

# (Τυπικές) Γλώσσες: Ορισμοί

---

Διδάσκοντες: **Φ. Αφράτη, Σ. Ζάχος, Δ. Σούλιου**

Επιμέλεια διαφανειών: **Δ. Φωτάκης**

Σχολή Ηλεκτρολόγων Μηχανικών  
και Μηχανικών Υπολογιστών

Εθνικό Μετσόβιο Πολυτεχνείο



# Αλφάβητα και Συμβολοσειρές

- **Αλφάβητο:** πεπερασμένο μη-κενό σύνολο  $\Sigma$ .
  - $\Sigma = \{0, 1\}$ ,  $\Sigma = \{a, b, \gamma, \dots, \omega\}$
- **Συμβολοσειρά:** πεπερασμένη ακολουθία συμβόλων του  $\Sigma$ .
  - Μήκος συμβολοσειράς: #συμβόλων.
  - Κενή συμβολοσειρά  $\epsilon$  ή  $\varepsilon$ : (μοναδική) συμβ/ρά μήκους 0.
- $\Sigma^k$ : σύνολο συμβολοσειρών  $\Sigma$  μήκους  $k$ .
  - $\{0, 1\}^2 = \{00, 01, 10, 11\}$
- $\Sigma^*$ : σύνολο όλων των συμβολοσειρών  $\Sigma$ .
  - $\{0, 1\}^* = \{\epsilon, 0, 1, 00, 01, 10, 11, 000, \dots\}$ .
- Το  $\Sigma^*$  είναι αριθμησιμο.
  - Για  $k = 0, 1, 2, \dots$ , λεξικογραφική διάταξη του  $\Sigma^k$ .

$$\Sigma^* = \bigcup_{k \in \mathbb{N}} \Sigma^k$$

# Πράξεις με Συμβολοσειρές

- **Παράθεση**  $x$  και  $y$ :  $x \circ y$  ή  $xy$ .
  - Π.χ. παράθεση των  $αβγδ$  και  $εζηθ$  είναι  $αβγδεζηθ$ .
- **Επανάληψη** (ή παράθεση) του  $w$  για  $k$  φορές:  $w^k = \overbrace{w \circ \dots \circ w}^{k \text{ φορές}}$ 
  - Π.χ.  $(αβ)^3 = αβαβαβ$ ,  $(010)^4 = 010010010010$
- **Αντίστροφη**  $w^R$  συμβολοσειράς  $w$ : συμβολοσειρά που προκύπτει διαβάζοντας  $w$  από «τέλος» προς «αρχή» (αντίστροφα).
  - $(αβγδ)^R = δγβα$ ,  $(10010)^R = 01001$ ,  $(αβγδγβα)^R = αβγδγβα$ .
  - **Παλινδρομική** (καρκινική) συμβολοσειρά  $w$  αν  $w = w^R$ .
  - Νδο (επαγωγικά)  $\forall x, y \in \Sigma^*$ ,  $(x \circ y)^R = y^R \circ x^R$
- $x$  **υποσυμβολοσειρά**  $w$  αν  $\exists y, z \in \Sigma^*$  :  $w = yxz$ .
  - $x$  **πρόθεμα**  $w$  αν  $y = \epsilon$ .  $x$  **κατάληξη**  $w$  αν  $z = \epsilon$ .

# (Τυπικές) Γλώσσες

---

- (Τυπική) γλώσσα  $L$ : οποιοδήποτε υποσύνολο  $\Sigma^*$ .
  - Αριθμήσιμο σύνολο συμβολοσειρών.
- Παραδείγματα γλωσσών στο  $\Sigma = \{0, 1, 2\}$ :
  - $L_1 = \{012, 021, 120, 102, 210, 201\}$
  - $L_2 = \{\varepsilon, 1, 11, 111, 1111, \dots\}$
  - $L_3 = \{w \in \Sigma^* : \text{το τελευταίο ψηφίο του } w \text{ είναι } 0\}$
  - $L_4 = \{\} = \emptyset$
  - $L_5 = \{\varepsilon\}$
- Σύνολο όλων των γλωσσών  $2^{\Sigma^*}$  μη αριθμήσιμο (διαγωνιοποίηση).

# Πράξεις με Γλώσσες

---

- Πράξεις συνόλων:
  - Ένωση, τομή, διαφορά, συμπλήρωμα.
- Παράθεση γλωσσών  $L_1$  και  $L_2$ :  $L_1 \circ L_2$  ή  $L_1 L_2$   
 $L_1 \circ L_2 = \{w \in \Sigma^* : w = x_1 \circ x_2 \text{ για κάποια } x_1 \in L_1 \text{ και } x_2 \in L_2\}$ 
  - $L_1 = \{0, 00, 11\}$  και  $L_2 = \{\varepsilon, 1\}$ , τότε  
 $L_1 L_2 = \{0, 00, 11, 01, 001, 111\}$
- $L^2 = L \circ L$ : παράθεση  $L$  με τον εαυτό της.
- $L^k = \overbrace{L \circ \dots \circ L}^{k \text{ φορές}}$ : παράθεση  $L$  με τον εαυτό της  $k$  φορές.

# Πράξεις με Γλώσσες

- **Kleene star**  $L^*$  γλώσσας  $L$ : αποτελείται από παράθεση οποιουδήποτε (πεπερασμένου) αριθμού συμβολοσειρών της  $L$ .  
$$L^* = \{w \in \Sigma^* : w = x_1 \circ \dots \circ x_k \text{ για } k \geq 0, x_1, \dots, x_k \in L\}$$
$$= \bigcup_{k \in \mathbb{N}} L^k$$
  - Αν  $L = \{1, 00\}$ , τότε  
 $L^* = \{\varepsilon, 1, 00, 11, 100, 001, 0000, 1100, 10000, 1001, 00100, \dots\}$
  - Αν  $L = \{\varepsilon\}$ , τότε  $L^* = \{\varepsilon\}$ .
  - Αν  $L = \emptyset$ , τότε  $L^* = \{\varepsilon\}$  (για κάθε  $L$ ,  $\varepsilon \in L^*$ ).
- $L^+$  : αποτελείται από παράθεση **θετικού** αριθμού συμβ/ρών της  $L$ .  
$$L^+ = \{w \in \Sigma^* : w = x_1 \circ \dots \circ x_k \text{ για } k \geq 1, x_1, \dots, x_k \in L\}$$
$$= \bigcup_{k \geq 1} L^k$$
  - Πότε οι  $L^*$  και  $L^+$  είναι διαφορετικές; Ποια είναι η διαφορά τους;