ΓΡΑΜΜΑΤΙΚΕΣ ΧΩΡΙΣ ΣΥΜΦΡΑΖΟΜΕΝΑ

Ορισμός:

Μία **γραμματική Χωρίς Συμφραζόμενα** είναι μια τετράδα: $G = (V, \Sigma, S, P)$:

- V το σύνολο των μεταβλητών
- Σ το σύνολο των τερματικών συμβόλων ($V \cap \Sigma = \emptyset$)
- $S \in V$ είναι η αρχική μεταβλητή
- P το σύνολο κανόνων με κάθε κανόνα να είναι της μορφής $W \to w$ με
  - $W \in V$ (είναι μία μεταβλητή) και
  - $w \in (V \cup \Sigma)^*$ (παράθεση μεταβλητών και μη τερματικών συμβόλων)

Παράδειγμα 1: Η Γραμματική Ανεξάρτητη Συμφραζομένων για την γλώσσα $L = \{0^n 1^n \mid n \geq 0\}$

$$\begin{cases} S \to 0S1 \\ S \to \varepsilon \end{cases}$$

Σχόλια:

- Τα παραπάνω λέγονται κανόνες της γραμματικής διότι ξεκινώντας από την μεταβλητή S μπορούμε να παράγουμε με διαδοχική χρήση των κανόνων οποιαδήποτε συμβολοσειρά της γλώσσας.
- Ο 1ος κανόνας $S \to 0S1$ λέγεται και αναδρομικός κανόνας διότι επανεμφανίζει την μεταβλητή S
- Ο 2ος κανόνας $S \to \varepsilon$ λέγεται και τερματικός κανόνας διότι σταματά τις εμφανίζεις μεταβλητών.

Παραδείγματα Παραγωγών:

S $\Rightarrow \varepsilon$	S $\Rightarrow 0S1$ $\Rightarrow 0\varepsilon 1 = 01$	S $\Rightarrow 0S1$ $\Rightarrow 00S11$ $\Rightarrow 00\varepsilon 11 = 0011$	S $\Rightarrow 0S1$ $\Rightarrow 00S11$ $\Rightarrow 000S111$ $\Rightarrow 000\varepsilon 111 = 000111$	S $\Rightarrow 0S1$ $\Rightarrow 00S11$ $\Rightarrow 000S111$ $\Rightarrow 0000S1111$ $\Rightarrow 0000\varepsilon 1111 = 00001111$	......

Παράδειγμα 2: Η Γραμματική για την γλώσσα $L = \{0^n 1^m 0^m 1^n \mid n, m \geq 0\}$

$$\begin{cases} S \to 0S1 \mid X \\ X \to 1X0 \mid \varepsilon \end{cases}$$

Σχόλια: Το | διαβάζεται ή (ή διαζευκτικό)

Ιδιότητα	Γραμματική Χωρίς Συμφραζόμενα
Ισότητα $\{0^n 1^n \mid n \geq 0\}$	$S \to 0S1 \mid \varepsilon$
Αναλογία $\{0^{2n} 1^{3n} \mid n \geq 0\}$	$S \to 00S111 \mid \varepsilon$
Παλινδρομ/τα $\{wcw^R \mid w \in \{a,b\}^*\}$	$S \to aSa \mid bSb \mid c$
Ανισότητα $\{a^n b^m \mid n \leq m\}$	$S \to aSb \mid X,\ X \to bX \mid \varepsilon$
$\{a^n b^m \mid n < m\}$	$S \to aSb \mid X,\ X \to bX \mid b$
$\{a^n b^m \mid n > m\}$	$S \to aSb \mid X,\ X \to aX \mid a$
Συμμετρία στο Κέντρο $\{a^n b^m c^m d^n \mid n,m \geq 0\}$	$S \to aSd \mid X,\ X \to bXc \mid \varepsilon$
$\{a^{n+m} b^m c^n \mid n,m \geq 0\}$ $\{a^i b^j c^k \mid i = j + k\}$	$S \to aSc \mid X,\ X \to aXb \mid \varepsilon$
$\{a^i b^j c^k \mid i > j + k\}$	$S \to aSc \mid X,\ X \to aXb \mid Y$ $Y \to aY \mid a$
Παράθεση $\{a^n b^n c^m d^m \mid n,m \geq 0\}$	$S \to XY$ $X \to aXb \mid \varepsilon$ $Y \to cYd \mid \varepsilon$
$\{a^n b^{n+m} c^n \mid n,m \geq 0\}$ $\{a^i b^j c^k \mid j = i + k\}$	$S \to XY$ $X \to aXb \mid \varepsilon$ $Y \to bYc \mid \varepsilon$
Διάζευξη Συμβ/ρών $\{a^i b^j c^k \mid i = j \text{ η } j = k\}$	$S \to S_1 \mid S_2$ $S_1 \to X_1 X_2 \quad X_1 \to aX_1 b \mid \varepsilon \quad X_2 \to cX_2 \mid \varepsilon$ $S_2 \to Y_1 Y_2 \quad Y_1 \to aY_1 \mid \varepsilon \quad Y_2 \to bY_2 c \mid \varepsilon$
Κανονικές $\{a^n \mid n \geq 0\}$	$S \to aS \mid \varepsilon$
$\{a^n \mid n > 0\}$	$S \to aS \mid a$

ΚΑΝΟΝΙΚΗ ΓΡΑΜΜΑΤΙΚΗ

Ορισμός Κανονικής Γραμματικής:

- Μία γραμματική χωρίς συμφραζόμενα θα λέγεται **Κανονική Γραμματική** αν και μόνο αν οι κανόνες της έχουν αποκλειστικά και μόνο τη μορφή:

$$X \to \sigma \quad ή \quad X \to \sigma Y$$

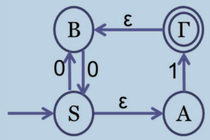
- όπου
  - $X, Y \in V$ (είναι μεταβλητές)
  - $\sigma \in \Sigma$ (είναι τερματικά σύμβολα, δηλαδή σύμβολα του αλφαβήτου ή η κενή συμβολοσειρά)

Λήμμα: Κάθε Κανονική Γραμματική είναι και Γραμματική Χωρίς Συμφραζόμενα

Κανόνες Μετατροπής ΜΠΑε,ΜΠΑ,ΝΠΑ σε Κανονική Γραμματική

- Κάθε κατάσταση γίνεται μεταβλητή. Ειδικά την αρχική κατάσταση την ονομάζουμε S.
- Βάζουμε τον κανόνα $X \to \sigma Y$ αν και μόνο αν από την κατάσταση Χ μεταβαίνουμε στην Υ με το σύμβολο σ
- Βάζουμε τον κανόνα $X \to Y$ αν και μόνο αν από την κατάσταση Χ μεταβαίνουμε στην Υ με ε−κίνηση
- Βάζουμε τον κανόνα $X \to \varepsilon$ αν η Χ είναι τελική κατάσταση.

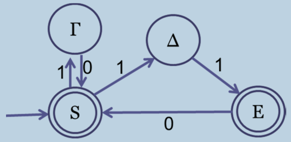
Παράδειγμα: Στο ακόλουθο ΜΠΑ-ε



αντιστοιχεί η κανονική γραμματική

$$\begin{cases} S \to A \mid 0B \\ A \to 1\Gamma \\ B \to 0S \\ \Gamma \to B \mid \varepsilon \end{cases}$$

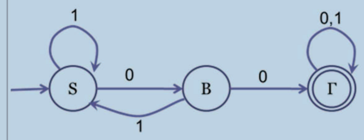
Παράδειγμα: Στο ακόλουθο ΜΠΑ



αντιστοιχεί η κανονική γραμματική

$$\begin{cases} S \to 1\Gamma \mid 1\Delta \mid \varepsilon \\ \Gamma \to 0S \\ \Delta \to 1E \\ E \to 0S \mid \varepsilon \end{cases}$$

Παράδειγμα: Στο ακόλουθο ΝΠΑ



αντιστοιχεί η κανονική γραμματική

$$\begin{cases} S \to 0B \mid 1S \\ B \to 0\Gamma \mid 1S \\ \Gamma \to 0\Gamma \mid 1\Gamma \mid \varepsilon \end{cases}$$

ΝΤΕΤΕΡΜΙΝΙΣΤΙΚΑ ΑΥΤΟΜΑΤΑ ΣΤΟΙΒΑΣ

Ορισμός:

Ένα **Ντετερμινιστικό Αυτόματο Στοίβας** είναι μία 7-άδα $M = (Q, \Sigma, \Gamma, q_0, Z_0, \delta, F)$

Όπου:

- Q είναι το σύνολο των καταστάσεων
- Σ είναι το αλφάβητο των συμβόλων εισόδου
- Γ είναι το αλφάβητο των συμβόλων στοίβας
- q_0 είναι η αρχική κατάσταση
- Z_0 είναι το αρχικό σύμβολο του σωρού
- δ είναι η συνάρτηση μετάβασης (π.χ. $\delta(q_1, \sigma, \sigma') = (q_2, w)$ που σημαίνει ότι είμαστε στην q_1 διαβάζουμε σ από την είσοδο και η στοίβα έχει πάνω-πάνω το σ', το αφαιρούμε πάμε στην q_2 και βάζουμε στην στοίβα την w).
- F είναι το σύνολο των τελικών καταστάσεων

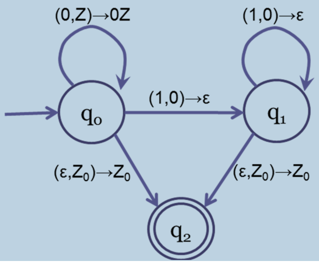
Να κατασκευαστεί Ντετερμινιστικό Αυτόματο Στοίβας που να αναγνωρίζει τις συμβολοσειρές της γλώσσας: $L = \{0^n 1^n \mid n \geq 0\}$

ΛΥΣΗ:

Αλγόριθμος Διαχείρισης Στοίβας

- Για κάθε 0 που διαβάζουμε, θα προσθέτουμε ένα 0 στη στοίβα.
- Έπειτα για κάθε 1 που διαβάζουμε, θα αφαιρούμε ένα 0 από την στοίβα.

Σχηματικά:



Το αυτόματο είναι η 7άδα: $M = (Q, \Sigma, \Gamma, q_0, Z_0, \delta, F)$ όπου:

- $Q = \{q_0, q_1, q_2\}$
- $\Sigma = \{0,1\}$
- $\Gamma = \{Z_0, 0\}$
- q_0 είναι η αρχική κατάσταση
- Z_0 είναι το αρχικό σύμβολο του σωρού
- δ είναι η συνάρτηση μετάβασης που περιγράφεται από τον ακόλουθο πίνακα μετάβασης.
- $F = \{q_2\}$

Ο πίνακας μετάβασης είναι:

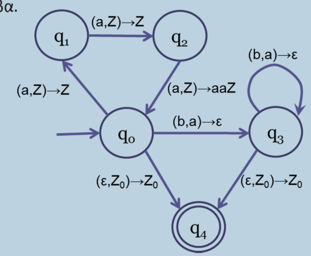
Αριθμός	Κατ/ση	Σύμβολο Εισόδου	Σύμβολο Σωρού	Κίνηση	Επεξήγηση
1	q_0	0	Z	$(q_0, 0Z)$	Διάβαζουμε 0 από την είσοδο, προσθέτουμε 0 στην στοίβα
2	q_0	1	0	(q_1, ε)	Διάβαζουμε το πρώτο 1, Αφαιρούμε 0 από τη στοίβα.
3	q_1	1	0	(q_1, ε)	Διάβαζουμε επόμενο 1, Αφαιρούμε 0 από τη στοίβα.
4	q_1	ε	Z_0	(q_2, Z_0)	Αποδοχή.
5	q_0	ε	Z_0	(q_2, Z_0)	Αποδοχή (κενή συμβολοσειρά).
Οι υπόλοιποι συνδυασμοί					ΤΙΠΟΤΑ

ΙΔΙΟΤΗΤΕΣ ΓΛΩΣΣΩΝ με ΝΤΕΤ.ΑΥΤ.ΣΤΟΙΒΑΣ

ΑΝΑΛΟΓΙΑ (π.χ. 3: 2) $L = \{a^{3n} b^{2n} \mid n \geq 0\}$

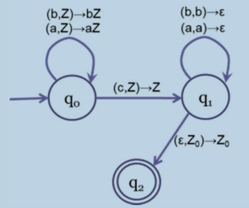
Αλγόριθμος Διαχείρισης της Στοίβας

- Για κάθε **τρία** a που διαβάζουμε, θα προσθέτουμε δύο a στη στοίβα.
- Έπειτα για **κάθε b** που διαβάζουμε, θα αφαιρούμε ένα a από την στοίβα.



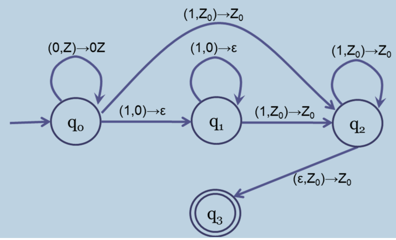
ΠΑΛΙΝΔΡΟΜΙΚΟΤΗΤΑ $L = \{wcw^R \mid w \in \{a,b\}^*\}$

- Κάθε σύμβολο που διαβάζουμε το **βάζουμε** στην στοίβα
- Διαβάζουμε το c
- **Ταυτίζουμε** τα σύμβολα που διαβάζουμε με τα σύμβολα που υπάρχουν στη στοίβα



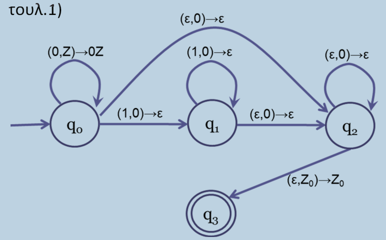
ΑΝΙΣΟΤΗΤΑ $L = \{0^n 1^m \mid n < m\}$

- Για κάθε 0 που διαβάζω, βάζω ένα 0 στη στοίβα
- Για κάθε 1 που διαβάζω, αφαιρώ ένα 0 από στοίβα
- Διαβάζω τα επόμενα 1 (πρέπει να είναι τουλ. 1)



ΑΝΙΣΟΤΗΤΑ $L = \{0^n 1^m \mid n > m\}$

- Για κάθε 0 που διαβάζω, βάζω ένα 0 στη στοίβα
- Για κάθε 1 που διαβάζω, αφαιρώ ένα 0 από στοίβα
- Αφαιρώ τα 0 που έχουν απομείνει στη στοίβα (πρέπει να είναι τουλ.1)



ΜΗ ΝΤΕΤΕΡΜΙΝΙΣΤΙΚΑ ΑΥΤΟΜΑΤΑ ΣΤΟΙΒΑΣ

<u>Ορισμός:</u>
Ένα **Μη Ντετερμινιστικό Αυτόματο Στοίβας** είναι μία 7-άδα $M = (Q, \Sigma, \Gamma, q_0, Z_0, \delta, F)$
Όπου:

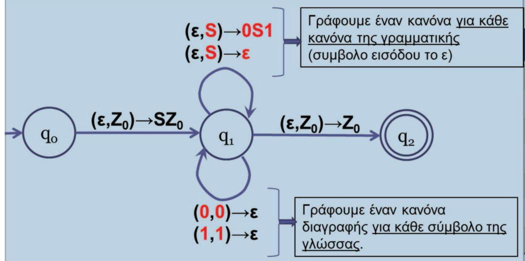
- Q είναι <u>το σύνολο των καταστάσεων</u>
- Σ είναι το <u>αλφάβητο των συμβόλων εισόδου</u>
- Γ είναι το <u>αλφάβητο των συμβόλων στοίβας</u>
- q_0 είναι <u>η αρχική κατάσταση</u>
- Z_0 είναι <u>το αρχικό σύμβολο του σωρού</u>
- δ είναι <u>η συνάρτηση μετάβασης</u> (π.χ. $\delta(q_1, \sigma, \sigma') = (q_2, w)$ που σημαίνει ότι είμαστε στην q_1 διαβάζουμε σ από την είσοδο και η στοίβα έχει πάνω-πάνω το σ', το αφαιρούμε πάμε στην q_2 και βάζουμε στην στοίβα την w).
- F είναι <u>το σύνολο των τελικών καταστάσεων</u>

Να κατασκευαστεί Μη Ντετερμινιστικό Αυτόματο Στοίβας που να αναγνωρίζει τις συμβολοσειρές της γλώσσας: $L = \{0^n 1^n \mid n \geq 0\}$

ΛΥΣΗ:

<u>Το Αυτόματο Στοίβας Προσομοιώνει τη λειτουργία της Γραμματικής Χωρίς Συμφραζόμενα που παράγει τις συμβολοσειρές της γλώσσας</u>

<u>Σχηματικά:</u>



Το αυτόματο είναι η 7άδα: $M = (Q, \Sigma, \Gamma, q_0, Z_0, \delta, F)$ όπου:

- $Q = \{q_0, q_1, q_2\}$
- $\Sigma = \{0,1\}$
- $\Gamma = \{Z_0, 0, 1, S\}$
- q_0 είναι η αρχική κατάσταση
- Z_0 είναι το αρχικό σύμβολο του σωρού
- δ είναι η συνάρτηση μετάβασης που περιγράφεται από τον ακόλουθο πίνακα μετάβασης.
- $F = \{q_2\}$

Ο πίνακας μετάβασης είναι:

Αριθμός	Κατ/ση	Σύμβολο Εισόδου	Σύμβολο Σωρού	Κίνηση	Επεξήγηση
1	q_0	ε	Z_0	(q_1, SZ_0)	Αρχικοποίηση
2.1	q_1	ε	S	$(q_1, 0S1)$	Κανόνας $S \to 0S1$
2.2	q_1	ε	S	(q_1, ε)	Κανόνας $S \to \varepsilon$
3.1	q_1	0	0	(q_1, ε)	Ταίριασμα 0
3.2	q_1	1	1	(q_1, ε)	Ταίριασμα 1
4	q_1	ε	Z_0	(q_2, Z_0)	Αποδοχή
Οι υπόλοιποι συνδυασμοί					ΤΙΠΟΤΑ

ΚΛΕΙΣΤΟΤΗΤΕΣ ΠΡΑΞΕΩΝ στις Γ.Χ.Σ.

ΓΛΩΣΣΕΣ ΧΩΡΙΣ ΣΥΜΦΡΑΖΟΜΕΝΑ www.psounis.gr

Θεώρημα: Οι **γλώσσες χωρίς συμφραζόμενα** είναι κλειστές στις πράξεις: **Ένωση, Παράθεση, Αστέρι Kleene.**

<u>**Κλειστότητα** των Γ.Χ.Σ στην **Ένωση**</u>

- Η L_1 είναι Ανεξάρτητη Συμφραζομένων, άρα παράγεται από μία γραμματική χωρίς συμφραζόμενα, έστω με αρχικό κανόνα S_1. Η L_2 είναι Ανεξάρτητη Συμφραζομένων, άρα παράγεται από μία γραμματική χωρίς συμφραζόμενα, έστω με αρχικό κανόνα S_2
- Η $L_1 \cup L_2$ παράγεται από την γραμματική χωρίς συμφραζόμενα με αρχικό κανόνα $\boldsymbol{S \to S_1 \mid S_2}$ άρα είναι χωρίς συμφραζόμενα

<u>**Κλειστότητα** των Γ.Χ.Σ στην **Παράθεση**</u>

- Η L_1 είναι Ανεξάρτητη Συμφραζομένων, άρα παράγεται από μία γραμματική χωρίς συμφραζόμενα, έστω με αρχικό κανόνα S_1. Η L_2 είναι Ανεξάρτητη Συμφραζομένων, άρα παράγεται από μία γραμματική χωρίς συμφραζόμενα, έστω με αρχικό κανόνα S_2
- Η $L_1 L_2$ παράγεται από την γραμματική χωρίς συμφραζόμενα με αρχικό κανόνα $\boldsymbol{S \to S_1 S_2}$ άρα είναι χωρίς συμφραζόμενα

<u>**Κλειστότητα** των Γ.Χ.Σ στο **Αστέρι Kleene**</u>

- Η L είναι Ανεξάρτητη Συμφραζομένων, άρα παράγεται από μία γραμματική χωρίς συμφραζόμενα, έστω με αρχικό κανόνα S_1
- Η L* παράγεται από την γραμματική χωρίς συμφραζόμενα με αρχικό κανόνα $\boldsymbol{S \to S_1 S \mid \varepsilon}$ άρα είναι χωρίς συμφραζόμενα.

Θεώρημα: Οι **γλώσσες χωρίς συμφραζόμενα ΔΕΝ** είναι κλειστές στις πράξεις: **Συμπλήρωμα, Τομή**

<u>**ΟΧΙ Κλειστότητα** των Γ.Χ.Σ στο **Συμπλήρωμα.**</u>
<u>ΑΝΤΙΠΑΡΑΔΕΙΓΜΑ:</u>Πράγματι αν:

- $L1 = \{w \in \{a,b,c\}^* \mid w$ δεν έχει ίσα a και b$\}$
- $L2 = \{w \in \{a,b,c\}^* \mid w$ δεν έχει ίσα b και c$\}$

που είναι και οι δύο χωρίς συμφραζόμενα (γιατί;).
Τότε η ένωση τους είναι η γλώσσα
$L' = \{w \in \{a,b,c\}^* \mid w$ δεν έχει ίσα a και b ή δεν έχει ίσα b και c $\}$
και είναι γλώσσα χωρίς συμφραζόμενα (κλειστότητα της ένωσης στις ΓΧΣ).
Τότε το συμπλήρωμα της L' είναι η γλώσσα:
$\overline{L'} = \{w \in \{a,b,c\}^* \mid w$ έχει ίσα a, b και c $\}$
που δεν είναι χωρίς συμφραζόμενα (αποδεικνύεται με το λήμμα της άντλησης για γλώσσες χωρίς συμφραζόμενα).

<u>**ΟΧΙ Κλειστότητα** των Γ.Χ.Σ στην **Τομή.**</u>
<u>ΑΝΤΙΠΑΡΑΔΕΙΓΜΑ:</u>Πράγματι αν:

- $L_1 = \{a^n b^n c^m \mid n, m \geq 0\}$
- $L_2 = \{a^m b^n c^n \mid n, m \geq 0\}$

Που είναι και οι δύο χωρίς συμφραζόμενα (έχουν γραμματική χωρίς συμφραζόμενα)
Η τομή τους είναι η γλώσσα:

$$L_1 \cap L_2 = \{a^n b^n c^n \mid n \geq 0\}$$

που όπως θα δούμε στο επόμενο μάθημα δεν είναι χωρίς συμφραζόμενα (αποδεικνύεται με το λήμμα της άντλησης για γλώσσες χωρίς συμφραζόμενα).

ΛΗΜΜΑ ΑΝΤΛΗΣΗΣ για Γ.Χ.Σ.

ΓΛΩΣΣΕΣ ΧΩΡΙΣ ΣΥΜΦΡΑΖΟΜΕΝΑ www.psounis.gr

<u>Το Λήμμα Άντλησης για Γλώσσες Ανεξάρτητες Συμφραζομένων</u>
Έστω L μια άπειρη γλώσσα ανεξάρτητη συμφραζομένων. Τότε υπάρχει ένας αριθμός n (μήκος άντλησης) τέτοιος ώστε κάθε $s \in L$ με $|s| \geq n$ να μπορεί να γραφεί στην μορφή $s = uvwxy$ όπου για τις συμβολοσειρές u, v, w, x και y ισχύει:

- ➢ $|vwx| \leq n$
- ➢ $|vx| > 0$
- ➢ $uv^m wx^m y \in L$ για κάθε φυσικό $m \geq 0$

(1) Επιλέγουμε μια **συμβολοσειρά s** που ανήκει στην γλώσσα που
- (α) <u>**όλα** τα σύμβολα είναι </u>υψωμένα τουλάχιστον στην p
- (β) ανήκει οριακά στην γλώσσα

(2) Υπολογίζουμε το μήκος της συμβολοσειράς που επιλέξαμε στο (1)

$L_1 = \{0^n 1^n 2^n \mid n \geq 0\}$ – ΑΠΟΔΕΙΞΗ
Η L είναι άπειρη. Υποθέτουμε ότι είναι ανεξάρτητη συμφραζομένων. Έστω p το μήκος άντλησής της.
Η συμβολοσειρά $s = 0^p 1^p 2^p$ ανήκει στην γλώσσα και έχει μήκος $3p \geq p$. Η συμβολοσειρά μπορεί να γραφεί στην μορφή $s = uvwxy$ με τις ιδιότητες του λήμματος άντλησης.
Επειδή $|vwx| \leq p$ και $|vx| > 0$ έπεται ότι τουλάχιστον ένα από τα v,x θα περιέχει τουλάχιστον ένα σύμβολο. Διακρίνω τις περιπτώσεις για τα v,x:

1. Να περιέχουν μόνο 0. Τότε $uv^2wx^2y \notin L$, διότι προστίθενται μηδενικά άρα π.χ. τα 0 δεν είναι ίσα με τα 1
2. Να περιέχουν 0 και 1. Τότε $uv^2wx^2y \notin L$, διότι προστίθενται μηδενικά και άσσοι άρα π.χ. τα 0 δεν είναι ίσα με τα 2
3. Να περιέχουν μόνο 1. Τότε $uv^2wx^2y \notin L$, διότι προστίθενται άσσοι άρα π.χ. τα 1 δεν είναι ίσα με τα 2
4. Να περιέχουν 1 και 2. Τότε $uv^2wx^2y \notin L$, διότι προστίθενται άσσοι και δυάρια άρα π.χ. τα 2 δεν είναι ίσα με τα 0
5. Να περιέχουν μόνο 2. Τότε $uv^2wx^2y \notin L$, διότι προστίθενται δυάρια άρα π.χ. τα 2 δεν είναι ίσα με τα 0.

Άτοπο από το λήμμα άντλησης. Συνεπώς η γλώσσα δεν είναι ανεξάρτητη συμφραζομένων.

(3) Εντοπίζουμε περιπτώσεις ανάλογα με το που περιέχεται το vwx δεδομένου ότι έχει μήκος το πολύ p. Χρήσιμο θα φανεί να κάνουμε ένα βοηθητικό σχήμα (βλέπε δεξιά)

$$s = \overbrace{00\ldots00}^{p}\,\overbrace{11\ldots11}^{p}\,\overbrace{22\ldots22}^{p}$$