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

ΜΕΡΟΣ **1**

Κατάλογος συμβόλων και συντομογραφίες

- $\mathcal{L}$ 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 $\mathrm{M} = (\mathrm{Q}, \Sigma, \delta, \mathrm{s}, \mathrm{F})$
- 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
- $\ldots$ "και ούτω καθεξής", συνέχιση ακολουθίας βάση εμφανούς μοτίβου
- $|$ ή $:$ ή $\ni$ Έτσι Ώστε
- $\mathcal{P}(Q)$ Δυναμοσύνολο συνόλου Q.
  Για $Q = \{a, b, c\} \overset{\emptyset \subseteq \text{all sets}}{\rightarrow} \mathcal{P}(Q) = \{\emptyset, a, b, c, ab, ac, bc, abc\}$
  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 **1**η - Γραμματικές χωρίς συμφραζόμενα**:**

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

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

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

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

1 - α Απάντηση Υποερωτήματος (α)

Έχουμε CFL $L_1 = \{a^n, b^m : n, m \in \mathbb{N}_0, 3n \le m \le 6n\}$ 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: $\rightarrow a\,bbb\,b,\ a\,bbb\,bb,\ a\,bbb\,bbb$.
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:

  $S \rightarrow aM \mid \varepsilon$

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

- $|w| = \aleph_0$:

  $S \rightarrow aSM \mid \varepsilon$

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

- Πιο ευανάγνωστη παραλλαγή:

  $S \rightarrow aSbbM|\varepsilon$

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

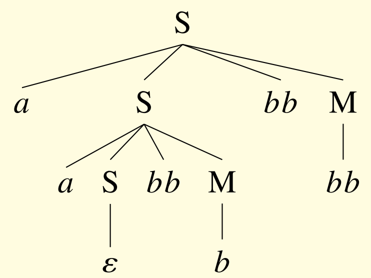
- Καμία αριστερή αναδρομή.

Γραμματική και συντακτικό δένδρο

Είμαστε έτοιμοι να προχωρήσουμε στην πλήρη περιγραφή της γραμματικής μας:

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

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)

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: $b\,aaa\,aa$, $a\,b\,aaa\,a$, ..., $aa\,aaa\,b$.
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 $|w|_b \in \aleph_0$ 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:

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

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

- $|w| = \aleph_0$:

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

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

- 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:

- $bb\,aaa\,aaa\,aa$: $S \rightarrow RaRaR \rightarrow (baaa)aRaR \rightarrow (bRaaa)aRaR \rightarrow (b(bRaaa)aaa)aRaR \rightarrow (b(b\varepsilon aaa)aaa)aRaR \rightarrow (bb\,aaa\,aaa)a\varepsilon a\varepsilon \rightarrow bb\,aaa\,aaa\,aa$
- $a\,bb\,aaa\,aaa\,a$: $S \rightarrow RaRaR \rightarrow \varepsilon aRa\varepsilon \rightarrow a(bRaaa)a \rightarrow a(b(bRaaa)aaa)a \rightarrow a(b(b\varepsilon aaa)aaa)a \rightarrow a\,bb\,aaa\,aaa\,a$
- $aa\,bbaaa\,aaa$: $S \rightarrow RaRaR \rightarrow \varepsilon a\varepsilon aR \rightarrow aa(bRaaa) \rightarrow aa(b(bRaaa)aaa) \rightarrow aa(b(b\varepsilon aaa)aaa) \rightarrow aa\,bb\,aaa\,aaa$
- ...
- $aa\,aaa\,bb\,aaa$: $S \rightarrow RaRaR \rightarrow \varepsilon a\varepsilon aR \rightarrow aa(aaabR) \rightarrow aa(aaab(bRaaa)) \rightarrow aa(aaab(b\varepsilon aaa)) \rightarrow aa\,aaa\,bb\,aaa$
- ...
- $aaa\,aaa\,bb\,aa$: $S \rightarrow RaRaR \rightarrow Ra\varepsilon a\varepsilon \rightarrow (aaab)aa \rightarrow (aaaRb)aa \rightarrow (aaa(aaaRb)b)aa \rightarrow (aaa(aaa\varepsilon b)b)aa \rightarrow aaa\,aaa\,bb\,aa$
- 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, $baab$, ..., $baaaaaaaab$, 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 (πχ $\in \mathbb{N}_0$ 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 πχ $\{\forall w \neq \varepsilon \mid w_{min} = xy,\ \exists w_{double} = xxyy \cup xyxy \cup xyyx \cup yxxy \cup yxyx \cup yyxx\}$

Υπόψιν ότι εκτός κάποιον 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.

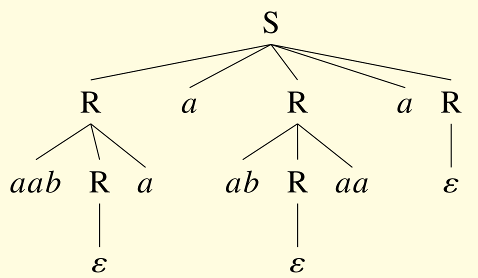
Γραμματική και συντακτικό δένδρο

Είμαστε έτοιμοι να προχωρήσουμε στην πλήρη περιγραφή της γραμματικής μας:

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

$R \to bRaaa \mid abRaa \mid aabRa \mid aaaRbR \mid \varepsilon\}, S)$

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

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

2 Άσκηση **2**η **-** Αυτόματα στοίβας**:**

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

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

(β) $L_2 = \{w \in \{a, b\}^* :$ η w περιέχει διπλάσιο αριθμό b 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 $w \in L_1$:

- $|w|_{min} = 0 \Longrightarrow n,\ k = 0 \Longrightarrow w = \varepsilon, \quad |w|_{max} = \aleph_0, \quad |w| = 2(n+k) =$ Άρτιο
- $n + k - 2nk = 1 \Longrightarrow |w|_{min} = 2 \Longrightarrow w = ba \cup ab$
- $n, k = 1 \Longrightarrow |w|_{min} = 4 \Longrightarrow w = baba, \qquad L \subseteq \{xx^R | x \in \Sigma^*\}$
- 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 τέσσερεις 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.

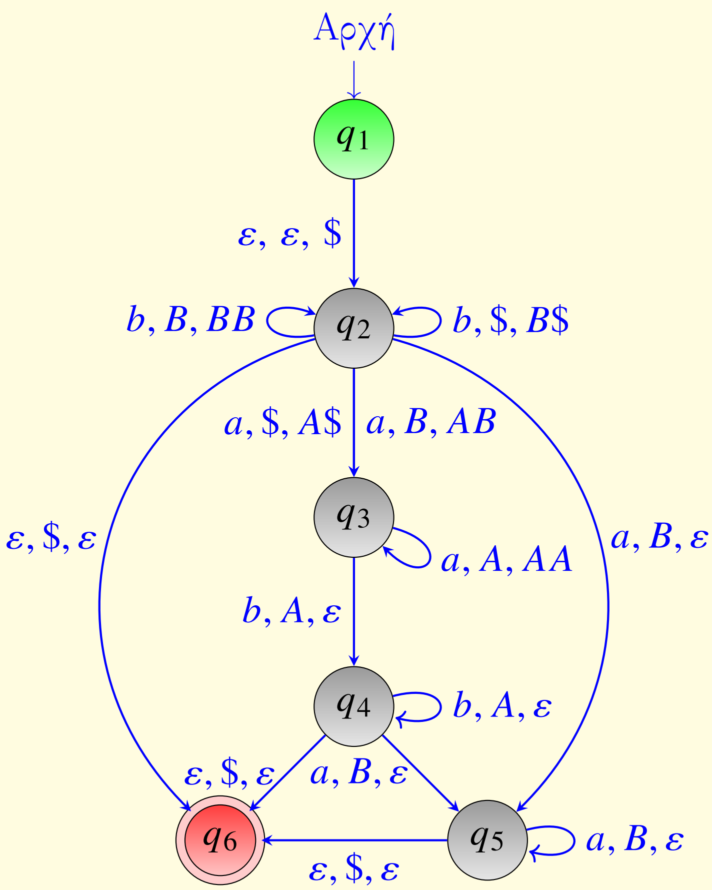
Τέλος αφού $n, k \in \mathbb{N}_k$ τότε μπορούν όλοι αυτοί οι χαρακτήρες να είναι (μετρήσιμα) άπειροι αλλά υπό τον όρο ότι τηρείται το γενικό μοτίβο.

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

- $K = \{q1, q2, q3, q4, q5, q6\}$
- $\Sigma = \{a, b\}$
- $\Gamma = \{A, B, \$\}$
- $\Delta =$
    1. $(q1, \varepsilon, \varepsilon) \rightarrow (q2, \$)$
    2. $(q2, b, \$) \rightarrow (q2, B\$)$
    3. $(q2, b, B) \rightarrow (q2, BB)$
    4. $(q2, \varepsilon, \$) \rightarrow (q6, \varepsilon)$
    5. $(q2, a, \$) \rightarrow (q3, A\$)$
    6. $(q2, a, B) \rightarrow (q3, AB)$
    7. $(q2, a, B) \rightarrow (q5, \varepsilon)$
    8. $(q3, a, A) \rightarrow (q3, AA)$
    9. $(q3, b, A) \rightarrow (q4, \varepsilon)$
    10. $(q4, b, A) \rightarrow (q1, \varepsilon)$
    11. $(q4, a, B) \rightarrow (q5, \varepsilon)$
    12. $(q4, \varepsilon, \$) \rightarrow (q6, \varepsilon)$
    13. $(q5, a, B) \rightarrow (q5, \varepsilon)$
    14. $(q5, \varepsilon, \$) \rightarrow (q6, \varepsilon)$
- $s = q_1$
- $F = q_6$

DFA $M_{\text{τελικό}}$



Υπολογισμός αποδοχής $bbbabaaa$

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

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

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

$$\vdash (q6, \varepsilon, \varepsilon) \qquad \text{Επιτυχής Ανάγνωση}$$

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 $w \in L_1$:

- $|w|_{min} = 0 \Longrightarrow w_{min} = \varepsilon, \quad |w|_{max} = \aleph_0, \quad |w| \equiv |w|_a \pmod 2$
- $\{w \in \{a,\ b\}^+ \mid |w| = 3,\ |w|_a = 1,\ |w|_b = 2\} \Longrightarrow |w| = 3|w|_a$
- $|w| = 3\ :\ abb, bab, bba$
- $|w| > 3\ :\ aabbbb,\ bbbbaa,\ abbbba,\ babbab,\ ababbb,\ bbbaba,\ bbaabb,\ \ldots$
- 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)).

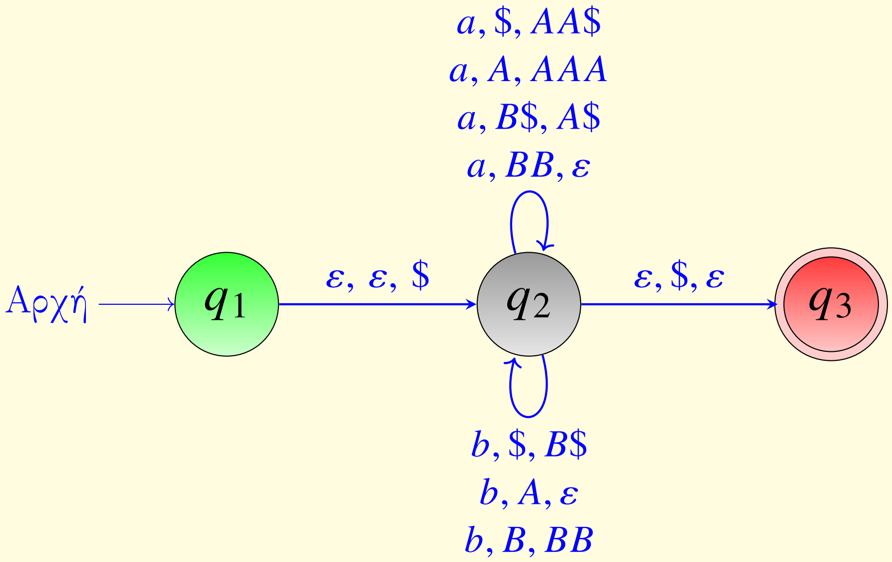
Αυτό με τη σειρά του δείχνει ότι πιθανώς να μην χρειαστεί κάποια διακλάδωση και να μπορούν να γίνουν όλες οι διαδικασίες αναδρομικά στον ίδιο κόμβο.

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

- $K = \{q1, q2, q3\}$
- $\Sigma = \{a, b\}$
- $\Gamma = \{A, B, \$\}$
- $\Delta =$
  1. $(q2, \varepsilon, \varepsilon) \rightarrow (q2, \$)$
  2. $(q2, b, \$) \rightarrow (q2, B\$)$
  3. $(q2, b, B) \rightarrow (q2, BB)$
  4. $(q2, b, A) \rightarrow (q2, \varepsilon)$
  5. $(q2, a, \$) \rightarrow (q2, AA\$)$
  6. $(q2, a, A) \rightarrow (q2, AAA)$
  7. $(q2, a, BB) \rightarrow (q2, \varepsilon)$
  8. $(q2, a, B\$) \rightarrow (q2, A\$)$
  9. $(q2, \varepsilon, \$) \rightarrow (q3, \varepsilon)$
- $s = q_1$
- $F = q_3$

DFA $M_{\text{τελικό}}$



Υπολογισμός αποδοχής $abbbab$

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

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

$$\vdash (q3, \varepsilon, \varepsilon) \qquad \text{Επιτυχής Ανάγνωση}$$

3 Άσκηση **3**η **-** 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

- $\mathcal{P} := \{\forall L_{DCF} \in \mathcal{L}_{\text{DCF}},\ \forall L_{REG} \in \mathcal{L}_{\text{REG}} :\ L_{DCF} \cap L_{REG} \in \mathcal{L}_{\text{CF}}\}$ **(1)**
- $\mathcal{L}_{FIN} \subset \mathcal{L}_{REG} \subset \mathcal{L}_{DCF} \subset \mathcal{L}_{NCF} = \mathcal{L}_{CF}$ **(2)**
- $A \subset B \Rightarrow A \cap B = A \subset B$ **(3)**
- $A \subset B \subset \ldots \subset N \Rightarrow A \subset N$ **(4)**
- $\overset{(2)(4)}{\Longrightarrow} \mathcal{L}_{FIN} \subset \mathcal{L}_{DCF}$ **(5)**
- $\mathcal{L}_{FIN} \cap \mathcal{L}_{DCF} \overset{(3)(5)}{=} \mathcal{L}_{FIN} \subset \mathcal{L}_{DCF} \subset \mathcal{L}_{CF} \overset{(4)}{\Rightarrow} \mathcal{L}_{FIN} \subset \mathcal{L}_{CF} \Rightarrow \mathcal{L}_{FIN} \in \mathcal{L}_{CF}$ **(6)**
- $\mathcal{Q} \overset{(6)}{:=} \{\forall L_{DCF} \in \mathcal{L}_{\text{DCF}},\ \forall L_{REG} \in \mathcal{L}_{\text{REG}} :\ L_{DCF} \cap L_{REG} \in \mathcal{L}_{\text{CF}}\}$ **(7)**
- $\mathcal{P}^{(1)} = \mathcal{Q}^{(7)} = \text{True} \Rightarrow \mathcal{P} = \text{True}$

3 - β Απάντηση Υποερωτήματος (β)

Η άσκηση μας ζητά να αποδείξουμε εάν η παρακάτω γλώσσα είναι **CF** ή όχι:

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

Υπάρχουν πάρα πολλοί τρόποι να το κάνουμε αυτό. Μεταξύ των οποίων είναι η απόπειρα κατασκευής **CFG** της γλώσσας, όπου αν τα καταφέρουμε αυτόματα αποδεικνύουμε ότι είναι **CFL**, ενώ μία άλλη μέθοδος είναι μέσω λήμματος άντλησης. Υπάρχουν αρκετοί ακόμη μέθοδοι αλλά δεν θα τις αναφέρουμε ούτε θα τις δείξουμε όλες.

Απόδειξη με απόπειρα κατασκευής **CFG**

- Έχουμε συγκεκριμένη θέση έκαστου συμβόλου δηλαδή πρώτα όλα τα a, έπειτα όλα τα b και τελευταία όλα τα c.

- Συγκεκριμένη η θέσης και ουσιαστικά το κάθε σύμβολο έχει συγκεκριμένη προτεραιότητα (δεν αναμιγνύονται πχ δεν υπάρχει bbabbcb 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 τον δύο και συγκεκριμένα:
  $\{w \in L,\, w_1 \in L_{LHS},\, w_2 \in L_{RHS} \,|\, w = w_1 w_2\}$

- Για $L_{LHS} \;:\; S \to \varepsilon \,|\, A, \quad A \to \varepsilon \,|\, aAbbb$

- Για $L_{RHS} \;:\; S \to \varepsilon \,|\, C, \quad C \to \varepsilon \,|\, bbCc$

- Άρα : $S \to \varepsilon \,|\, AC, \quad A \to \varepsilon \,|\, aAbbb, \quad C \to \varepsilon \,|\, bbCc$

- $G = (\{a,\, b,\, c\}, \quad \{\varepsilon\},$
  $\{S \to \varepsilon \,|\, AC,\; A \to \varepsilon \,|\, aAbbb,\; C \to \varepsilon \,|\, bbCc\}, \quad S)$

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)

Ο ισχυρισμός είναι λανθασμένος και ως απόδειξη δείχνοντας έστω μία περίπτωση όπου ένωση μετρήσιμης με μή κανονικής γλώσσας παράγει μή κανονική γλώσσα. Αφού ο ισχυριζόμαστε ότι από $\mathcal{L}_{fin} \cup \mathcal{L}_{irr}$ ΠΑΝΤΑ παράγεται κανονική άρα εάν βρεθεί οποιαδήποτε εξαίρεση ο ισχυρισμός απορρίπτεται ως λανθασμένος.

Απόδειξη - μέρος 1/3

- Ισχύει $\mathcal{L}_{fin} \subset \mathcal{L}_{reg} \xrightarrow{\mathcal{L}_{reg}=\mathcal{L}_{irr}^c} \mathcal{L}_{fin} \cap \mathcal{L}_{irr} = \emptyset$ **(1)**

- Ο ισχυρισμός του ερωτήματος ισοδυναμεί με $\xrightarrow{(1)}$ $\mathcal{P}$:= "Η ένωση κανονικής γλώσσας με μή κανονικής είναι πάντα κανονική γλώσσα" $\rightarrow$
$\mathcal{P} := \{\forall\ L_{reg} \cup L_{irr} \mid (L_{reg} \in \mathcal{L}_{reg},\ L_{irr} \in \mathcal{L}_{irr})\} \in \mathcal{L}_{reg}$. **(2)**

- $\mathcal{Q}$:="Υπάρχει τουλάχιστον μία πεπερασμένη γλώσσα με μία μή κανονική γλώσσα με την ένωση τους να μην είναι κανονική γλώσσα"$\xrightarrow{(1)}$
$\mathcal{Q} := \{\exists\ L_{reg} \cup L_{irr} \mid (L_{reg} \in \mathcal{L}_{reg},\ L_{irr} \in \mathcal{L}_{irr})\} \in \mathcal{L}_{irr}$. **(3)**

- Γλώσσα όπου εφαρμογή λήμματος άντλησης οδηγεί σε άτοπο είναι μή κανονική. **(4)**

- Ως λήμμα άντλησης[(4)] κανονικής γλώσσας L ορίζεται η εξής μέθοδος:
$$\{L \subseteq \Sigma^* \mid \exists n \geq 1 \ni (\forall w \in L,\ |w| \geq n \ni$$
$$[\exists x, y, z \in \Sigma^* \ni (w = xyz,\ y \neq \varepsilon,\ |xy| \leq n) \rightarrow (i \geq 0 \ni \exists xy^iz \in L)])\}$$
(5)

Απόδειξη - μέρος 2/3

- Θεωρούμε ότι έχουμε DFA M με $M = (K, \Sigma, \Delta, s, F) = (\{q0, q1\}, \{a,\ b\},$ $\{\ \delta(q0, a) = q1,\ \delta(q0, b) = q1,\ \delta(q1, a) = q1,\ \delta(q1, b) = q1\}, q0, \{q0\})$ που αναγνωρίζει τη γλώσσα $L(M)$(εξ ορισμού κανονική) $= L_{fin} \in \mathcal{L}_{fin}$.

  Η γλώσσα μας είναι $L_{fin} = \{w \in \{\varepsilon\}\}$, προφανέστατα πεπερασμένη αφού $|L| = 1 \leq \aleph_0$. Συνεπάγεται ότι είναι και κανονική[(1)] οπότε ορίζουμε αντίγραφο της με σύμβολο $L_{reg} \in \mathcal{L}_{reg} = L_{fin}$.

  (6)

- Θεωρούμε γλώσσα $L = w \in (a^x, b^x) : x \in \mathbb{N}$ (η οποία είναι προφανώς μετρήσιμα άπειρη) και ορίζουμε την $L_{irr} = L$. Εφαρμόζουμε το θεώρημα άντλησης[(5)] υποθέτοντας ότι η γλώσσα μας είναι κανονική ώστε να αποδείξουμε με εις άτοπον απαγωγή ότι είναι μή κανονική:

  $w \in L_{irr},\ x = n,\ w = a^n b^n,\ |w| = 2n \geq n^{\ (\beta\lambda\ |w| \geq n\ (5))},\ w = xyz,\ x = a^j,$

  $y = a^k,\ y \neq \emptyset \to k \geq 1,\ |xy| = j + k \leq n,\ z = a^{n-j-k} b^n$

  $w = xy^i z = a^j a^i k a^{n-j-k} b^n \Longrightarrow a^{n-j-k+j+ik} b^n \Longrightarrow a^{n-k+jk} b^n \Longrightarrow n-k+jk =$ $n \Longrightarrow jk - k = 0 \Longrightarrow jk = k \Longrightarrow j = 1$

  Άρα για οποιοδήποτε $j \neq 1$ οδηγούμαστε σε άτοπο για παράδειγμα:

  $i = 2 \to w = xy^2 z = a^{n-k+2k} b^n \Longrightarrow a^{n+k} b^n$ αλλά $k \geq 1$ αφού $|y| \neq \emptyset$ τότε $n + k \neq n$ δηλαδή $|w|_a \neq |w|_b$ άρα αποδείξαμε ότι η $L_{irr} \in \mathcal{L}_{irr}$ [(4)].

  (7)

- $L_{union} = L_{reg} \cup L_{irr} \overset{(7) \wedge (6)}{\Longrightarrow} L_{union} = \{\varepsilon\} \cup \{a^n b^n : n \in \mathbb{N}^0\}$

  αλλά $\{\varepsilon\} = \{a^0 b^0\} \overset{n \in \mathbb{N}^0}{\to} \{\varepsilon\} \in \{a^n b^n\ : n \in \mathbb{N}^0\}$ και άρα

  $L_{union} = \{a^n b^n | n \in \mathbb{N}^0\} = L_{irr} \neq$ κανονική γλώσσα $\to Q = T$.

  (8)

Απόδειξη - μέρος **3/3**

- $\mathcal{P} = \neg Q$ αλλά ήδη αποδείξαμε[8] ότι $Q = T$ και άρα:

  $\mathcal{P} = \neg Q \Longrightarrow \mathcal{P} = \neg T \Longrightarrow \mathcal{P} = F$

  Άρα η πρόταση του αρχικού ισχυρισμού αποδείχθηκε λανθασμένη.

4 - *β* Απάντηση Υποερωτήματος (β)

Ο ισχυρισμός είναι σωστός.

Αιτιολόγηση

Οι (ιδιότυπες) κλάσεις (σύνολα) ισοδυναμίας καθορίζονται αυστηρά και μόνο από ιδιότητες της γλώσσας. Καθώς το πρότυπο DFA $\sim M$ είναι ελάχιστο και δίχως απρόσιτη κατάσταση, οι καταστάσεις του είναι οι ίδιες οι κλάσεις ισοδυναμίας 1-1 προς τις κλάσεις ισοδυναμίας της $\approx L$. Δηλαδή και στις δύο περιπτώσεις διαβάζοντας οποιαδήποτε συμβολοσειρά οδηγούμαστε σε ίδια κλάση ισοδυναμίας. Θα πρέπει επίσης να θυμόμαστε ότι δεν περιγράφουν βέβαια του ίδιου τύπου δεδομένα μία και το ένα περιγράφεται από κλάσεις ισοδυναμίας καταστάσεων ενώ το άλλο κλάσεις ισοδυναμίας συμβολοσειρών.