Comp 330 - Lecture 2 - September 5th 2023

Admin

1. Ed & Crowdmark
2. Discord / Whatsapp
3. TA OHs
4. A1 released at 10AM

Plan for today

1. Alphabets, strings, languages
2. Intro to Automata Theory
   ↳ Deterministic Finite Automaton

Alphabet

Def An alphabet is a finite set of symbols / letters. $\Sigma \neq \emptyset$

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

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

$\Sigma = \{a,b,c,d,\ldots,z\} \rightarrow$ English letters

String

Def Given alphabet Σ, a string is a sequence of symbols from Σ.

Ex $\Sigma = \{a, b\}$

$w_1 = abb$ $\quad$ $w_2 = abc$ → Keyboard

$w_1 = \cancel{X}abb\cancel{X}$ $\quad$ valid

Remark $\Sigma = \{a, b\}$ How do we differentiate between symbol a and a string a

symbol $\qquad$ string

To do this: Use start and end of string marker

$\cancel{X}\bar{a}\cancel{X}$ vs a

string $\qquad$ symbol

"a"

String properties and operations

String length

Def Consider a string w defined using alphabet Σ i.e.

$$w = a_1 a_2 a_3 \dots a_n \qquad a_i \in \Sigma,\ 1 \leq i \leq n$$

Then the length of w, $|w|$, is the number of symbols in w. $|w| = n$

Remark Comp 330, strings are always of finite length

Def (Empty string) Given Σ, the empty string is the unique string with length 0. Typically, denote the empty string as ε.

$\hookrightarrow$ $x = $ ""

Ex Suppose that $\Sigma = \emptyset$. What are the possible strings that can be created using this Σ?

It's ε!

$\emptyset \subseteq \emptyset, \quad \varepsilon \notin \emptyset$

String concatenation

$x =$ "comp" $\quad y =$ "330"

$x + y =$ "comp330"

string concat

Def Given Σ, $w = a_1 \dots a_n$, $v = b_1 \dots b_m$

$a_i \in \Sigma$, $1 \leq i \leq n$ $\qquad$ $b_j \in \Sigma$, $1 \leq j \leq m$

The string concatenation of w and v is the operation by which you append v to w:

$w \cdot v = wv = a_1 \dots a_n b_1 \dots b_m$

Ex What happens if $x = \varepsilon$ $y = \varepsilon$?

$xy = \varepsilon\varepsilon = \varepsilon$

Remark 1. String concatenation is not commutative

Σ, u, v $\quad u \cdot v \neq v \cdot u$

String concatenation is associative

Σ, u, v, w $\quad (u \cdot v) \cdot w = u \cdot (v \cdot w)$

Remark 2 The empty string acts as an identity for string concatenation

Σ, x, $\quad \varepsilon \cdot x = x$

$x \cdot \varepsilon = x$

Remark 3 Σ, x, y $\quad |x \cdot y| = |x| + |y|$

Notation

$\Sigma = \{a, b\}$ $\quad w = \underbrace{aaa \cdots a}_{330 \text{ times}} = a^{330}$ => Symbol exponentiation

Formally a^n, $n \in \mathbb{N}$ $\quad$ Defined as $\triangleq$

$a^0 := \varepsilon$

$a^{n+1} := a^n \cdot a$

Inductive / recursive definitions

Formally (String exponentiation)

Σ, x is a string, $\quad x^n = x \cdot x \cdot x \cdot \ldots \cdot x$

n times

$$x^0 := \varepsilon \quad , \quad x^{n+1} := x^n \cdot x$$

Ex $a^n, a^n b^n, \ldots$

Exercise $\Sigma = \{a, b\}$ $\quad w = ab \quad v = ba$

What is $w^2 v^2 w^0$?

$$w^2 v^2 w^0 = (ab)^2 (ba)^2 \cdot \varepsilon$$

$$= abab\,baba$$

$\hookrightarrow$ Palindrome

Languages

Def 1 A language is a set of strings defined over some Σ.

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

$$\Sigma^* = \underbrace{\{\varepsilon\}}_{\text{length } 0} \cup \underbrace{\{a, b\}}_{\text{length } 1} \cup \{aa, ab, ba, bb\} \cup \ldots$$

Σ^*: language containing all strings defined using Σ

Def 2 (Given Σ, A language L is a subset of Σ^*.

$$L \subseteq \Sigma^*$$

(*) Strings always finite length

Ex $\Sigma = \{a, b\}$

$L_1 = \{a, b, aa, ab, ba, bb\}$ ← Finite,

$L_2 = \{a^n : n \in \mathbb{N},\ n \text{ is even}\}$ ← Infinite / Regular

$L_3 = \{a^n b^n : n \in \mathbb{N}\}$ ← Infinite / Not regular

Ex $\Sigma = \emptyset$ $\quad \Sigma^{\boxed{*}} = \{\varepsilon\}$

Ex $L = \boxed{\emptyset \neq \{\varepsilon\}}$ (A)

Who cares about languages?

Decision Problem: A problem where the answer is Y / N

Ex: "Given a number $n \in \mathbb{N}$, is n even?"

$\Updownarrow$

Languages and language membership are DPs

L_2, is $\underline{a^n \in L_2} \iff \underline{\text{is } n \text{ even}}$?

Languages offer more flexibility:

Σ, $\underbrace{w_i \in L_1, L_2, L_3}_{\text{"easy to solve for language membership"}}$ $\Rightarrow \underbrace{w \in (L_1 \cup L_2 \cup L_3)}_{\text{"also easy, don't need an explicit solution"}}$ → True for regular languages

Language operations

Operations from set theory : $\cup, \cap, -$

Language concatenation

Def Σ, $L_1, L_2 \subseteq \Sigma^*$, the language concatenation

$L_1 \cdot L_2 = \{xy : x \in L_1, y \in L_2\}$

$(L_1 \cdot L_2 \neq L_2 \cdot L_1)$

Ex $\Sigma = \{a, b\}$ $L_1 = \{a\}$ $L_2 = \{b^n : n \in \mathbb{N}\}$ 4

$L_1 \cdot L_2 = \{a \cdot b^n : n \in \mathbb{N}\}$

$L_2 \cdot L_1 = \{b^n \cdot a : n \in \mathbb{N}\}$

Notation Language exponentiation

$n \in \mathbb{N}$, $L^n = \underbrace{L \cdot L \cdot L \cdot \ldots \cdot L}_{n \text{ times}}$

Def $L^0 := \{\varepsilon\}$ $\rightarrow$ ⊛ True even if $L = \emptyset$

$L^{n+1} := L^n \cdot L$ $\qquad \emptyset^0 \neq \emptyset$

Ex $\Sigma = \{a, b\}$, $L = \{a, b\}$, $n \geq 2$

$L^n \neq \{a^n, b^n\}$ 4

$L^n = \underbrace{\{a, b\} \cdot \{a, b\} \cdot \ldots \cdot \{a, b\}}_{n \text{ times}}$

$x \in L^n$, $x = a_1 a_2 \dots a_n$
$\quad a_i \in \{a, b\}$

$L^n = \{x \in \{a,b\}^* : |x| = n\}$

<u>Exercise</u> (For you!) $\Sigma = \{a\}$, $L \subseteq \Sigma^*$, L is finite,
If $n \geq 2$, $|L^n| = |L|^n$? (True / False)

<u>Star operator</u> & <u>Plus operator</u> (Kleene $\to$ Stephen <u>Kleene</u>)

<u>Def</u> Σ, $L \subseteq \Sigma^*$

$$L^* := L^0 \cup L^1 \cup L^2 \cup \dots$$
$$= \bigcup_{i \geq 0} L^i$$
$$:= \{w_1 w_2 \dots w_n : \underset{n=0 \Rightarrow \varepsilon}{\underline{n \geq 0}},\ w_i \in L,\ 1 \leq i \leq n\}$$

$\boxed{\Sigma^*}$ exactly the star operator applied Σ.

<u>Def</u> Σ, $L \subseteq \Sigma^*$

$$L^+ := L^1 \cup L^2 \cup L^3 \cup \dots$$
$$:= L \cdot L^*$$

Language concatenation distributes over language union.

$\varepsilon \notin L^+$, $\varepsilon \in L^+ \iff \varepsilon \in L$

Ex How many strings of length 0, 1, 2 are in L^*/L^+ for $L = \{a, ab\}$?

$L^0 = \{\varepsilon\}$

$L^1 = \{a, ab\}$

$L^2 = \{aa, aab, aba, abab\}$ → Check why.

$L^3 =$

	Strings of len 0	1	2
L^+	×	a	ab aa
L^*	ε	a	ab aa

Some properties about $*$:

1. $(L^*)^* = L^*$
2. $\emptyset^* = \{\varepsilon\}$