# BU CS 332 – Theory of Computation

https://forms.gle/HrsxQnUohJh4JrAo8

## Lecture 14:

- Uncountability
- Undecidability

Reading:

Sipser Ch 4.2, 5.1

Mark Bun

March 19, 2025

# 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

i.e. $\exists$ a bijection $f: \mathbb{N} \to S$

"countably infinite"

$f: \mathbb{N} \to (\mathbb{N} \times \mathbb{N})$

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

$f(1)=$, $f(2)=$, $f(3)=$, $f(4)=$, $f(5)=$, $f(6)=$

| | | | | | |
|---|---|---|---|---|---|
| $(1,1)$ | $(2,1)$ | $(3,1)$ | $(4,1)$ | $(5,1)$ | … |
| $(1,2)$ | $(2,2)$ | $(3,2)$ | $(4,2)$ | $(5,2)$ | … |
| $(1,3)$ | $(2,3)$ | $(3,3)$ | $(4,3)$ | $(5,3)$ | … |
| $(1,4)$ | $(2,4)$ | $(3,4)$ | $(4,4)$ | $(5,4)$ | … |
| $(1,5)$ | $(2,5)$ | $(3,5)$ | $(4,5)$ | $(5,5)$ | |
| | | | | | ⋱ |

$f$ constructed this way is a bijection

# 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 $x + y = n + 1$

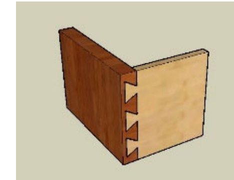
…

- 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

$\{\emptyset, \{\varepsilon\}, \{\varepsilon, 0\}, \{0\}, \{01, 011, 10\}, \dots \{0010100101\}, \dots\}$

Sketch 1

Enc(L) : string obtained by writing down all elts of L delimited by #

Ex: $L = \{01, 011, 10\}$ $\quad Enc(L) = 01\#011\#10$ $\quad 0101110$

$\{Enc(L) \mid L \text{ is finite}\} \subseteq \{0,1,\#\}^* \leftarrow$ countable

$\{01, 011, 10\}$

$\{010, 1110\}$

Sketch 2: Enumerate languages of $\mathcal{F}$ in stages

Define $m(L)$ = max length of a string in L

$|L|$ = # of strings in L

Stage 1: Enumerate all L s.t. $m(L) \leq 1$ and $|L| \leq 1$

$\emptyset, \{\varepsilon\}, \{0\}, \{1\}$

Stage 2: Enumerate all L s.t. $m(L) \leq 2$ and $|L| \leq 2$

$\emptyset, \{\varepsilon\}, \{0\}, \{1\}, \{00\}, \{01\}, \{10\}, \{11\}, \{\varepsilon, 0\}, \{\varepsilon, 1\}, \{\varepsilon, 00\}, \dots, \{\varepsilon, 11\}, \{0, 1\}, \{0, 00\}, \dots$

⋮ Stage n: Enumerate all L s.t. $m(L) \leq n$ and $|L| \leq n$

Every finite L appears in some stage k, where $k = \max\{m(L), |L|\}$

$\Rightarrow f(i) = i$'th distinct language enumerated is a bijection.

Why does this work?

Stage 1: Enumerate all language $L$ s.t $|L| \leq 1$

$\phi$, $\{\varepsilon\}$, $\{0\}$, $\{1\}$, $\{00\}$, $\{01\}$, $\{10\}$, $\{11\}$, $\{000\}$, $\{001\}$...

$\vdots$

$\{1010101101101 \ldots 11\}$

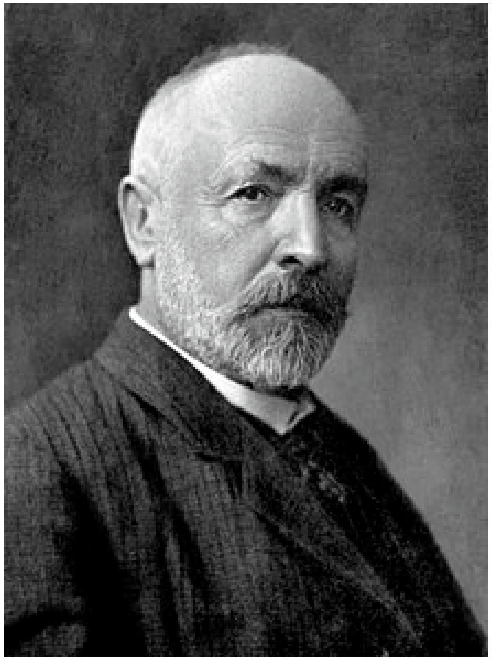
Stage n: " " s.t. $|L| \leq n$

$f(i) =$ i'th language enumerated

never gets to stage 2!

# 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 $\mathbb{N} \to [0,1]$.

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

first digit of decimal expansion of f(1)

2nd digit "

| $n$ | $f(n)$ |
|---|---|
| 1 | $0\,.\,d_1^1\ d_2^1\ d_3^1\ d_4^1\ d_5^1$ ... |
| 2 | $0\,.\,d_1^2\ d_2^2\ d_3^2\ d_4^2\ d_5^2$ ... |
| 3 | $0\,.\,d_1^3\ d_2^3\ d_3^3\ d_4^3\ d_5^3$ ... |
| 4 | $0\,.\,d_1^4\ d_2^4\ d_3^4\ d_4^4\ d_5^4$ ... |
| 5 | $0\,.\,d_1^5\ d_2^5\ d_3^5\ d_4^5\ d_5^5$ ... |

Ex:

$b \neq f(1) = 0.271828\ldots$

$b \neq f(2) = 0.314159\ldots$

$b \neq f(3) = 0.867530\ldots$

$\vdots$

$b = 0.328\ldots$

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

– Then $f$ can't be a surjection! $\Rightarrow$ No surjection $f: \mathbb{N} \to [0,1]$ exists, so $[0,1]$ not countable.

$b = 0.b_1 b_2 b_3\ldots$ where $b_n \neq d_n^n$ (digit $n$ of $f(n)$)

$b_n = d_n^n + 1 \pmod{10}$

# 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

WTS no surjection $f: \mathbb{N} \to T$

1) Let $f: \mathbb{N} \to T$ 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 $b = 0.b_1 b_2 b_3 \ldots$ 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 $\mathbb{N} \to$ ~~[illegible]~~ $T$ 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 $f: X \to P(X)$ be arbitrary. We'll show that $f$ is not onto (not a surjection)

<u>What should we do to show $f$ isn't onto?</u>

a) 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 $x \in X$

c) Construct a set $S \in P(X)$ and two distinct $x, x' \in X$ such that $f(x) = f(x') = S$

# Diagonalization argument

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

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

# Diagonalization argument

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

| $x$ | $x_1 \in f(x)$? | $x_2 \in f(x)$? | $x_3 \in f(x)$? | $x_4 \in f(x)$? | … |
|---|---|---|---|---|---|
| $x_1$ | Y | N | Y | Y | |
| $x_2$ | N | N | Y | Y | |
| $x_3$ | Y | Y | Y | N | |
| $x_4$ | N | N | Y | N | |
| ⋮ | | | | | ⋱ |

Define $S$ by flipping the diagonal:

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

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

# Example

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

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 | ~~Y~~ N | Y | N |
| 2 | N | ~~N~~ Y | N |
| 3 | N | Y | ~~N~~ Y |

$f(1) \neq S$ b/c $1 \in f(1)$, $1 \notin S$

$f(2) \neq S$ b/c $2 \notin f(2)$, $2 \in S$

$f(3) \neq S$ b/c $3 \notin f(3)$, $3 \in S$

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

# 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 $X$.

WTS no surjection $f: X \to P(X)$ exists

**Proof:** Let $f: X \to P(X)$ be an arbitrary function.

Define

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

If $S = f(y)$ for some $y \in X$,

then $y \in S$ if and only if $y \notin S$ ⨳

Claim:
- $S \in P(X)$
- $S \neq f(y)$ for any $y \in X$

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

∃ a language $L \subseteq \{0,1\}^*$ that is not decidable

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

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!

$\Rightarrow$ 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!

$\Rightarrow$ 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 $x \in X$

3) Conclude that $f$ is not onto

# 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 $L: X \to P(\{0,1\}^*)$

   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 onto

# An explicit undecidable language

| TM $M$ | | | | | |
|---|---|---|---|---|---|
| $M_1$ | | | | | |
| $M_2$ | | | | | |
| $M_3$ | | | | | |
| $M_4$ | | | | | |
| ⋮ | | | | | |

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

# An explicit undecidable language

| TM $M$ | $M(\langle M_1 \rangle)$? | $M(\langle M_2 \rangle)$? | $M(\langle M_3 \rangle)$? | $M(\langle M_4 \rangle)$? | | $D(\langle D \rangle)$? |
|---|---|---|---|---|---|---|
| $M_1$ | Y | N | Y | Y | … | |
| $M_2$ | N | N | Y | Y | | |
| $M_3$ | Y | Y | Y | N | | |
| $M_4$ | N | N | Y | N | | |
| $\vdots$ | | | | | $\ddots$ | |
| $D$ | | | | | | |

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

Claim: $UD$ is undecidable