

# 15-251

## Great Theoretical Ideas in Computer Science

### Cantor's Legacy: Infinity and Diagonalization

Lecture 23 (November 9, 2010)



#### Ideas from the course

Induction

Numbers, Number theory and Algebra

Representation

Finite Counting and Probability

Automata and Computation

#### A hint of the infinite

Infinite row of dominoes

Infinite sums (formal power series)

Infinite choice trees, and infinite probability

#### Infinite RAM Model

##### Platonic Version:

One memory location for each  
natural number 0, 1, 2, ...



##### Aristotelian Version:

Whenever you run out of memory,  
the computer contacts the factory.  
A maintenance person is flown by  
helicopter and attaches 1000 Gig of  
RAM and all programs resume their  
computations, as if they had never  
been interrupted.



**The Ideal Computer:**  
no bound on amount of memory  
no bound on amount of time

Ideal Computer is defined as a  
computer with infinite RAM.

You can run a Java program and never have  
any overflow, or out of memory errors.

#### An Ideal Computer

It can be programmed to print out:

2: 2.000000000000000000000000...

1/3: 0.333333333333333333333333...

$\phi$ : 1.6180339887498948482045...

e: 2.7182818284559045235336...

$\pi$ : 3.14159265358979323846264...

## Printing Out An Infinite Sequence..

A program  $P$  prints out the infinite sequence

$s_0, s_1, s_2, \dots, s_k, \dots$

if when  $P$  is executed on an ideal computer, it outputs a sequence of symbols such that

-The  $k^{\text{th}}$  symbol that it outputs is  $s_k$

-For every  $k \in \mathbb{N}$ ,  $P$  eventually outputs the  $k^{\text{th}}$  symbol. I.e., the delay between symbol  $k$  and symbol  $k+1$  is not infinite.

## Computable Real Numbers

A real number  $R$  is computable if there is a program that prints out the decimal representation of  $R$  from left to right.

Thus, each digit of  $R$  will eventually be output.



Are all real numbers computable?

## Describable Numbers

A real number  $R$  is describable if it can be denoted unambiguously by a finite piece of English text.

2: "Two."

$\pi$ : "The area of a circle of radius one."

Are all real numbers describable?



Is every computable real number, also a describable real number?

And what about the other way?

Computable  $R$ : some program outputs  $R$   
Describable  $R$ : some sentence denotes  $R$



## Computable $\Rightarrow$ describable

Theorem:

Every computable real is also describable

Proof:

Let  $R$  be a computable real that is output by a program  $P$ . The following is an unambiguous description of  $R$ :

"The real number output by the following program:"  $P$

**MORAL:** A computer program can be viewed as a description of its output.


Syntax: The text of the program  
Semantics: The real number output by  $P$



Are all reals describable?  
Are all reals computable?

We saw that  
computable  $\Rightarrow$   
describable,  
but do we also have  
describable  $\Rightarrow$   
computable?

Questions we will answer in this (and next) lecture...



## Correspondence Principle

If two finite sets can be placed into 1-1 onto (bijective) correspondence, then they have the same size.

## Correspondence Definition

In fact, we can use the correspondence as the definition:

Two finite sets are defined to have the same size if and only if they can be placed into 1-1 onto (bijective) correspondence.

## Georg Cantor (1845-1918)



## Cantor's Definition (1874)

Two sets are defined to have the same size if and only if they can be placed into 1-1 onto correspondence.

If there exists a bijection between them.

## Cantor's Definition (1874)

Two sets are defined to have the same cardinality if and only if they can be placed into 1-1 onto correspondence.

If there exists a bijection between them.

Do  $\mathbb{N}$  and  $\mathbb{E}$  have the same cardinality?

$$\mathbb{N} = \{ 0, 1, 2, 3, 4, 5, 6, 7, \dots \}$$

$$\mathbb{E} = \{ 0, 2, 4, 6, 8, 10, 12, \dots \}$$

The even, natural numbers.

$\mathbb{E}$  and  $\mathbb{N}$  do not have the same cardinality!  $\mathbb{E}$  is a proper subset of  $\mathbb{N}$  with plenty left over.



The attempted correspondence  $f(x)=x$  does not take  $\mathbb{E}$  onto  $\mathbb{N}$ .

$\mathbb{E}$  and  $\mathbb{N}$  do have the same cardinality!

$$\mathbb{N} = 0, 1, 2, 3, 4, 5, \dots$$

$$\mathbb{E} = 0, 2, 4, 6, 8, 10, \dots$$

$f(x) = 2x$  is 1-1 onto.



Lesson:

Cantor's definition only requires that *some* 1-1 correspondence between the two sets is onto, not that all 1-1 correspondences are onto.

This distinction never arises when the sets are finite.



## Cantor's Definition (1874)

Two sets are defined to have the same size if and only if they can be placed into 1-1 onto correspondence.

You just have to get used to this slight subtlety in order to argue about infinite sets!



Do  $\mathbb{N}$  and  $\mathbb{Z}$  have the same cardinality?

$$\mathbb{N} = \{ 0, 1, 2, 3, 4, 5, 6, 7, \dots \}$$

$$\mathbb{Z} = \{ \dots, -2, -1, 0, 1, 2, 3, \dots \}$$



No way!  $\mathbb{Z}$  is infinite in two ways: from 0 to positive infinity and from 0 to negative infinity.

Therefore, there are far more integers than naturals.

**Actually, no!**

$\mathbb{N}$  and  $\mathbb{Z}$  do have the same cardinality!

$$\mathbb{N} = 0, 1, 2, 3, 4, 5, 6 \dots$$

$$\mathbb{Z} = 0, 1, -1, 2, -2, 3, -3, \dots$$

$$f(x) = \begin{cases} \lceil x/2 \rceil & \text{if } x \text{ is odd} \\ -x/2 & \text{if } x \text{ is even} \end{cases}$$



### Transitivity Lemma

**Lemma:** If

$f: A \rightarrow B$  is a bijection, and

$g: B \rightarrow C$  is a bijection.

Then  $h(x) = g(f(x))$  defines a function

$h: A \rightarrow C$  that is a bijection too.

Hence,  $\mathbb{N}$ ,  $\mathbb{E}$ , and  $\mathbb{Z}$  all have the same cardinality.

Do  $\mathbb{N}$  and  $\mathbb{Q}$  have the same cardinality?

$$\mathbb{N} = \{ 0, 1, 2, 3, 4, 5, 6, 7, \dots \}$$

$\mathbb{Q}$  = The Rational Numbers




No way!

The rationals are dense: between any two there is a third. You can't list them one by one without leaving out an infinite number of them.


**Don't jump to conclusions!**

There is a clever way to list the rationals, one at a time, without missing a single one!



First, let's warm up with another interesting example:

$\mathbb{N}$  can be paired with  $\mathbb{N} \times \mathbb{N}$



**Theorem:  $\mathbb{N}$  and  $\mathbb{N} \times \mathbb{N}$  have the same cardinality**

**Theorem:  $\mathbb{N}$  and  $\mathbb{N} \times \mathbb{N}$  have the same cardinality**

...						
4	o	o	o	o	o	
3	o	o	o	o	o	
2	o	o	o	o	o	
1	o	o	o	o	o	
0	o	o	o	o	o	
	0	1	2	3	4	...

The point (x,y) represents the ordered pair (x,y)

**Theorem:  $\mathbb{N}$  and  $\mathbb{N} \times \mathbb{N}$  have the same cardinality**

...						
4	o	o	o	o	o	
3	6	o	o	o	o	
2	3	7	o	o	o	
1	1	4	o	o	o	
0	0	2	5	o	o	
	0	1	2	3	4	...

The point (x,y) represents the ordered pair (x,y)

**The first few tuples output...**

(0,0)  
 (0,1), (1,0)  
 (0,2), (1,1), (2,0)  
 (0,3), (1,2), (2,1), (3,0)  
 ...

### Defining bijection $f: \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$

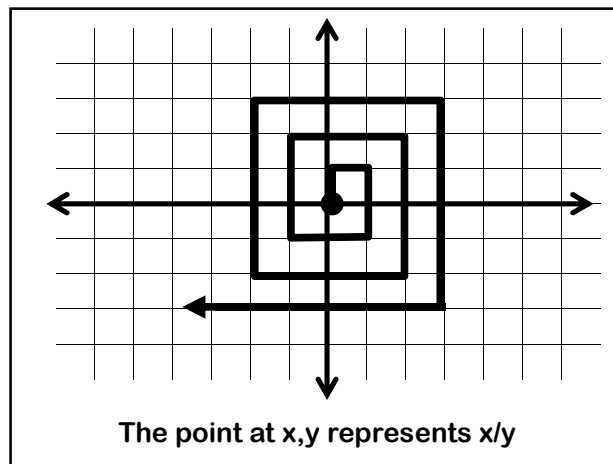
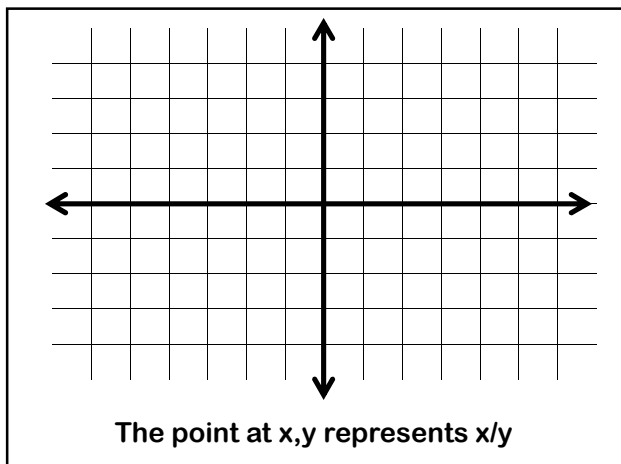
```

let i := 0;    //will range over N

for (sum = 0 to forever) {
  //generate all pairs with this sum
  for (x = 0 to sum) {
    y := sum-x
    define f(i) := the point (x,y)
    i++;
  }
}

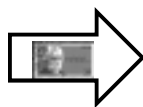
```

Onto the Rationals!



Cantor's 1877 letter to Dedekind:

*"I see it, but I don't believe it!"*



### Countable Sets

We call a set countable if it can be placed into 1-1 onto correspondence with the natural numbers  $\mathbb{N}$ .

Hence

$\mathbb{N}$ ,  $\mathbb{E}$ ,  $\mathbb{Q}$  and  $\mathbb{Z}$  are all countable.

Do  $\mathbb{N}$  and  $\mathbb{R}$  have the same cardinality?

$\mathbb{N} = \{ 0, 1, 2, 3, 4, 5, 6, 7, \dots \}$

$\mathbb{R}$  = The Real Numbers

No way!

You will run out of natural numbers long before you match up every real.



Now hang on a minute!

You can't be sure that there isn't some clever correspondence that you haven't thought of yet.



I am sure!  
Cantor proved it.

To do this, he invented a very important technique called "Diagonalization"



**Theorem: The set  $\mathbb{R}_{[0,1]}$  of reals between 0 and 1 is not countable.**

**Proof: (by contradiction)**

Suppose  $\mathbb{R}_{[0,1]}$  is countable.

Let  $f$  be a bijection from  $\mathbb{N}$  to  $\mathbb{R}_{[0,1]}$ .

Make a list  $L$  as follows:

0: decimal expansion of  $f(0)$   
1: decimal expansion of  $f(1)$   
2: decimal expansion of  $f(2)$   
...  
k: decimal expansion of  $f(k)$   
...

**Theorem: The set  $\mathbb{R}_{[0,1]}$  of reals between 0 and 1 is not countable.**

**Proof: (by contradiction)**

Suppose  $\mathbb{R}_{[0,1]}$  is countable.

Let  $f$  be a bijection from  $\mathbb{N}$  to  $\mathbb{R}_{[0,1]}$ .

Make a list  $L$  as follows:

0: 0.333333333333333333...  
1: 0.314159265657839593...  
2: 0.125912591259125912...  
...  
k: 0.235094385543905834...  
...



Position after decimal point

L	0	1	2	3	4	...
0						
1						
2						
3						
...						

Position after decimal point

L	0	1	2	3	4	...
0	3	3	3	3	3	3
1	3	1	4	1	5	9
2	1	2	5	9	1	2
3	4	1	2	5	6	8
...						

digits along the diagonal

L	0	1	2	3	4	...
0	$d_0$					
1		$d_1$				
2			$d_2$			
3				$d_3$		
...						...

L	0	1	2	3	4
0	$d_0$				
1		$d_1$			
2			$d_2$		
3				$d_3$	
...					...

Define the following real number  
 $\text{Confuse}_L = . C_0 C_1 C_2 C_3 C_4 C_5 \dots$

L	0	1	2	3	4
0	$d_0$				
1		$d_1$			
2			$d_2$		
3				$d_3$	
...					...

Define the following real number  
 $\text{Confuse}_L = . C_0 C_1 C_2 C_3 C_4 C_5 \dots$

$$C_k = \begin{cases} 5, & \text{if } d_k=6 \\ 6, & \text{otherwise} \end{cases}$$

L	0	1	2	3	4
0	$C_0 \neq d_0$	$C_1$	$C_2$	$C_3$	$C_4$
1		$d_1$			
2			$d_2$		
3				$d_3$	
...					...

$$C_k = \begin{cases} 5, & \text{if } d_k=6 \\ 6, & \text{otherwise} \end{cases}$$

L	0	1	2	3	4
0	$d_0$				
1	$C_0$	$C_1 \neq d_1$	$C_2$	$C_3$	$C_4$ ...
2			$d_2$		
3				$d_3$	
...					...

$$C_k = \begin{cases} 5, & \text{if } d_k=6 \\ 6, & \text{otherwise} \end{cases}$$

L	0	1	2	3	4
0	$d_0$				
1		$d_1$			
2	$C_0$	$C_1$	$C_2 \neq d_2$	$C_3$	$C_4$ ...
3				$d_3$	
...					...

$$C_k = \begin{cases} 5, & \text{if } d_k=6 \\ 6, & \text{otherwise} \end{cases}$$

### Diagonalized!

By design, Confuse<sub>L</sub> can't be on the list L!

Confuse<sub>L</sub> differs from the k<sup>th</sup> element on the list L in the k<sup>th</sup> position.

This contradicts the assumption that the list L is complete; i.e., that the map  $f: \mathbb{N} \rightarrow \mathbb{R}_{[0,1]}$  is surjective.

The set of reals is uncountable!  
(Even the reals between 0 and 1.)

An aside: you can set up a correspondence between  $\mathbb{R}$  and  $\mathbb{R}_{[0,1]}$ .



Hold it!  
Why can't the same argument be used to show that the set of rationals  $\mathbb{Q}$  is uncountable?



The argument is the same for  $\mathbb{Q}$  until the punchline.

However, since CONFUSE<sub>L</sub> is not necessarily rational, so there is no contradiction from the fact that it is missing from the list L.



## Another diagonalization proof

### Problem from a 15-251 final:

Show that the set of real numbers in  $[0, 1]$  whose decimal expansion has the property that every digit is a prime number (2, 3, 5, or 7) is uncountable.

E.g., 0.2375 and 0.5555... are in the set, but 0.14555... and 0.3030303... are not.

## Another diagonalization proof

Show that the set of real numbers in  $[0, 1]$  whose decimal expansion has the property that every digit is a prime number (2, 3, 5, or 7) is uncountable.

## Another diagonalization proof

Show that the set of real numbers in  $[0, 1]$  whose decimal expansion has the property that every digit is a prime number (2, 3, 5, or 7) is uncountable.

## Another diagonalization proof

Show that the set of real numbers in  $[0, 1]$  whose decimal expansion has the property that every digit is a prime number (2, 3, 5, or 7) is uncountable.

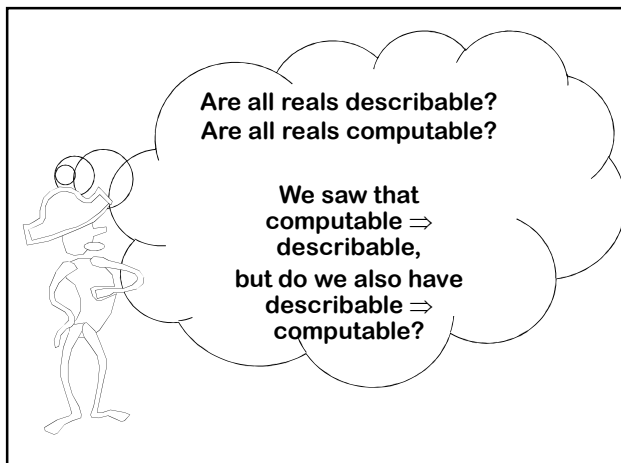
- A) Assume this set is countable and therefore it can be placed in a list  $L$ .
- B) Given  $L$ , show how to define a number called Confuse.
- C) Show that Confuse is not in  $L$ .
- D) Explain why Confuse not being in  $L$  implies the set is not countable.

## Countable and Uncountable

$\mathbb{N}$ ,  $\mathbb{E}$ ,  $\mathbb{Q}$  and  $\mathbb{Z}$  are all countable sets

$\mathbb{R}$  is an uncountable set

**Back to the questions  
we were asking earlier**



## Standard Notation

$\Sigma =$  Any finite alphabet  
Ex.:  $\{a,b,c,d,e,\dots,z\}$  or  $\{a,b\}$  or  $\{0,1\}$

$\Sigma^* =$  All finite strings of symbols from  $\Sigma$   
including the empty string  $\epsilon$

### Theorem: Every subset $S$ of $\Sigma^*$ is countable

Try #1:

Sort  $S$  alphabetically, map first word to 0, second word to 1, and so on...

What if  $S =$

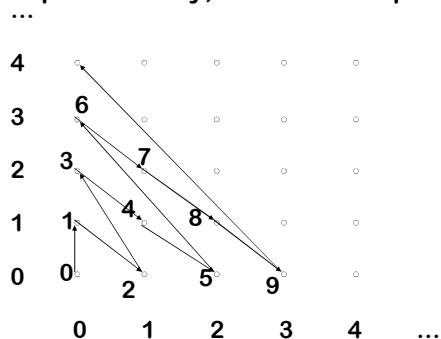
$\{a, b, aa, bb, aaa, bbb, aaaa, bbbb, \dots\}$  ?

### Theorem: Every subset $S$ of $\Sigma^*$ is countable

Proof: Sort  $S$  by first by length and then alphabetically.

Map the first word to 0, the second to 1, and so on....

This sorting on length, and then alphabetically, is similar in spirit to...



## Stringing Symbols Together

$\Sigma =$  The symbols on a standard keyboard

For example:


The set of all possible Java programs is a subset of  $\Sigma^*$

The set of all possible finite pieces of English text is a subset of  $\Sigma^*$

**Thus:**


The set of all possible Java programs is countable.

The set of all possible finite length pieces of English text is countable.



There are countably many Java programs and uncountably many reals.


Hence,  
Most reals are not computable!



**I see!**

There are countably many descriptions and uncountably many reals.


Hence:  
Most real numbers are not describable!



Are all reals describable? **NO**


Are all reals computable? **NO**

We saw that computable  $\Rightarrow$  describable, but do we also have describable  $\Rightarrow$  computable?




Is there a real number that can be described, but not computed?

Wait till the next lecture!



We know there are at least 2 infinities. (the number of naturals, the number of reals.)

Are there more?



### Definition: Power Set

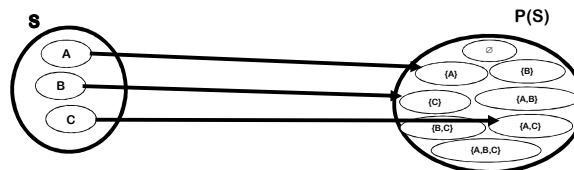
The power set of  $S$  is the set of all subsets of  $S$ .

The power set is denoted as  $\mathcal{P}(S)$ .

**Proposition:**

If  $S$  is finite, the power set of  $S$  has cardinality  $2^{|S|}$

**Theorem:**  $S$  can't be put into bijection with  $\mathcal{P}(S)$



Suppose  $f: S \rightarrow \mathcal{P}(S)$  is a bijection.

Let  $\text{CONFUSE}_f = \{x \mid x \in S, x \notin f(x)\}$

Since  $f$  is onto, exists  $y \in S$  such that  $f(y) = \text{CONFUSE}_f$ .  
Is  $y$  in  $\text{CONFUSE}_f$ ?

YES: Definition of  $\text{CONFUSE}_f$  implies no

NO: Definition of  $\text{CONFUSE}_f$  implies yes

This proves that there are at least a countable number of infinities.

The first infinity is called:

$\aleph_0$



$\aleph_0, \aleph_1, \aleph_2, \dots$

Are there any more infinities?



$\aleph_0, \aleph_1, \aleph_2, \dots$

Let  $S = \{\aleph_k \mid k \in \mathbb{N}\}$   
 $\mathcal{P}(S)$  is provably larger than any of them.



In fact, the same argument can be used to show that no single infinity is big enough to count the number of infinities!



$\aleph_0, \aleph_1, \aleph_2, \dots$ 

Cantor wanted to show that the number of reals was  $\aleph_1$



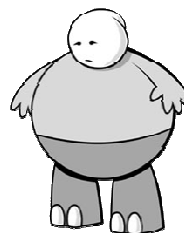
Cantor called his conjecture that  $\aleph_1$  was the number of reals the "Continuum Hypothesis."

However, he was unable to prove it. This helped fuel his depression.



The Continuum Hypothesis can't be proved or disproved from the standard axioms of set theory!

This has been proved!



Here's What You Need to Know...

**Cantor's Definition:**  
Two sets have the same cardinality if there exists a bijection between them.

E, N, Z and Q all have same cardinality (and proofs)

Proof that there is no bijection between N and R

Countable versus Uncountable

Power sets and their properties