Lecture 5

Last time:
- Closure properties.
- Equivalence of NFAs, DFAs and regular expressions

Today:
- Proving that a language is not regular: pumping lemma

Sofya Raskhodnikova
Conversion procedures

DFA ↔ NFA

Regular Language ↔ Regular Expression

definition
Design an NFA for the language:

\{0^n1^n \mid 0 < n \leq 2\}

\{0^n1^n \mid 0 < n \leq k\}

\{0^n1^n \mid n > 0\}?

(For R a regexp, \(R^2\) means \(RR\), and \(R^n\) means \(RR\ldots R\))
SOME LANGUAGES ARE NOT REGULAR!

\[ B = \{0^n1^n \mid n \geq 0\} \text{ is NOT regular!} \]
Proof (by contradiction)

Let M be a k-state DFA that recognizes B.

Consider the path M takes on $0^k1^k$:

$q_0q_1q_2... q_iz_{i+1}q_jq_k... q_{2k} \in F$

$$0000...00..0..011111...11$$

There must be $i < j \leq k$ such that $q_i = q_j$

M accepts $0^{k-(j-i)}1^k \notin B$!

So M does not recognize the language B.
REGULAR OR NOT?

\[ C = \{ w \mid w \text{ has equal number of 1s and 0s}\} \]

NOT REGULAR

\[ D = \{ w \mid w \text{ has equal number of occurrences of 01 and 10}\} \]

\[ (0\Sigma^*0) \cup (1\Sigma^*1) \cup 1 \cup 0 \cup \varepsilon \]
THE PUMPING LEMMA

Let $L$ be a regular language with $|L| = \infty$

Then there exists a length $p$ such that

if $w \in L$ and $|w| \geq p$ then

$w$ can be split into three parts $w=xyz$ where:

1. $|y| > 0$
2. $|xy| \leq p$
3. $xy^iz \in L$ for all $i \geq 0$
THE PUMPING LEMMA

Example:

Let \( L = 0^*1^* \); \( p = 1 \)

\( w = 011 \)

\( x = \varepsilon \)

\( y = 0 \)

\( z = 11 \)

if \( w \in L \) and \( |w| \geq p \)

then \( w = xyz \), where:

1. \( |y| > 0 \)
   Let \( L = (0 \cup 1)^2^* \); \( p = 2 \)

2. \( |xy| \leq p \)
   \( w = 12 \)

3. \( xy^iz \in L \) for all \( i \geq 0 \)

\( x = 1 \)

\( y = 2 \)

\( z = \varepsilon \)
Let $M$ be a DFA that recognizes $L$. Let $p$ be the number of states in $M$. Assume $w \in L$ is such that $|w| \geq p$.

We show $w = xyz$

1. $|y| > 0$
2. $|xy| \leq p$
3. $xy^i z \in L$ for all $i \geq 0$

There must be $j > i$ such that $q_i = q_j$. 

Proof of the pumping lemma
Use the pumping lemma to prove that $B = \{0^n1^n \mid n \geq 0\}$ is not regular.

**Hint:** Assume $B$ is regular. Let $p$ be the pumping length. Try pumping $w = 0^p1^p$.

If $B$ is regular, $w$ can be split into $w = xyz$, where

1. $|y| > 0$
2. $|xy| \leq p$
3. $xy^iz \in B$ for all $i \geq 0$

$y$ is all 0s: $xyyz$ has more 0s than 1s

**Contradiction!**
Proof by contradiction: assume $L$ is regular.

Then there is a pumping length $p$.

Find a string $w \in L$ with $|w| \geq p$.

Show that no matter how you choose $xyz$, $w$ cannot be pumped!

Conclude that $L$ is not regular.
Pumping lemma as a game

1. **YOU** pick the language $L$ to be proved nonregular.
2. **ADVERSARY** picks $p$, but doesn't reveal to **YOU** what $p$ is; **YOU** must devise a play for all possible $p$'s.
3. **YOU** pick $w$, which should depend on $p$ and which must be of length at least $p$.
4. **ADVERSARY** divides $w$ into $x, y, z$, obeying conditions stipulated in the pumping lemma: $|y| > 0$ and $|xy| \leq p$. Again, **ADVERSARY** does not tell **YOU** what $x, y, z$ are, although they must obey the constraints.
5. **YOU** win by picking $i$, which may be a function of $p, x, y, z$, such that $xy^iz$ is not in $L$. 

Sofya Raskhodnikova; based on slides by Nick Hopper
PALINDROMES = \{ \text{ww}^R | w \in \{0,1\}^* \} is not regular.

Proof: Assume … pumping length \( p \)

Find a \( w \in \text{PALINDROMES} \) longer than \( p \)

\[
0^p1^p1^p0^p
\]

Show that \( w \) cannot be pumped:

\[
w = 00\ldots0011\ldots1100\ldots00
\]

\( y \) must be in this part

\[
xyyz = 00\ldots00011\ldots1100\ldots00
\]

\[
> p \quad 2p \quad p
\]
PALINDROMES = \{ w w^R \mid w \in \{0,1\}^* \} is not regular.

Proof: Assume … pumping length p
Find a \( w \in \text{PALINDROMES} \) longer than p
\[ 0^p 1^p 1^p 0^p \]

Show that w cannot be pumped:

If \( w = xyz \) with \( |xy| \leq p \) then
\( y = 0^J \) for some \( J > 0 \).
Then \( xyyz = 0^{p+J} 1^2 0^p \not\in \text{PALINDROMES} \)

Contradiction!
Exercise

Prove $C = \{ 0^i1^j \mid i > j \geq 0 \}$ is not regular.

Proof: Assume ... pumping length $p$

What string can you choose in your next move?

A. 00011
B. $0^p1^p$
C. $0^{p/2}1^{p/2-1}$
D. $0^{p+1}1^p$
E. $0^{p+2}1^p$
F. More than one choice above works.
Prove \( C = \{ 0^i1^j \mid i > j \geq 0 \} \) is not regular.

Proof: Assume \( \ldots \) pumping length \( p \)

Find a \( w \in C \) of length at least \( p \)

\[ w = 0^p1^p \]

Show that \( w \) cannot be pumped:

\[ w = 00\ldots0011\ldots11 \]

\( y \) must be in this part

\[ xyyz = 00\ldots00011\ldots11 \quad xz = 0\ldots0011\ldots11 \]

> \( p+1 \)

\( \leq p \)
Prove $C = \{ 0^i1^j \mid i > j \geq 0 \}$ is not regular.
Proof: Assume ... pumping length $p$
Find a $w \in C$ of length at least $p$
$0^{p+1}1^p$

If $w = xyz$ with $|xy| \geq p$ then $y = 0^j$ for some $J \geq 1$.
Then $xy^0z = xz = 0^{p+1-J}1^p \notin C$

Contradiction!
Let $\text{BALANCED} = \{ w \mid w \text{ has an equal # of } 1\text{s and } 0\text{s} \}$

Assume ... there is a $p$

Find a $w \in \text{BALANCED}$ of length at least $p$

$(01)^p$ \text{\sout{}} $0^p1^p$

Show that $w$ cannot be pumped:

If $w = xyz$ with $|xy| \leq p$ then $y = 0^j$ for some $J > 0$.

Then $xyyzz = 0^{p+j}1^p \notin \text{BALANCED}$
Pumping a language can be lots of work…
Let’s try to reuse that work!

\[ \{0^n1^n \mid n \geq 0\} = \text{BALANCED} \cap 0^*1^* \]

If BALANCED is regular then so is \( \{0^n1^n \mid n \geq 0\} \)
Prove: A is not regular

any of \{\circ, \cup, \cap\} or, for one language, \{\neg, R, *\}

If A is regular, then A \cap C (= B) is regular.

But B is not regular so neither is A.
Prove $A = \{0^i1^j \mid i \neq j\}$ is not regular using $B = \{0^n1^n \mid n \geq 0\}$

$\neg A = B \cup \{\text{strings that mix 0s and 1s}\}$

$\neg A \cap 0^*1^* = B$