

# BU CS 332 – Theory of Computation

<https://forms.gle/MBf73tYPrgZNcfYm9>



## Lecture 13:

- Uncountability
- Undecidability

Reading:

Sipser Ch 4.2, 5.1

Alexander Poremba & Mark Bun

March 19, 2026

# Where we are and where we're going

Church-Turing thesis: TMs capture all algorithms

Consequence: studying the limits of TMs reveals the limits of computation

**Last time:** Sizes of infinite sets, countability

**Today:** Uncountable sets

Existential proof that there are undecidable and unrecognizable languages

An explicit undecidable language?

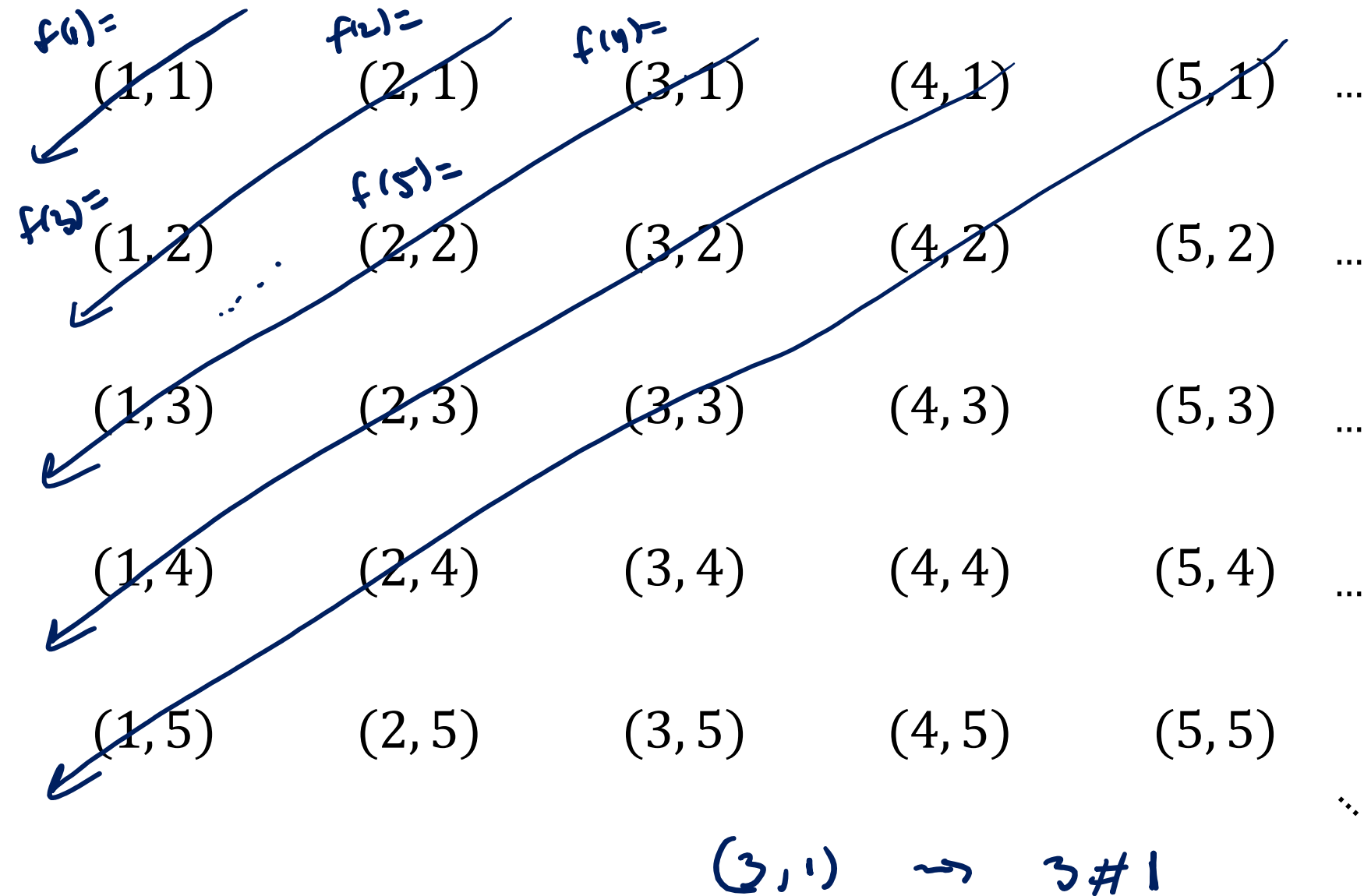
# How can we compare sizes of infinite sets?

**Definition:** Two sets have **the same size** if there is a bijection between them

A set is **countable** if either

- it is a finite set, or
- it has the same size as  $\mathbb{N}$ , the set of natural numbers

# How to show that $\mathbb{N} \times \mathbb{N}$ is countable?



# How to argue that a set $S$ is countable

- Describe how to “list” the elements of  $S$ , usually in stages:

**Ex:** Stage 1) List all pairs  $(x, y)$  such that  $x + y = 2$

Stage 2) List all pairs  $(x, y)$  such that  $x + y = 3$

...

Stage  $n$ ) List all pairs  $(x, y)$  such that  $x + y = n + 1$

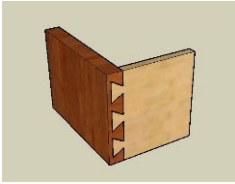
...

- Explain why every element of  $S$  appears in the list

**Ex:** Any  $(x, y) \in \mathbb{N} \times \mathbb{N}$  will be listed in stage  $x + y - 1$

- Define the bijection  $f: \mathbb{N} \rightarrow S$  by  $f(n) =$  the  $n$ 'th element in this list (ignoring duplicates if needed)

# Another version of the dovetailing trick



**Ex:** Show that  $\mathcal{F} = \{L \subseteq \{0, 1\}^* \mid L \text{ is finite}\}$  is countable  
 $\{\epsilon, 0, 011\} \in \mathcal{F}$        $\{0^n \mid n \geq 0\} \notin \mathcal{F}$

Idea that doesn't work:

stage 0:	List all languages $L$ s.t. $ L  = 0$	$\emptyset$
stage 1:	" " " " $ L  = 1$	$\{\epsilon\}, \{0\}, \{1\}, \{00\}, \{01\}, \dots$
stage 2:	" " " " $ L  = 2$	
...		

↙ will never get to stage 2

Proof. For a finite language  $L$ , define  $m(L) = \text{length of longest string in } L$   
 e.g.  $m(\{\epsilon, 0, 011\}) = 4$

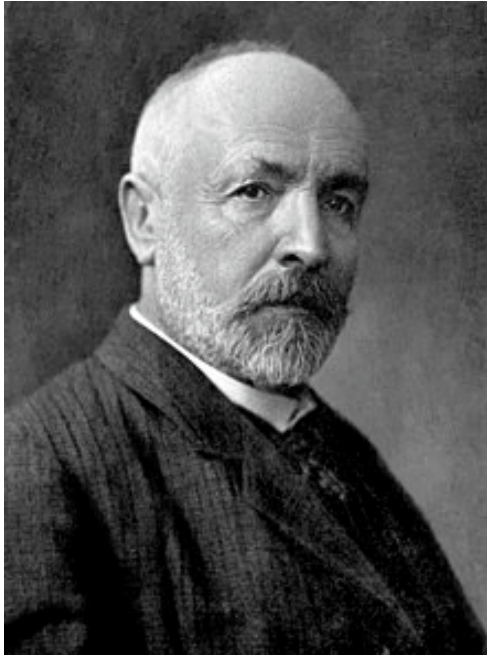
List elements of  $\mathcal{F}$  in stages:

<u>stage 1:</u>	List all languages $L$ s.t. $ L  \leq 1$ and $m(L) \leq 1$	$\emptyset, \{\epsilon\}, \{0\}, \{1\}$
<u>stage 2:</u>	" " " " $ L  \leq 2$ and $m(L) \leq 2$	$\emptyset, \{\epsilon\}, \{0\}, \{1\}, \{00\}, \{01\}, \{10\}, \{11\}$
<u>stage n:</u>	" " " " $ L  \leq n$ and $m(L) \leq n$	$\{\epsilon, 0\}, \{\epsilon, 1\}, \{\epsilon, 00\}, \dots$

Define bijection  $f: \mathbb{N} \rightarrow \mathcal{F}$  by  $f(i) = i$ th distinct language in the list.

So what *isn't* countable?

# Cantor's Diagonalization Method



Georg Cantor 1845-1918

- Invented set theory
- Defined countability, uncountability, cardinal and ordinal numbers, ...

Some praise for his work:

“Scientific charlatan...renegade...corruptor of youth”  
–L. Kronecker

“Set theory is wrong...utter nonsense...laughable”  
–L. Wittgenstein

# Uncountability of the reals

**Theorem:** The real interval  $[0, 1]$  is uncountable.

**Proof:** We'll show that there is no surjection  $\mathbb{N} \rightarrow [0,1]$ .

Let  $f: \mathbb{N} \rightarrow [0,1]$  be an arbitrary function

$n$	$f(n)$
1	$0.\boxed{d_1^1} d_2^1 d_3^1 d_4^1 d_5^1 \dots$
2	$0.d_1^2 \boxed{d_2^2} d_3^2 d_4^2 d_5^2 \dots$
3	$0.d_1^3 d_2^3 \boxed{d_3^3} d_4^3 d_5^3 \dots$
4	$0.d_1^4 d_2^4 d_3^4 \boxed{d_4^4} d_5^4 \dots$
5	$0.d_1^5 d_2^5 d_3^5 d_4^5 \boxed{d_5^5} \dots$
$\vdots$	

Ex:

$$f(1) = 0.\boxed{3}141592\dots$$

$$f(2) = 0.2\boxed{7}1828\dots$$

$$f(3) = 0.86\boxed{7}530\dots$$

$$f(4) = 0.161\boxed{8}\dots$$

$$\vdots$$

$$b = 0.4889\dots$$

Construct  $b \in [0,1]$  which does not appear as any  $f(n)$

– Then  $f$  can't be a surjection!

$b = 0.b_1b_2b_3\dots$  where  $b_n \neq d_n^n$  (digit  $n$  of  $f(n)$ )

e.g. let  $b_n = d_n^n + 1 \pmod{10}$

# Diagonalization

This process of constructing a counterexample by “contradicting the diagonal” is called **diagonalization**

# Structure of a diagonalization proof

Say you want to show that a set  $T$  is uncountable

- 1) Let  $f: \mathbb{N} \rightarrow T$  be an arbitrary function
- 2) “Flip the diagonal” to construct an element  $b \in T$  such that  $f(n) \neq b$  for every  $n$

**Ex:** Let  $b = 0.b_1b_2b_3\dots$  where  $b_n \neq d_n^n$   
(where  $d_n^n$  is digit  $n$  of  $f(n)$ )

- 3) Conclude that  $f$  is not a surjection. Since  $f$  was arbitrary, there is no surjection from  $\mathbb{N} \rightarrow T$  so  $T$  is not countable

# A general theorem about set sizes

**Theorem:** Let  $X$  be any set. Then the power set  $P(X)$  does **not** have the same size as  $X$ .  
 $= \{s \mid s \subseteq X\}$

**Proof:** Let  $f: X \rightarrow P(X)$  be arbitrary. We'll show that  $f$  is not a surjection



What should we do to show  $f$  isn't a surjection?

- Show that for every  $S \in P(X)$ , there exists  $x \in X$  such that  $f(x) = S$
- b) Construct a set  $S \in P(X)$  (meaning,  $S \subseteq X$ ) that cannot be the output  $f(x)$  for any  $x \in X$
- Construct a set  $S \in P(X)$  and two distinct  $x, x' \in X$  such that  $f(x) = f(x') = S$

# Diagonalization argument

Let  $f: X \rightarrow P(X)$  be an arbitrary function

$x$					
$x_1$					
$x_2$					
$x_3$					
$x_4$					
$\vdots$					

# Diagonalization argument

$I_3 \quad x_3 \in f(x_2) ?$

Let  $f: X \rightarrow P(X)$  be an arbitrary function

$x$	$x_1 \in f(x) ?$	$x_2 \in f(x) ?$	$x_3 \in f(x) ?$	$x_4 \in f(x) ?$	...
$x_1$	<del>Y</del> N	N	Y	Y	
$x_2$	N	<del>N</del> Y	Y	Y	
$x_3$	Y	Y	<del>Y</del> N	N	
$x_4$	N	N	Y	<del>N</del> Y	
$\vdots$					$\ddots$

Define  $S$  by flipping the diagonal:

$S = \{x_2, x_4, \dots\}$

Put  $x_i \in S \iff x_i \notin f(x_i)$

# Example

Let  $X = \{1, 2, 3\}$ ,  $P(X) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}$

Ex.  $f(1) = \{1, 2\}$ ,  $f(2) = \emptyset$ ,  $f(3) = \{2\}$

$x$	$1 \in f(x)?$	$2 \in f(x)?$	$3 \in f(x)?$
1	<del>Y</del> N	Y	N
2	N	<del>N</del> Y	N
3	N	Y	<del>N</del> Y

Construct  $S = \{2, 3\}$

$S \neq f(1)$  because  $1 \in f(1) = \{1, 2\}$   
 $1 \notin S$

$S \neq f(2)$  because  $2 \notin f(2) = \emptyset$   
 $2 \in S$

$S \neq f(3)$  because  $3 \notin f(3) = \{2\}$   
 $3 \in S$

# A general theorem about set sizes

**Theorem:** Let  $X$  be any set. Then the power set  $P(X)$  does **not** have the same size as  $X$ .

**Proof:** Let  $f: X \rightarrow P(X)$  be an arbitrary function. *Complete solution ... don't need to describe diagonalization table*

Define

$$S = \{x \in X \mid x \notin f(x)\} \in P(X)$$

Suppose, for contradiction, that  $S = f(y)$  for some  $y \in X$

Then  $y \in S$  if and only if  $y \notin S$

*$y \in S$  means  $y \in f(y) \Rightarrow y \notin S$   
 $y \notin S$  means  $y \notin f(y) \Rightarrow y \in S$*

Hence  $S \in P(X)$  cannot be the output  $f(x)$  for any  $x \in X$ , so  $f$  is not a surjection.

# Undecidable Languages

# Undecidability / Unrecognizability

**Definition:** A language  $L$  is **undecidable** if there is no TM deciding  $L$

**Definition:** A language  $L$  is **unrecognizable** if there is no TM recognizing  $L$

# An existential proof

$$L \subseteq \{0, 1\}^*$$

**Theorem:** There exists an undecidable language over  $\{0, 1\}$

**Proof:**

Set of all encodings of TM deciders:  $X \subseteq \{0, 1\}^*$

Set of all languages over  $\{0, 1\}$ :

a)  $\{0, 1\}$

b)  $\{0, 1\}^*$

c)  $P(\{0, 1\}^*)$  : The set of all subsets of  $\{0, 1\}^*$

d)  $P(P(\{0, 1\}^*))$  : The set of all subsets of the set of all subsets of  $\{0, 1\}^*$



# An existential proof

**Theorem:** There exists an undecidable language over  $\{0, 1\}$

**Proof:**

Set of all encodings of TM deciders:  $X \subseteq \{0, 1\}^*$

Set of all languages over  $\{0, 1\}$ :  $P(\{0, 1\}^*)$

There are more languages than there are TM deciders!

⇒ There must be an undecidable language

# An existential proof

**Theorem:** There exists an **unrecognizable** language over  $\{0, 1\}$

**Proof:**

Set of all encodings of **TMs**:  $X \subseteq \{0, 1\}^*$

Set of all languages over  $\{0, 1\}$ :  $P(\{0, 1\}^*)$

There are more languages than there are TM **recognizers!**

⇒ There must be an **unrecognizable** language

# “Almost all” languages are undecidable



But how do we actually find one?

# An Explicit Undecidable Language

# Our power set size proof

**Theorem:** Let  $X$  be any set. Then the power set  $P(X)$  does **not** have the same size as  $X$ .

1) Let  $f: X \rightarrow P(X)$  be an arbitrary function

2) “Flip the diagonal” to construct a set  $S \in P(X)$  such that  $f(x) \neq S$  for every  $x \in X$

$$S = \{x \in X \mid x \notin f(x)\}$$

3) Conclude that  $f$  is not a surjection

# Specializing the proof

**Theorem:** Let  $X$  be the set of all TM deciders. Then there exists an undecidable language in  $P(\{0, 1\}^*)$

- 1) Consider the function  $L: X \rightarrow P(\{0, 1\}^*)$   
 $L(M) =$  language decided by  $M$
- 2) “Flip the diagonal” to construct a language  $UD \in P(\{0, 1\}^*)$  such that  $L(M) \neq UD$  for every  $M \in X$
- 3) Conclude that  $L$  is not a surjection

# An explicit undecidable language

TM $M$					
$M_1$					
$M_2$					
$M_3$					
$M_4$					
$\vdots$					

Why is it possible to enumerate all TMs like this?

- a) The set of all TMs is finite
- b) The set of all TMs is countably infinite
- c) The set of all TMs is uncountable



$M(\langle R \rangle)$  is what is set by running TM  $M$  on input  $\langle R \rangle$

# An explicit undecidable language

Y if  $M_2(\langle M_3 \rangle)$  accepts  
 N if  $M_2(\langle M_3 \rangle)$  does not accept

TM $M$	$M(\langle M_1 \rangle)$ ?	$M(\langle M_2 \rangle)$ ?	$M(\langle M_3 \rangle)$ ?	$M(\langle M_4 \rangle)$ ?		$D(\langle D \rangle)$ ?
$M_1$	<del>Y</del> N	N	Y	Y	...	
$M_2$	N	<del>N</del> Y	Y	Y		
$M_3$	Y	Y	<del>Y</del> N	N		
$M_4$	N	N	Y	<del>N</del> Y		
⋮					⋮	
$D$						<del>Y</del> N   <del>N</del> Y

$UD = \{ \langle M \rangle \mid M \text{ is a TM that does not accept on input } \langle M \rangle \}$

**Claim:**  $UD$  is undecidable. Suppose  $D$  decides  $UD$

Case 1:  $D$  accepts  $\langle D \rangle \Rightarrow \langle D \rangle \notin UD$  so  $D$  made a mistake on input  $\langle D \rangle$  ✗

Case 2:  $D$  rejects  $\langle D \rangle \Rightarrow \langle D \rangle \in UD$  so  $D$  made a mistake on input  $\langle D \rangle$  ✗

# An explicit undecidable language

**Theorem:**  $UD = \{\langle M \rangle \mid M \text{ is a TM that does not accept on input } \langle M \rangle\}$  is undecidable

**Proof:** Suppose for contradiction that TM  $D$  decides  $UD$