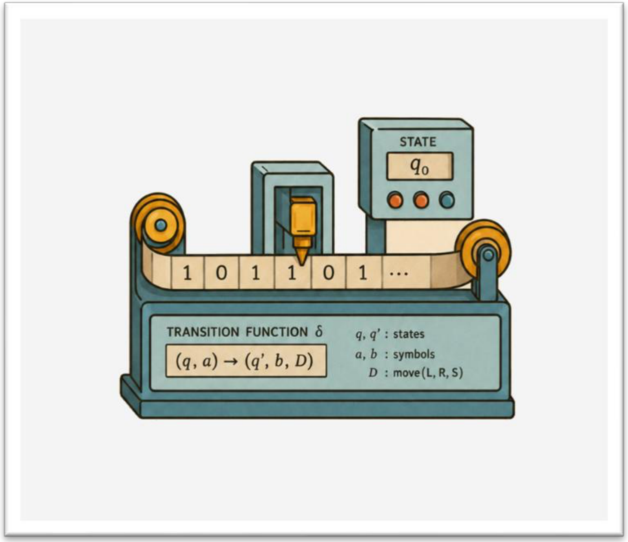
# BU CS 332 – Elements of the Theory of Computation

- **Lecture 2:**
  - Sets
  - Strings
  - Languages
  - **Quiz**




**Alexander Poremba**

September 8, 2026

# What Is A (Computational) Problem?

**For us:** a (computational) problem is the task of determining whether a string is in a language.

**Example:**

Primality: given a natural number $x$ (represented in binary), is $x$ a prime number?

$$\Sigma = \{0,1\}$$

$$L = \{ x \in \{0,1\}^* \mid x \text{ is a binary repr. of a prime} \}$$

# Strings and languages

# String Theory

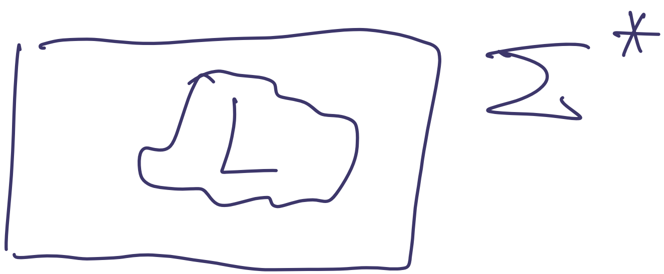
- **Symbol:** **Ex.** a, b, 0, 1
- **Alphabet**: A finite set $\Sigma$ of symbols **Ex.** $\Sigma = \{a, b\}$
- **String:** A finite concatenation of alphabet symbols **Ex.** bba, ababb

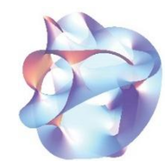
  $\varepsilon$ denotes empty string, length 0

  $\Sigma^*$ = set of all strings using symbols from $\Sigma$

  **Ex.** $\{a, b\}^* = \{\varepsilon, a, b, aa, ab, ba, bb, \ldots\}$

- **Language:** A set $L \subseteq \Sigma^*$ of strings



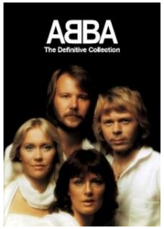



# String Theory

- **Length** of a string, written $|x|$, is the number of symbols

Ex. $|\text{abba}| =$ 4 $\qquad |\varepsilon| =$ 0

- **Concatenation** of strings $x$ and $y$, written $xy$, is the symbols from $x$ followed by the symbols from $y$

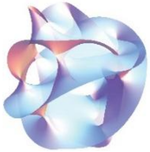



Ex. $x = \text{ab}, y = \text{ba} \quad \Rightarrow \quad xy =$ abba

$x = \text{ab}, y = \varepsilon \quad \Rightarrow \quad xy =$ ab

- **Reversal** of string $x$, written $x^R$, consists of the symbols of $x$ written backwards

Ex. $x = \text{aab} \quad \Rightarrow \quad x^R =$ baa



# Fun with String Operations

**Question:** What is $(xy)^R$?

Ex. $x = \text{aba}, y = \text{bba} \quad \Rightarrow \quad xy =$ aba bba

$\Rightarrow \quad (xy)^R =$ abbaba

a) $x^R y^R$

b) $y^R x^R$

c) $(yx)^R$

d) $xy^R$

# Fun Proofs with String Operations

**Claim:** $(xy)^R = y^R x^R, \ \forall x, y \in \Sigma^*$

**Proof:** Let $x = x_1 x_2 \dots x_n$ and $y = y_1 y_2 \dots y_m$

Then $(xy)^R = (x_1 x_2 \cdots x_n y_1 y_2 \cdots y_m)^R$

$= y_m y_{m-1} \cdots y_1 x_n x_{n-1} \cdots x_1$

$= \underline{(y_m y_{m-1} \cdots y_1)}\underline{(x_n x_{n-1} \cdots x_1)}$

$= y^R x^R.$

$\square$

# Languages

A language $L$ is a set of strings over an alphabet $\Sigma$

i.e., $L \subseteq \Sigma^*$


$\Sigma^*$

$x \in \Sigma^*$

**Languages = computational (decision) problems**

<u>Input:</u> String $x \in \Sigma^*$

<u>Output:</u> Is $x \in L$? (Yes or No?)

# Some Simple Languages

$L \subseteq \Sigma^*$

| | $\Sigma = \{0, 1\}$ | $\Sigma = \{a, b, c\}$ |
|---|---|---|
| Ø (Empty set) | $\{\ \}$ | $\{\ \}$ |
| $\Sigma^*$ (All strings) | $\{0,1\}^* = \{\varepsilon, 0, 1, 00, 01, \ldots\}$ | $\{a,b,c\}^* = \{\varepsilon, a, b, c, aa, ab, ac, \ldots\}$ |
| $\Sigma^n = \{x \in \Sigma^* \mid \lvert x \rvert = n\}$ (All strings of length $n$) | $n=2$: $\{0,1\}^2 = \{00, 01, 11, 10\}$ | $n=2$: $\{a,b,c\}^2 = \{aa, bb, cc, ab, ac, bc, ca, cb, ba\}$ |

## Some More Interesting Languages

- $L_1$ = The set of strings $x \in \{a, b\}^*$ that have an equal number of a's and b's

$L_1 = \{ x \in \{a,b\}^* \mid \#a\text{'s is equal to } \#b\text{'s} \}$

- $L_2$ = The set of strings $x \in \{a, b\}^*$ that start with (0 or more) a's and are followed by an equal number of b's

$L_2 = \{ yz \mid y \in \{a\}^*, \ z \in \{b\}^* \text{ s.t. } |y| = |z| \}$

- $L_3$ = The set of strings $x \in \{0, 1\}^*$ that contain the substring "0100"

$L_3 = \{ abc \mid a \in \{0,1\}^*, \ b = 0100, \ c \in \{0,1\}^* \}$

# Some More Interesting Languages

- $L_4 =$ The set of strings $x \in \{\mathrm{a}, \mathrm{b}\}^*$ of length at most 4

$$L_4 = \{ x \in \{a,b\}^* \mid |x| \leq 4 \}$$

- $L_5 =$ The set of strings $x \in \{\mathrm{a}, \mathrm{b}\}^*$ that contain at least two a's

$$L_5 = \{ xayaz \mid x, y, z \in \{a,b\}^* \}$$

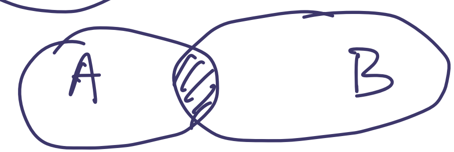
# New Languages from Old

$L_6$ = The set of strings $x \in \{\mathrm{a}, \mathrm{b}\}^*$ that have an equal number of a's and b's and length greater than 4

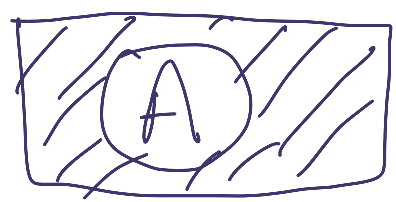
Since languages are just sets of strings, can build them using set operations:

$A \cup B$ "union"


$A \cap B$ "intersection"




$\bar{A}$ "complement"



# New Languages from Old

$L_6$ = The set of strings $x \in \{\mathrm{a}, \mathrm{b}\}^*$ that have an equal number of a's and b's and have length greater than 4

- $L_1$ = The set of strings $x \in \{\mathrm{a}, \mathrm{b}\}^*$ that have an equal number of a's and b's
- $L_4$ = The set of strings $x \in \{\mathrm{a}, \mathrm{b}\}^*$ of length at most 4

$$\Rightarrow L_6 = L_1 \cap \overline{L_4}$$

# Operations Specific to Languages

- **Reverse:** $L^R = \{x^R \mid x \in L\}$

  Ex. $L = \{\varepsilon, \text{a}, \text{ab}, \text{aab}\}$ $\Rightarrow L^R =$ $\{\varepsilon, a, ba, baa\}$

- **Concatenation:** $L_1 \circ L_2 = \{xy \mid x \in L_1, y \in L_2\}$

  Ex. $L_1 = \{\text{ab}, \text{aab}\}$ $L_2 = \{\varepsilon, \text{b}, \text{bb}\}$

  $\Rightarrow L_1 \circ L_2 =$

$\{ab, abb, abbb, aab, aabb, aabbb\}$

## A Few "Traps"

$L \subseteq \Sigma^*$

String, language, or something else?

$\varepsilon$ string

$\emptyset$ language

$\{\varepsilon\}$ language

$\{\emptyset\}$ something else!

$\{\emptyset\} \overset{?}{\subseteq} \Sigma^*$ ← set of strings

↑ set of sets

# Languages

**Languages = computational (decision) problems**

Input: String $x \in \Sigma^*$

Output: Is $x \in L$? (Yes or No? I.e., Accept or Reject?)

We say that the language **recognized** by a program is the set of strings $x \in \Sigma^*$ that it *accepts*.

$$x \in \Sigma^* \longrightarrow \boxed{P} \longrightarrow \text{accept / reject}$$

$$L = \{ x \in \Sigma^* \mid P \text{ accepts } x \}$$

# Question: What Language Does This Program Recognize?

Alphabet $\Sigma = \{\mathrm{a}, \mathrm{b}\}$

On input $x = x_1 x_2 \dots x_n$:

count = 0

For $i = 1, \dots, n$:

If $x_i = \mathrm{a}$:

count = count + 1

If count ≤ 4: accept

Else: reject

a) $\{x \in \Sigma^* \mid |x| > 4\}$

b) $\{x \in \Sigma^* \mid |x| \leq 4\}$

c) $\{x \in \Sigma^* \mid |x| = 4\}$

d) $\{x \in \Sigma^* \mid x \text{ has more than } 4 \text{ a's}\}$

e) $\{x \in \Sigma^* \mid x \text{ has at most } 4 \text{ a's}\}$

f) $\{x \in \Sigma^* \mid x \text{ has exactly } 4 \text{ a's}\}$

# Quiz time!