Regular Expressions

A *regular expression* is an expression that consists of sets and regular operations, which are used for pattern recognition and can be used to represent regular languages. Specifically, a regular expression over Σ is defined as

1. $\mathtt{a} \in \Sigma$, ε, and $\emptyset$ are regular expressions.
2. If P and Q are regular expressions, then $P \cup Q$, $P \circ Q$ and P^* are regular expressions.

and it has the following properties and shorthands

- The order of precedence of its operations in descending order is parentheses, star, concatenation, and union.
- The regular expression $\mathtt{a}$ where $\mathtt{a} \in \Sigma$ is the shorthand for $\{\mathtt{a}\}$.
- The regular expresion AB is the shorthand for $A \circ B$ where A and B are regular expressions.
- The regular expression Σ is shorthand for the union of all the symbols in the alphabet.
- $\emptyset^* = \varepsilon$.
- $R \circ \emptyset = \emptyset$ where R is a regular expression.
- $R \circ \varepsilon = R$ where R is a regular expression.
- The regular expression $R^+ = R \circ R^*$ where R is a regular expression and $+$ is the derived operator.
- The regular expression $R^? = R \cup \varepsilon$ where R is a regular expression and ? is the derived operator.

Example:
Some examples of regular expressions and their meaning are:

$$\begin{aligned}
\mathtt{0}^*\mathtt{10}^* &= \{w : w \texttt{ has exactly one } \mathtt{1}\} \\
\Sigma^*\mathtt{1}\Sigma^* &= \{w : w \texttt{ has at least one } \mathtt{1}\} \\
\Sigma\Sigma\Sigma &= \{w : |w| = 3\} \\
\Sigma^*\mathtt{010}\Sigma^* &= \{w : w \texttt{ contains the substring } \mathtt{010}\} \\
\Sigma(\Sigma\Sigma)^* &= \{w : w \texttt{ contains an odd number of symbols}\}
\end{aligned}$$