Lecture 5: Using Inclusion-Exclusion, and the Pigeonhole Principle

Anup Rao

April 10, 2019

We discuss some examples using the inclusion-exclusion principle.

Last time, we proved the inclusion-exclusion formula. Given sets $A_1, \ldots, A_n$, and a subset $I \subseteq [n]$, let us write A_I to denote the intersection of the sets that correspond to elements of I:

$$A_I = \bigcap_{i \in I} A_i.$$

We proved:

Fact 1. $\left|\bigcup_{i\in[n]} A_i\right| = \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \cdot |A_I|$.

Example: Number of Derangements

How many ways are there to arrange n items so that for every j, the j'th item is not in the j'th position?

This may seem familiar: you had to calculate the number of derangements for small valules of n in the homework.

Define A_j to be the set of permutations where j is mapped to j. So $|A_I| = (n - |I|)!$. Then the set of permutations that *do not* leave an element in its positions is just

$$n! - \left|\bigcup_{j\in[n]} A_j\right| = n! - \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \cdot |A_I| \qquad \text{by the inclusion-exclusion principle}$$

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

Note that $\sum_{i=0}^{n} \frac{(-1)^i}{i!}$ is a truncation of the Taylor series expansion for $e^{-1} = \sum_{i=0}^{\infty} \frac{(-1)^i}{i!}$, so this quantity is approximately $\frac{n!}{e}$.

Example: Euler's Totient function

Given a positive integer N, with prime factorization $p_1^{e_1} p_2^{e_2} \cdots p_t^{e_t}$, how many numbers from 1 to N are relatively prime to N?

Two numbers N and M are relatively prime if the greatest common divisor of N and M is 1.

This quantity is called *Euler's Totient* function $\phi(N)$. Euler's totient function is important to estimate, because it tells us something about many crypto systems whose security relies on the difficulty of factoring large numbers. If many numbers M from 1 to N have common factors with N, then you can find a factor of N by computing the greatest common divisor of M and N using Euclid's algorithm.

So, for a crypto system to have good security, it better use a number N for which $\phi(N)/N$ is very close to 1. Otherwise, you could pick a random number M from 1 to N and compute the gcd with N to find a factor of N and break the cryptosystem.

Let A_i denote the set of numbers from 1 to N that are divisible by p_i. Then $|A_I| = \frac{N}{\prod_{j\in I} p_j}$. So the number of relatively prime numbers is

$$\begin{aligned} & N - \sum_{\emptyset \neq I \subseteq [t]} (-1)^{|I|+1} |A_I| \\ =& (-1)^{|\emptyset|} \frac{N}{\prod_{j\in\emptyset} p_j} - \sum_{\emptyset \neq I \subseteq [t]} (-1)^{|I|+1} \frac{N}{\prod_{j\in I} p_j}. \end{aligned}$$

the first term corresponds to N, and we substituted the value of $|A_I|$ in the second term

Now, we can combine the term corresponding to N with everything else to get:

$$= \sum_{I\subseteq[t]} (-1)^{|I|} \frac{N}{\prod_{j\in I} p_j} = N \cdot \sum_{I\subseteq[t]} (-1)^{|I|} \frac{1}{\prod_{j\in I} p_j}.$$

The sum we have is exactly the same as:

$$= N \cdot \prod_{i=1}^{t} (1 - 1/p_i).$$

The reason this works is that, for example when you multiply $(1 - 1/p_1)(1 - 1/p_2)$, you get 4 terms

$$1 - 1/p_1 - 1/p_2 + 1/(p_1 p_2),$$

corresponding to the 2 choices 1 or $-1/p_1$ from the first product term, and the 2 choices 1 or $-1/p_2$ from the second product term. In general, when you multiply t such product terms, you will get 2^t terms in the sum, and those are exactly the 2^t terms we obtained.

The formula we have obtained has a very natural interpretation: intuitively $1/p_1$ fraction of the numbers from 1 to N are divisible by p_1. After we eliminate these, we are left with $N(1 - 1/p_1)$ numbers. We should expect $1/p_2$ fraction of these to be divisible by p_2, which leaves $N(1 - 1/p_1)(1 - 1/p_2)$ numbers and so on.

This intuition is not a proof: the argument ought to work only when the numbers $p_1, p_2, \ldots$ are distinct prime divisors of N.

Example: Ryser's Formula for the Permanent

We did not have the time to discuss Ryser's formula in class, and you will not be tested on it. I describe it here because it is another cool application of the inclusion-exclusion principle that is relevant to computer science.

The permanent of an $n \times n$ matrix M is defined to be

$$\mathsf{perm}(M) = \sum_{\text{permutations } \pi : [n] \to [n]} \prod_{i=1}^{n} M_{i,\pi(i)}.$$

Computing the permanent is a very important fundamental problem in computer science. We do not know of any algorithms that

run in less than exponential time. Moreover, if we could compute the permanent in polynomial time, then we could solve all the algorithmic problems in the hard complexity class NP in polynomial time as well! For example, this would give us efficient algorithms for all machine learning problems. This would be true even if we could evaluate the permanent on matrices whose entries are either 0 or 1.

However, we do not know any fast algorithms for computing the permanent. The most naive algorithm would be to run over all the permutations $\pi : [n] \to [n]$ and compute the sum according to the formula for the permanent. This would take time about $n \cdot n!$. Here we show how to use the inclusion-exclusion principle to get a much faster algorithm that runs in time $2^{O(n)}$.

Ryser's formula says:

$$\mathsf{perm}(M) = (-1)^n \sum_{S \subseteq [n]} (-1)^{|S|} \prod_{i=1}^{n} \sum_{j \in S} M_{i,j}.$$

This formula can be evaluated in time proportional to $2^n \cdot n^2$. The formula holds for all matrices, but for simplicity let us just prove it for matrices M that have $0/1$ entries.

Then the formula

$$\mathsf{perm}(M) = \sum_{\text{permutations } \pi} \prod_{i=1}^{n} M_{i,\pi(i)}$$

just counts the number of sequences $j_1, j_2, \ldots, j_n \in [n]$ such that $j_1, \ldots, j_n$ is a permutation of $1, 2, \ldots, n$, and $M_{1,j_1} M_{2,j_2} \ldots M_{n,j_n} = 1$. If we did not have the restriction that $j_1, \ldots, j_n$ corresponds to a permutation, then the number of such sequences would be easy to compute, it would be equal to

$$\prod_{i=1}^{n} \sum_{j=1}^{n} M_{i,j}.$$

So, let A_i denote the set of sequences $j_1, \ldots, j_n$ where i does not appear in the sequence, and $M_{1,j_1} M_{2,j_2} \ldots M_{n,j_n} = 1$. Then we see that $|A_i| = \prod_{i=1}^{n} \sum_{j \neq i} M_{i,j}$. Similarly, given any set $I \subseteq [n]$, we have

$$\left| \bigcap_{i \in I} A_i \right| = \prod_{i=1}^{n} \sum_{j \notin I} M_{i,j}.$$

The permanent is just the total number of sequences, minus the elements of $\bigcup_{i \in [n]} A_i$, so by the inclusion-exclusion formula, it is

equal to:

$$\begin{aligned}
\mathsf{perm}(M) &= \prod_{i=1}^{n}\sum_{j=1}^{n} M_{i,j} - \left|\bigcup_{i\in[n]} A_i\right| \\
&= \prod_{i=1}^{n}\sum_{j=1}^{n} M_{i,j} - \sum_{\emptyset\neq I\subseteq[n]} (-1)^{|I|+1}\left|\bigcap_{i\in I} A_i\right| && \text{now use } \left|\textstyle\bigcap_{i\in I} A_i\right| = \prod_{i=1}^{n}\sum_{j\notin I} M_{i,j} \\
&= \sum_{I\subseteq[n]} (-1)^{|I|}\prod_{i=1}^{n}\sum_{j\notin I} M_{i,j} && \text{the first term corresponds to } I=\emptyset \\
&= \sum_{S\subseteq[n]} (-1)^{n-|S|}\prod_{i=1}^{n}\sum_{j\in S} M_{i,j} && \text{setting } S \text{ to be the complement of } I \\
&= (-1)^n \sum_{S\subseteq[n]} (-1)^{|S|}\prod_{i=1}^{n}\sum_{j\in S} M_{i,j}.
\end{aligned}$$