REU 2005 · Discrete Mathematics · Lecture 7

Instructor: László Babai
Scribe: T. Schedler

June 30, 2005 (Afternoon). Last updated June 30, 2005

1 Lecture 7

1.1 Inclusion-Exclusion

Let $A_1, \ldots, A_n \subseteq \Omega$ be a family of subsets of Ω. If we know the size of the intersection $\bigcap_{i \in S} A_i$ for every $S \subseteq \{1, \ldots, n\}$, we can calculate $|\overline{A_1 \cup \ldots \cup A_n}|$ using the inclusions-exclusion formula:

Theorem 1.1 (Inclusion-Exclusion). *Let $A_1, \ldots, A_n \subseteq \Omega$ be a family of subsets of Ω. $|\overline{A_1 \cup \ldots \cup A_n}| = S_0 - S_1 + S_2 - \cdots \pm S_n$, where*

$$S_0 = |\Omega|$$

$$S_1 = \sum_i |A_i|$$

$$S_2 = \sum_{i,j} |A_i \cap A_j|$$

$$S_3 = \sum_{i,j,k} |A_i \cap A_j \cap A_k|$$

$$\vdots$$

$$S_k = \sum_{\substack{S \subseteq \{1,\ldots,n\} \\ |S|=k}} \left| \bigcap_{i \in S} A_i \right|$$

Note that there are $\binom{n}{k}$ terms in the summation defining S_k.

Proof. We wish to know the number of points in Ω which do not lie in any A_i. We claim that every point that lies inside at least one A_i is counted zero times by the above formula. Suppose that a point $x \in \Omega$ lies in exactly r of the A_i, without loss of generality $A_1, \ldots, A_r$. Then x contributes one to S_0, since every element of Ω gets counted once. Since x lies in exactly r of the A_i, it contributes r to the sum S_1. In general, x will contribute $\binom{r}{k}$ to the sum S_k, because there are that many ways to choose k of the r A_i's. Therefore, x is counted $\sum_{k=0}^{r}(-1)^k\binom{r}{k}$. We will prove later that this sum is 1 if r is 0 and 0 otherwise. Therefore, x is counted once if it lies outside all the A_i, and zero times otherwise, which is what we want. $\square$

The method of inclusion-exclusion is best illustrated by an example:

Example 1.2. *We have a five-letter alphabet, $\{A, B, C, D, E\}$. A* word *in this alphabet is just a finite string made up of these symbols. How many n-letter words are there which contain all five letters?*

We can rephrase this in the language of inclusion-exclusion by letting W_A be the set of n-letter words which do not contain A. W_B, W_C, W_D, W_D, W_E are defined analogously. We want to know how many words contain every letter, which is $|\overline{W_A \cup W_B \cup W_C \cup W_D \cup W_E}|$ $|W_A|$ is the number of n-letter words formed using only 4 letters (B, C, D, E), which is 4^n, $|W_A \cap W_B|$ is the number of n-letter words formed using only 3 letters (C, D, E), which is 3^n, etc.

Thus, by the inclusion-exclusion formula, the number of words which contain all five letters is $5^n - 5 \cdot 4^n + \binom{5}{2} \cdot 3^n - \binom{5}{3}2^n + \binom{5}{4}1^n$.

Exercise* 1.3 (Lovász). *Let S, T be finite semigroups. Prove: if $S \times S \cong T \times T$ (where $\times$ means direct product), then $S \cong T$. (Hint: use Inclusion-Exclusion.)*

Exercise 1.4. *Find a counterexample to the previous exercise when S and T are infinite.*

1.2 The Matrix-Tree Theorem

Theorem 1.5. *(Matrix-tree theorem): Let X be a graph with a vertex marked root. The determinant of the reduced Laplacian of X (obtained by deleting the column and row corresponding to the root from the Laplacian) is equal to the number of spanning trees directed towards the root.*

Recall that the Laplacian of a graph is the matrix:

$$\begin{pmatrix} \deg^+(1) & \dots & -a_{ij} \\ & \ddots & \\ & & \deg^+(n) \end{pmatrix}$$

where $\deg^+(i)$ is the out-degree of vertex i, and $a_{ij} = \#(i \to j \text{ edges})$, and the reduced Laplacian is the matrix obtained by deleting the row and column corresponding to the root.

Recall V is the set of vertices and $V_0 = V \setminus \{\text{sink}\}$ is the set of **ordinary vertices**. The sink is reachable from all vertices and there are no loops.

There was an example graph with Laplacian:

$$\begin{pmatrix} 1 & & -1 & & & \\ -1 & 3 & -1 & & & -1 \\ & -2 & 2 & & & \\ & & & 1 & -1 & \\ & & -1 & & 2 & -1 \\ 0 & \dots & \dots & \dots & \dots & 0 \end{pmatrix}.$$

Now if $B = (b_{ij})_{n\times n}$, then recall that

$$\det B = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} b_{i,i^\sigma}.$$

For example we have

i	i^σ
1	2
2	3
3	1
4	4
5	5

and

$$\begin{array}{c|l} \operatorname{sgn}(\sigma) = 1 & +b_{12}b_{23}b_{31}b_{44}b_{54} \\ \operatorname{sgn}(\sigma^{-1}) = 1 & +b_{21}b_{32}b_{13}b_{44}b_{55}. \end{array}$$

Lemma 1.6. *A k-cycle is an even permutation iff k is odd.*

Proof. A k**-cycle** (i.e. something of the form $i_1 \to i_2 \to \dots \to i_k \to i_1$ for distinct $i_1, \dots, i_k$) is a product of $k-1$ transpositions. For example, $(i_1, \dots, i_k) = (i_1, i_2)\cdots(i_{k-1}, i_k)$. □

Now recall that a functional subgraph of X is a subgraph Y such that $\deg_Y^+(v) = 1$ for all $v \in V_0$ and $\deg_Y^+(\text{sink}) = 0$. Spanning trees directed to the sink are the same as functional subgraphs which have no cycles.

Lemma 1.7. *Let X be a digraph, and let Δ denote its reduced Laplacian. The number of functional subgraphs of X containing disjoint cycles $\sigma_1, \sigma_2, \dots, \sigma_k$ is $\prod_{i=1}^n |\Delta_{ii^\pi}|$, where $\pi = \sigma_1 \cdot \sigma_2 \cdots \sigma_k$. Also, if this number is nonzero, the sign of sgn$(\pi)\prod_{i=1}^n \Delta_{ii^\pi}$ is positive if k is even, and negative if k is odd.*

Proof. Let V_f be the set of vertices of V_0 that are fixed points of π ($i^\pi = i$), i.e., they do not appear in the cycles σ_j. Let V_c be the vertices of V_0 that do appear in some σ_j. Now, to obtain a functional subgraph, we need to choose exactly one out-edge for every vertex. For a vertex $i \in V_f$, we can choose any of their edges as out-edges, which gives us $\deg^+(i)$ choices, which is $\Delta_{ii} = \Delta_{ii^\pi}$. For vertex i in V_c, we can choose an out-edge only from the out-edges which lie along the σ_ℓ which i is contained in. This gives us $a_{ii'}$ choices, where i' is the next vertex along the cycle σ_j. $a_{ii'} = -\Delta_{ii'} = -\Delta_{ii^\pi} = |\Delta_{ii^\pi}|$. Multiplying the number of choices together for each vertex, we get $\prod_{i=1}^n |\Delta_{ii^\pi}|$ choices for a functional subgraph containing the disjoint cycles $\sigma_1, \sigma_2, \dots, \sigma_k$.

Now, recall that $\mathrm{sgn}(\pi) = \mathrm{sgn}(\sigma_1)\cdots\mathrm{sgn}(\sigma_k)$. Also, recall that an ℓ-cycle is an even permutation if and only if ℓ is odd. Since $\Delta_{ij} = -a_{ij}$ for $i \neq j$, this means that we get an even number of minuses in each even cycle, and an odd number of minuses in each odd cycle. So, for each σ_ℓ, the product of the $\mathrm{sgn}(\sigma_\ell)\prod_{i\in\sigma_\ell}\Delta_{ii^{\sigma_ell}} \leq 0$, where $i \in \sigma_\ell$ means that vertex i is on cycle σ_ℓ. Therefore, the overall product is positive if and only if the number of cycles is even. □

Now, the proof of the **Matrix-Tree Theorem:**

Proof (Matrix-Tree Theorem). We use inclusion-exclusion. The number of spanning trees directed to the sink is equal to # functional subgraphs −# functional subgraphs containing a cycle. If we let A_σ be the set of functional subgraphs containing a particular cycle σ, then we want $|\overline{\cup_\sigma A_\sigma}|$. So

$$\prod_{v\in V_0}\deg^+(v) - \sum_{\pi,\text{containing one nontrivial cycle}} \#\text{functional subgraphs involving the cycle of } \pi$$
$$+ \sum_{\pi,\text{containing two nontrivial cycles}} \#\text{ functional subgraphs involving both cycles} - \sum + \ldots.$$

The terms of these summations are precisely $\mathrm{sgn}(\pi)\prod_{i=1}^n \Delta_{ii^\pi}$, and they have the correct signs because the inclusion-exclusion formula tells us to add the terms corresponding to permutations with an even number of cycles and subtract the terms corresponding to permutations with an odd number of cycles. Therefore, the number of spanning trees directed towards the root is precisely the determinant of the reduced Laplacian,

$$\sum_{\pi\in S_n}\mathrm{sgn}(\pi)\prod_{i=1}^n a_{ii^\pi}.$$

□

1.3 Some Binomial Coefficient Identities

To complete the proof of the inclusion-exclusion formula, we need to prove the fact that

$$\sum_{k=0}^r (-1)^k\binom{r}{k} = \begin{cases}1 & r=0\\ 0 & \text{otherwise}\end{cases}.$$

This is equivalent to saying that the number of even subsets of a set with r elements is equal to the number of odd subsets.

One proof, which is combinatorial, is the following: we give an explicit bijection between even and odd subsets ("clubs") by picking a special element ("the mayor"), and switching the mayor's membership status in each club: make the mayor a member if she wasn't, and revoke her membership if she was. This is obviously a bijection, and maps even subsets to odd subsets and vice versa.

A second proof relies on the binomial theorem. $(1+x)^n = \sum_{i=0}^{n} \binom{n}{i} x^i$, so

$$\sum_{k=0}^{r} (-1)^k \binom{r}{k} = (1+(-1))^r = \begin{cases} 1 & r=0 \\ 0 & \text{otherwise} \end{cases}.$$

Corollary 1.8. *The number of even subsets of a set of n elements is $2^{n-1} = \frac{1}{2}2^n$.*

Let $R(n,k)$ be the number of subsets of a set of size n which have size a multiple of k. $R(n,2) = \frac{1}{2} \cdot 2^n$, by the above. Is this indicative of a general pattern? Is $R(n,4) := \binom{n}{0} + \binom{n}{4} + \binom{n}{8} + \dots$ equal to $\frac{1}{4}2^n = 2^{n-2}$? Also, is $R(n,3)$ equal to $\frac{1}{3}2^n$? (Obviously not, since 2^n is not divisible by 3.)

Consider the following equations:

$$\begin{aligned}
(1+1)^n &= \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \binom{n}{3} + \dots, \\
(1-1)^n &= \binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \binom{n}{3} + \dots,
\end{aligned}$$

By adding them, the odd terms cancel out, so we get $2 \cdot R(n,2)$. Therefore, $R(n,2) = \frac{1}{2}2^n$.

For 4 we could try substituting fourth-roots of unity for x in the formula $(1+x)^n = \sum_j x^j \binom{n}{j}$. Doing this, we get

$$\begin{aligned}
(1+i)^n &= \binom{n}{0} + i\binom{n}{1} - \binom{n}{2} - i\binom{n}{3} + \binom{n}{4} + \dots \\
(1-i)^n &= \binom{n}{0} - i\binom{n}{1} - \binom{n}{2} + i\binom{n}{3} + \binom{n}{4} + \dots \\
(1+1)^n &= \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \binom{n}{3} + \binom{n}{4} + \dots \\
(1-1)^n &= \binom{n}{0} + \binom{n}{1} + \binom{n}{2} - \binom{n}{3} + \binom{n}{4} + \dots,
\end{aligned}$$

which we can add up and divide by four to get

$$R(n,4) = \frac{(1+1)^n + (1+i)^n + (1-1)^n + (1-i)^n}{4}.$$

For 3 we can similarly get

$$R(n,3) = \frac{(1+1)^n + (1+\omega)^n + (1+\bar{\omega})^n}{3},$$

where ω is a primitive cube root of unity: $\omega^3 = 1, \omega \neq 1$. That is, $\omega = \frac{1}{2} \pm i\frac{\sqrt{3}}{2}$.

Exercise 1.9. *Prove: $R(n,k) = \frac{1}{k}\sum_{z^k=1}(1+z)^n$. Here z ranges over all k-th roots of unity.*

As n gets large, the dominant term in the sum is 2^n so we end up getting something similar to $\frac{2^n}{k}$. We can formalize this in the following exercise:

Exercise 1.10. *Estimate $|R(n,4) - 2^{n-2}|$. (Prove that it is $O(2^{n/2})$.)*

Exercise 1.11. *Prove: $|R(n,3) - \frac{2^n}{3}| \leq \frac{2}{3}$.*

1.4 The Confluence Theorem

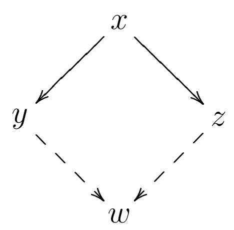
Definition 1.12. A *walk* in a digraph is a sequence of vertices $v_0, v_1, v_2, \ldots$ such that $v_i \to v_{i+1}$ is an edge for each i.

Definition 1.13. A *maximal walk* is either an infinite walk or a walk that ends in a *dead-end* (a vertex with out-degree 0).

Definition 1.14. A digraph (finite or infinite) is *confluent* if for each vertex x,

1. All maximal walks starting at x have the same length; and
2. If this length is finite, all maximal walks starting at x end at the same vertex.

Definition 1.15. The *diamond condition* for a vertex x is:



That is, if $x \to y, x \to z$ are edges, and $y \neq z$, then $\exists w : y \to w, z \to w$ are also edges.

Definition 1.16. A digraph is said to satisfy the diamond condition if the diamond condition is satisfied for every vertex.

Theorem 1.17 (Confluence Theorem). *If a digraph satisfies the diamond condition then it is confluent.*

Exercise 1.18. *Prove the Confluence Theorem.*

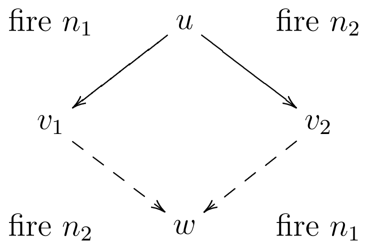
Corollary 1.19. *Suppose a chip-firing game starts at a given configuration C. If any avalanche starting from C has infinite length, then every avalanche starting from C has infinite length. If two avalanches take the same configuration C to stable configurations C_1 and C_2, then both avalanches had the same number of firings, and $C_1 = C_2$. Thus, the final, stable configuration depends only on the initial configuration.*

Proof. We consider a new digraph, with vertex set $V = \{$ chip configurations $\}$. We put an edge from u to v if the configuration v is reachable from the configuration u by firing one node. The dead ends in this digraph are exactly the stable configurations, in which no node can be fired. Walks in this digraph correspond to sequences of firings, and maximal walks correspond to avalanches.

Claim 1.20. *This digraph satifies the diamond condtion.*

Proof: Suppose we start at configuration u, and can move to two different configurations, v_1 and v_2. Suppose we can move to configuration v_1 by firing node n_1, or to configuration v_2 by firing node n_2. Then we can move to a common configuration w from both v_1 and v_2. We reach w from v_1 by firing node n_2, and from v_2 by firing node n_1.

Since n_1 and n_2 were both unstable in configuration u, n_2 is unstable in v_1 and n_1 is unstable in v_2, and we can perform these firing operations in either order. Since firing operations commute when they are defined, we reach a common configuration w.



This proves the claim.

So, by the Confluence Theorem, we conclude that the chip-firing graph is confluent, and that every avalanche from a given configuration has the same length. If this length is finite, then every avalanche ends at the same stable configuration. □

Definition 1.21. The *score vector* of an avalanche is a vector whose i-th coordinate is equal to the number of times vertex i fired during the avalanche.

Exercise 1.22. *Use the Confluence Theorem to prove that if any avalanche from a given configuration is finite then all avalanches are finite and have the same score vector.*

For those who know the Jordan-Hölder theorem in group theory, we have the

Exercise 1.23. *Infer the Jordan-Hölder theorem from the Confluence Theorem.*