Lecture 26: Unrecognizable languages

Last time:

- The language $L = \{x \in \{0, 1\}^* \mid x \text{ has an odd number of 1s} \}$ is DFA-recognizable
- The language $L = \{x \in \{0, 1\}^* \mid x \text{ does not contain 010} \}$ is DFA-recognizable
- If $L_1$ and $L_2$ are DFA-recognizable then so is $L_1 \cap L_2$

Have the following definitions handy for today:

- $L(M), \delta$

Today: is every language recognizable?

**Defn:** A language is a set of strings.

- $L: \mathbb{N} \to \{0, 1\}^*$ uncountably

So there are more languages than automata, so some languages have no automata.

$L: \text{DFA} \to \text{language}$ is not surjective.
An unrecognizable language

We saw that there exists an unrecognizable language. Here's an example:

Let \( L := \{0^n 1^n \mid n \in \mathbb{N} \} = \{0, 01, 0011, 000111, \ldots \} \)

**Claim:** \( L \) is not DFA-recognizable

**Proof:** Assume (for contradiction) that \( L \) is DFA-recognizable, so that \( \exists M \) with \( L(M) = L \). Let \( m \) be a state of \( M \).

Consider \( m \) processing \( 0^m \). Know \( m \notin L \), so \( \delta(q_0, 0) \notin A \).

\[ \text{Only } m \text{ distinct states to choose from:} \]

\[ \text{consider } i, j \text{ for some } i \neq j \]

\[ \text{let } w \text{ be first } i \text{ chars of } x, v \text{ next } j-i \]

\[ \text{chars, } w \text{ rest of } x. \]

Find a string \( y \) that is accepted by \( M \) but shouldn't be.

\[ \text{wwwv} \text{ is accepted, but has more } 0\text{s than } 1\text{s.} \]

\[ \text{wwww} \text{ is accepted, but has too many } 0\text{s.} \]

\[ \text{uw} \text{ is also accepted, but has too few } 0\text{s.} \]

**Contradiction.**
Pumping lemma

Claim (pumping lemma):

\[ \forall y \quad \exists n \in \mathbb{N}, \text{ such that} \]
\[ \forall x \in L \text{ with } \text{len}(x) \geq n, \]
\[ \exists u, v, w \in \Sigma^* \text{ such that} \]
\[ 1. \ x = uvw \quad 3. \ v \neq \varepsilon \]
\[ 2. \ \text{len}(uv) \leq n \quad 4. \ \forall k \in \mathbb{N}, \ uv^k w \in L \]

Example: Let \( L := \{0^n1^n \mid n \in \mathbb{N}\} \). Then \( L \) is unrecognizable.

Proof of example, using pumping lemma:

Assume \( L \) is DFA rec. Then \( \exists n \) as in PL.

Let \( x = 0^n1^n \). Then \( x \in L \) and \( \text{len}(x) \geq n \).

So \( \exists u,v,w \) as in PL. Since \( \text{len}(uv) \leq n \),
\( v \) must only contain 0's. Since \( v \neq \varepsilon \),
\( v \) has at least one 0. So \( uvw \) has
fewer 0's than 1's, so \( uvw \notin L \). But
PL says \( uv^k w \in L \), a contradiction.
Proof of pumping lemma:

Claim (pumping lemma):

1. For all DFA-recognizable language \( L \).
2. There exists a number \( n \in \mathbb{N} \) such that
3. For all \( x \in L \) with \( \text{len}(x) \geq n \).
4. There exists \( u, v, w \in \Sigma^* \) such that
   
   \[
   x = uvw, \quad v \neq \varepsilon, \quad \text{len}(vw) \leq n, \quad \forall k \in \mathbb{N}, \ uv^kw \in L
   \]

Proof of pumping lemma:

Choose an arb. DFA-recognizable language \( L \). Then \( \exists \) a DFA \( M \)
with \( L = L(M) \). Let \( n = n(M) \) be the number of states of \( M \).
Choose an arb. \( x \in L \) with \( \text{len}(x) \geq n \).
While processing the first \( n \) chars of \( x \),
we'll pass through all states of \( M \).

\[
\begin{array}{c}
q_0 \\
\downarrow \\
q_1 \\
\downarrow \\
q_2 \\
\downarrow \\
\vdots \\
\downarrow \\
q_i \\
\downarrow \\
q_j \\
\downarrow \\
q_k \\
\downarrow \\
q_{i+1} \\
\downarrow \\
\vdots \\
q_s \\
\end{array}
\]

So \( q_i \) and \( q_j \) are same.

\[
\begin{array}{c}
q_i \\
\downarrow \\
q_j \\
\downarrow \\
\vdots \\
\downarrow \\
q_s \\
\end{array}
\]

Let \( v \) be the string that gets to \( q_i \),
\( w \) be the rest.

Then \( x = uvw \) by construction.

\[
\text{len}(uv) \leq n \quad \text{by constr. (found while processing first } n \text{ chars)}
\]

\[
v \neq \varepsilon \quad \text{otherwise } \Rightarrow \text{Choose arb. } k \in \mathbb{N}, \text{ then}
\]

\[
M \text{ on } uv^kw \text{ will transition to } q_i, \text{ traverse } k \text{ times,}
\]

\[
\text{transition to an accept state.}
\]