CNG280

Formal Languages

&

Abstract Machines

Languages

A language is a set of strings

String: A sequence of letters

Examples: "cat", "dog", "house", ...

Defined over an alphabet:

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

Alphabets and Strings

We will use small alphabets: $\Sigma = \{a, b\}$

Strings

a	
ab	$u = ab$
$abba$	$v = bbbaaa$
$baba$	$w = abba$
$aaabbbaabab$	

String Operations

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

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

Concatenation

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

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

Reverse

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

String Length

$$w = a_1 a_2 \cdots a_n$$

Length: $|w| = n$

Examples:

$$|abba| = 4$$

$$|aa| = 2$$

$$|a| = 1$$

Length of Concatenation

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

Example: $u = aab, \quad |u| = 3$

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

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

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

Empty String

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

Observations: $|\lambda| = 0$

length of an empty string is 0

$$\lambda w = w\lambda = w$$

$$\lambda abba = abba\lambda = abba$$

Substring

we use these to matching, just like how do we understand that there is if, we look at the substring if.

Substring of string:

a subsequence of consecutive characters

String	Substring
$\underline{ab}bab$	ab
$\underline{abba}b$	$abba$
$a\underline{b}bab$	b
$a\underline{bbab}$	$bbab$

Prefix and Suffix

$abbab$

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

$$w = uv$$

prefix

suffix

You take all the possibilities bc we don't know where to divide the string

Another Operation

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

Example: $(abba)^2 = abbaabba$

Definition: $w^0 = \lambda$

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

The * Operation

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

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

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

means zero or more

The + Operation

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

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

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

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

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

means one or more

Languages

A language is any subset of Σ^*

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

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

Languages: $\{\lambda\}$

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

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

Note that:

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

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

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

String length $|\lambda| = 0$

Another Example

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

$$\left.\begin{matrix} \lambda \\ ab \\ aabb \\ aaaaabbbbb \end{matrix}\right\} \in L \qquad abb \notin L$$

Operations on Languages

The usual set operations

union is

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

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

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

doesn't always be like you imagine control the results

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

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

Reverse

hoca will not ask any proof if we didn't go through it in the class

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

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

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

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

Concatenation

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

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

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

More Examples

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

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

$$aabbaaabbb \in L^2$$

Star-Closure (Kleene *)

total union gibi

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

Example:

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

Positive Closure

total union without zero gibi

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

$$= L^* - \{\lambda\}$$

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