Infinite Cardinals

Mahesh Viswanathan

The main question that will concern us in this lecture is, "How do we compare the sizes of two sets A and B"? We will especially be interested in the case when the sets A and B have infinitely many elements. One natural way to compare the sizes of A and B is to pair up elements of A with elements of B. If every element of A is paired up with some element of B, then B has at least as many elements as A. In addition, if every element of B is also part of some pair, then A and B have the same size. Cantor formalized this intuition as a definition, and we will explore its consequences.

Before presenting the main definition of this lecture, let us recall that a function $f : A \to B$ is said to be *injective* or *one-to-one* if for any $a_1, a_2 \in A$, if $f(a_1) = f(a_2)$ then $a_1 = a_2$. In other words, distinct elements of A are mapped to distinct elements of B. In addition, f is said to be *surjective* or *onto* if for every $b \in B$, there is some $a \in A$ such that $f(a) = b$. That is, every element of B is the image (under f) of some element of A. Finally, f is *bijective* if it is both injective and surjective.

Definition 1 (Cantor)**.** *A set A is said to have the* same cardinality *as a set B, denoted $|A| = |B|$, if there is a bijective function $f : A \to B$.*

A set A has cardinality no more *than that of B, denoted $|A| \leq |B|$, if there is an injective function $f : A \to B$.*

Let us look at an example to highlight some of subtleties of this definition.

Example 2. *Recall that set of natural numbers $\mathbb{N} = \{0, 1, 2, \ldots\}$. Consider the subset of even natural numbers $2\mathbb{N} = \{2n \mid n \in \mathbb{N}\}$. One can show that $|\mathbb{N}| = |2\mathbb{N}|$ because the function* half $: 2\mathbb{N} \to \mathbb{N}$*, given by* $\text{half}(n) = n/2$*, is bijective. It is worth observing a couple of cautionary points about our intuitions from finite sets.*

1. *A set A and a strict subset B of A may have the same cardinality. This holds even if $A \setminus B$ has infinitely many elements. This is happens only when A and B are infinite; if A is finite and $B \subset A$ is such that $B \setminus A \neq \emptyset$, then there can be no bijection between A and B. In fact, Cantor used this property to define the notion of an infinite set — A is an infinite set if for any $a \in A$, $|A| = |A \setminus \{a\}|$.*

2. *Notice, the presence of* one *bijection between A and B is enough to conclude $|A| = |B|$. In particular, there are examples of injective functions $f : 2\mathbb{N} \to \mathbb{N}$ that are not onto, e.g., the function* id $: 2\mathbb{N} \to \mathbb{N}$*, given by* $\text{id}(n) = n$*, which is one-to-one but not onto. So demonstrating an injective function that is not onto does not prove that the sets don't have the same cardinality. This is again not true for finite sets. If A and B are finite sets, and there is an injective function $f : A \to B$ that is not onto, then one can conclude that there is no bijection between A and B.*

Example 3. *Let Σ be a finite set. Recall that Σ^* is the set of all finite strings/sequences/words over Σ and ϵ is the empty string, which is the unique string of length 0. We will observe that $|\Sigma^*| = |\mathbb{N}|$.*

Without loss of generality, take $\Sigma = \{1, 2, \ldots k\}$, for some k. The function num $: \Sigma^* \to \mathbb{N}$ *that maps a string $w = a_0 a_1 \cdots a_n$, where $a_i \in \Sigma$, to the* base-k bijective number *it represents is a bijection* [1]*. In other words,*

$$\text{num}(\epsilon) = 0$$
$$\text{num}(a_0 a_1 \cdots a_n) = \textstyle\sum_{i=0}^{n} a_i k^i$$

[1] See `https://en.wikipedia.org/wiki/Bijective_numeration` for more on bijective number systems.

is a bijection. Notice it is important that we use a bijective number system because otherwise the mapping above will not be injective because of the usual problem of "zeros in the most significant bits."

The cardinality of sets is often compared with the set of natural numbers $\mathbb{N}$, because $\mathbb{N}$ is the "smallest" infinite set.

Definition 4. *A set A is said to be* countable *if $|A| \leq |\mathbb{N}|$. If, in addition A is infinite, then A is said to be* countably infinite.

Thus, based on Examples 2 and 3, we can conclude that $2\mathbb{N}$ and Σ^* are countable, for any finite set Σ.

Definition 1 says that $|A| \leq |B|$ if there is an injective function from A to B, while A and B are said to have the same cardinality if there is a bijection between A and B. Are these definitions consistent? In other words, if $|A| \leq |B|$ and $|B| \leq |A|$ then is $|A| = |B|$? That is, if there are injective functions from A to B, and from B to A, is there a bijective function between A and B? This turns out to be true and is an important observation, attributed to Cantor, Schröder, and Bernstein.

Theorem 5 (Cantor-Schröder-Bernstein)**.** *For any two sets A, B, $|A| = |B|$ if and only if $|A| \leq |B|$ and $|B| \leq |A|$.*

Before proving this theorem, let us look at an example that highlights both the usefulness of this result, as well as why it's proof is non-trivial.

Example 6. *Recall that the set of rational numbers $\mathbb{Q}$, is the collection of all real numbers that can be written as a fraction a/b, where $a, b \in \mathbb{N}$. Classical examples of irrational numbers include $\sqrt{2}$, and π.*

Since $\mathbb{N} \subseteq \mathbb{Q}$, it is easy to see that $|\mathbb{N}| \leq |\mathbb{Q}|$; the function $\mathrm{id} : \mathbb{N} \to \mathbb{Q}$, defined as $\mathrm{id}(n) = n$, witnesses this fact. Similarly, one can construct an injective function from $\mathbb{Q}$ to $\mathbb{N}$ to demonstrate $|\mathbb{Q}| \leq |\mathbb{N}|$ as follows.

$$f(r) = 2^a 3^b \text{ where } a, b \in \mathbb{N},\ r = a/b,\ \text{and } \gcd(a, b) = 1.$$

The function f can be argued to be injective because every rational number has a unique representation as a/b, where $\gcd(a, b) = 1$, and because of the fact that every natural number has a unique prime factorization.

Notice, that because of the above observations, Theorem 5 allows us to conclude that $|\mathbb{Q}| = |\mathbb{N}|$. However, coming up with a bijective functions between the two sets is not as straightforward as coming up with the injective functions above. This illustrates both why the proof of Theorem 5 is non-trivial and why it is useful — it is often easier to come up with injective functions between sets than a bijective function.

Proof of Theorem 5. Let us start with some notation. For a function $h : D \to E$, and a subset $D' \subseteq D$, we use $h(D') \subseteq E$ to denote the set $\{h(d) \mid d \in D'\}$.

First observe that if $f : A \to B$ is a bijection (i.e., $|A| = |B|$) then $f : A \to B$ and $f^{-1} : B \to A$ are both injective. Thus, $|A| \leq |B|$ and $|B| \leq |A|$. The interesting direction is, therefore, to argue that if $|A| \leq |B|$ and $|B| \leq |A|$ then $|A| = |B|$.

Suppose $f : A \to B$ and $g : B \to A$ are injective functions. Let us inductively define the following sequence of subsets of A.

$$\begin{aligned} C_0 &= A \setminus g(B) \\ C_{i+1} &= g(f(C_i)) \quad \text{for } i \geq 0. \end{aligned}$$

Let $C_* = \cup_{i \in \mathbb{N}} C_i$, i.e., the elements of A that belong to some C_i. Consider the function $h : A \to B$ defined as follows.

$$h(a) = \begin{cases} f(a) & \text{if } a \in C_* \\ g^{-1}(a) & \text{otherwise} \end{cases}$$

It useful to observe that h is well-defined — since $A \setminus C_0 = g(B)$, we have $A \setminus C_* \subseteq g(B)$, and $g^{-1}(a)$ is defined when $a \in A \setminus C_*$. We will show that h is our desired bijection.

First we argue that h is injective. Consider $a_1, a_2 \in A$ such that $a_1 \neq a_2$. If $a_1, a_2 \in C_*$ then since $h(a_i) = f(a_i)$ (for $i \in \{1, 2\}$), $h(a_1) \neq h(a_2)$ because f is injective. Similarly, if $a_1, a_2 \notin C_*$ then $h(a_i) = g^{-1}(a_i)$ and $h(a_1) \neq h(a_2)$ because g is injective. The only case left is when exactly one out of a_1, a_2 is in

C_*; without loss of generality, assume that $a_1 \in C_*$ and $a_2 \notin C_*$. Since $a_1 \in C_*$, we know that $a_1 \in C_i$ for some i. Now if $h(a_1) = h(a_2)$ then we have $f(a_1) = g^{-1}(a_2)$, which means that $g(f(a_1)) = a_2$. Thus, $a_2 \in g(f(C_i))$, which means that $a_2 \in C_{i+1} \subseteq C_*$. This contradicts the assumption that $a_2 \notin C_*$.

Now, we show that h is surjective. Consider any $b \in B$. We need to show that there is some $a \in A$ such that $h(a) = b$. We consider two cases. Suppose $g(b) \notin C_*$. Then $h(g(b)) = g^{-1}(g(b)) = b$. Thus, in this case if we take $a = g(b)$, then $h(a) = b$. The second case we need to consider is if $g(b) \in C_*$. This means that $g(b) \in C_i$ for some i. Moreover, since $C_0 = A \setminus g(B)$, we have $g(b) \notin C_0$. Thus, $g(b) \in C_i$ for some $i > 0$. Or in other words, $g(b) \in C_i = g(f(C_{i-1}))$ for $i - 1 \geq 0$. Since g is injective, we have $b \in f(C_{i-1})$, or there is $a \in C_{i-1}$ such that $f(a) = b$. But for such an a, we have $h(a) = f(a) = b$ which completes the proof. □

Through Examples 2, 3, and 6, we have observed that the sets $2\mathbb{N}$, Σ^* (for any finite set Σ), and $\mathbb{Q}$ are all countable. It is natural to ask if there are any that are not countable [2]. It turns out that many sets are uncountable. The most famous example of an uncountable set is the set of real numbers $\mathbb{R}$. This observation is due to Cantor. Cantor, in fact, proved an even more powerful observation — given any set A, its powerset 2^A, namely the collection of all its subsets, has a strictly larger cardinality. Thus, we can construct larger and larger infinite sets by taking the powerset repeatedly. These observations were proved by Cantor using a powerful proof technique that he discovered called *diagonalization*. We illustrate this idea through the next proof.

Theorem 7 (Cantor)**.** *For any set A, there is no surjective function from A to 2^A. Thus, $|A| \neq |2^A|$.*

Proof. Consider any function $f : A \to 2^A$. We will show that f is not surjective by demonstrating a subset B of A such that B is not in the range of f. Define the set B as follows.

$$B = \{a \in A \mid a \notin f(a)\}.$$

To establish this claim, we need to show that $B \neq f(a)$ for any $a \in A$. If $a \in f(a)$ then we know (by definition), $a \notin B$. Therefore, $a \in f(a) \setminus B$. On the other hand, if $a \notin f(a)$ then (by definition) $a \in B$, and so $a \in B \setminus f(a)$. Since no matter what $a \in (B \setminus f(a)) \cup (f(a) \setminus B)$, we have $B \neq f(a)$. Thus, f is not surjective because B is not the image of any element under f. □

Observe that since the function $\mathrm{id} : A \to 2^A$ defined as $\mathrm{id}(a) = \{a\}$ is injective, we trivially have $|A| \leq |2^A|$. Since Theorem 7 shows that $|A| \neq |2^A|$, we have that $|A| < |2^A|$ (cardinality of A is strictly smaller than that of 2^A). Theorem 7 also gives an example of an uncountable set, namely, $2^{\mathbb{N}}$.

Theorem 7 has some simple consequences for computability. Recall that a decision problem is one where, given an input, the task is to compute a boolean valued function on the input. Inputs to computational problems can be (abstractly) thought of as being encoded as a string over some alphabet Σ like the set of ASCII characters. Thus, any decision problem is a function $f : \Sigma^* \to \{0, 1\}$, which can also be thought of as the set of input (strings) on which f evaluates to 1. Therefore, the collection of decision problems is the set 2^{Σ^*}. Can all of these problems be solved algorithmically? No matter what programming language we consider, we can make one fundamental assumption — any program is a finite string of instructions encoded in some alphabet Σ say the ASCII characters. Thus the size of the collection of all decision problems that can be solved by some program is no more that the size of the set Σ^*, which is countable. Since 2^{Σ^*} is larger set than Σ^*, this means that there are (many) computational problems that have no algorithmic solution.

[2] If a set is not countable, it is called *uncountable*.