

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

symbols

string:  
sequence of symbols

Examples of strings over  $\Sigma$ :

abc   a    $\epsilon$    b   ccccc

~~a~~ not over  $\Sigma$   
because d is  
not in  $\Sigma$

A string is over  $\Sigma$  if all of its symbols are in  $\Sigma$

A language over  $\Sigma$  is a set of strings over  $\Sigma$

Examples of languages over  $\Sigma$ :

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

$$b^2 = bb \quad b^3 = bbb \\ b^0 = \epsilon$$

We can use regular expressions to describe these languages.



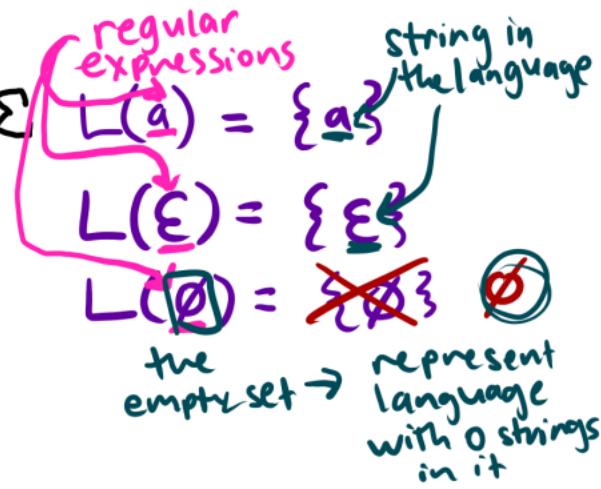
# Regular Expressions

Basis steps:

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

$\epsilon$  is a regular expression

$\emptyset$  is a regular expression



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

$b^*$   $L(b) = \{b\}$

$$L(b^*) = \{b^n \mid n \geq 0\}$$

$$\in b \quad bb \quad bbb \dots$$

$$\Sigma^* \quad L(\Sigma^*) =$$

the set of strings  
over  $\Sigma$

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

$$a \circ a \circ a \circ a \circ a \quad L(aaaa) =$$

(shorthand: aaaaa)  $\{aaaaaa\}$

$$c \circ b^* \quad L(cb^*)$$

(shorthand: cb<sup>\*</sup>)  $= \{cb^n \mid n \geq 0\}$

$$c \quad cb \quad cbb \quad cbbb \dots$$

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

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

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

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

# Regular Expressions

$cb^* \cup \text{aaaaaa}$

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

Implicitly:

$$(c(b^*)) \cup (\text{aaaaaa})$$

cb cbcbcbcb aaaaa

What happens when we evaluate in a different order?

$$(cb)^* \cup (\text{aaaaaa})$$

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

cbcbcbcbcb  $\epsilon$

examples of strings

→ The language is not the same!

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