CNG280

Formal Languages

Abstract Machines

Languages

Alphabets and Strings We will use small alphabets: $\Sigma = \{a, b\}$ A language is a set of strings Strings String: A sequence of letters Examples: "cat", "dog", "house", ... ab u = abv = bbbaaaabba Defined over an alphabet: baba w = abba $\Sigma = \{a, b, c, \dots, z\}$ aaabbbaabab

$$w = a_1 a_2 \cdots a_n$$
 $abba$ $w = a_1 a_2 \cdots a_n$ $ababaaabbb$ $bbbaaa$ $bbbaaa$

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

String Operations

Reverse $w^R = a_n \cdots a_2 a_1$

bbbaaababa

String Length
$$w = a_1 a_2 \cdots a_n$$
 Length of Concatenation
$$|uv| = |u| + |v|$$
 Length:
$$|w| = n$$
 Example: $u = aab$,
$$|u| = 3$$

$$v = abaab$$
,
$$|v| = 5$$

$$|aa| = 2$$

$$|a| = 1$$

$$|uv| = |aababaab| = 8$$

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

Empty String A string with no letters: λ (or e or ϵ)

Observations: $|\lambda| = 0$ remarks of an empty $|\lambda| = 0$

 $\lambda w = w\lambda = w$

 $\lambda abba = abba\lambda = abba$

abbab

String

Substring of string:

abbab abbab

abbab

a subsequence of consecutive characters

abba

bbab

Substring - we use these to matching, just like how do we understood

ab

that there is if , we look at the substring

Substring

Prefix and Suffix

abbab

Prefixes Suffixes

$$\lambda$$
 abbab

 a bbab

 ab prefix

 ab suffix

 ab suffix

 ab abbab

 ab suffix

 ab for take all the possibilities because don't know where to divide the string abbab

 $abbab$ λ

Another Operation $w^n = \underbrace{ww\cdots w}$

Example: $(abba)^2 = abbaabba$

Definition: $w^0 = \lambda$ $(abba)^0 = \lambda$

12

The * Operation Σ^* : the set of all possible strings from

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

alphabet
$$\Sigma$$

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

$$\Sigma^* \neq \{\lambda, a, b, aa, ab, ba, bb, aaa, aab, ...\}$$

$$* = \{\lambda, a, b, aa,$$

$$\Sigma^{+} = \Sigma^{*}$$

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

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

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

The + Operation

 Σ^+ : the set of all possible strings from

alphabet Σ except λ

$$\Sigma^{+} = \{a,b,aa,ab,ba,bb,aaa,aab,...\}$$
means one or more

Languages A language is any subset of Σ^*

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

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

Languages:
$$\{\lambda\}$$
 $\{a,aa,aab\}$

 $\{\lambda, abba, baba, aa, ab, aaaaaaa\}$

Sets

Set size

String length
$$|\lambda| = 0$$

Note that:

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

 $|\{\}| = |\emptyset| = 0$

 $\emptyset = \{\} \neq \{\lambda\}$

An infinite language $L = \{a^n b^n : n \ge 0\}$

abb ∉ L

$$\lambda$$

ab

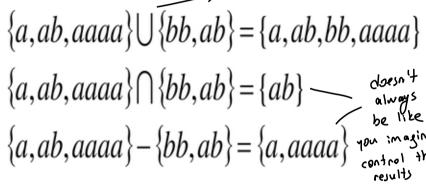
aabb

aaaaabbbbb.

$$angle \in L$$

The usual set operations
$$\{a,ab,aaaa\} \cup \widehat{\{bb,ab\}}$$

Complement:



Operations on Languages

$$\overline{L} = \Sigma$$

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

Definition: $L^R = \{w^R : w \in L\}$ Definition: $L_1L_2 = \{xy : x \in L_1, y \in L_2\}$ Examples: $\{ab, aab, baba\}^R = \{ba, baa, abab\}$

Concatenation

Example: $\{a,ab,ba\}\{b,aa\}$ $L = \{a^n b^n : n \ge 0\}$ $= \{ab, aaa, abb, abaa, bab, baaa\}$ $L^R = \{b^n a^n : n \ge 0\}$

Another Operation

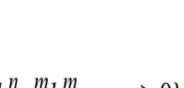
Definition:
$$L^n = LL \cdots L$$

$${a,b}^3 = {a,b}{a,b}{a,b} =$$

$$\lambda$$
}

Special case:
$$L^0 = \{\lambda\}$$

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



More Examples

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

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

 $aabbaaabbb \in L^2$

Definition:
$$L^* = I$$

Definition: $L^* = L^0 \cup L^1 \cup L^2 \cdots$

Star-Closure (Kleene *)

total

Definition: $L^+ = L^1 \cup L^2 \cup \cdots$ = $L^* - \{\lambda\}$

Positive Closure

 $\{a,bb\}^{+} = \begin{cases} a,bb, \\ aa,abb,bba,bbb, \\ aaa,aabb,abba,abbb, \dots \end{cases}$

total union

without zero

gibi