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: Είναι σχετικά εύκολο αν μας δοθεί μία κανονική έκφραση να κατασκευάσουμε ένα μη ντετερμινιστικό πεπερασμένο αυτόματο μέ την ίδια γλώσσα. Το αντίστροφο είναι πιο πολύπλοκο.

Έστω $M(K, \Sigma, \delta, s, F)$ ένα ντετερμινιστικό πεπερασμένο αυτόματο. Θα δείξουμε ότι υπάρχει μια κανονική έκφραση α τέτοια ώστε $L(M) = L(\alpha)$. Εστω $K = \{q_1, \ldots, q_n\}$ το σύνολο των καταστάσεων του M με $s = q_1$.

Ορίζουμε

$R(i, j, k)$ = το σύνολο των συμβολοσειρών που οδηγούν το M από την κατάσταση q_i στην q_j χωρίς να περάσει από τις καταστάσεις q_l με $l \geq k$.

Επειδή $L(M) = \cup_{q_j \in F} R(1, j, n+1)$, αρκεί να δείξουμε ότι κάθε $R(i, j, k)$ είναι γλώσσα μιας κανονικής έκφρασης.

Αλλά αυτό προκύπτει εύκολα από τον ορισμό της R. Έχουμε

$$\begin{aligned}
R(i, j, k+1) &= R(i, j, k) \cup R(i, k, k) R(k, k, k)^* R(k, j, k) \\
R(i, j, 0) &= \{\sigma : \delta(q_i, \sigma) = q_j\}, \quad i \neq j \\
R(i, i, 0) &= \{\sigma : \delta(q_i, \sigma) = q_i\} \cup \{\epsilon\}
\end{aligned}$$

Θεώρημα Άντλησης (Pumping Lemma)

Θεώρημα 1 (Θεώρημα Άντλησης, Pumping Lemma) *Έστω L μια άπειρη κανονική γλώσσα. Τότε υπάρχουν x, y, z με $y \neq \varepsilon$ τέτοια ώστε $xy^n z \in L$ για κάθε $n \geq 0$.*

Απόδειξη: Έστω M ένα ντετερμινιστικό ΠΑ με k καταστάσεις που δέχεται τη γλώσσα L. Αφού η L είναι άπειρη, υπάρχει κάποιο $w = \sigma_1 \cdots \sigma_l \in L$ με $l > k$. Όταν το M διαβάζει το w περνάει από τις καταστάσεις $s = q_0, q_1, \ldots, q_l$ όπου $q_l \in F$.

Τουλάχιστον δύο από τις καταστάσεις $q_0, q_1, \ldots, q_l$ είναι ίδιες (Pigeonhole Principle ή Αρχή του Περιστερεώνα). Έστω λοιπόν ότι $q_i = q_j$ γιά κάποια $i < j$. Έστω $x = \sigma_1 \cdots \sigma_i$ είναι το τμήμα του w μέχρι την κατάσταση q_i, $y = \sigma_{i+1} \cdots \sigma_j$ το τμήμα του w ανάμεσα στην q_i και στην q_j, και $z = \sigma_{j+1} \cdots \sigma_l$ το τμήμα μετά την q_j. Τότε $w = xyz$ και $y \neq \varepsilon$.

Το M δέχεται επίσης όλες τις συμβολοσειρές $xy^n z$ ως εξής: Γιά το x περνάει από τις καταστάσεις $q_0, \ldots, q_i$, για το y^n κάνει n κύκλους $q_i, \ldots, q_j$, και για το z περνάει από τις καταστάσεις $q_{j+1}, \ldots, q_l$ και καταλήγει στην $q_l \in F$.

Μη κανονικότητα γλωσσών: με το Θεώρημα Άντλησης

Να δειχτεί ότι η γλώσσα $L = \{0^n 1^n : n \geq 0\}$ δεν είναι κανονική.

Επειδή η L είναι άπειρη μπορούμε να εφαρμόσουμε το Θεώρημα Άντλησης. Έστω x, y, z τρείς συμβολοσειρές τέτοιες ώστε $xy^n z \in L$ για κάθε $n \geq 0$. Συγκεκριμένα η συμβολοσειρά xyz ανήκει στην L.

1. Αν το y αποτελείται μόνο από 0 τότε το $xy^0 z = xz$ έχει λιγότερα 0 από 1 (αφού $y \neq \varepsilon$) και δεν ανήκει στην L.
2. Το ίδιο ισχύει αν το y αποτελείται μόνο από 1.
3. Αν το y περιέχει και 0 και 1 τότε στο $xy^2 z$ κάποιο 1 ακολουθείται από 0 και επομένως δεν ανήκει στην L.

Και στις 3 περιπτώσεις $xy^n z \notin L$ για κάποιο n. Συνεπώς η L δεν είναι κανονική.

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.

Αλγόριθμοι για πεπερασμένα αυτόματα

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* $L(M_1) = L(M_2)$;