771 Η - Θεωρία Υπολογισμών και Αλγορίθμων

Σημειώσεις – Μέρος 3ο

Διδάσκων: Α. Ντελόπουλος

Το παρόν αποτελεί σημειώσεις που αντιστοιχούν σε μέρος των διαλέξεων για το μάθημα **771 Η - Θεωρία Υπολογισμών και Αλγορίθμων** του Τμήματος Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Αριστοτέλειου Πανεπιστήμιου Θεσσαλονίκης.

Το υλικό των διαλέξεων και των σημειώσεων στηρίζεται σε μεγάλο βαθμό στην ύλη του βιβλίου: *Harry R. Lewis and Christos H. Papadimitriou, "Elements of the Theory of Computation," 2nd edition, Prentice Hall, 1998*.

Γλώσσες Χωρίς Συμφραζόμενα

Ορισμοί, Ιδιότητες, Αντιστοιχία με τα Αυτόματα Στοίβας (Push-Down)

Γεννήτριες Γλωσσών

- **Γεννήτρια Γλώσσας** είναι ένας μηχανισμός (αλγόριθμος) που μπορεί να παράγει ένα συγκεκριμένο σύνολο συμβολοσειρών
- Το σύνολο αυτό είναι η γλώσσα της γεννήτριας
- Οι ΚΕ είναι **ένα** από τα είδη γεννητριών
  – π.χ. ΚΕ $R=a(a^*\cup b^*)b$
  – Πρώτα βάλε ένα a. Μετά κάνε ένα από τα ακόλουθα: Είτε βάλε έναν αριθμό από a είτε βάλε έναν οποιοδήποτε αριθμό από b. Στο τέλος βάλε ένα b.
- Μία **γενικότερη** μορφή γεννητριών είναι οι “Γραμματικές Χωρίς Συμφραζόμενα” (ΓΧΣ)

Γραμματικές Χωρίς Συμφραζόμενα (ΓΧΣ)

Ορισμός: Μία ΓΧΣ είναι μία τετράδα G=(V,Σ,R,S) όπου:

- V είναι ένα επεκτεταμένο αλφάβητο
- Σ είναι το σύνολο των Τερματικών Συμβόλων (ΤΣ) δηλαδή το κλασσικό μας αλφάβητο. $\Sigma \subseteq V$.
- R το σύνολο των κανόνων της γραμματικής δηλαδή $R \subseteq (V-\Sigma)\times V^*$ (όπου V* όλες οι συμβολοσειρές του V)
- S το Σύμβολο Εκκίνησης (ΣΕ) $\in V-\Sigma$
  - Το V-Σ είναι το σύνολο Μη Τερματικών Συμβόλων (ΜΤΣ)

- Μία ΓΧΣ G γεννά συμβολοσειρές αντικαθιστώντας ΜΤΣ με συνδιασμούς ΤΣ και ΜΤΣ (ή και με την κενή συμβολοσειρά $e \in V^*$) εφαρμόζοντας τους κανόνες R. Η αντικατάσταση σταματά όταν δεν υπάρχουν άλλα ΜΤΣ
- Για κάθε αντικατάσταση ΔΕΝ ελέγχουμε ποιά σύμβολα υπάρχουν "τριγύρω" = Χωρίς Συμφραζόμενα

Παράδειγμα Λειτουργίας ΓΧΣ - 1

- Έστω G=(V,Σ,R,S) με
  - V={S,M,A,B,a,b}
  - Σ={a,b}
  - R={(S,aMb), (M,A),(M,B),(A,e),(A,aA),(B,e),(B,bB)} ισοδύναμα γράφουμε R={S→aMb, M→A, M→B, A→e, A→aA, B→e, B→bB}
  - S=S
- Παραδείγματα γέννεσης (παραγωγής) συμβολοσειρών:

  $S \Rightarrow aMb \Rightarrow aAb \Rightarrow aaAb \Rightarrow aaaAb \Rightarrow aaab$

  $S \Rightarrow aMb \Rightarrow aBb \Rightarrow abBb \Rightarrow abb$

  $S \Rightarrow aMb \Rightarrow aAb \Rightarrow ab$
- Εύκολα φαίνεται ότι αυτή η ΓΧΣ G ταυτίζεται με την ΚΕ R= $a(a^* \cup b^*)b$
- ΠΡΟΣΟΧΗ: **Δεν** αντιστοιχούν όλες οι ΓΧΣ σε ΚΕ

Παράδειγμα Λειτουργίας ΓΧΣ - 2

- Έστω G=(V,Σ,R,S) με
  - V={S,a,b}
  - Σ={a,b}
  - R= {$S \rightarrow aSb$, $S \rightarrow e$}
  - S=S
- Παραδείγματα γέννεσης:

  $S \Rightarrow aSb \Rightarrow ab$

  $S \Rightarrow aSb \Rightarrow aaSbb \Rightarrow aabb$

  $S \Rightarrow aSb \Rightarrow aaSbb \Rightarrow aaaSbbb \Rightarrow aaabbb$
- Εύκολα φαίνεται ότι αυτή η ΓΧΣ G παράγει όλες τις συμβολοσειρές της μορφής $a^n b^n$, όπου n θετικός ακέραιος.
- Είναι γνωστό ότι καμμία ΚΕ δεν παράγει αυτές τις συμβολοσειρές!

Παραγωγή και Γλώσσες Χωρίς Συμφραζόμενα (ΓΛΧΣ)

- Συνοψίζοντας γράφουμε:

  "$A \rightarrow_G u$" όταν $A \in V\text{-}\Sigma$, $u \in V^*$ και $(A,u) \in R \subseteq (V\text{-}\Sigma) \times V^*$

  "$u \Rightarrow_G v$" όταν το v **παράγεται** από το u με την αντικατάσταση **ενός** ΜΤΣ του u

  "$u \Rightarrow_G^* v$" όταν το v **παράγεται** από το u με διαδοχικές αντικαταστάσεις ΜΤΣ του u (και των "παραγώγων" του)

  Ο δείκτης $_G$ στους παραπάνω συμβολισμούς μπορεί να παραληφθεί όταν είναι αυτονόητη η γραμματική G

- Οποιαδήποτε ακολουθία $w_0 \Rightarrow_G w_1 \Rightarrow_G \ldots \Rightarrow_G w_n$ λέγεται **παραγωγή** του w_n από το w_0 από την G
- ΓΛΧΣ $L(G)=\{w \in \Sigma^*: S \Rightarrow_G^* w\}$

Σχέση Κανονικών Γλωσσών και Γλωσσών Χωρίς Συμφραζόμενα

- Κάθε Κανονική Γλώσσα είναι και ΓΛΧΣ
  - Απόδειξη: Έστω κανονική γλώσσα L και M=(K,Σ,δ,s,F) το ΑΠΑ που την αποδέχεται. Τότε η ίδια γλώσσα παράγεται από τη ΓΧΣ G=(V,Σ,R,S) με
    - $V = K \cup \Sigma$ (αντιστοιχίζουμε τις καταστάσεις σε ΜΤΣ)
    - $S = s$
    - $R = \{q \rightarrow ap: \delta(q,a) = p\} \cup \{q \rightarrow e: q \in F\}$

  εύκολα φαίνεται ότι η G γεννά τις συμβολοσειρές που είναι αποδεκτές από το M.
- Υπάρχουν ΓΛΧΣ που ΔΕΝ είναι Κανονικές

Συντακτικά Δέντρα (Parse Trees)

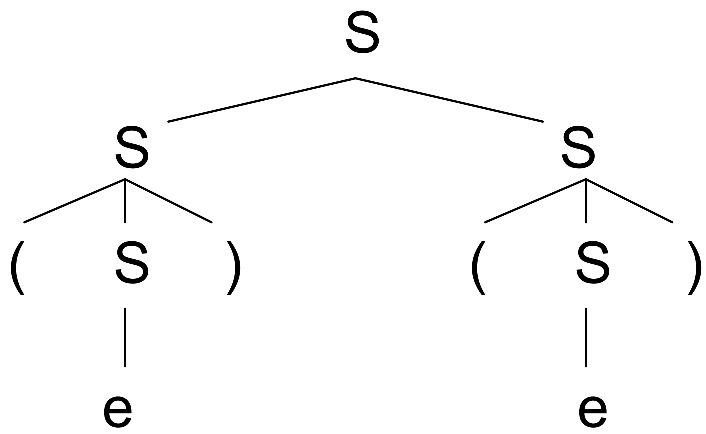
- Οι Παραγωγές συμβολοσειρών αναπαρίστανται μέσω Συντακτικών Δέντρων (ΣΔ):
  - Κάθε κόμβος του ΣΔ αντιστοιχεί σε ένα σύμβολο του V
  - Κάθε κλάδος του ΣΔ που ξεκινά από έναν κόμβο A καταλήγει σε κόμβους B_1, B_2, …, B_k αν το A αντικαθίσταται με χρήση του κανόνα $A \rightarrow B_1B_2\ldots B_k$
  - Το σύμβολο εκκίνησης αντιστοιχεί στη **ρίζα** του ΣΔ
  - Τα (ακραία) φύλλα του ΣΔ είναι Τερματικά Σύμβολα
  - Η συμβολοσειρά που σχηματίζεται από τα (ακραία) φύλλα όταν τα διαβάσουμε από τα αριστερά προς τα δεξιά είναι το **αποτέλεσμα** του ΣΔ

Συντακτικά Δέντρα - Παράδειγμα

- Έστω ΓΧΣ G=(V,Σ,R,S) με
  - V={S,(,)}
  - Σ={(,)}
  - R={S→e, S→SS, S→(S)}
  - S=S
- και οι παραγωγές:

1. S⇒SS⇒(S)S⇒()(S)⇒()()
2. S⇒SS⇒S(S)⇒(S)(S)⇒(S)() ⇒()()

- οι δύο αυτές παραγωγές αντιστοιχούν στο **ίδιο ΣΔ** που φαίνεται παρακάτω



Διάταξη μεταξύ Παραγωγών

- Η παραγωγή $D = x_1 \Rightarrow x_2 \Rightarrow \ldots \Rightarrow x_n$ **προηγείται** της $D' = x_1' \Rightarrow x_2' \Rightarrow \ldots \Rightarrow x_n'$ όταν
  - διαφέρουν **μόνο σε δύο** διαδοχικά βήματα
    - δηλαδή $\Rightarrow x_k \Rightarrow$ διαφέρει από το $\Rightarrow x_k' \Rightarrow$
  - στην D η αντικατάσταση των αριστερότερων ΜΤΣ γίνεται πριν από την αντίστοιχη αντικατάσταση στην D'
  - Γράφουμε: $D < D'$
- Παράδειγμα: Αν

$D_1 = S \Rightarrow SS \Rightarrow (S)S \Rightarrow ((S))S \Rightarrow (())S \Rightarrow (())(S) \Rightarrow (())()$

$D_2 = S \Rightarrow SS \Rightarrow (S)S \Rightarrow ((S))S \Rightarrow ((S))(S) \Rightarrow (())(S) \Rightarrow (())()$

$D_3 = S \Rightarrow SS \Rightarrow (S)S \Rightarrow ((S))S \Rightarrow ((S))(S) \Rightarrow ((S))() \Rightarrow (())()$

Τότε $D_1 < D_2$, $D_2 < D_3$, αλλά ΟΧΙ $D_1 < D_3$

Ομοιες Παραγωγές

- Δύο παραγωγές D_1 και D_2 λέγονται **όμοιες** όταν συνδέονται, απευθείας ή μέσω μιας αλυσίδας άλλων παραγωγών με τις σχέσεις "<" ή ">"
  – οι D_1 και D_3 του προηγούμενου παραδείγματος είναι όμοιες γιατί $D_1 < D_2 < D_3$
- Η ομοιότητα είναι σχέση ισοδυναμίας
- Δύο παραγωγές D_1 και D_2 είναι **όμοιες** αν και μόνο αν αντιστοιχούν στο **ίδιο ΣΔ**
- **Αριστερότερη** (**Δεξιότερη**) Παραγωγή είναι αυτή που προηγείται (έπεται) όλων των ομοίων της
  – Ισοδύναμα: Στο ΣΔ αντικαθίσταται πρώτα το πιο αριστερό (δεξιό) ΜΤΣ
- Είναι δυνατό μία συμβολοσειρά, w, να γεννιέται από περισσότερα του ενός ΣΔ = **ασαφής** γραμματική

Αυτόματα Στοίβας (push-down)

- Κεντρική Ιδέα:
  - Όπως τα *Πεπεπερασμένα Αυτόματα* (ΑΠΑ και ΜΑΠΑ) αντιστοιχούν 1-1 σε *Κανονικές Γλώσσες*, έτσι και τα *Αυτόματα Στοίβας* αντιστοιχούν 1-1 στις *Γλώσσες Χωρίς Συμφραζόμενα*
  - Όπως οι *Γλώσσες Χωρίς Συμφραζόμενα* είναι γενίκευση των *Κανονικών Γλωσσών* έτσι και τα *Αυτόματα Στοίβας* είναι γενίκευση των *Πεπερασμένων Αυτομάτων*
- Ένα Αυτόματο Στοίβας λειτουργεί όπως ένα Πεπερασμένο Αυτόματο μόνο που
  - διαθέτει μία επιπλέον μνήμη υπό τη μορφή στοίβας
  - η μετάβαση στη επόμενη κατάσταση εξαρτάται από την τρέχουσα κατάσταση, το σύμβολο εισόδου και το περιεχόμενο τις στοίβας
- Στις επόμενες ενότητες ορίζονται τα *Αυτόματα Στοίβας* και αναλύονται οι παραπάνω προτάσεις

(Μη Αιτιοκρατικά) Αυτόματα Στοίβας

- Ορισμός: Αυτόματο Στοίβας (ΑΣ) είναι μία εξάδα $M=(K,\Sigma,\Gamma,\Delta,s,F)$ όπου
  - K πεπερασμένο σύνολο καταστάσεων
  - Σ το αλφάβητο των συμβολοσειρών εισόδου
  - Γ το αλφάβητο της στοίβας
  - s η αρχική κατάσταση
  - F το σύνολο των τελικών καταστάσεων ($F \subseteq K$)
  - Δ η σχέση μετάβασης, $\Delta \subseteq (K\times(\Sigma\cup\{e\})\times\Gamma^*) \times (K\times\Gamma^*)$
- Τρόπος Λειτουργίας: Σε κάθε κύκλο το αυτόματο
  - διαβάζει 1 σύμβολο εισόδου, a, από το w ή κανένα σύμβολο (a=e)
  - σύμφωνα με το νόμο μετάβασης $((p,a,\beta),(q,\gamma))\in\Delta$
    - διαβάζει (και διαγράφει) από την κορυφή της στοίβας το $\beta\in\Gamma^*$ και γράφει στην κορυφή της στοίβας ένα νέο σύμβολο/συμβολοσειρά $\gamma\in\Gamma^*$
    - μεταβαίνει σε μία νέα κατάσταση, q, ανάλογα με το σύμβολο, a, που διαβάστηκε, την τρέχουσα κατάσταση, p, και το σύμβολο β
- **Συμπεριφέρεται Μή Αιτιοκρατικά (πολλαπλά ενδεχόμενα)**

(Μη Αιτιοκρατικά) Αυτόματα Στοίβας

- Ανα πάσα στιγμή, η συνέχεια της λειτουργίας του ΑΣ ΔΕΝ εξαρτάται από τα παρελθόντα σύμβολα εισόδου
- Οι μελλοντικές μεταβάσεις καθορίζονται από την **ολική κατάσταση** του ΑΣ: $C = (q,w,v) \in K \times \Sigma^* \times \Gamma^*$ όπου
  - q η τρέχουσα κατάσταση
  - w η υποσυμβολοσειρά που απομένει να διαβαστεί
  - v η συμβολοσειρά της στοίβας διαβασμένη από πάνω προς τα κάτω
- Οι σχέσεις "**παράγει σε ένα βήμα**" ($\Rightarrow_M$) και το ανακλαστικό-μεταβατικό της κλείσμο "**παράγει**" ($\Rightarrow_M^*$) ορίζονται μεταξύ ολικών καταστάσεων όπως και στα ΜΑΠΑ
- **Υπολογισμός n βημάτων**: Η σταδιακή παραγωγή ολικών καταστάσεων $\mathbf{C_0 \Rightarrow_M C_1 \Rightarrow_M \ldots \Rightarrow_M C_n}$
- **Αποδεκτή γλώσσα** του ΑΣ, L(M), είναι το σύνολο των συμβολοσειρών εισόδου που είναι αποδεκτές από το αυτόματο (όμοια με τα ΑΠΑ και ΜΑΠΑ)

Αυτόματα Στοίβας – παράδειγμα #1

- Έστω M=(K,Σ,Γ,Δ,s,F) με
  - K={s, f}, Σ = {a, b, c}, Γ = {a, b}, F = {f}
  - Το σύνολο Δ περιέχει τις ακόλουθες μεταβάσεις:
    1. ((s,a,e),(s,a))
    2. ((s,b,e),(s,b))
    3. ((s,c,e),(f,e))
    4. ((f,a,a),(f,e))
    5. ((f,b,b),(f,e))
- Ο πίνακας στην επόμενη σελίδα καταγράφει έναν υπολογισμό για είσοδο w=abbcbba

Αυτόματα Στοίβας – παράδειγμα #1 (συν.)

Κατάσταση	Συμβολοσειρά που απομένει να διαβαστεί	Στοίβα	Κανόνας Μετάβασης που χρησιμοποιήθηκε
s	abbcbba	e	-
s	bbcbba	a	1 (“push a”)
s	bcbba	ba	2 (“push b”)
s	cbba	bba	2 (“push b”)
f	bba	bba	3
f	ba	ba	5 (“pop b”)
f	a	a	5 (“pop b”)
f	e	e	4 (“pop a”)

Το αυτόματο αποδέχεται συμβολοσειρές της μορφής wcw^R όπου $w \in \{a,b\}^*$.

Αυτόματα Στοίβας – παράδειγμα #2

- Έστω M=(K,Σ,Γ,Δ,s,F) με
  - K={s, f}, Σ = {a, b}, Γ = {a, b}, F = {f}
  - Το σύνολο Δ περιέχει τις ακόλουθες μεταβάσεις:
    1. ((s,a,e),(s,a))
    2. ((s,b,e),(s,b))
    3. ((s,e,e),(f,e))
    4. ((f,a,a),(f,e))
    5. ((f,b,b),(f,e))
- Ίδιο με το προηγούμενο παράδειγμα εκτός του κανόνα 3
- Ο πίνακας στην επόμενη σελίδα καταγράφει έναν υπολογισμό για είσοδο w=abbbba

Αυτόματα Στοίβας – παράδειγμα #2 (συν.)

Βήμα	Κατάσταση	Συμβολοσειρά που απομένει να διαβαστεί	Στοίβα	Κανόνας Μετάβασης που χρησιμοποιήθηκε
0	s	abbbba	e	-
1	s	bbbba	a	1 (“push a”)
2	s	bbba	ba	2 (“push b”)
3	s	bba	bba	2 (“push b”)
4	f	bba	bba	3 (αλλαγή κατάστασης)
5	f	ba	ba	5 (“pop b”)
6	f	a	a	5 (“pop b”)
7	f	e	e	4 (“pop a”)

Το αυτόματο αποδέχεται συμβολοσειρές της μορφής ww^R όπου $w \in \{a,b\}^*$. Το αυτόματο **Μη Αιτιοκρατικά** επέλεξε στο βήμα 4 να χρησιμοποιήσει τον κανόνα 3 και όχι τον 2.

Αυτόματα Στοίβας και Γλώσσες Χωρίς Συμφραζόμενα

- Το σύνολο των ΓΛΧΣ ταυτίζεται με το σύνολο των γλωσσών που αποδέχονται τα ΑΣ. Ειδικότερα:
  - Κάθε ΓΛΧΣ είναι αποδεκτή από κάποιο ΑΣ
    - G=(V,Σ,R,S)
    - M=(K,Σ,Γ,Δ,s,F)({p,q},Σ,V,Δ,p,{q}) όπου p,q αυθαίρετες καταστάσεις και το σύνολο μεταβάσεων Δ περιέχει:
      - ((p,e,e),(q,S))
      - ((q,e,A),(q,x)) για κάθε κανόνα $A \rightarrow x$ της R
      - ((q,a,a),(q,e)) για κάθε σύμβολο $a \in \Sigma$
  - Κάθε γλώσσα αποδεκτή από ΑΣ είναι ΓΛΧΣ
    - Είναι δυνατή (αλλά πιο περίπλοκη η κατασκευή γραμματικής G που να αντιστοιχεί σε οποιοδήποτε ΑΣ M)

Παράδειγμα κατασκευής ΑΣ από ΓΛΧΣ

- Έστω ΓΧΣ G = (V,Σ,R,S) με
  - V={S,a,b,c}, Σ = {a, b,c}, R = {$S \rightarrow aSa$, $S \rightarrow bSb$, $S \rightarrow c$}
  - Η ΓΧΣ G παράγει τη γλώσσα L(G)={wcw^R, $w \in$ {a,b}*}
- Το αντίστοιχο ΑΣ είναι M=({p,q},Σ,V,Δ,p,{q}) και το Δ περιέχει τις ακόλουθες μεταβάσεις:
  1. ((p,e,e),(q,S)) – προετοιμασία για εκκίνηση
  2. ((q,e,S),(q,aSa)) – αντικατάσταση σύμφωνα με τον κανόνα 1
  3. ((q,e,S),(q,bSb)) – αντικατάσταση σύμφωνα με τον κανόνα 2
  4. ((q,e,S),(q,c)) – αντικατάσταση σύμφωνα με τον κανόνα 3
  5. ((q,a,a),(q,e)) – pop a
  6. ((q,b,b),(q,e)) – pop b
  7. ((q,c,c),(q,e)) – pop c

Παράδειγμα κατασκευής ΑΣ από ΓΛΧΣ (συν.)

Βήμα	Κατάσταση	Συμβολοσειρά που απομένει	Στοίβα	Κανόνας Μετάβασης που χρησιμοποιήθηκε
0	p	abbcbba	e	-
1	q	abbcbba	S	1
2	q	abbcbba	aSa	2
3	q	bbcbba	Sa	5 (“pop a”)
4	q	bbcbba	bSba	3
5	q	bcbba	Sba	6 (“pop b”)
6	q	bcbba	bSbba	3
7	q	cbba	Sbba	6 (“pop b”)
8	q	cbba	cbba	4
9	q	bba	bba	7 (“pop c”)
10	q	ba	ba	6 (“pop b”)
11	q	a	a	6 (“pop b”)
12	q	e	e	5 (“pop a”)

Πράξεις Γλωσσών Χωρίς Συμφραζόμενα

Θεώρημα:

- Το σύνολο των ΓΛΧΣ **είναι κλειστό** ώς προς τις (μεταξύ γλωσσών) πράξεις
  1. ένωση
  2. παράθεση
  3. Kleene star
- Το σύνολο των ΓΛΧΣ **ΔΕΝ είναι κλειστό** ώς προς τις (μεταξύ γλωσσών) πράξεις
  1. τομή
  2. συμπλήρωμα

Ένωση ΓΛΧΣ

- Έστω δύο ΓΛΧΣ $L(G_1)$ και $L(G_2)$ με $G_1=(V_1,\Sigma_1,R_1,S_1)$ και $G_2=(V_2,\Sigma_2,R_2,S_2)$ τότε υπάρχει ΓΧΣ $G=(V,\Sigma,R,S)$ τέτοια ώστε $L(G)= L(G_1)\cup L(G_2)$.
- Η παραπάνω γραμματική, G, κατασκευάζεται ως εξής:
  - $V= V_1 \cup V_2 \cup \{S\}$
  - $\Sigma= \Sigma_1 \cup \Sigma_2$
  - $R= R_1 \cup R_2 \cup \{S\rightarrow S_1,\ S\rightarrow S_2\}$ (αποδείξτε το)

Παράθεση ΓΛΧΣ

- Έστω δύο ΓΛΧΣ $L(G_1)$ και $L(G_2)$ με $G_1=(V_1,\Sigma_1,R_1,S_1)$ και $G_2=(V_2,\Sigma_2,R_2,S_2)$ τότε υπάρχει ΓΧΣ $G=(V,\Sigma,R,S)$ τέτοια ώστε $L(G)= L(G_1)L(G_2)$.
- Η παραπάνω γραμματική, G, κατασκευάζεται ως εξής:
  - $V= V_1 \cup V_2 \cup \{S\}$
  - $\Sigma= \Sigma_1 \cup \Sigma_2$
  - $R= R_1 \cup R_2 \cup \{S \rightarrow S_1 S_2\}$ (αποδείξτε το)

Kleene Star ΓΛΧΣ

- Έστω ΓΛΧΣ $L(G_1)$ με $G_1=(V_!,\Sigma_1,R_1,S_1)$ τότε υπάρχει ΓΧΣ $G=(V,\Sigma,R,S)$ τέτοια ώστε $L(G)= L(G_1)^*$.
- Η παραπάνω γραμματική, G, κατασκευάζεται ως εξής:
  - $V= V_1 \cup \{S\}$
  - $\Sigma= \Sigma_1$
  - $R= R_1 \cup \{S\rightarrow e, S\rightarrow SS_1\}$ (αποδείξτε το)

Κανονική Μορφή Chomsky (KMC)

- Η ΓΧΣ $G=(V,\Sigma,R,S)$ είναι στην KMC όταν

$$R \subseteq (V-\Sigma) \times V^2$$

  - δηλαδή όλοι οι κανόνες είναι στη μορφή **A→BC** όπου A ΜΤΣ ($A \in V-\Sigma$) και $B,C \in V$
  - υπενθύμιση: στη γενική περίπτωση ΓΧΣ είναι

$$R \subseteq (V-\Sigma) \times V^*$$

  - θα δειχθεί στη συνέχεια ότι:
    - για κάθε ΓΧΣ υπάρχει μία "σχεδόν" ισοδύναμη γραμματική KMC
    - αν η γραμματική G είναι στην KMC, είναι αλγοριθμικά εύκολος (πολυωνυμικής πολυπλοκότητας) ο έλεγχος του αν μία συμβολοσειρά $x \in L(G)$ με τη μέθοδο του Δυναμικού Προγραμματισμού

Μετατροπή σε KMC

Θεώρημα:

Για κάθε ΓΧΣ G υπάρχει ΓΧΣ G' στην KMC τέτοια ώστε $L(G') = L(G) - (\Sigma \cup \{e\})$.

Δηλαδή από την L(G') λείπουν μόνο όσες συμβολοσειρές w έχουν μήκος $|w| \leq 1$.

Απόδειξη:

Εντοπίζουμε στην G τους κανόνες που αντιβαίνουν τις συνθήκες της KMC, δηλαδή:

1. Μακριοί κανόνες: $A \rightarrow B_1 B_2 \ldots B_n$, $n \geq 3$
2. Κανόνες προς το κενό: $A \rightarrow e$
3. Κοντοί κανόνες: $A \rightarrow B$

και σταδιακά τους εξαλείφουμε (βλ. παρακάτω)

Μετατροπή σε ΚΜC – εξάλειψη μακριών κανόνων

- Αντικαθιστούμε τους κανόνες της μορφής $A \to B_1 B_2 \dots B_n$, $n \geq 3$ με τους ακόλουθους n-1 κανόνες:
  - $A \to B_1 A_1$
  - $A_1 \to B_2 A_2$
  - ...
  - $A_{n-2} \to B_{n-1} B_n$ όπου $A_1, A_2, \dots, A_n$ νέα ΜΤΣ
- Η ΓΧΣ που προκύπτει είναι ακριβώς ισοδύμανη με την αρχική.

Μετατροπή σε ΚΜC – εξάλειψη κανόνων προς το κενό

- Υπολογίζουμε το σύνολο των "υπό διαγραφή ΜΤΣ" $E=\{A\in V-\Sigma: A \Rightarrow^* e\}$

  WHILE υπάρχει κανόνας $A\rightarrow\alpha$ με $\alpha \in E^*$ AND $A\notin E$

  DO πρόσθεσε το A στο E

- Όταν υπολογιστεί το E
  - Διαγράφουμε όλους τους κανόνες της μορφής $A\rightarrow e$
  - Για κάθε κανόνα της μορφής $A\rightarrow BC$ ή $A\rightarrow CB$ με $B\in E$ προσθέτουμε έναν κανόνα $A\rightarrow C$
- Από τη ΓΛΧΣ που αντιστοιχεί στη νέα ΓΧΣ λείπει μόνο το e (δε μπορεί πλέον να παραχθεί)

Μετατροπή σε ΚΜC – εξάλειψη κοντών κανόνων

- Για κάθε $A \in V$ υπολογίζουμε το σύνολο των συμβόλων που παράγονται από το A: $D(A)=\{B \in V: A \Rightarrow^* B\}$

  Αρχικοποίηση: $D(A)=\{A\}$

  WHILE υπάρχει κανόνας $B \rightarrow C$ με $B \in D(A)$ AND $C \notin D(A)$

  DO πρόσθεσε το C στο D(A)

  – Παρατήρηση: $A \in D(A)$ για κάθε ΜΤΣ A και $D(\alpha)=\{\alpha\}$ για κάθε ΤΣ α
- Μετά
  – Διαγράφουμε όλους τους κοντούς κανόνες
  – Αντικαθιστούμε όλους τους κανόνες $A \rightarrow BC$ με όλους τους δυνατούς κανόνες $A \rightarrow B'C'$ με $B' \in D(B)$ και $C' \in D(C)$
  – Για κάθε κανόνα $A \rightarrow BC$ με $A \in D(S)-\{S\}$ προσθέτουμε τον κανόνα $S \rightarrow BC$
- Από τη ΓΛΧΣ που αντιστοιχεί στη νέα ΓΧΣ λείπουν μόνο τα $\alpha \in \Sigma$ (συμβολοσειρές μήκους 1 δε μπορεί πλέον να παραχθούν)

Μετατροπή σε ΚΜC – Παράδειγμα

- Εστω ΓΧΣ G με κανόνες R={$S \rightarrow SS$, $S \rightarrow (S)$, $S \rightarrow e$} που παράγει τη γλώσσα των ταιριασμένων παρενθέσεων.
- Εξάλειψη μακριών κανόνων:
  - Αντικαθιστούμε τον $S \rightarrow (S)$ με τους $S \rightarrow (S_1$ και $S_1 \rightarrow S)$,
  - οπότε προκύπτουν οι κανόνες R={$S \rightarrow SS$, $S \rightarrow (S_1$, $S_1 \rightarrow S)$, $S \rightarrow e$} .
- Εξάλειψη κανόνων προς το κενό
  - Υπολογίζουμε ότι E={S},
  - οπότε προκύπτουν οι κανόνες R={$S \rightarrow SS$, $S \rightarrow (S_1$, $S_1 \rightarrow S)$, $S \rightarrow S$, $S \rightarrow)$} και επειδή ο $S \rightarrow S$ δεν προσφέρει τίποτε έχουμε R={$S \rightarrow SS$, $S \rightarrow (S_1$, $S_1 \rightarrow S)$, $S_1 \rightarrow)$ }.
- Εξάλειψη κοντών κανόνων
  - Υπολογίζουμε ότι $D(S_1)=\{S_1,)\}$, $D(S)=\{S\}$, $D( ))=\{)\}$, $D(( )=\{(\}$,
  - οπότε λαμβάνοντας υπόψη μόνο το "μη τετριμμένο" $D(S_1)=\{S_1,)\}$ προκύπτουν οι κανόνες R={$S \rightarrow SS$, $S \rightarrow (S_1$, $S_1 \rightarrow S)$, $S \rightarrow ()$} όπου ο $S \rightarrow ()$ προστέθηκε επειδή $S \rightarrow (S_1$ και $) \in D(S_1)$.

Πολυπλοκότητα Μετατροπής σε ΚΜC

- Έστω ΓΧΣ G μεγέθους n
  - "μέγεθος, n, της γραμματικής G" εννούμε το μήκος μιας συμβολοσειράς που περιγράφει την G. Πρακτικά, n≈(το συνολικό μήκος των συμβολοσειρών που αναπαριστούν τους κανόνες της G)
- Εξάλειψη μακριών κανόνων
  - πολυπλοκότητα = O(n)
  - η νέα γραμματική έχει μέγεθος O(n)
- Εξάλειψη κανόνων προς το κενό
  - πολυπλοκότητα = $O(n^2)$ για τον υπολογισμό του E + O(n) για την προσθήκη των κανόνων
- Εξάλειψη κοντών κανόνων
  - πολυπλοκότητα = $O(n^3)$
- ΠΟΛΥΩΝΥΜΙΚΗ ΣΥΝΟΛΙΚΗ ΠΟΛΥΠΛΟΚΟΤΗΤΑ $O(n^3)$

Δυναμικός Προγραμματισμός

- Οι αλγόριθμοι Δυναμικού Προγραμματισμού λύνουν ένα πρόβλημα ξεκινώντας από μικρότερα προβλήματα και φτάνοντας, αναδρομικά, στη λύση του τελικού
- Αν G’ ΓΧΣ στην KMC τότε ο έλεγχος

$$\text{“}x \in L(G') \text{ ;”}$$

επιτυγχάνεται με Δυναμικό Προγραμματισμό

- Ελέγχουμε τις υποσυμβολοσειρές του x αρχίζοντας από αυτές μήκους 1
- Επειδή η G’ είναι στην KMC το $x \in L(G')$ αν και μόνο αν υπάρχουν $B,C \in L(G')$ ώστε $x=BC$
- Ο αλγόριθμος περιγράφεται στη συνέχεια ...

Έλεγχος “$x \in L(G')$;” για G’ στην KMC

- Έστω ότι η συμβολοσειρά $x=x_1x_2\ldots x_n$ ($x_i \in \Sigma$) με μήκος n.
- Ορίζουμε N[i, i+s] = {σύμβολα του V που παράγουν την υποσυμβολοσειρά $x_i\ldots x_{i+s}$}
- Υπολογίζουμε αναδρομικά τα N[i, i+s] για κάθε υποσυμβολοσειρά μήκους s+1 (s=1…n-1) με βάση τα ήδη υπολογισμένα, N[i,k], N[k+1,i+s], (που αντιστοιχούν σε ζεύγη υποσυμβολοσειρών, $x_i\ldots x_k$, $x_{k+1}\ldots x_{i+s}$, που την αποτελούν)

```
Αρχικοποίηση: FOR i=1…n DO N[i,i]={x_i}
FOR s=1…n-1 DO
   FOR i=1…n-s DO
     FOR k=1…i+s-1 DO
       IF υπάρχει κανόνας A→BC με B∈N[i,k] AND        C∈N[k+1,i+s]
       THEN πρόσθεσε το A στο N[i,i+s]
```

- Το $x \in L(G')$ αν και μόνο αν $S \in N[1,n]$
- **Πολυπλοκότητα: $O(|x|^3|G'|)$ Πολυωνυμική !**

Παράδειγμα ελέγχου

- Έστω η G' στην KMC που είχε παραχθεί στο προηγούμενο παράδειγμα με κανόνες

$$R=\{S\rightarrow SS,\ S\rightarrow(S_1,\ S_1\rightarrow S),\ S\rightarrow()\}$$

  και η συμβολοσειρά x = (( ) (( )))
- Κάθε κελί (i,j) του παρακάτω πίνακα περιέχει τα στοιχεία του συνόλου N(i,j) όπως υπολογίζονται από τον αλγόριθμο
  - Ο υπολογισμός ξεκινά από τα διαγώνια στοιχεία
  - προχωρά σε διαδοχικές δευτερεύουσες διαγωνίους
  - ο υπολογισμός κάθε κελιού χρησιμοποιεί μόνο τα κελιά που βρίσκονται
    - στην ίδια γραμμή και αριστερά
    - στην ίδια στήλη και προς τα πάνω
- Το κελί (1,n) που αντιστοιχεί στο N(1,n) συμπληρώνεται τελευταίο. Στο παράδειγμα περιέχει το S, άρα $x \in L(G')$.

Παράδειγμα ελέγχου (συν.)

Για παράδειγμα στον υπολογισμό του N(4,7) ελέγχουμε διαδοχικά αν υπάρχει κανόνας $A \rightarrow BC$ με

- $B \in N(4,4)$ και $C \in N(5,7)$
  - υπάρχει ο $S \rightarrow (S_1$
- $B \in N(4,5)$ και $C \in N(6,7)$
  - δεν υπάρχει
- $B \in N(4,6)$ και $C \in N(7,7)$
  - δεν υπάρχει

Άρα N(4,7) ={S}

8								)
7							)	∅
6						)	∅	∅
5					(	S	S_1	∅
4				(	∅	∅	S	S_1
3			)	∅	∅	∅	∅	∅
2		(	S	∅	∅	∅	S	S_1
1	(	∅	∅	∅	∅	∅	∅	**S**
	1	2	3	4	5	6	7	8

Αλγόριθμοι για ΓΛΧΣ

1. ΓΧΣ $G=(V,\Sigma,R,S) \rightarrow$ ΑΣ $M=(K,\Sigma,\Gamma,\Delta,s,F)$

   πολυπλοκότητα: πολυωνυμική

2. ΑΣ $M=(K,\Sigma,\Gamma,\Delta,s,F) \rightarrow$ ΓΧΣ $G=(V,\Sigma,R,S)$

   πολυπλοκότητα: πολυωνυμική

3. **έλεγχος για το αν η συμβολοσειρά $x \in L(G)$ όπου G=ΓΧΣ**

   πολυπλοκότητα: πολυωνυμική (=μετατροπή της G σε KMC + Δυναμικός προγραμματισμός)

Αιτιοκρατικά Αυτόματα Στοίβας (ΑΑΣ)

- Ένα ΑΣ M είναι αιτιοκρατικό όταν από κάθε ολική κατάσταση C_n μπορεί να προκύψει το πολύ μία επόμενη κατάσταση C_{n+1}
- Μία γλώσσα L λέγεται Αιτιοκρατική Γλώσσα Χωρίς Συμφραζόμενα (ΑΓΛΧΣ) όταν υπάρχει ΑΑΣ M τέτοιο ώστε

$$L\$ = L(M)$$

όπου το νέο σύμβολο (delimiter) $\$ \notin \Sigma$ σηματοδοτεί το τέλος των συμβολοσειρών που παραδίδονται στο αυτόματο

- Υπάρχουν ΑΣ που ΔΕΝ είναι (ούτε ισοδυναμούν με) ΑΑΣ
  - αντιπαράβαλε με τη σχέση ΑΠΑ/ΜΑΠΑ
- Υπάρχουν ΓΧΣ που ΔΕΝ είναι Αιτιοκρατικές