

Γενικά περί γραμματικών

Γραμματική: Γραμματική (γενικά) είναι μια τετράδα $G(V, \Sigma, R, S)$ όπου

- V είναι αλφάβητο,
- Σ είναι επίσης αλφάβητο με $\Sigma \subseteq V$,
- R είναι το σύνολο κανόνων, $R \subseteq V^* \times V^*$, και
- S είναι το αρχικό σύμβολο, $S \in V - \Sigma$.

Το Σ λέγεται σύνολο **τερματικών συμβόλων**, και το $V - \Sigma$ λέγεται σύνολο μη τερματικών συμβόλων (ή μεταβλητών).

Παράδειγμα γραμματικής:

- $V = \{0, 1, S, A\},$
- $\Sigma = \{0, 1\},$
- $R = \left\{ \begin{array}{l} S \rightarrow 0A1, \\ A1 \rightarrow 0A11, \\ 0A \rightarrow 00A1, \\ A \rightarrow \varepsilon \end{array} \right\}$

Γραμματικές χωρίς συμφραζόμενα (context-free)

Θα μελετήσουμε την ειδική κατηγορία των γραμματικών, τις **context-free** (χωρίς συμφραζόμενα) γραμματικές. Οι κανόνες των γραμματικών χωρίς συμφραζόμενα είναι της μορφής $(V - \Sigma) \times V^*$ (ενώ των γενικών γραμματικών είναι της μορφής $V^* \times V^*$).

Παράδειγμα γραμματικής χωρίς συμφραζόμενα:

- $V = \{0, 1, S\},$
- $\Sigma = \{0, 1\},$
- $R = \left\{ \begin{array}{l} S \rightarrow 0S1, \\ S \rightarrow \varepsilon \end{array} \right\}$

Γλώσσες χωρίς συμφραζόμενα

Έστω $u, v, w \in V^*$ για μια γραμματική G .

Αν $A \rightarrow w$ είναι κανόνας της G τότε λέμε ότι η uAv παράγει uwv και το συμβολίζουμε με

$$uAv \Rightarrow uwv.$$

Γλώσσες χωρίς συμφραζόμενα

Έστω $u, v, w \in V^*$ για μια γραμματική G .

Αν $A \rightarrow w$ είναι κανόνας της G τότε λέμε ότι η uAv παράγει uwv και το συμβολίζουμε με

$$uAv \Rightarrow uwv.$$

Ορίζουμε επίσης τη σχέση u παράγει v σε μηδέν ή περισσότερα βήματα και το συμβολίζουμε $u \Rightarrow^* v$, αν υπάρχουν $u_1, \dots, u_k$ όπου $k \geq 0$ τέτοια ώστε

$$u \Rightarrow u_1 \Rightarrow \dots \Rightarrow u_k \Rightarrow v.$$

Γλώσσες χωρίς συμφραζόμενα

Έστω $u, v, w \in V^*$ για μια γραμματική G .

Αν $A \rightarrow w$ είναι κανόνας της G τότε λέμε ότι η uAv παράγει uwv και το συμβολίζουμε με

$$uAv \Rightarrow uwv.$$

Ορίζουμε επίσης τη σχέση u παράγει v σε μηδέν ή περισσότερα βήματα και το συμβολίζουμε $u \Rightarrow^* v$, αν υπάρχουν $u_1, \dots, u_k$ όπου $k \geq 0$ τέτοια ώστε

$$u \Rightarrow u_1 \Rightarrow \dots \Rightarrow u_k \Rightarrow v.$$

Ορίζουμε τη γλώσσα $L(G)$ μιας γραμματικής $G(V, \Sigma, R, S)$ σαν το σύνολο των συμβολοσειρών του Σ^* που παράγονται από το αρχικό σύμβολο S :

$$L(G) = \{w : w \in \Sigma^* \text{ και } S \Rightarrow^* w\}.$$

Οι γλώσσες χωρίς συμφραζόμενα είναι οι γλώσσες κάποιας γραμματικής χωρίς συμφραζόμενα.

Παράδειγμα

Η γλώσσα της γραμματικής

- $V = \{0, 1, S\},$
- $\Sigma = \{0, 1\},$
- $R = \left\{ \begin{array}{l} S \rightarrow 0S1, \\ S \rightarrow \varepsilon \end{array} \right\}$

είναι $L(G) = \{0^n 1^n : n \geq 0\}.$

Συμπέρασμα: Υπάρχουν γλώσσες χωρίς συμφραζόμενα που δεν είναι κανονικές.

Παράδειγμα

Η γλώσσα της γραμματικής

- $V = \{ (,), S \},$
- $\Sigma = \{ (,) \},$
- $R = \left\{ \begin{array}{l} S \rightarrow SS, \\ S \rightarrow (S), \\ S \rightarrow \varepsilon \end{array} \right\}$

είναι $L(G) = \{ w : w \text{ αποτελείται από σωστά ζυγισμένες παρενθέσεις} \}.$

Παράδειγμα

Βρείτε μια γραμματική G χωρίς συμφραζόμενα με γλώσσα

$$L(G) = \{w : w \text{ είναι αριθμητική έκφραση με σύμβολα } x, y, +, *, (,)\}.$$

Απάντηση (F =factor, T =term, E =expression):

- $V = \{x, y, +, *, (,), S, F, T, E\},$

- $\Sigma = \{x, y, +, *, (,)\},$

- $R = \left\{ \begin{array}{l} F \rightarrow x, \\ F \rightarrow y, \\ F \rightarrow (E), \\ T \rightarrow F, \\ T \rightarrow F * T, \\ E \rightarrow T, \\ E \rightarrow T + E, \\ S \rightarrow E \end{array} \right\}$

Παραδειγμα

Να βρείτε γραμματική χωρίς συμφραζόμενα για τη γλώσσα
 $\mathcal{L} = \{ww^R : w \in \{0,1\}^*\}$. (Η γλώσσα των παλινδρόμων άρτιου μήκους).

$$S \longrightarrow \varepsilon$$

$$S \longrightarrow 0S0$$

$$S \longrightarrow 1S1$$

Για συντομία μερικές φορές γράφουμε

$$S \longrightarrow \varepsilon|0S0|1S1$$

Κανονικές γλώσσες και γραμματικές

Όπως είδαμε υπάρχουν γλώσσες χωρίς συμφραζόμενα που δεν είναι κανονικές. Τώρα θα δείξουμε ότι το σύνολο των γλωσσών χωρίς συμφραζόμενα περιέχει τις κανονικές γλώσσες. Δηλαδή

$$\text{KANON. ΓΛΩΣΣΕΣ} \subset \text{ΓΛΩΣΣΕΣ ΧΩΡΙΣ ΣΥΜΦΡ.}$$

Κανονική γραμματική: Μια γραμματική είναι κανονική αν οι κανόνες της έχουν τη μορφή $(V - \Sigma) \times \Sigma^*((V - \Sigma) \cup \varepsilon)$.

Παράδειγμα:

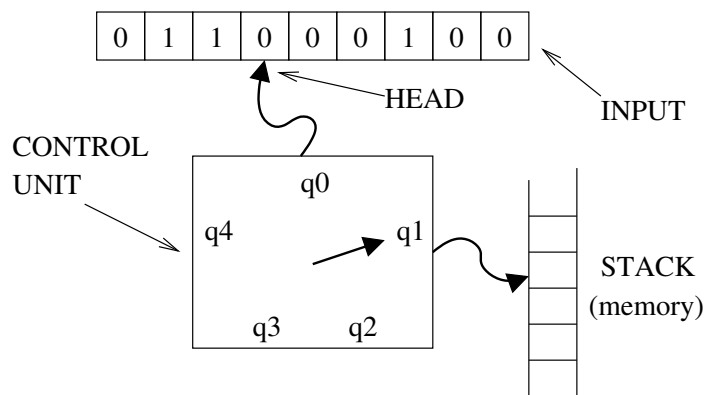
- $V = \{0, 1, S\},$
- $\Sigma = \{0, 1\},$
- $R = \left\{ \begin{array}{l} S \rightarrow 01S, \\ S \rightarrow 010S, \\ S \rightarrow \varepsilon \end{array} \right\}$

Η γλώσσα αυτής της γραμματικής είναι: $L(G) = (01 \cup 010)^*$.

Θεώρημα 1 Μια γλώσσα είναι κανονική αν και μόνο αν είναι γλώσσα μιας κανονικής γραμματικής.

Η απόδειξη είναι εύκολη. Για κάθε πεπερασμένο αυτόματο μπορούμε να κατασκευάσουμε μια κανονική γραμματική όπου το σύνολο των μη τερματικών συμβόλων είναι οι καταστάσεις του αυτομάτου. Και το αντίστροφο, για κάθε κανονική γραμματική μπορούμε να κατασκευάσουμε ένα (μη ντετερμινιστικό) πεπερασμένο αυτόματο, όπου οι καταστάσεις είναι το σύνολο των μη τερματικών συμβόλων της γραμματικής.

Αυτόματα Στοιβάς (Pushdown Automata)



Αυτόματο στοιβάς $M(K, \Sigma, \Gamma, \Delta, s, F)$

K	=	σύνολο καταστάσεων (πεπερασμένο)	} Όπως στα πεπερασμέ- να αυτόμα- τα
Σ	=	αλφάβητο εισόδου	
s	=	αρχική κατάσταση $s \in K$	
F	=	σύνολο τελικών καταστάσεων, $F \subseteq K$	
Γ	=	αλφάβητο στοιβάς	
Δ	=	σχέση μετάβασης, $\Delta \subseteq (K \times \Sigma^* \times \Gamma^*) \times (K \times \Gamma^*)$	

$((p, u, \beta), (q, \gamma)) \in \Delta$: Το M είναι στην κατάσταση p , η κορυφή της στοιβάς είναι β , διαβάζει από την είσοδο u και μεταβαίνει στην κατάσταση q αντικαθιστώντας το β με το γ .

Ένα αυτόματο στοίβας M **δέχεται** τη συμβολοσειρά w αν και μόνο αν, ξεκινώντας από την αρχική κατάσταση με **άδεια στοίβα**, επεξεργαστεί το w και καταλήξει σε κάποια τελική κατάσταση του F με **άδεια στοίβα πάλι**.

Δηλαδή

$$(S, w, \varepsilon) \vdash_M^* (q, \varepsilon, \varepsilon), \text{ με } q \in F$$

Η γλώσσα ενός αυτομάτου στοίβας είναι το σύνολο των συμβολοσειρών που δέχεται το αυτόματο.

$$\mathcal{L}(M) = \{w : M \text{ δέχεται } w\}$$

Παράδειγμα

Εστω το αυτοματο στοιβας με σχεση μεταβασης

$$\Delta = \{ \begin{array}{l} ((s, 0, \varepsilon), (s, 0)) \\ ((s, 1, \varepsilon), (s, 1)) \\ ((s, \#, \varepsilon), (f, \varepsilon)) \\ ((f, 0, 0), (f, \varepsilon)) \\ ((f, 1, 1), (f, \varepsilon)) \end{array} \} \quad \text{και } F = \{f\}$$

Το αυτοματο δεχεται τις συμβολοσειρες που ειναι της μορφης $w\#w^R$. Η γλωσσα δηλαδη του αυτοματου ειναι $L = \{w\#w^R : w \in \{0, 1\}^*\}$.

Παράδειγμα

$$\mathcal{L} = \{w : w \text{ έχει ζυγισμένες αγκύλες} \}$$

$$S \longrightarrow \varepsilon \mid [S] \mid SS$$

$$(s, [, \varepsilon)(s, [)$$

$$(s,], []) (s, \varepsilon)$$

Κατάσταση Εισοδος Στοιβά

s	$[[]] []$	ε
s	$[]] []$	$[$
s	$]] []$	$[[$
s	$] []$	$[$
s	$[]$	ε
s	$] $	$[$
s	ε	ε

Γραμματικές χωρίς συμφραζόμενα και αυτόματα στοίβας

Θεώρημα 2 Οι γλώσσες των αυτομάτων στοίβας είναι ακριβώς οι γλώσσες χωρίς συμφραζόμενα.

Το θεώρημα προκύπτει από τα Θεωρήματα 3 και 4:

Θεώρημα 3 Για κάθε γραμματική χωρίς συμφραζόμενα G υπάρχει αυτόματο στοίβας M , τέτοιο ώστε $\mathcal{L}(M) = \mathcal{L}(G)$.

Θεώρημα 4 Για κάθε αυτόματο στοίβας M υπάρχει γραμματική G χωρίς συμφραζόμενα, τέτοια ώστε $\mathcal{L}(G) = \mathcal{L}(M)$.