Extra Problem Set I

Countable and Uncountable Sets

These questions add detail to the discussion we had in class about different types of infinities.

We have a good intuition about counting finite sets (collections). For example, if you were asked how many elements were in the set $\{47, \pi, \sqrt{2}\}$ you would say "Three." If pressed for further justification, you might count the elements off saying "One" (while pointing to 47), "Two" (while pointing to π), and "Three" (while pointing to $\sqrt{2}$). What you are in effect doing is describing a *one-to-one* function

$$f : \{1, 2, 3\} \;\to\; \{47, \pi, \sqrt{2}\}$$

defined by $f(1) = 47$, $f(2) = \pi$, $f(3) = \sqrt{2}$. Saying that this counting function is *one-to-one* means two things: it does not *undercount* the elements of the target set (that is, every element of the target set is an output of f) and it does not *overcount* the elements of the target set (that is, there are no two numbers in the domain which are sent to the same output by f).

You should verify that there are 6 different one-to-one functions from $\{1, 2, 3\}$ to $\{47, \pi, \sqrt{2}\}$; that is, there are 6 different ways of counting off the elements of the set $\{47, \pi, \sqrt{2}\}$.

Here is an example of a function which is undercounting the elements in the target set:

$$g : \{1, 2, 3\} \;\to\; \{a, b, c, d\} : g(1) = a, \; g(2) = b, \; g(3) = c$$

There is no element in the domain which is being sent to d by the function g. This corresponds to our intuition that the set $\{a, b, c, d\}$ has 4 elements, not 3.

Here is an example of a function which is overcounting the elements in the target set:

$$h : \{1, 2, 3\} \;\to\; \{a, b\} : h(1) = a, \; h(2) = b, \; h(3) = b$$

The element b in the target set is being counted "twice" by h. This corresponds to our intuition that the set $\{a, b\}$ has 2 elements, not 3.

We use the idea of one-to-one functions to talk about infinite sets with the same number of elements. To get the discussion started, we introduce the infinite set

$$\mathbb{N} \;=\; \{1, 2, 3, 4, \ldots\}$$

of all counting numbers. We say that a set X is *countably infinite* if there is a one-to-one function from $\mathbb{N}$ to X.

Q1. Show that the set of all negative counting numbers

$$X \;=\; \{-1, -2, -3, \ldots\}$$

is countably infinite. Write down an explicit one-to-one function $f : \mathbb{N} \to X$ and verify that f doesn't overcount or undercount.

Q2. Show that the set of all positive even numbers

$$E = \{2, 4, 6, 8, \ldots\}$$

is countably infinite. Write down an explicit one-to-one function $f : \mathbb{N} \to E$ and verify that f doesn't overcount or undercount.

The example above might seem a little strange at first; after all E is a strictly smaller subset of $\mathbb{N}$. It may feel like we have shown that two times infinity is infinity.

Q3. Show that the cartesian product

$$\mathbb{N} \times \mathbb{N} = \{(p, q) \,|\, p, q \in \mathbb{N}\}$$

is countably infinite using the two steps below.

- *Step 1.* Write out the elements of $\mathbb{N} \times \mathbb{N}$ in an infinite array (for example, extending infinitely down and to the right as shown)
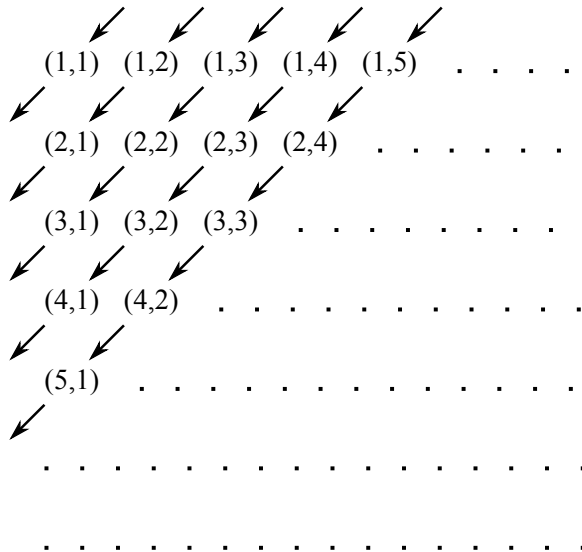
- *Step 2.* Define a function $f : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ by following the diagonal arrowed line segments from left to right as indicated

$$f(1) = (1, 1); \quad f(2) = (1, 2); \quad f(3) = (2, 1); \quad f(4) = (3, 1); \quad f(5) = (2, 2); \quad \cdots$$

(a) Write out the first 30 outputs of the function f.

(b) Verify that f is a one-to-one function (that is, show that f does not overcount and does not undercount).

This example above may feel like we have shown that infinity times infinity is infinity.

Q4. Show that the set of all positive rational numbers

$$\mathbb{Q}_+ = \{p/q \,|\, p, q \in \mathbb{N}\}$$

is countable.

Define a function $g : \mathbb{N} \to \mathbb{Q}_+$ using the function $f : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ of the previous problem as follows. Convert each output of f from an ordered pair (p, q) into a rational number p/q and omit output numbers that repeat. So $g(1) = 1/1 = 1$, $g(2) = 1/2$, $g(3) = 2/1 = 2$, $g(4) = 1/3$. The next fraction is $2/2$, but since this is equal to 1 which has already appeared as the output $g(1)$ we move on to the next fraction and write $g(5) = 3/1 = 3$.

(a) Write down the first 30 positive rational numbers that the function g counts. (the first 5 are written out above)

(b) Convince yourself that g is a one-to-one function. Write down an argument which says why g doesn't undercount, and why g doesn't overcount.

Q5. Show that the set of all real numbers in the interval $(0, 1)$ is *uncountably infinite*. Show that there is no one-to-one function $f : \mathbb{N} \to (0, 1)$. Do this by arguing that any function $f : \mathbb{N} \to (0, 1)$ which tries to count all the numbers in $(0, 1)$ is doomed to failure; by its nature it must undercount (by a huge amount).

(a) Every number in $(0, 1)$ can be represented by an infinite decimal of the form

$$0.a_1 a_2 a_3 \ldots \qquad (*)$$

where the digits a_i belong to $\{0, 1, \ldots, 9\}$.

For example $1/2$ is 0.5 and so $a_1 = 5$, $a_2 = 0$, $a_3 = 0$ etc.

Another example, $1/3$ is $0.33\ldots$ and so $a_1 = 3$, $a_2 = 3$, $a_3 = 3$, etc.

Write an argument which explains why this decimal expansion representation of a number is almost unique. The only ambiguity is that a decimal expansion with an infinite tail of 9's (say that $a_i = 9$ from some index $i = m$ onwards and a_{m-1} is not equal to 9) represents the same number as the decimal where a_1 upto a_{m-2} are unchanged, a_{m-1} is increased by 1, and a_i is set to 0 for all $i \geq m$.

So if we rule out decimal expansions with infinite tails of 9's, we have a unique representation of numbers in $(0, 1)$ and decimal expansion of the form described in $(*)$ above.

(b) We show that any putative counting function $f : \mathbb{N} \to (0, 1)$ is necessarily undercounting by showing that there are lots of decimal expansions of the form $(*)$ which are not targets of f. Because of the uniqueness of representation, this means that there are lots of numbers in $(0, 1)$ that are not being counted by f.

Complete the following argument.

Write out the list of outputs of the putative counting function $f : \mathbb{N} \to (0, 1)$ in decimal expansion as shown.

$$\begin{aligned}
f(1) &= 0.a_{1,1}a_{1,2}a_{1,3}a_{1,4}\ldots \\
f(2) &= 0.a_{2,1}a_{2,2}a_{2,3}a_{2,4}\ldots \\
f(3) &= 0.a_{3,1}a_{3,2}a_{3,3}a_{3,4}\ldots \\
f(4) &= 0.a_{4,1}a_{4,2}a_{4,3}a_{4,4}\ldots \\
&\vdots \qquad \vdots
\end{aligned}$$

Here each of the digits $a_{i,j}$ is a number in $\{0, 1, 2, \ldots, 9\}$, and no decimal $0.a_{i,1}a_{i,2}a_{i,3}\ldots$ ends in an infinite string of 9's.

i. Show that f undercounts, by describing a number in $(0, 1)$ (in decimal form) which is not in the list of outputs above.

ii. Show how to obtain many more numbers which are not in the list of outputs above.

Remarks. This is the start of several very interesting stories.

(a) We saw that the rational numbers are countably infinite, and that all the real numbers (in $(0, 1)$) are uncountably infinite. So what are the extra numbers? These are the irrational numbers. Questions 4 and 5 above combine to prove that irrational numbers exist, without telling you which numbers are irrational.

In fact most numbers are irrational. If you recall (or prove this directly) rational numbers are precisely those which have decimal expansions which terminate (like 0.5) or else have an infinite repeating pattern (like 0.23232323...). So it is easy to write down a decimal expansion of an irrational number; e.g.,

$$0.101001000100001000001\ldots$$

where the number of 0's in each successive grouping grows. Just ensure that your number does not terminate and has no repeating pattern.

(b) It is harder to determine whether a given number is irrational. For example $\sqrt{2}$ is irrational; π is irrational.

(c) It is possible to do a counting argument analogous to the one in problem 3 to show that *algebraic numbers* are countable. Algebraic numbers are generalizations of rational numbers defined as follows.

The rational number $x = p/q$ is the solution of a linear equation with integer coefficients; namely, $qx - p = 0$.

An algebraic number is a number which occurs as the solution of a polynomial equation with integer coefficients. For example, $\sqrt{2}$ is algebraic, since it is a solution of the quadratic equation $1x^2 + 0.x - 2 = 0$.

Combining the result that algebraic numbers are countably infinite with the result of question 5 above we see that there must exist numbers which are not the solutions of polynomial equations with integer coefficients. These numbers are called *transcendental*. For example π is transcendental.

(d) Every real number has a decimal expansion. We can chop off the tail of this expansion at m places after the decimal point and obtain a rational number. This rational number will approximate the original real number to within $\frac{1}{10^m}$. So every real number can be approximated arbitrarily closely by rational numbers. Nevertheless, the collection of all rational numbers can be covered by a collection of intervals of arbitrarily small total length. So the line gets its infinite length from the other numbers. In terms of length, there are many more irrational numbers than rational numbers. If you find this strangely compelling, you could check out an Introduction to Analysis class.

(e) A more sophisticated version of the diagonal argument in Question 5 above (due to the mathematician Georg Cantor) can be used to show that there is a tower of increasingly larger infinities. For example the set of all subsets of real numbers is a set that is truly larger that the set of real numbers; there is no one-to-one function between these two sets.