Πεπερασμενα αυτοματα και Κανονικες Εκφρασεις

Θεωρημα

Η κλαση των γλωσσων που γινονται δεκτες απο πεπερασμενα αυτοματα ειναι κλειστη ως προς:

(α) Ενωση

(β) Παραθεση

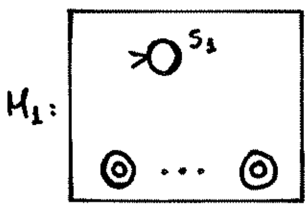
(γ) Kleene star

(δ) Συμπληρωση

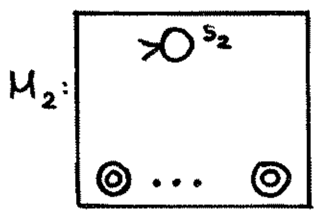
(ε) Τομη

Αποδειξη

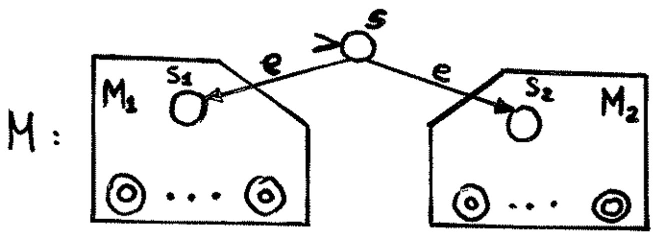
(α) Αν



και

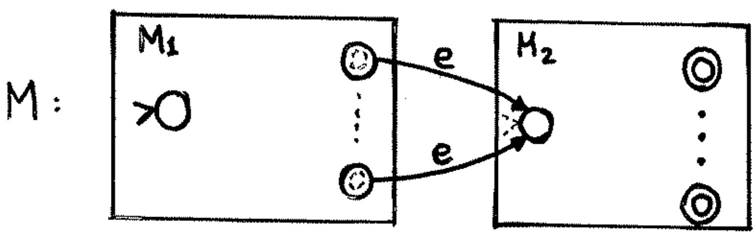


τοτε



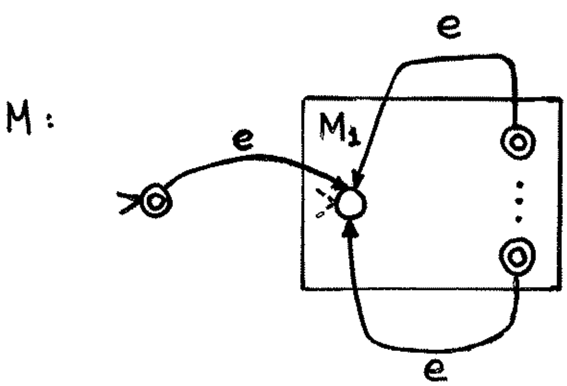
οπου $L(M) = L(M_1) \cup L(M_2)$

(β)



οπου $L(M) = L(M_1)L(M_2)$.

(γ)



οπου $L(M) = (L(M_1))^*$.

(δ) Συμπληρωση

Αν $M_1 = (K, \Sigma, \delta, s, F)$ ειναι <u>ντετερμινιστικο</u> πεπερασμενο αυτοματο και $M = (K, \Sigma, \delta, s, K-F)$ τοτε $L(M) = \Sigma^* - L(M_1)$.

(ε) Τομη

Ισχυει καθως $A \cap B = \overline{\bar{A} \cup \bar{B}}$.

Απο τα παραπανω προκυπτει οτι:

Καθε κανονικη γλωσσα γινεται αποδεκτη απο ενα πεπερασμενο αυτοματο.

Σημειωση

Παρατηρηστε οτι υπαρχουν πεπερασμενα αυτοματα που αποδεχονται το $\emptyset$ καθως και το $\{a\}$ $\forall a \in \Sigma$.

Θα δειξουμε στη συνεχεια και το αντιστροφο οποτε:

<u>Θεωρημα</u>

Μια γλωσσα ειναι κανονικη αν και μονον αν γινεται δεκτη απο ενα πεπερασμενο αυτοματο.

Λημμα

Εαν μια γλωσσα L γινεται δεκτη απο ενα πεπερασμενο αυτοματο τοτε η L μπορει να περιγραφει απο μια κανονικη εκφραση.

Αποδειξη

Εστω $M = (\{q_1, q_2, \ldots, q_n\}, \Sigma, \Delta, q_1, F)$ ενα Π.Α.

$R(i,j,k)$ = συνολο ολων των συμβολοσειρων του Σ^* που οδηγουν το M απο την q_i στην q_j χωρις να περασει απο καμια ενδιαμεση κατασταση με δεικτη $k+1$ ή μεγαλυτερο.

Τοτε: $R(i,j,n)$ = συνολο συμβολοσειρων που οδηγουν το M απο q_i στην q_j.

Για τον υπολογισμο εχουμε:

$$\text{(a)}\quad \underbrace{\text{συμβολοσειρες για τις οποιες η } q_k \text{ δεν ειναι ενδιαμεση κατασταση}}_{R(i,j,k-1)}$$

$$\text{(β)}\quad \underbrace{\text{συμβολοσειρες για τις οποιες η } q_k \text{ ειναι ενδιαμεση κατασταση}}_{R(i,k,k-1)\left(R(k,k,k-1)\right)^* R(k,j,k-1)}$$

$$\Longrightarrow R(i,j,k) = R(i,j,k-1) \cup R(i,k,k-1)\left(R(k,k,k-1)\right)^* R(k,j,k-1).$$

Επισης

$$R(i,j,0) = \begin{cases} \{a \in \Sigma \cup \{e\} : (q_i, a, q_j) \in \Delta\} & \text{αν } i \neq j \\ \{e\} \cup \{a \in \Sigma \cup \{e\} : (q_i, a, q_i) \in \Delta\} & \text{αν } i = j \end{cases}$$

και

$$L(M) = \bigcup_{q_j \in F} R(1, j, n).$$

Χρησιμες ισοτητες για κανονικες εκφρασεις

$(r \cup \phi^*)^* = r^*$ $\qquad$ $(r \cup \phi^*)^* (r \cup \phi^*) = r^*$

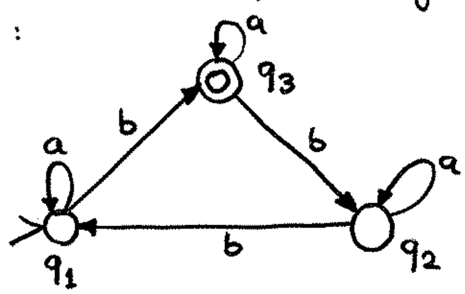
$(r^*)^* = r^*$ $\qquad$ $(r^* s^*)^* = (r \cup s)^*$

$(r^* \cup s^*)^* = (r \cup s)^*$

$r\phi = \phi$ $\qquad$ $r \cup \phi = r$

<u>Παραδειγμα</u>

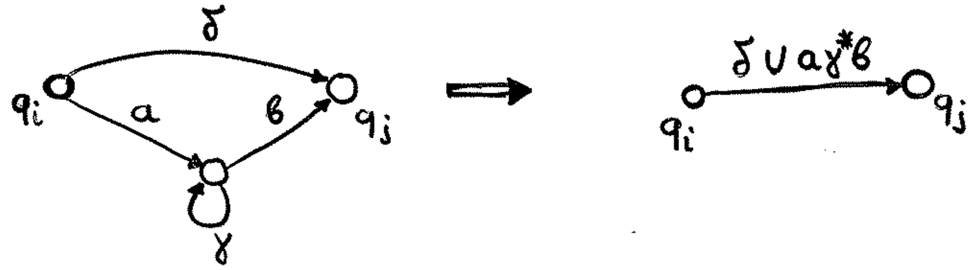
Να βρεθει κανονικη εκφραση για τη γλωσσα που δεχεται το αυτοματο:



Μπορουμε επισης να υπολογισουμε μια κανονικη εκφραση για τη γλωσσα που αποδεχεται ενα αυτοματο:

(α) φροντιζοντας ωστε <u>να μην υπαρχει μεταβαση προς την αρχικη κατασταση</u>, να υπαρχει <u>μία τελικη κατασταση</u> και <u>να μην υπαρχει μεταβαση απο την τελικη κατασταση</u>,

(β) λαμβανοντας υποψιν οτι



Κανονικοτητα ή μη των γλωσσων.

Για να δειξουμε οτι μια γλωσσα ειναι κανονικη χρησιμοποιουμε:

- το οτι οι κανονικες γλωσσες περιγραφονται απο κανονικες εκφρασεις,
- το οτι οι κανονικες γλωσσες γινονται αποδεκτες απο πεπερασμενα αυτοματα, και
- τις ιδιοτητες κλειστοτητας των κανονικων γλωσσων.

Παραδειγμα

Το συνολο των δεκαδικων παραστασεων των μη αρνητικων ακεραιων (χωρις αχρηστα 0 στην αρχη) που διαιρουνται ακριβως με 2 ή 3 ειναι μια κανονικη γλωσσα.

L_1 : μη αρνητικοι ακεραιοι

$$L_1 = \mathcal{L}\big(\, 0 \cup (1 \cup 2 \cup \ldots \cup 9)(0 \cup 1 \cup \ldots \cup 9)^* \,\big)$$

L_2 : δεκαδ. παραστασεις μη αρνητικων που διαιρουνται με το 2

$$L_2 = L_1 \cap \mathcal{L}\big((0 \cup 1 \cup 2 \cup \ldots \cup 9)^* (0 \cup 2 \cup 4 \cup 6 \cup 8) \big)$$

L_3 : δεκαδ. παραστασεις μη αρνητικων που διαιρουνται με το 3

$L_3 = L_1 \cap$ (δεκαδικες παραστασεις με αθροισμα ψηφιων πολλαπλασιο του 3)

αποδεκτη απο Πεπ. Αυτοματο

Τοτε,

η ζητουμενη γλωσσα ειναι η $L_2 \cup L_3$.

Για να δειξουμε οτι μια γλωσσα δεν ειναι κανονικη χρησιμοποιουμε το Λημμα Αντλησης για Κανονικες Γλωσσες.

Λημμα Αντλησης για Κανονικες Γλωσσες (Pumping lemma)

Για καθε (απειρη) κανονικη γλωσσα L,
υπαρχει σταθερα $n \geq 1$ τετοια ωστε
για καθε συμβολοσειρα $w \in L$ με μηκος $|w| \geq n$
υπαρχουν συμβολοσειρες x, y, z με $w = xyz$, $|xy| \leq n$ και $y \neq e$
τετοιες ωστε
για καθε $i \geq 0$ η συμβολοσειρα $xy^iz \in L$.

Αποδειξη

Θεωρουμε πεπερασμενο αυτοματο $M = (K, \Sigma, \Delta, q_0, F)$ που αποδεχεται την L.

Εαν $w = a_1 a_2 \cdots a_k$ οπου $\forall i,\ a_i \in \Sigma$, εχουμε

$$
\begin{array}{lccccccccccc}
 & a_1 & a_2 & \cdots & a_i & a_{i+1} & \cdots & a_j & a_{j+1} & \cdots a_n \cdots a_k & \\
 & \uparrow & & & \uparrow & & & \uparrow & & & \uparrow \\
M: & q_0 & & & q & & & q & & & q_f \in F
\end{array}
$$

Τοτε ομως $\forall t \geq 0$

$$a_1 a_2 \ldots a_i (a_{i+1} \ldots a_j)^t a_{j+1} \ldots a_n \ldots a_k \in L. \quad \blacksquare$$

Για να δειξουμε οτι μια γλωσσα L (που ειναι απειρη) <u>ΔΕΝ</u> ειναι κανονικη χρησιμοποιωντας το Θεωρημα Αντλησης για κανονικες γλωσσες:

1. a) Υποθετουμε οτι η L ειναι κανονικη.
   β) Υποθετουμε οτι η σταθερα του Θεωρηματος αντλησης για την L ειναι n.
   γ) Επιλεγουμε μια (καταλληλη) συμβολοσειρα $w \in L$ μηκους τουλαχιστον n.

2. <u>Για καθε</u> δυνατο χωρισμο της w σε x, y, z οπου

   $w = xyz$, $\quad |xy| \leq n$, $\quad$ και $\quad |y| \geq 1$

   βρισκουμε ενα $i \geq 0$ τετοιο ωστε $xy^iz \notin L$

$\Longrightarrow$ ΑΤΟΠΟ, δηλαδη, η L <u>δεν</u> ειναι κανονικη.

<u>Παραδειγματα</u>

- Η γλωσσα $\{a^i b^i \mid i \geq 0\}$ δεν ειναι κανονικη.
- Η γλωσσα $\{a^p \mid p$ πρωτος αριθμος$\}$ δεν ειναι κανονικη.
- Η γλωσσα $\{w \in \{a,b\}^* :$ η w εχει ισο πληθος απο a και b$\}$ δεν ειναι κανονικη.

Εκτος απο το Θεωρημα αντλησης για κανονικες γλωσσες, επισης το Θεωρημα **Myhill-Nerode** περιγραφει μια αναγκαια και ικανη συνθηκη για να ειναι μια γλωσσα κανονικη και ετσι μας επιτρεπει να δειξουμε οτι μια γλωσσα ειναι κανονικη ή οτι δεν ειναι κανονικη.

Αλγοριθμοι και Πεπερασμενα αυτοματα

Θεωρημα

α) Υπαρχει εκθετικος αλγοριθμος ο οποιος δοθεντος ενος μη ντετερμινιστικου Π. Α. κατασκευαζει ισοδυναμο ντετερμινιστικο Π. Α.

β) Υπαρχει πολυωνυμικος αλγοριθμος ο οποιος για δοθεισα κανονικη εκφραση r κατασκευαζει ενα Πεπ. Αυτοματο που αποδεχεται την $\mathcal{L}(r)$.

γ) Υπαρχει εκθετικος αλγοριθμος ο οποιος δοθεντος ενος πεπ. αυτοματου M κατασκευαζει κανονικη εκφραση r τετοια ωστε $L(M) = \mathcal{L}(r)$.

δ) Υπαρχει πολυωνυμικος αλγοριθμος ο οποιος δοθεντος ενος ντετ. πεπ. αυτοματου κατασκευαζει ισοδυναμο ντετ. πεπ. αυτοματο με το ελαχιστο πληθος καταστασεων.
[Moore, Hopcroft]

ε) Υπαρχει πολυωνυμικος αλγοριθμος ο οποιος αποφασιζει εαν δυο ντετ. πεπ. αυτοματα ειναι ισοδυναμα.

στ) Υπαρχει εκθετικος αλγοριθμος ο οποιος αποφασιζει εαν δυο μη ντετ. πεπ. αυτοματα ειναι ισοδυναμα.

Επισης, υπαρχουν αλγοριθμοι για τα εξης προβληματα:

α) για πεπ. αυτοματο M και συμβολοσειρα w, ισχυει $w \in L(M)$;

β) για πεπ. αυτοματο M, ισχυει οτι $L(M) \neq \emptyset$;