$$\Sigma = \overbrace{\{\underline{a}, \underline{b}, \underline{c}\}}^{\text{alphabet}}$$

symbols

string:
sequence of symbols

Examples of strings over Σ:

abc a ε b ccccc ~~abcd~~ not over Σ because d is not in Σ

A string is **over** Σ if all of its symbols are in Σ

A **language** over Σ is a set of strings over Σ

Examples of languages over Σ:

$b^2 = bb$ $b^3 = bbb$

$\{aa, bb, cc\}$ $\{abc, a, \varepsilon, b, ccccc\}$ $\{b^n \mid n \geq 0\}$

$b^0 = \varepsilon$

We can use **regular expressions** to describe these languages.

Regular Expressions

Basis steps:

a is a regular expression, for $a \in \Sigma$

ε is a regular expression

$\emptyset$ is a regular expression

regular expressions

string in the language

$L(a) = \{a\}$

$L(\varepsilon) = \{\varepsilon\}$

$L(\boxed{\emptyset}) = \cancel{\{\emptyset\}}$ $\emptyset$

the empty set → represent language with 0 strings in it

Recursive steps:

(R_1^*) is a regular expression when R_1 is a regular expression

$(R_1 \circ R_2)$ is a regular expression when R_1 and R_2 are regular expressions

$(R_1 \cup R_2)$ is a regular expression when R_1 and R_2 are regular expressions

$\Sigma = \{a, b, c\}$

Regular Expressions

(R_1^*) is a regular expression when R_1 is a regular expression

b^* $L(b) = \{b\}$
$L(b^*) = \{b^n \mid n \geq 0\}$
ε b bb bbb ...

Σ^* $L(\Sigma^*) =$ the set of strings over Σ

$(R_1 \circ R_2)$ is a regular expression when R_1 and R_2 are regular expressions

$a \circ a \circ a \circ a \circ a$
(shorthand: aaaaa)
$L(aaaaa) = \{aaaaa\}$

$c \circ b^*$
(shorthand: cb^*)
$L(cb^*) = \{cb^n \mid n \geq 0\}$
c cb cbb cbbb ...

$(R_1 \cup R_2)$ is a regular expression when R_1 and R_2 are regular expressions

$a \cup c$
$L(a \cup c)$
$= L(a) \cup L(c)$
$= \{a, c\}$

$cb^* \cup aaaaa$
$L(cb^* \cup aaaaa)$
$= L(cb^*) \cup L(aaaaa)$
$= \{cb^n \mid n \geq 0\} \cup \{aaaaa\}$

$\Sigma = \{a, b, c\}$

Regular Expressions

$cb^* \cup aaaaa$ $\quad L(cb^* \cup aaaaa) = \{cb^n \mid n \geq 0\} \cup \{aaaaa\}$

cb cbbbbb aaaaa

Implictly:

$(c(b^*)) \cup (aaaaa)$

What happens when we evaluate in a different order?

$\boxed{(cb)^*} \cup (aaaaa)$ $\quad L((cb)^*) = \{(cb)^n \mid n \geq 0\}$

$\boxed{cbcbcbcb \quad \varepsilon}$

examples of strings

→ The language is not the same!

Precedence order: First $*$, then $\circ$, then $\cup$