Computation, Encoding and Languages

- Computational Problems, Strings and Data Encoding
- Binary Encoding
- Language
- Versions of Computational Problems
- Decision Problems as Language Recognition
- Models of Computation – CPU + Memory

IMDAD ULLAH KHAN

Languages

Language

A language is a subset of strings from some alphabet

A Language on Σ : A set $L \subseteq \Sigma^*$

- $L_1 : \emptyset$ ▷ Empty language
- $L_2 : \Sigma^*$ ▷ All strings
- $L_3 : \Sigma^n = \{x \in \Sigma^* : |x| = n\}$ ▷ All strings of length n
- $L_4 : \Sigma^n = \{x \in \Sigma^* : |x| \leq 4\}$ ▷ All strings of length at most 4

Let $\Sigma = \{0, 1\}$.

1. List all the above languages
2. Find the number of strings in these languages

Repeat the above two tasks with $\Sigma = \{a, b, c\}$

Languages: Examples

A language is a subset of strings from some alphabet

- $L_5 \subseteq \{a, b\}^*$ of strings with an equal number of a's and b's
- $L_6 \subseteq \{a, b\}^*$ of strings that start with 0 or more a's and are followed by an equal number of b's

$$L_6 = \{a^n b^n : n \geq 0\}$$

- $L_7 \subseteq \{a, b\}^*$ of strings containing "$abaa$"

$$L_7 = \{xyz : x, z \in \{a, b\}^*, y = abaa\}$$

- $L_8 \subseteq \{a, b\}^*$ of strings containing at least two a's

$$L_8 = \{xayaz : x, y, z \in \{a, b\}^*\}$$

Operations on Languages

We can make new languages from old languages

Since languages are sets, we can apply set operations to form new languages

- $\overline{L_i} = \{w \in \Sigma^* : w \notin L_i\}$
- $L_i \cup L_j = \{w \in \Sigma^* : w \in L_i \vee w \in L_j\}$
- $L_i \cap L_j = \{w \in \Sigma^* : w \in L_i \wedge w \in L_j\}$

Let $|L_1| = n_1$ and $|L_2| = n_2$

What can we say about the cardinalities of the above languages in terms of n_1 and/or n_2?

Operations on Languages

We can make new languages from old languages

- $L_1 =$ set of strings $w \in \{a, b\}^*$ that have an equal number of a's and b's and have length at least 5
- $L_2 =$ set of strings $w \in \{a, b\}^*$ that have an equal number of a's and b's
- L_3 set of strings $w \in \{a, b\}^*$ that have length at most 4

$$L_1 = L_2 \cap \overline{L_3}$$

Operations on Languages

We can make new languages from old languages

Since languages are sets of strings, we can make new sets by performing string operations on elements of a language

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

$L = \{\epsilon, a, ab, aab, aaab, aaaab\} \implies L^R = \{\epsilon, a, baa, baaa, baaaa\}$

What is $|L^R|$ in terms of $|L|$?

Operations on Languages

We can make new languages from old languages

Since languages are sets of strings, we can make new sets by performing string operations on elements of a language

Concatenation $L_1 \circ L_2 = \{w_1 w_2 : w_1 \in L_1, w_2 \in L_2\}$

$L_1 = \{\epsilon, a, ab, aab\} \quad L_2 = \{b, ba\}$

$\implies L_1 \circ L_2 = \{b, ba, ab, aba, abb, abba, aabb, aabba\}$

Notice

$$|L_1 \circ L_2| \leq |L_1| * |L_2|$$

In what cases will we get $|L_1 \circ L_2| < |L_1| * |L_2|$?

Languages: Examples

- ϵ: The empty string
- $\emptyset$: The empty language
- $\{\epsilon\}$: the language containing the empty string
- $\{\emptyset\}$: A class of languages, (set containing the empty language)