ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ

ΑΝΟΙΚΤΑ ΑΚΑΔΗΜΑΙΚΑ ΜΑΘΗΜΑΤΑ

Θεωρία Υπολογισμού

Ενότητα 18: Λήμμα Άντλησης για ΓΧΣ

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

Χαρακτηριστικά Συντακτικού Δέντρου

- Οι ΓΧΣ έχουν πιο σύνθετη δομή επανάληψης από τις κανονικές γλώσσες, όπου αναπαραστούμε μόνο δομές απλής περιοδικότητας για την επαναληπτική εμφάνιση συμβολοσειρών.
- Η σύνθετη αυτή δομή αποτυπώνεται στα συντακτικά δέντρα.

Ορισμός 1 (Εύρος ΓΧΣ)

Εύρος $\phi(G)$ μιας ΓΧΣ G είναι ο μεγαλύτερος αριθμός συμβόλων που εμφανίζονται στο δεξί μέρος κάποιου κανόνα της.

Ορισμός 2 (Μονοπάτι Συντακτικού Δέντρου)

Μονοπάτι ΣΔ είναι μία ακολουθία κόμβων που συνδέονται με τον προηγούμενό τους με ένα ευθύγραμμο τμήμα, έτσι ώστε ο πρώτος κόμβος να αντιστοιχεί στη *ρίζα* και ο τελευταίος κόμβος σε ένα *φύλλο*. Μήκος μονοπατιού είναι ο αριθμός των ευθύγραμμων τμημάτων.

- **Ύψος** συντακτικού δέντρου είναι το μήκος του μεγαλύτερου μονοπατιού.

Μήκος συμβολοσειράς

Λήμμα 1

Οι παραγόμενες συμβολοσειρές μιας ΓΧΣ G με συντακτικό δέντρο ύψους h έχουν μήκος το πολύ $\phi(G)^h$.

Απόδειξη

Με επαγωγή ως προς το h.
Βασικό βήμα: για $h = 1$ το δέντρο είναι κανόνας της G και άρα η συμβολοσειρά του έχει το πολύ $\phi(G) = \phi(G)^h$ σύμβολα.
Επαγωγική υπόθεση: έστω ότι το λήμμα ισχύει για δέντρα ύψους $h \geq 1$.
Επαγωγικό βήμα:
Από τον ορισμό του ΣΔ, κάθε δέντρο ύψους $h + 1$ αποτελείται από μία ρίζα, που συνδέεται το πολύ με $\phi(G)$ υποδέντρα ύψους h.
Από την επαγωγική υπόθεση, όλα τα υποδέντρα έχουν παραγόμενες συμβολοσειρές μήκους το πολύ $\phi(G)^h$.
Άρα για όλο το ΣΔ η παραγόμενη συμβολοσειρά έχει μήκος το πολύ $\phi(G) \cdot \phi(G)^h = \phi(G)^{h+1}$. Άρα το λήμμα ισχύει για κάθε h.

Λήμμα Άντλησης

Θεώρημα 2 (Λήμμα Άντλησης)

Έστω $G = (V, \Sigma, R, S)$ μια ΓΧΣ. Για κάθε $w \in L(G)$ με μήκος μεγαλύτερο από $\phi(G)^{|V-\Sigma|}$ που μπορεί να γραφτεί επίσης ως $w = uvxyz$ έτσι ώστε μία τουλάχιστο εκ των v και y να είναι μη κενή, όλες οι συμβολοσειρές $uv^nxy^nz, n \geq 0$ θα ανήκουν επίσης στην $L(G)$.

Απόδειξη

Έστω μία τέτοια w και έστω T το συντακτικό της δέντρο με ρίζα S που έχει το μικρότερο δυνατό αριθμό φύλλων.

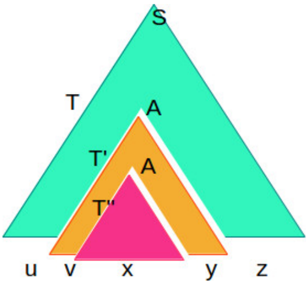
Αφού η w έχει μήκος μεγαλύτερο από $\phi(G)^{|V-\Sigma|}$ από το Λήμμα 1 προκύπτει ότι το T έχει ένα μονοπάτι μήκους τουλάχιστον $|V - \Sigma| + 1$ με $|V - \Sigma| + 2$ κόμβους, εκ των οποίων μόνο ένας, δηλ. το φύλλο, μπορεί να αντιστοιχεί σε τερματικό σύμβολο.

Κατά συνέπεια, θα υπάρχουν στο μονοπάτι δύο κόμβοι που θα αντιστοιχούν στο ίδιο μη τερματικό σύμβολο, αφού αυτό περιλαμβάνει περισσότερα σύμβολα εκτός του φύλλου από τα μη τερματικά της G.

Λήμμα Άντλησης

Απόδειξη (συνέχεια)

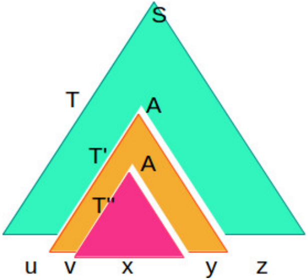
Έστω A το μη τερματικό σύμβολο που εμφανίζεται δύο φορές. Στο σχήμα φαίνονται οι παραγόμενες συμβολοσειρές του κάθε υποδέντρου.



- Η x παράγεται από το T'' με ρίζα το χαμηλότερο A.
- Η v είναι το τμήμα της συμβολοσειράς του δέντρου T' με ρίζα το υψηλότερο A και μέχρι το σημείο που ξεκινά η συμβ/ρά του T''.
- Η z είναι το υπόλοιπο της συμβολοσειράς του T.

Λήμμα Άντλησης

Απόδειξη (συνέχεια)



Το τμήμα του T' χωρίς το T'' μπορεί να εμφανίζεται μία ή περισσότερες φορές σχηματίζοντας συντακτικά δέντρα, που αντιστοιχούν σε συμβολοσειρές της μορφής $uv^nxy^nz, n \geq 0$.
Αν υπήρχε η περίπτωση $vy = \epsilon$, τότε θα υπήρχε δέντρο με ρίζα S και συμβολοσειρά w με λιγότερα φύλλα από το T (αφαιρείται από το T το T' χωρίς το T''), που οδηγεί σε άτοπο.

Παράδειγμα

Θα αποδειχθεί ότι η γλώσσα $L = \{\alpha^n b^n c^n : n \geq 0\}$ δεν είναι γλώσσα χωρίς συμφραζόμενα.

Λύση

Υποθέτουμε ότι για τη γλώσσα L υπάρχει ΓΧΣ $G = (V, \Sigma, R, S)$ τέτοια ώστε $L = L(G)$. Έστω ένα $n > \frac{\phi(G)^{|V-\Sigma|}}{3}$ και μία σύμβολοσειρά $w = \alpha^n b^n c^n$ που ανήκει στην $L(G)$ και γράφεται ως $w = uvxyz$ τέτοια ώστε η v ή η y να είναι μη κενή και η $uv^k xy^k z$ να ανήκει στην $L(G)$ για κάθε $k = 0, 1, 2, \ldots$. Διακρίνουμε δύο ενδεχόμενα:

- *Η vy περιέχει εμφανίσεις και των τριών συμβόλων α, b, c.* Τουλάχιστο μία από τις v, y θα έχει εμφανίσεις τουλάχιστο των δύο εκ των τριών συμβόλων. Τότε όμως η uv^2xy^2z περιέχει δύο εμφανίσεις σε λάθος σειρά, π.χ. ένα b πριν από ένα α και άρα δεν ανήκει στην L όπως προβλέπει το Λήμμα της Άντλησης.
- *Η vy περιέχει εμφανίσεις μόνο κάποιων από τα τρία σύμβολα.* Στην περίπτωση αυτή η uv^2xy^2z θα περιέχει διαφορετικό αριθμό από α, b και c και δεν ανήκει στην L όπως προβλέπει το Λήμμα.

Μη κλειστότητα

Θεώρημα 3

Οι γλώσσες χωρίς συμφραζόμενα δεν είναι κλειστές ως προς την τομή και τη συμπλήρωση.

Απόδειξη *Τομή*: αποδεικνύεται με ένα αντιπαράδειγμα
Για τις γλώσσες $\{a^n b^n c^m : m, n \geq 0\}$ και $\{a^m b^n c^n : m, n \geq 0\}$ υπάρχουν ΓΧΣ που τις παράγουν.
Παρόλα αυτά, για την τομή τους που είναι η $\{a^n b^n c^n : n \geq 0\}$ ξέρουμε από την εφαρμογή του Λήμματος της Άντλησης ότι δεν είναι ΓΧΣ.
Συμπλήρωση: επειδή ισχύει η

$$L_1 \cap L_2 = \overline{\overline{L_1} \cup \overline{L_2}}$$

αν οι ΓΧΣ ήτανε κλειστές ως προς τη συμπλήρωση, τότε θα ήτανε κλειστές και ως προς την τομή (είναι κλειστές ως προς την ένωση). Επειδή όμως δεν είναι κλειστές ως προς την τομή, δε μπορεί να είναι κλειστές ως προς τη συμπλήρωση.

Τέλος ενότητας

Επεξεργασία: Εμμανουέλα Στάχτιαρη
Θεσσαλονίκη, 24/07/2014