13 Infinite Sets

So you might be wondering how much is there to say about an infinite set other than, well, it has an infinite number of elements. Of course, an infinite set does have an infinite number of elements, but it turns out that not all infinite sets have the same size—some are bigger than others! And, understanding infinity is not as easy as you might think. Some of the toughest questions in mathematics involve infinite sets.

Why should you care? Indeed, isn't computer science only about finite sets? Not exactly. For example, we deal with the set of natural numbers $\mathbb{N}$ all the time and it is an infinite set. In fact, that is why we have induction: to reason about predicates over $\mathbb{N}$. Infinite sets are also important in Part IV of the text when we talk about random variables over potentially infinite sample spaces.

So sit back and open your mind for a few moments while we take a very brief look at *infinity*.

13.1 Injections, Surjections, and Bijections

We know from Theorem 7.2.1 that if there is an injection or surjection between two finite sets, then we can say something about the relative sizes of the two sets. The same is true for infinite sets. In fact, relations are the primary tool for determining the relative size of infinite sets.

Definition 13.1.1. Given any two sets A and B, we say that

A surj B	iff there is a surjection from A to B,
A inj B	iff there is an injection from A to B,
A bij B	iff there is a bijection between A and B, and
A strict B	iff there is a surjection from A to B but there is *no* bijection from B to A.

Restating Theorem 7.2.1 with this new terminology, we have:

Theorem 13.1.2. *For any pair of* finite *sets A and B,*

$$|A| \geq |B| \leftarrow \textit{iff} \quad A \text{ surj } B,$$
$$|A| \leq |B| \leftarrow \textit{iff} \quad A \text{ inj } B,$$
$$|A| = |B| \leftarrow \textit{iff} \quad A \text{ bij } B,$$
$$|A| > |B| \leftarrow \textit{iff} \quad A \text{ strict } B.$$

Theorem 13.1.2 suggests a way to generalize size comparisons to infinite sets; namely, we can think of the relation surj as an "at least as big" relation between sets, even if they are infinite. Similarly, the relation bij can be regarded as a "same size" relation between (possibly infinite) sets, and strict can be thought of as a "strictly bigger" relation between sets.

Note that we haven't, and won't, define what the size of an infinite set is. The definition of infinite "sizes" is cumbersome and technical, and we can get by just fine without it. All we need are the "as big as" and "same size" relations, surj and bij, between sets.

But there's something else to watch out for. We've referred to surj as an "as big as" relation and bij as a "same size" relation on sets. Most of the "as big as" and "same size" properties of surj and bij on finite sets do carry over to infinite sets, but *some important ones don't*—as we're about to show. So you have to be careful: don't assume that surj has any particular "as big as" property on *infinite* sets until it's been proved.

Let's begin with some familiar properties of the "as big as" and "same size" relations on finite sets that do carry over exactly to infinite sets:

Theorem 13.1.3. *For any sets, A, B, and C,*

1. A surj B *and* B surj C IMPLIES A surj C.
2. A bij B *and* B bij C IMPLIES A bij C.
3. A bij B IMPLIES B bij A.

Parts 1 and 2 of Theorem 13.1.3 follow immediately from the fact that compositions of surjections are surjections, and likewise for bijections. Part 3 follows from the fact that the inverse of a bijection is a bijection. We'll leave a proof of these facts to the problems.

Another familiar property of finite sets carries over to infinite sets, but this time it's not so obvious:

Theorem 13.1.4 (Schröder-Bernstein)**.** *For any pair of sets A and B, if A* surj B *and B* surj A*, then A* bij B.

The Schröder-Bernstein Theorem says that if A is at least as big as B and, conversely, B is at least as big as A, then A is the same size as B. Phrased this way, you might be tempted to take this theorem for granted, but that would be a mistake. For infinite sets A and B, the Schröder-Bernstein Theorem is actually pretty technical. Just because there is a surjective function $f : A \to B$—which need not be a bijection—and a surjective function $g : B \to A$—which also need not

be a bijection—it's not at all clear that there must be a bijection $h : \not\leftarrow A \not\rightarrow \not\leftarrow B$. The challenge is to construct h from parts of both f and g. We'll leave the actual construction to the problems.

13.1.1 Infinity Is Different

A basic property of finite sets that does *not* carry over to infinite sets is that adding something new makes a set bigger. That is, if A is a finite set and $b \notin A$, then $|A \cup \{b\}| = |A| + 1$, and so A and $A \cup \{b\}$ are not the same size. But if A is infinite, then these two sets *are* the same size!

Theorem 13.1.5. *Let A be a set and $b \notin A$. Then A is infinite iff A* bij $A \cup \{b\}$.

Proof. Since A is *not* the same size as $A \cup \{b\}$ when A is finite, we only have to show that $A \cup \{b\}$ *is* the same size as A when A is infinite.

That is, we have to find a bijection between $A \cup \{b\}$ and A when A is infinite. Since A is infinite, it certainly has at least one element; call it a_0. Since A is infinite, it has at least two elements, and one of them must not be equal to a_0; call this new element a_1. Since A is infinite, it has at least three elements, one of which must not equal a_0 or a_1; call this new element a_2. Continuing in this way, we conclude that there is an infinite sequence $a_0, a_1, a_2, \ldots, a_n, \ldots$, of different elements of A. Now it's easy to define a bijection $f : A \cup \{b\} \to A$:

$$\begin{aligned} f(b) &::= a_0, \\ f(a_n) &::= a_{n+1} && \text{for } n \in \mathbb{N}, \\ f(a) &::= a && \text{for } a \in A - \{b, a_0, a_1, \ldots\}. \end{aligned}$$

■

13.2 Countable Sets

13.2.1 Definitions

A set C is *countable* iff its elements can be listed in order, that is, the distinct elements in C are precisely

$$c_0, c_1, \ldots, c_n, \ldots.$$

This means that if we defined a function f on the nonnegative integers by the rule that $f(i) ::= c_i$, then f would be a bijection from $\mathbb{N}$ to C. More formally,

Definition 13.2.1. A set C is *countably infinite* iff $\mathbb{N}$ bij C. A set is *countable* iff it is finite or countably infinite.

Discrete mathematics is often defined as the mathematics of countable sets and so it is probably worth spending a little time understanding what it means to be countable and why countable sets are so special. For example, a small modification of the proof of Theorem 13.1.5 shows that countably infinite sets are the "smallest" infinite sets; namely, if A is any infinite set, then A surj $\mathbb{N}$.

13.2.2 Unions

Since adding one new element to an infinite set doesn't change its size, it's obvious that neither will adding any *finite* number of elements. It's a common mistake to think that this proves that you can throw in countably infinitely many new elements—just because it's ok to do something any finite number of times doesn't make it ok to do it an infinite number of times.

For example, suppose that you have two countably infinite sets $A = \{a_0, a_1, a_2, \ldots\}$ and $B = \{b_0, b_1, b_2, \ldots\}$. You might try to show that $A \cup B$ is countable by making the following "list" for $A \cup B$:

$$a_0, a_1, a_2, \ldots, b_0, b_1, b_2, \ldots \tag{13.1}$$

But this is not a valid argument because Equation 13.1 is not a list. The key property required for listing the elements in a countable set is that for any element in the set, you can determine its finite index in the list. For example, a_i shows up in position i in Equation 13.1, but there is no index in the supposed "list" for any of the b_i. Hence, Equation 13.1 is not a valid list for the purposes of showing that $A \cup B$ is countable when A is infinite. Equation 13.1 is only useful when A is finite.

It turns out you really can add a countably infinite number of new elements to a countable set and still wind up with just a countably infinite set, but another argument is needed to prove this.

Theorem 13.2.2. *If A and B are countable sets, then so is $A \cup B$.*

Proof. Suppose the list of distinct elements of A is $a_0, a_1, \ldots$, and the list of B is $b_0, b_1, \ldots$. Then a valid way to list all the elements of $A \cup B$ is

$$a_0, b_0, a_1, b_1, \ldots, a_n, b_n, \ldots. \tag{13.2}$$

Of course this list will contain duplicates if A and B have elements in common, but then deleting all but the first occurrence of each element in Equation 13.2 leaves a list of all the distinct elements of A and B. ■

Note that the list in Equation 13.2 does not have the same defect as the purported "list" in Equation 13.1, since every item in $A \cup B$ has a finite index in the list created in Theorem 13.2.2.

	b_0	b_1	b_2	b_3	$\ldots$
a_0	c_0	c_1	c_4	c_9	
a_1	c_3	c_2	c_5	c_{10}	
a_2	c_8	c_7	c_6	c_{11}	
a_3	c_{15}	c_{14}	c_{13}	c_{12}	
$\vdots$					$\ddots$

Figure 13.1 A listing of the elements of $C = A \times B$ where $A = \{a_0, a_1, a_2, \ldots\}$ and $B = \{b_0, b_1, b_2, \ldots\}$ are countably infinite sets. For example, $c_5 = (a_1, b_2)$.

13.2.3 Cross Products

Somewhat surprisingly, cross products of countable sets are also countable. At first, you might be tempted to think that "infinity times infinity" (whatever that means) somehow results in a larger infinity, but this is not the case.

Theorem 13.2.3. *The cross product of two countable sets is countable.*

Proof. Let A and B be any pair of countable sets. To show that $C = A \times B$ is also countable, we need to find a listing of the elements

$$\{(a, b) \mid a \in A, b \in B\}.$$

There are many such listings. One is shown in Figure 13.1 for the case when A and B are both infinite sets. In this listing, (a_i, b_j) is the kth element in the list for C where

$$\begin{aligned} &a_i \text{ is the } i\text{th element in } A,\\ &b_j \text{ is the } j\text{th element in } B, \text{ and}\\ &k = \max(i, j)^2 + i + \max(i - j, 0). \end{aligned}$$

The task of finding a listing when one or both of A and B are finite is left to the problems at the end of the chapter.

■

13.2.4 $\mathbb{Q}$ Is Countable

Theorem 13.2.3 also has a surprising Corollary; namely that the set of rational numbers is countable.

Corollary 13.2.4. *The set of rational numbers $\mathbb{Q}$ is countable.*

Proof. Since $\mathbb{Z}\times\mathbb{Z}$ is countable by Theorem 13.2.3, it suffices to find a surjection f from $\mathbb{Z}\times\mathbb{Z}$ to $\mathbb{Q}$. This is easy to to since

$$f(a,b) = \begin{cases} a/b & \text{if } b \neq 0 \\ 0 & \text{if } b = 0 \end{cases}$$

is one such surjection. ■

At this point, you may be thinking that every set is countable. That is *not* the case. In fact, as we will shortly see, there are many infinite sets that are uncountable, including the set of real numbers $\mathbb{R}$.

13.3 Power Sets Are Strictly Bigger

It turns out that the ideas behind Russell's Paradox, which caused so much trouble for the early efforts to formulate Set Theory, also lead to a correct and astonishing fact discovered by Georg Cantor in the late nineteenth century: infinite sets are *not all the same size*.

Theorem 13.3.1. *For any set A, the power set $\mathcal{P}(A)$ is strictly bigger than A.*

Proof. First of all, $\mathcal{P}(A)$ is as big as A: for example, the partial function $f :\leftarrow \mathcal{P}(A) \to A$ where $f(\{a\}) ::= a$ for $a \in A$ is a surjection.

To show that $\mathcal{P}(A)$ is strictly bigger than A, we have to show that if g is a function from A to $\mathcal{P}(A)$, then g is not a surjection. So, mimicking Russell's Paradox, define

$$A_g ::= \{\, a \in A \mid a \notin g(a) \,\}.$$

A_g is a well-defined subset of A, which means it is a member of $\mathcal{P}(A)$. But A_g can't be in the range of g, because if it were, we would have

$$A_g = g(a_0)$$

for some $a_0 \in A$. So by definition of A_g,

$$a \in g(a_0) \quad \text{iff} \quad a \in A_g \quad \text{iff} \quad a \notin g(a)$$

for all $a \in A$. Now letting $a = a_0$ yields the contradiction

$$a_0 \in g(a_0) \quad \text{iff} \quad a_0 \notin g(a_0).$$

So g is not a surjection, because there is an element in the power set of A, namely the set A_g, that is not in the range of g. ■

13.3.1 $\mathbb{R}$ Is Uncountable

To prove that the set of real numbers is uncountable, we will show that there is a surjection from $\mathbb{R}$ to $\mathcal{P}(\mathbb{N})$ and then apply Theorem 13.3.1 to $\mathcal{P}(\mathbb{N})$.

Lemma 13.3.2. $\mathbb{R}$ surj $\mathcal{P}(\mathbb{N})$.

Proof. Let $A \subset \mathbb{N}$ be any subset of the natural numbers. Since $\mathbb{N}$ is countable, this means that A is countable and thus that $A = \{a_0, a_1, a_2, \ldots\}$. For each $i \geq 0$, define $\text{bin}(a_i)$ to be the binary representation of a_i. Let x_A be the real number using only digits 0, 1, 2 as follows:

$$x_A ::= 0.2\,\text{bin}(a_0)2\,\text{bin}(a_1)2\,\text{bin}(a_2)2\ldots \tag{13.3}$$

We can then define a surjection $f : \mathbb{R} \to \mathcal{P}(\mathbb{N})$ as follows:

$$f(x) = \begin{cases} A & \text{if } x = x_A \text{ for some } A \in \mathbb{N}, \\ 0 & \text{otherwise.} \end{cases}$$

Hence $\mathbb{R}$ surj $\mathcal{P}(\mathbb{N})$. ■

Corollary 13.3.3. $\mathbb{R}$ *is uncountable.*

Proof. By contradiction. Assume $\mathbb{R}$ is countable. Then $\mathbb{N}$ surj $\mathbb{R}$. By Lemma 13.3.2, $\mathbb{R}$ surj $\mathcal{P}(\mathbb{N})$. Hence $\mathbb{N}$ surj $\mathcal{P}(\mathbb{N})$. This contradicts Theorem 13.3.1 for the case when $A = \mathbb{N}$. ■

So the set of rational numbers and the set of natural numbers have the same size, but the set of real numbers is strictly larger. In fact, $\mathbb{R}$ bij $\mathcal{P}(N)$, but we won't prove that here.

Is there anything bigger?

13.3.2 Even Larger Infinities

There are lots of different sizes of infinite sets. For example, starting with the infinite set $\mathbb{N}$ of nonnegative integers, we can build the infinite sequence of sets

$$\mathbb{N},\ \mathcal{P}(\mathbb{N}),\ \mathcal{P}(\mathcal{P}(\mathbb{N})),\ \mathcal{P}(\mathcal{P}(\mathcal{P}(\mathbb{N}))), \ldots$$

By Theorem 13.3.1, each of these sets is strictly bigger than all the preceding ones. But that's not all, the union of all the sets in the sequence is strictly bigger than each set in the sequence. In this way, you can keep going, building still bigger infinities.

13.3.3 The Continuum Hypothesis

Georg Cantor was the mathematician who first developed the theory of infinite sizes (because he thought he needed it in his study of Fourier series). Cantor raised the question whether there is a set whose size is strictly between the "smallest" infinite set, $\mathbb{N}$, and $\mathcal{P}(\mathbb{N})$. He guessed not:

Cantor's Continuum Hypothesis. *There is no set A such that $\mathcal{P}(\mathbb{N})$ is strictly bigger than A and A is strictly bigger than $\mathbb{N}$.*

The Continuum Hypothesis remains an open problem a century later. Its difficulty arises from one of the deepest results in modern Set Theory—discovered in part by Gödel in the 1930s and Paul Cohen in the 1960s—namely, the ZFC axioms are not sufficient to settle the Continuum Hypothesis: there are two collections of sets, each obeying the laws of ZFC, and in one collection, the Continuum Hypothesis is true, and in the other, it is false. So settling the Continuum Hypothesis requires a new understanding of what sets should be to arrive at persuasive new axioms that extend ZFC and are strong enough to determine the truth of the Continuum Hypothesis one way or the other.

13.4 Infinities in Computer Science

If the romance of different size infinities and continuum hypotheses doesn't appeal to you, not knowing about them is not going to lower your professional abilities as a computer scientist. These abstract issues about infinite sets rarely come up in mainstream mathematics, and they don't come up at all in computer science, where the focus is generally on countable, and often just finite, sets. In practice, only logicians and set theorists have to worry about collections that are too big to be sets. In fact, at the end of the 19th century, even the general mathematical community doubted the relevance of what they called "Cantor's paradise" of unfamiliar sets of arbitrary infinite size.

That said, it is worth noting that the proof of Theorem 13.3.1 gives the simplest form of what is known as a "diagonal argument." Diagonal arguments are used to prove many fundamental results about the limitations of computation, such as the undecidability of the Halting Problem for programs and the inherent, unavoidable inefficiency (exponential time or worse) of procedures for other computational problems. So computer scientists do need to study diagonal arguments in order to understand the logical limits of computation. Ad a well-educated computer scientist will be comfortable dealing with countable sets, finite as well as infinite.

6.042J / 18.062J Mathematics for Computer Science
Fall 2010