[CS1961: Lecture 6] Sperner's Theorem, Möbius Inversion

Instructor: Chihao Zhang;
Scribed by Xiaoyang Lu, Yunran Yang, Xianruo Yu, Yuchen He

1 Sperner's Theorem

Recall that we have shown set inclusion is a canonical partial order (at least for finite posets) in the last lecture. Consider a typical poset $\left(2^{[n]}, \subseteq\right)$. What is the width (or equivalently, the size of the maximum anti-chain) of this poset?

Theorem 1 (Sperner's Theorem) *The size of the maximum anti-chain of $\left(2^{[n]}, \subseteq\right)$ is $\binom{n}{\lfloor \frac{n}{2} \rfloor}$.*

Proof. Suppose $\{S_1, S_2, \cdots S_w\}$ is the largest anti-chain of $\left(2^{[n]}, \subseteq\right)$. Note that the collection $\binom{[n]}{\lfloor \frac{n}{2} \rfloor}$ forms an anti-chain. Therefore $w \geqslant \binom{n}{\lfloor \frac{n}{2} \rfloor}$. We only need to show that $w \leqslant \binom{n}{\lfloor \frac{n}{2} \rfloor}$.

For each $i \in [w]$, let C_i be the set of maximal chains containing S_i [1]. Then we have $|C_i| = (n - |S_i|)! \cdot |S_i|!$ since there are $(n - |S_i|)!$ paths from $[n]$ to S_i and $|S_i|!$ paths from S_i to $\varnothing$.

[1] In other words, each element in C_i is a path in the Hasse diagram from $\varnothing$ to $[n]$ that passes through S_i.

Since for any $i \neq j$ and $i, j \in [w]$, $C_i \cap C_j = \varnothing$, we have

$$n! \geqslant \sum_{i=1}^{w} |C_i| = \sum_{i=1}^{w} (n - |S_i|)!\,|S_i|! = n! \sum_{i=1}^{w} \frac{(n - |S_i|)!\,|S_i|!}{n!} = n! \sum_{i=1}^{w} \binom{n}{|S_i|}^{-1}.$$

Note that $\binom{n}{|S_i|} \leqslant \binom{n}{\lfloor \frac{n}{2} \rfloor}$ for any $S_i \subseteq [n]$. Therefore,

$$1 \geqslant \sum_{i=1}^{w} \binom{n}{|S_i|}^{-1} \geqslant w \cdot \binom{n}{\lfloor \frac{n}{2} \rfloor}^{-1},$$

which yields that $w \leqslant \binom{n}{\lfloor \frac{n}{2} \rfloor}$. □

Sperner's theorem can be applied to prove some interesting results in combinatorics. Let's see an example.

For more applications of Sperner's theorem, see *Combinatorics: The Rota Way.*

Theorem 2 (Erdős, 1945) *Suppose $x_1, x_2, \cdots, x_n$ are n real numbers satisfying $|x_i| \geqslant 1$ for each $i \in [n]$. Let $\varepsilon = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)$ where each $\varepsilon_i \in \{-1, 1\}$. Then for an open interval I of length 2, there are at most $\binom{n}{\lfloor \frac{n}{2} \rfloor}$ ε's that satisfy $\sum_{i=1}^{n} \varepsilon_i x_i \in I$.*

Proof. W.l.o.g., assume that each $x_i \geqslant 0$. We can identify each ε with a unique set $A = \{i \in [n] \mid \varepsilon_i = 1\}$. Let $S(A) \triangleq \sum_{i \in A} x_i - \sum_{j \notin A} x_j = \sum_{i=1}^{n} \varepsilon_i x_i$. Then for any $A \subsetneq B \subseteq [n]$, $|S(A) - S(B)| \geqslant 2$.

If the two sets A and B satisfy $S(A) \in (y, y+2)$ and $S(B) \in (y, y+2)$ for some fixed $y \in \mathbb{R}$, we have $|S(A) - S(B)| < 2$. Therefore, neither

$A \subsetneq B$ nor $B \subsetneq A$ holds. This indicates that the sets in $\{A \subseteq [n] \mid S(A) \in I\}$ must have no inclusion relation with each other and form an anti-chain of $(2^{[n]}, \subseteq)$. Thus, the number of such ε's that $\sum_{i=1}^{n} \varepsilon_i x_i \in I$ is no larger than $\binom{n}{\lfloor \frac{n}{2} \rfloor}$ according to Theorem 1. □

2 Möbius Inversion

2.1 Inclusion-Exclusion Principle Revisit

Let $B = \{B_1, B_2, \cdots, B_m\}$ be a set of bad sets. For $S \subseteq B$, $N_=(S)$ is defined as the number of elements belong to exactly the sets in S (not in $B \setminus S$) and $N_\geqslant(S)$ is the number of elements belong to at least the sets in S.

It follows from the definition that $N_\geqslant(S) = \sum_{J:\, S \subseteq J} N_=(J)$ and the inclusion-exclusion principle states that

$$N_=(S) = \sum_{J:\, S \subseteq J} (-1)^{|J \setminus S|} N_\geqslant(J).$$

For example, let $B = \{B_2, B_3, B_5\}$ where $B_2 = \{x \in [100] \mid 2|x\}$, $B_3 = \{x \in [100] \mid 3|x\}$ and $B_5 = \{x \in [100] \mid 5|x\}$. Then $N_=(\{B_2, B_3\}) = |\{\text{numbers divisible by 2 and 3 but not by 5}\}|$ and $N_\geqslant(\{B_2, B_3\}) = |\{\text{numbers divisible by 2 and 3}\}|$.

We can use a matrix to represent the relationship between $N_=$ and $N_\geqslant$. Let Z be a square matrix that

$$Z[I, J] = \begin{cases} 1 & \text{if } I \subseteq J \\ 0 & \text{otherwise} \end{cases}.$$

It is easy to see that Z is an upper triangular matrix and each element on the diagonal is 1. Thus the determinant of Z is positive and Z is invertible. If we regard $N_=$ and $N_\geqslant$ as two vectors of length 2^m, then $Z \cdot N_= = N_\geqslant$. According to this viewpoint, the inclusion-exclusion principle can be described as $Z^{-1} \cdot N_\geqslant = N_=$. Such matrix Z is called the zeta matrix and Z^{-1} is called the Möbius matrix.

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} N_=(\varnothing) \\ N_=(\{B_2\}) \\ N_=(\{B_3\}) \\ N_=(\{B_5\}) \\ N_=(\{B_2, B_3\}) \\ N_=(\{B_2, B_5\}) \\ N_=(\{B_3, B_5\}) \\ N_=(\{B_2, B_3, B_5\}) \end{bmatrix} = \begin{bmatrix} N_\geqslant(\varnothing) \\ N_\geqslant(\{B_2\}) \\ N_\geqslant(\{B_3\}) \\ N_\geqslant(\{B_5\}) \\ N_\geqslant(\{B_2, B_3\}) \\ N_\geqslant(\{B_2, B_5\}) \\ N_\geqslant(\{B_3, B_5\}) \\ N_\geqslant(\{B_2, B_3, B_5\}) \end{bmatrix}$$

2.2 Incidence Algebra

We generalize the set of bad sets B to any poset $(P, \leqslant)$ and generalize the functions $N_\geqslant : 2^B \to \mathbb{R}$, $N_= : 2^B \to \mathbb{R}$ to any functions $N_\geqslant : P \to \mathbb{R}$,

$N_=: P \to \mathbb{R}$ satisfying $N_=(x) = \sum_{y:\, x \leqslant y} N_{\geqslant}(y)$ for any $x \in P$.

Then we have a natural generalization of the zeta matrix to encode the partial order.

Definition 3 (Zeta function)

$$\zeta(x, y) = \begin{cases} 1, & x \leqslant y, \\ 0, & otherwise. \end{cases}$$

For a locally finite poset[2] $(P, \leqslant)$, we can define the corresponding incidence algebra . For every functions $\alpha, \beta: P \times P \to \mathbb{R}$,

[2] A locally finite poset is a poset that for every $x \leqslant y$, the closed interval $[x, y] \triangleq \{z \mid x \leqslant z \leqslant y\}$ is finite.

- $(\alpha + \beta)(x, y) = \alpha(x, y) + \beta(x, y)$;
- for any $c \in \mathbb{R}$, $(c\alpha)(x, y) = c \cdot \alpha(x, y)$;
- $(\alpha\beta)(x, y) = \sum_{x \leqslant z \leqslant y} \alpha(x, z)\beta(z, y)$.

The multiplicative identity of the incidence algebra is defined as

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

which is called the delta function.

With the definition of identity, we have the notion of multiplicative inverse. The inverse of the zeta function is the Möbius function. It can also be defined inductively:

$$\mu(x, y) = \begin{cases} 0, & x \nleqslant y \\ 1, & x = y \\ -\sum_{z:\, x \leqslant z < y} \mu(x, z), & x < y \end{cases}.$$

Theorem 4 shows that these two definitions of Möbius function are the same.

Theorem 4 $\mu\zeta = \delta$.

Proof.

$$\begin{aligned} (\mu\zeta)(x, y) &= \sum_{x \leqslant z \leqslant y} \mu(x, z)\zeta(z, y) \\ &= \begin{cases} 0, & x \nleqslant y \\ 1, & x = y \\ \mu(x, y) + \sum\limits_{x \leqslant z < y} \mu(x, z), & x < y \end{cases} \\ &= \begin{cases} 0, & x \neq y \\ 1, & x = y \end{cases}. \end{aligned}$$

□

2.3 Möbius Inversion

Then we can generalize the inclusion-exclusion principle to any locally finite poset.

Theorem 5 (Möbius inversion) *Let* $(P, \leqslant)$ *be a locally finite poset. Then for any functions* $N_{\geqslant}$ *and* $N_{=}$ *satisfying* $N_{\geqslant}(x) = \sum_{y\,:\,x\leqslant y} N_{=}(y)$*, we have* $N_{=}(x) = \sum_{y\,:\,x\leqslant y} \mu(x,y) N_{\geqslant}(y)$. [3]

[3] Here we assume all summations are finite.

Proof.

$$
\begin{aligned}
\sum_{y\,:\,x\leqslant y} N_{\geqslant}(y)\cdot\mu(x,y) &= \sum_{y\,:\,x\leqslant y}\left(\sum_{z\,:\,y\leqslant z} N_{=}(z)\right)\cdot\mu(x,y)\\
&\overset{(1)}{=} \sum_{y\,:\,x\leqslant y}\left(\sum_{z\in P}\zeta(y,z)N_{=}(z)\right)\cdot\mu(x,y)\\
&= \sum_{z\in P} N_{=}(z)\cdot\left(\sum_{y\,:\,x\leqslant y}\mu(x,y)\zeta(y,z)\right)\\
&\overset{(2)}{=} \sum_{z\in P} N_{=}(z)\cdot\delta(x,z)\\
&= N_{=}(x),
\end{aligned}
$$

where we use the finite summation condition to change the summation order at (1) and the equation at (2) follows from Theorem 4. □

We will see some applications of above theorem in the next lecture.