MAT120: Lecture 5 Handout

Number and Counting

Today we will cover our second of three lectures on algebra. In fact, this lecture is a bit of a fun one, because we will go away from algebra slightly and instead study the concepts of "number" and "counting". This might sound like a silly thing to talk about, however I assure you that it is not. By the end of this lecture you will have an understanding of how cutting-edge mathematicians think about these concepts, and I will also introduce probably some of the most profoundly unexpected mathematics that you have ever heard.

1 Counting in Other Number Systems

In previous lectures, I've discussed the difference in language between "symbols" and "meaning" (the fancy words for this are "syntax" and "semantics"). In any language there are symbols, which carry an interpretation. We saw that Propositional Logic was no different: in this "formal language" there were symbols (letters, connectives, brackets) and these had meaning in terms of the truth-value that they carried. In a way, our systems of numbers again fall into this category of a language: there are symbols (1,2,3, and so on) and these carry a meaning, e.g. the symbol "3" means three. In this section we will see just how arbitrary our numerals are.

1.1 A Brief History of Numerals

Throughout history there have been many other ways to represent numbers. The earliest evidence we have suggests that the first representation of numbers was via "tally systems", which are essentially a collection of lines that count quantity. One of the earliest examples of tally systems is from around 20,000-25,000 years ago: the Ishango Bone, discovered in modern-day Congo, is an animal bone with many notches carved into it. A removable quartz crystal sits at one end of the bone, indicating a clear intentional practice. It is theorized that these notches represent the lunar cycle, or perhaps used to track menstrual cycles. Although we will ultimately never know the mysteries of history, we could have some fun and speculate by saying that the first known mathematicians were the women of Africa.

Fast-forwarding through history, many other types of counting systems emerged independently throughout the world. In Ancient China, prototypes of the modern writing system were emerging, again as a type of tallying system to represent smaller quantities. Other systems for writing numbers have been found from Babylon, Ancient Greece, Ancient Rome and India, and they all look very different from each other. However, lots of these systems start off by denoting small numbers in terms of a tally system. After about the number 4 it becomes bothersome and inefficient to represent a number via its quantity, so more abstract symbols were introduced. In the table below, observe the similarity between the writing systems of the ancient world for the numbers $1, 2$ and 3.

Culture	Earliest evidence	1	2	3	4	5
Ishango Bone	~20000 BC	I	II	III	IIII	IIIII
Babylon	~2000 BC	𒁹	𒈫	𒐈	𒐉	𒐊
Egypt (hieratic)	~2500 BC	|	||	|||	||||	ʕ
China (oracle bone)	~1200 BC	一	二	三	亖	X
China (counting rods)	~200 BC	I	II	III	IIII	—
Athens (Attic)	~500 BC	I	II	III	IIII	Π
Rome	~500 BC	I	II	III	IV	V
Mesoamerica	~300 BC	•	••	•••	••••	▬

1.2 Numerical Bases

In the modern world, the most common writing system uses the familiar symbols: $0, 1, 2, 3, 4, 5, 6, 7, 8, 9$. This writing system is called the "Hindu-Arabic numeral system", since it is a mixture of the old Arabic and Indian systems that passed into Europe in the Middle Ages. The symbols that we use to represent numbers are arbitrary – they could have been completely different and the number system would still work just fine. The strength of the Hindu-Arabic numerals is not in the symbols themselves, but in the *system* that underlies them. We will now explore this in some detail.

Consider something like the number 2173. In our heads, we read this as "two thousand, one hundred and seventy three". We can break this down mathematically to see that

$$2173 = 2 \times 1000 + 1 \times 100 + 7 \times 10 + 3.$$

So, in fact, the way in which we represent numbers is in powers of 10. To make this a bit more clear: remember that $1 = 10^0$, and $100 = 10^2$ and $1000 = 10^3$. Then we can rewrite 2173 as a combination of powers of ten:

$$2173 = 2 \times 1000 + 1 \times 100 + 7 \times 10 + 3 = 2 \times 10^3 + 1 \times 10^2 + 7 \times 10^1 + 3 \times 10^0.$$

In fact, when we write our numbers we place them in columns of increasing powers of 10, starting from the right. Such a mathematical system is called "base-10", since 10 is the fundamental number that defines what to do in each column. In a base-10 counting system, there are ten unique symbols for numbers (i.e. $0, 1, \cdots, 9$) that run through a single column.

1.2.1 The Alien

There is a good reason that we perform mathematics within a base-10 system: we humans have 10 fingers, so running across our fingers to count is pretty natural for us humans. However, from a mathematical perspective there is absolutely no special reason to commit to base-10. To see this, let's consider the following idea.

Imagine a distant planet among the stars. There are aliens living on this planet, and these aliens are weird and their bodies are different: they have two arms, but their hands only have 4 fingers on them. We can imagine that this alien species might learn to count in the same way that we learn: by running across their alien fingers and stopping once they get to their final finger before moving to a new column. But, since they only have 4 fingers on each hand, these aliens would stop at their 8th finger, and they would start all over again with the next column in their writing system.

As an example, we will now demonstrate how the aliens would write the number 2173 in a totally different way. Instead of representing 2173 as a sum of powers of 10, they would decompose it into a sum of powers of 8:

$$\begin{aligned} 2173 &= 2048 + 64 + 56 + 5 \\ &= 4 \times 8^3 + 1 \times 8^2 + 7 \times 8^1 + 5 \times 8^0 \end{aligned}$$

Each extra power of 8 gives us a new column to write in and the number we multiply by will give us a numerical representation of the same number that we call "2173". According to the alien, "two thousand one hundred and seventy three" would therefore be written "4175" in their base-8 system.

According to these aliens, they would call us "human weirdos" strange for counting in base-10! They might say something like "glib glob glib glob", or in English "wow those humans count weird... they have two extra symbols after 7 and they use them to keep on counting!". So, who is right here – us or them? The answer is neither of us is using a "universal" system, instead we are just committing to a system that works for us.

1.2.2 Other Bases

Hopefully the previous example demonstrates that there is absolutely no reason to commit to a base-10 system. Generally speaking, we could also pick some other base number, let's call it b, and convert "2173" into a base-b number by writing it out as a sum of increasing powers of b:

$$2173 = a_k \times b^k + a_{k-1} \times b^{k-1} + \cdots + a_1 \times b + a_0 \times b^0.$$

Depending on the size of b, there will be more or less powers of b needed to make it up to 2173. Then, the translation of 2173 into base-b would be: $a_k a_{k-1} \cdots a_1 a_0$ (base-b). Also note that the only possible numbers we use in base-b are $0, 1, \ldots, (b-1)$.

Example

Let's consider the number 211 (base-10), and we will write it in base-5. To write 211 in base-5, we need to first represent it as a sum of increasing powers of 5. Firstly, we recall that $5^2 = 25$, and $5^3 = 125$ and $5^4 = 625$. Since the number 211 is smaller than 5^4, we do not need to go all the way up to the 4^{th} power of 5 in order to represent 211 in base-5. So, we expect our sum to be:

$$\begin{aligned} 211 &= a_3 \times 5^3 + a_2 \times 5^2 + a_1 \times 5^1 + a_0 \times 5^0 \\ &= a_3 \times 125 + a_2 \times 25 + a_1 \times 5 + a_0 \times 1. \end{aligned}$$

From here, we notice that $2 \times 125 = 250$, which is larger than 211, so we will only need to use one multiple of 125 in our sum – this means that $a_3 = 1$. Next, we take the difference $211 - 125 = 86$ and we try to write

$$86 = a_2 \times 5^2 + a_1 \times 5^1 + a_0 \times 5^0.$$

Again, we start with the largest available power, which is $5^2 = 25$. We observe that $4 \times 25 = 100$ yet $3 \times 25 = 75$, so to make up 86 we need to use 3 multiples of $5^2 = 25$. This means that $a_2 = 3$. We now repeat the reasoning by taking the difference $86 - 75 = 11$. We now want to represent $11 = a_1 \times 5^1 + a_0 \times 5^0$. Fortunately, we see that $11 = 2 \times 5 + 1$, which means that $a_1 = 2$ and $a_0 = 1$. Putting this all together, we conclude that

$$\begin{aligned} 211 &= 1 \times 5^3 + 3 \times 5^2 + 2 \times 5^1 + 1 \times 5^0 \\ &= 1 \times 125 + 3 \times 25 + 2 \times 5 + 1 \times 1 \end{aligned}$$

and therefore we write 211 as "1321" in base-5.

1.3 Binary Counting

We will look at the idea of binary counting. The word "binary" ultimately comes from the Latin word "bini", which means something like "two things together". In modern English, this is where we get the prefix "bi" from; used in bicycle, biannual, binoculars, and so on. In our context, a binary counting system is one that counts in base-2. This means that $b = 2$ and therefore there are only two numbers used for counting: 0 and 1. Put differently, in a binary system we only use combinations of 0s and 1s in order to represent all other numbers. For example "11" is how we write "3", and "101" is how we write "5". Of course, every other number that we can usually write in base-10 can also be "translated" into base-2.

1.3.1 Converting Base-10 to Base-2

We saw earlier that

$$2173 = 2 \times 10^3 + 1 \times 10^2 + 7 \times 10^1 + 3 \times 10^0.$$

In this situation, the number 2173 can be described by adding powers of 10 together, and it's a unique description. If we wanted to change to base-2, we would have to replace these "10^n" terms with "2^n" terms. Since 2 is smaller than 10, we can expect that we might need more than 3 powers of 2 to make it all the way up to 2173. As a matter of fact, using a calculator we may confirm that

$$2173 = 2048 + 64 + 32 + 16 + 8 + 4 + 1.$$

Written out more explicitly, we have to go all the way up to the 11th power of 2 in order to express 2173:

$$\begin{aligned} 2173 = 1 \times 2^{11} + 0 \times 2^{10} + 0 \times 2^9 + 0 \times 2^8 + 0 \times 2^7 \\ + 1 \times 2^6 + 1 \times 2^5 + 1 \times 2^4 + 1 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0. \end{aligned}$$

So, to rewrite the number 2173 in base-2, we have to use twelve numbers in a row. The base-2 expression of 2173 is therefore 100001111101 (base-2).

Notice that each multiple in front of the power of 2 is only ever a 1 or a 0. Why? Because if we tried to multiply by a bigger number (like 2), then that would just jump the power of 2 up by a step and get absorbed by the next power of 2.

1.3.2 General Rule

The general rule for converting to base-2 is to try and write the number in decreasing powers of 2. The number in front of each power of 2 will be either a 1 or a 0, and these numbers become the base-2 expression of the base-10 number that you started with.

To convert numbers into base-2 we follow the general method outlined in the example of Section 1.2.2. For instance, suppose that we wanted to write 17 in base-2. The powers of 2 are $1, 2, 4, 8, 16, 32, \ldots$. As we can see, 17 is in between 16 and 32, so we do not need to go all the way up to 32 in order to express 17 in base-2. In fact, we can see that $17 = 16 + 1$, so writing out 17 in terms of all of the base-2 numbers is:

$$17 = 1 \times 2^4 + 0 \times 2^3 + 0 \times 2^2 + 0 \times 2^1 + 1 \times 2^0.$$

Therefore, the binary expression of 17 is 10001.

Exercise

Write these numbers in base-2.

1. 4

2. 7

3. 5

Solution

$$4 = 1 \times 2^2 + 0 \times 2^1 + 0 \times 2^0 \quad \text{so it is written as 100 (base-2)}$$
$$7 = 1 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 \quad \text{so it is written as 111 (base-2)}$$
$$5 = 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 \quad \text{so it is written as 101 (base-2)}$$

1.3.3 The Benefit of Binary

Let's count up to 3 in binary: 0,1,2,3 becomes 00, 01, 10, 11. If we write these out into a row, we suddenly have something that looks awfully like a truth table. Making this a bit more clear: let's pretend that 0 = F and 1 = T. Then "FF, FT, TF, TT" is counting from 0 to 3 in binary!

This is no coincidence: the binary numbers are used in computers as a way to count: 0 = "off" and 1 = "on", which are two states that an electrical signal may be in (high/low voltage). A computer interprets these on/off signals using logic. There are only a small collection of rules for how to combine signals and use their combinations to create new signals, so modern-day computers are actually made up of millions of tiny little machines with a special name: they are called logic gates. There are only a few of these logic gates, and they work in exactly the same way that our truth tables did. That is why we used Torii gates to explain what truth tables are: inside computers there are literally millions of tiny little parts that behave exactly like truth tables. So, this whole idea of base-2 counting is very helpful for computers.

1.4 Other Base Systems

Historically, there were many different types of base systems: the Babylonians, for example, counted in base-60. This may seem silly now, but at the time they were very productive. The Babylonians produced all sorts of mathematical and scientific systems that we still use today, thousands of years later. For example, the Babylonians are the reason that a circle has 360 degrees, and why hours are divided into 60 minutes, and minutes divided into 60 seconds.

Computers also use another type of counting system: base-16, also known as "hexadecimal". The reason for this is that computers like to count in combinations of powers of two, since it is more efficient for them to operate in these terms. In this case, we have a funny problem: since we only count in base-10 usually, we only have 10 symbols for writing numbers, namely 0,1,2,3,4,5,6,7,8 and 9. If we count in base-16, we observe that we have run out of symbols to use. So, we have to use other symbols to represent the numbers from 10–15 in a base-16 system. The most common system is to use capital letters A,B,C,D,E and F, i.e. A=ten, B=eleven and so on. That is why in computer programs like Photoshop, colours are represented with strings of data like "0FA123" – this is actually a number in base-16.

2 Counting Combinations

Suppose that you have just had your breakfast, and you are deciding what to wear for the day. You have a choice between 5 different shirts and 3 different trousers. How many possible outfits can you make from these options?

Let's spell this out with some mathematical notation. We will label the five different shirts by s_1, s_2, s_3, s_4, s_5, and the 3 different trousers by t_1, t_2, t_3. An outfit would consist of a choice of one of each, so it can be written as a pair, for instance (s_3, t_2) would mean that you chose the third shirt and the second pair of trousers. Suppose now that we commit to the first shirt, s_1. We see that there are three options for the choice of outfits, which correspond to the three different trousers: $(s_1, t_1), (s_1, t_2)$ and (s_1, t_3). Similarly, if we were to pick the second shirt instead, then we would also have three possible outfits to build from this, namely $(s_2, t_1), (s_2, t_2)$ and (s_2, t_3). This line of reasoning works for all of the possible shirts, so we can therefore arrange all of the options into a grid:

$$\begin{array}{ccccc} (s_1,t_1) & (s_2,t_1) & (s_3,t_1) & (s_4,t_1) & (s_5,t_1) \\ (s_1,t_2) & (s_2,t_2) & (s_3,t_2) & (s_4,t_2) & (s_5,t_2) \\ (s_1,t_3) & (s_2,t_3) & (s_3,t_3) & (s_4,t_3) & (s_5,t_3) \end{array}$$

As you can see, the total number in this grid is precisely 15, which is simply the number of possible shirts multiplied by the number of possible trousers.

This "multiplication rule" applies in general: if you are making a choice in two stages, then the number of possible outcomes will be given as the product of the choices made in the first stage and the choices made in the second stage. In symbols: if each of the m-many choices in the first step has n-many choices in the second step, then the total number of combinations is given by $m \cdot n$. This also holds for choices made up of more than two stages, as the following example demonstrates.

Example

Consider a propositional formula $(p \wedge q) \to (r \wedge (s \vee t))$. In this case, there are 5 propositional variables that may take a value of either T or F. The truth table for $(p \wedge q) \to (r \wedge (s \vee t))$ considers all possible combinations of Ts and Fs that the propositional variables can take, and then evaluates the truth-value of the formula in each of these situations. In this case, we have two possible "choices" per variable letter, namely T or F, and we have 5 letters in total. Therefore, the total number of possibilities will be $2 \times 2 \times 2 \times 2 \times 2 = 2^5 = 32$. This means that the truth table associated to the formula $(p \wedge q) \to (r \wedge (s \vee t))$ will need 32 rows.

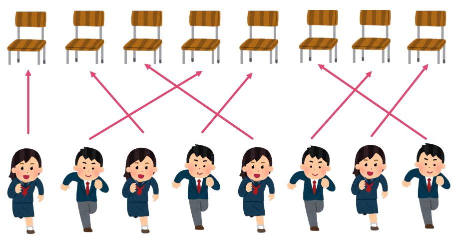
In what follows, we will explore more complicated situations in which successive choices change the number of possibilities, and therefore the number of possible combinations.

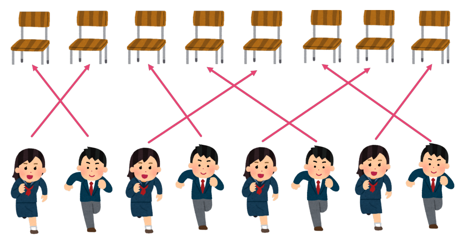
Factorial Notation

Before getting to anything else, we need to first introduce some new notation: the factorial. In short, the factorial is an operation that multiplies a positive integer by every positive integer less than it, going all the way down to 1. For example, "4 factorial" would be $4 \times 3 \times 2 \times 1 = 24$, and "8 factorial" would be $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 40320$. We stop at 1 because if we were to stop at zero then we would be multiplying by zero and therefore the answer would also be zero. We write the factorial notation with an exclamation point, that is, "n factorial" is written $n!$ mathematically. For values of n greater than 2, the factorial $n!$ will always be bigger than n itself, and in fact, usually it will be much bigger.

2.1 Counting Permutations Using 8 Students and 8 Chairs

Suppose that Dr. O'Connell has a new class with 8 students in it. In his classroom there are 8 chairs. Dr. O'Connell decides to create a "seating arrangement" by assigning students to the chairs in the room. For example, the image below contains two possible seating arrangements.





How many ways are there of assigning students to the 8 chairs in the classroom? In other words, how many possible "seating arrangements" are there? We will now calculate this.

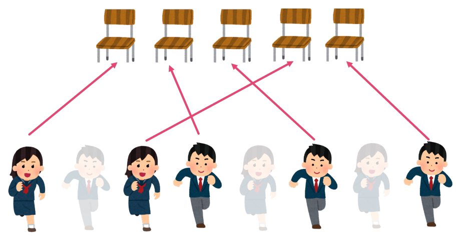
- **The first chair:** since there are 8 students, there are 8 different choices of students to sit in the first chair. In other words: there are 8 possible choices for the first chair. So, Dr. O'Connell selects a student and sits them in the first chair.
- **The second chair:** since one student has just sat down in the first chair, there are now only 7 students that can possibly sit in the second chair. Again, Dr. O'Connell selects a student and sits them down.
- **The general process:** if we repeat this process over and over again, sitting one student in each chair, we see that the process constantly reduces our number of choices by one. Eventually, after 7 chairs have been assigned, there is only one possible choice left: the final student must sit in the last remaining chair. Therefore, the number of possible choices is given by the factorial: $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 8!$

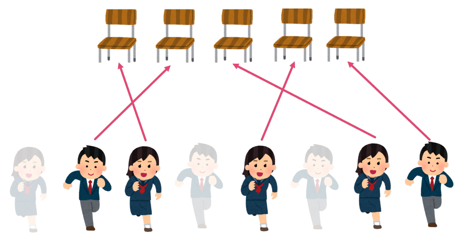
In general, if we have n-many things and we are looking at ways to arrange or rearrange them, then there are $n!$ possible combinations. For example, there are 52 cards in a normal deck of cards. If I shuffle the cards, I have created a new arrangement in which the cards are in a different order. How many different ways are there to shuffle these cards? The answer is 52! which is a number so overwhelmingly large that it's hard to imagine. The number 52! is approximately equal to 8.07×10^{67}. For reference: the number of grains of sand on the Earth is "only" 7.6×10^{18}, the number of stars in the universe is about 4×10^{23}, and the number of atoms in a glass of water is 2.5×10^{25}.

2.2 Counting Permutations Using 8 Students and 5 Chairs

Suppose now that Dr. O'Connell only has 5 chairs in the classroom. So, if he tries to sit 8 students down, then only 5 of them will be able to sit down and the other 3 of the students will have to stand.

Our next question is: what is the number of possible ways that we can sit 5 of the 8 students down? Two such arrangements are depicted below.





We can repeat the same reasoning as before, and we see that the first chair has 8 choices, the second chair has 7 choices and so on. However, since there are now only 5 possible chairs to assign, we must stop our procedure after 5 chairs instead of going all the way to 8. This means that our answer will not be 8!, because we don't care about how those last 3 students are arranged. Instead, the answer will only multiply the first 5 terms, because there are only 5 chairs available. The answer is: $8 \times 7 \times 6 \times 5 \times 4$ possible seating arrangements.

Generally speaking, let n be the number of students and r be the number of chairs. The number of ways that we can arrange these n-many students into r-many chairs can be calculated with the formula:

$$\frac{n!}{(n-r)!}.$$

The reason that we divide by $(n-r)!$ is because the $n!$ term counts up more than what we want: we don't care about ordering all of the $(n-r)$-many students that have to stand up. So, we divide by $(n-r)!$ to*remove* these combinations from the overall count. To double-check, let $n = 8$ and $r = 5$. Then:

$$\frac{n!}{(n-r)!} = \frac{8!}{3!} = \frac{8 \times 7 \times 6 \times 5 \times 4 \times (3 \times 2 \times 1)}{(3 \times 2 \times 1)} = 8 \times 7 \times 6 \times 5 \times 4,$$

which is the same answer that we got before. These "rearrangements" are actually called "permutations" in mathematics. The number of all permutations of n-many objects into r-many positions is written $P_{n,r}$, or sometimes $P(n,r)$, and the formula is given by:

$$P(n,r) = \frac{n!}{(n-r)!}.$$

2.3 Counting Choices

When discussing permutations, we cared about the order in which we arranged the students. That is, Dr. O'Connell would line up all the chairs in order and assign different students by saying "you go first, you go second, ...". We can imagine that there might be another way to do things: more generally, perhaps Dr. O'Connell could simply select r-many students and not care in which order they are placed in chairs. This would amount to simply choosing 5 students from the collection of 8. Two such choices are pictured below.

In fact, we can also count up "all the possible ways to choose r-many objects from n-many". These choices aren't permutations, because in this situation we don't care about their order. Instead, the choices are just picking some members of the group and that's it.

2.3.1 Formula for Combinations

There is a formula for all the possible choices of r-many objects amongst a collection of n-many. In mathematics, we call all of these unordered choices "combinations". We use the notation $C_{n,r}$ to denote the number of ways to choose r-many things from a collection of size n. The formula for $C_{n,r}$ is very similar to that of $P(n, r)$. It is given by:

$$C_{n,r} = \frac{n!}{r!(n-r)!}.$$

Here, we basically have the formula for $P(n, r)$ but we divide by an extra $r!$ term. Why do we do this? Well, in the case of $P(n, r)$ we are asking for ordered combinations, and the same choice of r-many objects can be ordered in $r!$ many different ways. However, according to unordered combinations, we don't actually care about those $r!$ many things, so we divide the overall formula by $r!$ to remove all of those choices from our count.

2.3.2 Pascal's Triangle

For low values of n and r, we can arrange the values of $C_{n,r}$ into a triangle. At the top of the triangle is $C_{0,0} = 1$, the next row is $C_{1,0} = 1$ and $C_{1,1} = 1$. But, things start to get interesting once we add more rows. In the figure below, the values for each $C_{n,r}$ is written in the form of a triangle.

Row 0:	1	$C_{0,0}$
Row 1:	1 1	$C_{1,0}$ $C_{1,1}$
Row 2:	1 2 1	$C_{2,0}$ $C_{2,1}$ $C_{2,2}$
Row 3:	1 3 3 1	$C_{3,0}$ $C_{3,1}$ $C_{3,2}$ $C_{3,3}$
Row 4:	1 4 6 4 1	$C_{4,0}$ $C_{4,1}$ $C_{4,2}$ $C_{4,3}$ $C_{4,4}$
Row 5:	1 5 10 10 5 1	$C_{5,0}$ $C_{5,1}$ $C_{5,2}$ $C_{5,3}$ $C_{5,4}$ $C_{5,5}$
⋮	⋮	⋮

Observe that each number in the triangle is given by the sum of the two numbers above it. For example, the "6" in Row 4 (coloured red) is given by the sum of the "3" and the "3" that are above it in Row 3. By applying this addition rule, we can form as many new rows of the triangle as we like.

Observe that each row is numbered by the "n" in $C_{n,r}$, and the entries for that row are given by running across the different values of r between 0 and n. This means that we don't need to calculate the value of $C_{n,r}$ – instead we can simply read off its value using the triangle.

3 Counting Subsets

A few lectures ago we saw an example of God lining up all of the objects in the world and placing them into boxes. For instance, God has a box full of all of the guitars, a box full of all of the pickles, and so on. We used this to motivate a discussion of “quantifiers” which are little language tools that allow us to “look inside these boxes and check things out”. As a matter of fact, this analogy was left intentionally informal. In reality, the boxes in our mind don’t have clear and precise “edges” to them, so it’s hard to properly categorize objects. For example: what goes in the box of all chairs? Is a stool a chair? Is a drawing of a chair a chair? Is a tree stump a chair? What about broken chairs?

In mathematics, everything is precisely defined in an idealized way. So, we have the benefit of avoiding ill-defined concepts and we have definite, well-defined boxes that we can use to categorize objects. In the simplest interpretation (the one I will present), these “boxes” are called “sets”, and they are the foundation of modern mathematics.

3.1 What is a Set?

Naively speaking, a set is just an unordered collection of objects, much like a box. Sets have essentially no structure, and they are only defined by the elements that they contain. We write sets with curly brackets $\{$ and $\}$. So, the set containing the numbers $2, 4$ and 5 would be written $\{2, 4, 5\}$. Order doesn’t matter at all, so we can write the same set as $\{4, 2, 5\}$ or $\{5, 2, 4\}$. If an object is a member of a set, we call it an “element” of the set. So, in the example I just gave, the elements of the set $\{2, 4, 5\}$ are $2, 4$ and 5 (and that’s it).

3.2 Rules for Sets

Sets have many rules and interesting properties, and the study of sets and their rules is called “set theory”. It’s a very interesting topic, however for this class we are only interested in counting these things, so we won’t be able to talk too much about set theory. But, for our purposes, there are a few important rules that we need to talk about.

Rule 1: sets are only defined by their objects, and the same object can’t appear twice. This means that the set $\{1, 1, 2\}$ is the same as the set $\{1, 2\}$. So our first rule is that we don’t write the same object multiple times.

Rule 2: there is a fundamental object, which is the set containing nothing at all – informally, you can imagine it like this: $\{\}$. This is called the “empty set”, because there is nothing inside of it. We denote the empty set using the symbol $\emptyset$.

Rule 3: sets can contain basically any mathematical objects, including sets themselves! We write this using multiple brackets: the notation $\{1, 2, \{3, 4\}\}$ means a set with three elements: 1, 2 and the set $\{3, 4\}$.

Rule 4: a set cannot contain itself as an element.

In reality, there are actually quite a lot of rules for sets – the interested reader should look into *axiomatic set theories* such as Zermelo-Frankel set theory. However, since we are only discussing things very lightly, we will only commit to an informal treatment of sets.

3.3 Counting Subsets

As the name may suggest, a "subset" is another set that contains only some of the elements we started with. To be a bit more precise: a set X is a subset of a set Y if every element of X is also in Y. For example, the set $\{2, 4\}$ is a subset of $\{2, 4, 5\}$.

Exercise

Let $X = \{a, b, c, d\}$. Are the following sets subsets of X or not? Answer with a yes or no.

1. $\{a, b\}$
2. $\{c, e\}$
3. $\{a, b, c, d\}$
4. $\emptyset$

Solutions: Yes, no, yes, yes.

From any set X we can make another set, which is the collection of all of its subsets. For example, if we have a set $X = \{1, 2, 3\}$ then there are 8 subsets in total:

$$\emptyset,\ \{1\},\ \{2\},\ \{3\},\ \{1,2\},\ \{1,3\},\ \{2,3\},\ \{1,2,3\}.$$

The power set of X, which we write $\mathcal{P}(X)$, is the set containing these 8 subsets.

How should we count the size of the power set? It seems like the power set of X will always be bigger than X itself, since we are looking at all the possible subsets. As a matter of fact, the size of a power set will always be 2 raised to the power of the size of X. In the previous example, $X = \{1, 2, 3\}$ has size=3. The power set of X has 8 elements, i.e. $8 = 2^3$. This works in general: a set of size 4 has a power set of size $2^4 = 16$, and a set of size 1000 has a power set of size 2^{1000}.

3.3.1 How to Count Subsets of a Given Size

In a way, picking a subset of a particular size is just like choosing r-many elements amongst n-many choices. Sets don't care about ordering, so in fact we can use the general rule: if a set X has size n, then the number of subsets of X of size r is given by the formula $C_{n,r}$. If we take $C_{n,r}$ and add up all these numbers for all the possible choices of r, then we will get 2^n. This becomes more clear once we sum along the rows in Pascal's triangle.

Row 0:						1						= 1
Row 1:					1	+	1					= 2
Row 2:				1	+	2	+	1				= 4
Row 3:			1	+	3	+	3	+	1			= 8
Row 4:		1	+	4	+	6	+	4	+	1		= 16
Row 5:	1	+	5	+	10	+	10	+	5	+	1	= 32
⋮						⋮						⋮

As you can see, the sums of each row of Pascal's triangle give us increasing powers of 2, which is one way to describe the size of power sets. To summarise, given a set X of size n,

- The number of subsets of X with size r is given by $C_{n,r}$, which can be read directly from Pascal's triangle by looking at the r^{th} term in the n^{th} row.[1]
- The total number of all subsets of X is given by 2^n, which can be calculated by summing all of the terms that appear in row n of Pascal's triangle.

Example

Consider a set $X = \{a, b, c, d\}$, which is a set of size 4. We can use $C_{4,r}$ to count all of the subsets of size r, where r ranges from 0 to 4.

- $C_{4,0} = 1$ counts the single subset of size 0, namely the empty set $\emptyset$.
- $C_{4,1} = 4$ counts the subsets of size 1, namely the sets $\{a\}$, $\{b\}$, $\{c\}$ and $\{d\}$.
- $C_{4,2} = 6$ counts the subsets of size 2, namely the sets $\{a, b\}$, $\{a, c\}$, $\{a, d\}$, $\{b, c\}$, $\{b, d\}$, and $\{c, d\}$.
- $C_{4,3} = 4$ counts the subsets of size 3, namely the sets $\{a, b, c\}$, $\{a, b, d\}$, $\{a, c, d\}$, and $\{b, c, d\}$.
- $C_{4,4} = 1$ counts the single subset of size 4, which is the entire set X itself.

Observe that the sum of all of the $C_{4,r}$ terms adds to: $1+4+6+4+1 = 16$, which is exactly 2^4.

3.4 To Infinity and Beyond!

We are now going to talk about something special: infinity. So far we've been talking about the size of sets on a really simple level, e.g. $X = \{a, b, c\}$ is a set of size 3, or $Y = \{2, 3, 5\}$ is also a set of size 3. Now we are going to consider the set of *all* natural numbers, i.e. the set

$$\mathbb{N} = \{0, 1, 2, 3, \cdots\}.$$

Clearly this set is different from any finite set – it's somehow "infinitely big". This shouldn't be too controversial: clearly the set $\mathbb{N}$ cannot be of finite size, since it contains *all* of the natural numbers, $\mathbb{N}$ must have a size that is bigger than every finite number.

[1] Here, we start counting rows from 0, i.e. row n has $(n+1)$-many numbers in it, corresponding to $r = 0, 1, 2, \cdots, n-1, n$.

Consider now the following question: what about the power set of $\mathbb{N} = \{0, 1, 2, 3, \dots\}$? How big is that?

Now, the answer is weird, but correct: the power set of $\mathbb{N}$, written $\mathcal{P}(\mathbb{N})$ is also infinitely big. But, it is literally bigger in size than $\mathbb{N}$ itself, and we can prove this mathematically. The argument is actually quite simple, so I will repeat it now.

3.4.1 A Diagonalization Argument

Let's suppose that $\mathbb{N}$ and its power set are actually the same size. This means that if I take the contents of both sets, and line them up, then I can match up the contents of these sets one-to-one. So, let's do this: recall that $\mathbb{N} = \{0, 1, 2, 3, \dots\}$, and let's write its power set out:

$$\mathcal{P}(\mathbb{N}) = \{A_0, A_1, A_2, A_3, \dots\}.$$

Each one of these A_n is a subset of $\mathbb{N}$, so it is itself a set that has a bunch of numbers in it. So, we can go through the numbers in $\mathbb{N}$ and ask "is this particular number in the particular subset A_n"? The answer will always be either yes or no. This means that we can create a giant table containing all of this information about the contents of each subset A_n. For illustration purposes, let's suppose that

$$\begin{aligned} A_0 &= \{2, 3, 4\} \\ A_1 &= \{0, 2, 4, 6, 8, \cdots\} \\ A_2 &= \{1, 2, 3, 9, 27, 1000\} \\ A_3 &= \{3\} \\ A_4 &= \emptyset \\ &\vdots \end{aligned}$$

Then, our table will look something like this:

	0	1	2	3	4	$\cdots$
A_0	×	×	✓	✓	✓	$\cdots$
A_1	✓	×	✓	×	✓	$\cdots$
A_2	×	✓	✓	✓	×	$\cdots$
A_3	×	×	×	✓	×	$\cdots$
A_4	×	×	×	×	×	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Now here is where the magic happens: we are going to create a new subset B of $\mathbb{N}$ that is not contained in the table. Since B is a subset of $\mathbb{N}$, we are going to define it by specifying its numbers one by one. Firstly, we look in the first column of the first row. If the answer is "no" then that means that 0 is not in the subset A_0. So, we decide to include "0" in our new subset B. Therefore, B is different from A_0 because A_0 doesn't contain 0 but B does. Now, we move down one row and across one column. Again, there is a "no" in this column, which means that the number 1 is not present in the subset A_1. So, again, we will intentionally include the number 1 in our set B. This means that, by construction, B is different from A_1 since they don't contain exactly the same numbers. Moving down one spot and across to the right, we see that 2 is indeed contained in the subset A_2. So, we intentionally exclude the number 2 from B. This guarantees that B is also different from A_2.

We can repeat this process indefinitely by going down the diagonal line of the table, and we keep adding or excluding elements to/from B to force B to be different from every $A_0, A_1, A_2, \ldots$. The result is a new subset B that is ultimately different from every subset in the table.

	0	1	2	3	4	$\cdots$
A_0	×	×	✓	✓	✓	$\cdots$
A_1	✓	×	✓	×	✓	$\cdots$
A_2	×	✓	✓	✓	×	$\cdots$
A_3	×	×	×	✓	×	$\cdots$
A_4	×	×	×	×	×	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

$\longrightarrow$

	0	1	2	3	4	$\cdots$
B	✓	✓	×	×	✓	$\cdots$

But, the list $A_0, A_1, A_2, \ldots$ was supposed to contain *all* of the subsets of $\mathbb{N}$! How is this possible? The answer is that it is not possible, so our original assumption that $\mathbb{N}$ matches up one-to-one with $\mathcal{P}(\mathbb{N})$ must be false. In other words: $\mathcal{P}(\mathbb{N})$ is strictly larger than $\mathbb{N}$. We have just demonstrated that there are infinitely-big objects that are literally larger than other infinitely-big things.