Θεωρία Υπολογισμού και Πολυπλοκότητα

Κεφάλαιο 7.
Κατηγορηματικές Γραμματικές

11 Μαρτίου 2008

Δρ. Παπαδοπούλου Βίκη

Κατηγορηματικές Γραμματικές Ή Γραμματικές Χωρίς Συμφραζόμενα

Παράδειγμα. $L = a\,(a^{*} \cup b^{*})\,b$

- Μια λέξη μπορούμε να παραχθεί ως εξής:
  - S = συμβολοσειρά ή μη τερματικό
  - $S \rightarrow aMb$ (R1)
  - $M \rightarrow A \mid (ή)\ B$ (R2)
  - $A \rightarrow e \mid aA$ (R3)
  - $B \rightarrow e \mid bB$ (R4)

  Κανόνες R
- a, b, M, S : **αλφάβητο**
- M, S, A, B : **μη τερματικά ή μεταβλητές** a, b : **τερματικά σύμβολα**
- Η aaab μπορεί να παραχθεί ως εξής:

$$
\begin{aligned}
S &\Rightarrow a\,M\,b && (R1)\\
&\Rightarrow a\,A\,b && (R2)\\
&\Rightarrow a\,aA\,b && (R3)\\
&\Rightarrow a\,a\,a\,b && (R3)
\end{aligned}
$$

Γραμματικές Χωρίς Συμφραζόμενα

- Μια Γραμματική χωρίς Συμφραζόμενα G είναι μια τετράδα (V, Σ, R, S):
  - V = αλφάβητο
  - Σ = τερματικά, υποσύνολο του V
  - V-Σ = μη τερματικά σύμβολα ή μεταβλητές
  - R = σύνολο από κανόνες $(V\text{-}\Sigma) \rightarrow V^*$
  - S =αρχικό σύμβολο , $\in V\text{-}\Sigma$ (μη τερματικό)
  - Για οποιεσδήποτε λέξεις $\alpha_1, \alpha_2 \in V^*$, $\alpha_1 \Rightarrow_G \alpha_2$ αν:
    - υπάρχει μεταβλητή $A \in V\text{-}\Sigma$ και λέξεις $\alpha, \beta, \gamma \in V^*$ τέτοιες ώστε
      - $\alpha_1 = \beta\, A\, \gamma$,
      - $A \rightarrow \alpha$ και
      - $\alpha_2 = \beta\, \alpha\, \gamma$ Π.χ. $S \rightarrow aMb$, $M \rightarrow A$, $S \Rightarrow a\, A\, b$
    - Δηλ. $\alpha_1 \Rightarrow_G \alpha_2$: το α_2 προκύπτει από το α_1 αν αντικαταστήσουμε την μεταβλητή A που περιέχεται στο a_1.

Παράδειγμα 2.

- Παράδειγμα.
- $G = (V, \Sigma, R, S)$, $\Sigma = \{a,b\}$, $R = \{ S \rightarrow aSb , S \rightarrow e \}$
  - Μπορεί να παραχθεί η aabb?
  - Μπορεί να παραχθεί η aab?
  - Ποια γλώσσα παράγει?
    - Την $L(G) = \{ a^n b^n : n \geq 0\}$
    - η οποία δεν είναι κανονική γλώσσα!
  - Παράδειγμα παραγωγής: a a a b b b
    - $S \Rightarrow aSb \Rightarrow a\, aSb\, b \Rightarrow a\, a\, aSb\, b\, b \Rightarrow a\, a\, a\, b\, b\, b$

- Σχόλιο. Λέγονται γραμματικές χωρίς συμφραζόμενα επειδή στους κανόνες αντικαταστούμε ένα μη τερματικό Α ασχέτως με την υπολέξη πριν και μετά το μη-τερματικό Α.

Κατηγορηματικές Γλώσσες

- Η $\Rightarrow^*_G$ είναι η ανακλαστική, μεταβατική θήκη της $\Rightarrow_G$
- Η γλώσσα που παράγεται από την G είναι η :
  - $L(G) = \{ w \in \Sigma^* : S \Rightarrow^*_G w \}$
  - w: μόνο τερματικά σύμβολα

Παράδειγμα 4.

- Κατασκευάστε μια κατηγορηματική γραμματική για την παρακάτω γλώσσα:
  - L = {w\in {a,b}*, στην w το a και b παρουσιάζονται τον ίδιο αριθμό φορών}

  - $G = (V, \Sigma, R, S)$,
  - $V=\{a, b, S, A, B\}$
  - $\Sigma = \{a,b\}$,
  - R =
    - $S \to aB$ (αν παρουσιαστεί a χρωστά στην B)
      $\mid bA$ (αν παρουσιαστεί a χρωστά στην B)
      $\mid e$
    - $A \to aS \mid bAA$
      (ή θα δώσει το a που χρωστά ή θα χρωστά ένα a περισσότερο)
    - $B \to bS \mid aBB$
      (αντίστοιχα)

Παράδειγμα 3.

- Κατασκευάστε μια κατηγορηματική γραμματική για την παρακάτω γλώσσα: $L = \{a^m b^n : m \geq n\}$

  - $G = (V, \Sigma, R, S)$,
  - $V=\{a, b, S\}$
  - $\Sigma = \{a,b\}$,
  - $R = \{ S\rightarrow aSb , S\rightarrow aS, S\rightarrow e \}$

Σχέσεις κλειστότητας

- Θεώρημα 1. Οι γλώσσες χωρίς συμφραζόμενα είναι κλειστές ως προς την ένωση, την παράθεση και την θήκη Kleene.
- Απόδειξη.
- Ένωση. $G_1 = (V_1, \Sigma_1, R_1, S_1)$, $G_2 = (V_2, \Sigma_2, R_2, S_2)$, έστω $V_1 - \Sigma_1 \cap V_2 - \Sigma_2 = \emptyset$.
  - $L(G_1) \cup L(G_2) = L(G) = (V_1 \cup V_2 \cup \{S\}, \Sigma_1 \cup \Sigma_2, R, S)$,
  - $R = R_1 \cup R_2 \cup \{S \rightarrow S_1, S \rightarrow S_2 \}$
- Παράθεση.
  - $L(G_1)L(G_2) = L(G) = (V_1 \cup V_2 \cup \{S\}, \Sigma_1 \cup \Sigma_2, R_1 \cup R_2 \cup \{S \rightarrow S_1 S_2 \}, S)$
- θήκη Kleene.
  - $(L(G_1))^* = L(G) = (V_1 \cup \{S\}, \Sigma_1, R_1 \cup \{S \rightarrow e , SS_1 \}, S)$

Τομή

- Θεώρημα 1(b). Οι γλώσσες χωρίς συμφραζόμενα δεν είναι είναι κλειστές ως προς την τομή.
- Απόδειξη.

Επόμενη διάλεξη.

Κανονικές Γλώσσες

- Θεώρημα 2. Όλες οι κανονικές γλώσσες είναι γλώσσες χωρίς συμφραζόμενα.

Απόδειξη1.

- Οι κανονικές γλώσσες ορίζονται επαγωγικά χρησιμοποιώντας τις πράξεις τις ένωσης, της παράθεση και της θήκη Kleene και είναι κλειστές ως προς τις πράξεις αυτές.
- Οι γλώσσες χωρίς συμφραζόμενα είναι κλειστές ως προς την ένωση, την παράθεση και την θήκη Kleene.

$\Rightarrow$ Όλες οι κανονικές γλώσσες είναι γλώσσες χωρίς συμφραζόμενα.

Απόδειξη2. Άμεση κατασκευή.

Κανονικές Γλώσσες

- Θεώρημα 3. Όλες οι κανονικές γλώσσες είναι γλώσσες χωρίς συμφραζόμενα.

Απόδειξη2. Άμεση κατασκευή.

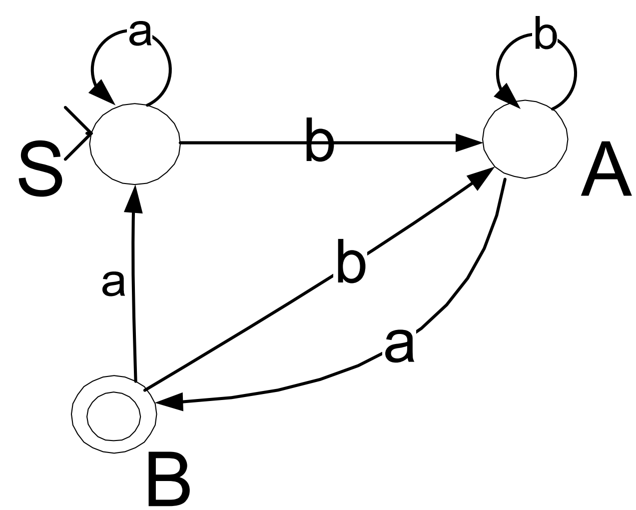
- Έστω μια κανονική γλώσσα που γίνεται δεκτή από το ντετερμινιστικό πεπερασμένο αυτόματο $M=(Q,\Sigma, \delta, s, F)$.
- Η ίδια γλώσσα παράγεται από την γραμματική $G(M)=(V, \Sigma, R, S)$:
  - $V = (Q \cup \Sigma)$
  - $\Sigma = \Sigma$: Τερματικά σύμβολα είναι το αλφάβητο Σ του M
  - $V\text{-}\Sigma = Q$: τα μη τερματικά σύμβολα είναι οι καταστάσεις Q του M
  - $S = s$
  - (για κάθε μεταβίβαση από την q στην p με είσοδο a έχουμε ένα κανόνα, τον $q \rightarrow ap$)

    $R = \{ q \rightarrow ap : \forall a \in \Sigma, \; \delta(q, a) = p \} \cup \{ q \rightarrow e : q \in F \}$

Παράδειγμα

- Έστω μια κανονική γλώσσα που γίνεται δεχτή από το αυτόματο αριστερά. Υπολογίστε την κατηγορηματική γλώσσα την οποία δέχεται.
- G(M)=(V, Σ, R, S):
  - $V = (Q \cup \Sigma) =$(S, A, B,a, b)
  - V-Σ : μη τερματικά
  - Σ ={a,b} : Τερματικά
  - Κανονες R:

  (για κάθε μεταβίβαση από την q στην p με είσοδο a έχουμε ένα κανόνα, τον $q \rightarrow ap$)
    - $S \rightarrow aS$
    - $S \rightarrow bA$
    - $A \rightarrow aB$
    - $A \rightarrow bA$
    - $B \rightarrow aS$
    - $B \rightarrow bA$
    - $B \rightarrow e$



Κανονικές Γραμματικές

- Ορισμός. Μια κατηγορηματική γραμματική (γλώσσα) είναι κανονική αν
  1. το δεξιό μέλος κάθε κανόνα περιέχει το πολύ μία μεταβλητή A (είναι γραμμική).
  2. η μεταβλητή A βρίσκεται στην τελευταία δεξιότερη θέση.
- Μία γλώσσα L καλείται γλώσσα κανονικής γραμματικής αν υπάρχει κανονική γραμματική G τέτοια ώστε $L = L(G)$.

Κανονικές Γλώσσες

Θεώρημα 4. Μια γλώσσα είναι κανονική αν και μόνο υπάρχει μια κανονική γραμματική που την παράγει.

Απόδειξη.

- Αν η γλώσσα είναι κανονική υπάρχει μια κανονική γραμματική που την παράγει:
  - (Θεώρημα 3): Όλες οι κανονικές γλώσσες είναι κατηγορηματικές.
  - Προσέξτε: η κατηγορηματική γραμματική που κατασκευάζεται στο Θεώρημα είναι κανονική.
- Αν έχουμε μια κανονική γραμματική $G(M)=(V, \Sigma, R, S)$ υπάρχει ένα μη-ντετερμινιστικό αυτόματο που την εκφράζει.
  - το αυτόματο $M = (V\text{-}\Sigma \cup \{f\}, \Sigma, \delta, S, \{f\})$, R: παράγει ένα τερματικό κάθε φορά
  - $\delta = \{ \delta(A, a) = B : \{ (A \rightarrow aB) \in R,\ a \in \{\Sigma \cup e\}$
    $\cup \{\delta(A, a) = f : \{(A \rightarrow a) \in R,\ a \in \{\Sigma \cup e\} \}$ ■

Παράδειγμα

- Η γραμματική G(M)=(V, Σ, R, S), V={a,b,A,B,S}, Σ={a,b},

  R = {S→ abA, S→ B, A→ bS, B→ aS, A→ b, S→ e }   είναι κανονική.

- Υπολογίστε ένα πεπερασμένο αυτόματο που να δέχεται την γλώσσα που παράγει η γραμματική.

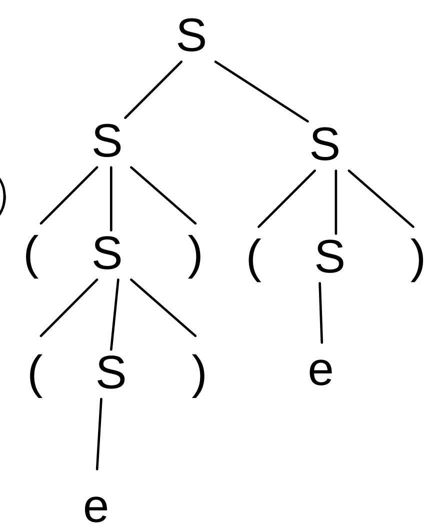
R’ = {S→ aA_1, A_1→ bA, S→ B, A→ bS, B→ aS, A→ b, S→ e }.

  - S=αρχική κατάσταση, f = τελική κατάσταση
  - δ (S,a) = A_1
  - δ (A_1,b) = A
  - δ (S,e) = B
  - δ (A,b) = S
  - δ (B,a) = S
  - δ (A,b) = f
  - δ (S,e) = f

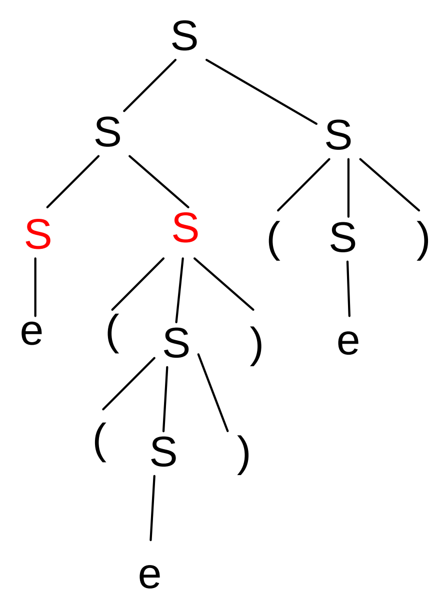
- Θεώρημα 5. Η τομή μιας κατηγορηματικής γλώσσας με μια κανονική γλώσσα είναι κατηγορηματική γλώσσα.

Συντακτικά Δένδρα

- Παραγωγή της (())():
  - D_1 = SS⇒ (S)S ⇒ ((S))S ⇒ (())S ⇒ (())(S) ⇒ (())()
  - D_2 = SS⇒ (S)S ⇒ ((S))S ⇒ ((S))(S) ⇒ (())(S) ⇒ (())()



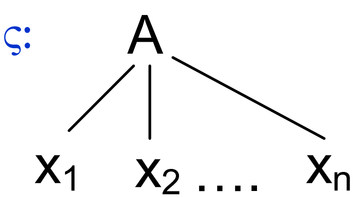
- Ισοδύναμες παραγωγές:
  - Εφαρμογή ίδιου συνόλου κανόνων στα ίδια σημεία.
  - Διαφορά στην σειρά εφαρμογής των κανόνων



- D_3 = SS⇒ SSS ⇒ S(S)S
  ⇒ $\mathcal{S}$((S))S ⇒ $\mathcal{S}$(())(S)
  ⇒ $\mathcal{S}$(())() ⇒(())()
- Εφαρμόστηκε διαφορετικό σύνολο κανόνων.

Συντακτικά Δένδρα

- **Ορισμός.** Για μια γραμματική χωρίς συμφραζόμενα $G=(V, \Sigma, R, S)$ ένα συντακτικό δένδρο μπορεί να κατασκευαστεί ως εξής:
  1. Έστω S η ρίζα του δένδρου.
  2. Επαναλαμβάνουμε την παρακάτω διαδικασία μέχρι κάθε φύλλο του δένδρου είναι είτε η κενή λέξη ε, ή τελικό σύμβολο.
     - Για κάθε φύλλο $A \in V$ του δένδρου, επιλέγουμε ένα κανόνα $A \rightarrow x_1 x_2 \ldots x_n$, όπου $x_i \in V \cup \Sigma$ για $i = 1, \ldots, n$, και
     - αντικαθιστούμε τον κόμβο A στο V ως εξής:



- Αν η διαδοχική παράθεση των φύλλων του δένδρου είναι η λέξη $x \in \Sigma^*$, τότε το δένδρο είναι το συντακτικό δένδρο της x.
- Για κάθε λέξη $x \in L(G)$, υπάρχει ένα συντακτικό δένδρο στο οποίο διαδοχική παράθεση των φύλλων ισούται με την x.

Παράδειγμα

- Για την γραμματική $G=(V, \{a,b,c\}, R, S)$,
  $R=\{S\rightarrow SbS$ (1), $S\rightarrow ScS$ (2), $S\rightarrow a$ (3) $\}$. Δώστε παραγωγή της λέξης abaca.

Αριστερή (Δεξιά) Παραγωγή

- Μια παραγωγή της x∈ L(G) είναι αριστερή (δεξιά) παραγωγή $\Rightarrow^{*\,L}$ αν κάθε φορά αντικαταστούμε με το μη τερματικό σύμβολο στα αριστερότερα (δεξιότερα) της λέξης.
- Για κάθε λέξη x∈ L(G) υπάρχει μόνο μια αριστερή (δεξιά) παραγωγή
  - Πχ. $G_1 = (S, \{a,b,c\}, R, S)$, $R = \{S \rightarrow SbS \mid ScS \mid a\}$
  - abaca $\in L(G)$: $S \Rightarrow^{L} SbS \Rightarrow^{L} abS \Rightarrow^{L} abScS \Rightarrow^{L} abacS \Rightarrow^{L} abaca$.
- Συμπερασματικά, για μια γραμματική χωρίς συμφραζόμενα G=(V, Σ, R, S), και A∈ VΣ, οι παρακάτω προτάσεις είναι ισοδύναμες:
  - $A \Rightarrow^{*} w$
  - Υπάρχει συντακτικό δένδρο με ρίζα A και παράγει την w.
  - Υπάρχει αριστερή παραγωγή $A \Rightarrow^{*L} w$.
  - Υπάρχει αριστερή παραγωγή $A \Rightarrow^{*L} w$.

Συντακτικά Δένδρα

- Ερώτηση. Για μια γραμματική $G=(V, \Sigma, R, S)$, η λέξη w ανήκει στη γλώσσα?
- Π.χ. $G=(V, \{a,b,c\}, R, S)$, $R=\{S\rightarrow SbS\ (1), S\rightarrow ScS\ (2), S\rightarrow a\ (3)\}$. Η λέξη abaca ανήκει στην γλώσσα?
  - Φτιάχνω ένα δένδρο ξεκινώντας από το S
  - για κάθε λέξη w (με τερματικά και μη τερματικά σύμβολα) που μπορεί να παραχθεί από το S
  - Προσθέτω σαν παιδιά όλες τις λέξεις που μπορεί να παραχθούν από την w.
  - Συνεχίζω έως ότου φτάσω σε τερματικά σύμβολα.