Regular Languages and Regular Sets

CS712
Dept. of Computer Science
Univ. of New Hampshire
A language is a set of strings.

Individual elements that make up a string are chosen from a finite set called the alphabet.

e.g. alphabet \( A = \{a, b\} \)

\[ L_1 = \{a, b\} \]

\[ L_2 = \{a^n b^n, \ n \geq 0\} \]
If $A$ is an alphabet, then $A^*$ is the set of all strings over $A$.

Any language over $A$ is a subset of $A^*$.

Notes:

- the empty string is a string, denoted as $\lambda$
- $\emptyset \subset A^*$ is a language
- the empty set is a language, denoted as $\emptyset$
Concatenation of strings:

\[ (aabb)(bb) = (aabb) \]

Product of languages \( L \) and \( M \):

\[ (L)(M) = \{ s(t) \mid s \in L \text{ and } t \in M \} \]

\( L^0 = \emptyset \)

\( L^n = \{ s_1s_2...s_n \mid s_k \in L \text{ for } 1 \leq k \leq n \} \)

\( L^* = L^0 \cup L^1 \cup L^2 \cup ... \cup L^n \cup \ldots \)

\( L \rightarrow \) the closure of \( L \)
note: $\mathcal{E}\mathcal{L}\mathcal{E}^* = \mathcal{E}\mathcal{L}\mathcal{E}$

$\phi^* = \phi \cup \phi \cup \phi \cup \ldots$

$\mathcal{E}\mathcal{L}\mathcal{E} \cup \phi \cup \phi \cup \phi \cup \ldots$

$\mathcal{E}\mathcal{L}\mathcal{E}$

$(\mathcal{E}\mathcal{L}\mathcal{E})(\mathcal{L}) = \mathcal{L}$  $(\mathcal{L})(\mathcal{E}\mathcal{L}\mathcal{E}) = \mathcal{L}$

$(\phi)(\mathcal{L}) = \phi$  $(\mathcal{L})(\phi) = \phi$

$\phi \cup \mathcal{L} = \mathcal{L}$
The recognition problem: is string $s$ in language $L$?
example

\[ A = 39.63 \]

\[ (\langle 9 \rangle 1 (\langle 9 \rangle \rangle^*) = \langle 1, 9, 6, 9, 69, 69, 69, 69, 69, 69, 69, ... \rangle \]
A is the alphabet.

**Regular Language:**
- \( \emptyset \), \( \{ \epsilon \} \), \( \{ \alpha \} \quad \forall \alpha \in A \)

**Regular Set:**
- \( \emptyset, \lambda, \alpha \quad \forall \alpha \in A \)

**Base Case:**
- If \( L \) and \( M \) are regular languages then the following are regular languages:
  - i) \( L \cup M \)
  - ii) \( (L)(M) \)
  - iii) \( L^* \)

**Recursion:**
- If \( R \) and \( S \) are regular sets then the following are regular sets:
  - i) \( (R) \)
  - ii) \( R \cdot S \)
  - iii) \( RS \)
  - iv) \( R^* \)

i.e.: a regular set is notation for describing a regular language.