



ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ
TECHNICAL UNIVERSITY OF CRETE

ΉΜΜΥ

ΠΛΗ 402 - Θεωρία Υπολογισμού

1η Σειρά Ασκήσεων

Συγγραφέας:

Βασιλειάδης Σταύρος

ΑΜ:

2019030023

Διδάσκων:

Λαγουδάκης Μιχαήλ

12 Μαΐου 2025

Περιεχόμενα

Περιεχόμενα	2
1 Κατάλογος συμβόλων και συντομογραφίες	i
2 Λύσεις Ασκήσεων	1
1 Άσκηση 1η - Γραμματικές χωρίς συμφραζόμενα:	2
1 - α Απάντηση Υποερωτήματος (α)	3
1 - β Απάντηση Υποερωτήματος (β)	5
2 Άσκηση 2η - Αυτόματα στοίβας:	9
2 - α Απάντηση Υποερωτήματος (β)	10
2 - β Απάντηση Υποερωτήματος (β)	13
3 Άσκηση 3η - Γλώσσες χωρίς συμφραζόμενα	16
3 - α Απάντηση Υποερωτήματος (α)	17
3 - β Απάντηση Υποερωτήματος (β)	18
4 Άσκηση 4η - Αναγνώριση γλωσσών χωρίς συμφραζόμενα:	19
4 - α Απάντηση Υποερωτήματος (α)	20
4 - β Απάντηση Υποερωτήματος (β)	24

ΜΕΡΟΣ 1

Κατάλογος συμβόλων και συντομογραφίες

ΜΕΡΟΣ 1. ΚΑΤΑΛΟΓΟΣ ΣΥΜΒΟΛΩΝ ΚΑΙ ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ

- \mathcal{L} Σύνολο γλωσσών με κάποιες κοινές ιδιότητες (πχ βλ παρακάτω)
- \mathcal{L}_{reg} Σύνολο κανονικών γλωσσών
- \mathcal{L}_{irr} Σύνολο μη κανονικών γλωσσών
- \mathcal{L}_{fin} Σύνολο πεπερασμένων γλωσσών
- $\approx L$ Ισοδυναμία γλώσσας κατά Myhill-Nerode
- $\sim M$ Ισοδυναμία DFA κατά Myhill-Nerode
- Σ Αλφάβητο
- Q ή K Σύνολο καταστάσεων (αυτομάτου)
- s Αρχική κατάσταση (αυτομάτου)
- F Σύνολο τελικών καταστάσεων (αυτομάτου)
- Δ Σύνολο σχέσεων μεταβάσεων (αυτομάτου)
- w, x, y, z Συμβολοσειρά / Υποσυμβολοσειρές
- DFA Ντετερμινιστικό Πεπερασμένο Αυτόματο $M = (Q, \Sigma, \delta, s, F)$
- NFA Μη Ντετερμινιστικό Πεπερασμένο Αυτόματο
- $\varepsilon - NFA$ Μη Ντετερμινιστικό Πεπερασμένο Αυτόματο με κενές μεταβάσεις
- M Αυτόματο
- M' Πρότυπο/Ελάχιστο Αυτόματο
Εναλλακτικά DFA παραγόμενο ενός άλλου DFA / NFA.

- \mathbb{N}^0 Σύνολο φυσικών αριθμών (συμπεριλαμβανομένου και του μηδενός)
- \aleph_0 Πληθάριθμος συνόλου φυσικών αριθμών
Ο ελάχιστος μετρήσιμα άπειρος αριθμός
- $|w|$ πληθάριθμος συνόλου w / συμβολοσειράς w
- $|w|_c$ πληθάριθμος συνόλου w / συμβολοσειράς w για σύμβολο c
- $|w|_{min}$ ελάχιστος δυνατός πληθάριθμος συνόλου w / συμβολοσειράς w
- c Ακριβώς μία εμφάνιση συμβόλου c .
Αντίστοιχα για $w = cc..c$, $|w|_c = n$ τότε ακριβώς n εμφανίσεις
συμβόλου c
- c^+ Τουλάχιστον μία εμφάνιση συμβόλου c
- c^* Τουλάχιστον μηδέν εμφανίσεις συμβόλου c
- ε ή e Κενό σύμβολο
- \exists Υπάρχει
- \forall Για όλα
- \in Ανήκει στο
- \dots "και ούτω καθεξής", συνέχιση ακολουθίας βάσει εμφανούς μοτίβου
- $| \text{ ή } : \text{ ή } \ni$ Έτσι ώστε
- $\mathcal{P}(Q)$ Δυναμοσύνολο συνόλου Q .
Για $Q = \{a, b, c\} \xrightarrow{\emptyset \subseteq \text{all sets}} \mathcal{P}(Q) = \{\emptyset, a, b, c, ab, ac, bc, abc\}$
Συνεπάγεται ότι $|\mathcal{P}(Q)| = 2^{|Q|}$

- \mathcal{P} ή \mathcal{Q} κοκ Λογικές προτάσεις
- T / F (Λογικό) Αληθές / Ψευδές
- \cup Ένωση
- \cap Τομή
- $-$ ή $/$ Διαφορά
- \subseteq Υποσύνολο
- \subset Γνήσιο υποσύνολο
- \wedge Λογική Σύζευξη
- \vee Λογική Διάζευξη
- \neg Λογική Άρνηση
- Σ^* Σύνολο πεπερασμένου μήκους συμβολοσειρών αλφάβητου Σ
συμπεριλαμβανομένης της κενής συμβολοσειράς ε

ΜΕΡΟΣ 2

Λύσεις Ασκήσεων

1 Άσκηση 1η - Γραμματικές χωρίς συμφραζόμενα:

[30%] Κατασκευάστε γραμματικές χωρίς συμφραζόμενα για τις παρακάτω γλώσσες:

(α) $L_1 = \{a^n b^m : n, m \in \mathbb{N}_k, 3n \leq m \leq 6n\}$

(β) $L_2 = \{w \in \{a, b\}^* : \eta \ w \text{ περιέχει ακριβώς } 2 \text{ περισσότερα } a \text{ από το τριπλάσιο του πλήθους των } b\}$

Για κάθε μία από τις συμβολοσειρές $aabbbbbbb \in L_1$, $aabaaabaaa \in L_2$ και την αντίστοιχη γραμματική, δώστε το αντίστοιχο συντακτικό δέντρο.

1 - α Απάντηση Υποερωτήματος (α)

Έχουμε CFL $L_1 = \{a^n, b^m : n, m \in \mathbb{N}_0, 3n \leq m \leq 6n\}$ άρα $|w|_{\min} = 0$ για $n = 0$ με a να προηγούνται των b . Οπότε: $\xrightarrow{\text{Ελάχιστη περίπτωση}} \varepsilon$ και $\xrightarrow{\text{Ελάχιστη μή κενή περίπτωση}} abbb$.

Αλλά με το ίδιο a έχουμε επίσης: $\rightarrow abbb b, abbb bb, abbb bbb$.

Δεν πρέπει όμως να ξεχνάμε ότι μπορούμε να έχουμε έως $|w|_a = \aleph_0$ και αντίστοιχα τα b που αναλογούν σε αυτά για να είναι σωστή η συμβολοσειρά. Οτιδήποτε όμως από αυτά είναι ουσιαστικά απλή επανάληψη του κανόνα που δημιουργεί όλες τις ενός a συμβολοσειρές. Άρα δεν χρειαζόμαστε κάποιο επιπρόσθετο κανόνα παραγωγής αλλά έχουμε αναδρομή.

Συμψηφίζοντας τώρα αυτές τις παρατηρήσεις μπορούμε να περάσουμε στην κατασκευή των κανόνων παραγωγής. Θα το προσεγγίσουμε σταδιακά:

- Ελάχιστη περίπτωση:

$$S \rightarrow \varepsilon$$

- Ελάχιστα a :

$$S \rightarrow aM \mid \varepsilon$$

$$M \rightarrow bbb \mid bbbb \mid bbbbbb \mid bbbbbb$$

- $|w| = \aleph_0$:

$$S \rightarrow aSM \mid \varepsilon$$

$$M \rightarrow bbb \mid bbbb \mid bbbbbb \mid bbbbbb$$

- Πιο ευανάγνωστη παραλλαγή:

$$S \rightarrow aSbbM \mid \varepsilon$$

$$M \rightarrow b \mid bb \mid bbb \mid bbbb$$

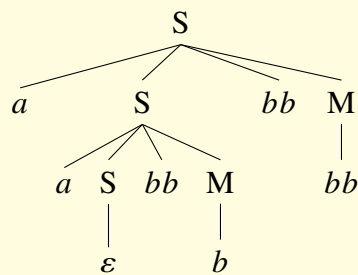
- Καμία αριστερή αναδρομή.

Γραμματική και συντακτικό δένδρο

Είμαστε έτοιμοι να προχωρήσουμε στην πλήρη περιγραφή της γραμματικής μας:

$$G_1 = (\{a, b\}, \{\varepsilon\}, \{S \rightarrow aSbbM \mid \varepsilon, M \rightarrow b \mid bb \mid bbb \mid bbbb\}, S)$$

Ήρθε η στιγμή να δώσουμε το συντακτικό δένδρο για συμβολοσειρά (τα κενά για ευκολότερη ανάγνωση) $w = aa\ bbb\ bbb\ b$, $|w| = 9$, $|w|_a = 2$, $|w|_b = 7$:



Που δίνει:

$a \quad a \quad \quad bb \quad b \quad bb \quad bb$

Εναλλακτικά θα μπορούσαμε να το διαβάσουμε ως $a(a(\varepsilon)bb(bb))bb(b)$.

1 - β Απάντηση Υποερωτήματος (β)

$\text{CFL } L_1 = \{w \in \{a, b\}^* : |w|_a = 3|w|_b + 2\}$ άρα $|w|_{\min} = 2$ και συγκεκριμένα aa .

Στην περίπτωση όμως που έχουμε έστω και ένα b έχουμε: $aaa aa b$.

Παρατηρούμε ότι η γλώσσα δεν ορίζει συγκεκριμένη σειρά μεταξύ a και b και άρα θα πρέπει να μπορούν να παραχθούν όλοι οι συνδυασμοί και με βάση τη θέση των συμβόλων και άρα επιπρόσθετες συμβολοσειρές: $b aaa aa, a b aaa a, \dots, aa aaa b$.

Για να οδηγηθούμε σε κάποιες από αυτές το ελάχιστο aa διαχωρίστηκε, με το b να παρεμβάλλεται ανάμεσα. Με βάση την ελάχιστη περίπτωση καταλαβαίνουμε ότι παρόλα αυτά, τα δύο πρόσθετα a θα πρέπει να απομονωθούν σε δικό τους κανόνα.

Έχουμε όμως $|w|_b \in \mathbb{N}_0$ και άρα τα ανάλογα a , υπονοώντας χρήση αναδρομής και συγκεκριμένα όλων των κανόνων παραγωγής εκτός του ελάχιστου aa .

Συμψηφίζοντας μπορούμε να περάσουμε στην κατασκευή των κανόνων παραγωγής:

- Ελάχιστη περίπτωση:

$$S \rightarrow aa$$

- Ελάχιστα b :

$$S \rightarrow aaR \mid aRa \mid Raa$$

$$R \rightarrow baaa \mid abaa \mid aaba \mid aaab \mid \varepsilon$$

- $|w| = \mathbb{N}_0$:

$$S \rightarrow aaR \mid aRa \mid Raa$$

$$R \rightarrow RbRaaa \mid aRbRaa \mid aaRbR \mid aaaRbR \mid \varepsilon$$

- Σύμπτυξη:

$$S \rightarrow RaRaR$$

$$M \rightarrow bRaaa \mid abRaa \mid aabRa \mid aaaRbR \mid \varepsilon$$

Για να καταλάβουμε για πιο λόγο έφυγε η αναδρομή πριν από το b σε όλες τις παραλλαγές του R εκτός την $aaaRbR$ και άρα να υπάρχει κάποια αιτιολόγηση θα πρέπει να δώσουμε κάποια βασικά παραδείγματα:

- $bb\ aaa\ aaa\ aa$: $S \rightarrow RaRaR \rightarrow (baaa)aRaR \rightarrow (bRaaa)aRaR \rightarrow (b(bRaaa)aaa)aRaR \rightarrow (b(b\epsilon aaa)aaa)aRaR \rightarrow (bb\ aaa\ aaa)a\epsilon a\epsilon \rightarrow bb\ aaa\ aaa\ aa$
- $a\ bb\ aaa\ aaa\ a$: $S \rightarrow RaRaR \rightarrow \epsilon aRa\epsilon \rightarrow a(bRaaa)a \rightarrow a(b(bRaaa)aaa)a \rightarrow a(b(b\epsilon aaa)aaa)a \rightarrow a\ bb\ aaa\ aaa\ a$
- $aa\ bbaaa\ aaa$: $S \rightarrow RaRaR \rightarrow \epsilon a\epsilon aR \rightarrow aa(bRaaa) \rightarrow aa(b(bRaaa)aaa) \rightarrow aa(b(b\epsilon aaa)aaa) \rightarrow aa\ bb\ aaa\ aaa$
- ...
- $aa\ aaa\ bb\ aaa$: $S \rightarrow RaRaR \rightarrow \epsilon a\epsilon aR \rightarrow aa(aaabR) \rightarrow aa(aaab(bRaaa)) \rightarrow aa(aaab(b\epsilon aaa)) \rightarrow aa\ aaa\ bb\ aaa$
- ...
- $aaa\ aaa\ bb\ aa$: $S \rightarrow RaRaR \rightarrow Ra\epsilon a\epsilon \rightarrow (aaab)aa \rightarrow (aaaRb)aa \rightarrow (aaa(aaabR)b)aa \rightarrow (aaa(aaab\epsilon b)b)aa \rightarrow aaa\ aaa\ bb\ aa$
- Είδαμε ότι για τις συγκεκριμένες υποπεριπτώσεις χρειαζόμαστε $bRaaa$, $aaaRb$ και $aaabR$. Θυμόμαστε ότι τα R δεν είναι παρά αναδρομές στον ίδιο κανόνα παραγωγής και παρατηρούμε ότι το $aaa(R)b$ και το $aaab(R)$ είναι ουσιαστικά η ίδια (σταθερή) συμβολοσειρά απλά με αναδρομή είτε πριν, είτε μετά το b . Άρα

μπορούμε να αντικαταστήσουμε και τα δύο με το να πάρουμε τα σταθερά σύμβολα, όπου είναι κοινά και να τοποθετήσουμε αναδρομή και στις δύο πλευρές του b παίρνοντας έτσι $aaaRbR$.

- Για να σιγουρευτούμε για το αν όντως δεν χρειάζονται κανόνες τύπου $RbRaaa$, $aRbRaa$, $aaRbRa$, $aaaRbR$, έχουμε αναλύσει όλες τις παραλλαγές για ορθές συμβολοσειρές με δύο " b " και σταθερές υποσυμβολοσειρές bb , bab , $baab$, ..., $baaaaaaaaaab$, δλδ με τα υπόλοιπα a να μετακινούνται σε όλες τις πιθανές ακραίες θέσεις, καθώς και άλλες, συμπεριλαμβανομένων και κάποιων βασικών μη ορθών, αλλά δεν τις συμπεριλαμβάνουμε για προφανείς λόγους (άνω των 150 παραλλαγών). Αυτό δεν σημαίνει ότι δεν γίνεται να μας έχει ξεφύγει κάτι.
- Τέλος αν πούμε ότι η CFG μας δεν περιέχει αριστερή αναδρομή.

Γενικά θα πρέπει να έχουμε κατά νου ότι είναι πάρα πολύ χρήσιμο έως και απαραίτητο, εκτός των ελάχιστων περιπτώσεων βάση συνόλου ($\pi\chi \in \mathbb{N}_0$ ή και ελάχιστων περιπτώσεων για μη μηδενικό σύνολο (ουσιαστικά λαμβάνοντας υπόψιν αποδοχή κενής συμβολοσειράς εάν επιτρέπεται και ελάχιστης μή κενής), καλό θα είναι να ελέγχουμε και για διπλές εμφανίσεις των ελάχιστων μή κενών $\pi\chi$ $\{\forall w \neq \varepsilon \mid w_{min} = xy, \exists w_{double} = xxxy \cup xyxy \cup xygx \cup yxxy \cup yxyx \cup ygyx\}$

Υπόψιν ότι εκτός κάποιων πολύ απλών περιπτώσεων, στις CFL δεν μπορούμε πάντα να είμαστε σίγουροι ότι η CFG μας καλύπτει όλες τις πιθανές ορθές συμβολοσειρές της γλώσσας ή ότι είναι ελάχιστη.

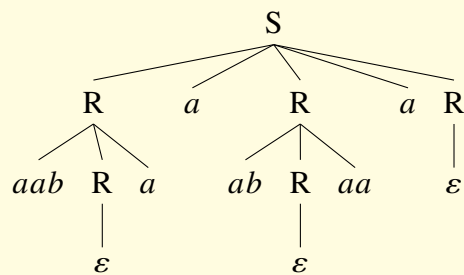
Γραμματική και συντακτικό δένδρο

Είμαστε έτοιμοι να προχωρήσουμε στην πλήρη περιγραφή της γραμματικής μας:

$$G_2 = (\{a, b\}, \{aa, \varepsilon\}, \{S \rightarrow RaRaR$$

$$R \rightarrow bRa aa \mid abRa a \mid aabRa \mid aaaRbR \mid \varepsilon\}, S)$$

Ήρθε η στιγμή να δώσουμε το συντακτικό δένδρο για συμβολοσειρά (τα κενά για ευκολότερη ανάγνωση) $w = aa \ b \ aaa \ b \ aaa$, $|w| = 10$, $|w|_a = 8$, $|w|_b = 2$:



Που δίνει: $aab \quad a \ a \ ab \quad aa \ a$

Εναλλακτικά θα μπορούσαμε να το διαβάσουμε ως $(\varepsilon) \ a \ (ab(\varepsilon)aa) \ a \ (b(\varepsilon)aaa)$.

2 Άσκηση 2η - Αυτόματα στοίβας:

[30%] Κατασκευάστε αυτόματα στοίβας για τις παρακάτω γλώσσες:

$$(\alpha) L_1 = \{b^n a^k b^k a^n : k, n \in \mathbb{N}^0\}$$

$$(\beta) L_2 = \{w \in \{a, b\}^* : \eta \ w \text{ περιέχει διπλάσιο αριθμό } b \text{ απ'ότι } a\}$$

Για κάθε μία από τις συμβολοσειρές $bbbabaaa \in L_1$, $abbbab \in L_2$ δώστε έναν υπολογισμό αποδοχής στο αντίστοιχο αυτόματο στοίβας χρησιμοποιώντας συνολικές καταστάσεις.

2 - α Απάντηση Υποερωτήματος (β)

Έχουμε να κατασκευάσουμε PDA για CFL $L_1 = \{b^n a^k b^k a^n : k, n \in \mathbb{N}^0\}$

Από τον τύπο δυνάμεθα να εξάγουμε τα εξής δεδομένα, για συμβολοσειρά $w \in L_1$:

- $|w|_{\min} = 0 \implies n, k = 0 \implies w = \varepsilon, \quad |w|_{\max} = \aleph_0, \quad |w| = 2(n+k) = \text{Άρτιο}$
- $n+k-2nk=1 \implies |w|_{\min}=2 \implies w = ba \cup ab$
- $n, k=1 \implies |w|_{\min}=4 \implies w = baba, \quad L \subseteq \{xx^R | x \in \Sigma^*\}$
- Για μεγαλύτερα n, k έχουμε συμβολοσειρές όπως $bbabaa, baabba, bbaabbaa...$

Από αυτά κάποια δεν μας λένε κάτι ιδιαίτερα χρήσιμο, άλλα όμως μας φανερώνουν την οδό που θα πρέπει να ακολουθήσουμε. Για παράδειγμα, κατά αντιστοιχία με τις αντίστοιχες τεχνικές που ακολουθήσαμε στις λύσεις ασκήσεων κανονικών εκφράσεων, τα διάφορα ελάχιστα μας δίνουν την βάση πάνω στην οποία θα χτιστεί το (PD) αυτόματο.

Ως εκ τούτου έχουμε φαινομενικά τέσσερις ελάχιστες περιπτώσεις, που η συγκεκριμένη γλώσσα επιτρέπει: $w_0 = \varepsilon \vee w_1 = ab \vee w_2 = ba \vee w_3 = baba$ αλλά προσοχή, το $w = baba$ αποτελεί παρεμβολή του w_1 εντός του w_2 και όχι $(ba)^*$

Αυτό με τη σειρά του δείχνει ότι ουσιαστικά, το w_3 αποτελεί επέκταση της w_2 μετά αρχικού b και άρα δεν είναι μέρος της καθολικής βάσης του αυτομάτου. Θα μας χρειαστεί όμως, αφού δείχνει ότι όταν ο αλγόριθμος θα μπαίνει στη διαδρομή που διαβάζει ba , θα διακλαδώνεται σε δύο εναλλακτικές βάση του αν αυτό που έπεται του b είναι a ή aba .

Τέλος αφού $n, k \in \mathbb{N}_*$ τότε μπορούν όλοι αυτοί οι χαρακτήρες να είναι (μετρήσιμα) άπειροι αλλά υπό τον όρο ότι τηρείται το γενικό μοτίβο.

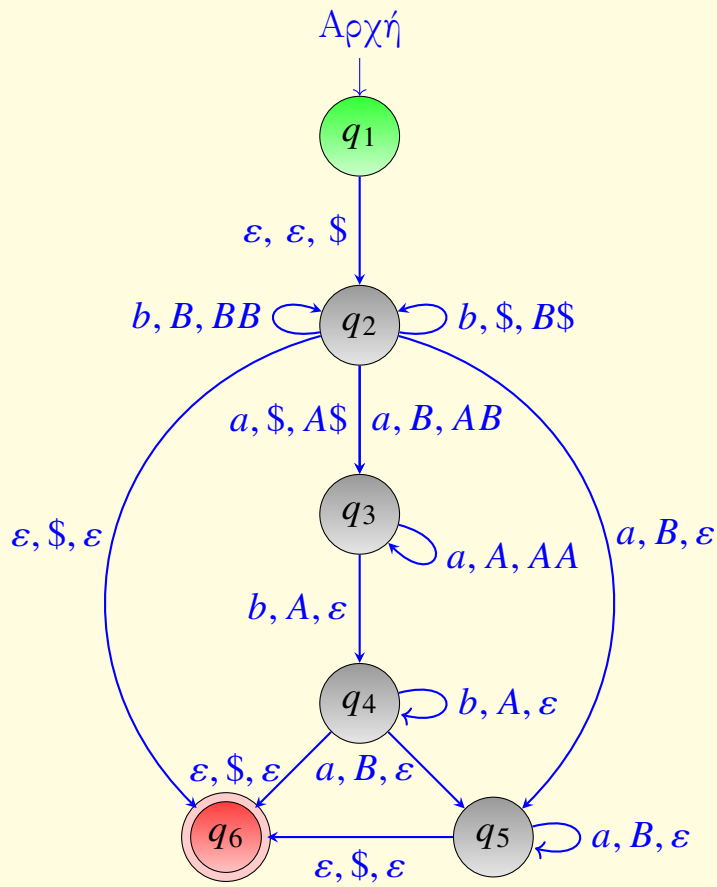
Άρα, υποθετικά, έχουμε τρεις αρχικούς κλάδους και σε έναν από αυτούς μία διάσπαση σε δύο. Θα δούμε ότι ακριβώς αυτό συμβαίνει και στην πράξη (τουλάχιστον στη συγκεκριμένη επίλυση) και ότι γενικά είναι καλή πρακτική. Ξεκινάμε κατασκευή αυτομάτου, αρχικά δίνοντας μαθηματική περιγραφή και κατόπιν αντίστοιχο διάγραμμα:

PDA $M = (K, \Sigma, \Gamma, \Delta, s, F)$

- $K = \{q_1, q_2, q_3, q_4, q_5, q_6\}$
- $\Sigma = \{a, b\}$
- $\Gamma = \{A, B, \$\}$
- $\Delta =$

1. $(q_1, \varepsilon, \varepsilon) \rightarrow (q_2, \$)$	8. $(q_3, a, A) \rightarrow (q_3, AA)$
2. $(q_2, b, \$) \rightarrow (q_2, B\$)$	9. $(q_3, b, A) \rightarrow (q_4, \varepsilon)$
3. $(q_2, b, B) \rightarrow (q_2, BB)$	10. $(q_4, b, A) \rightarrow (q_1, \varepsilon)$
4. $(q_2, \varepsilon, \$) \rightarrow (q_6, \varepsilon)$	11. $(q_4, a, B) \rightarrow (q_5, \varepsilon)$
5. $(q_2, a, \$) \rightarrow (q_3, A\$)$	12. $(q_4, \varepsilon, \$) \rightarrow (q_6, \varepsilon)$
6. $(q_2, a, B) \rightarrow (q_3, AB)$	13. $(q_5, a, B) \rightarrow (q_5, \varepsilon)$
7. $(q_2, a, B) \rightarrow (q_5, \varepsilon)$	14. $(q_5, \varepsilon, \$) \rightarrow (q_6, \varepsilon)$
- $s = q_1$
- $F = q_6$

DFA $M_{\text{τελικό}}$



Υπολογισμός αποδοχής $bbbabaaa$

$(q1, bbbabaaa, \varepsilon) \vdash (q2, bbbabaaa, \$) \vdash (q2, bbabaaa, B\$) \vdash (q2, babaaa, BB\$)$

$$\vdash (q2, abaaa, BBB\$) \vdash \begin{cases} (q3, baaa, ABBS\$) \\ (q5, baaa, BB\$) \neq \text{Αποτυχία Κλώνου} \end{cases}$$

$\vdash (q4, aaa, BBB\$) \vdash (q5, aa, BB\$) \vdash (q5, a, B\$) \vdash (q5, \varepsilon, \$)$

$\vdash (q6, \varepsilon, \varepsilon)$ Επιτυχής Ανάγνωση

2 - β Απάντηση Υποερωτήματος (β)

Έχουμε να κατασκευάσουμε PDA για CFL $L_2 = \{w \in \{a, b\}^* : |w|_b = 2|w|_a\}$
Από τον τύπο δυνάμεθα να εξάγουμε τα εξής δεδομένα, για συμβολοσειρά $w \in L_1$:

- $|w|_{min} = 0 \implies w_{min} = \varepsilon, \quad |w|_{max} = \aleph_0, \quad |w| \equiv |w|_a \pmod{2}$
- $\{w \in \{a, b\}^+ \mid |w| = 3, |w|_a = 1, |w|_b = 2\} \implies |w| = 3|w|_a$
- $|w| = 3 : abb, bab, bba$
- $|w| > 3 : aabbbb, bbbbaa, abbbba, babbab, ababbb, bbbaba, bbaabb, \dots$
- Κανένας περιορισμός οποιουδήποτε συμβόλου ως προς την θέση.

Έχουμε φαινομενικά τέσσερεις ελάχιστες περιπτώσεις, που η συγκεκριμένη γλώσσα επιτρέπει: $w_0 = \varepsilon \vee w_1 = abb \vee w_2 = bab \vee w_3 = bba$ αλλά όπως ήδη είπαμε δεν υπάρχει περιορισμός ως προς την θέση αντίθετα με το προηγούμενο πρόβλημα στην άσκηση 2.1 (υποερώτημα (α)).

Αυτό με τη σειρά του δείχνει ότι πιθανώς να μην χρειαστεί κάποια διακλάδωση και να μπορούν να γίνουν όλες οι διαδικασίες αναδρομικά στον ίδιο κόμβο.

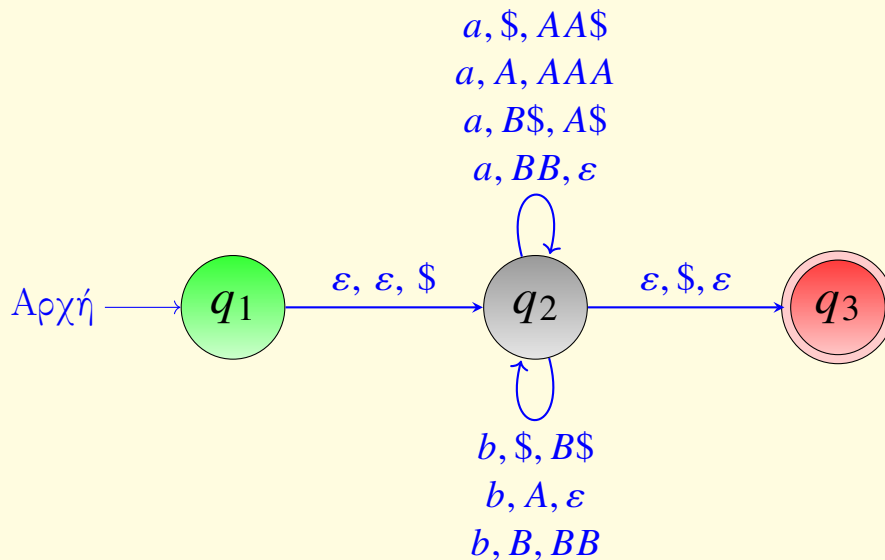
Άρα, υποθέτουμε ένα κόμβο για όλη την ανάγνωση βασιζόμενοι σε σύνθετες αναδρομές, συν επιπρόσθετους βοηθητικούς κόμβους (Εναρκτήριο με μετάβαση αρχικοποίησης σωρού με δείκτη τελευταίου κελιού και τελικό με μετάβαση προς αυτόν όπου αδειάζει τον δείκτη από τον σωρό). Ξεκινάμε την κατασκευή του αυτομάτου, αρχικά δίνοντας την μαθηματική περιγραφή και κατόπιν το αντίστοιχο διάγραμμα:

PDA $M = (K, \Sigma, \Gamma, \Delta, s, F)$

- $K = \{q_1, q_2, q_3\}$
- $\Sigma = \{a, b\}$
- $\Gamma = \{A, B, \$\}$
- $\Delta =$

1. $(q_1, \varepsilon, \varepsilon) \rightarrow (q_2, \$)$	6. $(q_2, a, A) \rightarrow (q_2, AAA)$
2. $(q_2, b, \$) \rightarrow (q_2, B\$)$	7. $(q_2, a, BB) \rightarrow (q_2, \varepsilon)$
3. $(q_2, b, B) \rightarrow (q_2, BB)$	8. $(q_2, a, B\$) \rightarrow (q_2, A\$)$
4. $(q_2, b, A) \rightarrow (q_2, \varepsilon)$	9. $(q_2, \varepsilon, \$) \rightarrow (q_3, \varepsilon)$
5. $(q_2, a, \$) \rightarrow (q_2, AA\$)$	
- $s = q_1$
- $F = q_3$

DFA $M_{\text{τελικό}}$



Υπολογισμός αποδοχής *abbbab*

$(q1, abbbab, \varepsilon) \vdash (q2, abbbab, \$) \vdash (q2, bbbab, AA\$) \vdash (q2, bbab, A\$)$

$\vdash (q2, bab, \$) \vdash (q2, ab, B\$) \vdash (q2, b, A\$) \vdash (q2, \varepsilon, \$)$

$\vdash (q3, \varepsilon, \varepsilon)$ Επιτυχής Ανάγνωση

3 Άσκηση 3η - Γλώσσες χωρίς συμφραζόμενα

(α) [5%] Αποφανθείτε αν ο παρακάτω ισχυρισμός είναι σωστός ή λανθασμένος και αιτιολογήστε: Η τομή μίας ντετερμινιστικής γλώσσας χωρίς συμφραζόμενα με μία πεπερασμένη γλώσσα είναι πάντα γλώσσα χωρίς συμφραζόμενα.

(β) [5%] Αποδείξτε εάν η παρακάτω γλώσσα είναι ή δεν είναι γλώσσα χωρίς συμφραζόμενα:

$$L = \{a^m b^n c^k : n, k, m \in \mathbb{N}_0, n = 3m + 2k\}$$

3 - α Απάντηση Υποερωτήματος (α)

Η πρόταση "η τομή μίας ντετερμινιστικής γλώσσας χωρίς συμφραζόμενα με μία πεπερασμένη γλώσσα είναι πάντα γλώσσα χωρίς συμφραζόμενα" είναι αληθής όπως θα αποδείξουμε παρακάτω:

Απόδειξη

$$\bullet \mathcal{P} := \{\forall L_{DCF} \in \mathcal{L}_{DCF}, \forall L_{REG} \in \mathcal{L}_{REG} : L_{DCF} \cap L_{REG} \in \mathcal{L}_{CF}\} \quad (1)$$

$$\bullet \mathcal{L}_{FIN} \subset \mathcal{L}_{REG} \subset \mathcal{L}_{DCF} \subset \mathcal{L}_{NCF} = \mathcal{L}_{CF} \quad (2)$$

$$\bullet A \subset B \Rightarrow A \cap B = A \subset B \quad (3)$$

$$\bullet A \subset B \subset \dots \subset N \Rightarrow A \subset N \quad (4)$$

$$\bullet \stackrel{(2)(4)}{\Rightarrow} \mathcal{L}_{FIN} \subset \mathcal{L}_{DCF} \quad (5)$$

$$\bullet \mathcal{L}_{FIN} \cap \mathcal{L}_{DCF} \stackrel{(3)(5)}{=} \mathcal{L}_{FIN} \subset \mathcal{L}_{DCF} \subset \mathcal{L}_{CF} \stackrel{(4)}{\Rightarrow} \mathcal{L}_{FIN} \subset \mathcal{L}_{CF} \Rightarrow \mathcal{L}_{FIN} \in \mathcal{L}_{CF} \quad (6)$$

$$\bullet \mathcal{Q} \stackrel{(6)}{:=} \{\forall L_{DCF} \in \mathcal{L}_{DCF}, \forall L_{REG} \in \mathcal{L}_{REG} : L_{DCF} \cap L_{REG} \in \mathcal{L}_{CF}\} \quad (7)$$

$$\bullet \mathcal{P}^{(1)} = \mathcal{Q}^{(7)} = \text{True} \Rightarrow \mathcal{P} = \text{True}$$

3 - β Απόδειξη Υποερωτήματος (β)

Ζητείται να αποδείξουμε ότι η $L = \{a^m b^n c^k : n, k, m \in \mathbb{N}_0, n = 3m + 2k\}$ είναι CF.

Αυτό μπορεί να γίνει μεταξύ άλλων με απόπειρα κατασκευής CFG της γλώσσας L , ενώ μία άλλη μέθοδος είναι μέσω λήμματος άντλησης κα.

Απόδειξη με απόπειρα κατασκευής CFG

- Προτεραιότητα συμβόλων: $w = \{xyz \mid x = \{a\}^m, y = \{b\}^{3m+2k}, z = \{c\}^k\}$.
- Λόγο συγκεκριμένης προτεραιότητας τα σύμβολα δεν αναμιγνύονται πχ δεν υπάρχει $bbabbc$ παρόλο που η αναλογία συμβόλων είναι ορθή). Ακριανά σύμβολα (a, c) καθορίζουν το πλήθος των μεσαίων (b). Άρα μπορούμε να σπάσουμε όλες τις συμβολοσειρές της γλώσσας σε δύο μέρη: ένα αριστερό:

$$L_{LHS} = \{w = \{a^m b^{3m}\}^* : m \in \mathbb{N}_0\}$$

και ένα δεξιό:

$$L_{RHS} = \{w = \{b^{2k} c^k\}^* : k \in \mathbb{N}_0\}$$

Η λύση θα είναι η σύνθεση αυτών των δύο και συγκεκριμένα:

$$\{w \in L, w_1 \in L_{LHS}, w_2 \in L_{RHS} \mid w = w_1 w_2\}$$

- Για $L_{LHS} : S \rightarrow \varepsilon \mid A, \quad A \rightarrow \varepsilon \mid aAbbb$
- Για $L_{RHS} : S \rightarrow \varepsilon \mid C, \quad C \rightarrow \varepsilon \mid bbCc$
- Άρα : $S \rightarrow \varepsilon \mid AC, \quad A \rightarrow \varepsilon \mid aAbbb, \quad C \rightarrow \varepsilon \mid bbCc$
- $G = (\{a, b, c\}, \{\varepsilon\},$
 $\{S \rightarrow \varepsilon \mid AC, A \rightarrow \varepsilon \mid aAbbb, C \rightarrow \varepsilon \mid bbCc\}, S)$

4 Άσκηση 4η - Αναγνώριση γλωσσών χωρίς συμφραζόμενα:

[30%] Έστω η γραμματική χωρίς συμφραζόμενα

$G = (V, \Sigma, R, S)$, όπου $V = \{S, A, L, D, a, b\}$, $\Sigma = \{a, b\}$ και

$R = \{S \rightarrow A, S \rightarrow L, A \rightarrow LbADa, A \rightarrow LaD, L \rightarrow Da, L \rightarrow e, D \rightarrow b\}$.

(α) Μετατρέψτε τη γραμματική G σε κανονική μορφή Chomsky.

(β) Εφαρμόστε τον αλγόριθμο δυναμικού προγραμματισμού για συντακτική ανάλυση στη συμβολοσειρά $w = babba$. Δώστε τον πλήρη πίνακα N καθώς και όλα τα συντακτικά δέντρα της w .

4 - α Απάντηση Υποερωτήματος (α)

Ζητούμενο η μετατροπή σε κανονική μορφή Chomsky (CNF) CFG's:

$G = (V, \Sigma, R, S)$, όπου $V = \{S, A, L, D, a, b\}$, $\Sigma = \{a, b\}$ και

$R = \{S \rightarrow A, S \rightarrow L, A \rightarrow LbADa, A \rightarrow LaD, L \rightarrow Da, L \rightarrow e, D \rightarrow b\}$.

Καταρχήν να θυμηθούμε τον ορισμό του τί εννοούμε "κανονική μορφή Chomsky":

- Ξεκινώντας να αναφέρω ότι, ότι γράφουμε από εδώ και έπειτα εκτός την δοθέντα CFG, θα είναι με βάση της εναλλακτικής μαθηματικής περιγραφής CFG / CNF που θεωρεί ότι το V ως σύνολο μη τερματικών συμβόλων.

Δηλαδή $V \cap \Sigma = \emptyset$, όπως περιγράφεται σε εξίσου μεγάλο μέρος της βιβλιογραφίας (πχ Sipser). Υπό αυτό το πρίσμα το V μεταφράζεται ως "(V)ariables" σε αντίθεση με όταν ισχύει $V \cap \Sigma = \Sigma$ όπου μεταφράζεται ως "(V)ocabulary". Αυτό διαφέρει από την περιγραφή που δίδεται στο μάθημα, αλλά δεν είναι λιγότερο ή περισσότερο ορθή, απλά διαφορετική.

- $A, B, C \in V$
- $a \in \Sigma$
- Επιτρέπονται μόνο οι παρακάτω τύπου παραγωγές:

1. $A \rightarrow BC$

2. $A \rightarrow a$

3. Σε ειδικές περιπτώσεις απαιτείται δημιουργία νέου εναρκτήριου κανόνα παραγωγής $S_0 \rightarrow S$

4. $\varepsilon \in L(G) \Leftrightarrow S \rightarrow \varepsilon$ ή $S_0 \rightarrow S | \varepsilon$ εάν το S εμφανίζεται στην παραγωγή κάποιου κανόνα.

Μετατροπή $CFG \rightarrow CNF$ 1/3

- Εισαγωγή νέας κατάστασης

1. Στη συγκεκριμένη γραμματική ισχύει $S \in S, S \rightarrow L \rightarrow \varepsilon \Leftrightarrow \varepsilon \in L(G)$ αλλά το S δεν υπάρχει στη δεξιά μεριά κανενός κανόνα παραγωγής και άρα δεν απαιτείτε κατασκευή νέου κανόνα παραγωγής $S_0 \rightarrow S \mid \varepsilon$.

- Απαλοιφή κενής παραγωγής

$$1. S \rightarrow A \mid L$$

$$1. S \rightarrow A \mid L \mid \varepsilon$$

$$2. A \rightarrow LbADa \mid LaD$$

$$2. A \rightarrow LbADa \mid LaD \mid$$

$$bADa \mid aD$$

$$3. L \rightarrow Da \mid \varepsilon$$

$$3. L \rightarrow Da$$

$$4. D \rightarrow b$$

$$4. D \rightarrow b$$

- Απαλοιφή μοναδιαίων παραγωγών (μη τερματικών συμβόλων)

$$1. S \rightarrow A \mid L \mid \varepsilon$$

$$1. S \rightarrow LbADa \mid LaD \mid bADa \mid aD \mid Da \mid \varepsilon$$

$$2. A \rightarrow LbADa \mid LaD \mid$$

$$2. A \rightarrow LbADa \mid LaD \mid bADa \mid$$

$$bADa \mid aD$$

$$aD$$

$$3. L \rightarrow Da$$

$$3. L \rightarrow Da$$

$$4. D \rightarrow b$$

$$4. D \rightarrow b$$

- Απαλοιφή μη προσβάσιμων συμβόλων

1. Όλα τα σύμβολα είναι προσβάσιμα \rightarrow καμία αλλαγή.

Μετατροπή $CFG \rightarrow CNF$ 2/3

- Απομόνωση τερματικών συμβόλων

$$1. S \rightarrow \mathbf{LbADa} \mid \mathbf{LaD} \mid \mathbf{bADa} \mid \mathbf{aD} \mid \mathbf{Da} \mid \varepsilon$$

$$2. A \rightarrow \mathbf{LbADa} \mid \mathbf{LaD} \mid \mathbf{bADa} \mid \mathbf{aD}$$

$$3. L \rightarrow \mathbf{Da}$$

$$4. D \rightarrow b$$

$$1. S \rightarrow \mathbf{LDADD_a} \mid \mathbf{LD_aD} \mid \mathbf{DADD_a} \mid \mathbf{D_aD} \mid \mathbf{DD_a} \mid \varepsilon$$

$$2. A \rightarrow \mathbf{LDADD_a} \mid \mathbf{LD_aD} \mid \mathbf{DADD_a} \mid \mathbf{D_aD}$$

$$3. L \rightarrow \mathbf{DD_a}$$

$$4. D \rightarrow b$$

$$5. \mathbf{D_a} \rightarrow \mathbf{a}$$

- Απαλοιφή τριαδικών+ παραγωγών (1ο πέρασμα)

$$1. S \rightarrow \mathbf{LDADD_a} \mid \mathbf{LD_aD} \mid \mathbf{DADD_a} \mid D_aD \mid DD_a \mid \varepsilon$$

$$2. A \rightarrow \mathbf{LDADD_a} \mid \mathbf{LD_aD} \mid \mathbf{DADD_a} \mid D_aD$$

$$3. L \rightarrow DD_a$$

$$4. D \rightarrow b$$

$$5. D_a \rightarrow a$$

$$1. S \rightarrow \mathbf{LX_1X_2} \mid \mathbf{LX_3} \mid \mathbf{X_1X_2} \mid D_aD \mid DD_a \mid \varepsilon$$

$$2. A \rightarrow \mathbf{LX_1X_2} \mid \mathbf{LX_3} \mid \mathbf{X_1X_2} \mid D_aD$$

$$3. L \rightarrow DD_a$$

$$4. D \rightarrow b$$

$$5. D_a \rightarrow a$$

$$6. \mathbf{X_1} \rightarrow \mathbf{DA}$$

$$7. \mathbf{X_2} \rightarrow \mathbf{DD_a}$$

$$8. \mathbf{X_3} \rightarrow \mathbf{D_aD}$$

Μετατροπή $CFG \rightarrow CNF$ 3/3

- Απαλοιφή τριαδικών+ παραγωγών (2ο πέρασμα)

$$1. S \rightarrow \mathbf{LX_1X_2} \mid LX_3 \mid$$

$$X_1X_2 \mid D_aD \mid DD_a \mid \varepsilon$$

$$2. A \rightarrow \mathbf{LX_1X_2} \mid LX_3 \mid$$

$$X_1X_2 \mid D_aD$$

$$3. L \rightarrow DD_a$$

$$4. D \rightarrow b$$

$$5. D_a \rightarrow a$$

$$6. X_1 \rightarrow DA$$

$$7. X_2 \rightarrow DD_a$$

$$8. X_3 \rightarrow D_aD$$

$$1. S \rightarrow \mathbf{LX_4} \mid LX_3 \mid$$

$$X_1X_2 \mid D_aD \mid DD_a \mid \varepsilon$$

$$2. A \rightarrow \mathbf{LX_4} \mid LX_3 \mid$$

$$X_1X_2 \mid D_aD$$

$$3. L \rightarrow DD_a$$

$$4. D \rightarrow b$$

$$5. D_a \rightarrow a$$

$$6. X_1 \rightarrow DA$$

$$7. X_2 \rightarrow DD_a$$

$$8. X_3 \rightarrow D_aD$$

$$9. \mathbf{X_4 \rightarrow X_1X_2}$$

- $G_{CNF} = (\{S, A, L, D, D_a, X_1, X_2, X_3, X_4\}, \{a, b\},$

$$\{S \rightarrow \varepsilon, S \rightarrow LX_4, S \rightarrow X_1X_2, S \rightarrow LX_3, S \rightarrow D_aD, S \rightarrow DD_a,$$

$$A \rightarrow LX_4, A \rightarrow X_1X_2, A \rightarrow LX_3, A \rightarrow D_aD,$$

$$L \rightarrow DD_a, D \rightarrow b, D_a \rightarrow a, X_1 \rightarrow DA, X_2 \rightarrow DD_a, X_3 \rightarrow D_aD, X_4 \rightarrow X_1X_2\}, S)$$

4 - β Απάντηση Υποερωτήματος (β)

Χρησιμοποιούμε το παρακάτω ως αναφορά για τους κανόνες παραγωγής καθώς με αρίθμηση για να μπορούμε να διαχωρίσουμε αυτούς που ξεκινάνε από ίδιο μη τερματικό σύμβολο:

$$1. S \rightarrow LX_4$$

$$10. A \rightarrow D_aD$$

$$2. S \rightarrow LX_3$$

$$11. L \rightarrow DD_a$$

$$3. S \rightarrow X_1X_2$$

$$12. D \rightarrow b$$

$$4. S \rightarrow D_aD$$

$$13. D_a \rightarrow a$$

$$5. S \rightarrow DD_a$$

$$14. X_1 \rightarrow DA$$

$$6. S \rightarrow \varepsilon$$

$$15. X_2 \rightarrow DD_a$$

$$7. A \rightarrow LX_4$$

$$16. X_3 \rightarrow D_aD$$

$$8. A \rightarrow LX_3$$

$$17. X_4 \rightarrow X_1X_2$$

$$9. A \rightarrow X_1X_2$$

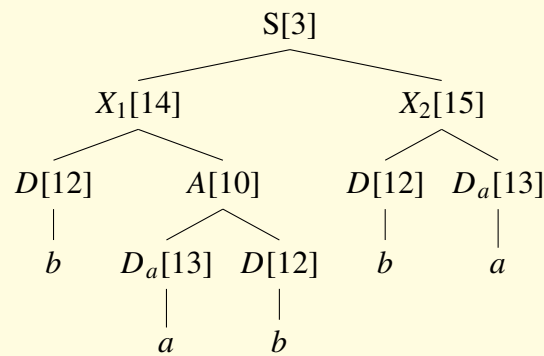
Πίνακας συντακτικής ανάλυσης για $w = babba$

- Κατασκευάζουμε πίνακα $|w|$ στηλών και για κάθε στήλη $n \in [1, |w|_{max}]$ ξεκινώντας να μετράμε από αριστερά ως 1η, έχουμε n κελιά.
- Στο κορυφαίο κελί κάθε στήλης (δηλαδή το n), τοποθετούμε το αντίστοιχο σύμβολο της λέξης (δηλαδή το w_n).
- Κάθε ένα από αυτά τα σύμβολα ουσιαστικά αποτελεί κανόνα παραγωγής τερματικού συμβόλου.
- Ανά στήλη, σε κελί k , $k \in [2, n]$ με k_{max} να βρίσκεται στη βάση του πίνακα και k_{min} ακριβώς στο κελί κάτω από αυτό στην κορυφή της στήλης του, τοποθετούμε τους κανόνες παραγωγής (τα μη τερματικά σύμβολα) τα οποία δύνανται να οδηγήσουν σε αριστερότερο τμήμα της w μήκους υποσυμβολοσειράς x με $|x| = k$, $x = \sum_{i=1}^k w_i$
- Όταν ολοκληρώσουμε περιμένουμε να βρούμε στη κάτω δεξιά γωνία κάποιο κανόνα παραγωγής, από το αρχικό (μη τερματικό προφανώς) σύμβολο. Εάν αυτό υπάρχει τότε $w \in G$, ειδάλως $w \notin G$.

				a
			b	L[11], X ₂ [15], S[5]
		b	∅	∅
	a	S[4], A[10], X ₃ [16]	∅	∅
b	L[11], X ₂ [15], S[5]	X ₁ [14]	∅	A[9], S[3], X ₄ [17]

Δέντρο συντακτικής ανάλυσης για $w = babba$

- Κατασκευάζουμε δέντρο συντακτικής ανάλυσης βάση αντίστροφης χρήσης των αποτελεσμάτων του πίνακα, δηλαδή ξεκινώντας από την κάτω δεξιά γωνία και συγκεκριμένα την αντίστοιχη εκδοχή αρχικού σύμβολου εφόσον η w όντως ανήκει στη γραμματική μας.
- Σε κάθε βήμα ακολουθούμε κλάδο βάση κανόνων παραγωγής αυτής της εκδοχής συμβόλου και ακολουθείτε από την εκάστοτε εκδοχή επόμενων με τερματικών συμβόλων που έχουν βρεθεί στο αντίστοιχο βήμα στο πίνακα, έως να φτάσουμε διαδοχικά στα κατάλληλα τερματικά σύμβολα που κατασκευάζουν τη λέξη που μας δόθηκε.



Που δίνει:

$b \quad a \quad b \quad b \quad a$

