BU CS 332 – Theory of Computation

https://forms.gle/8DBrFvVU4M7Twrsk7

Lecture 8:

- More on non-regularity
- Test 1 Review

Reading:

"Myhill-Nerode" note

Mark Bun

February 14, 2024

Last Time: Distinguishing Set Method

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

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.

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

Corollary: If language L has an infinite pairwise distinguishable set, then L is not regular.

Reusing a Proof

Finding a distinguishing set can take some work…

Let's try to reuse that work!

How might we show that

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

is not regular?

Not regular $\quad$ $= L(0^*1^*)$ $\quad$ regular

$$\boxed{\{0^n1^n \mid n \geq 0\}} = BALANCED \cap \{w \mid \text{all 0s in } w \text{ appear before all 1s}\}$$

Claim. BALANCED is not regular.

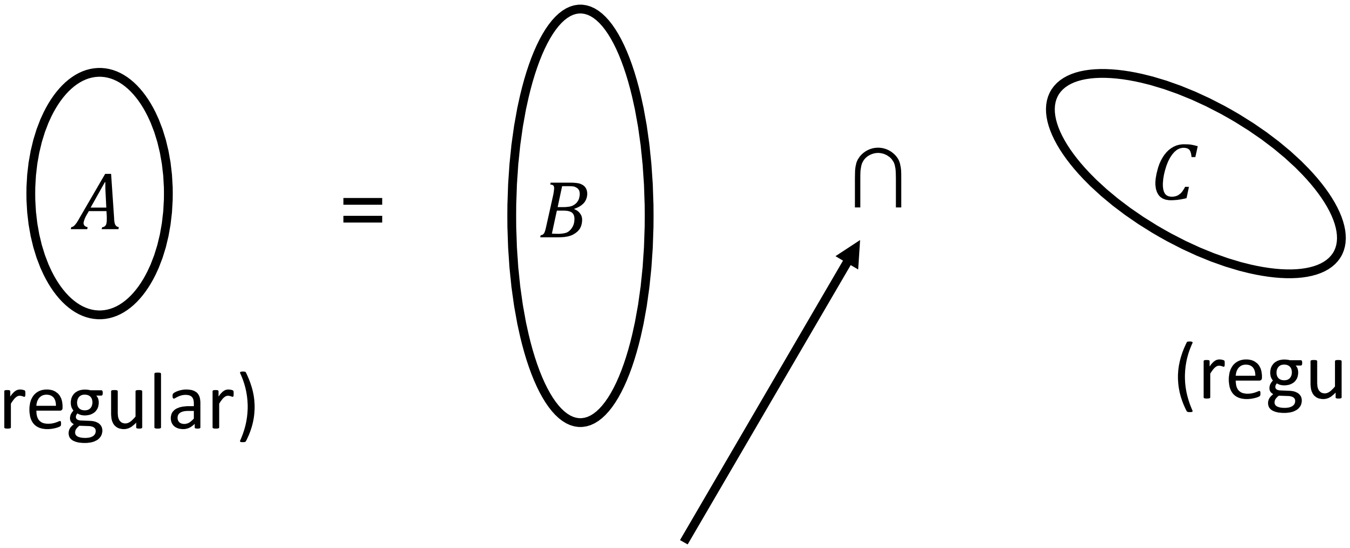
Proof. Assume for contradiction that BALANCE is regular

$\Rightarrow$ BALANCED $\cap$ $L(0^*1^*)$ is regular (because reg. langs. closed under $\cap$)

$\Rightarrow$ $\{0^n1^n \mid n \geq 0\}$ is regular $\quad$ ✗ $\Rightarrow$ conclude BALANCED is not regular

Using Closure Properties

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



(not regular)

(regular)

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

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

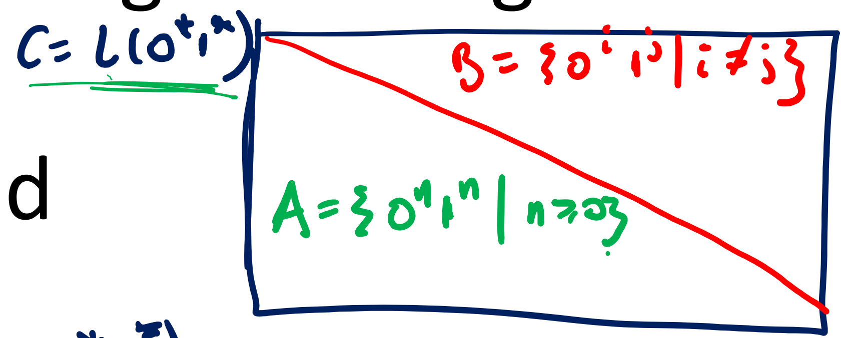
But A is not regular so neither is B!

Example

Prove $B = \{0^i 1^j | i \neq j\}$ is not regular using

- Nonregular language

  $A = \{0^n 1^n | n \geq 0\}$ and



- Regular language

  $L(0^*1^*)$

  $C = \{w \mid$ all 0s in w appear before all 1s$\}$

$A = C \setminus B = \bar{B} \cap C$

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

a) $A = B \cap C$

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

c) $A = B \cup C$

d) $A = \bar{B} \cup C$

Proof that B is nonregular

Assume for the sake of contradiction that B is regular

We know: $A = \bar{B} \cap C$ know C is regular because $C = L(0^*1^*)$

B regular $\Rightarrow$ $\bar{B}$ regular (reg. langs. closed under complement)

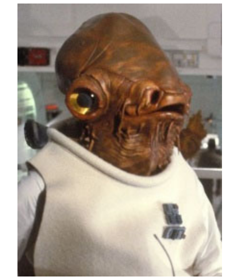
$\Rightarrow$ $\bar{B} \cap C$ regular (" closed under intersection)

$\Rightarrow$ A regular

which contradicts fact that A is non-regular

so conclude B must be nonregular.

!DANGER!



Let $B = \{0^i 1^j | i \neq j\}$ and write $B = A \cup C$ where

$B' = L(0^* 1^*)$

- Nonregular language

$$A = \{0^i 1^j | i > j \geq 0\} \text{ and}$$

$A' = \{0^i 1^j | i \geq j \geq 0\}$

$C = \{0^i 1^j | j > i \geq 0\}$

- Nonregular language

$$C = \{0^i 1^j | j > i \geq 0\} \text{ and}$$

Does this let us conclude B is nonregular?

Issue: Nonregular languages are not closed under union

so. it is possible for the union of two nonregular languages to be regular

A' and C are such an example

Let D be any nonregular language over Σ. Then $\overline{D}$ is nonregular.

$D \cup \overline{D} = \Sigma^*$ is regular

(Because if $\overline{D}$ were regular, then $D = \overline{(\overline{D})}$ would be regular)

Test 1 Topics

Sets, Strings, Languages (0)

- Know the definition of a string and of a language (and the difference between them)
- Understand operations on strings: Concatenation, reverse

  $x \circ y = xy = x_1 \dots x_n y_1 \dots y_n$

- Understand operations on languages: Union, intersection, concatenation, reverse, star, complement

  $L_1 \circ L_2 = \{ xy \mid x \in L_1, y \in L_2 \}$

- Know the difference between ∅ and ε

Deterministic FAs (1.1)

- Given an English or formal description of a language L, draw the state diagram of a DFA recognizing L (and vice versa)
- Know the formal definition of a DFA (A DFA is a 5 tuple…) and convert between state diagram and formal description
- Know the formal definition of how a DFA computes
- Construction for closure of regular languages under complement

Nondeterministic FAs (1.2)

- Given an English or formal description of a language L, draw the state diagram of an NFA recognizing L (and vice versa)
- Know the formal definition of an NFA
- Know the subset construction for converting an NFA to a DFA
- Proving closure properties: Know the constructions for union, concatenation, star
- Know how to prove your own closure properties

Regular Expressions (1.3)

- Given an English or formal description of a language L, construct a regex generating L (and vice versa)
- Formal definition of a regex

  Base regexes: ε, ϕ, alphabet symbo a

  $R_1 \cup R_2$ $R_1 \circ R_2$ R_1^*
- Know how to convert a regex to an NFA
- Know how to convert a DFA/NFA to a regex

Limitations of DFAs (Myhill-Nerode Note)

- Understand the statements of the distinguishing set method for proving DFA size lower bounds / non-regularity
- Understand the proof of why the distinguishing set method works, and be able to use it to prove similar statements
- Know how to apply the method to specific languages
- Note: I won't ask you to show anything is non-regular, since you didn't have any homework problems on this yet

Test format

Problem 1: "Check your type checker"

E.g., Is aabba a string, language, or a regex?

How about {ab} U {aab}?

language

Problem 2: True/false with **justification**

Either provide a convincing explanation or a specific counterexample

Problems 3-5(?) Homework-style problems

Test tips

- You may cite without proof any result…
  - Stated in lecture + Myhill-Nerode note
  - Stated and proved in the main body of the text (Ch. 0-1.3)
  - These include worked-out examples of state diagrams, regexes
- Not included above: homework problems, discussion problems, (solved) exercises/problems in the text

- Showing your work / explaining your answers will help us give you partial credit
- Make sure you're interpreting quantifiers (for all / there exists) correctly and in the correct order

Practice Problems

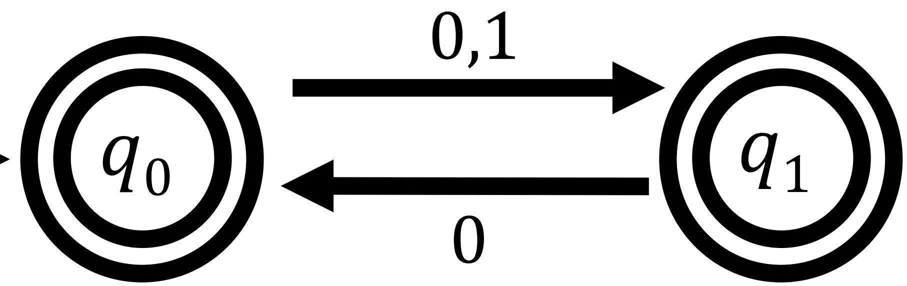
Name six operations under which the regular languages are closed

Prove or disprove: The **non-**regular languages are closed under complement

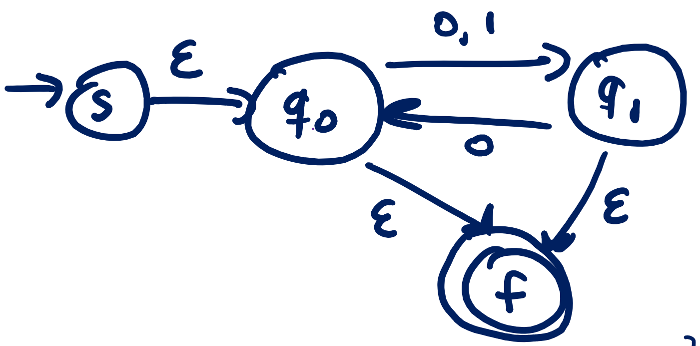
Prove or disprove: The **non-**regular languages are closed under intersection

Give the state diagram of an NFA recognizing the language $(01 \cup 10)^* \circ 1$

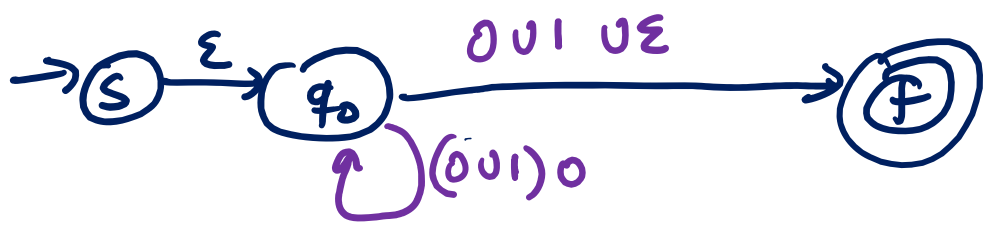
Give an equivalent regular expression for the following NFA



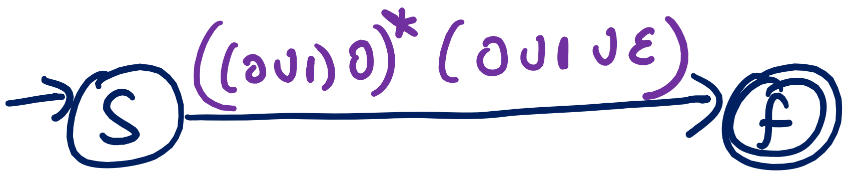
1) Convert to GNFA



2) Rip out q_1



3) Rip out q_0

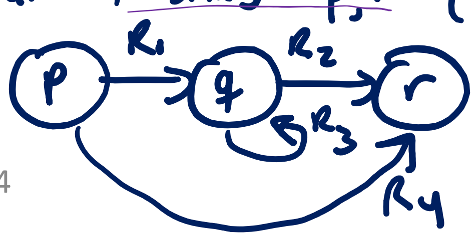


Final regex: $((0 \cup 1)0)^*(0 \cup 1 \cup \varepsilon)$

To rip out q:

For every pair of states p, r ($\neq q$):

Replace



with

For a language L over $\{0,1\}$, define the operation $\mathbf{split}(L) = \{x\#y \mid x, y \in L\}$. Show that the regular languages are closed under split

over $\{0,1,\#\}$

Let A be regular. Then $\exists$ regex R generating A.

$$\text{split}(A) = \cancel{L\left((0\cup 1)^* \# (0\cup 1)^*\right)}$$

$$L(R \# R)$$

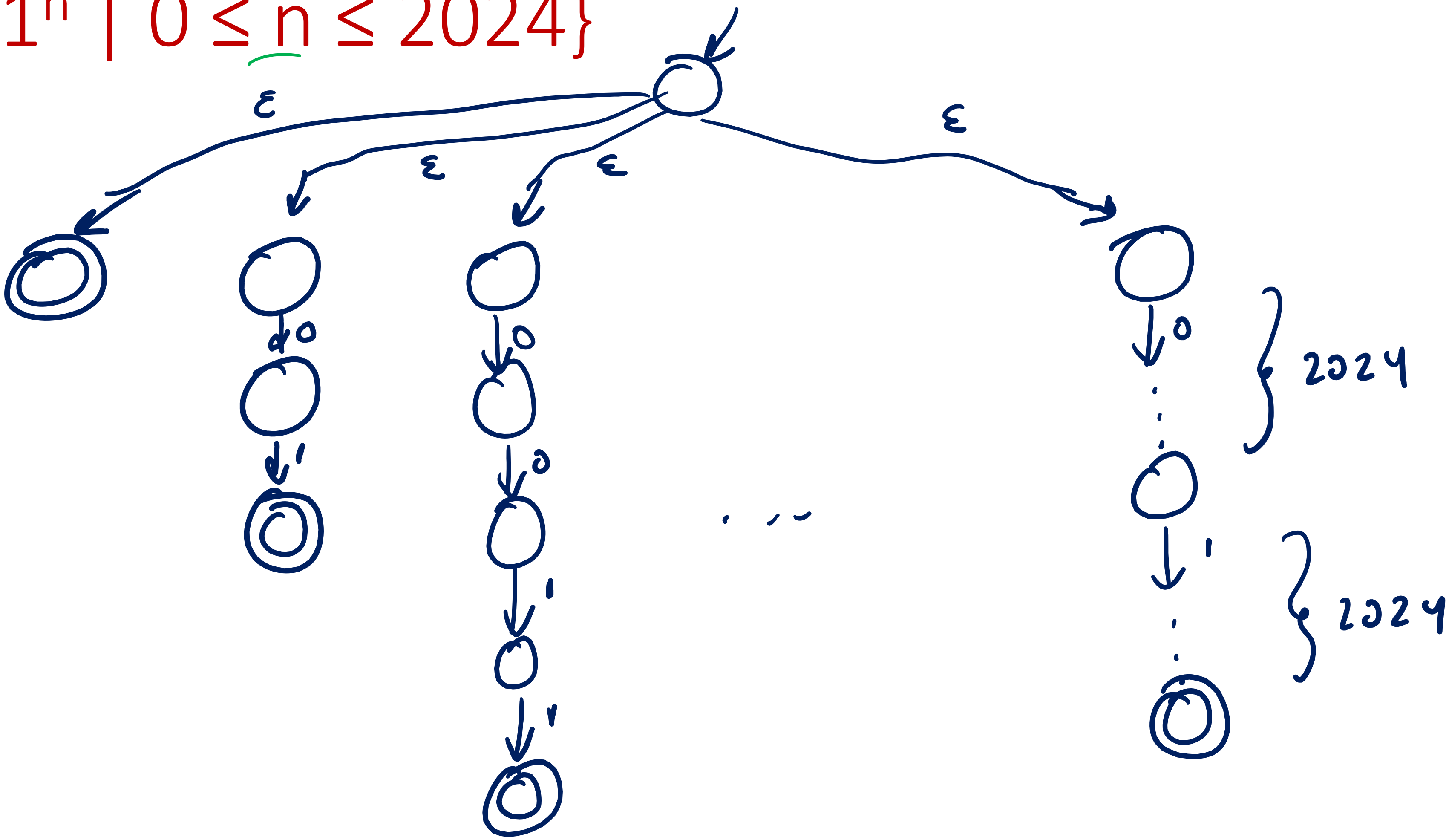
$\Rightarrow$ split(A) is regular

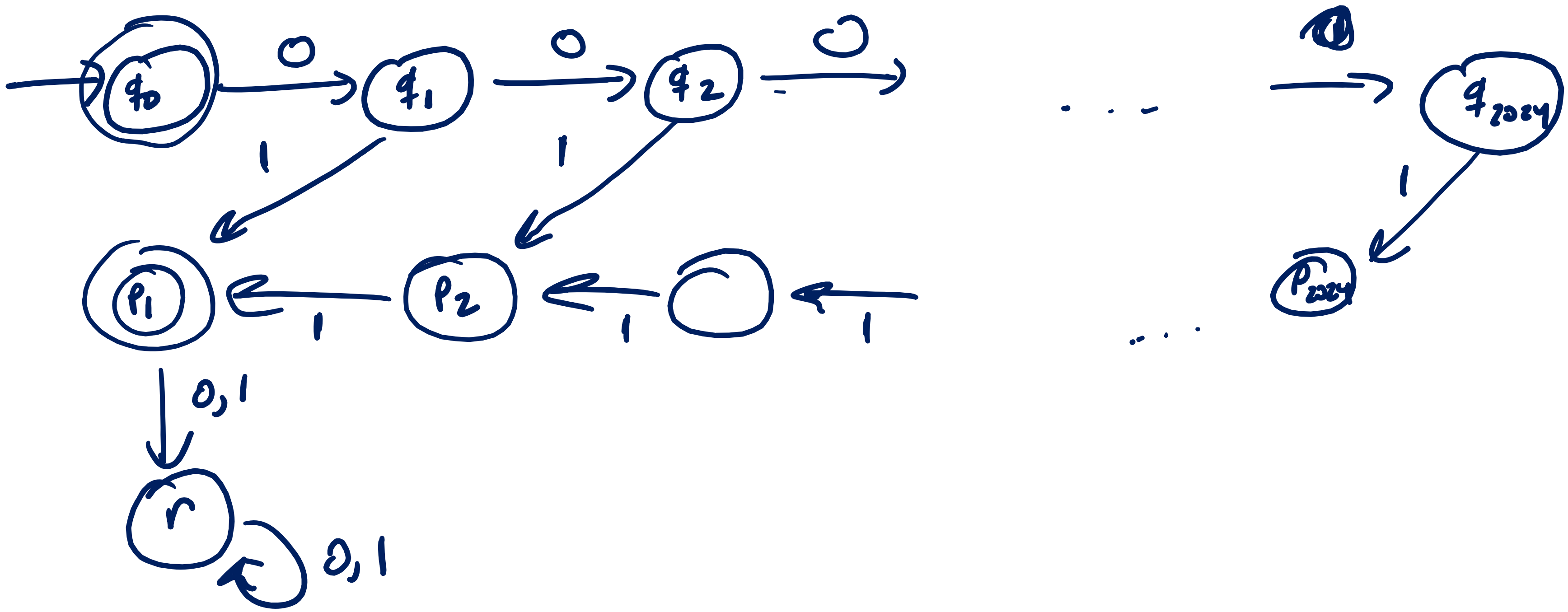
$\Rightarrow$ reg langs. closed under split.

For a language L over alphabet Σ, define the operation $\mathbf{drop}(L) = \{xyz \mid xyz \in L, xy \in \Sigma^*, z \in \Sigma\}$. Show that the regular languages are closed under $\mathbf{drop}$.

Is the following language regular?

$\{0^n 1^n \mid 0 \leq n \leq 2024\}$





q_0
0
q_1
0
q_2
0
0
q_{2024}
1
1
1
p_1
1
p_2
1
1
p_{2024}
0, 1
r
0, 1

How many states does a DFA recognizing $L = \{0^n 1^n \mid 0 \le n \le 2024\}$ require?

Goal: Construct a "big" distinguishing set for L.

Claim: $S = \{0^n \mid 0 \le n \le 2024\}$ is pairwise dist. for L.

WTS: $\forall\, x \ne y \in S$, $\exists\, z$ s.t. exactly one of xz, yz is in L

Proof. Let $x = 0^m$ $y = 0^n$ where $0 \le m, n \le 2024$, $m \ne n$

Set $z = 1^m$. Then $xz = 0^m 1^m \in L$

$yz = 0^n 1^m \notin L$

$\Rightarrow$ conclude any DFA for L needs ≥ 2025 states.

Claim: $T = \{0^n \mid 0 \le n \le 2024\} \cup \{0^n 1 \mid 0 \le n \le 2024\}$ is pairwise dist by L??

$\{0^n 1^m \mid 0 \le n \le 2024,\ 0 < m \le n\}$

can't work: bigger than DFA we already constructed