Ιδιότητες των γλωσσών των πεπερασμένων αυτομάτων

Θεώρημα: Η κλάση των γλωσσών των πεπερασμένων αυτομάτων είναι κλειστή ως προς:

1. Ένωση
2. Παράθεση
3. **Kleene star**
4. Συμπλήρωμα
5. Τομή

Κανονικές εκφράσεις και ΠΑ

Θεώρημα: Μία γλώσσα είναι κανονική αν και μόνο αν είναι γλώσσα κάποιου πεπερασμένου αυτομάτου.

Απόδειξη: Θα δείξουμε το ευθύ, οτι δηλαδή αν μας δοθεί μία κανονική έκφραση μπορούμε να κατασκευάσουμε ένα μη ντετερμινιστικό πεπερασμένο αυτόματο που δέχεται την ίδια γλώσσα. Το αντίστροφο είναι αρκετά πιο πολύπλοκο και η απόδειξη του παραλείπεται.

Η κλάση των κανονικών γλωσσών είναι η ελάχιστη κλάση γλωσσών που περιέχει το κενό σύνολο, τα μονοσύνολα $\{a\}$, $a \in \Sigma$, και που είναι κλειστή ως προς την ένωση, την παράθεση και το Kleene star.

Το κενό σύνολο καθώς και τα μονομελή σύνολα γίνονται δεκτά από πεπερασμένα αυτόματα. Επιπλέον, οι γλώσσες των πεπερασμένων αυτομάτων είναι κλειστές ως προς την ένωση, την παράθεση και το Kleene star.

Επομένως, κάθε κανονική γλώσσα γίνεται δεκτή από κάποιο πεπερασμένο αυτόματο.

Θεώρημα Άντλησης (Pumping Lemma)

Θεώρημα 1 (Θεώρημα Άντλησης, Pumping Lemma) *Έστω L μια άπειρη κανονική γλώσσα. Υπάρχει ακέραιος $n \geq 1$ τέτοιος ώστε κάθε συμβολοσειρά $w \in L$ με $|w| \geq n$ να μπορεί να γραφτεί διαφορετικά ως $w = xyz$ έτσι ώστε $y \neq \epsilon$, $|xy| \leq n$ και $xy^iz \in L$ για κάθε $i \geq 0$.*

Απόδειξη: Έστω M ντετερμινιστικό ΠΑ με n καταστάσεις που δέχεται την L. Έστω $w \in L$ συμβολοσειρά μήκους $m \geq n$. Τα πρώτα n βήματα του υπολογισμού του M για τα πρώτα n σύμβολα $w_1 \cdots w_n$ της w είναι:

$$(q_0, w_1 \cdots w_n) \vdash_M (q_1, w_2 \cdots w_n) \vdash_M \cdots \vdash_M (q_n, \epsilon)$$

Τουλάχιστον δύο από τις καταστάσεις $q_0, q_1, \ldots, q_n$ είναι ίδιες (Pigeonhole Principle). Έστω λοιπόν ότι $q_k = q_j$ γιά κάποια $k < j \leq n$. Έστω $x = w_1 \cdots w_k$ το τμήμα του w που φέρνει το M στην κατάσταση q_k, $y = w_{k+1} \cdots w_j$ το τμήμα του w ανάμεσα στην q_k και στην q_j, και $z = w_{j+1} \cdots w_m$ το τμήμα μετά την q_j. Τότε $w = xyz$, $y \neq \varepsilon$ και $|xy| \leq n$ (γιατί;).

Το M δέχεται επίσης όλες τις συμβολοσειρές xy^iz ως εξής: Γιά το x περνάει από τις καταστάσεις $q_0, \ldots, q_k$, για το y^i κάνει i κύκλους $q_k, \ldots, q_j$, και για το z συνεχίζει και καταλήγει σε μια τελική κατάσταση $q_m \in F$.

Μη κανονικότητα γλωσσών: με το Θεώρημα Άντλησης

Να δειχτεί ότι η γλώσσα $L = \{a^m b^m : m \geq 0\}$ δεν είναι κανονική.

Έστω ότι η L είναι κανονική. Τότε, μπορούμε να εφαρμόσουμε το Θεώρημα Άντλησης. Έστω n ο φυσικός αριθμός που καθορίζεται από το Θεώρημα.

Έστω τώρα η συμβολοσειρά $w = a^n b^n \in L$. Σύμφωνα με το Θεώρημα, μπορούμε να γράψουμε τη w στη μορφή: $w = xyz$ με $|xy| \leq n$ και $y \neq \epsilon$. Τότε όμως, $y = a^r$ για κάποιο $r > 0$ (γιατί;).

Το Θεώρημα Άντλησης μας λέει όμως ότι $xy^i z \in L$ για κάθε $i \geq 0$. Αν το εφαρμόσουμε για $i = 0$ παίρνουμε ότι $xz = a^{n-r} b^n \in L$ (άτοπο).

Μη κανονικότητα γλωσσών: με τις ιδιότητες κλειστότητας

Έστω L η γλώσσα όλων των συμβολοσειρών του αλφαβήτου $\Sigma = \{(,)\}$ με ταιριασμένες παρενθέσεις (δηλαδή κάθε δεξιά παρένθεση αντιστοιχεί σε μία και μοναδική προηγούμενη αριστερή παρένθεση). Π.χ. $(()())(()) \in L$ αλλά $()))(() \notin L$. Να δειχτεί ότι η γλώσσα L δεν είναι κανονική.

Δέν θα χρησιμοποιήσουμε το Θεώρημα Άντλησης αλλά θα εκμεταλευτούμε το γεγονός ότι η γλώσσα

$$L_1 = \{(^n)^n : n \geq 0\}$$

δεν είναι κανονική (η L_1 είναι σχεδόν ίδια με τη γλώσσα του προηγουμένου παραδείγματος). Εστω

$$L_2 = \{(^*)^*\}$$

που προφανώς είναι κανονική γλώσσα. Οι τρεις γλώσσες έχουν την εξής σχέση:

$$L_1 = L \cap L_2.$$

Αν η L ήταν κανονική τότε και η $L \cap L_2$ θα ήταν κανονική (η τομή δύο κανονικών γλωσσών είναι κανονική). Αλλά τότε και η $L_1 = L \cap L_2$ θα ήταν κανονική, άτοπο.

Θεώρημα 2 *Υπάρχουν αλγόριθμοι που απαντούν στα παρακάτω ερωτήματα για πεπερασμένα αυτόματα:*

1. *Δίνεται M και w. Ανήκει το w στην $L(M)$;*
2. *Δίνεται M. Ισχύει $L(M) = \emptyset$;*
3. *Δίνεται M. Ισχύει $L(M) = \Sigma^*$;*
4. *Δίνονται M_1, M_2. Ισχύει $L(M_1) \subseteq L(M_2)$;*
5. *Δίνονται M_1, M_2. Ισχύει $L(M_1) = L(M_2)$;*