Ιδιότητες κλειστότητας των γλωσσών χωρίς συμφραζόμενα

Θεώρημα 1 *Οι γλώσσες χωρίς συμφραζόμενα είναι κλειστές ως προς τις πράξεις*

1. *Ένωση.*
2. *Παράθεση.*
3. *Kleene star.*

Θα αποδείξουμε αργότερα πως η κλάση των γλωσσών χωρίς συμφραζόμενα δεν είναι κλειστή ως προς την τομή και το συμπλήρωμα. Όμως...

Θεώρημα 2 *Η τομή μιας γλώσσας χωρίς συμφραζόμενα και μιας κανονικής γλώσσας είναι γλώσσα χωρίς συμφραζόμενα.*

Απόδειξη: Έστω αυτόματο στοίβας M_1 τέτοιο ώστε $L = L(M_1)$ με $M_1 = (K_1, \Sigma_1, \Gamma_1, \Delta_1, s_1, F_1)$ και ντετερμινιστικό ΠΑ M_2 τέτοιο ώστε $R = L(M_2)$ με $M_2 = (K_2, \Sigma_2, \delta_2, s_2, F_2)$. Θα κατασκευάσουμε αυτόματο στοίβας M που να προσομοιώνει παράλληλα τα M_1 και M_2 και δέχεται την είσοδο αν και μόνο αν θα τη δέχονταν και τα δύο. Ορίζουμε $M = (K, \Sigma, \Gamma, \Delta, s, F)$ όπου

- $K = K_1 \times K_2$.
- $\Sigma = \Sigma_1 \cup \Sigma_2$.
- $\Gamma = \Gamma_1$.
- $s = (s_1, s_2)$.
- $F = F_1 \times F_2$.

και

$$\left(\, ((q_1, q_2), u, \beta)\, , ((p_1, p_2), \gamma)\, \right) \in \Delta \text{ αν και μόνο αν}$$

$$((q_1, u, \beta), (p_1, \gamma)) \in \Delta_1 \text{ ΚΑΙ } \quad (q_2, u) \vdash^*_{M_2} (p_2, \epsilon)$$

Συντακτικά Δέντρα

Ένα συντακτικό δέντρο (**parse tree**) απεικονίζει τον τρόπο παραγωγής μιας συμβολοσειράς από μία δοσμένη γραμματική $G = (V, \Sigma, R, S)$.

Ιδιότητες ενός συντακτικού δέντρου

1. Η ρίζα επιγράφεται με το αρχικό σύμβολο S.
2. Κάθε κόμβος επιγράφεται με ένα σύμβολο του V.
3. Κάθε φύλλο επιγράφεται με στοιχείο του Σ ή ϵ.
4. Με παράθεση των επιγραφών των φύλλων από αριστερά προς τα δεξιά παίρνουμε την παραγόμενη συμβολοσειρά του συντακτικού δέντρου.

Θεώρημα Άντλησης για γλώσσες χωρίς συμφραζόμενα

Θεώρημα Άντλησης: Έστω G μια γραμματική χωρίς συμφραζόμενα. Υπάρχει αριθμός K που εξαρτάται από την G έτσι ώστε κάθε $w \in L(G)$ με μήκος μεγαλύτερο από K μπορεί να ξαναγραφτεί ως $w = uvxyz$ με τέτοιο τρόπο ώστε είτε η v είτε η y είναι μη κενές και η uv^nxy^nz ανήκει στην $L(G)$ για κάθε $n \geq 0$.

Παράδειγμα: δείξτε πως η γλώσσα $L = \{a^m b^m c^m : m \geq 0\}$ δεν είναι χωρίς συμφραζόμενα.

Ας υποθέσουμε πως υπάρχει γραμματική G τέτοια ώστε $L = L(G)$. Έστω $m > K/3$ και $w = a^m b^m c^m$. Τότε $w \in L$ και $w = uvxyz$ ώστε είτε η v είτε η y είναι μη κενές και η uv^nxy^nz ανήκει στην $L(G)$ για κάθε $n \geq 0$. Διακρίνουμε δύο περιπτώσεις.

Περίπτωση 1: η vy περιέχει εμφανίσεις και των τριών συμβόλων a, b, c.
Περίπτωση 2: η vy περιέχει εμφανίσεις το πολύ δύο από τα σύμβολα a, b, c.

Και οι δύο περιπτώσεις οδηγούν σε άτοπο.

Συνέπειες για την κλειστότητα

Θεωρείστε τις γλώσσες $L_1 = \{a^m b^m c^n : m \geq 0, n \geq 0\}$, $L_2 = \{a^n b^m c^m : m \geq 0, n \geq 0\}$. Και οι δύο είναι χωρίς συμφραζόμενα (γιατί;).

$$\{a^m b^m c^m : m \geq 0\} = L_1 \cap L_2$$

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

Τι συμπεραίνετε από τις δύο ισότητες;

Οι γλώσσες χωρίς συμφραζόμενα δεν είναι κλειστές ως προς

1. την τομή.
2. το συμπλήρωμα.

Απόδειξη του Θεωρήματος Άντλησης

Το **εύρος** $\phi(G)$ της $G = (V, \Sigma, R, S)$ ορίζεται ως ο μεγαλύτερος αριθμός συμβόλων στο δεξί μέλος ενός κανόνα της G. Το **ύψος** ενός συντακτικού δέντρου είναι το μήκος του μεγαλύτερου μονοπατιού του.

Λήμμα 1 *Η παραγόμενη συμβολοσειρά κάθε συντακτικού δέντρου της G με ύψος h έχει μήκος το πολύ $\phi(G)^h$.*

Θεώρημα Άντλησης: Έστω G μια γραμματική χωρίς συμφραζόμενα. Κάθε $w \in L(G)$ με μήκος μεγαλύτερο από $\phi(G)^{|V-\Sigma|}$ μπορεί να ξαναγραφτεί ως $w = uvxyz$ με τέτοιο τρόπο ώστε είτε η v είτε η y είναι μη κενές και η uv^nxy^nz ανήκει στην $L(G)$ για κάθε $n \geq 0$.

Απόδειξη: Έστω w μια τέτοια συμβολοσειρά και T το συντακτικό δέντρο που την παράγει και έχει τον ελάχιστο αριθμό κόμβων. Από το λήμμα, το T έχει ένα μονοπάτι μήκους τουλάχιστον $|V - \Sigma| + 1$, δηλ. με τουλάχιστον $|V - \Sigma| + 2$ κόμβους. Τουλάχιστον δύο από αυτούς επιγράφονται με το ίδιο μη τερματικό σύμβολο του $V - \Sigma$.

Το θεώρημα προκύπτει αν εξετάσει κανείς προσεκτικά τη δομή του μονοπατιού αυτού.