Lecture 7:

• Distinguishing sets
• Non-regular languages

Reading:
“Myhill-Nerode” note

Mark Bun
September 23, 2021
Motivating Questions

• How can we tell if we’ve found the smallest DFA recognizing a language?
  
  Last time: Introduced distinguishing set method

• Are all languages regular? How can we prove that a language is not regular?
A General Technique

Definition: Strings $x$ and $y$ are distinguishable by $L$ if there exists a “distinguishing extension” $z$ such that exactly one of $xz$ or $yz$ is in $L$.

Ex. $x = \varepsilon$, $y = 0$

Definition: A set of strings $S$ is pairwise distinguishable by $L$ if every pair of distinct strings $x, y \in S$ is distinguishable by $L$.

Ex. $S = \{\varepsilon, 0, 01\}$

Theorem: If $S$ is pairwise distinguishable by $L$, then every DFA recognizing $L$ needs at least $|S|$ states
A General Technique

**Theorem:** If $S$ is pairwise distinguishable by $L$, then every DFA recognizing $L$ needs at least $|S|$ states

**Proof:** Let $M$ be a DFA with $< |S|$ states. By the pigeonhole principle, there are $x, y \in S$ such that $M$ ends up in same state on $x$ and $y$
Another Example

\[ B = \{ w \in \{0, 1\}^* \mid |w| = 2 \} \]

**Theorem:** If \( S \) is pairwise distinguishable by \( L \), then every DFA recognizing \( L \) needs at least \( |S| \) states

\[ S = \{ \} \]
Distinguishing Extension

Which of the following is a distinguishing extension for \( x = 0 \) and \( y = 00 \) for language \( B = \{ w \in \{0, 1\}^* \mid |w| = 2 \} \) ?

a) \( z = \varepsilon \)
b) \( z = 0 \)
c) \( z = 1 \)
d) \( z = 00 \)
Historical Note

Converse to the distinguishing set method:

If $L$ has no distinguishing set of size $> k$, then $L$ is recognized by a DFA with $k$ states

Myhill-Nerode Theorem (1958): $L$ is recognized by a DFA with $\leq k$ states iff $L$ does not have a distinguishing set of size $> k$
Non-Regularity

**Theorem:** If $S$ is pairwise distinguishable by $L$, then every DFA recognizing $L$ needs at least $|S|$ states.

**Corollary:** If $S$ is an **infinite** set that is pairwise distinguishable by $L$, then no DFA recognizes $L$. 
h/t Islam
The Classic Example

**Theorem:** $A = \{0^n1^n \mid n \geq 0\}$ is not regular

**Proof:** We construct an infinite pairwise distinguishable set
Palindromes

Theorem: $L = \{w \in \{0,1\}^* \mid w = w^R\}$ is not regular

Proof: We construct an infinite pairwise distinguishable set
Now you try!

Use the distinguishing set method to show that the following languages are not regular

\[ L_1 = \{0^i1^j \mid i > j \geq 0\} \]
Now you try!

Use the distinguishing set method to show that the following languages are not regular

\[ L_2 = \{ww \mid w \in \{0,1\}^*\} \]
Now you try!

Use the distinguishing set method to show that the following languages are not regular

\[ L_3 = \{1^{n^2} \mid n \geq 0\} \]
Reusing a Proof

Finding a distinguishing set can take some work...
Let’s try to reuse that work!

How might we show that

\[ \text{BALANCED} = \{w \mid w \text{ has an equal # of 0s and 1s}\} \]

is not regular?

\[ \{0^n1^n \mid n \geq 0\} = \text{BALANCED} \cap \{w \mid \text{all 0s in } w \text{ appear before all 1s}\} \]
Using Closure Properties

If $A$ is not regular, we can show a related language $B$ is not regular

By contradiction: If $B$ is regular, then $B \cap C (= A)$ is regular.

But $A$ is not regular so neither is $B$!
Example

Prove $B = \{0^i1^j | i \neq j\}$ is not regular using

- nonregular language
  \[ A = \{0^n1^n | n \geq 0\} \text{ and} \]

- regular language
  \[ C = \{w \mid \text{all } 0\text{s in } w \text{ appear before all } 1\text{s}\} \]

Which of the following expresses $A$ in terms of $B$ and $C$?

a) $A = B \cap C$

b) $A = \overline{B} \cap C$

c) $A = B \cup C$

d) $A = \overline{B} \cup C$
Proof that $B$ is nonregular

Assume for the sake of contradiction that $B$ is regular
We know: \[ A = \overline{B} \cap C \]
Let $B = \{0^i1^j \mid i \neq j\}$ and write $B = A \cup C$ where

- nonregular language
  
  $A = \{0^i1^j \mid i > j \geq 0\}$ and

- nonregular language
  
  $C = \{0^i1^j \mid j > i \geq 0\}$ and

Does this let us conclude $B$ is nonregular?