Alphabets, Strings, and Languages
Key Concepts

Definition 1. An ***alphabet*** is a finite non-empty set. The elements of an alphabet are often called ***symbols***.

Definition 2. A ***string***, over an alphabet Σ, is a finite sequence of elements of Σ.

Definition 3. The ***length*** of a string ω, over an alphabet Σ, is the same as its length when it is considered as a sequence. The length of a string ω is often written as $|\omega|$.

- The empty string (that is, the string with length zero) will be denoted by λ.

Definition 4. For any alphabet Σ, $\mathbf{\Sigma^\star}$ is the set of *all* strings over Σ.

Definition 5. Suppose that

$$\mu = a_1 a_2 \ldots a_n \qquad \text{and} \qquad \nu = b_1 b_2 \ldots b_m$$

are strings over an alphabet Σ (that is, suppose that $\mu, \nu \in \Sigma^\star$) with lengths n and m, respectively, so that

$$\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_m \in \Sigma.$$

Then the ***concatenation*** of μ and ν, denoted by $\mu \cdot \nu$, is the string

$$\alpha_1 \alpha_2 \ldots \alpha_n \beta_1 \beta_2 \ldots \beta_m$$

over Σ, with length $n + m$, obtained by listing the symbols in ν after the symbols in μ.

Definition 6. If $\mu, \omega \in \Sigma^\star$, for an alphabet Σ, then μ is a ***substring*** of ω if there exist strings $\nu, \varphi \in \Sigma^\star$ such that $\omega = \nu \cdot \mu \cdot \varphi$.

Definition 7. Once again, suppose that $\mu, \omega \in \Sigma^\star$, for an alphabet Σ. Then μ is a ***prefix*** of ω if there exists a string $\nu \in \Sigma^\star$ such that

$$\omega = \mu \cdot \nu,$$

and μ is a ***suffix*** of ω if there exists a string $\nu \in \Sigma^\star$ such that

$$\omega = \nu \cdot \mu.$$

Definition 8. A ***language*** over an alphabet Σ is a subset of $\Sigma^\star$.