1 Introduction

The first section of this course will consider combinatorics, which, loosely speaking is the study of counting sizes of sets. The questions we can ask of this manner are quire varied – how big are certain sets, how many ways can we arrange elements of certain sets, or choose elements of sets, and so forth. We will later study generating functions and graph theory, which have numerous applications in computer science, operations research, and other fields.

2 Sets

Definition 2.1 (Empty Set). *The empty set, denoted $\emptyset$, is the (only) set containing no elements.*

Definition 2.2. *Let A and B be sets. Then the intersection of A and B, denoted $A \cap B$, is the set of all objects that are in both A and B. The union of A and B, denoted $A \cup B$, is the set of all objects that are in A or in B. A and B are disjoint if $A \cap B$ is the empty set. We denote the number of elements in A by $|A|$.*

If we have a collection of sets, $A_1, \ldots, A_n$, we'll use notation such as $\bigcup_{i=1}^n A_i$ and $\bigcap_{i=1}^n A_i$ to mean the union and intersection of all of them, respectively.

Today, we'll look at ways of stating $|\bigcup_{i=1}^n A_i|$ in terms of each of the $|A_i|$.

3 The Sum Principle

Let's say we have a collection $A_1, \ldots, A_n$ of sets.

We'd like to ask how many objects is in at least one of the sets if we know how many objects are in each of the sets.

In other words, if we know $|A_1|, \ldots, |A_n|$, we'd like to know $|\bigcup_{i=1}^n A_i|$.

This is not always so easy. If there are objects that is in several of the A_i, then calculating $|\bigcup_{i=1}^n A_i|$ may be more tricky. (We'll see how to compute this later).

But, if we know that no object is in more than one of the A_i, then we have the following:

Fact 3.1 (Sum Principle). *If $A_1, \ldots, A_n$ is a collection of sets, so that any two of them are disjoint, then:*

$$\left|\bigcup_{i=1}^n A_i\right| = \sum_{i=1}^n |A_i|$$

This isn't very surprising The above fact by itself, simply states something like "If there are 30 big cities in the U.S and 5 in Canada, then there are 35 in either the U.S. or in Canada".

But we can, with a little more cleverness, prove some more interesting statements. A more interesting way of looking at the sum principle is, given a larger set whose size we wish to determine, how can we break it down into smaller sets that are mutually disjoint, and whose sizes can we easily compute?

3.1 An example: Electing president and Vice president

Suppose that a club has 20 members. We must elect from the club a president and vice-president among its members. How many possible choices are there? Some of you may know the answer is 380, but let us derive the answer from the sum principle.

We can transform this problem into one of finding the size of a union of disjoint sets as follows:

Let's consider the problem as counting the size of the set of all lists of two members from the club, where the first member of each list is the president, and the second is the vice-president.

How can we break this into a union of disjoint sets? We sort the lists into piles so that each pile has all the lists with a given person as president and no other lists. Then we have 20 such piles, and each pile has 19 lists. So there are $20 \cdot 19 = 380$ such arrangements.

4 The Product Principle

In the above example, we applied the sum principle to a large number of sets, each of which were the same size. This situation occurs frequently enough that we give it a name, called the product principle.

Fact 4.1 (Product Principle). *Suppose that we have a collection of m mutually disjoint sets, each of size n. Then there are mn elements in the union of these sets.*

Proof. A direct application of the sum principle. □

In formal notation, we call a list with two elements an *ordered pair*. Using the product principle, we have the following:

Theorem 4.2. *If the set M has m elements and the set N has n elements, then there are nm ordered pairs whose first entry is in M and whose second entry in in N.*

Proof. We split the set of all such ordered pairs into sets, classified by their first element. There will then be m such sets, each with n elements. Using the product principle, we then see that there are mn such sets. □

Definition 4.3. *The set of all ordered pairs whose first element in M and whose second element is in N is the cartesian product of M and N, denoted by $M \times N$.*

This notation is appropriate, in the sense that we know from the above theorem that $|M \times N| = |M| \cdot |N|$.

4.1 Lists of more than two elements

Now, let us generalize the product principle. Let's say that we also need to elect a treasurer for our club.

Then we have lists of three elements. How many ways can we pick the three people? We can sort the lists into piles so that each pile consists of all the lists with a given person as president and as vice president. Then there are 380 piles, each of which contain 18 lists. So we have $20 \cdot 19 \cdot 18$ such lists.

This suggests a more general form of the product principle that we can use to help us count the number of lists of length k.:

Theorem 4.4 (Generalized Product Principle). *Suppose that we have a set of lists S of length k. Suppose that there are m_1 different first elements of lists in S, and when $i-1$ entries for our list have been chosen, there are always m_i possibilities for the next entry. Then S contains $\prod_{i=1}^{k} m_k$ elements.*

Proof. A formal proof would use induction, which will be introduced shortly. Intuitively, we use the product principle. Formal proof left as an exercise. □

5 More List Formulas

Definition 5.1. *The number $n!$, called n factorial, is $1 \cdot 2 \cdot \ldots \cdot n = \prod_{i=1}^{n} i$. We also define $0! = 1$.*

Theorem 5.2. *The number of ways of listing all the elements of an n-element set without repeating any elements is $n!$.*

Proof. By the generalized product principle. Also later on we'll see a more "direct" proof using induction. □

Theorem 5.3. *If S has n elements, then the set of all k-element lists of S (repetition forbidden) is $n!/(n-k)!$.*

Proof. A consequence of the general product principle □

6 Using Induction

So far we have used a lot of "further" ideas. To formalize this concept, and to prove it "really" works, we'll use induction.

Definition 6.1 (Induction). *The principle of mathematical induction works as follows:*

Suppose we have a statement $P(n)$ about the natural number n. To show that $P(n)$ is true for all n, it is enough to show:

1. *$P(1)$ is true. That is, the statement P is true if 1 is used instead of n.*
2. *Whenever $P(k)$ is true, then $P(k+1)$ is true also.*

We can more formally prove some of the above theorems by induction

Example: Show that if S has n elements, then there are $n!$ ways of arranging the members of S:

Proof. By induction on n.

Base Case: If $n = 1$, then there is only one way of arranging the elements of S, and $1! = 1$.

Inductive Step: Suppose the statement is true for all sets of size k (This is the inductive hypothesis). Let S be a set of size $k + 1$.

Consider the set of all arrangements of elements of S. We may break them up into categories by sorting them by their first element. That is, we place them into piles of arrangements, where each pile has all the arrangements with a given object as the first element, and no others. Then by the inductive hypothesis, each pile is a list of all arrangements of a set of k elements (why?). Therefore each pile has $k!$ lists, and there are $(k + 1)$ piles, so by the product principle there are a total of $(k + 1)k! = (k + 1)!$ arrangements. □

7 Stirling's Formula

Here's an approximation of $n!$:

Theorem 7.1 (Stirling's Formula). *$n!$ can be approximated by $\sqrt{2\pi n}\left(\frac{n}{e}\right)^n$. In other words, the ratio $n!/(\sqrt{2\pi n}\left(\frac{n}{e}\right)^n)$ approaches 1 as n approaches infinity.*

8 Lists with Repetitions allowed

Fact 8.1. *There are n^k lists of length k, with repetitions allowed, chosen from a set of n elements.*

Proof. For each position in the list, there are n choices, so the product principle tess us there are n^k lists. □

9 Functions and Relations

Intuitively, we can think of a function that takes an input $x$from one set A, and associates with that input exactly one output $f(x)$ from a set B.

Let's consider the case where $A = \{-1, 0, 1, 2\}$, and $B = \{1, 2, 3, 4\}$. Then the rule $f(x) = x + 2$ is a rule associating each member of A to a member of B. And the rule $g(x) = x^4 - 2z^3 - x^2 + 3x + 2$, is another rule associating each member of A to a member of B. Even further, $f(x) = g(x)$ for all $x \in A$.

So we ask ourselves the question: Are f and g *really* different functions? For our purposes, we simply classify a function by the associations is creates and not by the rule itself. In other words, we view f and g as simply a set of ordered pairs describing the relation. That is, $f = \{(-1,1),(0,2),(1,3),(2,4)\}$, and $g = \{(-1,1),(0,2),(1,3),(2,4)\}$.

Therefore, f and g are the same set, and therefore are the same function.

Now we will precisely define these concepts.

9.1 Relations

Definition 9.1 (Relation). *A relation from A to B is a set of ordered pairs whose first entries are in A and whose second entries are in B. In other words, just a subset of $A \times B$. If (a,b) is in a relation, we say that b is related to a or the relation associates b with a.*

Example: Let $A = B = \{1,2,3,4\}$, and let $S = \{(1,1),(2,2),(3,3),(4,4)\}$. Then S is a relation from A to B, and $(a,b) \in S$ iff $a = b$. We can similarly define a less than or greater than relation.

Definition 9.2 (Function). *f is a function from A to B, denoted $f : A \to B$, if f is a relation from A to B, so that for each $a \in A$, there is a unique $b \in B$ so that $(a,b) \in f$. We say $f(a) = b$ if $(a,b) \in f$. We call A the domain and B the codomain.*

f and g, as we saw above, are examples of functions.

With this more precise definition, we can now ask counting questions about functions and relations.

9.2 Counting Functions

Theorem 9.3. *If A has k elements and B has n elements, then there are n^k functions from $A \to B$.*

Proof. Use the formula for lists with repetitions allowed. Details left as an exercise. □

Definition 9.4. *We denote by B^A the set of all functions from A to B. Therefore, $\left|B^A\right| = |B|^{|A|}$*

Example: We have 4 different chairs to paint and five different colors of paint. In how many ways can we paint the chairs?

Answer: This can be viewed as a function assigning a color to each chair. By the theorem, then, there are $5^4 = 625$ possibilities.

10 One-to-One functions; Pigeonhole Principle

Definition 10.1. *A function $f : A \to B$ is one-to-one or an injection iff $f(a_1) = f(a_2)$ means that $a_1 = a_2$.*

Example: The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^3$ is one-to-one. But the function $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = x^2$ is *not* one-to-one.

In these terms. we can view a list of k elements from a set B (called a k-element permuation), is a one-to-one function from the set $K = \{1, \ldots, k\}$ to B.

Suppose that K was larger than B. Then we cannot have a one-to-one function from K to B – this is called the *pigeonhole principle*: If we have n pigeons and less than n holes to put them in, then some hole will have more than one pigeon.

On the other hand:

Theorem 10.2. *If A is a set with k elements, and B a set with n elements, with $k \leq n$, then there are $\frac{n!}{(n-k)!}$ functions from a k-element set A to an n-element set B.*

Proof. This is the same as lists without repetitions allowed. □

So, if in our example above, we require that each chair be a *different* color, then we have $5 * 4 * 3 * 2 = 120$ chairs. But if we instead had five chairs and four colors, then we can't get the job done at all.

Applications of the pigeonhole principle can be more subtle:

Example: The powers of two are all even, and the powers of five all end in five. Show that for any other prime p, among the first five powers of p, one of these numbers must end in one.

Answer: Since p is not a power of two or a power of five, then for any integer k, p^k cannot end with an even number or five. This means that, there are 5 possible powers of p invovled, and only four possible ending digits. Therefore, by the pigeonhole principle, among the first five powers of p, there are two, call them p^m and p^n, where $n > m$, so that p^m and p^n end in the same digit.

Since $p^n - p^m$ is a multiple of 10, factoring out p^m gives $p^m(p^{n-m} - 1)$ which is the same number and thus also multiple of 10. Since p^m has no multiples of two or five, we know that $p^{n-m} - 1$ is a multiple of 10. Thus, p^{n-m} ends in zero, so p^{n-m} ends in 1.

11 The Extended Pigeonhole Principle

Theorem 11.1 (Extended Pigeonhole Principle). *If f is a function from a set of $mk + 1$ elements to a set of m elements, then there is a set of at least $k + 1$ elements of the domain on which f is constant.*

In other words, if we try to stuff $mk + 1$ pigeons into m holes, then there must be at least one hole with at least $k + 1$ pigeons in it.

11.1 An example – Ramsey Numbers

In a group of n people, either they all know each other, or else there are two that do not know each other. This isn't interesting. *But* did you know that in any group of six people, there is a group of three, *all* of whom know each other, or a group of three, *all* of whom do not know each other (i.e. they are all strangers). (We're assuming that if A knows B, then B knows A, and vice versa)

Here's a proof:

Fact 11.2. *In any group of six people, there are three mutual acuqaintances, or three mutual strangers.*

Proof. Let's single out a person A. Then there are five other people, so by the pigeonhole principle, A must know 3 other people, or must be a stranger with 3 other people. To see this, Apply the Extended pigeonhole principle to the graph where $k = 2$ and $m = 2 - m$ represents the two possibilities of "knowing" and "stranger" $- mk + 1 = 5$ which is the number of people involved "other" than A.

Case 1: A knows three other people: Let B, C, D be three other people that A knows. Then if B, C, and D are all strangers, then this gives us a group of three mutual strangers. If any two of B, C, and D know each other, then these two, taken together with A, form a group of three mutual acquaintances.

Case 2: A is a stranger to three other people: This case works almost identically. If B, C, D are three other people that A is a stranger with, then if B, C, and D are all acquainted with each other, then this is a group of three mutual acquaintances. If not, then there are some two that know each other. Therefore, these two, taken together with A, forms a group of three mutual strangers. □

It turns out that no matter which m and n we choose, we can prove that if we take a large enough group, then no matter which people are in the group, we can find m mutual acquantinaces or n mutual strangers. We denote by $R(m, n)$ the smallest number of people that we need to pick to *guarantee* that there will be m mutual acquaintances or n mutual strangers.

Note that $R(n, 2) = R(2, n) = n$. To see that $R(n, 2) = n$, if we take any group of n people, either all of them know each other, in which case we have a group of n people that all know one another, or some two are strangers, in which case we have a group of 2 people that are all mutual strangers. The proof that $R(2, n) = n$ works similarly.

Note that to prove that $R(m, n)$ exists for all choices of m, n, with $m, n \geq 2$, we need only to prove that for each m, n, we can find some number, K so that if we choose K people, we are guaranteed that we will find some group of m acquaintances or n strangers. We then know that $R(m, n) \leq K$, but do not necessarily know that $R(m, n) = K$.

11.1.1 Double Induction

To perform this proof, we will use a procedure called *Double Induction*

Essentially, here's how we proceed:

We have a statement about two numbers now instead of 1. Let's call it $P(m,n)$. Let's say we wish to show that $P(m,n)$ is true for all m,n with $m \geq K_1$ and $n \geq K_2$. It is then enough to show that:

1. $P(K_1,n)$ is true for all n and $P(n,K_2)$ is true for all n.
2. $P(m,n)$ is true whenever $P(m-1,n)$ and $P(m,n-1)$ are both true.

We'll give an example of this, using our Ramsey numbers. This example will also incorporate an adaptation of the pigeonhole principle.

11.1.2 Proof of Ramsey's Theorem

Theorem 11.3 (Ramsey's Theorem). *$R(m,n)$ actually exists for all m and n with $m,n \geq 2$. Furthermore, $R(m,n) \leq R(m-1,n) + R(m,n-1)$.*

Proof. We know that $R(2,n) = R(n,2) = n$ for all n. So let us consider the case $R(m,n)$ where $m,n \geq 3$. Suppose that $R(m-1,n)$ and $R(m,n-1)$ both exist. Let $k = R(m-1,n) + R(m,n-1)$.

Take any group of k people, and fix a person A from that group. Now, either A knows $R(m-1,n)$ people, or is a stranger to $R(m,n-1)$ other people. This is an adaptation of the pigeonhole principle.

Case 1: A knows $R(m-1,n)$ people. Consider a set of $R(m-1,n)$ people that A knows. Either n of these are mutual strangers or $m-1$ of these are all mutual acquaintances. If the former is true, then we have our group of n strangers. If the latter is true then such a group of $m-1$ acquaintances, taken together with A, forms a group of m mutual acquaintances, since we said that all of these $R(m-1,n)$ people know A.

Case 2: A is a stranger to $R(m,n-1)$ people. So take a set of $R(m,n-1)$ people that are all strangers with A. Either m of these are mutual acquaintances or $n-1$ of these are mutual strangers. If the former is true then we have our group of $R(m,n-1)$ people. If the latter is true then these $n-1$ mutual strangers, taken together with A, gives us our group n mutual strangers.

So, in either case, there are either m mutual acquaintances or n mutual strangers. □

Finding exact values of Ramsey numbers is surprisingly difficult. For instance, it has been proven that $R(3,3) = 6$ and $R(4,4) = 18$. However, the exact value of $R(5,5)$ is not known – we only know that $R(5,5)$ is between 42 and 55.

To illustrate how difficult a task this is, Paul Erdos, one of the more well-known mathemeticians, related the following anecdote:

Aliens invade the earth and threaten to obliterate it in a year's time unless human beings can find the value $R(5,5)$. We could marshal the world's best minds and fastest computers, and within a year we could probably calculate the value. If the aliens demanded the Ramsey number $R(6,6)$, however, we would have no choice but to launch a preemptive attack.