COMS W3261 — Lecture 5 — Context-Free Languages (Using the Pumping Lemma)

Teaser: Consider the language

$$A = \underbrace{\{0^n 1^n \mid n \geq 0\}}_{\text{nonregular.}} \cup \underbrace{\{0^n 1^m \mid n \geq 0, m \geq 0\}}_{0^* 1^*}.$$

Is our language A regular?

$$A = \{0^n 1^m \mid n \geq 0, m \geq 0\} = 0^* 1^*;$$

A is regular.

Announcements: HW #3 due Monday, 7/19/21 @ 11:59PM EST.

One example (proving $\{0^n 1^n \mid n \geq 0\}$ irregular using the pumping lemma) in lecture 4 on YT, in lecture 5 in class.

Readings: Sipser 2.1. (1.4 for Pumping Lemma review)

Today:

1. Review
2. More examples of proving languages nonregular (PL, closure properties)
3. Context-Free Grammars, parse trees, derivations, etc. Context-Free Languages.

1. Review

- We now know that a language is regular

  $\longleftrightarrow$ some DFA recognizes it (by definition)

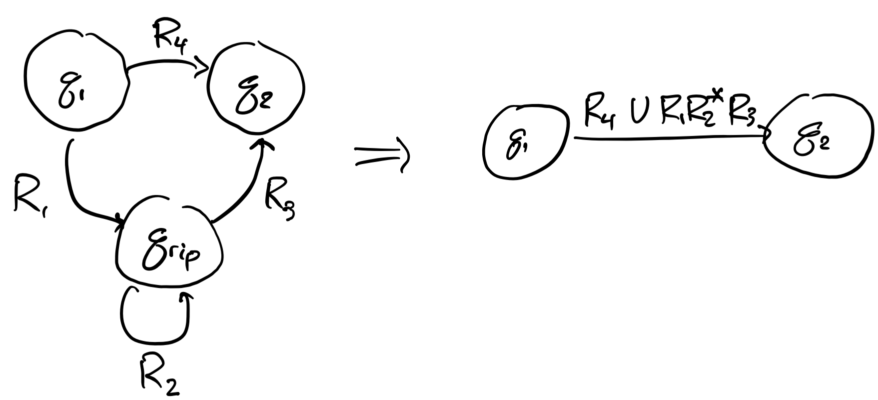
$\longleftrightarrow$ some NFA recognizes it (NFA $\leftrightarrow$ DFA)

$\longleftrightarrow$ some regular expression evaluates to the language.

(reg. ex. $\longrightarrow$ NFA, using inductive defn of reg. ex and closure properties)

(regular language $\rightarrow$ DFA $\rightarrow$ GNFA $\rightarrow$ reg. ex.)

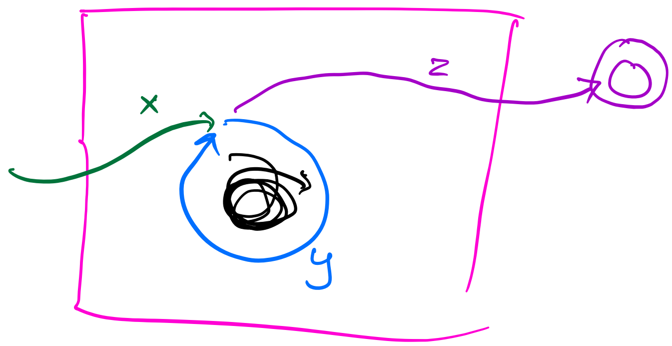
- GNFAs: like NFAs, but transition on regular expressions.
  - Exactly one start state q_{start}, one accept state q_{accept}.
  - Exactly one transition between every pair of states in $(Q - \{q_{accept}\}) \times (Q - \{q_{start}\})$
  - We showed DFA $\rightarrow$ GNFA by adding states and adding $\emptyset$-edges
  - Lec showed how to boil down GNFA $\rightarrow$ reg. ex by iteratively ripping out states.



Pumping Lemma. If A is a regular language, there is some "pumping length" p such that every string $s \in A$ with length $|s| \geq p$ can be divided into three substrings x, y, and z satisfying

- $xy^iz \in A$, $i \geq 0$ // intuition: loopiness.
- $|y| > 0$

- $|xy| \le p$.



<u>Example.</u> Show that $F = \{ww \mid w \in \{0,1\}^*\}$ is nonregular.

<u>Proof.</u> (1) Assume for contradiction that F is regular.

(2) Thus F satisfies the pumping lemma and there exists a pumping length p, such that any string $s \in F$ w/ $|s| \ge p$ can be divided into substrings $s = xyz$ such that $|y| > 0$, $|xy| \le p$, and $xy^iz \in F$ for all $i \ge 0$.

(3) Consider the string

$$s = \underline{0^p}\,1\,\underline{0^p}\,1. \qquad (s \ne 0^p0^p = 0^{2p}$$

By assumption, $|xy| \le p$. thus $|z| \ge 2p+2-p = p+2$. As a result, z contains the substring 10^p1. Thus y consists of all 0's, and $xyyz = 0^{p+|y|}10^p1$.

However, $xyyz \notin F$, because $|y| > 0$ by assumption so this string can't be divided into identical substrings of equal length. Thus s cannot be pumped; F does not satisfy the pumping lemma and thus F is nonregular.

Example (PL #2.)

Show that $D = \{1^{n^2} \mid n \geq 0\}$ is nonregular.

Proof. (1) Assume D is regular

(2) By assumption, D satisfies the pumping lemma: $\exists p$ such that for all $s \in D$, $|s| \geq p$, s can be divided into xyz satisfying $|y| > 0$, $|xy| \leq p$, $xy^iz \in D$ for all $i \geq 0$.

(3) Choose $w = 1^{p^2}$.

Suppose we divide w into xyz, $|xy| \leq p$, $|y| > 0$.

$$|xyz| = |w| = p^2.$$

$$|xyyz| = |xyz| + |y| = p^2 + |y| > p^2, \text{ as } |y| > 0.$$

$$p^2 + |y| \leq p^2 + |xy| \leq p^2 + p < p^2 + 2p + 1 = (p+1)^2.$$

We have $p^2 < |xyyz| < (p+1)^2$, so the length of $xyyz$ is not a perfect square. $xyyz \notin D$. D fails the pumping lemma and thus D is nonregular. ∎

Example. Show $\{0^n 1^n \mid n \geq 3\}$ is nonregular.

Proof. We know $\{0^n 1^n \mid n \geq 0\}$ is nonregular.

We observe that $\underbrace{\{0^n 1^n \mid n \geq 0\}}_{\text{nonreg}} = \underbrace{\{0^n 1^n \mid n \geq 3\}}_{\text{target}} \cup \underbrace{\{0^n 1^n \mid n < 3\}}_{\text{Reg}}$

$\{0^n 1^n \mid n < 3\} = \{\varepsilon, 01, 0011\}$. ✓ Regular language because finite.

If $\{0^n 1^n \mid n \geq 3\}$ is regular, then its union with $\{0^n 1^n \mid n < 3\}$ is regular by the closure of regular languages under union. This is a contradiction, so our target language is nonregular. $\blacksquare$

Next: Context-Free Grammars