$$

for $1 \leqslant i \leqslant n$.

The vector $\mathbf{0}_n$ whose components are all equal to 0 is the characteristic vector of the empty subset $\emptyset$ of S.

Definition

Let S be a set. A *sequence of length n on S* is a mapping $\mathbf{s} : \{1, \ldots, n\} \longrightarrow S$. The set of sequences of length n on S is denoted by $\mathbf{Seq}_n(S)$.
An *ordered pair on S* is a sequence of length 2 on S; a *singleton* is a sequence of length 1.
If $\mathbf{s}$ is a sequence of length n on S and $\mathbf{s}(i) = x_i$ for $1 \leqslant i \leqslant n$, we write $\mathbf{s} = (x_1, \ldots, x_n)$. The elements $x_1, \ldots, x_n$ are the *components* of $\mathbf{s}$.

The length of a sequence $\mathbf{s}$ is denoted by $|\mathbf{s}|$.

Example

A sequence of natural numbers of length 6 is $\mathbf{s} = (6, 5, 2, 4, 9, 6)$. Note that in a sequence the same element of S may occur on multiple positions.

Counting Sequences

If S is a finite set containing m elements, then there are m^n sequences of length n for any $n \geqslant 1$. We extend the definition of sequences on S be defining the *null sequence on S* as the sequence λ that has no components, $\lambda = ()$. Note that there exists exactly one such sequence on S and this is consistent with the fact that $m^0 = 1$ for every $m \geqslant 1$. The *set of sequences of elements of S* is the set

$$\mathbf{Seq}(S) = \bigcup \{\mathbf{Seq}_n(S) \mid n \geqslant 0\}.$$

Operations with Sequences

If $\mathbf{s} = (s_1, s_2, \ldots, s_n)$ is a sequence in S, we refer to the sequence $\tilde{s} = (s_n, \ldots, s_2, s_1)$ as the *reversal* of the sequence s. Clearly $\tilde{\lambda} = \lambda$.
If $\mathbf{s} = (s_1, \ldots, s_n)$ and $\mathbf{t} = (t_1, \ldots, t_m)$ are two sequences on a set S, their *concatenation* is the sequence $\mathbf{st} = (s_1, \ldots, s_n, t_1, \ldots, t_m)$. For the null sequence we define $\lambda\mathbf{s} = \mathbf{s}\lambda = \mathbf{s}$ for every $\mathbf{s} \in \mathbf{Seq}(S)$. Note that $|\mathbf{st}| = |\mathbf{s}| + |\mathbf{t}|$ for all sequences $\mathbf{s}, \mathbf{t} \in \mathbf{Seq}(S)$.
Note that sequence concatenation is not a commutative operation in general.

Example

Let $\mathbf{s} = (1, 2, 3), \mathbf{t} = (4, 5)$. We have

$$\mathbf{st} = (1, 2, 3, 4, 5) \text{ and } \mathbf{ts} = (4, 5, 1, 2, 3),$$

so $\mathbf{st} \neq \mathbf{ts}$.

Sequence concatenation is an associative operation on $\mathbf{Seq}(S)$, that is $(\mathbf{st})\mathbf{u} = \mathbf{s}(\mathbf{tu})$ for every $\mathbf{s}, \mathbf{t}, \mathbf{u} \in \mathbf{Seq}(S)$.

Operations with Collections of Sets

Let $\mathcal{C} = \{S_i \mid i \in I\}$ be a collection of sets. Its union is the set U defined as

$$U = \bigcup_{i \in I} S_i.$$

Note that $\mathcal{C} \subseteq \mathcal{C}'$ implies $\bigcup \mathcal{C} \subseteq \bigcup \mathcal{C}'$.
Unlike the union, the intersection is defined only for collections that consist of subsets of a set S.
If $\mathcal{C}$ is a collection of subsets of S, that is, if $\mathcal{C} \subseteq \mathcal{P}(S)$. the intersection of $\mathcal{C}$ is the set of all elements of S that belong to every set of $\mathcal{C}$. The intersection of $\mathcal{C}$ is denoted by $\bigcap \mathcal{C}$.

If $\mathcal{C}$ and $\mathcal{C}'$ are two collections of subsets of a set S and $\mathcal{C} \subseteq \mathcal{C}'$, then $\bigcap \mathcal{C}' \subseteq \bigcap \mathcal{C}$. If $\emptyset$ is the empty collection of subsets of S, we define $\bigcap \emptyset = S$.

Definition

A *closure system on the set S* is a collection $\mathcal{K}$ of subsets of S such that for every collection of subsets $\mathcal{C}$ such that $\mathcal{C} \subseteq \mathcal{K}$ we have $\bigcap \mathcal{C} \in \mathcal{K}$.

Note that if $\mathcal{K}$ is a closure system on a set S, then $S \in \mathcal{K}$ because S is the intersection of the empty collection of subsets of $\mathcal{K}$.

Definition

Let $\mathcal{K}$ be a closure system on a set S and let T be a subset of S. The *closure of T relative to the closure system $\mathcal{K}$* is the set $\mathbf{K}(T) = \bigcap\{U \in \mathcal{K} \mid T \subseteq U\}$.

For every set T the collection $\mathcal{C}_T = \{U \in \mathcal{K} \mid T \subseteq U\}$ is non-empty because it includes at least S. The set $\bigcap \mathcal{C}_T$ is denoted by $\mathbf{K}(T)$ and is referred to as the *closure* of T.
To emphasize that the closure of T is computed relative to the closure system $\mathcal{K}$ we may denote this closure by $\mathbf{K}_{\mathcal{K}}(T)$.

Example

A subset E of $\mathbb{R}$ is said to be symmetric if $x \in E$ if and only if $-x \in E$. Let $\{E_i \mid i \in I\}$ be a collection of symmetric subsets of $\mathbb{R}$. It is easy to see that $\bigcap\{E_i \mid i \in I\}$ is a symmetric set. Note that $\mathbb{R}$ itself is symmetric. Thus, the collection $\mathcal{E}$ of symmetric subsets of $\mathbb{R}$ is a closure system. For a subset T of $\mathbb{R}$ the set $\mathbf{K}_{\mathcal{E}}(T)$ is the smallest symmetric set that includes T.

Definition

A *relation* on the set S is a set of ordered pairs of S.

The set of relations on S is denoted by $\mathbf{rel}(S)$.
Since relations on S are sets of pairs on S they can be involved in the usual set-theoretical operations: union, intersection, difference, etc. If $\rho, \sigma \in \mathbf{rel}(S)$, the union, intersection, and difference of ρ and σ are denoted by $\rho \cup \sigma, \rho \cap \sigma$, and $\rho - \sigma$, respectively.
Also, $\rho \subseteq \sigma$ denotes the inclusion of the set of pairs ρ into the set of pairs σ.

Two important relations on S are the *diagonal relation*

$$\iota_S = \{(x,x) \mid x \in S\},$$

and the *total relation*

$$\theta_S = \{(x,y) \mid x,y \in S\}.$$

Definition

Let $\rho, \sigma \in \mathbf{rel}(S)$. The *product* of ρ and σ is the relation $\rho\sigma$ given by

$$\rho\sigma = \{(x,z) \in \mathbf{Seq}_2(S) \mid (x,y) \in \rho \text{ and } (y,z) \in \sigma\}.$$

Definition

A relation $\rho \in \mathbf{rel}(S)$ is:

- *reflexive*, if $\iota_S \subseteq \rho$;
- *symmetric*, if $(x, y) \in \rho$ is equivalent to $(y, x) \in \rho$;
- *antisymmetric*, if $(x, y) \in \rho$ and $(y, x) \in \rho$ implies $x = y$;
- *transitive*, if $(x, y), (y, z) \in \rho$ implies $(x, z) \in \rho$,

for all $x, y, z \in S$.

If $\rho \in \mathbf{rel}(S)$, the *inverse* of ρ is the relation

$$\rho^{-1} = \{(y, x) \in S \times S \mid (x, y) \in \rho\}.$$

The *n^{th} power of a relation* ρ, where $\rho \subseteq S \times S$ is defined inductively as

$$\begin{aligned} \rho^0 &= \iota_S, \\ \rho^{n+1} &= \rho^n \rho \end{aligned}$$

for $n \geqslant 0$.
If ρ is a relation on S, then $(x, x) \in \rho^0$ for every $x \in S$. An easy argument by induction on $n \in \mathbb{N}$ shows that $(x, y) \in \rho^n$ if and only if there exists a sequence $\mathbf{z} = (z_0, z_1, \ldots, z_n)$ of length $n + 1$ such that $x = z_0$, $(z_i, z_{i+1}) \in \rho$ for $0 \leqslant i \leqslant n - 1$ and $z_n = y$.

Properties of relations can be expressed using the operations just introduced. For example, a relation ρ on a set S is symmetric if and only if $\rho^{-1} = \rho$; a relation ρ is transitive if $\rho^2 \subseteq \rho$.

Definition

An *equivalence relation* on a set S is a relation ρ, $\rho \subseteq S \times S$ that is reflexive, symmetric, and transitive.
The set of equivalence relations on S is denoted by $\mathrm{EQ}(S)$.

Example

Both ι_S and θ_S are equivalence relations on S; moreover, for any equivalence $\rho \in \mathrm{EQ}(S)$ we have $\iota_S \subseteq \rho \subset \theta_S$.

Example

Let m be a positive integer. Define the relation $\equiv_m$ on $\mathbb{Z}$ as consisting of those pairs $(p, q) \in \mathbb{Z} \times \mathbb{Z}$ if $p - q = km$ for some $k \in \mathbb{Z}$. In other words, we have $(p, q) \in \equiv_m$ if $p - q$ is divisible by m.

Note that $(r, r) \in \equiv_m$ because $r - r = 0$ is divisible by m. If $p - q = km$ for some $k \in \mathbb{Z}$, then $q - p = (-k)m$, so $(p, q) \in \equiv_m$ implies $(q, p) \in \equiv_m$. Finally suppose that $(p, q) \in \equiv_m$ and $(q, s) \in \equiv_m$. Since $p - q = km$ and $q - s = hm$, we have $p - s = (k + h)m$, hence $(p, s) \in \equiv_m$. Thus, $\equiv_m$ is an equivalence relation on $\mathbb{Z}$.

Following common practice, for an equivalence ρ on a set S and for $x, y \in S$ we write $x\rho y$ for $(x, y) \in \rho$.

Definition

Let ρ be an equivalence relation on a set S. The *equivalence class* of an element x of S is the set

$$[x]_\rho = \{u \in S \mid (x, u) \in \rho\}.$$

By the reflexivity of ρ, $(x, x) \in \rho$ for every $x \in S$. Thus, $x \in [x]_\rho$, hence each equivalence class is non-empty.

Lemma

Let ρ *be an equivalence relation on a set* S. *We have* $y \in [x]_\rho$ *if and only if* $[y]_\rho = [x]_\rho$.

Proof

Suppose that $y \in [x]_\rho$ and that $u \in [y]_\rho$. Then, we have $(x, y) \in \rho$ and $(y, u) \in \rho$. By transitivity, $(x, u) \in \rho$, that is, $u \in [x]_\rho$, which implies $[y]_\rho \subseteq [x]_\rho$.
If $v \in [x]_\rho$, then $(x, v) \in \rho$. Since $(x, y) \in \rho$, by the symmetry and transitivity of ρ we obtain $(y, v) \in \rho$, hence $v \in [y]_\rho$, so $[x]_\rho \subseteq [y]_\rho$. This implies $[x]_\rho = [y]_\rho$.
Conversely, if $[y]_\rho = [x]_\rho$, we have $y \in [x]_\rho$ because $y \in [y]_\rho$.

Theorem

Let ρ be an equivalence relation on a set S. If $[x]_\rho \neq [y]_\rho$, then $[x]_\rho \cap [y]_\rho = \emptyset$.

Proof.

Let $x, y \in S$ be such that $[x]_\rho \neq [y]_\rho$ and suppose that $z \in [x]_\rho \cap [y]_\rho$. Since $z \in [x]_\rho$ we have $[z]_\rho = [x]_\rho$; similarly, since $z \in [y]_\rho$ we have $[z]_\rho = [y]_\rho$, which means that $[x]_\rho = [y]_\rho$. This contradicts the hypothesis. □

Definition

Let S be a non-emptyset. A *partition* on S is a non-empty collection $\pi = \{B_i \mid i \in I\}$ such that

- $B_i \neq \emptyset$ for $i \in I$;
- $i, j \in I$ and $i \neq j$ implies $B_i \cap B_j = \emptyset$;
- $\bigcup_{i \in I} B_i = S$.

The sets B_i are the *blocks* of the partition π.
The set of partitions of a set S is denoted by $PART(S)$; the set of partitions of S that have k blocks, where $1 \leqslant k \leqslant |S|$ is denoted by $PART_k(S)$.
The partitions in $PART_2(S)$ are referred to as *bipartitions*.

Clearly, $PART(S) = \bigcup_{k=1}^{|S|} PART_k(S)$.

Example

The partition of a set S that consists of all singletons $\{x\}$, where $x \in S$ is denoted by α_S; the partition of S that contains one block, namely S, is denoted by ω_S. We have $PART_{|S|} = \{\alpha_S\}$ and $PART_1(S) = \{\omega_S\}$.

Example

Let ρ be an equivalence relation on a set S. The set of equivalence classes of ρ is a partition of the set S. Indeed, we saw that $S = \bigcup_{x \in S} [x]_\rho$, no equivalence class is empty and, as we saw, any two equivalence classes are disjoint.

The set of equivalence classes of an equivalence relation is known as the *quotient set* of S by ρ and is denoted by S/ρ. The partition generated by the equivalence relation is also denoted by π_ρ.

Example

Let $m \in \mathbb{P}$ and let B_i be the set of all members n of $\mathbb{P}$ such that the remainder of the division of n by m equals i, where $0 \leqslant i \leqslant m-1$. It is immediate that the collection $\{B_0, B_1, \ldots, B_{m-1}\}$ is a partition of the set $\mathbb{P}$. For instance, if $m = 3$, we have $B_0 = \{3, 6, 9, 12 \ldots\}$, $B_1 = \{1, 4, 7, 10, \ldots\}$, and $B_2 = \{2, 5, 8, 11, \ldots\}$.

Theorem

Let $\pi = \{B_i \mid i \in I\}$ *be a partition of the set* S. *The relation* ρ_π *defined by*

$$\rho_\pi = \{(x, y) \in S \times S \mid \{x, y\} \subseteq B_i \in \pi\}$$

is an equivalence on S.

Proof

Each x belongs to a block B_i of π, so $(x, x) \in \rho_\pi$ for every $x \in S$, which means that ρ_π is reflexive.
If $(x, y) \in \rho_\pi$, then $\{x, y\} \subseteq B_i$, which obviously implies $(y, x) \in \rho_\pi$, so ρ_π is symmetric.
Finally, if $(x, y) \in \rho_\pi$ and $(y, z) \in \rho_\pi$, there exist $B_i, B_j \in \pi$ such that $\{x, y\} \subseteq B_i$ and $\{y, z\} \subseteq B_j$. Thus, $B_i \cap B_j \neq \emptyset$ (because both contain y), which implies $B_i = B_j$. Therefore, $\{x, z\} \subseteq B_i = B_j$, hence $(x, z) \in \rho_\pi$, which allows us to conclude that ρ_π is an equivalence relation.

Corollary

Let $\pi \in PART(S)$ *and let* $\rho \in \mathrm{EQ}(S)$ $\rho = \rho_{\pi_\rho}$ *and* $\pi = \pi_{\rho_\pi}$.

Proof.

The equalities follow easily from the definitions of π_ρ and ρ_π. □

Example

Note that $\pi_{\iota_S} = \alpha_S$, $\pi_{\theta_S} = \omega_S$ and $\rho_{\alpha_S} = \iota_S$, $\rho_{\omega_S} = \theta_S$.

We write $x \equiv y(\pi)$ to denote that $(x, y) \in \rho_\pi$.

Denote by $(x)_n$ the n-degree polynomial

$$(x)_n = x(x-1)\cdots(x-n+1).$$

The coefficients of this polynomial

$$(x)_n = s(n,n)x^n + s(n,n-1)x^{n-1} + \cdots + s(n,i)x^i + \cdots + s(n,0)$$

are known as *the Stirling numbers of the first kind*.

Theorem

We have:

$$\begin{aligned} s(n,0) &= 0, \\ s(n,n) &= 1, \\ s(n+1,k) &= s(n,k-1) - ns(n,k). \end{aligned}$$

Proof.

The verification of the first two equalities is immediate. The third equality follows by observing that $(x)_{n+1} = (x)_n(x-n)$ and seeking the coefficient of x^k on both sides. □

Let S be a set having n elements. We are interested in the number of partitions of S that have k blocks. We begin by counting the number of onto functions of the form $f : A \longrightarrow B$, where $|A| = n$, $|B| = k$, and $n \geqslant k$.

Lemma

Let A and B be two sets, where $|A| = n$, $|B| = k$, and $n \geqslant k$. The number of surjective functions from A to B is given by

$$\sum_{j=0}^{k-1} (-1)^j \binom{k}{j} (k-j)^n.$$

Proof

There are k^n functions of the form $f : A \longrightarrow B$.
We begin by determining the number of functions that are not surjective. Suppose that $B = \{b_1, \ldots, b_k\}$, and let $F_j = \{f : A \longrightarrow B \mid b_j \not\in f(A)\}$ for $1 \leqslant j \leqslant k$. A function is not surjective if it belongs to one of the sets F_j. Thus, we need to evaluate $|\bigcup_{j=1}^{k} F_j|$. By using the inclusion-exclusion principle, we can write:

$$\begin{aligned}\left|\bigcup_{j=1}^{k} F_j\right| &= \sum_{j_1=1}^{k} |F_{j_1}| - \sum_{j_1,j_2=1}^{k} |F_{j_1} \cap F_{j_2}| \\ &+ \sum_{j_1,j_2,j_3=1}^{k} |F_{j_1} \cap F_{j_2} \cap F_{j_3}| - \cdots - +(-1)^k |F_1 \cap F_2 \cap \cdots \cap F_k|.\end{aligned}$$

Proof (cont'd)

Note that the set $|F_{j-1} \cap F_{j_2} \cap \cdots \cap F_{j_p}|$ is actually the set of functions defined on A with values in the set $B - \{y_{j_1}, y_{j_2}, \ldots, y_{j_p}\}$, and there are $(k-p)^n$ such functions. Since there are $\binom{k}{p}$ choices for the set $\{j_1, j_2, \ldots, j_p\}$, it follows that there are

$$\binom{k}{1}(k-1)^n - \binom{k}{2}(k-2)^n + \binom{k}{3}(k-3)^n - \cdots + (-1)^k \binom{k}{k-1}$$

functions that are not surjective.

Proof (cont'd)

Thus, we can conclude that there are

$$\sum_{j=0}^{k-1}(-1)^j\binom{k}{j}(k-j)^n = k^n - \binom{k}{1}(k-1)^n + \binom{k}{2}(k-2)^n - \cdots + (-1)^{k-1}\binom{k}{k-1}$$

surjective functions from A to B.

Theorem

The number of partitions of a set S that have k blocks ($k \leq n$) is given by

$$\frac{1}{k!}\sum_{j=0}^{k-1}(-1)^j\binom{k}{j}(k-j)^n.$$

Proof

Note that there are $k!$ distinct onto functions that have the same kernel partition. Indeed, given a surjective function $f : A \longrightarrow B$, one can obtain a function g that has the same partition as f by defining $g(a) = p(f(a))$, where p is a permutation of the set B, that is, a bijection $p : B \longrightarrow B$. Since there are $k!$ such bijections, it follows that the number of partitions is $\frac{1}{k!}\sum_{j=0}^{k-1}(-1)^j\binom{k}{j}(k-j)^n$.

The numbers $S(n,k)$ defined by

$$S(n,k) = \begin{cases} \frac{1}{k!}\sum_{j=0}^{k-1}(-1)^j\binom{k}{j}(k-j)^n & \text{if } n \geqslant k > 0, \\ 1 & \text{if } n = k = 0, \\ 0 & \text{in other cases.} \end{cases}$$

for $n, k \in \mathbb{N}$ and are known as the *Stirling numbers of the second kind*.

Example

Note that $S(n,1) = 1$ and $S(n,n) = 1$ because only one partition of a set with n elements, ω_S, has one block, and only one partition of a set with n elements, α_S has n blocks which are singletons.
The number of partitions of a 4-element set having two blocks is

$$\begin{aligned} S(4,2) &= \frac{1}{2!}\sum_{j=0}^{1}\binom{2}{j}(2-j)^4 \\ &= \frac{1}{2!}\left(\binom{2}{0}\cdot 2^4 - \binom{2}{1}\cdot 1^4\right) = 7. \end{aligned}$$

Namely, these partitions are:

$$\{\{1\},\{2,3,4\}\}, \{\{2\},\{1,3,4\}\}, \{\{3\},\{1,2,4\}\}, \{\{4\},\{1,2,3\}\},$$
$$\{\{1,2\},\{3,4\}\}, \{1,3\},\{2,4\}\}, \{\{1,4\},\{2,3\}\}.$$

We claim that

$$S(n, k) = kS(n-1, k) + S(n-1, k-1).$$

Indeed, observe that a partition π of the set $\{1, \ldots, n-1\}$ can be transformed into a partition of $\{1, \ldots, n\}$ be adjoining n to one of the blocks of π or by increasing the number of blocks by 1 and making $\{n\}$ a block.

Theorem

For every $n \geqslant 1$ *we have* $m^n = \sum_{j=1}^{n} S(n,j)(m)_j$.

Proof

Let A and B be two finite sets such that $|A| = n$ and $|B| = m$. There are m^n functions $f : A \longrightarrow B$. These functions can be classified depending on the size of their range $f(A)$. If $g : A \longrightarrow B$ is a function such that $|g(A)| = j$, then g can be regarded as a surjection from A to $g(A)$. Since there are $j!S(n,j)$ such surjective functions and there are $\binom{m}{j}$ subsets of B that have j elements, we can write

$$\begin{aligned} m^n &= \sum_{j=1}^{m} \binom{m}{j} j! S(n,j) \\ &= \sum_{j=1}^{m} m(m-1)\cdots(m-j+1)S(n,j) = \sum_{j=1}^{m} (m)_n S(n,j) \end{aligned}$$

for every $m \geqslant 1$.

The *Bell number* B_n is the total number of partitions of a set of n objects, that is,

$$B_n = \sum_{k=1}^{n} S(n,k).$$

Example

For $n = 4$, we have shown that there exist 7 partitions having two blocks, one partition with one block and one partition with 4 blocks. It is easy to see that there are 6 partitions with 3 blocks, so $B_4 = 1 + 7 + 6 + 1 = 15$.

The first 10 values of the Bell numbers are given below.

n	1	2	3	4	5	6	7	8	9	10
B_n	1	2	5	15	52	203	877	4140	21147	115975

Definition

A relation ρ is a *partial order* on a set S if ρ is reflexive, antisymmetric and transitive.
A *partially ordered set* (or, a *poset*) is a pair (S, ρ), where ρ is a partial order on S.

In general, we denote partial orders using the symbol "$\leqslant$" or similar symbols; furthermore, instead of writing $(x, y) \in \leqslant$, we write $x \leqslant y$.

Example

Let T be a set. The set of subsets of T, $\mathcal{P}(T)$ equippeed with the set inclusion "$\subseteq$" yields the poset $(\mathcal{P}(T), \subseteq)$.

Example

The pair $(\mathbb{P}, |)$, where "$|$" is the divisibility relation is a poset defined by $p|q$ if there exists $k \in \mathbb{P}$ such that $q = pk$. Indeed, we have $p|p$ for every $p \in \mathbb{P}$, so "$|$" is reflexive. If $p|q$ and $q|p$, we have $q = pk$ and $p = qh$, hence $hk = 1$ which implies $h = k = 1$. Thus, $p = q$, which shows that "$|$" is antisymmetric. Finally, if $p|q$ and $q|r$ we have $q = pk$ and $r = qh$ for some $k, h \in \mathbb{P}$. Thus, $r = pkh$, so $p|r$.

If (S, ρ) is a poset and $T \subseteq S$, it is easy to see that the relation $\rho_T = \rho \cap (T \times T)$ is itself a partial order; we will refer to it as the *trace of (S, ρ) on T*. Often, we will use the same symbol ρ instead of ρ_T to denote the partial order on T.

Example

Let $S \subseteq \mathbb{P}$ be the set $\{1, 2, 3, 4, 5, 6\}$. The trace of $\mathbb{P}$ on S consists of the pairs:

$$\begin{array}{l}(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),\\(2,2),(2,4),(2,6),(3,3),(3,6),(4,4),\\(5,5),(6,6).\end{array}$$

UMASS BOSTON

Definition

A *totally ordered set* is a pair (S, ρ), where ρ is a partial order with the additional property that for all $x, y \in S$ we have either $(x, y) \in \rho$, or $(y, x) \in \rho$. The relation ρ is refered to as a *total order*.

Example

The real numbers $\mathbb{R}$ equipped with the standard less-than-or-equal relation $\leqslant$ is a totally ordered set.

Definition

A sequence $\mathbf{x} \in \mathbf{Seq}(S)$ is a *subsequence of a sequence* $\mathbf{y} \in \mathbf{Seq}(S)$, if $\mathbf{y} = \mathbf{uxv}$ for some sequences $\mathbf{u}, \mathbf{v} \in \mathbf{Seq}(S)$. This is denoted by $\mathbf{x} \sqsubseteq \mathbf{y}$.

Example

Let $S = \{0, 1\}$. The sequence $\mathbf{y} = 1011$ is a subsequence of $\mathbf{x} = 010110110101100$.

The relation "$\sqsubseteq$" is a partial order on $\mathbf{Seq}(S)$.

Definition

Let $(P, \leqslant)$ be a poset. An element y *covers* an element x of P if $x \leqslant y$ and there is no $z \in P$, $z \neq x$ and $z \neq y$ such that $x \leqslant z \leqslant y$.
We denote the fact that y covers x by $x \prec y$.

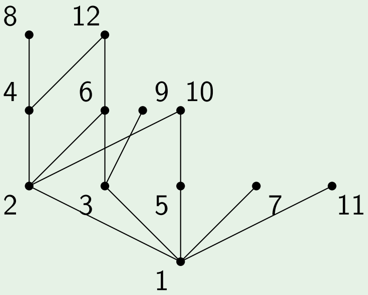
Example

Let $(\mathbb{P}, |)$ be the poset of positive numbers equipped with the divisibility relation. We have $p \prec y$ if x is none of the largest divisors of y. For example, we have $6 \prec 12$ beacuse there is no number z distinct from 6 and 12 such that $6|z$ and $z|12$. Note that $3|12$ but $3 \not\prec 12$.

Finite posets can be represented graphically using *Hasse diagrams*. Each element is represented by a dot. If x, y are elements of a poset $(P, \leqslant)$ and $x \prec y$, then the dot representing y is placed at a greater height than x and a link between the dots is drawn.

Example

The Hasse diagram of the poset $(\{1, \ldots, 12\}, |)$ is given below:



Definition

Let $(S, \leqslant)$ be a poset and let T be a subset of S. The *set of upper bounds* of T is the set

$$T^s = \{y \in S \mid \text{ for all } x \in T \text{ we have } x \leqslant y\}.$$

The *set of lower bounds* of T is the set

$$T^i = \{y \in S \mid \text{ for all } x \in T \text{ we have } y \leqslant x\}.$$

If T_1, T_2 are two subsets of S, $T_1 \subseteq T_2$ implies $T_2^s \subseteq T_1^s$ and $T_2^i \subseteq T_1^i$.

Theorem

Let $(S, \leqslant)$ be a poset and let T be a subset of S. The sets $T \cap T^s$ and $T \cap T^i$ contain at most one element of S.

Proof.

Suppose that $x, y \in T \cap T^s$. Since $x \in T$ and $y \in T^s$, it follows that $x \leqslant y$. On other hand, since $x \in T^s$ and $y \in T$ we have $y \leqslant x$. Therefore, $x = y$, which implies that the set $T \cap T^s$ contains at most one element. The argument for $T \cap T^i$ is similar. □

Definition

Let $(S, \leqslant)$ be a poset and let T be a subset of S. If $T \cap T^s = \{u\}$, then u is the *largest element* of set T.
If $T \cap T^i = \{v\}$, then v is the *least element* of set T.

Example

Not every subset of a poset has a least or a greatest element. The subset $\{1, 2, 3, 6\}$ of the poset $(\{1, \ldots, 12\}, |)$ considered before has 1 as its least element and 6 as the largest element. In contrast, the set $\{4, 5, 6\}$ has neither a least nor a largest element.

If T is a subset of a poset $(S, \leqslant)$ we will consider the sets $(T^s)^i$ and $(T^i)^s$ denoted by T^{si} and T^{is}, respectively.
Observe that the set $T^s \cap T^{si} = T^s \cap (T^s)^i$ may contain at most one element, by a previous observation applied to the set T^s. Similarly, the set $T^i \cap T^{is}$ may contain at most one element.

Definition

Let $(S, \leqslant)$ be a poset and let T be a subset of S. If $T^s \cap T^{si} = \{u\}$, u is the *supremum* of the set T.
If $T^i \cap T^{is} = \{v\}$, v is the *infimum* of T.

The supremum and infimum of a set T (if they exist) are unique and are denoted by $\sup T$ and $\inf T$, respectively.

Example

In the poset $(\mathcal{P}(T), \subseteq)$ introduced in before, for every $\mathcal{C} \in \mathcal{P}(X)$ we have

$$\inf \mathcal{C} = \bigcap \mathcal{C} \text{ and } \sup \mathcal{C} = \bigcup \mathcal{C}.$$

Example

In the poset $(\mathbb{P}, |)$, we have

$$\inf\{p, q\} = \gcd(p, q) \text{ and } \sup\{p, q\} = \mathrm{lcm}(p, q),$$

where $\gcd(p, q)$ is the greatest common divisor of p and q, and $\mathrm{lcm}(p, q)$ is the least common multiple of p and q.

Definition

A poset $(S, \leqslant)$ is a *lattice* if for every two elements $x, y \in S$ there exist $\inf\{x, y\}$ and $\sup\{x, y\}$. If $(S, \leqslant)$ is a lattice we use the notations

$$x \wedge y = \inf\{x, y\} \text{ and } x \vee y = \sup\{x, y\}.$$

The element $x \wedge y$ is referred to as the *meet* of x and y; $x \vee y$ is the *join* of x and y.
A poset $(S, \leqslant)$ is a *complete lattice* if for every $X \in \mathcal{P}(S)$ there exist $\inf X$ and $\sup X$.

Example

- $(\mathbb{P}, |)$ is a lattice;
- $(\mathcal{P}(T), \subseteq)$ is a complete lattice.

Note that if $(S, \leqslant)$ is a complete lattice and $S \neq \emptyset$, then this poset has a least element $0_S = \inf S$, and a greatest element $1_S = \sup S$.

Theorem

Let $(S, \leqslant)$ be a complete lattice and let W be a subset of S such that $1_S \in W$ and $T \subseteq W$ implies that $\inf T$ in S belongs to W. Then W is a complete lattice.

Proof

For every nonvoid subset T of W, $\inf T \in W$ and is the infimum of T in S. Let U be a subset of W defined as $U = T^s$. We have $U \neq \emptyset$ because $1_S \in W$. Then, $\inf U \in W$ is also an upper bound of T, and is actually the least upper bound of U. Thus, $(W, \leqslant)$ is a complete lattice.

Corollary

Let $\mathcal{K}$ be a closure system on a set S. The subsets of S in $\mathcal{K}$ form a complete lattice in which $\inf \mathcal{C} = \bigcap \mathcal{C}$ *and* $\sup \mathcal{C} = \bigcap \{T \in \mathcal{P} \mid C \subseteq T \text{ for every } C \in \mathcal{K}\}$.

Let π, σ be two partitions of S. We write $\pi \leqslant \sigma$ if each block B of π is included in a block C of σ.

Theorem

The pair $(PART(S), \leqslant)$ is a partially ordered set.

Proof

The relation "$\leqslant$" is obviously reflexive.
Suppose that we have both $\pi \leqslant \sigma$ and $\sigma \leqslant \pi$. Then, a block B of π is included in a block C of σ, and C, in turn, is included in a block B' of C. Thus, $B \subseteq C \subseteq B'$, which implies $B = C = B'$ because no block of π can be included into another block. Thus, $\pi \subseteq \sigma$. In the same manner, starting from a block C of σ we can show that $\sigma \subseteq \pi$, so $\pi = \sigma$. This shows that the relation "$\leqslant$" is antisymmetric. It is immediate that "$\leqslant$" is transitive

Theorem

Let $\pi, \sigma \in PART(S)$ be two partitions, $\pi = \{B_i \mid i \in I\}$ and $\sigma = \{C_j \mid j \in J\}$. We have $\pi \leqslant \sigma$ if and only if for each $j \in J$ there exists a subset I_j of I such that $C_j = \bigcup\{B_i \mid i \in I_j\}$.

Suppose that $\pi \leqslant \sigma$ and let $C \in \sigma$. Suppose that $B \cap C \neq \emptyset$. Since each block B of π is included in a block C' of σ we must have $C' = C$ because, otherwise C' and C would have a non-empty intersection. Thus, if a block B of π has a non-empty intersection with a block C of σ we must have $B \subseteq C$. This implies that a block C of σ is a union of block of π.
The converse implication is immediate.

Example

If $\pi \in PART(S)$ we have $\alpha_S \leqslant \pi \leqslant \omega_S$. Thus, α_S is the smallest element of $(PART(S), \leqslant)$ and ω_S is its largest element.

Definition

Let π, σ be two partitions of a set S. The partition σ covers π if $\pi < \sigma$ and there is no partition $\tau \in PART(S)$ such that $\pi < \tau < \sigma$.

Theorem

Let π, σ be two partitions of a set S. The partition σ covers π if and only if there exists a block C of σ that is the union of two blocks B and B' of π, and every other block of σ that is distinct of C is a block of π.

Proof

Suppose that σ is a partition that covers the partition π. Since $\pi \leqslant \sigma$, every block of σ is a union of blocks of π. Suppose that there exists a block E of σ that is the union of more than two blocks of π; that is, $E = \bigcup\{B_i \mid i \in I\}$, where $|I| \geqslant 3$, and let $B_{i_1}, B_{i_2}, B_{i_3}$ be three blocks of π included in E. Consider the partitions

$$\begin{aligned}\sigma_1 &= \{C \in \sigma \mid C \neq E\} \cup \{B_{i_1}, B_{i_2}, B_{i_3}\},\\ \sigma_2 &= \{C \in \sigma \mid C \neq E\} \cup \{B_{i_1} \cup B_{i_2}, B_{i_3}\}.\end{aligned}$$

It is easy to see that $\pi \leqslant \sigma_1 < \sigma_2 < \sigma$, which contradicts the fact that σ covers π. Thus, each block of σ is the union of at most two blocks of π.

Proof (cont'd)

Suppose that σ contains two blocks C' and C'' that are unions of two blocks of π, namely $C' = B_{i_0} \cup B_{i_1}$ and $C'' = B_{i_2} \cup B_{i_3}$. Define the partitions

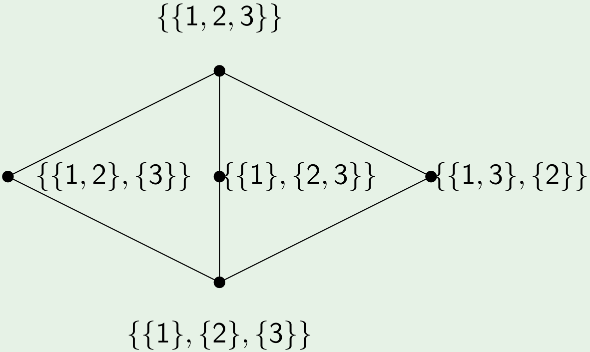
$$\begin{aligned} \sigma' &= \{C \in \sigma \mid C \notin \{C', C''\}\} \cup \{C', B_{i_2}, B_{i_3}\}, \\ \sigma'' &= \{C \in \sigma \mid C \notin \{C', C''\}\} \cup \{B_{i_1}, B_{i_2}, C''\}. \end{aligned}$$

Since $\pi < \sigma', \sigma'' < \sigma$, this contradicts the fact that σ covers π. Thus, we obtain the conclusion of the theorem.

Hasse diagrams

Example

The Hasse diagram of $(PART(\{1,2,3\}), \leqslant)$ is given below:



Theorem

The posets $(\text{EQUIV}(S), \subseteq)$ *and* $(PART(S), \leqslant)$ *are isomorphic.*

Let $f : \text{EQUIV}(S) \longrightarrow PART(S)$ be the mapping defined by $f(\rho) = S/\rho$. We need to show that f is a monotonic bijective mapping and that its inverse mapping f^{-1} is also monotonic.
The bijectivity of f follows immediately from the remarks that precede the theorem. Let ρ_0, ρ_1 be two equivalences such that $\rho_0 \subseteq \rho_1$ and let $S/\rho_0 = \{B_i \mid i \in I\}$, $S/\rho_1 = \{C_j \mid j \in J\}$. Let B_i be a block in S/ρ_0 and assume that $B_i = [x]_{\rho_0}$. We have $y \in B_i$ if and only if $(x, y) \in \rho_0$, so $(x, y) \in \rho_1$. Therefore, $y \in [x]_{\rho_1}$, which shows that every block $B \in S/\rho_0$ is included in a block $C \in \rho_1$. This shows that $f(\rho_0) \leqslant f(\rho_1)$, so f is indeed monotonic.

Let $\{\rho_i \mid i \in I\} \subseteq \text{EQUIV}(S)$ be a collection of equivalences. Then, $\inf\{\rho_i \mid i \in I\} = \bigcap_{i \in I} \rho_i$.

Definition

Let S be a set and let $\rho, \tau \in \text{EQUIV}(S)$. A *$(\rho,\tau)$-alternating sequence that joins x to y* is a sequence $(s_0, s_1, \ldots, s_n)$ such that $x = s_0$, $y = s_n$, $(s_i, s_{i+1}) \in \rho$ for every even i and $(s_i, s_{i+1}) \in \tau$ for every odd i, where $0 \leqslant i \leqslant n-1$.

Lemma

Let S be a set and let $\rho, \tau \in \mathrm{EQUIV}(S)$. If **s** *and* **z** *are two (ρ, τ)-alternating sequences joining x to y and y to z, respectively, then there exists a (ρ, τ)-alternating sequence that joins x to z.*

Proof

Let $(s_0, \ldots, s_n)$ be a (ρ, τ)-alternating sequences joining x to y and $(w_0, \ldots, w_m)$ a (ρ, τ)-alternating sequences joining y to z, where $x = s_0$, $s_n = w_0 = y$ and $w_m = z$. If $(s_{n-1}, s_n) \in \tau$, then the sequence $(s_0, \ldots, s_n, w_1, \ldots, w_m)$ is a (ρ, τ)-alternating sequence joining x to z. Otherwise, that is, if $(s_{n-1}, s_n) \in \rho$, then taking into account the reflexivity of τ we have $(s_n, w_0) = (s_n, s_n) \in \tau$. In this case, $(s_0, \ldots, s_n, s_n, w_1, \ldots, w_m)$ is a (ρ, τ)-alternating sequence joining x to z.

Theorem

Let S be a set and let $\rho, \tau \in \mathrm{EQUIV}(S)$. If ξ is the relation that consists of all pairs $(x, y) \in S \times S$ that can be joined by a (ρ, τ)-alternating sequence, then $\xi = \sup\{\rho, \tau\}$.

It is easy to verify that ξ is indeed an equivalence relation. Note that we have both $\rho \subseteq \xi$ and $\tau \subseteq \xi$. Indeed, if $(x, y) \in \rho$, then (x, y, y) is a (ρ, τ)-alternating sequence joining x to y. If $(x, y) \in \tau$, then (x, x, y) is the needed alternating sequence.
Let $\zeta \in \text{EQUIV}(S)$ such that $\rho \subseteq \zeta$ and $\tau \subseteq \zeta$. If $(x, y) \in \xi$, and $(s_0, s_1, \ldots, s_n)$ is a (ρ, τ)-alternating sequence such that $x = s_0$, $y = s_n$, then each pair (s_i, s_{i+1}) belongs to ζ. By the transitivity property, $(x, y) \in \zeta$, so $\xi \subseteq \zeta$. This implies that $\xi = \sup\{\rho, \tau\}$.

If $\pi, \sigma \in PART(S)$ both the infimum and the supremum of the set $\{\pi, \sigma\}$ exist and their description follows from the corresponding results that refer to the equivalence relations. Namely, if $\pi, \sigma \in PART(S)$, where $\pi = \{B_i \mid i \in I\}$ and $\sigma = \{C_j \mid j \in J\}$, the partition $\inf\{\pi, \sigma\}$ exists and is given by

$$\inf\{\pi, \sigma\} = \{B_i \cap C_j \mid i \in I, j \in J \text{ and } B_i \cap C_j \neq \emptyset\}.$$

The partition $\inf\{\pi, \sigma\}$ will be denoted by $\pi \wedge \sigma$.

A block of the partition $\sup\{\pi, \sigma\}$, denoted by $\pi \vee \sigma$, is an equivalence class of the equivalence $\theta = \sup\{\rho_\pi \wedge \rho_\sigma\}$. We have $y \in [x]_\theta$ if there exists a sequence $(s_0, \ldots, s_n) \in \mathbf{Seq}(S)$ such that $x = s_0$, $s_n = y$ and successive sets $\{s_i, s_{i+1}\}$ are included, alternatively, in a block of π or in a block of σ.