NDMI113 – Extremal Combinatorics

Mykhaylo Tyomkyn

Lecture 1 - Introduction. Set systems, Sperner's theorem.

A major part of this course will deal with combinatorics of set systems. Here, a *set system* is a family of subsets of a finite set. Given a set X, we denote its power set by 2^X.

Definition 1 (Set system) *A* set system *is a set $\mathcal{F} \subseteq 2^X$ where X is a (typically, finite) set. We call X the* ground set*.*

Without loss of generality we can assume our ground set to be $[n] = \{1, \ldots, n\}$. In particular, $2^{[n]}$ is the set of all subsets of $[n]$, also known as the *boolean (hyper)cube*. Its cardinality is, naturally, 2^n.

Example 1 *$\mathcal{F} = \{A \subseteq [n] : |A| \equiv 0 \text{ mod } 2\}$ is a set system.*

The next example is a prominent one.

Example 2 *The family of all sets of a given size. In notation:*

$$\binom{[n]}{k} := \{A \subseteq [n] : |A| = k\}.$$

$\binom{[n]}{k}$ is referred to as a *layer* in the boolean cube, alternatively, as the *complete k-uniform hypergraph* on $[n]$. The cardinality of $\binom{[n]}{k}$ is, naturally, $\binom{n}{k}$.

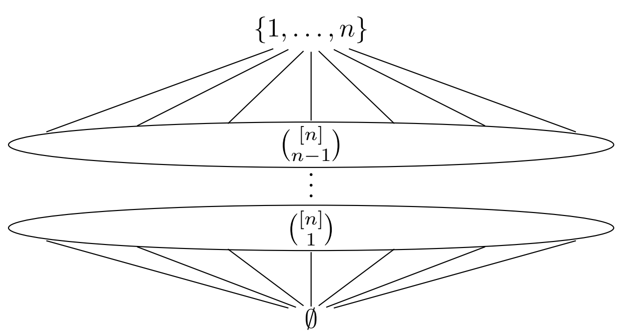

Figure 1: A graphical representation of $2^{[n]}$.

A set system $\mathcal{F} \subseteq \binom{[n]}{k}$ is called *k-uniform*, or a *k-uniform hypergraph*. Note that 2-uniform set systems are just graphs. [1]

[1] Set systems and hypergraphs are the same. The reason we have duplicate notation is historical. 'Hypergraphs' were first considered in an attempt to extend classical graph theory to higher uniformities, while 'set systems' were defined for questions that do not have natural graph analogues. We will mainly be dealing with the latter in this course.

Definition 2 *A* Sperner family*, or an* antichain *in the boolean cube, is a set system* $\mathcal{F} \subseteq 2^{[n]}$ *such that* $A \not\subseteq B$ *holds for all distinct* $A, B \in \mathcal{F}$*.*

Clearly, each layer $\binom{[n]}{k}$ is a Sperner family, but there are many more examples.

Example 3 *Let* $n = 3$*. The set system* $\mathcal{F} = \{\{1\}, \{2, 3\}\}$ *is a Sperner family.*

It is natural to ask, how large a Sperner family on $[n]$ can be. Each every layer $\binom{[n]}{k}$ is a Sperner family of size $\binom{n}{k}$, and by properties of binomial coefficients $\binom{[n]}{\lfloor n/2 \rfloor}$ is the largest among them, of size $\binom{n}{\lfloor n/2 \rfloor}$. Note that when n is even there is a unique 'middle layer' $\binom{[n]}{n/2}$, while for odd n there are two middle layers $\binom{[n]}{\lfloor n/2 \rfloor}$ and $\binom{[n]}{\lceil n/2 \rceil}$ of equal size. Sperner's theorem asserts that $\binom{n}{\lfloor n/2 \rfloor}$ is indeed the maximum.

Theorem 1 (Sperner, 1928) *The largest size of any Sperner family on* $[n]$ *is* $\binom{n}{\lfloor n/2 \rfloor}$.

To give a proof, let us first recall Hall's matching theorem.

Theorem 2 (Hall, '35) *Let* $G = G[X, Y]$ *be a bipartite graph with* $|Y| \geq |X|$*. There is a matching of size* $|X|$ *in* G *if and only if every subset* $S \subseteq X$ *satisfies* $|N(S)| \geq |S|$ *(Hall's condition).*[2]

For a proof of Hall's theorem we refer to Combinatorics and Graphs 1 course and the literature.

Proof [of Sperner's theorem] As size $\binom{n}{\lfloor n/2 \rfloor}$ is attained by the middle layer $\binom{[n]}{\lfloor n/2 \rfloor}$, we need to show that every antichain is of size at most $\binom{n}{\lfloor n/2 \rfloor}$. Define a *chain* to be a set system $\mathcal{A} = \{A_1, \ldots, A_m\} \subseteq 2^{[n]}$ such that $A_1 \subsetneq A_2 \subsetneq \cdots \subsetneq A_m$. We aim to partition $2^{[n]}$ into $\binom{n}{\lfloor n/2 \rfloor}$ disjoint[3] chains. Since each chain can contain at most one member of $\mathcal{F}$, this would readily imply $|\mathcal{F}| \leq \binom{n}{\lfloor n/2 \rfloor}$.

So, let $2k < n$ and consider two consecutive layers $\mathcal{X} = \binom{[n]}{k}$ and $\mathcal{Y} = \binom{[n]}{k+1}$ of $2^{[n]}$. Define the *inclusion graph* $G = G[\mathcal{X}, \mathcal{Y}]$ to be the bipartite graph between $\mathcal{X}$ and $\mathcal{Y}$ in which $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$ are adjacent whenever $X \subset Y$. We claim that G has a matching of size $|X|$. Indeed, note that in G the degree of each $X \in \mathcal{X}$ is $n - k$, while the degree of each $Y \in \mathcal{Y}$ is $k + 1$. Thus, for any $S \subseteq \mathcal{X}$, double counting, we obtain

$$|N(S)| \geq \frac{(n-k)||S|}{k+1} \geq |S|.$$

An analogous argument when $2k > n$ yields an inclusion preserving matching from $\binom{[n]}{k}$ to $\binom{[n]}{k-1}$. The desired partition into chains is then obtained by combining the matchings between all pairs of layers. (Working your way up from $\emptyset$ in the bottom half of the cube (down from $[n]$ in the top half), there may be isolated vertices in the middle layer, but this is fine, since we are allowed to have one-element chains.) If n is even, some lower-half and upper-half chains will meet in the unique middle layer, and we glue them there. If n is odd, they meet in one of the two middle layers. In either case we have constructed the desired partition. □

[2] $N(S) = N_G(S)$ is the neighbourhood of S, i.e. $\{y \in Y : ys \in E(G) \text{ for some } s \in S\}$.

[3] It is quite common nowadays to write "pairwise disjoint". To me "disjoint" already means that. It gets even worse when you see "pairwise distinct" being used.

Example 4 *One partition of* $2^{[3]}$ *into* $\binom{3}{1} = 3$ *chains is* $\{\emptyset, \{1\}, \{1,2\}, \{1,2,3\}\}, \{\{2\}, \{2,3\}\}, \{\{3\}, \{1,3\}\}$.

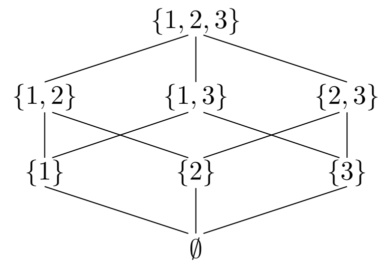

Figure 2: The boolean cube $2^{[3]}$.

Since the terminology of "chains/antichains" stems from the area of partially ordered sets, it will be fitting at this point to discuss Dilworth's theorem.

Definition 3 *A* partially ordered set *(poset)* $(X, \leq)$ *is a set* X *endowed with a binary relation* $\leq \subseteq X \times X$*, satisfying the* partial order axioms

- *Reflexivity:* $\forall x \in X \ x \leq x$,
- *Anti-symmetry:* $\forall x, y \in X \ (x \leq y \wedge y \leq x) \Rightarrow x = y$,
- *Transitivity* $\forall x, y, z \in X \ (x \leq y \wedge y \leq z) \Rightarrow x \leq z$.

By analogy with the $\leq$*-notation for real numbers, we write* $x \geq y$ *when* $y \leq x$*. If for* $x, y \in X$ *we can write* $x \leq y$ *or* $y \leq x$*, then* x *and* y *are said to be* comparable*. Otherwise, they are* incomparable*.*

Definition 4 *An* antichain *in a poset* $(X, \leq)$ *is a set* $F \subseteq X$ *such that no two elements of* F *are comparable.*

Definition 5 *A* chain *in a poset* $(X, \leq)$ *is a set* $C \subseteq X$ *such that any two elements of* C *are comparable.*

Theorem 3 (Dilworth) *In a finite poset* $(X, \leq)$ *the largest antichain equals in size the smallest partition of* X *into chains.*

Proof One direction is immediate: since an antichain and a chain can intersect in at most one element, if X can be partitioned into m chains, the largest antichain can have size at most m. We now need to prove the harder direction: if the largest antichain is of size m, then X can be partitioned into m chains.

To do so, we proceed by induction on $|X|$. If $|X| = 0$, there is nothing to prove. So, suppose $|X| \geq 1$ and that the statement holds for all smaller sets. Let S be a maximal chain in X (not a proper subset of any other chain). If $X \setminus S$ has no antichain of size m, then we are done by induction. So, suppose that $\{x_1, \ldots, x_m\}$ is an antichain in $(X \setminus S, \leq)$. Define C^+ and C^- to be

$$C^+ := \{x \in X \mid x \geq x_i \text{ for some } 1 \leq i \leq m\}$$
$$C^- := \{x \in X \mid x \leq x_i \text{ for some } 1 \leq i \leq m\}.$$

Observe that $C^+ \neq X$ and $C^- \neq X$ since $\max(S) \notin C^-$ and $\min(S) \notin C^+$, by maximality of S. On the other hand, we have $C^+ \cup C^- = X$ as otherwise X would contain an antichain of size strictly larger than m and $C^+ \cap C^- = \{x_1, \ldots, x_m\}$ as otherwise $x_i \leq x_j$ for some i, j. Now since $|C^+|, |C^-| < |X|$, we can apply the induction hypothesis on C^+ and C^- to partition both into m chains, $C_1^+, \ldots, C_m^+$ and $C_1^-, \ldots, C_m^-$. WLOG let $x_i \in C_i^+ \cap C_i^-$ for all i. We can now string the chains together by setting $C_i = C_i^+ \cup C_i^-$ for all i, obtaining the desired partition. □

Remark 1 *In the poset $(2^{[n]}, \subseteq)$ Dilworth's Theorem gives us a 'guideline' to proving Sperner's Theorem by asserting that if $\binom{n}{\lfloor n/2 \rfloor}$ is indeed the maximum size of an antichain, a partition into $\binom{n}{\lfloor n/2 \rfloor}$ chains must exist.*

Another consequence of Dilworth's Theorem is a proof of a classical theorem of Erdős and Szekeres.

Corollary 1 (Erdős-Szekeres '35) *In every sequence of $rs+1$ distinct real numbers, there is either an increasing subsequence of length $r+1$ or a decreasing subsequence of length $s+1$.*

Proof We define the following partial order on the underlying set of numbers: $x \leq y$ if x is less than y on the number line and x appears before y in the sequence. Under this partial order, a chain is a monotone increasing subsequence whereas an antichain is a monotone decreasing subsequence. By Dilworth's theorem, there is either an antichain of length $s+1$ or the sequence can be partitioned into at most s chains. By pigeonhole the longest of them will be of length at least $r+1$. □

The related *Erdős-Szekeres conjecture* in combinatorial geometry, from the same paper, states that among any $2^{n-2}+1$ points in the plane in general position one can always find n points in convex position. It was near-resolved in 2017 by Andrew Suk, who proved that $2^{n+\Omega(n)}$ are always sufficient. The proof makes crucial use of Dilworth's theorem.

Back to set systems, let us discuss a famous application of Sperner's theorem.

The Littlewood-Offord Problem: Let $a_1, \ldots, a_n$ be non-zero real numbers, and let $\vec{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)$ be the Bernoulli ± 1 vector, that is, each ε_i is $+1$ or -1 with probability $1/2$ each, independently. Consider

$$\langle \vec{a}, \vec{\varepsilon} \rangle = \sum_{i=1}^{n} a_i \varepsilon_i.$$

How large can $\mathbb{P}(\langle \vec{a}, \vec{\varepsilon} \rangle = c)$ get? Using Sperner's theorem, Erdős proved the tight bound of $\gamma/\sqrt{n}$.

Proof The problem amounts to finding, how many among the 2^n values of $\pm a_1 \pm \cdots \pm a_n$ can maximally attain the same value. By symmetry we may assume that all a_i are strictly positive. Apply the affine transformation $x \to \frac{1}{2}(x + \sum_{i=1}^n a_i)$. Then we are looking at partial sums $\sum_{I \subseteq [n]} a_i$. Notice now that for any given $c \in \mathbb{R}$, the family $\mathcal{F}_c := \{I \subseteq [n] : \sum_I a_i = c\}$ is a Sperner family on $[n]$. So, by Sperner's theorem, we have

$$|\mathcal{F}_c| \leq \binom{n}{\lfloor n/2 \rfloor} \sim \frac{2^n}{\sqrt{n}}.$$

□