CSE 312: Foundations of Computing II
Advanced Topics Session #1

Special Topics in Combinatorics: Ramsey Theory & Proofs by Counting

Lecturer: Alex Tsun
Date: April 4, 2017

0 Introduction

Probability has become more and more prominent and useful in computer science, especially in machine learning and artificial intelligence. Every week, we will discuss some interesting topic (which you will be able to understand we progress through CSE 312), which either I find interesting, or which is extremely relevant to fields such as machine learning, or both! Today we discuss Ramsey Theory and the Proofs By Counting, but first cover multinomial coefficients and a review of graph terminology.

1 Multinomial Coefficients

Recall that $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ is the number of ways to choose k items out of n, where order does not matter. For example, to count the number of ways we can get from $(0,0)$ to (r, s) with steps only allowed going north one unit or east one unit is in one-to-one correspondence with sequences of characters N and E of length $r + s$ with exactly r N's and s E's. The number of different orderings of N's and E's is simply $\binom{r+s}{r} = \binom{r+s}{s} = \frac{(r+s)!}{r!s!}$ in which you can choose the location of the N's or E's.

Now consider a more complicated example, $BUBBLE$. How many ways can we rearrange these letters? Suppose we numbered the B's so that they were distinct, so we are counting the orderings of $B_1UB_2B_3LE$, of which there are $6!$. However any permutation of the B's is counted $3!$ times. For example, $B_1UB_2B_3LE = B_1UB_3B_2LE = \cdots = B_3UB_2B_1LE$. So to get our result, we divide by $3!$ To get $\frac{6!}{3!}$.

Finally consider the word $GODOGGY$. How many ways can we rearrange these letters? Let's number the G's and O's as we did earlier and get $G_1O_1DO_2G_2G_3Y$. There are $7!$ ways to arrange these distinct letters. For the G's, we are overcounting by a factor of $3!$ and for the O's, we are overcounting by a factor of $2!$. So our answer is $\frac{7!}{3!2!}$.

In full generality: Suppose there are n objects, but only k are distinct, with $k \leq n$. (For example, "godoggy" has $n = 7$ objects [characters] but $k = 4$ distinct objects, $\{g, o, d, y\}$). Let n_i for $i = 1, \ldots, k$ be the number of times object i appears. (For example, $\{n_1 = 3, n_2 = 2, n_3 = 1, n_4 = 1\}$ continuing the "godoggy" example). The number of ways to arrange the n objects is then

$$\frac{n!}{n_1!\, n_2! \ldots .\, n_k!} = \binom{n}{n_1, n_2, \ldots, n_k}$$

Multinomial Theorem: Let $x_1, \dots, x_m \in \mathbb{R}, n \in \mathbb{N}$.

$$(x_1 + \cdots + x_m)^n = \sum_{\substack{k_1+\cdots+k_m=n \\ k_1,\dots,k_m \geq 0}} \binom{n}{k_1, \dots, k_m} \prod_{t=1}^{m} x_t^{k_t}$$

The multinomial theorem is a generalization of the binomial theorem. The multinomial coefficient and theorem will be useful later when we talk about the multinomial distribution, and it generalizes our previous example about counting the number of ways to walk from the origin to some point in $\mathbb{R}^2$ to $\mathbb{R}^m$.

2 Graphs

A **graph** $G = (V, E)$ is a collection of **vertices** $V = \{v_1, \dots, v_n\}$ and **edges** $E = \{e_1, \dots, e_m\}$. We typically have $|V| = n$ and $|E| = m$. A **directed graph** is one in which the edges $e_i = (v_j, v_k)$ are ordered pairs – that is, the edge goes from v_j to v_k. An **undirected graph** is one in which the edges $e_i = \{v_j, v_k\}$ are sets of two vertices, in which there is an undirected edge between v_j and v_k (note that this definition does not allow for self-loops). Here we will only consider undirected graphs.

Exercise: Suppose G is an undirected graph on the vertex set V, where $|V| = n$. How many different graph structures are there? Graph $G_1 = (V, E_1)$ is different from $G_2 = (V, E_2)$ (with the same vertex set V) if and only if $E_1 \neq E_2$.
Solution: There are $\binom{n}{2}$ possible edges – one for each pair of distinct vertices. For each edge, there are two possibilities: it is either in the graph or not. So there are $2^{\binom{n}{2}}$ different graph structures.

We denote K_n as the **complete (undirected) graph on n vertices**, which has all $\binom{n}{2}$ edges present. We denote C_n as the **cycle on n vertices**, which has exactly n edges when $n \geq 3$.

Define a **clique** in a graph $G = (V, E)$ as a subset of vertices $U \subseteq V$ in which all vertices of U are connected ($\binom{|U|}{2}$ edges). Define an **independent set** in a graph $G = (V, E)$ as a subset of vertices $U \subseteq V$ in which there is no edge present between any two vertices of U. A **k-clique** is a clique of size k (in terms of vertices). An **ℓ-independent set** is an independent set of size ℓ.

Define $\omega(G) = \max\{|A|: A \subseteq V, A \text{ is a clique in G}\}$, the size of the largest clique in G. We call $\omega(G)$ the **clique number of G**. Define $\alpha(G) = \max\{|B|: B \subseteq V, B \text{ is an independent set in G}\}$, the size of the largest independent set in G. We call $\alpha(G)$ the **independence number of G**.

Exercise: Find $\omega(G)$ and $\alpha(G)$ for the following graphs: $G = K_n$, $G = C_5$.
Solution: $\omega(K_n) = n$ and $\alpha(K_n) = 1$ since all edges are present. $\omega(C_5) = \alpha(C_5) = 2$ (draw a picture).

Define $\mathcal{N}(U)$, where $U \subseteq V$ to be all vertices which are connected by an edge to some vertex of U. We call $\mathcal{N}(U)$ the **neighbors** or **neighboring set** of U. So if $U = \{v\}$ the singleton set with one vertex, $\mathcal{N}(\{v\}) = \{u: \{u, v\} \in E\}$.

3 Ramsey Theory

Ramsey Theory is a very important generalization of the pigeonhole principle.
Very informally, Ramsey's Theorem says that for an arbitrary graph that is "large enough", it will contain at least either a large clique or a large independent set (or both).

Ramsey's Theorem (for graphs): Let $k, \ell \in \mathbb{N}$ and let $G = (V, E)$ be an arbitrary graph with at least $\binom{k+\ell-2}{k-1}$ vertices. Then $\omega(G) \geq k$ or $\alpha(G) \geq \ell$.
Proof: Not given
Ramsey's Theorem implies that there exists some N such that all graphs with N vertices have at least a clique of size k or an independent set of size ℓ. Define $r(k, \ell)$ to be the smallest such value of N (the r is for Ramsey). In other words, $r(k, \ell)$ is the smallest value such that any graph with $r(k, \ell)$ vertices has a k-clique or an ℓ-independent set (or both).

Exercise: Show that $r(3,3) = 6$. That is, any graph on 6 vertices has a 3-clique or 3-independent set.
Solution: We need to show that $r(3,3) \geq 6$ and $r(3,3) \leq 6$. To show the former, consider C_5. Recall that $\omega(C_5) = \alpha(C_5) = 2$, so C_5 is a graph on 5 vertices with no 3-clique nor 3-independent set. Now to show that $r(3,3) \leq 6$, let G be any graph on 6 vertices $\{v_1, \dots, v_6\}$. Consider v_1 – by the pigeonhole principle, it has either at least 3 neighbors or at least 3 non-neighbors.
Case 1: $|\mathcal{N}(\{v_1\})| \geq 3$ (at least 3 neighbors)
Take any 3 of the neighbors, say $\{v_2, v_3, v_4\}$ WLOG (without loss of generality). If there are no edges between these three edges, then they form an independent set of size 3. If there is even one edge between these three edges, say again WLOG $\{v_2, v_3\}$, then $\{v_1, v_2, v_3\}$ form a clique of size 3.
Case 2: $|\mathcal{N}(\{v_1\})| \leq 2$ (at least 3 non-neighbors)
Take any 3 of the non-neighbors, say $\{v_2, v_3, v_4\}$ again WLOG. If all edges are present between them, then $\{v_2, v_3, v_4\}$ form a 3-clique. If there is even one edge missing, say between $\{v_2, v_3\}$ WLOG, then $\{v_1, v_2, v_3\}$ has no edges between them and form a 3-independent set.

Q.E.D.

Exercise: What is $r(k, 2)$?
Solution: $r(k, 2) = k$. If $G = K_{k-1}$, then G has a clique of size $k - 1$ but not k, and has no independent set of size 2, so $r(k, 2) > k - 1$. Now consider any graph on k vertices. If $G = K_k$, G has a clique of size k. Otherwise, there is at least one edge missing, saying from $\{v_1, v_2\}$ and then those two vertices form an independent set.

This is currently an open problem in combinatorics! We know the values of $r(4,3)$ and $r(4,4)$, but no one yet knows what $r(5,5)$ and $r(4,6)$ are.

Theorem (Generalized Ramsey's): Let $m \geq 2$ and $k_1, \dots, k_m \in \mathbb{N}$. Then there exists some minimum number $N = r(k_1, \dots, k_m)$ such that if we color all <u>edges</u> of K_N with colors $1,2, \dots, m$, then there will always exist either: a subgraph K_{k_1} all of whose edges are of color 1, a subgraph K_{k_2} all of whose edges are of color 2, etc.

Exercise: Show that $r(3,3,2) = 6$. That is, if we color edges of K_6 in three colors, say red, green, and blue, there will be a red 3-clique, a green 3-clique, or a blue 2-clique.
Solution: First we show that $r(3,3,2) > 5$. Again, consider C_5 alternating colors red and green. Then there is neither a red 3-clique, a green 3-clique, nor a blue 2-clique. If there is any blue edge at all, say between v_1 and v_2, then those two vertices form a blue 2-clique and we are done. Otherwise, all edges of G are red or green. Since $r(3,3) = 6$, there exists either a red 3-clique or a green 3-clique and we are done.

The Ramsey number $r(k, \ell)$ gives the solution to the party problem: what is the minimum number of guests that we need to invite to a party such that either k of them all know each other or ℓ of them all don't know each other?

This is fascinating in computer science as well – because we work with graphs a lot – and this guarantees in sufficiently large (and probably confusing) graphs, we can find some structure (large cliques/independent sets)!

Ramsey Theory has lots of applications, both in computer science and mathematics. If you are interested, here is where you can find some: http://www.cs.umd.edu/~gasarch/TOPICS/ramsey/ramsey.html.

4 Proofs By Counting

Recall that the **power set** of a set X is $\mathcal{P}(X)$, the set of all subsets of X.
Let X be any set – define the notation $\binom{X}{k} = \{A \in \mathcal{P}(X): |A| = k\}$, the set of all subsets of X of size k.

Consider a finite set X and family (set) of subsets $\mathcal{F} = \{F_1, \dots, F_m\} \subseteq \binom{X}{k}$ of k-subsets of X. We say that $\mathcal{F}$ is **two-colorable** if there exists a coloring of elements of X in (say red and blue) such a way that no set F_i is monochromatic (that is, consists of elements of the same color).

Exercise: Let $X = \{1,2,3\}$. Let $\mathcal{F}_1 = \{\{1,2\},\{1,3\}\}$ and $\mathcal{F}_2 = \{\{1,2\},\{2,3\},\{1,3\}\}$. Is $\mathcal{F}_1$ two-colorable? Is $\mathcal{F}_2$?
Solution: $\mathcal{F}_1$ is two-colorable – color 1 red and $2,3$ blue. $\mathcal{F}_2$ is not. If 1 is red, then 2 has to be blue because of the first subset $\{1,2\}$. Because of the second $\{2,3\}$, 3 must be red. But then $\{1,3\}$ is monochromatic (all red).

By the exercise, it seems like if $\mathcal{F}$ is too large, it may not be two-colorable.

Lemma (Union Bound): Let $A_1, \dots, A_n$ be finite sets. Then $|A_1 \cup \dots \cup A_n| \le \sum_{i=1}^{n} |A_i|$.

Theorem (Erdos): If $\mathcal{F} = \{F_1, \dots, F_m\}$ is a family of k-element subsets of X and if $m < 2^{k-1}$, then $\mathcal{F}$ is two-colorable.
Proof: Let Ω be the set of all two-colorings of X in red and blue. Then $|\Omega| = 2^{|X|}$. (Ex. If $X = \{1,2\}$, then $\Omega = \{rr, rb, br, bb\}$). We want to avoid colorings that make one of the F_i's monochromatic. For $j = 1, \dots, m$, let $A_j^{red} = \{c \in \Omega: c \text{ assigns color red to all elements of } F_j\}$. Define A_j^{blue} similarly.

Notice that the "bad" sets (the ones that make at least one of the F_i's monochromatic) are precisely described by the set:

$$\bigcup_{j=1}^{m}\left(A_j^{red} \cup A_j^{blue}\right)$$

So we are done if we show the following: $\left|\bigcup_{j=1}^{m}\left(A_j^{red} \cup A_j^{blue}\right)\right| < |\Omega|$. We would be done since then there would be some coloring in Ω which **doesn't** assign all elements of any F_j to the same color.

Notice that $\left|A_j^{red}\right| = 2^{|X|-k}$ since all k elements of F_j are red, and we have 2 choices for the other $|X| - k$ elements in $X \setminus F_j$. Similarly, $\left|A_j^{blue}\right| = 2^{|X|-k}$.

So

$$\left|\bigcup_{j=1}^{m}\left(A_j^{red} \cup A_j^{blue}\right)\right| \leq \sum_{j=1}^{m} 2 \cdot 2^{|X|-k} \ [union\ bound]$$
$$= 2m \cdot 2^{|X|-k} < 2 \cdot 2^{k-1} \cdot 2^{|X|-k} \ [inequality\ because\ m < 2^{k-1}]$$
$$= 2^{|X|} = |\Omega|$$

Q.E.D.

Usually to prove existence of something, we need to construct it. This is a very interesting technique because we can show existence of something by counting instead of construction!

5 Conclusion

CSE 312 only covers combinatorial analysis in the beginning, as a way to start thinking about probability. However, even with one week of basic combinatorics, we were able to discuss Ramsey Theory and the Proofs by Counting! Hopefully you can see that the study of combinatorics is interesting and relevant to both probability and computer science. In the future, you will see even more how probability can be applied in computer science: most prominently in machine learning, artificial intelligence, and to make algorithms more efficient, even in other subfields you may not expect!

6 Acknowledgements

Thanks to Martin Tompa for helping me arrange these weekly lectures, and to Isabella Novik for the use of lecture notes in making these notes!