Math/CS 120: Intro. to Math — Professor: Padraic Bartlett

Lecture 3: Sizes of Infinity

Week 2 — *UCSB 2014*

1 Sizes of Infinity

On one hand, we know that the real numbers contain "more" elements than the rational numbers: things like $\sqrt{2}$ are in $\mathbb{R}$ but not in $\mathbb{Q}$, for example. But how much more?

In this lecture, we discuss how we can come up with a rigorous way of studying the above question. Let's start with the most basic thing we can ask: what does it mean for two sets to be the same size? In the finite case, this question is rather trivial; for example, we know that the two sets

$$A = \{1, 2, 3\}, \qquad\qquad B = \{A, B, \text{emu}\}$$

are the same size because they both have the same *number* of elements – in this case, 3.

But what about infinite sets? For example, look at the sets

$$\mathbb{N}, \qquad \mathbb{Q}, \qquad \mathbb{R}, \qquad \mathbb{C};$$

are any of these sets the same size? Are any of them larger? By how much?

In the infinite case, the tools we used for the finite – counting up all of the elements – don't work. In response to this, we are motivated to try to find another way to count: in this case, one that involves **functions**.

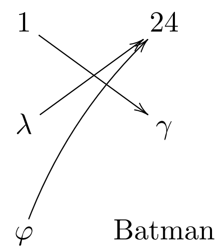
1.1 Functions (formally defined)

Recall the definition we made for functions in our last set of notes:

Definition. A **function** f with domain A and codomain B, formally speaking, is a collection of pairs (a, b), with $a \in A$ and $b \in B$, such that there is exactly one pair (a, b) for every $a \in A$. More informally, a function $f : A \to B$ is just a map which takes each element in A to some element of B.

Examples.

- $f : \mathbb{Z} \to \mathbb{N}$ given by $f(n) = 2|n| + 1$ is a function.
- $g : \mathbb{N} \to \mathbb{N}$ given by $f(n) = 2|n| + 1$ is also a function. It is in fact a different function than f, because it has a different domain!
- The function h depicted below by the three arrows is a function, with domain $\{1, \lambda, \varphi\}$ and codomain $\{24, \gamma, \text{ Batman}\}$:

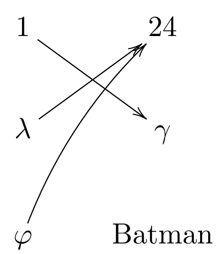


This may seem like a silly example, but it's illustrative of one key concept: functions are just **maps between sets!** Often, people fall into the trap of assuming that functions have to have some nice "closed form" like $x^3 - \sin(x)$ or something, but that's not true! Often, functions are either defined piecewise, or have special cases, or are generally fairly ugly/awful things; in these cases, the best way to think of them is just as a collection of arrows from one set to another, like we just did above.

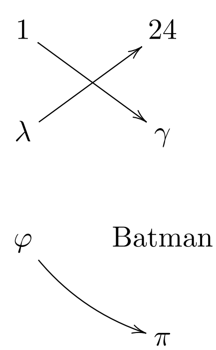
Now that we've formally defined functions and have a grasp on them, let's introduce a pair of definitions that will help us with our question of "size:"

Definition. We call a function f **injective** if it never hits the same point twice – i.e. for every $b \in B$, there is **at most one** $a \in A$ such that $f(a) = b$.

Example. The function h from before is not injective, as it sends both λ and φ to 24:



However, if we add a new element π to our codomain, and make φ map to π, our function is now injective, as no two elements in the domain are sent to the same place:



One observation we can quickly make about injective functions is the following:

Proposition. If $f : A \to B$ is an injective function and A, B are finite sets , then $\text{size}(A) \leq \text{size}(B)$.

The reasoning for this, in the finite case, is relatively simple:

1. If f is injective, then each element in A is being sent to a different element in B.
2. Thus, you'll need B to have at least $|A|$-many elements to provide that many targets.

A converse concept to the idea of injectivity is that of **surjectivity**, as defined below:

Definition. We call a function f **surjective** if it hits every single point in its codomain – i.e. if for every $b \in B$, there is **at least one** $a \in A$ such that $f(a) = b$.

Alternately: define the **image** of a function as the collection of all points that it maps to. That is, for a function $f : A \to B$, define the image of f, denoted $f(A)$, as the set $\{b \in B \mid \exists a \in A \text{ such that } f(a) = b\}$.

Then a surjective function is any map whose image is equal to its codomain: i.e. $f : A \to B$ is surjective if and only if $f(A) = B$.

Example. The function h from before is not injective, as it doesn't send anything to Batman:

$$\begin{tikzcd}
1 \arrow[dr] & 24 \\
\lambda \arrow[ur] & \gamma \\
\varphi \arrow[uur, bend left] & \text{Batman}
\end{tikzcd}$$

However, if we add a new element ρ to our domain, and make ρ map to Batman, our function is now surjective, as it hits all of the elements in its codomain:

$$\begin{tikzcd}
1 \arrow[dr] & 24 \\
\lambda \arrow[ur] & \gamma \\
\varphi \arrow[uur, bend left] & \text{Batman} \\
\rho \arrow[ur, bend right] &
\end{tikzcd}$$

As we did earlier, we can make one quick observation about what surjective functions imply about the size of their domains and codomains:

Proposition. If $f : A \to B$ is an surjective function and A, B are finite sets , then $|A| \geq |B|$.

Basically, this holds true because

1. Thinking about f as a collection of arrows from A to B, it has precisely $|A|$-many arrows by definition, as each element in A gets to go to precisely one place in B.

2. Thus, if we have to hit every element in B, and we start with only $|A|$-many arrows, we need to have $|A| \geq |B|$ in order to hit everything.

So: in the finite case, if $f : A \to B$ is injective, it means that $|A| \leq |B|$, and if f is surjective, it means that $|A| \geq |B|$. This motivates the following definition and observation:

Definition. We call a function **bijective** if it is both injective and surjective.

Proposition. If $f : A \to B$ is an bijective function and A, B are finite sets , then $|A| = |B|$.

Unlike our earlier idea of counting, this process of "finding a bijection" seems like something we can do with any sets – not just finite ones! As a consequence, we are motivated to make this our **definition** of size! In other words, we have the following definition:

Definition. We say that two sets A, B are the same size (formally, we say that they are of the same **cardinality**,) and write $|A| = |B|$, if and only if there is a bijection $f : A \to B$.

1.2 The Natural Numbers

Armed with a definition of size that can actually deal with infinite sets, let's start with some calculations to build our intuition:

Question. Are the sets $\mathbb{N}$ and $\mathbb{N} \cup \{\text{lemur}\}$ the same size?

Answer. Well: we know that they can be the same size if and only if there is a bijection between one and the other. So: let's try to make a bijection! In the typed notes, the suspense is somewhat gone, but (at home) imagine yourself taking a piece of paper, and writing out the first few elements of $\mathbb{N}$ on one side and of $\mathbb{N} \cup \{\text{lemur}\}$ on the other side. After some experimentation, you might eventually find yourself with the following map:

$$
\begin{array}{ccc}
\boxed{\mathbb{N}} & & \boxed{\mathbb{N} \cup \{\text{lemur}\}} \\
1 & \longrightarrow & \text{lemur} \\
2 & \longrightarrow & 1 \\
3 & \longrightarrow & 2 \\
4 & \longrightarrow & 3 \\
5 & \longrightarrow & 4 \\
6 & \longrightarrow & 5 \\
\vdots & & \vdots
\end{array}
$$

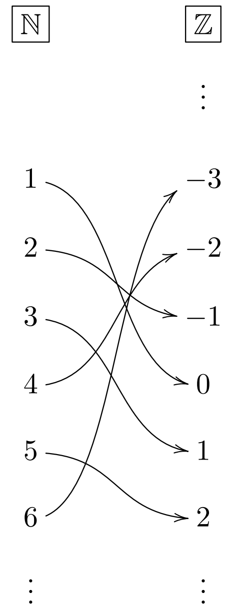
i.e. the map which sends 1 to the lemur and sends $n \to n - 1$, for all $n \geq 2$. This is clearly a bijection; so these sets are the same size!

In a rather crude way, we have shown that adding one more element to a set as "infinitely large" as the natural numbers doesn't do anything to it! – the extra element just gets lost amongst all of the others.

This trick worked for one additional element. Can it work for infinitely many? Consider the next proposition:

Proposition. The sets $\mathbb{N}$ and $\mathbb{Z}$ are the same cardinality.

Proof. Consider the following map:



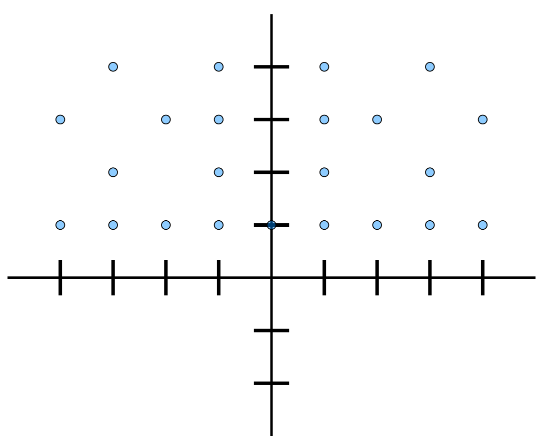
i.e. the map which sends $n \to (n-1)/2$ if n is odd, and $n \to -n/2$ if n is even. This, again, is clearly a bijection; so these sets are the same cardinality. $\square$

So: we can in some sense "double" infinity! Strange, right? Yet, if you think about it for a while, it kind of makes sense: after all, don't the natural numbers contain two copies of themselves (i.e.the even and odd numbers?) And isn't that observation just what we used to turn $\mathbb{N}$ into $\mathbb{Z}$?

After these last two results, you might be beginning to feel like all of our infinite sets are the same size. In that case, the next result will hardly surprise you:

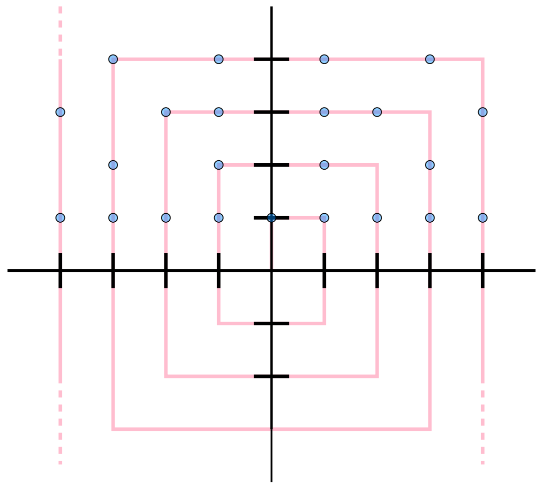
Proposition. The sets $\mathbb{N}$ and $\mathbb{Q}$ are the same cardinality.

Proof. First, take every rational number p/q with $GCD(p,q) = 1$, $p > 0$, and draw a point at (p,q) in the integer lattice $\mathbb{Z}^2$:



In the picture on the previous page, every rational number has exactly one unique representative by one of our blue dots.

Now, on this picture, draw a spiral that starts at $(0,0)$ and goes through every point of $\mathbb{Z} \times \mathbb{Z}$, as depicted below:



We use this spiral to define our bijection from $\mathbb{N}$ to $\mathbb{Q}$ as follows:

$f(n) =$ *the n-th rational point found by starting at (0,0) and walking along the depicted spiral pattern.*

This function hits every rational number exactly once by construction; thus, it is a bijection from $\mathbb{N}$ to $\mathbb{Q}$. Consequently, $\mathbb{N}$ and $\mathbb{Q}$ are the same size. $\square$

1.3 The Reals

At this point, it almost seems inevitable that **every** infinte set will wind up having the same size!

This is false.

Theorem. The sets $\mathbb{N}$ and $\mathbb{R}$ have different cardinalities.

Proof. (This is Cantor's famous diagonalization argument.) Suppose not – that they were the same cardinalities. As a result, there is a bijection between these two sets! Pick such a bijection $f : \mathbb{N} \to \mathbb{R}$.

For every $n \in \mathbb{N}$, look at the number $f(n)$. It has a decimal representation. Pick a number $a_{n,\text{trash}}$ corresponding to the integer part of $f(n)$, and $a_{n\_1}, a_{n\_2}, a_{n\_3}, \ldots$ that correspond to the digits after the decimal place of this decimal representation – i.e. pick numbers a_{n_i} such that

$$f(n) = a_{n\_\text{trash}}.a_{n\_1}a_{n\_2}a_{n\_3}\ldots$$

For example, if $f(4) = 31.125$, we would pick $a_{4\_\text{trash}} = 31, a_{4\_1} = 1, a_{4\_2} = 2, a_{4\_3} = 5$, and $0 = a_{4\_4} = a_{4\_5} = a_{4\_6} = \ldots$, because the integer part of $f(4)$ is 31, its first three digits after the decimal place are 1,2, and 5, and the rest of them are zeroes.

Now, get rid of the $a_{n_{\text{trash}}}$ parts, and write the rest of these numbers in a table, as below:

$$\begin{array}{c|ccccc}
 & & & & & \\
\hline
f(1) & a_{1_1} & a_{1_2} & a_{1_3} & a_{1_4} & \cdots \\
f(2) & a_{2_1} & a_{2_2} & a_{2_3} & a_{2_4} & \\
f(3) & a_{3_1} & a_{3_2} & a_{3_3} & a_{3_4} & \\
f(4) & a_{4_1} & a_{4_2} & a_{4_3} & a_{4_4} & \\
\vdots & \vdots & & & & \ddots
\end{array}$$

In particular, look at the entries $a_{1\_1}a_{2\_2}a_{3\_3}\dots$ on the diagonal. We define a number B using these digits as follows:

- Define $b_i = 2$ if $a_{i\_i} \neq 2$, and $b_i = 8$ if $a_{i\_i} = 2$.
- Define B to the be the number with digits given by the b_i – i.e.

$$B = .b_1b_2b_3b_4\dots$$

Because B has a decimal representation, it's a real number! So, because our function f is a bijection, it must have some value of n such that $f(n) = B$. But the n-th digit of $f(n)$ is $a_{n,n}$ by construction, and the n-th digit of B is b_n – by construction, these are different numbers! So $f(n) \neq B$, because they disagree at their n-th decimal place!

This is a contradiction to our original assumption that such a bijection existed. Therefore, we know that no such bijection can exist: as a result, we've shown that the natural numbers are of a strictly "different" size of infinity than the real numbers. □

Crazy! This is how in a sense about how half of the mathematics we do as researchers goes: by combining relatively mundane ideas (the concepts of bijection and proof by contradiction) we can get to remarkably strange results (there are different sizes of infinity!)

This is one of the beautiful things of mathematics; it's surprising! However, it's not always surprising in the way that we've just seen. Instead, some results flip the script from above on its head; instead of proving surprising results using simple methods, we sometimes have to create tricky proofs to prove seemingly-obvious statements! (In a sense, you can think of your proofs as telling you that "obviousness" is not an easy quality to judge in a problem; plenty of horrible-looking claims have remarkably elegant proofs, while many simple problems[1] are completely open!)

We finish this section with an example of such a result. Recall, from your very early days of elementary school, the following theorem:

Theorem. Suppose that a and b are a pair of natural numbers such that $a \leq b$ and $b \leq a$. Then $a = b$.

Given that we've built all of our intuition for sets from the finite setting, we might hope that this result would hold true for cardinalities! In other words, we would hope for the following result:

[1] Like the Collatz conjecture!

Theorem. (Cantor-Schroeder-Bernstein) Suppose that A and B are a pair of sets such that there is an injection $f : A \to B$ and another injection $g : B \to A$. Then $|A| = |B|$: in other words, there is a bijection from A to B.

This is true! However, its proof is surprisingly nonobvious. Most people, when presented with this theorem on (say) their homework, construct the following "proof:"

1. Because there is an injection from A to B, we know that $|A| \leq |B|$.
2. Similarly, because there is an injection from B to A, we know that $|B| \leq |A|$.
3. Therefore, because $|A| \leq |B|$ and $|B| \leq |A|$, we can conclude that $|A| = |B|$.

You may have noticed that the word "proof" is in quotes above: this is because the above outline doesn't prove anything! (Take a second to think about why that is before reading on.)

So: why is this not a proof? Well: let's consider point 1. What do we mean when we say that $|A| \leq |B|$? For finite sets this is easy: it means that the number represented by $|A|$ is no greater than the number represented by $|B|$. For infinite sets, however, this is tricky! We have a notion of "inequality" for sizes of infinite sets — two sets have different cardinalities if and only if it is impossible to construct a bijection between them — but we actually don't have a notion for which set is **larger**.

We can fix this by making a definition:

Definition. Given any two sets A, B, we say that A has cardinality less than B, and write $|A| < |B|$, if and only A and B are different cardinalities, and also there is an injection from A into B.

This captures the intuitive idea we have for inequalities — the natural numbers should be a "smaller" set than the reals, because on one hand they are sets of different cardinalities, while on the other the natural numbers are a subset of the reals! So this is a good definition, and with it we've made points 1 and 2 in our "proof" of Cantor-Schroeder-Bernstein completely airtight.

The issue, as it turns out, is in point 3. When we write

$$(|A| \leq |B|) \wedge (|B| \leq |A|) \Rightarrow (|A| = |B|),$$

we are actually assuming precisely what we want to show: that if we have injections from A to B and from B to A, there is a bijection from A to B! It seems plausible, because for natural numbers we are used to

$$(x \leq y) \wedge (y \leq x) \Rightarrow (x = y)$$

being a valid deduction. But this property above holds true only because the natural numbers are a **totally ordered set**! In other words, for any two natural numbers x, y with $x \neq y$, exactly one of the two possible relations is true:

- $x \leq y$,
- $x \geq y$.

This is not at all an obvious property for cardinalities of sets! In general, for an relation $\sim$ on some collection of objects, one can easily imagine a world in which knowing that $x \sim y$ and $y \sim x$ does not tell you that $x = y$! For a trivial example, consider the "always-true" relation on any set, that says that $x \sim y$ is true for any two elements in your set! This relation fails the total ordering property in a spectacular fashion; one can easily imagine other more natural relations (like "is the same height as," on the set of people) $\sim$ where knowing that $x \sim y$ and $y \sim x$ does not tell you that $x = y$.

If cardinalities lived in that world, our point 3 would completely fail. We could have two sets A, B such that $|A| \leq |B|$ and $|B| \leq |A|$; however, without knowing that sets are totally ordered under cardinalities, we would not be able to deduce that we must have $|A| = |B|$!

To be clear: this imagined world is **false**. The collection of all cardinal numbers (that is, possible cardinalities of sets) is a totally ordered collection under the $\leq$ relation defined above. But the above discussion has hopefully illustrated why this is not an obvious claim!

With this done, we start the actual proof here. We open by introducing a remarkably useful concept that will come in handy in this proof: the **preimage** of a set.

Definition. Suppose that X, Y are a pair of sets, and that $f : X \to Y$ is a function from X to Y. Take any subset $B \subset Y$. We define the set $f^{-1}(B)$, called the **preimage** of B under the function f, as follows:

$$f^{-1}(B) = \{x \in X \mid f(x) \in B\}.$$

When B is a set consisting only of a single element $\{b\}$, we will often omit the braces, and simply use $f^{-1}(b)$ to denote the set $\{x \in X \mid f(x) = b\}$. Similarly, whenever $f^{-1}(b)$ is a set consisting of a single element, we will often treat it as just that single element rather than as a set containing that single element, and write things like $f^{-1}(b) = a$ instead of $f^{-1}(b) = \{a\}$.

Notice that a function $f : X \to Y$ is injective if and only if $|f^{-1}(y)| \leq 1$, for any $y \in Y$, and similarly that a function $f : X \to Y$ is surjective if and only if $f^{-1}(y)| \geq 1$, for any $y \in Y$.

Using this concept, we proceed with our proof:

Proof. First, start by assuming that the two sets A and B are **disjoint**; i.e. that they have no elements in common. (If this is not the case, simply make them disjoint by marking all of the elements in A with some appropriate subscript. This clearly does not change the size or nature of the set A, nor our desired goal to make a bijection between A and B; the only reason that we want to have these sets disjoint is because it makes our argument somewhat cleaner.)

Recall that because f, g are injective functions, that the sets $f^{-1}(b), g^{-1}(a)$ are either empty or contain single elements as discussed earlier.

Using this idea, we classify set elements in $A \cup B$ as follows:

- If an element $a \in A$ is such that $g^{-1}(a)$ is the empty set, we call a an A-stopping-element.
- Similarly, if an element $b \in B$ is such that $f^{-1}(b)$ is the empty set, we call b a B-stopping element.

- Otherwise, we call these elements boring elements.

With this classification, we break $A \cup B$ into disjoint subsets as follows:

- Take any A-stopping element a. By using the functions f and g, form the following infinite sequence, which we call the A-stopping **chain** corresponding to a:

$$\Big(a, f(a), g(f(a)), f(g(f(a))), g(f(g(f(a)))), \ldots\Big).$$

- Similarly, take any B-stopping element b, and form the B-stopping **chain** corresponding to b as follows:

$$\Big(b, g(b), f(g(b)), g(f(g(b))), f(g(f(g(b)))), \ldots\Big)$$

- Now, take each element $a \in A$ that doesn't show up in any A or B-stopping chain. Form the following double-sided sequence, which we call the **doubly-infinite** chain corresponding to a:

$$\Big(\ldots, g^{-1}(f^{-1}(g^{-1}(a))), f^{-1}(g^{-1}(a)), g^{-1}(a), a, f(a), g(f(a)), f(g(f(a))), \ldots\Big).$$

  Note that the sequence above is well-defined and does continue forever to the left, because a does not show up in any A- or B-stopping chain.

- For each $b \in B$ that doesn't show up in any A or B-stopping chain, define the doubly-infinite chain corresponding to b in the same fashion to a:

$$\Big(\ldots, f^{-1}(g^{-1}(f^{-1}(b))), g^{-1}(f^{-1}(b)), f^{-1}(b), b, g(b), f(g(b)), g(f(g(b)))\ldots\Big).$$

Notice that if any two doubly-infinite chains overlap at any point, then they are completely identical (if this is unclear, work out why this is true!) Moreover, no elements in any A- or B-stopping chain show up in any doubly-infinite chain, because those chains are infinite to their left while A- and B-stopping chains are finite to their left. Finally, no A- or B-stopping chains overlap, because their furthest-left end points are in different sets!

Take all of the chains we've just made, and delete all of the duplicated chains. The result is a collection of chains, such that every element of $A \cup B$ is in exactly one chain!

From here, our proof is remarkably easy. Take any A-chain or any doubly-infinite chain, and break its elements into two sets $A' \subset A, B' \subset B$. Take our function f, and restrict it to A': notice that it is now a bijection $A' \to B'$! This is because for any $b \in B'$, there was some earlier element $a \in A'$ from our chain such that $f(a) = b$.

Similarly: take any B-chain, break its elements into two sets $A' \subset A, B' \subset B$, and consider the function g^{-1} on the set A'. Because this is a B-stopping chain, g^{-1} is defined on all of the elements of A': moreover, g^{-1} is a bijection $A' \to B'$, because for any $a \in A'$ there is earlier $b \in B'$ such that $g(b) = a$.

By combining these two functions, we can make a bijection $A \to B$! Specifically, consider the following function $h : A \to B$:

$$h(a) = \begin{cases} f(a), & \text{for } a \text{ in a doubly-infinite or } A\text{-stopping chain,} \\ g^{-1}(a), & \text{otherwise.} \end{cases}$$

This function is a bijection on each chain; therefore, because the union of the chains gives us all of $A \cup B$ and all of the chains are disjoint, this function is in fact a bijection from A to B! Consequently, $|A| = |B|$, as claimed. $\square$