# Bastardized Fundamentals of Computer Science

## Part 1: Countability

Marc Bender
lemma@0x90.org

This is truly abstract math—you've been warned. Proceed only if you like exercising your brain every once in a while. Did you know that there are different sizes of infinity? Well, there are. So here's a simple proof of that fact, using nothing but the natural numbers and the real numbers.

The easiest type of infinity to understand is *countable* infinity. Not that you could actually count everything in an infinite bunch—that's stupid. But let's say you're standing on the side of an infinite highway, and you're infinitely bored. You notice that along the side of the highway is an infinite line of telephone poles, with the first one right beside you. Seeing that you've got nothing better to do, you set out walking down the side of this highway counting poles. Better add the assumption here that you're not drunk or psychotic and your eyes work all right. Those poles are an example of countable infinity. They're also an example of bad analogy.

The key thing here is that you can devise a method by which, given an infinite amount of time, you *could* count all the objects, i.e. you're not missing any of them along the way. You could walk along with a marker and write "1" on the first one, "2" on the second one, and so on, and you can picture the process going on until you've marked as many of them as you want. What you're effectively doing is mapping each object to a natural (counting) number 0, 1, 2, 3, . . . .

There are all kinds of sets that are countable. Take the integers for example. To count them, you just have to write them like this:

$$0, -1,\ 1, -2,\ 2, -3,\ 3, \ldots$$

Following that method, you can get any integer you want... eventually. A more interesting example is rational numbers, that is fractions. We can count those by laying them out like this, and following the arrows along the potentially infinite path:

$$\begin{array}{ccccccccc}
0 & \rightarrow & \frac{1}{2} & & \frac{1}{3} & \rightarrow & \frac{1}{4} & & \frac{1}{5} \rightarrow \cdots \\
 & \swarrow & & \nearrow & & \swarrow & & \nearrow & \\
\frac{2}{1} & & \frac{2}{2} & & \frac{2}{3} & & \frac{2}{4} & & \\
\downarrow & \nearrow & & \swarrow & & \nearrow & & & \\
\frac{3}{1} & & \frac{3}{2} & & \frac{3}{3} & & & & \\
 & \swarrow & & \nearrow & & & & & \\
\frac{4}{1} & & \frac{4}{2} & & & & & & \\
\downarrow & \nearrow & & & & & & & \\
\frac{5}{1} & & & & & & & &
\end{array}$$

We could keep going like this, and since we can get any numerator/denominator combination, we can get any fraction. Speaking in infinite terms, there are the same "number" of fractions as there are natural numbers. Whoa. (Yes, you can easily include the negative rationals, basically the same way as we did the integers above)

What I'm going to show now is that there are some kinds of infinity that you *can't* count. We're going to do a bit of abstract thinking here, so maybe you should drink a couple of beers before you read this (works for me anyway). Like I mentioned before, we'll use the real numbers as an example.

First off, we'll write all real numbers with an infinite number of decimal places. That's not some kind of cheap trick, because you can easily write 4 as $4.00000\cdots$, and for that matter tacking zeroes onto the end of any terminating real number is cool (by cool I mean identity-preserving). To make things easier, we'll deal only with the positive real numbers between 0 and 1, because if you can't count those then you sure as shit can't count them all.

For argument's sake let's say you *could* count all the real numbers between 0 and 1. Then you'd be able to list them from the first one onward, like we did above for the integers and fractions. The list could look something like $\frac{1}{\pi}$, 0.5, 0, $0.\overline{857}, \ldots$, or in decimal notation:

$$\begin{array}{rcccl}
x_1 = & \frac{1}{\pi} & = .3183098\cdots \\
x_2 = & 0.5 & = .5000000\cdots \\
x_3 = & 0 & = .0000000\cdots \\
x_4 = & 0.\overline{857} & = .8578578\cdots \\
\vdots & \vdots & \vdots
\end{array}$$

I snuck in some variables there, because that's what we're going to be dealing with from now on. Before you start thinking too much, remember that the list above is just an *absurd example* of a list of real numbers, the numbers I chucked in there basically came directly out of my ass. But hell, if you want to imagine that my ass is some sort of ideal random real number generator, that's up to you.

Now we have to deal with the actual decimal places of a particular number. Take our example value for $x_4$ above, i.e. $x_4 = 0.\overline{857}$. We want to break it down one step further to be able to reference particular decimal places within $x_4$. Square brackets should do the trick; so let's have $x_4[i]$

represent the $i$th decimal place of $x_4$. So we write

$$x_4[1] = 8,\ x_4[2] = 5,\ x_4[3] = 7,\ x_4[4] = 8,\ \ldots$$

This is where we get purely abstract. Take the list of real numbers above, $x_1, x_2, x_3, \ldots$ Instead of thinking of each individual variable $x_i$ separately, think of the *whole list* as one variable, i.e. $x = (x_1, x_2, x_3, \ldots)$. That one variable, $x$, can now represent *any possible list of real numbers* between 0 and 1; that's the key to the whole hootenanny. If there *is* a way of counting these real numbers the way we did the integers and rationals, then $x$ can represent it.

We've got our notation down now, so we can draw up our list. Remember, $x_1, x_2, x_3, \ldots$ represents *any* possible list of real numbers between 0 and 1.

$$\begin{array}{rcccccl} x_1 = & .\,x_1[1] & x_1[2] & x_1[3] & x_1[4] & \cdots \\ x_2 = & .\,x_2[1] & x_2[2] & x_2[3] & x_2[4] & \cdots \\ x_3 = & .\,x_3[1] & x_3[2] & x_3[3] & x_3[4] & \cdots \\ x_4 = & .\,x_4[1] & x_4[2] & x_4[3] & x_4[4] & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

Time for some action. Introducing our all-purpose supervariable, $y$ (okay, it's just a normal variable). We're going to turn it into a real number between 0 and 1 that's not in the list. We'll deal with what exactly that means afterwards.

Like the $x_i$'s, we'll expand $y$ to $y = .\,y[1]\ y[2]\ y[3] \cdots$ where $y[i]$ is the $i$th decimal place of $y$. What we're going for here is simple: we want to make $y$ different from every $x_i$ in the list, thereby showing that $y$ *can't possibly be in the list*. But how the hell can you make $y$ different than *every* number in that infinite list? By bustin' this sick trick called diagonalisation, yo.

Let's start by making sure that $y$ is different than $x_1$. For two real numbers to be different, they have to differ in *at least one* decimal place, right? The first decimal place is a good place to start. So let's make $y \neq x_1$ by making $y$ and $x_1$ different in their first decimal places ($y[1] \neq x_1[1]$). One way of doing this is by defining $y[1] = x_1[1] + 1$, or in the case where $x_1[1] = 9$, $y[1] = 0$. Now we've guaranteed that $y \neq x_1$, and we've got ourselves the first decimal place of $y$.

Next step: guarantee that $y \neq x_2$. We can do *this* by making sure that $y$ and $x_2$ are different in the *second* decimal place. Using the same construction as above, set $y[2] = x_2[2] + 1$ if $x_2 < 9$, or $y[2] = 0$ if $x_2 = 9$. Bam! There's the second decimal place taken care of. Pulling the same stunt to make sure that $y \neq x_3$, we get $y[3] = x_3[3] + 1$ if $x_3[3] < 9$, $y[3] = 0$ otherwise. There's a pattern emerging here, and I'm going to make it clear in a second.

$$\begin{array}{rcccccl} x_1 = & .\,\boxed{x_1[1]} & x_1[2] & x_1[3] & x_1[4] & \cdots \\ x_2 = & .\,x_2[1] & \boxed{x_2[2]} & x_2[3] & x_2[4] & \cdots \\ x_3 = & .\,x_3[1] & x_3[2] & \boxed{x_3[3]} & x_3[4] & \cdots \\ x_4 = & .\,x_4[1] & x_4[2] & x_4[3] & \boxed{x_4[4]} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

We're effectively *building* $y$ out of the $x_i$'s, while at the same time guaranteeing that $y$ can't be equal to any of them! We do this by moving simultaneously down the list ($x_1$ to $x_2$ to $x_3$ to ...) and across the decimal places ([1] to [2] to [3] to ...). This is demonstrated by the boxed values in the list above. That's why it's called diagonalisation, dude.

The way we did the first through third decimal places above can be generalised to the $i$th decimal place by defining $y$ like this:

$$y[i] = \begin{cases} x_i[i] + 1 & \text{if } x_i[i] < 9 \\ 0 & \text{otherwise.} \end{cases}$$

or better, using modulo:

$$y[i] = (x_i[i] + 1) \bmod 10$$

Now we've got what we were looking for. For every $i$, we're 100% sure that $y \neq x_i$, because we're 100% sure that $y[i] \neq x_i[i]$. Notice that we've only guaranteed that $y$ and $x_i$ are different in one decimal place, but they could very well be different in a shitload of other decimal places. That just doesn't have any bearing on the proof.

Since $y$ can't be equal to *any* number in the list $x_1, x_2, \ldots$, it follows that it can't be in the list at all. But remember the mental gymnastics that got us here; $x_1, x_2, \ldots$ represents *any possible listing* of real numbers between 0 and 1. So if we managed to find a number outside of this infinite list, that can only mean one thing... the list is *too small* to hold all of the real numbers between 0 and 1! In other words, the real numbers between 0 and 1 constitute a "bigger" infinity than the natural numbers, integers or rationals. Extending this to cover *all* of the real numbers is just a matter of playing with decimal point placement.

I wrote this paper because I think that diagonalisation is a truly awesome mathematical tool and I wanted to try to explain it better than it first was explained to me. If I've failed you, then shit dude, I owe you a metabeer.

I guess there's only one more thing to say… booyah! What a waste of 3 (correction—5) perfectly good hours.

## References

[1] Kleene, S.C. *Introduction to Metamathematics*, North-Holland, New York, 1952.

[2] Zucker, J. and L. Pretorius. Introduction to Computability Theory, *South African Computer Journal*, 9, April 1993, 3-30.