Last time
- Decidable languages
- Designing deciders

Today
- Undecidable languages
- Diagonalization
Classes of languages

CFL
regular

recognizable
decidable
### Problems in language theory

<table>
<thead>
<tr>
<th>$A_{DFA}$</th>
<th>$A_{CFG}$</th>
<th>$A_{TM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>decidable</td>
<td>decidable</td>
<td>?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$E_{DFA}$</th>
<th>$E_{CFG}$</th>
<th>$E_{TM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>decidable</td>
<td>decidable</td>
<td>?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$EQ_{DFA}$</th>
<th>$EQ_{CFG}$</th>
<th>$EQ_{TM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>decidable</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>
We will prove that there are some undecidable languages:
• i.e., problems a computer cannot solve no matter how long it computes

The proof idea is “simple:”

There are more languages than there are Turing Machines.
A language $L$ is **undecidable** if there is no TM that decides $L$.

If $L$ is undecidable, then every TM must either:

1. Accept (infinitely many) strings $s \notin L$.
2. Reject (infinitely many) strings $s \in L$.
3. Loop forever on (infinitely many) strings.
Let \( \mathbb{N} = \{1,2,\ldots\} \) be the natural numbers.
Let \( \mathbb{E} = \{2,4,6,\ldots\} \) be the even natural numbers.
Let \( \mathbb{Z} = \{\ldots,-2,-1,0,1,2,\ldots\} \) be the integers.

Which one is largest?

A. \( \mathbb{N} \)
B. \( \mathbb{E} \)
C. \( \mathbb{Z} \)
D. the same size.
Are there more blue or yellow dots?
A function $f : A \to B$ is

- **1-to-1 (or injective)** if
  
  
  $$f(a) \neq f(b) \text{ for } a \neq b.$$  

- **onto (or surjective)** if for all $b \in B$, some $a \in A$ maps to $b$: $f(a) = b$.

- **correspondence (or bijective)** if
  
  it is 1-to-1 and onto, i.e.,
  
  each $a \in A$ maps to a unique $b \in B$, and each $b \in B$ has a unique $a \in A$ mapping to it.
How to compare sizes of infinite sets?

- Two sets are **the same size** if there is a bijection between them.
- A set is **countable** if it is
  - finite or
  - it has the same size as \( \mathbb{N} \), the set of natural numbers
Examples of countable sets

\[
\emptyset, \{0\}, \{0,1\}, \{0,1, \ldots, 255\}
\]

\[
E = \{2,4,6,\ldots\}
\]

\[
O = \{1,3,5,7,\ldots\}
\]

SQUARES = \{1,4,9,16,25\ldots\}

POWERS = \{1,2,4,8,16,32\ldots\}

|POWERS| = |SQUARES| = |E| = |O| = |\mathbb{N}|
There is a bijection between $\mathbb{N}$ and $\mathbb{N} \times \mathbb{N}$. 

\[
\begin{align*}
(0,0) & \quad (0,1) & \quad (0,2) & \quad (0,3) & \quad (0,4) & \ldots \\
(1,0) & \quad (1,1) & \quad (1,2) & \quad (1,3) & \quad (1,4) & \ldots \\
(2,0) & \quad (2,1) & \quad (2,2) & \quad (2,3) & \quad (2,4) & \ldots \\
(3,0) & \quad (3,1) & \quad (3,2) & \quad (3,3) & \quad (3,4) & \ldots \\
(4,0) & \quad (4,1) & \quad (4,2) & \quad (4,3) & \quad (4,4) & \ldots 
\end{align*}
\]
\{0,1\}^* \text{ is countable}

\{ \langle M \rangle \mid M \text{ is a TM} \} \text{ is countable}

\mathbb{Q}^+ = \{ \frac{p}{q} \mid p,q \in \mathbb{Z}^+ \} \text{ is countable!}

Is any set \textit{uncountable}?
Creator of Set Theory

Georg Cantor
1845--1918
Theorem. There is no bijection from the positive integers to the real interval (0,1).

Proof: Suppose f is such a function:

<table>
<thead>
<tr>
<th>n</th>
<th>f(n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.28347279…</td>
</tr>
<tr>
<td>2</td>
<td>0.88388384…</td>
</tr>
<tr>
<td>3</td>
<td>0.77635284…</td>
</tr>
<tr>
<td>4</td>
<td>0.11111111…</td>
</tr>
<tr>
<td>5</td>
<td>0.12345678…</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
</tr>
</tbody>
</table>

Construct $b \in (0,1)$ that does not appear in the table.

$b=0.d_1d_2d_3 \ldots$, where $d_i \neq \text{digit } i \text{ of } f(i)$. 

Sofya Raskhodnikova; based on slides by Nick Hopper
The process of constructing a counterexample by “contradicting the diagonal” is called **DIAGONALIZATION**
What if we try this argument on \( \mathbb{Q} \) instead of \( \mathbb{R} \)?

**Proof:** Suppose \( f \) is such a function:

<table>
<thead>
<tr>
<th>( n )</th>
<th>( f(n) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.28347279…</td>
</tr>
<tr>
<td>2</td>
<td>0.8388384…</td>
</tr>
<tr>
<td>3</td>
<td>0.77635284…</td>
</tr>
<tr>
<td>4</td>
<td>0.11111111…</td>
</tr>
<tr>
<td>5</td>
<td>0.12345678…</td>
</tr>
<tr>
<td>( \vdots )</td>
<td>( \vdots )</td>
</tr>
</tbody>
</table>

Construct \( b \in (0, 1) \) that does not appear in the table.

\[
b = 0.d_1d_2d_3\ldots, \text{ where } d_i \neq \text{ digit } i \text{ of } f(i).
\]
Exercise

What if we try Cantor’s diagonalization argument on \( \mathbb{Q} \) instead of \( \mathbb{R} \)?

A. It works.
B. It fails because there are some rational numbers that cannot be represented in decimal point notation.
C. It fails because the \( i \)-th number might have no digit in the \( i \)-th position after the decimal point.
D. It fails because the constructed number is not rational.
E. None of the above.
Let $L$ be any set and $P(L)$ be the power set of $L$.

**Theorem:** There is no bijection from $L$ to $P(L)$

**Proof:** Assume, for a contradiction, that there is bijection $f : L \rightarrow P(L)$

We construct a set $S$ that cannot be the output, $f(y)$, for any $y \in L$. 
Use diagonalization

Define set $S$ by flipping the diagonal:

$$(y_i \in S) = Y \text{ iff } (y_i \in f(y_i)) = N$$
Let \( L = \{0,1,2\} \). Then \( P(L) = \{\emptyset, \{0\}, \{1\}, \{2\}, \{0,1\}, \{0,2\}, \{1,2\}, \{0,1,2\}\} \)

Let \( f(0) = \{1\}, f(1) = \emptyset, f(2) = \{0,2\} \). Then:

\[
\begin{array}{c|c|c|c}
 x & 0 \in f(x) & 1 \in f(x) & 2 \in f(x) \\
\hline
0 & N & Y & N \\
1 & N & N & N \\
2 & Y & N & Y \\
\end{array}
\]

\( S = \{0,1\} \)
Let $L$ be any set and $\mathcal{P}(L)$ be the power set of $L$

**Theorem:** There is no bijection from $L$ to $\mathcal{P}(L)$

**Proof:** Assume, for a contradiction, that there is bijection $f : L \rightarrow \mathcal{P}(L)$

We construct a set $S$ that cannot be the output, $f(y)$, for any $y \in L$.

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

If $S = f(y)$ then $y \in S$ if and only if $y \notin S$
For all sets $L$, $P(L)$ has more elements than $L$. 
Not all languages over \( \{0,1\} \) are decidable

**TM Deciders**

Strings of 0s and 1s

**Languages over \( \{0,1\} \)**

Sets of strings of 0s and 1s

\[
L \\
\text{P}(L)
\]
Not all languages over \{0,1\} are recognizable.

Turing Machines

Strings of 0s and 1s

\[ L \]

Languages over \{0,1\}

Sets of strings of 0s and 1s

\[ P(L) \]
A specific undecidable language

$A_{TM} = \{ \langle M, w \rangle \mid M \text{ is a TM, } w \text{ is a string, and } M \text{ accepts } w \}$
Theorem. $A_{TM}$ is undecidable.

Proof: For contradiction, suppose a TM $H$ decides $A_{TM}$.

$$H(\langle M, w \rangle) = \begin{cases} 
\text{accept} & \text{if } M \text{ accepts } w \\
\text{reject} & \text{if } M \text{ doesn't accept } w 
\end{cases}$$

Idea: Use $H$ to check what TM $M$ does on its own description (and do the opposite).

**TM D** = ```On input $\langle M \rangle$, where $M$ is a TM:
1. Run $H$ on input $\langle M, \langle M \rangle \rangle$.
2. Accept if it rejects. O.w. reject.”```

D is a decider. What does it do on $\langle D \rangle$?
Is it diagonalization again? Does M accept $\langle M \rangle$?

<table>
<thead>
<tr>
<th>TMs</th>
<th>$\langle M_1 \rangle$</th>
<th>$\langle M_2 \rangle$</th>
<th>$\langle M_3 \rangle$</th>
<th>$\langle M_4 \rangle$</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1$</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>$M_2$</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>$M_3$</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>$M_4$</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>...</td>
</tr>
</tbody>
</table>

D accepts $\langle M_i \rangle$ iff entry $(i, i)$ is $\mathbf{N}$.
Is it diagonalization again?

Does $M$ accept $\langle M \rangle$?

<table>
<thead>
<tr>
<th>TMs</th>
<th>$\langle M_1 \rangle$</th>
<th>$\langle M_2 \rangle$</th>
<th>$\langle M_3 \rangle$</th>
<th>$\langle M_4 \rangle$</th>
<th>...</th>
<th>$\langle D \rangle$</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1$</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M_2$</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M_3$</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M_4$</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>?</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A movie about undecidability of the Halting Problem

https://www.youtube.com/watch?v=92WHN-pAFCs