1. Posets and Möbius inversion

Previously we considered partially ordered sets with the anti symmetric relation $\preceq$ which satisfies

- $x \preceq x$
- $x \preceq y, y \preceq z \Rightarrow x \preceq z$
- $x \preceq y, y \preceq x \Rightarrow y = x$

We then defined a function $\zeta(x,y)$ in the following way

$$\zeta(x,y) = \begin{cases} 1 & \text{if } x \preceq y \\ 0 & \text{otherwise} \end{cases}$$

This $\zeta$ then forms an upper triangular matrix should we order rows and columns properly. This will prove valuable as many problems we wish to consider such as inclusion exclusion can be thought of as simply solving $\zeta x = y$ which then simply gives us $x = \zeta^{-1} y$

The catch here is that we know $\zeta$ has an inverse, we still need to actually compute it. For example, if our poset was the subsets of $\{1, \ldots, n\}$ then $\zeta$ has dimensions $2^n \times 2^n$ and while upper triangular matrices are easy to invert, the size of the matrix makes even a nice matrix difficult to invert. If our poset is structured, however, we may be able to use this structure to compute $\mu$ without inverting the full matrix.

1.1. Summing over Intervals. Note that if $\mu(x,y)$ is the $x,y$ entry of $\zeta^{-1}$ then all it really means to be the inverse is that $\mu$ should satisfy

$$\sum_z \zeta(x,z)\mu(z,y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{otherwise} \end{cases}$$

Last time we sketched the proof that $\mu(x,y) = 0$ unless $x \preceq y$. This allows us to then restrict the sum for $x \preceq z, z \preceq y$

We then have, provided $x \neq y$

$$\sum_{x \preceq z \preceq y} \mu(z,y) = 0$$

and via the same argument we may obtain again provided $x \neq z$

$$\sum_{x \preceq z \preceq y} \mu(x,z) = 0$$

This can be thought of as saying that the sum of the Möbius function over any interval is 0

1.2. Example: subsets of the two element set.

We construct the following Hasse diagram

```
{1,2}  {1}  {2}
\downarrow \quad \downarrow \\
\{1\}  \quad \{2\}
\downarrow \\
0
```

We have

$$\mu(\emptyset, \emptyset) = 1$$
and it follows that
\[ \mu(\emptyset, \{1\}) + \mu(\emptyset, \emptyset) = 0 \]
so
\[ \mu(\emptyset, \{1\}) = -1 \]
following this logic we obtain further that
\[ \mu(\emptyset, \{2\}) = -1 \]
Applying the interval argument one more time with \( y = \{1, 2\} \) we can further obtain
\[ \mu(\emptyset, \{1, 2\}) = -1 \]
Continuing onwards, \( \mu \) has the form
\[
\begin{array}{cccc}
0 & \{1\} & \{2\} & \{1, 2\} \\
\{1\} & 0 & 1 & 1 \\
\{2\} & 0 & 0 & 1 \\
\{1, 2\} & 1 & 0 & 1 \\
\end{array}
\]
One way of speeding up the calculation is to observe that the Möbius function only cares \( \mu(x, y) \) only cares about what happens in the part of the Poset between \( x \) and \( y \) (since the recursive calculation only involves that part of the Poset). For example, since the two chains
\[
\{1, 2\} \quad \{1\} \\
\downarrow \quad \downarrow \\
\{1\} \quad \{\emptyset\}
\]
are isomorphic, we have \( \mu(\emptyset, \{1\}) = \mu(\{1\}, \{1, 2\}) \).
This suggests that posets such as the subset lattice which contain enormous amounts of symmetry may have particularly simply Möbius functions.

1.3. Example: The Integers from 1 to \( n \).
Take \( 1, \ldots, n \) with standard \( \leq \).
\[
\{n\} \\
\downarrow \\
\vdots \\
\{2\} \\
\downarrow \\
\{1\}
\]
We have \( \mu(1, 1) = 1 \) (recall that \( \mu(x, x) = 1 \) always, since \( \mu \) is the inverse of \( \zeta \) and \( \zeta \) is upper triangular with 1 on the diagonal. Since any interval must add up to 0 we have \( \mu(1, 2) = -1 \) and \( \mu(1, y) = 0 \) for all \( y > 2 \).
Observing the isomorphism,
\[
y \quad \Rightarrow \quad (y - x - 1), \\
x \quad 1
\]
in general we have
\[
\mu(x, y) = \begin{cases} 
1 & \text{if } x = y \\
-1 & \text{if } x = y - 1 \\
0 & \text{otherwise}
\end{cases}
\]
What the inverse formula then says is that

\[ g(x) = \sum_{y=1}^{x} f(y) \]

, then

\[ f(x) = \sum_{y=1}^{x} \mu(y, x) g(y) = g(x) - g(x-1) \]

. 

1.4. Product posets and the Möbius function of the subset lattice.

**Definition.** Given two posets, \( S \) and \( T \) with relations \( \preceq_S \) and \( \preceq_T \) the **product poset** \( S \times T \) with the relation \( (s_1, t_1) \preceq (s_2, t_2) \iff s_1 \preceq_S s_2 \) and \( t_1 \preceq_T t_2 \) so

\[
\begin{array}{cc}
1 & 1 \\
| & \times \\
0 & 0
\end{array}
\]

can be thought of as

\[
(1,1) \quad \searrow \quad (0,1) \\
\nearrow \quad (1,0) \\
(0,0)
\]

This can also be thought of as a representation of the subsets of \( \{1, 2\} \), where the first coordinate asks if 1 is in the set, and the second coordinate asks if 2 is in the set. More generally, we can think of the subset lattice of \( \{1, \ldots, n\} \) as

\[
\begin{array}{cc}
1 & 1 \\
| & \times \ldots \times \\
0 & 0
\end{array}
\]

n times

Ie is 1 in my subset, is 2 in my subset, all the way up to n. The key thing about this is that it turns out that if we know the Möbius function for each part of the product we can write the Möbius function for the whole product.

**Theorem.** In the product poset \( S \times T \)

\[ \mu((s_1, t_1), (s_2, t_2)) = \mu_S(s_1, t_1) \mu_T(s_2, t_2) \]

**Proof.** It suffices to show that the \( \mu \cdot \zeta \) matrix is the identity for \( \mu \) as defined in this theorem. Consider the \( ((s_1, t_1), (s_2, t_2)) \) entry of the this matrix, which is given by

\[
\sum_{(s_3, t_3)} \mu((s_1, t_1), (s_3, t_3)) \zeta((s_3, t_3), (s_2, t_2))
\]

The summand is only nonzero in the range \( (s_1, t_1) \preceq (s_3, t_3) \preceq (s_2, t_2) \). Once we restrict to that range \( \zeta \) is always 1 so we may remove it from the summand and obtain

\[
\sum_{(s_1, t_1) \preceq (s_3, t_3) \preceq (s_2, t_2)} \mu((s_1, t_1), (s_3, t_3))
\]


And by our assumption our $\mu$ factors as it is a product, as well as our $\preceq$ relation factors, so the whole sum will factor.

$$\sum_{s_1 \preceq s_3 \preceq s_2} \mu_S(s_1, t_1) \mu_T(s_2, t_2)$$

$$\sum_{t_1 \preceq t_3 \preceq t_2} \mu_S(s_1, t_1) \sum_{t_1 \preceq t_3 \preceq t_2} \mu_T(s_2, t_2)$$

which is simply the product

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

The central observation to this procedure is that we can build bigger posets from smaller posets and if we understand how the Möbius function behaves on the smaller posets, we understand how it behaves on the bigger poset, and with this understanding we can obtain more complex inversion formulas.

Returning to our example we may now consider the Möbius function on the subset lattice. We know that $\mu(S, T) = 0$ unless $S \subseteq T$. If $S \subset T$ we have $\mu(S, T) = (-1)^{|T \setminus S|}$, since when we multiply the Möbius functions, we get a $(-1)$ for each time $T$ differs from $S$. The inversion formula then becomes

$$M(S) = \sum_{T \supseteq S} f(T) \quad \Rightarrow \quad f(S) = \sum_{T \supseteq S} M(T) (-1)^{|T \setminus S|}$$

In particular observe that if $S = \emptyset$ then

$$f(\emptyset) = \sum_T M(T) (-1)^{|T|}$$

Which is simply inclusion exclusion.

1.5. **Example: Divisor lattice.**

We now consider a new lattice that will again be the product of lattices we understand.

**Definition.** The divisor lattice of $n$ is the poset whose elements are all the positive integers which divide $n$, having the order relation $x \preceq y$ if and only if $x$ is a divisor of $y$.

For example, here is the divisor lattice of 12.

```
12
  / \  \
 4  6 
 /   \ 
2  3 
  \ / 
  1 
```

We can think of this as being isomorphic to

```
  4
 /|
2 x 3
| |
1 1
  
 4
```
In general if \( n = p_1^{a_1} p_2^{a_2} \ldots p_k^{a_k} \), then the divisor lattice is isomorphic to

\[
\begin{array}{cccc}
| & & | \\
\vdots & \times \ldots \times & \vdots \\
| & & | \\
0 & & 0 & & k \text{ copies}
\end{array}
\]

Using the previously computed M"obius function for the integers under \( \leq \), we immediately obtain

\[
\mu(x, y) = \begin{cases} 
0 & \text{if } x \mid y \\
0 & \text{if } \frac{x}{y} \text{ is not square free} \\
(-1)^k & \text{if } \frac{x}{y} \text{ is the product of } k \text{ distinct primes}
\end{cases}
\]

We will sometimes write \( \mu(1, d) \) as simply \( \mu(d) \), corresponding to the classical M"obius function from number theory which it equals. Note also that \( \mu(d, n) = \mu \left( \frac{n}{d} \right) \).

The M"obius inversion formula for this lattice then becomes that

\[
f(n) = \sum_{d \mid n} g(d) \implies g(n) = \sum_{d \mid n} f(d) \mu \left( \frac{n}{d} \right)
\]

We will now look at some familiar number theoretic identities that can be obtained via this formula. Let \( \phi \) denote Euler’s function. We first note

**Claim.** \( \sum_{d \mid n} \phi(d) = n \)

**Proof.** Consider a map \( \theta \) defined on \( \{1, \ldots, n\} \) defined by \( \theta(x) = \gcd(x, n) \). If \( d \mid n \) then the elements in the preimage of \( d \) are of the form \( dy \) where \( \gcd(y, \frac{n}{d}) = 1 \). In particular, the preimage of \( d \) has size \( \phi \left( \frac{n}{d} \right) \).

Everything must be mapped somewhere, and as there are \( n \) things total we have

\[
n = \sum_{d \mid n} \phi \left( \frac{n}{d} \right)
\]

Applying M"obius inversion to the above claim, we obtain

\[
\phi(n) = \sum_{d \mid n} \mu \left( \frac{n}{d} \right) \cdot d = \sum_{d \mid n} \frac{n}{d} \mu(d)
\]

which can be written elegantly as

\[
\frac{\phi(n)}{n} = \sum_{d \mid n} \frac{\mu(d)}{d}
\]

1.6. **Counting circular sequences of 0’s and 1’s.**

Our goal is to count how many circular sequences of 0’s and 1’s can be obtained if we consider two sequences that are rotations of each other to be the same. For example, for \( n = 4 \) we have the following sequences

\[
\begin{array}{cccccccc}
0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 1
\end{array}
\]

If it weren’t for the rotation equivalence, we would simply have \( 2^n \) sequences. For any sequence \( S \) having period \( d \), there are precisely \( d \) sequences (including \( S \)) equivalent to \( d \). Let’s define \( f(d) \) equal the number of rotation-inequivalent sequences of period \( d \). By the above observations, we know that

\[
\sum_{d \mid n} df(d) = 2^n.
\]
Applying Möbius inversion to this relation, we obtain
\[ nf(n) = \sum_{d|n} \mu(d)2^{\frac{n}{d}} \]
The expression we really care about is
\[ \sum_{d|n} f(d) \]
Applying the above relation to each summand, we obtain
\[ \sum_{d|n} f(d) = \sum_{d|n} \frac{1}{d} \sum_{j|d} \mu(j)2^{\frac{n}{j}} \]
This is a sum over chains \( j \mid d \mid n \). We can rewrite it as a sum parameterized by \( j \) and \( \ell := \frac{d}{j} \) (so that \( d = j\ell \)). Doing so (and reordering the sum so \( \ell \) is on the outside), we obtain
\[ \sum_{d|n} f(d) = \sum_{\ell|n} \sum_{j|\ell} \frac{1}{j\ell} \mu(j)2^{\frac{n}{\ell}} = \sum_{\ell|n} \frac{2^\ell}{\ell} \sum_{j|\ell} \mu(j) \frac{n}{j}. \]
Applying our claim from the previous section, we see that
\[ \frac{1}{n} \sum_{\ell|n} 2^\ell \phi \left( \frac{n}{\ell} \right). \]
If \( n \) is prime this has the particularly nice form \( \frac{1}{n} \left( 2^n + 2(n-1) \right) \). Note that when we applied our Möbius inversion after a bunch of cancelation we were left with a formula where all the summands are positive! This suggests that maybe there is an alternate approach where each term of the sum has some sort of combinatorial meaning, which we’ll see in the next section.

2. Counting orbits of a group \( G \) acting on a set \( X \)

We now are going to think of these problems in more algebraic terms. Suppose that we have a group \( G \) and it is acting on a set \( X \) by permuting it.

**Theorem.** *(Burnside’s Lemma)* For each \( g \in G \) let \( \psi(g) \) be the number of fixed points of \( G \) acting on \( X \). Then the number of orbits of \( G \)’s action on \( X \) satisfies
\[ \# \text{ orbits} = \frac{1}{|G|} \sum_{g \in G} \psi(g) \]
In other words, the number of orbits is just the average number of fixed points.

**Proof.** We count the number of pairs \( (x, g) \in X \times G \) where \( g(x) = x \) two different ways.

**Method 1**
\( \forall g \in G \) has \( \psi(g) \) solutions so the total is simply \( \sum_g \psi(g) \)

**Method 2**
take any \( x \in X \) if \( x \) is in the orbit of size \( O_x \) then \( x \) is fixed by \( \frac{|G|}{|O_x|} \) permutations so the total number of pairs is
\[ \sum_x \frac{|G|}{O_x} = |G| \sum_{\text{orbits } x} \sum \frac{1}{O_x} = 1 \]
thus
\[ \sum_{g \in G} \psi(g) = |G| \cdot \# \text{number of orbits} \]

\( \square \)

Note that in Method 2 we needed that \( G \) was a group in order to guarantee that \( x \) was spread evenly over its orbit by \( G \)’s action.
2.1. Example: \( \mathbb{Z}/n\mathbb{Z} \) acting on colorings. Here \( G \) is \( \mathbb{Z}/n\mathbb{Z} \), and \( X \) is the collection of all \( 2^n \) functions. \( G \) acts by rotation. If \( G \) rotates our set by \( k \), then the fixed points are precisely those functions with \( f(x) = f(x + k) \) for all \( x \). It follows from elementary number theory that this happens precisely when \( f \) is periodic with period a divisor of \( \gcd(k, n) \), so by Burnside’s lemma the number of orbits is

\[
\frac{1}{n} \sum_{k=0}^{n-1} 2^{\gcd(k, n)}
\]

Since for each \( d \mid n \) there are \( \phi\left(\frac{n}{d}\right) \) choices of \( k \) with \( \gcd(k, n) = d \), we can rewrite this as

\[
\frac{1}{n} \sum_{d \mid n} 2^d \phi\left(\frac{n}{d}\right)
\]

a simpler way of getting at the formula from the previous section.

3. Weighted Colorings

3.1. The Cycle Index. Our goal here is to generalize Burnside’s lemma to allow for counting of weighted colorings.

Definition. The cycle index of a group \( G \) acting on a set \( X \) is given by

\[
Z_g(x_1, \ldots, x_n) = \frac{1}{|G|} \sum_{g \in G} x_1^{z_1(g)} x_2^{z_2(g)} \cdots x_n^{z_n(g)},
\]

where \( z_j(g) \) denotes the number of cycles of length \( j \) in the permutation of \( X \) by \( g \).

Despite the notation, the cycle index depends both on the group and on the action itself. For example, consider the permutation of \( S_4 \) acting on \( \{1, 2, 3, 4\} \) in the standard way. The conjugacy classes here are

<table>
<thead>
<tr>
<th>Conjugacy Class</th>
<th>Number of Perm.</th>
<th>Cycle Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1234)</td>
<td>6</td>
<td>one 4-cycle</td>
</tr>
<tr>
<td>(12)(34)</td>
<td>8</td>
<td>one 3-cycle, one 1-cycle</td>
</tr>
<tr>
<td>(12)(3)(4)</td>
<td>3</td>
<td>two 2-cycles</td>
</tr>
<tr>
<td>(1)(2)(3)(4)</td>
<td>6</td>
<td>one 2-cycle, two 1-cycles</td>
</tr>
<tr>
<td>(1)</td>
<td>1</td>
<td>four 1-cycles</td>
</tr>
</tbody>
</table>

\[
Z_{S_4}(x_1, x_2, x_3, x_4) = \frac{1}{24}(x_1^4 + 6x_1^2x_2 + 3x_2^2 + 8x_1x_3 + 8x_1x_3 + 6x_4)
\]

3.2. New framework: Colorings. Let \( G \) be any group, \( A \) be any set and let \( B \) be a set finite set of colors. For each \( b \in B \) associate a weight \( w(b) \) To each function \( f : A \to B \) associate a weight \( W(f) = \prod_{a \in A} w(f(a)) \) We say two functions are equivalent if \( f_1(a) = f_2(g(a)) \) for some \( g \in G \) and all \( a \in A \). Our previous example corresponded to the case where all weights were equal to 1.

We now want to compute

\[
\sum_{\text{classes}} W(f)
\]

Note that that \( f_1 \sim f_2 \implies W(f_1) = W(f_2) \), so it makes sense to talk about the weight of an equivalent class.

Theorem. (Weighted Burnside’s)

\[
\sum_{\text{classes}} W(f) = \frac{1}{|G|} \sum_{g \in G} \sum_{f \in G} \Phi_W(g),
\]

where \( \Phi_W(g) \) denotes the total weight of the fixed points of \( g \).

Again, this reduces down to the previous version of Burnside’s lemma in the case where \( W \) is identically 1.

Proof. (sketched) Count the total weight of pairs \((g, f)\) where \( f \) is a function fixed by \( g \) two different ways in the exact same way as in the proof of the unweighted version.
Note that the fixed points of $g$ are precisely those colorings which are constant on every cycle of $g$’s action on $A$. In the simplest case where $g$ has one cycle of length $k$, then the fixed points are just the constant colorings, and the total weight of such colorings is
\[ \Phi_W(g) = \sum_b w(b)^k. \]

If instead $g$ has cycles of lengths $k_1, k_2, \ldots, k_\ell$, then we can write
\[ \Phi_W(g) = \sum_{(b_1, \ldots, b_\ell)} \prod_{i=1}^\ell w(b_i)^{k_i}. \]
\[ = \prod_{i=1}^\ell \left( \sum_b w(b)^{k_i} \right) \]
\[ = \left( \sum_b w(b)^1 \right)^{z_1(g)} \left( \sum_b w(b)^2 \right)^{z_2(g)} \cdots \left( \sum_b w(b)^n \right)^{z_n(g)}, \]
where $z_j(g)$ is the number of cycles of length $j$ in the action of $g$. Plugging this into Burnside’s Lemma, we see

**Theorem. (Pólya)**
\[ \sum_{\text{classes}} W(f) = Z_G \left( \sum_{b \in B} w(b), \sum_{b \in B} w(b)^2, \ldots, \sum_{b \in B} w(b)^n \right) \]

3.3. **Returning to circular sequence of 0’s and 1’s.** Now let us consider the circular coloring problem from before, now with $w(0) = 1$ and $w(1) = x$. We start by computing the cycle index of $\mathbb{Z}/n\mathbb{Z}$. We return to the familiar fact that $\phi(d)$ elements have $\gcd(n, k) = d$. It follows that the weighted sum of classes is
\[ Z_G(1 + x, 1 + x^2, \ldots, 1 + x^n) = \frac{1}{n} \sum_{d \mid n} \phi(d)(1 + x^d)^\frac{n}{d}. \]

Applying the binomial theorem to this, we obtain
\[ \frac{1}{n} \sum_{d \mid n} \phi(d) \cdot \sum_{r=0}^\frac{n}{d} \binom{\frac{n}{d}}{r} x^{rd} \]
As with generating functions, we can obtain additional information by extracting individual coefficients from this.
For the $x^k$ coefficient, $d$ is then forced to divide $k$. Since our sum was already restricted to terms which also divide $n$, $d$ necessarily divides $\gcd(n, k)$. Thus our $x^k$ coefficient is.
\[ \frac{1}{n} \sum_{d \mid \gcd(k, n)} \phi(d) \binom{\frac{n}{d}}{\frac{k}{d}} \]
Combinatorially, this is counting circular sequences zeroes and ones containing exactly $k$ ones.