CARDINALITY

1. Introduction

As our focus in this class is elsewhere, we don't have the lecture time to discuss more set theory. However, as you learn more mathematics, questions about the *size* of sets are often relevant. The notion of the size of a set is formalized in the concept of *cardinality*. In particular, the trichotomy of finite/countably infinite/uncountably infinite recurs with some regularity in both analysis and algebra.

In this handout I will record the basic definitions and prove two foundational results: the Cantor-Bernstein-Schröder theorem, which ensures that our notion of cardinality is reasonable, and Cantor's theorem, which tells us that the cardinality of a set is strictly smaller than the cardinality of its power set. This latter result in particular can be used to show that the set of real numbers is uncountably infinite.

2. Definitions

Once we have defined the natural numbers, it is easy enough to figure out what it means for a set S to have size n for any $n \in \mathbf{Z}_{\geq 0}$: to say that a set has n elements is to say that we can write a list that is n items long consisting of elements of S, such that every element of S is on the list and no element is repeated. To make this precise, note that a list of elements of S with n items is nothing other than a *function*

$$f : \{i \in \mathbf{Z}_{\geq 0} \mid 0 < i \leq n\} = \{1, \ldots, n\} \to S,$$

and the conditions that every element is included on the list and no element is repeated mean exactly that we require f to be a bijection. We proved in class that this notion is reasonable, in the sense that if there exists such an f (so S is *finite*), then the number n is unique. We can therefore define the size of S to be the natural number n.

This works just fine for finite sets, but for infinite sets we obviously run out of numbers to "index" the size of our sets. The solution is to do something more abstract: instead of defining when a set has a certain size, we will instead first define what it means for two sets to have the same size:

Definition 2.1. We say that two sets S and T *have the same cardinality* and write $|S| = |T|$ if there exists a bijection $f : S \to T$. If no such bijection exists, we say that *have different cardinality* and write $|S| \neq |T|$.

What we have done here is to define a relation, "having the same cardinality," on sets. Note that as the set of all sets does not exist by Russell's paradox, this is not strictly speaking a relation under our definition. We will abuse terminology and call it one anyway.

Proposition 2.2 (Properties of the "same cardinality" relation)**.** *For all sets S, T, and U, we have*

(1) $|S| = |S|$,
(2) $|S| = |T|$ *if and only if* $|T| = |S|$, *and*
(3) *If* $|S| = |T|$ *and* $|T| = |U|$, *then* $|S| = |U|$.

Proof. For (1), use the identity function $\mathrm{id}_S : S \to S$, which is always a bijection. For (2), note that if $f : S \to T$ is a bijection, then by our theorem from class there exists an inverse $f^{-1} : T \to S$. It is easy to see, either directly or by using the proposition I stated but did not prove in class regarding persistence of injectivity and surjectivity, that f^{-1} is a bijection as well. The converse is proved in exactly the same way. For (3), let $f : S \to T$ and $g : T \to U$ be bijections. Then by the proposition in class, $f \circ g : S \to U$ is a bijection. □

These properties are known as *reflexivity*, *symmetry*, and *transitivity*, respectively. A relation satisfying all three is called an *equivalence relation*. Equivalence relations are relations that behave formally like equality (hence the notation in our particular case!).

Example 2.3. Let $S = \mathbf{Z}_{\geq 0}$ and $T = \mathbf{Z}$. Then $|S| = |T|$. One bijection that witnesses this is the following: let $f : S \to T$ be the function defined by

$$f(n) = \begin{cases} \frac{n}{2} & \text{if } n \text{ is even,} \\ -\frac{n+1}{2} & \text{if } n \text{ is odd.} \end{cases}$$

Then it is easy to check that f is a bijection. In particular, although it might seem that T is "twice as big" as S in some sense, they have the same cardinality. With a little more thought, one sees also that $|\mathbf{Z}| = |\mathbf{Z} \times \mathbf{Z}|$ and $|\mathbf{Z}| = |\mathbf{Q}|$, for example.

Example 2.4. Assuming some notation that we haven't discussed yet, let $S = [0, 1]$ and $T = [0, 3]$ as subsets of $\mathbf{R}$. Then $g : S \to T$ given by $g(x) = 3x$ is a bijection. Therefore $|S| = |T|$.

Example 2.5. Perhaps more dramatically, let $S = (-1, 1)$ and $T = \mathbf{R}$ and let $h : S \to T$ be defined by

$$h(x) = \tan\left(\frac{\pi}{2}x\right).$$

Therefore $|S| = |T|$.

We will see later that it is possible to use the Cantor-Bernstein-Schröder theorem to prove certain sets have the same cardinality even without producing an explicit bijection.

Example 2.6. It is more difficult to prove that two sets have different cardinality, because we somehow have to rule out all possible functions between them being bijections. In certain cases we can do this "by hand;" i.e., it is clear by examining all the possible functions between them that $\{1\}$ and $\{1, 2\}$ have different cardinalities. We will see another way of doing this when we prove Cantor's theorem. A third way is to exploit some additional feature of the sets at hand: for example, given a positive integer n, let $S = \{1, \ldots, n\}$ and $T = \mathbf{Z}_{\geq 0}$. Suppose that $f : S \to T$ is a function; we will prove that it is not a bijection. Consider the maximum M of $\{f(1), f(2), \ldots, f(n)\}$. Then $M + 1$ is not in the image of f, so f is not surjective. We conclude that $|S| \neq |T|$.

Just as a bijection exhibits that two sets have the same size, an injection exhibits that one set is at least as big as another. We formalize this as follows:

Definition 2.7. Let S and T be sets. We say that S has *smaller or equal cardinality* as T and write $|S| \leq |T|$ or $|T| \geq |S|$ if there exists an injective function $f : S \to T$.

Definition 2.8. Let S and T be sets. We say that S has *smaller cardinality* than T and write $|S| < |T|$ or $|T| > |S|$ if $|S| \leq |T|$ and $|S| \neq |T|$.

Example 2.9. By using the inclusion function, we see that $|\mathbf{Z}| \leq |\mathbf{R}|$, but we don't know yet whether these two sets have the same cardinality (by Cantor's theorem it will turn out that they do not, so in fact $|\mathbf{Z}| < |\mathbf{R}|$).

3. Further properties

We do not yet know that the symbols we have defined behave as we expect them to. For example, the following is not trivial:

Theorem 3.1 (Cantor-Bernstein-Schröder)**.** *Let S and T be sets. If $|S| \leq |T|$ and $|T| \leq |S|$, then $|T| = |S|$.*

Before writing a formal proof, here is one way to think about what is going on. Think of elements of S as cats and elements of T as dogs. By the assumptions of the theorem, we have injective functions f from cats to dogs and g from dogs to cats, and we want to construct a bijection; i.e., we want to perfectly pair off the cats and dogs. We interpret $f(s) = t$ as "the cat s is chasing the dog t" and $g(t) = s$ as "the dog t is chasing the cat s." As the functions are everywhere defined, each dog is chasing a cat and each cat is chasing a dog. Furthermore by injectivity of f and g, no two dogs are chasing the same cat and no two cats are chasing the same dog. Therefore there are a limited number of configurations of cats and dogs that can possibly exist: we can have doubly infinite chains that looks like

$$\cdots \to \text{cat} \to \text{dog} \to \text{cat} \to \text{dog} \to \cdots,$$

we can have singly infinite chains that looks like

$$\text{cat} \to \text{dog} \to \text{cat} \to \text{dog} \to \cdots$$

or

$$\text{dog} \to \text{cat} \to \text{dog} \to \text{cat} \to \cdots,$$

and we can have loops of dogs and cats with an even number of animals in each loop, perhaps looking something like

$$\text{dog} \rightleftarrows \text{cat} \longrightarrow \text{dog} \longrightarrow \text{cat}$$

In each case, we can pair off the cats and dogs: if we have a doubly infinite chain then we can, say, pair off each dog with the cat it is chasing, if we have a singly infinite chain we can do that if there is a dog at the

end of the line or the opposite if a cat is at the end of the line, and for a finite loop we can use the same recipe again. This basic argument almost constitutes a proof; we just have to make it precise:

Proof. By assumption there exist injective functions $f : S \to T$ and $g : T \to S$. We wish to construct a bijection from S to T. As a matter of notation, if n is a positive integer and $h : U \to U$ a function from a set to itself, let's use h^n to denote the n-times composition of h with itself; e.g. $h^2 = h \circ h$ (we can define this notation inductively). If $n = 0$ we can just let h^0 denote the identity function id_U.

Any injective function is bijective onto its image, so in particular $g : T \to g(T)$ is a bijection and we may consider its inverse $g^{-1} : g(T) \to T$. (Of course, we can't define g^{-1} on *all* of S, just on the image of g.) Consider the subset

$$V = \bigcup_{n=0}^{\infty} (g \circ f)^n (S - g(T)) \subseteq S,$$

the infinite union of a bunch of images under many compositions of $g \circ f$. Let $W = S - V$, so S is the disjoint union of V and W (i.e., $S = V \cup W$ and $V \cap W = \emptyset$). Define a function $h : S \to T$ by

$$h(s) = \begin{cases} f(s) & \text{if } s \in V, \\ g^{-1}(s) & \text{if } s \in W. \end{cases}$$

This makes sense because if $s \in W$, then $s \notin V$, so in particular $s \notin (S - g(T))$; i.e., $s \in g(T)$. Therefore the inverse to g is defined for such s.

We show that h is injective and surjective.

For injective, assume that $h(s) = h(s')$. There are three cases to consider. If $s, s' \in V$, then we know that $s = s'$ because f itself is injective. If $s, s' \in W$, then $g^{-1}(s) = g^{-1}(s')$, so applying g to both sides yields $s = s'$. Finally the third case is that one element is in V and the other is in W; without loss of generality we say $s \in V$ and $s' \in W$. By the definition of V, $s = (g \circ f)^n(z)$ for some $n \geq 0$ and some $z \in S - g(T)$. Now $h(s) = h(s')$ implies that $f(s) = g^{-1}(s')$, so $f((g \circ f)^n(z)) = g^{-1}(s')$. Apply g to both sides; we get $(g \circ f)^{n+1}(z) = s'$. Therefore $s' \in V$, contrary to assumption. Therefore this case cannot occur.

For surjective, let $t \in T$. Either $g(t) \in V$ or $g(t) \in W$. In the second case, $h(g(t)) = g^{-1}(g(t)) = t$, so $g(t)$ is an element of S mapping to t via h. In the first case, by definition of V we have $g(t) = (g \circ f)^n(z)$ for some $n > 0$ and some $z \in S - g(T)$. Let's rewrite this as $g(t) = g\left(f\left((g \circ f)^{n-1}(z)\right)\right)$. As g is injective, we conclude that $t = f\left((g \circ f)^{n-1}(z)\right)$. Letting $s = (g \circ f)^{n-1}(z)$, we see that as $s \in V$ we have $h(s) = f(s) = t$. Therefore $s \in S$ maps to t via h, and we are done. □

In the above formal proof, we defined a bijection h from "cats" to "dogs." The condition $s \in V$ is the condition that the "chain" that s lies in is singly infinite and starts with a cat; in this case we map the cat s to the dog it is chasing, which is $f(s)$. In all other cases $s \in W$, and we can map the cat to the dog that is chasing it, which is $g^{-1}(s)$.

Example 3.2. We can use Cantor-Bernstein-Schröder to prove that certain sets are bijective without writing down tricky equations. For example, $S = (0, 1)$ and $T = [0, 1)$ have the same cardinality. It might not be immediately obvious how to write down a bijection, but we can certainly write down the injections $f : S \to T$ defined by $f(x) = x$ and $g : T \to S$ defined by $g(x) = \frac{1}{2}x + \frac{1}{4}$, say. The theorem then implies that $|S| = |T|$.

Denote the power set of a set S by $\mathcal{P}(S)$. As a more involved example, we have

Proposition 3.3. *The sets* $\mathbf{R}$ *and* $\mathcal{P}(\mathbf{Z}_{\geq 0})$ *have the same cardinality.*

Proof. By the Theorem, it suffices to exhibit injections in both directions. We will use decimal representations of real numbers, which we have certainly not proved to exist yet, but let's assume we already know how they work for the purposes of this proof. In particular, any number has a decimal expansion, and such an expansion is unique if we furthermore require that the expansion does not terminate in an infinite string of 9's (unfortunately, we have equalities like $0.99\overline{9} = 1.00\overline{0}$ that ruin uniqueness if we do not enforce this condition). Let h be the bijection from Example 2.5 taking $\mathbf{R}$ to $(-1, 1)$, and let $h' : \mathbf{R} \to (0, 1)$ be the composition of h with the function that adds one and then divides by two. In particular h' is injective, and because compositions of injective functions are injective it suffices in one direction to find an injection $(0, 1) \to \mathcal{P}(\mathbf{Z}_{\geq 0})$ and in the other direction to find an injection $\mathcal{P}(\mathbf{Z}_{\geq 0}) \to \mathbf{R}$.

Let us define $f : (0, 1) \to \mathcal{P}(\mathbf{Z}_{\geq 0})$ by taking a real number with decimal expansion $0.a_1a_2a_3\ldots$ (requiring no infinite string of 9's) to the set $\{10a_1, 100a_2, 1000a_3, \ldots\}$. For example, $f(0.\overline{34}) = \{30, 400, 3000, 40000, \ldots\}$. This function is injective: if $0.a_1a_2a_3\ldots$ is not equal to $0.b_1b_2b_3\ldots$, then by uniqueness of decimal expansions we have $a_n \neq b_n$ for some n. But then $10^n a_n$ will lie in $f(0.a_1a_2a_3\ldots)$ but not in $f(0.b_1b_2b_3\ldots)$, so the two sets are distinct.

In the opposite direction, let $g : \mathcal{P}(\mathbf{Z}_{\geq 0}) \to \mathbf{R}$ be the function taking a set S of natural numbers to the decimal expansion $0.a_1a_2a_3\ldots$ where $a_n = 1$ if $n \in S$ and $a_n = 0$ otherwise. (This is very similar to the maneuver that we did on Exercise 8 of the second homework.) For example, $g(\emptyset) = 0$. If $S \neq T$, then they differ at one place; say $n \in S$ but $n \notin T$. Then $g(S)$ and $g(T)$ differ at the nth decimal place, and are consequently not equal. Therefore g is injective. Putting everything together, we get the result. □

There is one more foundational issue to address: with the tools we have, it is difficult to actually abstractly produce functions between sets. In particular, we do not know that given any two sets S and T, there exists either an injection $S \to T$ or an injection $T \to S$, so we do not know that any two sets are "comparable" in the sense that either $|S| \leq |T|$ or $|T| \leq |S|$. It turns out that this requires a genuinely new tool in set theory, the *axiom of choice*. So as not to get bogged down in set theory, we will merely note that the axiom of choice in its usual formulation is actually equivalent to the statement that for any two sets S and T, we have $|S| \leq |T|$ or $|T| \leq |S|$. Therefore we will skip the intermediate steps and simply declare this as a new axiom of set theory:

Axiom 3.4 (Equivalent formulation of the axiom of choice)**.** If S and T are sets, then $|S| \leq |T|$ or $|T| \leq |S|$.

With this axiom and the Cantor-Bernstein-Schröder theorem, our notation for comparing cardinality behaves exactly as we would wish. In particular, given any two sets S and T, exactly one of $|S| < |T|$, $|S| = |T|$, or $|S| > |T|$ is true.

There is still something slightly unsatisfying about all of this: although we know how to compare the sizes of sets, we have no definition of "what cardinality is;" i.e., it is meaningless to ask what the cardinality of a set S is. We can fix this, but only via a sort of linguistic trick. We define the *equivalence class* of any set S to be the collection of all sets T such that $|S| = |T|$. By Proposition 2.2, this partitions all sets into equivalence classes, the classes of sets that are all of the same cardinality. We then *define* the cardinality of a set to be its equivalence class; then two sets have the same cardinality in the sense of Definition 2.1 if and only if they have the same cardinality in this new sense.

The same construction works for an arbitrary equivalence relation (as defined after Proposition 2.2), and is often very useful.

4. Cantor's theorem and countability

Theorem 4.1 (Cantor)**.** *Let S be a set. Then $|S| < |\mathcal{P}(S)|$.*

Proof. Certainly there is an injective function $f : S \to \mathcal{P}(S)$ defined by $f(s) = \{s\}$, so $|S| \leq |\mathcal{P}(S)|$. We must show that there is no bijection $S \to \mathcal{P}(S)$, which we will do by showing that there is no surjection.

The method of proof is reminiscent of Russell's paradox. Suppose that $g : S \to \mathcal{P}(S)$ is surjective. As g takes elements of S to subsets of S, for any $s \in S$ we can ask whether $s \in g(s)$ or not. Let

$$T = \{s \in S : s \notin g(s)\} \subseteq S.$$

By surjectivity of g, there is some $x \in S$ such that $g(x) = T$. We may ask whether $x \in T$ or not.

In the first case, $x \in T$, then the definition of T implies that $x \notin g(x)$. But $g(x) = T$, so $x \notin T$. This is a contradiction.

In the second case, $x \notin T$, so by the definition of T we have $x \in g(x)$. But again $g(x) = T$, so $x \in T$. This is a contradiction.

Therefore no such g exists and we conclude. □

Corollary 4.2. *We have $|\mathbf{Z}_{\geq 0}| < |\mathbf{R}|$.*

Proof. By Theorem 4.1, we have $|\mathbf{Z}_{\geq 0}| < |\mathcal{P}(\mathbf{Z}_{\geq 0})|$ and by Proposition 3.3 we have $|\mathcal{P}(\mathbf{Z}_{\geq 0})| = |\mathbf{R}|$, so we immediately conclude. □

The set of real numbers is "larger" than the set of natural numbers in quite a serious way.

Cantor's theorem gives us the means to construct a number of large sets: take any set, such as $\mathbf{Z}_{\geq 0}$, and iterate the power set operation as many times as you like. Each time you do, you get a new and bigger set. The basic terminology is the following:

Definition 4.3. A set S is *finite* if $|S| < |\mathbf{Z}_{\geq 0}|$, *countably infinite* if $|S| = |\mathbf{Z}_{\geq 0}|$, and *uncountably infinite* or simply *uncountable* if $|S| > |\mathbf{Z}_{\geq 0}|$.

So the above corollary states that the real numbers are uncountable. It is possible, but not entirely trivial, to show that this definition of finite coincides with the definition given on Homework 3.

A natural question at this point is the following: are there any cardinalities between that of the natural numbers and that of the reals? That is, does there exist a set S such that $|\mathbf{Z}_{\geq 0}| < |S| < |\mathbf{R}|$? The statement that no such set exists is known as the *continuum hypothesis*. Deep work of Kurt Gödel in the 1930s and Paul Cohen in the 1960s has established that the continuum hypothesis is independent of the usual axioms of set theory (i.e., Zermelo-Fraenkel set theory with the axiom of choice included). That is, if set theory itself is consistent, it is also consistent to add an axiom stating that the continuum hypothesis is true, and equally consistent to add an axiom stating that the continuum hypothesis is false.