ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ

ΑΝΟΙΚΤΑ ΑΚΑΔΗΜΑΙΚΑ ΜΑΘΗΜΑΤΑ

Θεωρία Υπολογισμού

Ενότητα 15: Συντακτικά Δέντρα

Επ. Καθ. Π. Κατσαρός
Τμήμα Πληροφορικής

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Άδειες Χρήσης

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «Ανοικτά Ακαδημαϊκά Μαθήματα στο Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης» έχει χρηματοδοτήσει μόνο την αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

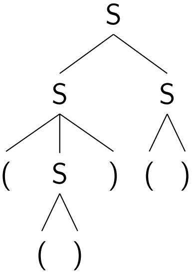
1. Τι περιγράφει ένα ΣΔ
   - Δομή συντακτικών δέντρων
   - Παραγωγές συμβολοσειρών
   - Τυπικός ορισμός ΣΔ
2. ΣΔ και παραγωγές
   - Όμοιες παραγωγές
   - Αριστερότερες και δεξιότερες παραγωγές
   - Θεώρημα Συντακτικού Δέντρου
   - Διφορούμενες Γραμματικές

Τι περιγράφει ένα Συντακτικό Δέντρο

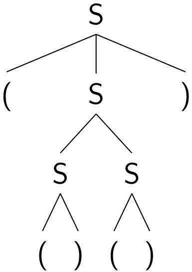
- Τα **Συντακτικά Δέντρα** είναι δεντρικές δομές, που οι κόμβοι τους επισημειώνονται με σύμβολα μιας γραμματικής χωρίς συμφραζόμενα.
- Οι εσωτερικοί κόμβοι επισημειώνονται με σύμβολα μεταβλητών. Για κάθε κόμβο-γονέα οι κόμβοι-απόγονοι επισημειώνονται με τα σύμβολα του δεξιού μέρους ενός κανόνα της γραμματικής.
- Ο κόμβος της **ρίζας** επισημειώνεται με το σύμβολο αρχής της γραμματικής.
- Τα **φύλλα** επισημειώνονται με τα τερματικά σύμβολα της **παραγόμενης συμβολοσειράς** ή με ϵ.

Παραδείγματα Συντακτικών Δέντρων

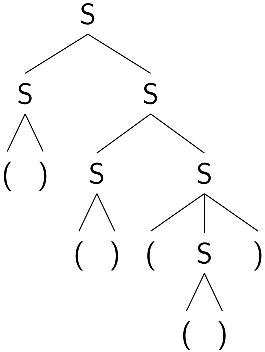
$$S \rightarrow S\ S \mid (\ S\ ) \mid (\ )$$



συμβολοσειρά:
(())()



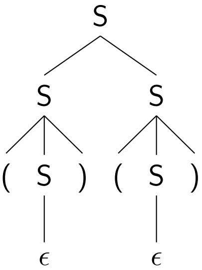
συμβολοσειρά:
(()())



συμβολοσειρά:
()()(())

Παραγωγές συμβολοσειρών

- Η ίδια συμβολοσειρά μπορεί να παραχθεί με διαφορετικές παραγωγές, αν εφαρμοστούν οι ίδιοι κανόνες με διαφορετική σειρά.
  - $S \Rightarrow SS \Rightarrow (S)S \Rightarrow ()S \Rightarrow ()(S) \Rightarrow ()()$
  - $S \Rightarrow SS \Rightarrow S(S) \Rightarrow (S)(S) \Rightarrow (S)() \Rightarrow ()()$
- Οι δύο αυτές παραγωγές μπορούν να απεικονιστούν σε ένα συντακτικό δέντρο:



Τυπικός Ορισμός ΣΔ

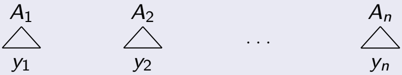
Ορισμός 1 (Συντακτικό δέντρο)

Για μια ΓΧΣ $G = (V, \Sigma, R, S)$:

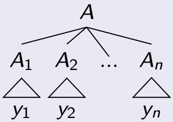
- Για κάθε $\alpha \in \Sigma$ ο γράφος με ένα μόνο κόμβο, α είναι συντακτικό δέντρο με ρίζα και μοναδικό φύλλο τον κόμβο α.
- Αν $A \longrightarrow \epsilon$ είναι κανόνας του R, τότε το

$$\begin{array}{c} A \\ | \\ \epsilon \end{array}$$

  είναι συντακτικό δέντρο με ρίζα τον κόμβο A και μοναδικό φύλλο τον ϵ.
- Αν τα



  είναι συντακτικό δέντρα με ρίζες $A_1, A_2, ..., A_n$ και με παραγόμενες συμβ/ρές τις $y_1, y_2, ..., y_n$ και αν $A \longrightarrow A_1 A_2 ... A_n$ είναι κανόνας του R, τότε το



  είναι συντακτικό δέντρο με ρίζα A και παραγόμενη συμβ/ρά την $y_1 y_2 ... y_n$.

Σύγκριση παραγωγών

- Τα συντακτικά δέντρα αναπαριστούν κλάσεις ισοδυναμίας παραγωγών, δηλ. παραγωγές που διαφέρουν μόνο στη σειρά εφαρμογής των ίδιων κανόνων.

Ορισμός 2 (Σχέση «προηγείται»)

Για $G = (V, \Sigma, R, S)$ έστω $D \Rightarrow x_1 \Rightarrow x_2 \Rightarrow \ldots \Rightarrow x_n$ και $D' \Rightarrow x_1' \Rightarrow x_2' \Rightarrow \ldots \Rightarrow x_n'$ δύο παραγωγές της G με $x_i, x_i' \in V^*$, $i = 1, \ldots, n$ και $x_1, x_1' \in V - \Sigma$, x_n, x_n' $\in \Sigma^*$. Λέμε ότι η D **προηγείται** της D' ($D \prec D'$) αν $n > 2$ και υπάρχει k, $1 \leq k \leq n$, τέτοιος ώστε

1. για κάθε $i \neq k$ ισχύει $x_i = x_i'$
2. $x_{k-1} = x_{k-1}' = uAvBw$, με $u, v, w \in V^*$ και $A, B \in V - \Sigma$
3. $x_k = uyvBw$ για $A \to y \in R$ και $x_k' = uAvzw$ για $B \to z \in R$
4. $x_{k+1} = x_{k+1}' = uyvzw$

δηλαδή προηγείται η παραγωγή στην οποία το μη τερματικό σύμβολο που βρίσκεται πιο αριστερά αντικαθιστάται πριν από το άλλο.

Παράδειγμα σύγκρισης παραγωγών

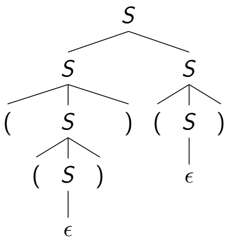
- Έστω οι παραγωγές:

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

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

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

  Ισχύει $D_1 \prec D_2$ και $D_2 \prec D_3$
- Δεν ισχύει η $D_1 \prec D_3$, επειδή οι δύο παραγωγές διαφέρουν σε περισσότερες από μία ενδιάμεσες συμβολοσειρές.
- Όλες οι παραγωγές αντιστοιχούν στο συντακτικό δέντρο



Ομοιότητα παραγωγών

Ορισμός 3 (Όμοιες παραγωγές)

Δύο παραγωγές D και D' είναι **όμοιες**

- αν το ζεύγος (D,D') ανήκει στην *ανακλαστική, συμμετρική, μεταβατική κλειστότητα* της $\prec$ ή
- αν μπορούν να μετασχηματιστούν η μία στην άλλη μέσω μιας ακολουθίας εναλλαγών στη σειρά εφαρμογής των κανόνων της γραμματικής.

- Οι D_1, D_2 και D_3 του προηγούμενου παραδείγματος είναι όμοιες.
- Όμως μπορεί να παραχθεί η ίδια συμβολοσειρά με παραγωγές που δεν είναι όμοιες (αναπαριστώνται από διαφορετικά συντακτικά δέντρα)
  - $D = S \Rightarrow SS \Rightarrow (S)S \Rightarrow ((S))S \Rightarrow (())S \Rightarrow (())(S) \Rightarrow (())()$
  - $F = S \Rightarrow SS \Rightarrow SSS \Rightarrow S(S)S \Rightarrow S((S))S \Rightarrow S(())S \Rightarrow S(())(S) \Rightarrow S(())() \Rightarrow (())()$

Αριστερότερες και δεξιότερες παραγωγές

- Κάθε συντακτικό δέντρο (δηλ. κλάση ισοδυναμίας παραγωγών) έχει μία μοναδική παραγωγή που είναι *μέγιστη* ως προς την $\prec$ (καμία άλλη παραγωγή δεν προηγείται αυτής) και ονομάζεται **αριστερότερη** παραγωγή.
- Κατασκευάζεται ως εξής: ξεκινάμε από την επιγραφή της ρίζας και σε κάθε συμβολοσειρά που προκύπτει αντικαθιστούμε το αριστερότερο μη τερματικό σύμβολο.
- Η **δεξιότερη** παραγωγή κατασκευάζεται αντιστοίχως, με αντικατάσταση σε κάθε βήμα του δεξιότερου μη τερματικού.
- Κάθε συντακτικό δέντρο έχει ακριβώς μία αριστερότερη και μία δεξιότερη παραγωγή, αφού σε κάθε βήμα αντικαθίσταται ένα και μόνο ένα μη τερματικό.
- Αν $x_1 \Rightarrow x_2 \Rightarrow \ldots \Rightarrow x_n$ είναι αριστερότερη παραγωγή, τότε τη γράφουμε ως $x_1 \overset{L}{\Rightarrow} x_2 \overset{L}{\Rightarrow} \ldots \overset{L}{\Rightarrow} x_n$
- Αντίστοιχα για μία δεξιότερη παραγωγή χρησιμοποιούμε το $\overset{R}{\Rightarrow}$

Θεώρημα για συντακτικά δέντρα

Θεώρημα 4 (ΣΔ και παραγωγές)

Έστω $G = (V, \Sigma, R, S)$ μία γραμματική χωρίς συμφραζόμενα, και έστω $A \in V - \Sigma$ και $w \in \Sigma^*$. Τότε οι παρακάτω προτάσεις είναι ισοδύναμες:

- $A \Rightarrow^* w$
- *Υπάρχει συντακτικό δέντρο με ρίζα A που παράγει την w*
- *Υπάρχει αριστερότερη παραγωγή $A \overset{L}{\Rightarrow}{}^* w$*
- *Υπάρχει δεξιότερη παραγωγή $A \overset{R}{\Rightarrow}{}^* w$*

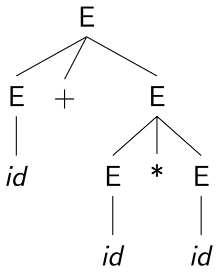
Διφορούμενες Γραμματικές

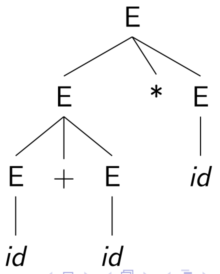
Ορισμός 5 (Διφορούμενη Γραμματική Χωρίς Συμφραζόμενα)

Μία γραμματική χωρίς συμφραζόμενα είναι **διφορούμενη** ή **ασαφής** αν υπάρχει συμβολοσειρά της γλώσσας της, που μπορεί να παράγεται με δύο τουλάχιστο συντακτικά δέντρα.

Παράδειγμα: $E \to E + E \quad E \to E * E \quad E \to (E) \quad E \to id$

Συμβολοσειρά: $id + id * id$





Διφορούμενες Γραμματικές

- Η ασάφεια είναι ιδιότητα των γραμματικών και όχι των γλωσσών.
- Οι γλώσσες που περιγράφονται με διφορούμενες γραμματικές μπορεί να περιγραφούν από ισοδύναμες μη διφορούμενες γραμματικές.

Παράδειγμα:
Έστω η γραμματική $S \rightarrow SS \mid (S) \mid \epsilon$ και έστω η συμβολοσειρά $(())()$.
Δύο αριστερότερες παραγωγές:

$$S \Rightarrow SS \Rightarrow (S)S \Rightarrow ((S))S \Rightarrow (())S \Rightarrow (())()$$

$$S \Rightarrow SS \Rightarrow SSS \Rightarrow SS \Rightarrow (S)S \Rightarrow ((S))S \Rightarrow (())S \Rightarrow (())()$$

- Το πρόβλημα είναι ότι ο κανόνας $S \rightarrow SS$ επιτρέπει την παραγωγή πολλών S που στη συνέχεια απαλείφονται.
- Μπορούμε να μετασχηματίσουμε τη γραμματική σε μία ισοδύναμη μη διφορούμενη

$$S \rightarrow T \mid \epsilon$$
$$T \rightarrow () \mid (T) \mid TT$$

Τέλος ενότητας

Επεξεργασία: Εμμανουέλα Στάχτιαρη
Θεσσαλονίκη, 24/07/2014