Lecture – 8 4/5/20 (1)

Language L denotes a set of strings over certain alphabet.

- ϕ and $\{\epsilon\}$ are languages

  L without a string → empty box inside a box.

examples

$L_1 = \{ab, aabb, aaabbb\}$ is a language over alphabet $\{a, b\}$.

$L_2 = \{\epsilon, a, aa, aaa, \ldots\}$ is a language over alphabet $\{a\}$

$L_3 = \{a^n b^n c^n : n > 1\}$ is a language over $\{a, b, c\}$.

- An **alphabet** is a finite and non-empty set of symbols denoted by Σ.

  e.g $\Sigma = \{0, 1\}$ denotes machine language

- A string is a finite sequence of symbols

Useful Operations on Languages

- Languages are represented by sets
- so all operations applicable on sets are also applicable on the languages.

e.g; Union, Concatenation, Intersection, Complement

$L = L_1 \cup L_2$ $L = L_1 L_2$ $L = L_1 \cap L_2$ $L = \Sigma^* - L_1$

Formal languages (FL)

- is an abstraction of the general characteristics of programming languages
- FL consists of a set of symbols and some rules of formation by which these symbols can be combined into entities called sentences.
- ∴ FL is the set of all sentences permitted by the rules of the formation. (called productions)

Grammar

- To study languages mathematically, we need a mechanism to describe them
- e.g, a grammar for English language can tell us whether 'ď' is well formed or not

Formal Grammar

- The set of rules to form a sentence which in turn ~~are~~ used to generate a language is called a grammar

~~A Grammar G is a 4-tuple~~ (3)

$G = (V, T, P, S)$ where

$V \Rightarrow$ finite set of objects called variables / non-terminals

$T \Rightarrow$ finite set of objects called terminal symbols

$S \in V$ is a special symbol called start symbol

P is a finite set of productions

(Assumption set V and T are non-empty and disjoint).

(a) Each production is of the form $\alpha \to \beta$, where $\underline{\alpha}$ is a non-empty string of terminals and or non-terminals, and β is a string of nonterminals including the null-string.

α is a string $(V \cup T)^+$

β is a string in $(V \cup T)^*$

(b) production rules are the heart of a grammar,

- They specify how the grammar transforms one string into another, thereby defining a language associated with the grammar.

Ex

$S \to aCa$
$C \to aCa \mid b$

$G = (V, T, P, S)$ where

$V = \{S, C\}$

$T = \{a, b\}$

$P \to \{$
$S \to aCa$
$C \to aCa \mid b$
$\}$

S is the start symbol

~~[illegible]~~

$$L(G) = \{ w \mid S \overset{*}{\Rightarrow} w \text{ and } w \in T^* \}$$

Q) $G = (V, T, P, S)$
$V = \{S, C\}$
$T = \{a, b\}$

Productions

$S \to aCa$ — 1
$C \to aCa$ — 2
$C \to b$ — 3

consider the derivation

$S \Rightarrow aCa \to$ 1
$\Rightarrow aaCaa \to$ 2
$\to aaaCaaa \to$ 2
$\to aaabaaa \to$ 3
$= a$

apply 2 n times
$a^n C a^n$
$= a^n b a^n$

$\therefore L(G) = \{a^n b a^n \mid n >$

~~Obtain a grammar to generate integers~~ 5

$+, -, \epsilon$

$S \to + \mid - \mid \epsilon$

$D \to 0 \mid 1 \mid 2 \mid 3 \ldots \mid 9.$

+ 1
- 12
13

⇒ A number can be recursively defined as follows :→

1. A digit is a number ($N \to D$)
2. The number followed by a digit is a number ($N \to ND$)

or

digit followed by a number is also a number

($N \to DN$)

$N \to D$

$N \to ND \mid DN.$

Let I be an integer (number)

$I \to N \mid SN$

$G = (V, T, P, S)$

$V = \{D, S, N, I\}$

$T = \{+, -, 0, 1, 2, \ldots 9\}$

$P = \{ \quad I \to N \mid SN$

$N \to D \mid ND \mid DN$

$S \to + \mid - \mid \epsilon$

$D \to 0 \mid 1 \mid 2 \mid 3 \mid \ldots 9$

$\}$

$S = I.$

1965 / +1965

$\Rightarrow N$	$I \Rightarrow SN$
$\Rightarrow ND$	$\Rightarrow +N$
$\Rightarrow N5$	$\Rightarrow +ND$
$\Rightarrow ND5$	$\Rightarrow +N5$
$\Rightarrow N65$	$\Rightarrow +ND5$
$\Rightarrow ND65$	$\Rightarrow +N65$
$\Rightarrow N965$	$\Rightarrow +ND65$
$\Rightarrow D965$	$\Rightarrow +N965$
$\Rightarrow 1965$	$\Rightarrow +\cancel{D}1965$
	$\Rightarrow +1965$

Ex $\Sigma = \{a, b\}$ Generate a set of all palindromes.

Sol

Recursive definition of a palindrome is

1. ϵ is a palindrome.
2. a and b are palindromes.
3. If w is a palindrome, then the strings awa and bwb are also palindromes

Form productions for 1, 2, 3

$S \to \epsilon \qquad S \to a|b \qquad S \to aSa|bSb.$

$V = \{S\}$

$T = \{a, b\}$

$P \to \{ S \to \epsilon$ — ①

$S \to a|b$ ②

$S \to aSa|bSb$ — ③

$\}$

S is the start symbol.