Γλώσσες Ανεξάρτητες Συμφραζομένων

Γραμματικές Ανεξάρτητες Συμφραζομένων,
Αυτόματα Στοίβας,
Γλώσσες Ανεξάρτητες Συμφραζομένων,
Γλώσσες που δεν είναι Ανεξάρτητες Συμφραζομένων

Γραμματικές Ανεξάρτητες Συμφραζομένων

Κανονικές Γλώσσες

→ **Κανονικές Εκφράσεις, Πεπερασμένα Αυτόματα**
βρίσκουν εφαρμογές στη **Λεκτική Ανάλυση**, …

Γλώσσες Ανεξάρτητες Συμφραζομένων

→ **Γραμματικές Ανεξάρτητες Συμφραζ., Αυτόματα Στοίβας**
βρίσκουν εφαρμογές στη **Συντακτική Ανάλυση**, …

***Παράδειγμα*:** Γραμ. Α. Σ. για τη γλώσσα με Κ.Ε. $\mathbf{a\ (a^* \cup b^*)\ b}$

$$S \to \mathrm{a}\, M\, \mathrm{b}$$
$$M \to A \mid B$$
$$A \to \mathrm{a}\, A \mid e$$
$$B \to \mathrm{b}\, B \mid e$$

a, b: τερματικά σύμβολα

S, M, A, B: μεταβλητές

Γραμματικές Ανεξάρτητες Συμφραζομένων

Μια **γραμματική ανεξάρτητη συμφραζομένων** (***context-free grammar***) είναι μια τετράδα (V, Σ, R, S) όπου

[LP]: V σύνολο μεταβλητών **και τερματικών συμβόλων**

- V: ένα **πεπερασμένο** σύνολο **μεταβλητών**
- Σ: **πεπερασμένο** σύνολο **τερματικών συμβόλων** $(V \cap \Sigma = \emptyset)$
- R: ένα **πεπερασμένο** σύνολο **κανόνων** της μορφής $A \rightarrow w$
  όπου $A \in V$, $w \in (V \cup \Sigma)^*$
- S: **εναρκτήρια μεταβλητή** ή **αρχικό σύμβολο** $(S \in V)$

Γραμματικές Ανεξάρτητες Συμφραζομένων

Παραγωγές

- Εάν στη Γρ.Α.Σ. G έχουμε $A \rightarrow u$ τότε $xAy \Rightarrow xuy$

  Λέμε: το xAy **παράγει (αποδίδει) σε 1 βήμα** το xuy

- Εάν $u_0 \Rightarrow u_1 \Rightarrow \ldots \Rightarrow u_m$ ($m \geq 1$ βήματα εφαρμογής κανόνων)

  ή $u_0 = u_m$ (0 βήματα εφαρμογής κανόνων)

  τότε $u_0 \Rightarrow^* u_m$ *Λέμε*: το u_0 **παράγει** το u_m

  $\Rightarrow^*$: ανακλαστική και μεταβατική κλειστότητα της σχέσης $\Rightarrow$

Η γλώσσα $L(G)$ που **παράγεται** από μια Γρ.Α.Σ. G είναι το σύνολο των συμβολοσειρών που **παράγονται από την αρχική μεταβλητή** της G.

Μια γλώσσα είναι **ανεξάρτητη συμφραζομένων** εάν υπάρχει μια γραμματική Α.Σ. που την παράγει.

Γραμματικές Ανεξάρτητες Συμφραζομένων

Παραδείγματα

$\{a^n b^n \mid n \geq 0\}$: $(\{S\}, \{a,b\}, \{S \rightarrow aSb, S \rightarrow e\}, S)$

$$S \Rightarrow aSb \Rightarrow aaSbb \Rightarrow aaaSbbb \Rightarrow aaaebbb = aaabbb$$

$G = (\,$**Μεταβλητές** $= \{S, P, A, N, V\}, \Sigma, R, S)$ όπου

$\Sigma = \{$ **Jim, big, green, cheese, ate** $\}$

$R = \{\ S \rightarrow P\ V\ P$

$P \rightarrow A\ P \mid N$

$A \rightarrow$ **big** | **green**

$N \rightarrow$ **Jim** | **cheese**

$V \rightarrow$ **ate** $\}$

Γραμματικές Ανεξάρτητες Συμφραζομένων

Παραδείγματα

$G = ($ **Μεταβλητές** $= \{ E, T, F \}, \ \Sigma, \ R, \ E)$ όπου

$$\Sigma = \{ +, \times, (,), \mathbf{id} \}$$

$$R = \{ \begin{array}{lll} E \rightarrow E + T & \mid & T \\ T \rightarrow T \times F & \mid & F \\ F \rightarrow (E) & \mid & \mathbf{id} \quad \} \end{array}$$

Αριστερότερη / Δεξιότερη Παραγωγή

Αριστερότερη Παραγωγή

κάθε φορά αντικαθιστούμε την **αριστερότερη** μεταβλητή

$$E \Rightarrow E + T \Rightarrow T + T \Rightarrow F + T \Rightarrow \mathbf{id} + T \Rightarrow \mathbf{id} + T \times F$$
$$\Rightarrow \mathbf{id} + F \times F \Rightarrow \mathbf{id} + (E) \times F \Rightarrow \mathbf{id} + (T) \times F$$
$$\Rightarrow \mathbf{id} + (F) \times F \Rightarrow \mathbf{id} + (\mathbf{id}) \times F \Rightarrow \mathbf{id} + (\mathbf{id}) \times \mathbf{id}$$

Δεξιότερη Παραγωγή

κάθε φορά αντικαθιστούμε τη **δεξιότερη** μεταβλητή

$$E \Rightarrow E + T \Rightarrow E + T \times F \Rightarrow E + T \times \mathbf{id} \Rightarrow E + F \times \mathbf{id}$$
$$\Rightarrow E + (E) \times \mathbf{id} \Rightarrow E + (T) \times \mathbf{id} \Rightarrow E + (F) \times \mathbf{id}$$
$$\Rightarrow E + (\mathbf{id}) \times \mathbf{id} \Rightarrow T + (\mathbf{id}) \times \mathbf{id} \Rightarrow F + (\mathbf{id}) \times \mathbf{id}$$
$$\Rightarrow \mathbf{id} + (\mathbf{id}) \times \mathbf{id}$$

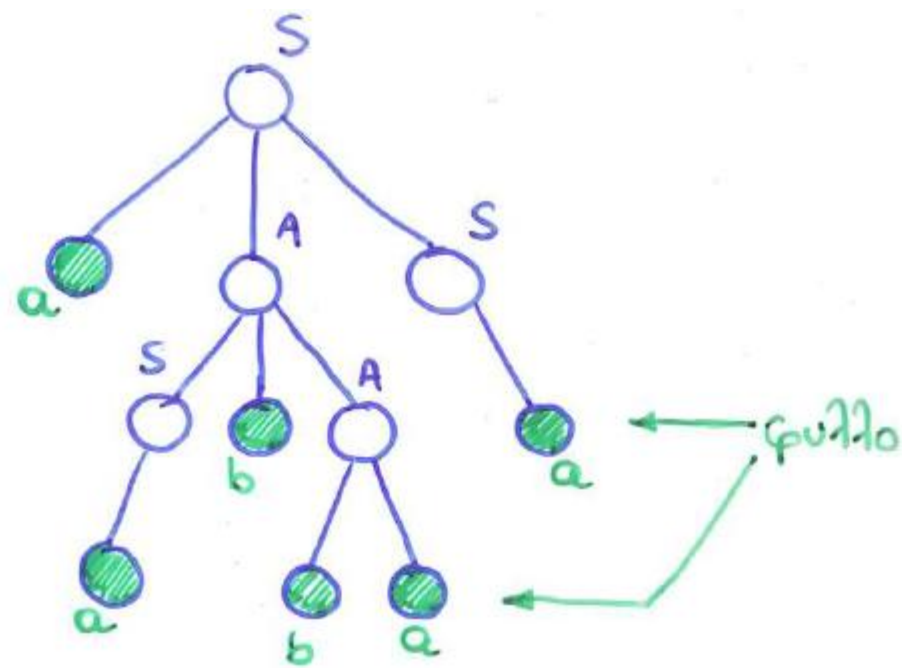
Δένδρα Παραγωγής ή Συντακτικά Δένδρα

Έστω $G = (\{S, A\}, \{a, b\}, R, S)$

όπου $R = \{\ S \rightarrow a\,A\,S \mid a$

$A \rightarrow S\,b\,A \mid S\,S \mid ba\ \}$

$$
\begin{aligned}
S &\Rightarrow a\,A\,S\\
&\Rightarrow a\,A\,a\\
&\Rightarrow a\,S\,b\,A\,a\\
&\Rightarrow a\,a\,b\,A\,a\\
&\Rightarrow a\,a\,b\,b\,a\,a
\end{aligned}
$$



Δένδρα Παραγωγής ή Συντακτικά Δένδρα

Ένα **δένδρο παραγωγής** ή **συντακτικό δένδρο** για τη Γρ.Α.Σ. $G = (V, \Sigma, R, S)$ είναι ένα **διατεταγμένο** δένδρο όπου

- **Κάθε** κόμβος του δέντρου φέρει **επιγραφή** $\in V \cup \Sigma \cup \{ e \}$.
- Η **επιγραφή** της **ρίζας** είναι S.
- Αν ένας **εσωτερικός κόμβος** έχει **επιγραφή** A, τότε $A \in V$.
- Αν ένας **εσωτερικός κόμβος** έχει **επιγραφή** A και οι **κόμβοι-παιδιά** του έχουν **επιγραφές** $t_1, t_2, \ldots, t_n$ από **αριστερά προς τα δεξιά**, τότε ο $A \rightarrow t_1\ t_2 \ldots t_n$ είναι **κανόνας** της G.
- Ένας κόμβος με επιγραφή e είναι **φύλλο** και **δεν έχει αδέλφια**.

Αποτέλεσμα (yield)

η **παράθεση** των **επιγραφών των φύλλων** από **αριστερά προς τα δεξιά**

Γραμματικές Ανεξάρτητες Συμφραζομένων

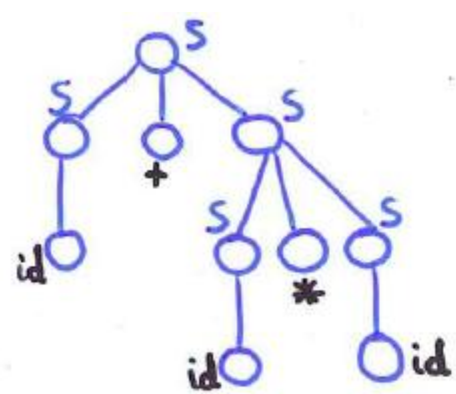
Διφορούμενη (ασαφής ή **πολύτροπη) Γραμματική Α.Σ.**

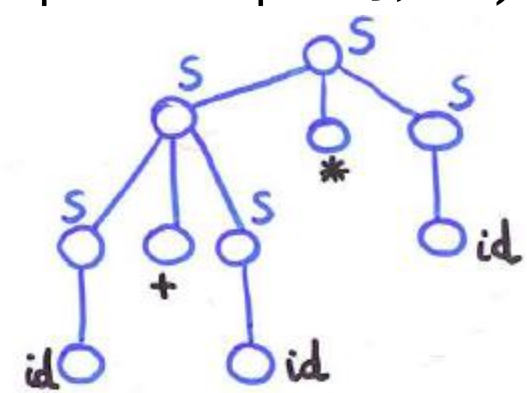
υπάρχουν τουλάχιστον **δύο διαφορετικά συντακτικά δέντρα** για κάποια **συμβολοσειρά** που παράγεται από τη γραμματική

Παράδειγμα

$\boldsymbol{G = (\, V = \{\, S \,\},\ \Sigma = \{\, +, \times, \mathrm{id} \,\},\ R = \{\, S \to S + S \mid S \times S \mid \mathrm{id} \,\},\ S \,)}$

Τότε για τη συμβολοσειρά **id + id × id** έχουμε:

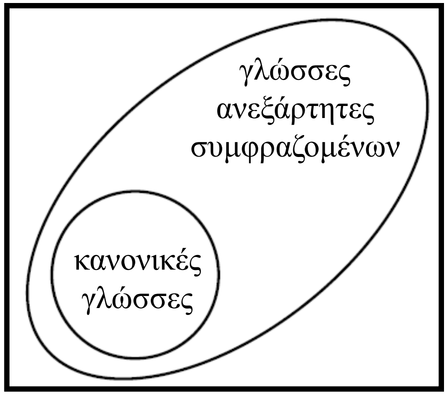




Εγγενώς διφορούμενη (πολυτροπική) γλώσσα ⇔ **κάθε** Γρ.Α.Σ. **διφορούμενη**

Π.χ., $\boldsymbol{\{\, a^n b^n c^m d^m \mid n, m \geq 1 \,\} \cup \{\, a^n b^m c^m d^n \mid n, m \geq 1 \,\}}$

Κανονικές Γλώσσες και Γλώσσες Α.Σ.



Λήμμα

Κάθε κανονική γλώσσα είναι **γλώσσα Α.Σ.**

Απόδειξη

Έστω ***L* κανονική** γλώσσα.

Τότε υπάρχει **Αιτ.Π.Α.** $\boldsymbol{M = (Q, \Sigma, q_0, \delta, F)}$ που αναγνωρίζει την $\boldsymbol{L}$.

Η γλώσσα ***L* παράγεται** από τη γραμματική $\boldsymbol{G = (Q, \Sigma, R, q_0)}$

όπου $\boldsymbol{R = \{ q \rightarrow a\,p \mid q,p \in Q,\ a \in \Sigma,\ \delta(q,a) = p \}}$

$\boldsymbol{\cup \{ q \rightarrow e \mid \forall q \in F \}}$.

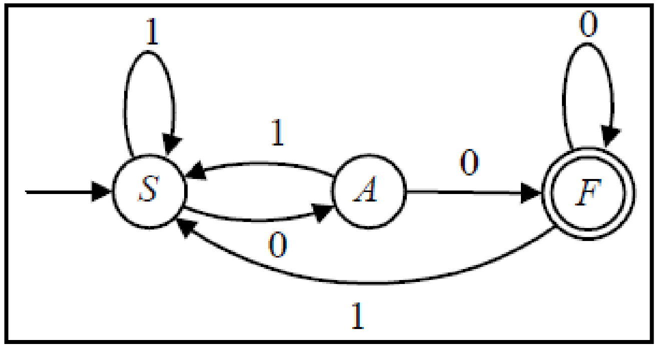
κανονική ή δεξιά γραμμική γραμματική

Παράδειγμα

$\boldsymbol{S \rightarrow 0\,A \mid 1\,S}$

$\boldsymbol{A \rightarrow 0\,F \mid 1\,S}$

$\boldsymbol{F \rightarrow 0\,F \mid 1\,S \mid e}$



Αυτόματα Στοίβας

Στοίβα

άπειρη σε μέγεθος
για εγγραφή
και ανάγνωση

Μη Αιτιοκρατικό Πεπερασμένο Αυτόματο

μια 5-άδα $(Q, \Sigma, q_0, \Delta, F)$ όπου

Q : ένα **πεπερασμένο σύνολο καταστάσεων**,

Σ : το **αλφάβητο** των συμβόλων **εισόδου**,

$q_0 \in Q$: η **αρχική κατάσταση**,

$\Delta \subseteq Q \times (\Sigma \cup \{e\}) \times Q$: η **σχέση μετάβασης** και

$F \subseteq Q$: το σύνολο των **τελικών καταστάσεων**.

Αυτόματο Στοίβας

μια 6-άδα $(Q, \Sigma, \Gamma, q_0, \Delta, F)$ όπου

Q : ένα **πεπερασμένο σύνολο καταστάσεων**,

Σ : το **αλφάβητο** των συμβόλων **εισόδου**,

Γ : το **αλφάβητο** των συμβόλων **στοίβας**,

$q_0 \in Q$: η **αρχική κατάσταση**,

$\Delta \subseteq (Q \times (\Sigma \cup \{e\}) \times \Gamma^*) \times (Q \times \Gamma^*)$: η **σχέση μετάβασης**

$F \subseteq Q$: το σύνολο των **τελικών καταστάσεων**.

Αυτόματα Στοίβας

Η **μετάβαση ((*p*, *a*, *w*), (*q*, *w'*))**, όπου ***p*, *q*** $\in$ ***Q*, *a*** $\in$ **Σ, *w*, *w'*** $\in$ **Γ***, εκτελείται **εάν** το αυτόματο στοίβας βρίσκεται στην κατάσταση ***p***, στην είσοδο διαβάζεται ***a*** και στην κορυφή της στοίβας υπάρχει ***w*** και έχει **ως αποτέλεσμα**, το αυτόματο να πάει στην κατάσταση ***q*** και να **αντικαταστήσει** το ***w*** στην κορυφή της στοίβας με το ***w'***.

Παρόμοια για μεταβάσεις **((*p*, *e*, *w*), (*q*, *w'*))**.

((*p*, *a*, *e*), (*q*, *aaa*)): **ώθηση** των συμβόλων ***aaa*** στη στοίβα

((*p*, *a*, *bbb*), (*q*, *e*)): **απώθηση** των συμβόλων ***bbb*** από τη στοίβα

((*p*, *a*, *ab*), (*q*, *ba*)): **αντικατάσταση** των συμβόλων ***ab*** στην κορυφή της στοίβας από τα ***ba***

!!! Το ***e*** σε μεταβάσεις **((*p*, *a*, *e*), (*q*, *x*))** σημαίνει ότι **αδιαφορούμε** για το περιεχόμενο της στοίβας **και** **όχι** ότι αυτή είναι άδεια.

Αυτόματα Στοίβας

Έστω Αιτ. Π. Α. $\boldsymbol{M = (Q, \Sigma, \Gamma, q_0, \Delta, F)}$.

Συνολική κατάσταση: ένα στοιχείο $(\boldsymbol{q}, \boldsymbol{w}, \boldsymbol{x})$ του $\boldsymbol{Q}\times\boldsymbol{\Sigma^*}\times\boldsymbol{\Gamma^*}$
όπου $\boldsymbol{q}$ είναι η **τρέχουσα κατάσταση** του αυτόματου M,
$\boldsymbol{w}$ είναι το τμήμα της **εισόδου που δεν έχει διαβαστεί** και
$\boldsymbol{x}$ είναι το **τρέχον περιεχόμενο της στοίβας**.

Μια συνολική κατάσταση $(\boldsymbol{p}, \boldsymbol{aw}, \boldsymbol{\beta x})$ όπου $\boldsymbol{p} \in \boldsymbol{Q}$, $\boldsymbol{a} \in \boldsymbol{\Sigma} \cup \{\boldsymbol{e}\}$, $\boldsymbol{w} \in \boldsymbol{\Sigma^*}$, $\boldsymbol{\beta}, \boldsymbol{x} \in \boldsymbol{\Gamma^*}$ **παράγει** (**δίνει**) **σε ένα βήμα** τη συνολική κατάσταση $(\boldsymbol{q}, \boldsymbol{w}, \boldsymbol{\gamma x})$ αν $((\boldsymbol{p}, \boldsymbol{a}, \boldsymbol{\beta}), (\boldsymbol{q}, \boldsymbol{\gamma})) \in \boldsymbol{\Delta}$ με $\boldsymbol{q} \in \boldsymbol{Q}$, $\boldsymbol{\gamma} \in \boldsymbol{\Gamma^*}$.
Γράφουμε: $(\boldsymbol{p}, \boldsymbol{aw}, \boldsymbol{\beta x}) \vdash (\boldsymbol{q}, \boldsymbol{w}, \boldsymbol{\gamma x})$

Έστω $\vdash^*$ η **ανακλαστική & μεταβατική κλειστότητα** της $\vdash$.

Αυτόματα Στοίβας

Ένα Αυτόματο Στοίβας $M = ((Q, \Sigma, \Gamma, q_0, \Delta, F)$ **αναγνωρίζει** μια γλώσσα L **<u>ανν</u>** $\forall\, w \in L,\ (q_0, w, e) \vdash^* (q, e, e)$ **και** $q \in F$.

Αποδοχή με **τελική κατάσταση** **και** **κενή στοίβα**

Παράδειγμα

Δώστε αυτόματο στοίβας για $L = \{\, x\, c\, x^R \mid x \in \{a, b\}^* \,\}$.

Τότε: $M = ((Q, \Sigma, \Gamma, s, \Delta, F)$

όπου $Q = \{s, f\}$, $\Sigma = \{a, b, c\}$, $\Gamma = \{a, b\}$, $F = \{f\}$

και $\Delta = \{$ $((s, a, e), (s, a))$,
$((s, b, e), (s, b))$,
$((s, c, e), (f, e))$,
$((f, a, a), (f, e))$,
$((f, b, b), (f, e))$ $\}$

Ώθηση *συμβόλων του* $\boldsymbol{x}$

Ταίριασμα *συμβόλων του* $\boldsymbol{x^R}$ *με τα σύμβολα της στοίβας*

Αυτόματα Στοίβας

Αιτιοκρατικά Αυτόματα Στοίβας

Τα **αιτιοκρατικά** και τα **μη αιτιοκρατικά** αυτόματα στοίβας **δεν είναι ισοδύναμα** διότι **δεν** υπάρχει **αιτιοκρατικό** αυτόματο στοίβας για τη γλώσσα $\{ x x^R \mid x \in \{a, b\}^* \}$.

Αποδοχή με **τελική κατάσταση** **(*σύγγραμμα Sipser*)**

Ένα Αυτόματο Στοίβας $M = ((Q, \Sigma, \Gamma, q_0, \Delta, F)$ **αναγνωρίζει** μια γλώσσα L **αννν** $\forall w \in L$, $(q_0, w, e) \vdash^* (q, e, x)$ **και** $q \in F$.

Αποδοχή με **κενή στοίβα**

Ένα Αυτόματο Στοίβας $M = ((Q, \Sigma, \Gamma, q_0, \Delta, F)$ **αναγνωρίζει** μια γλώσσα L **αννν** $\forall w \in L$, $(q_0, w, e) \vdash^* (q, e, e)$.

Οι **3 τρόποι αναγνώρισης** γλωσσών είναι **ισοδύναμοι**.

Αυτόματα Στοίβας

Παράδειγμα

Δώστε αυτόματο στοίβας που αναγνωρίζει την $\boldsymbol{L = \{\, 0^n\, 1^n \mid n \geq 0 \,\}}$.

Αποδοχή με **τελική κατάσταση**

Τότε: $\boldsymbol{M = ((Q, \Sigma, \Gamma, q_1, \Delta, F)}$

όπου $\boldsymbol{Q = \{q_1, q_2, q_3, q_4\}}$, $\boldsymbol{\Sigma = \{0, 1\}}$, $\boldsymbol{\Gamma = \{0, \$\}}$, $\boldsymbol{F = \{q_1, q_4\}}$

και η σχέση μετάβασης **Δ** όπως περιγράφεται παρακάτω.

Είσοδος:	0			1			ε		
Στοίβα:	0	\$	ε	0	\$	ε	0	\$	ε
q_1									$\{(q_2, \$)\}$
q_2			$\{(q_2, 0)\}$	$\{(q_3, \varepsilon)\}$					
q_3				$\{(q_3, \varepsilon)\}$				$\{(q_4, \varepsilon)\}$	
q_4									

Αυτόματα Στοίβας

Παράδειγμα

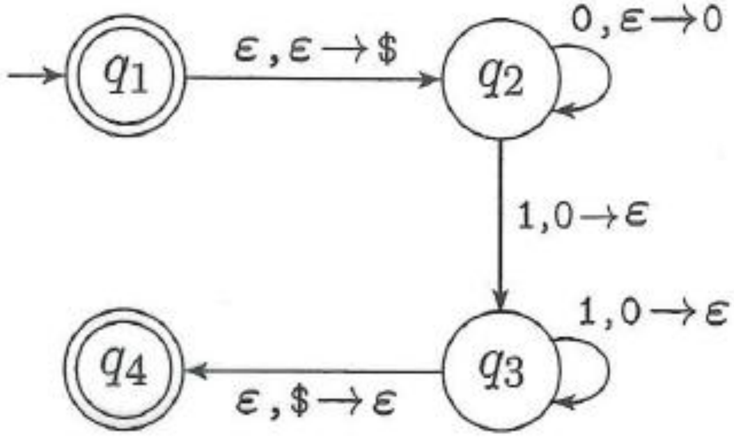
Δώστε αυτόματο στοίβας που αναγνωρίζει την $L = \{ 0^n 1^n \mid n \geq 0 \}$.

Αποδοχή με **τελική κατάσταση**

Τότε: $M = ((Q, \Sigma, \Gamma, q_1, \Delta, F)$

όπου $Q = \{q_1, q_2, q_3, q_4\}$, $\Sigma = \{0, 1\}$, $\Gamma = \{0, \$\}$, $F = \{q_1, q_4\}$

και η σχέση μετάβασης Δ όπως περιγράφεται παρακάτω.



Γλώσσες Ανεξ. Συμφρ. & Αυτόματα Στοίβας

Θεώρημα

Κάθε γλώσσα ανεξάρτητη συμφραζομένων *L* αναγνωρίζεται από (***μη ντετερμινιστικό***) **Αυτόματο Στοίβας**.

Απόδειξη

<u>*Βασική Ιδέα*</u>: Το αυτόματο στοίβας κατασκευάζει **μη αιτιοκρατικά** στη στοίβα του μια **αριστερότερη παραγωγή** της συμβολοσειράς εισόδου και κάθε φορά που προκύπτει **τερματικό σύμβολο στην κορυφή της στοίβας**, το **ταιριάζει** με το αντίστοιχο σύμβολο της συμβολοσειράς εισόδου.

Έστω $G = (V, \Sigma, S, R)$ Γρ.Α.Σ για την L. Τότε $L(M) = L$ όπου

$M = (\, Q = \{q, f\},\ \Sigma,\ \Gamma = V \cup \Sigma,\ \Delta,\ q,\ F = \{f\}\,)$ με

$$\Delta = \{ ((q, e, e), (f, S)) \}$$
$$\cup \{ ((f, e, A), (f, x^R)) \mid \text{για κάθε κανόνα } A \to x \in R \}$$
$$\cup \{ ((f, a, a), (f, e)) \mid \text{για κάθε τερμ. σύμβολο } a \in \Sigma \}.$$

Αποδοχή με Τ.Κ. και κενή στοίβα

Γλώσσες Ανεξ. Συμφρ. & Αυτόματα Στοίβας

Παράδειγμα

$\{\, 0^k\, 1^k \mid k \geq 0 \,\}$: $\quad (\{S\}, \{0,1\}, S, \{S \rightarrow 0S1,\ S \rightarrow e\})$

$M = (\ Q = \{q, f\},\ \Sigma = \{0,1\},\ \Gamma = \{S,0,1\},\ \Delta,\ q, F = \{f\}\ )$ με

$\Delta = \{\ ((q, e, e), (f, S)),$

$((f, e, S), (f, \begin{matrix}0\\S\\1\end{matrix})),$ /* κανόνας $S \rightarrow 0S1$ */

$((f, e, S), (f, e)),$ /* κανόνας $S \rightarrow e$ */

$((f, 0, 0), (f, e)),$ /* τερμ. σύμβολο **0** */

$((f, 1, 1), (f, e))$ /* τερμ. σύμβολο **1** */

$\}$

Γλώσσες Ανεξ. Συμφρ. & Αυτόματα Στοίβας

Θεώρημα
Κάθε γλώσσα ***L*** που **αναγνωρίζεται από αυτόματο στοίβας** **είναι γλώσσα ανεξάρτητη συμφραζομένων**.

Θεώρημα
Μια γλώσσα είναι **ανεξάρτητη συμφραζομένων** **αν και μόνον αν** **αναγνωρίζεται από αυτόματο στοίβας**.

Ιδιότητες Κλειστότητας για Γλ.Α.Σ.

Θεώρημα
Οι γλώσσες ανεξάρτητες συμφραζομένων είναι **κλειστές** ως προς τις πράξεις: **ένωση**, **παράθεση**, **Kleene star**.

Ένωση:
Έστω $G_1 = (V_1, \Sigma, S_1, R_1)$ και $G_2 = (V_2, \Sigma, S_2, R_2)$ δύο **Γρ.Α.Σ.**
με $V_1 \cap V_2 = \varnothing$ και $S \notin V_1 \cup V_2$.
Τότε η Γρ.Α.Σ. (V, Σ, S, R) όπου

$$V = V_1 \cup V_2 \cup \{S\}$$

$$R = R_1 \cup R_2 \cup \{ S \rightarrow S_1 \mid S_2 \}$$

παράγει τη γλώσσα $L(G_1) \cup L(G_2)$.

Ιδιότητες Κλειστότητας για Γλ.Α.Σ.

Παράθεση:
Έστω $G_1 = (V_1, \Sigma, S_1, R_1)$ και $G_2 = (V_2, \Sigma, S_2, R_2)$ δύο **Γρ.Α.Σ.**
με $V_1 \cap V_2 = \varnothing$ και $S \notin V_1 \cup V_2$.
Τότε η Γρ.Α.Σ. (V, Σ, S, R) όπου

$$V = V_1 \cup V_2 \cup \{S\} \qquad R = R_1 \cup R_2 \cup \{ S \rightarrow S_1 S_2 \}$$

παράγει τη γλώσσα $L(G_1)\, L(G_2)$.

Kleene star:
Έστω $G_1 = (V_1, \Sigma, S_1, R_1)$ μια **Γρ.Α.Σ.** με $S \notin V_1$.
Τότε η Γρ.Α.Σ. (V, Σ, S, R) όπου

$$V = V_1 \cup \{S\} \qquad R = R_1 \cup \{ S \rightarrow S_1 S \mid e \}$$

παράγει τη γλώσσα $L(G_1)^*$.

Ιδιότητες Κλειστότητας για Γλ.Α.Σ.

Παράδειγμα

Βρείτε Γρ.Α.Σ. για τη γλώσσα $\boldsymbol{L = \{ 0^i 1^j 2^k \mid j = i + k \}}$.

Σημείωση: $\boldsymbol{L = \{ 0^i 1^j 2^k \mid j = i + k \} = \{ 0^i 1^i 1^k 2^k \}}$

$\boldsymbol{G = (V=\{S_1, S_2, S\}, \Sigma=\{0,1,2\}, S, R)}$

όπου

$$\boldsymbol{R = \{ \quad S_1 \rightarrow 0 S_1 1 \mid e \quad S_2 \rightarrow 1 S_2 2 \mid e \quad S \rightarrow S_1 S_2 \quad \}}$$

Ιδιότητες Κλειστότητας για Γλ.Α.Σ.

Λήμμα

Οι Γλ.Α.Σ. **δεν** είναι **κλειστές** ως προς την **τομή**.

Απόδειξη

Εάν οι Γλ.Α.Σ. ήταν κλειστές ως προς την **τομή**, τότε και η

$$\{ 0^k 1^k 2^k \mid k \geq 0 \} = \{ 0^k 1^k 2^p \mid k, p \geq 0 \} \cap \{ 0^q 1^k 2^k \mid k, q \geq 0 \}$$

θα ήταν Γλ.Α.Σ. **Άτοπο.**

Θα δούμε ότι η $\{ 0^k 1^k 2^k \}$ ***δεν*** *είναι Γλ.Α.Σ.*

Λήμμα

Οι Γλ.Α.Σ. **δεν** είναι **κλειστές** ως προς τη **συμπλήρωση**.

Απόδειξη

Εάν οι Γλ.Α.Σ. ήταν κλειστές ως προς τη **συμπλήρωση**, τότε θα ήταν και ως προς την τομή καθώς $L_1 \cap L_2 = \overline{\left(\overline{L_1} \cup \overline{L_2} \right)}$.

Άτοπο.

Ιδιότητες Κλειστότητας για Γλ.Α.Σ.

Θεώρημα

Οι Γλ.Α.Σ. είναι **κλειστές** ως προς την **τομή** με **κανονικές γλώσσες**.

Απόδειξη

Έστω **Γλ.Α.Σ.** L_1 και **κανονική γλώσσα** L_2,

αυτόματο στοίβας $M_1 = (Q_1, \Sigma, \Gamma, \Delta_1, q_1, F_1) : L(M_1) = L_1$ &

αιτιοκρατικό Π.Α. $M_2 = (Q_2, \Sigma, \delta, q_2, F_2) : L(M_2) = L_2$.

Τότε, $L(M) = L_1 \cap L_2$ για το αυτόματο στοίβας

$M = (Q_1 \times Q_2, \Sigma, \Gamma, \Delta, (q_1,q_2), F_1 \times F_2)$

εκτέλεση των M_1 & M_2 *παράλληλα*

όπου $\forall p \in Q_1, \forall q \in Q_2, \forall a \in \Sigma$:

$$\Delta((p, q), a, w) = \{ ((p', q'), w') \mid \forall (p', w') \in \Delta_1(p, a, w) \text{ και για } q' = \delta(q, a) \}$$

$$\Delta((p, q), e, w) = \{ ((p', q), w') \mid \forall (p', w') \in \Delta_1(p, e, w) \}$$

Κανονική Μορφή Chomsky

Ορισμός

Μια Γρ.Α.Σ. $\boldsymbol{G = (V, \Sigma, S, R)}$ είναι σε **κανονική μορφή Chomsky** (***Chomsky normal form – CNF***) εάν **κάθε** κανόνας της είναι σε μια από τις εξής μορφές:

$A \rightarrow B\, C$ A, B, C **μεταβλητές** και B, $C \neq S$

$A \rightarrow a$ A **μεταβλητή**, $a \in \Sigma$

Επιπλέον, επιτρέπουμε τον κανόνα $S \rightarrow e$.

Θεώρημα

Κάθε Γλ.Α.Σ. παράγεται από κάποια Γρ.Α.Σ. **σε κανονική μορφή Chomsky**.

Avram **Noam Chomsky**. Αμερικανός γλωσσολόγος, φιλόσοφος, ιστορικός, πολιτικός ακτιβιστής, κοινωνικός κριτικός.

Κανονική Μορφή Chomsky

Αλγόριθμος Μετασχηματισμού σε CNF

Απόδειξη. Έστω Γρ.Α.Σ. $G = (V, \Sigma, S, R)$ που παράγει τη γλώσσα.

1. **Προσθέτουμε νέα αρχική μεταβλητή S_0**
   και κανόνα $S_0 \to S$

2. **Αφαίρεση** κανόνων $A \to e$

   Για **κάθε** κανόνα $A \to e$ όπου $A \neq S_0$

   Αφαιρούμε $A \to e$

   Για **κάθε** κανόνα $X \to u\,A\,v$

   Για **κάθε** συνδυασμό εμφανίσεων του A στο δεξί μέλος του κανόνα, **προσθέτουμε** τον νέο κανόνα που προκύπτει αν **αντικαταστήσουμε** αυτές τις εμφανίσεις του A με e

***Παράδειγμα*:** για $A \to e$ & $X \to u\,A\,v\,A\,w$, **αφαιρούμε** $A \to e$
και **προσθέτουμε** $X \to u\,v\,A\,w \;\mid\; u\,A\,v\,w \;\mid\; u\,v\,w$

Κανονική Μορφή Chomsky

Αλγόριθμος Μετασχηματισμού σε CNF

Απόδειξη. (συνέχεια)

3. Αφαίρεση κανόνων $A \rightarrow B$ *(μοναδιαίοι κανόνες)*

Για **κάθε** κανόνα $A \rightarrow B$

Αφαιρούμε $A \rightarrow B$

Για **κάθε** κανόνα $B \rightarrow u$

προσθέτουμε τον κανόνα $A \rightarrow u$ **εκτός αν** αυτός ο κανόνας είναι ένας από τους μοναδιαίους κανόνες που αφαιρέσαμε

4. Αφαίρεση κανόνων $A \rightarrow x_1\, x_2 \ldots x_k$ $k \geq 3$, $x_1, \ldots, x_k \in V \cup \Sigma$

Για **κάθε** κανόνα $A \rightarrow x_1\, x_2 \ldots x_k$ $(k \geq 3)$

Αφαιρούμε τον κανόνα

Προσθέτουμε: $A \rightarrow x_1\, Y_1$ $Y_1 \rightarrow x_2 Y_2$ $\ldots$ $Y_{k-2} \rightarrow x_{k-1}\, x_k$

όπου $Y_1, \ldots, Y_{k-2}$ **νέες** μεταβλητές

Κανονική Μορφή Chomsky

Αλγόριθμος Μετασχηματισμού σε CNF

***Απόδειξη*. (*συνέχεια*)**

5. Αφαίρεση κανόνων $A \to x_1\, x_2$ όπου $x_1 \in \Sigma$ ή $x_2 \in \Sigma$

Για **κάθε** κανόνα $A \to x_1\, B$ όπου $x_1 \in \Sigma$

Αφαιρούμε τον κανόνα

Προσθέτουμε: $A \to U_1\, B$ και $U_1 \to x_1$ με U_1 **νέα** μεταβλητή

*Ομοίως, για **κάθε** κανόνα* $A \to B\, x_2$ *όπου* $x_2 \in \Sigma$.

Για **κάθε** κανόνα $A \to x_1\, x_2$ όπου $x_1, x_2 \in \Sigma$

Αφαιρούμε τον κανόνα

Προσθέτουμε: $A \to U_1\, U_2$ $\quad U_1 \to x_1$ $\quad U_2 \to x_2$

με U_1, U_2 **νέες** μεταβλητές

***Παράδειγμα*:** Μετατρέψτε σε CNF τη Γρ.Α.Σ. $G = (V, \Sigma, S, R)$ με
$V = \{S, A, B\}$ & $R = \{\, S \to ASA \mid aB,\quad A \to B \mid S,\quad B \to b \mid e \,\}$

Γλώσσες Όχι Ανεξάρτητες Συμφραζομένων

Λήμμα Άντλησης για Γλ. Α. Σ.

Για κάθε **άπειρη** γλώσσα **ανεξάρτητη συμφραζομένων** L,
υπάρχει ακέραιος $n \geq 1$ τέτοιος ώστε
για **κάθε** συμβολοσειρά $w \in L$ με $|w| \geq n$
υπάρχουν συμβολοσειρές u, v, x, y, z:

$$w = uvxyz \quad \text{με} \quad |vxy| \leq n, \quad vy \neq e$$

και για **κάθε** ακέραιο $i \geq 0$, $uv^i xy^i z \in L$.

Λήμμα Άντλησης για Κανονικές Γλώσσες

Για **κάθε άπειρη κανονική** γλώσσα L,
υπάρχει ακέραιος $n \geq 1$ τέτοιος ώστε
για **κάθε** συμβολοσειρά $w \in L$ με $|w| \geq n$
υπάρχουν συμβολοσειρές x, y, z:

$$w = xyz, \quad |xy| \leq n, \quad y \neq e$$

και για **κάθε** ακέραιο $i \geq 0$, $xy^i z \in L$.

Γλώσσες Όχι Ανεξάρτητες Συμφραζομένων

Απόδειξη του Λήμματος Άντλησης για Γλ. Α. Σ.

Έστω $L = L(G)$ με $G = (V, \Sigma, S, R)$ και $n = \varphi(G)^{|V|} - 1$ όπου $\varphi(G)$ το **μέγιστο** μήκος **δεξιού μέλους** κανόνα στο σύνολο R.

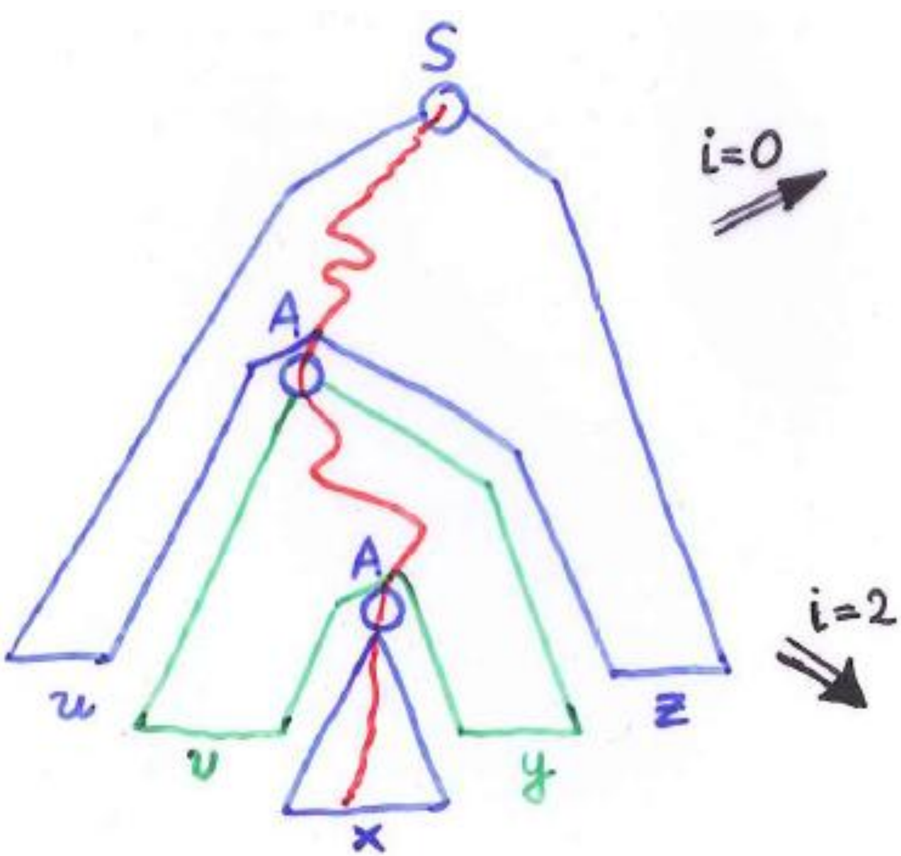
Καθώς **κάθε** δέντρο βαθμού d και ύψους h έχει $\leq d^h$ φύλλα, το αποτέλεσμα (παραγόμενη **συμβολοσειρά**) **κάθε** συντακτικού δέντρου της G με ύψος h έχει **μήκος** $\leq \varphi(G)^h$.

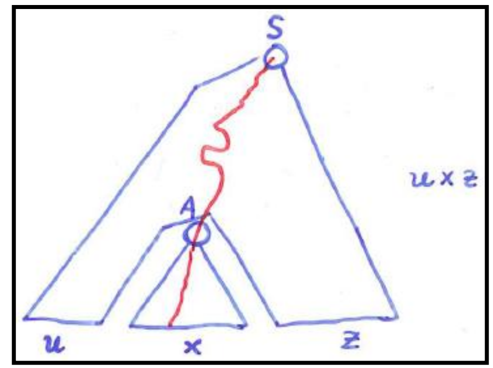
Αφού $|w| \geq n = \varphi(G)^{|V|} - 1 \Rightarrow |w| > \varphi(G)^{|V|}$, το συντακτικό δέντρο T_w για τη συμβολοσειρά w έχει **ύψος** $> |V|$.

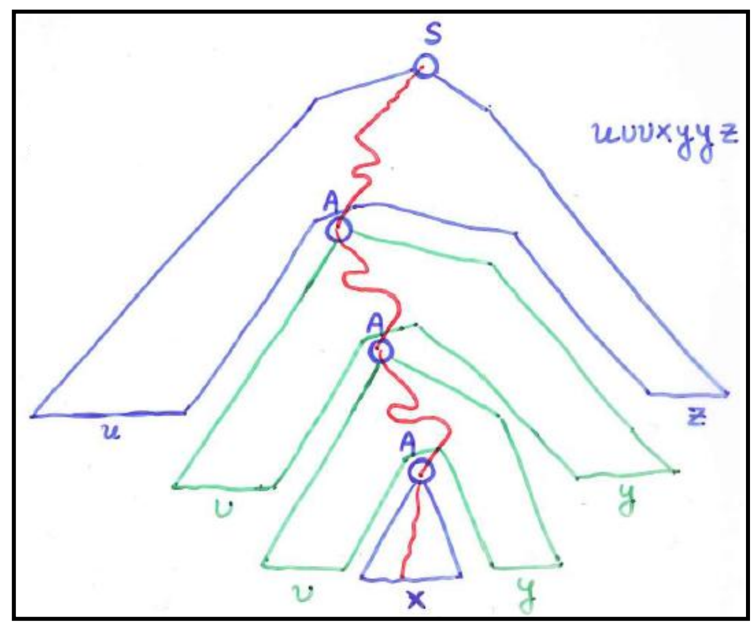
Άρα το δέντρο T_w έχει μια διαδρομή από τη ρίζα του σε κάποιο φύλλο με **μήκος** $\geq |V| + 1$, δηλαδή περνά από $\geq |V| + 1$ εσωτ. κόμβους. $\Rightarrow$ Υπάρχουν (τουλάχιστον) **2 κόμβοι** σε αυτή τη διαδρομή με την **ίδια επιγραφή**, έστω A.

Γλώσσες Όχι Ανεξάρτητες Συμφραζομένων

Απόδειξη του Λήμματος Άντλησης για Γλ. Α. Σ. (*συνέχεια*)







Χρησιμοποιώντας το Λήμμα Άντλησης

Γενική μέθοδος **απόδειξης** ότι μια **άπειρη** γλώσσα ***L*** **δεν είναι ανεξάρτητη συμφραζομένων**

Εις άτοπον απαγωγή

- Υποθέτουμε ότι η ***L*** είναι γλώσσα **ανεξάρτητη συμφραζομένων** και έστω ***n*** η σταθερά του Λ.Α. για Γλ.Α.Σ. για την ***L***.
- **Επιλέγουμε** συμβολοσειρά $w \in L$ με $|w| \geq n$.
- Για **κάθε** χωρισμό της w σε $w = u\,v\,x\,y\,z$
  όπου $|v\,x\,y| \leq n$, $|v\,y| \geq 1$
  βρίσκουμε $i \geq 0$ τ.ώ. $u\,v^i\,x\,y^i\,z \notin L$.
- **Άτοπο** $\Rightarrow$ ***L*** **δεν** είναι **γλώσσα ανεξάρτητη συμφραζομένων**.

***Σημείωση*:** $i \neq 1$

Χρησιμοποιώντας το Λήμμα Άντλησης

***Παράδειγμα*:** Δείξτε ότι η $L = \{ 0^k\, 1^k\, 2^k \mid k \geq 0 \}$ **δεν** είναι Γλ.Α.Σ.

Εις άτοπον απαγωγή.

- Υποθέτουμε ότι η L είναι **Γλ.Α.Σ.** και
  έστω n η σταθερά του Λ.Α. για Γλ.Α.Σ. για την L.
- **Επιλέγουμε** συμβολοσειρά $w = 0^n\, 1^n\, 2^n \in L$ και $|w| \geq n$.
- **Περίπτωση 1:** Το v ή το y περιέχει **διαφορετικά σύμβολα**.
  Π.χ., έστω $v = 0^p\, 1^q$ με $p, q > 0$. Τότε $v^2 = 0^p\, 1^q\, 0^p\, 1^q$ και
  $u\, v^2\, x\, y^2\, z \notin L$. *Παρομοίως, για τις άλλες περιπτώσεις.*
- **Περίπτωση 2:** Το v & y **δεν** περιέχει **διαφορετικά** σύμβολα.
  Τότε, στη $u\, v^2\, x\, y^2\, z$, (τουλάχιστον) **ένα** από τα 0,1,2 **αυξάνει** σε
  μήκος & (τουλάχιστον) **ένα διατηρεί το μήκος** του $\Rightarrow u\, v^2\, x\, y^2\, z \notin L$
- **Άτοπο** $\Rightarrow$ η L **δεν** είναι **Γλ.Α.Σ.**

Χρησιμοποιώντας το Λήμμα Άντλησης

Παρατήρηση

Για κάθε άπειρη Γλ.Α.Σ. L, το Λήμμα Άντλησης συνεπάγεται ότι υπάρχουν συμβολοσειρές $w_i \in L$ ($i \geq 0$) τέτοιες ώστε $|w_i| = \lambda + i\,\mu$ για ακέραιους $\lambda = |u\,x\,z|$, $\mu = |v\,y|$.

Με βάση αυτήν την παρατήρηση, μπορούμε να δείξουμε ότι η $L = \{\, a^k \mid k = j^2 \text{ για ακέραιο } j \,\}$ **δεν** είναι Γλ.Α.Σ.

Σημειώστε ότι $L = \{\, e, a, a^4, a^9, a^{16}, a^{25}, a^{36}, \ldots \}$

Εάν η σταθερά του Λ.Α. είναι n, επιλέγουμε $w = a^{n^2}$.

Τότε: $v = a^p$, $y = a^q$ (όπου $1 \leq p+q \leq n$), $u\,v^2\,x\,y^2\,z = a^{n^2+p+q}$.

Αλλά $n^2 < n^2 + p + q < n^2 + 2n + 1 = (n+1)^2$

$\Rightarrow u\,v^2\,x\,y^2\,z = a^{n^2+p+q} \notin L$. Άτοπο $\Rightarrow$ η L δεν είναι Γλ.Α.Σ.

Γλώσσες Όχι Ανεξάρτητες Συμφραζομένων

Γλώσσες που δεν είναι Ανεξάρτητες Συμφραζομένων

- Γλώσσες με συμβολοσειρές **τουλάχιστον 3 μέρη των οποίων συσχετίζονται** ή υπάρχουν **τουλάχιστον 2 «τεμνόμενα» ζεύγη συσχετιζόμενων μερών**
  Π.χ., $\{\ \mathbf{0}^k\, \mathbf{1}^k\, \mathbf{2}^k \mid \boldsymbol{k} \geq 0\ \}$, $\{\ \mathbf{0}^k\, \boldsymbol{x}\, \mathbf{0}^{2k}\, \boldsymbol{x}^R \mid \boldsymbol{k} \geq 0,\ \boldsymbol{x} \in \{1,2\}^*\ \}$
- Γλώσσες στις οποίες **η διαφορά μηκών διαδοχικών συμβολο-σειρών** (σε μια διάταξη κατά **αύξον μήκος**) συνεχώς **αυξάνει**
  Π.χ., $\{\ \boldsymbol{a}^k \mid \boldsymbol{k} = \boldsymbol{j}^2$ για ακέραιο $j\ \}$, $\{\ \boldsymbol{babaab \ldots ba^k b} \mid \boldsymbol{k} \geq 1\ \}$
- **Άλλες** γλώσσες
  Π.χ., $\{\ \boldsymbol{w}\boldsymbol{w} \mid \boldsymbol{w} \in \{0,1\}^*\ \}$, $\{\ \boldsymbol{a}^p \mid \boldsymbol{p}$ πρώτος αριθμός $\}$

Προσοχή: οι $\{\ \mathbf{0}^k\, \mathbf{1}^k\, \mathbf{1}^k \mid \boldsymbol{k} \geq 0\ \}$, $\{\ \mathbf{0}^k\, \mathbf{1}^s\, \mathbf{1}^k\, \mathbf{2}^s \mid \boldsymbol{s}, \boldsymbol{k} \geq 0\ \}$ είναι Γλ.Α.Σ.

Αλγόριθμοι Απόφασης για Γρ.Α.Σ.

Θεώρημα

Η γλώσσα L των συμβολοσειρών που παράγονται από μια γραμματική Α.Σ. G:

- **δεν είναι κενή αν και μόνο αν** υπάρχει **συμβολοσειρά** $\in L$ με **μήκος μικρότερο** του n,
- είναι **άπειρη αν και μόνο αν** υπάρχει **συμβολοσειρά** $w \in L$ με **μήκος** $n \leq |w| \leq 2n$

για **σταθερά** n που εξαρτάται από την G.

Έλεγχος εάν συμβολοσειρά $w \in L(G)$

αλγόριθμος δυναμικού προγραμματισμού χρόνου $O(|w|^3 |G|)$
(Cocke, Younger, Kasami (‘70))

Αλγόριθμοι Απόφασης για Γρ.Α.Σ.

Προβλήματα που δεν έχουν λύση (αλγορίθμους)

- Δοθεισών δύο γραμματικών, **παράγουν την ίδια γλώσσα**;
- Για δοθείσα γλώσσα ανεξάρτητη συμφραζομένων, είναι το **συμπλήρωμά** της **γλώσσα ανεξάρτητη συμφραζομένων**;
- Για δοθείσα γραμματική ανεξάρτητη συμφραζομένων, είναι αυτή **διφορούμενη**;
- …

Ενδιαφέρον Θεώρημα (*απόρροια του Λήμματος Άντλησης*)

Κάθε Γλ.Α.Σ. με **μονοσύνολο αλφάβητο Σ** (δηλαδή $|\Sigma| = 1$) είναι **κανονική**.