CSC 339 – Theory of Computation
Fall 2023-2024

2. Languages

Outline

- Alphabet and strings
- String operations
- Languages
- Operations on languages
- The membership problem

2

General concepts

- Language: a set of strings
- String: a sequence of symbols from some alphabet
- Example:

  Strings: cat, dog, house

  Language: {cat, dog, house}

  Alphabet: $\Sigma = \{a, b, c, \ldots, z\}$

General Concepts

Languages are used to describe computation problems

$$PRIMES = \{2,3,5,7,11,13,17,\ldots\}$$

$$EVEN = \{0,2,4,6,\ldots\}$$

Alphabet: $\Sigma = \{0,1,2,\ldots,9\}$

4

Alphabets and Strings

An alphabet is a set of symbols

Example Alphabet: $\Sigma = \{a, b\}$

A string is a sequence of symbols from the alphabet

Example Strings:

a

ab

$abba$

$aaabbbaabab$

$u = ab$

$v = bbbaaa$

$w = abba$

Alphabets and Strings

Examples of strings

Decimal numbers alphabet $\Sigma = \{0,1,2,\ldots,9\}$

102345

567463386

Binary numbers alphabet $\Sigma = \{0,1\}$

100010001

101101111

Unary numbers alphabet $\Sigma = \{1\}$

1 11 111 1111

String Operations

$w = a_1 a_2 \cdots a_n$ $\qquad$ $abba$

$v = b_1 b_2 \cdots b_m$ $\qquad$ $bbbaaa$

Concatenation

$wv = a_1 a_2 \cdots a_n b_1 b_2 \cdots b_m$ $\qquad$ $abbabbbaaa$

String Operations

$w = a_1 a_2 \cdots a_n$ $\qquad$ $ababaaabbb$

Reverse

$w^R = a_n \cdots a_2 a_1$ $\qquad$ $bbbaaababa$

String Length

$$w = a_1 a_2 \cdots a_n$$

- Length: $|w| = n$
- Examples:

$$|abba| = 4$$

$$|aa| = 2$$

$$|a| = 1$$

9

Length of Concatenation

$$|uv| = |u| + |v|$$

- Examples: $u = aab, \quad |u| = 3$

  $v = abaab, \quad |v| = 5$

  $|uv| = |aababaab| = 8$

  $|uv| = |u| + |v| = 3 + 5 = 8$

10

Empty String
ε or λ

- A string with no letters
  is denoted: $|\varepsilon| = 0$
- Observations:

$$\varepsilon w = w\varepsilon = w$$

$$\varepsilon abba = abba\varepsilon = ab\varepsilon ba = abba$$

11

Substring

- Substring of string:
  - a subsequence of consecutive characters

String	Substring
abbab	*ab*
abbab	*abba*
abbab	*b*
abbab	*bbab*

12

Prefix and Suffix

String: $abbab$

Prefixes	Suffixes
ε	$abbab$
a	$bbab$
ab	bab
abb	ab
$abba$	b
$abbab$	ε

$$w = uv$$

prefix

suffix

Another Operation

$$w^n = \underbrace{ww\cdots w}_{n}$$

- Example: $(abba)^2 = abbaabba$

- Definition: $w^0 = \varepsilon$

$$(abba)^0 = \varepsilon$$

The * Operation

- Σ^* : the set of all possible strings from alphabet Σ

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

$$\Sigma^* = \{\varepsilon, a, b, aa, ab, ba, bb, aaa, aab, \ldots\}$$

15

The + Operation

Σ^+: the set of all possible strings from alphabet Σ except ε

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

$$\Sigma^* = \{\varepsilon, a, b, aa, ab, ba, bb, aaa, aab, \ldots\}$$

$$\Sigma^+ = \Sigma^* - \varepsilon$$

$$\Sigma^+ = \{a, b, aa, ab, ba, bb, aaa, aab, \ldots\}$$

16

Languages

- A language over alphabet Σ is any subset of Σ^*
- Examples:

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

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

Language: $\{\varepsilon\}$

Language: $\{a, aa, aab\}$

Language: $\{\varepsilon, abba, baba, aa, ab, aaaaaa\}$

More Language Examples

Alphabet $\Sigma = \{a, b\}$ $\quad L = \{a^n b^n : n \geq 0\}$

- An infinite language

$$\left.\begin{array}{l} \varepsilon \\ ab \\ aabb \\ aaaaabbbbb \end{array}\right\} \in L \qquad abb \notin L$$

18

More Language Examples

Prime numbers

Alphabet: $\Sigma = \{0,1,2,\ldots,9\}$

Language:

$$PRIMES = \{x : x \in \Sigma^* \text{ and } x \text{ is prime}\}$$

$$PRIMES = \{2,3,5,7,11,13,17,\ldots\}$$

More Language Examples

Even and odd numbers

Alphabet: $\Sigma = \{0,1,2,\ldots,9\}$

$$EVEN = \{x : x \in \Sigma^* \text{ and } x \text{ is even}\}$$

$$EVEN = \{0,2,4,6,\ldots\}$$

$$ODD = \{x : x \in \Sigma^* \text{ and } x \text{ is odd}\}$$

$$ODD = \{1,3,5,7,\ldots\}$$

More Language Examples

Unary Addition

Alphabet: $\Sigma = \{1, +, =\}$

Language:

$$ADDITION = \{x + y = z \ : x = 1^n, y = 1^m, z = 1^k, \\ n + m = k\}$$

$$11 + 111 = 11111 \in ADDITION$$

$$111 + 111 = 111 \notin ADDITION$$

21

More Language Examples

Squares

Alphabet: $\Sigma = \{1, \#\}$

Language:

$$SQUARES = \{x\#y \; : x = 1^n, y = 1^m, m = n^2\}$$

$$11\#1111 \in SQUARES$$

$$111\#1111 \notin SQUARES$$

22

Important Notes

Sets $\varnothing = \{\ \} \neq \{\varepsilon\}$

Set size $|\{\ \}| = |\varnothing| = 0$

Set size $|\{\varepsilon\}| = 1$

String length $|\varepsilon| = 0$

23

Operations on Languages

- The usual set operations

$$\{a, ab, aaaa\} \cup \{bb, ab\} = \{a, ab, bb, aaaa\}$$

$$\{a, ab, aaaa\} \cap \{bb, ab\} = \{ab\}$$

$$\{a, ab, aaaa\} - \{bb, ab\} = \{a, aaaa\}$$

- Complement: $\overline{L} = \Sigma^* - L$

$$\overline{\{a, ba\}} = \{\varepsilon, b, aa, ab, bb, aaa, \ldots\}$$

Reverse

- Definition: $L^R = \{w^R : w \in L\}$
- Examples:

$$\{ab, aab, baba\}^R = \{ba, baa, abab\}$$

$$L = \{a^n b^n : n \geq 0\} \quad L^R = \{b^n a^n : n \geq 0\}$$

Concatenation

- Definition: $L_1L_2 = \{xy : x \in L_1, y \in L_2\}$
- Example:

$$\{a, ab, ba\}\{b, aa\}$$

$$= \{ab, aaa, abb, abaa, bab, baaa\}$$

Another Operation

- Definition: $L^n = \underbrace{LL \cdots L}_{n}$

$$\{a,b\}^3 = \{a,b\}\{a,b\}\{a,b\} =$$
$$\{aaa, aab, aba, abb, baa, bab, bba, bbb\}$$

- Special case:

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

$$\{a, bba, aaa\}^0 = \{\varepsilon\}$$

27

Another Operation

- Example:

$$L = \{a^n b^n : n \geq 0\}$$

$$L^2 = \{a^n b^n a^m b^m : n, m \geq 0\}$$

$$aabbaaabbb \in L^2$$

Star-Closure (Kleene *)

- All strings that can be constructed from L
- Definition: $L^* = L^0 \cup L^1 \cup L^2 \cdots$
- Example:

$$\{a, bb\}^* = \left\{ \begin{array}{l} \varepsilon, \\ a, bb, \\ aa, abb, bba, bbbb, \\ aaa, aabb, abba, abbbb, \ldots \end{array} \right\}$$

29

Positive Closure

- Definition: $L^+ = L^1 \cup L^2 \cup \cdots$

  Same with L^* but without the ε

$$\{a, bb\}^+ = \left\{ \begin{matrix} a, bb, \\ aa, abb, bba, bbbb, \\ aaa, aabb, abba, abbbb, \ldots \end{matrix} \right\}$$

The Membership Problem

*Given a string $w \in \sum$*and a language L over $\sum$, decide whether or not $w \in L$.*

Example:

Let w = 100011

Question)

Is w $\in$ the language of strings with an equal number of 0s and 1s?

31