

HY-280 Θεωρία Υπολογισμού

Φροντιστήριο 3

10/10/2024

Άσκηση 1. Για τις ακόλουθες κανονικές εκφράσεις,

$$r_1 = 0^* \cup 1^*, \quad r_2 = 01^* \cup 10^* \cup 1^*0 \cup (0^*1)^*$$

1. βρείτε μια συμβολοσειρά που αντιστοιχεί στη r_2 αλλιά όχι στη r_1 και
2. βρείτε μια συμβολοσειρά που αντιστοιχεί και στις δύο r_1 και r_2 .

Λύση.

1. Οποιαδήποτε συμβολοσειρά που αποτελείται μόνο από 0 ή μόνο από 1 και η κενή συμβολοσειρά είναι στη r_1 . Έτσι, πρέπει να βρούμε συμβολοσειρές της r_2 που περιέχουν τουλάχιστον ένα 0 και τουλάχιστον ένα 1. Για παράδειγμα, οι 01 και 10 είναι τέτοιες συμβολοσειρές.
2. Μια συμβολοσειρά που αντιστοιχεί στη r_1 αποτελείται μόνο από 0 ή μόνο από 1 ή την κενή συμβολοσειρά. Οι μόνες συμβολοσειρές που αντιστοιχούν στη r_2 και αποτελούνται μόνο από 0 ή 1 είναι οι 0, 1 και οι συμβολοσειρές που αποτελούνται μόνο από 1 (από το $(0^*1)^*$).

□

Άσκηση 2. Έστω r_1 και r_2 αυθαίρετες κανονικές εκφράσεις πάνω σε κάποιο αλφάβητο. Βρείτε μια απλή (ή συντομότερη και με τα λιγότερο εμφωλιασμένα * και $\cup$) κανονική έκφραση που είναι ίση με κάθε μία από τις παρακάτω κανονικές εκφράσεις.

$$(r_1 \cup r_2 \cup r_1 r_2 \cup r_2 r_1)^*, \quad (r_1(r_1 \cup r_2)^*)^+.$$

Λύση. Θα χρησιμοποιήσουμε το αλφάβητο $\Sigma = \{0, 1\}$. Μια γενική στρατηγική για να προσεγγίσουμε αυτό το είδος ερώτησης είναι να προσπαθήσουμε να δούμε αν είναι ίσες με απλές κανονικές εκφράσεις που μας είναι γνωστές, όπως 0, 1, 0^+ , $(0 \cup 1)$, $(0 \cup 1)^+$ κ.λ.π.

1. Εφόσον $(r_1 \cup r_2)^*$ αντιπροσωπεύει όλες τις συμβολοσειρές που αποτελούνται από συμβολοσειρές του r_1 και/ή του r_2 , το $r_1 r_2 \cup r_2 r_1$ στην δεδομένη κανονική έκφραση είναι περιττό, δηλαδή, δεν παράγουν καμία συμβολοσειρά που δεν αντιπροσωπεύεται από το $(r_1 \cup r_2)^*$. Έτσι, το $(r_1 \cup r_2 \cup r_1 r_2 \cup r_2 r_1)$ ανάγεται σε $(r_1 \cup r_2)$.
2. Το $(r_1(r_1 \cup r_2)^*)^+$ σημαίνει ότι όλες οι συμβολοσειρές που αντιπροσωπεύονται από αυτό πρέπει να αποτελούνται από μία ή περισσότερες συμβολοσειρές του $(r_1(r_1 \cup r_2)^*)$. Ωστόσο, οι συμβολοσειρές του $(r_1(r_1 \cup r_2)^*)$ ξεκινούν με μια συμβολοσειρά του r_1 ακολουθούμενη από οποιονδήποτε αριθμό συμβολοσειρών που λαμβάνονται αυθαίρετα από το r_1 και/ή το r_2 . Έτσι, οτιδήποτε έρχεται μετά το πρώτο r_1 στο $(r_1(r_1 \cup r_2)^*)^+$ αντιπροσωπεύεται από το $(r_1 \cup r_2)^*$. Επομένως, το $(r_1(r_1 \cup r_2)^*)$ αντιπροσωπεύει επίσης τις συμβολοσειρές του $(r_1(r_1 \cup r_2)^*)^+$, και αντίστροφα το $(r_1(r_1 \cup r_2)^*)^+$ αντιπροσωπεύει τις συμβολοσειρές που αντιπροσωπεύονται από το $(r_1(r_1 \cup r_2)^*)$. Έτσι, το $(r_1(r_1 \cup r_2)^*)^+$ ανάγεται σε $(r_1(r_1 \cup r_2)^*)$.

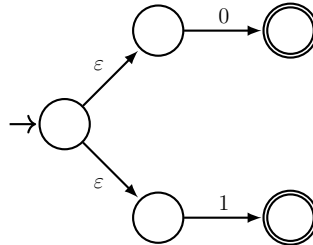
□

Άσκηση 3. Μετατρέψτε την κανονική έκφραση $(0 + 1)^*100$ σε NFA.

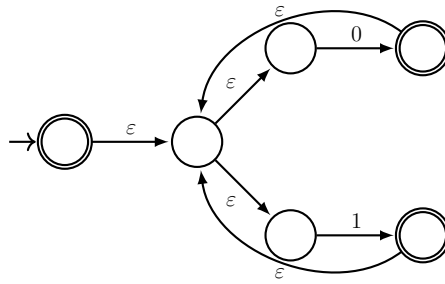
Λύση. Ξεκινάμε από τα σύμβολα του αλφαβήτου,



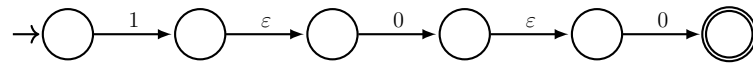
Κατασκευάζουμε εν συνεχεία την κανονική έκφραση $0 + 1$,



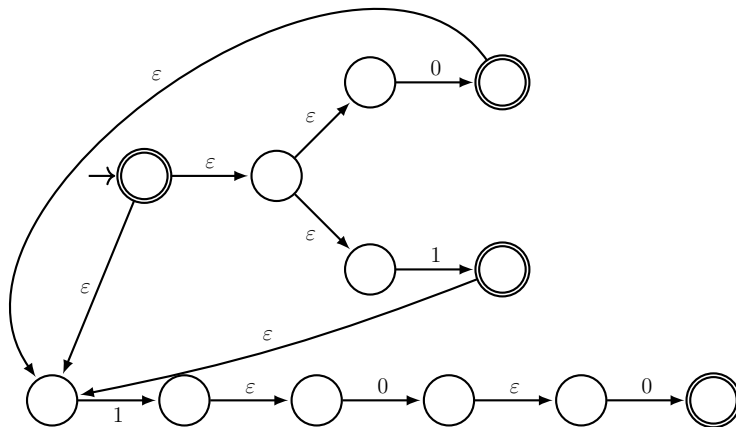
Προχωράμε με την κατασκευή της κανονικής έκφρασης $(0 + 1)^*$,



Η κανονική έκφραση 100 δίνει το παρακάτω αυτόματο,

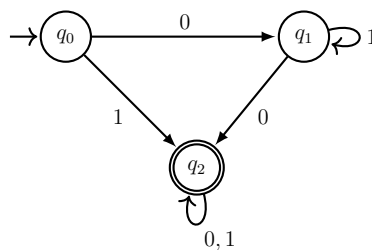


Παραθέτουμε τις δύο τελευταίες κανονικές εκφράσεις, παράγοντας την κανονική έκφραση $(0 + 1)^*100$. Επομένως, το ισοδύναμο πεπερασμένο αυτόματο θα είναι,

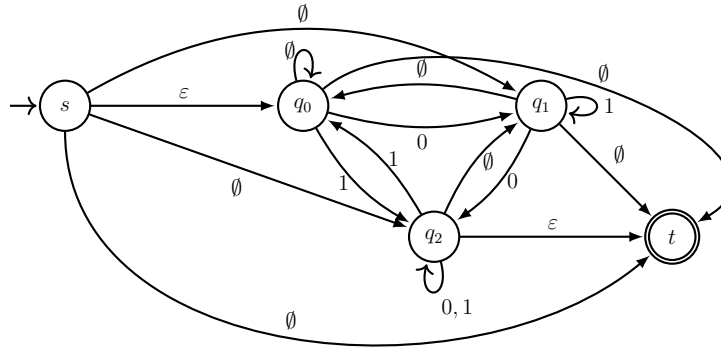


□

Άσκηση 4. Μετατρέψτε το παρακάτω GNFA σε ισοδύναμη κανονική έκφραση.



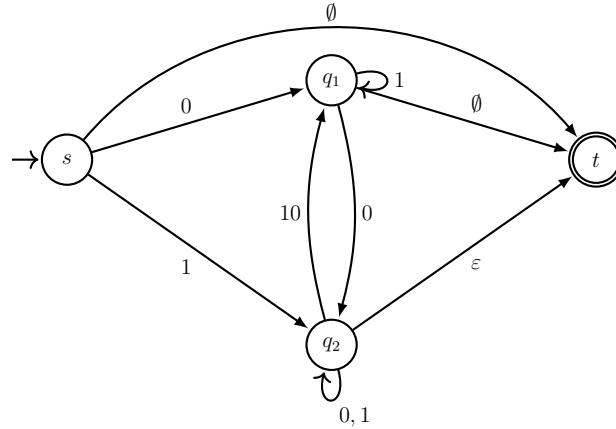
Λύση. Αρχικά κανονικοποιούμε το αυτόματο και προσθέτουμε όλες τις $\emptyset$ -μεταβάσεις,



Διαγράφουμε την κατάσταση q_0 ,

- Στο υπογράφημα sq_0q_1 , η s θα συνδέεται με την q_1 μέσω της κανονικής έκφρασης $\emptyset \cup \varepsilon\emptyset^*0 = \varepsilon 0 = 0$.
- Στο υπογράφημα sq_0q_2 , η s θα συνδέεται με την q_2 μέσω της κανονικής έκφρασης $\emptyset \cup \varepsilon\emptyset^*1 = \varepsilon 1 = 1$.
- Στο υπογράφημα sq_0t , η s θα συνδέεται με την t μέσω της κανονικής έκφρασης $\emptyset \cup \varepsilon\emptyset^*\emptyset = \varepsilon\emptyset = \emptyset$.
- Στο υπογράφημα $q_1q_0q_2$, η q_1 θα συνδέεται με την q_2 μέσω της κανονικής έκφρασης $0 \cup \emptyset\emptyset^*1 = 0$.
- Στο υπογράφημα q_1q_0t , η q_1 θα συνδέεται με την t μέσω της κανονικής έκφρασης $\emptyset \cup \emptyset\emptyset^*\emptyset = \emptyset$.
- Στο υπογράφημα $q_2q_0q_1$, η q_2 θα συνδέεται με την q_1 μέσω της κανονικής έκφρασης $\emptyset \cup 1\emptyset^*0 = 10$.
- Στο υπογράφημα q_2q_0t , η q_2 θα συνδέεται με την t μέσω της κανονικής έκφρασης $\varepsilon \cup 1\emptyset^*\emptyset = \varepsilon$.

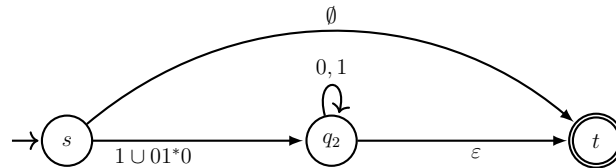
Σύμφωνα με την παραπάνω ανάλυση, επαναδρομολογούμε τις μεταβάσεις και το αυτόματο γίνεται,



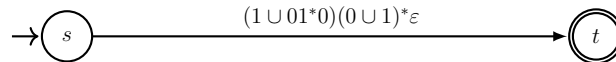
Διαγράφουμε την κατάσταση q_1 ,

- Στο υπογράφημα sq_1q_2 , η s θα συνδέεται με την q_2 μέσω της κανονικής έκφρασης $1 \cup 01^*0$.
- Στο υπογράφημα sq_1t , η s θα συνδέεται με την t μέσω της κανονικής έκφρασης $\emptyset \cup 01^*\emptyset = \emptyset$.
- Στο υπογράφημα q_2q_1t , η q_2 θα συνδέεται με την t μέσω της κανονικής έκφρασης $\varepsilon \cup 101^*\emptyset = \varepsilon$.

Εάν επαναδρομολογούμε τις μεταβάσεις και το αυτόματο γίνεται,



Τέλος, διαγράφουμε και την κατάσταση q_2 . Επομένως καταλήγουμε στο αυτόματο,



Η ισοδύναμη κανονική έκφραση είναι η $R = (1 \cup 01^*0)(0 \cup 1)^*$. $\square$

Άσκηση 5. Έστω η γλώσσα $A = \{w \mid w = 1^{k^2}, k \in \mathbb{N}\}$. Να δείξετε ότι η A δεν είναι κανονική.

Λύση. Η γλώσσα A αποτελείται από όλες τις συμβολοσειρές 1 που το μήκος τους είναι τέλειο τετράγωνο. Έστω ότι η A είναι κανονική γλώσσα, άρα ισχύει το Λήμμα Άντλησης. Για $p \geq 1$ διαλέγουμε την συμβολοσειρά $s = 1^{p^2}$. Παρατηρούμε ότι, $|s| = p^2 \geq p$. Άρα από το Λήμμα Άντλησης υπάρχει διαμέριση του $s = xyz$ τ.ω. $|xy| \leq p$, $y \neq \emptyset$, και $xy^iz \in A$ για κάθε $i \geq 0$.

Εφόσον το αλφάβητο έχει μόνο ένα σύμβολο, θα πρέπει $y = 1^k$ για κάποιο $1 \leq k \leq p$, και $xy^iz = 1^{p^2+(i-1)k}$ για κάθε $i \geq 0$.

Αν πάρουμε $i = 2$, θα έχουμε $xy^2z = xygz = 1^{p^2+k}$. Όμως, δεν μπορεί το p^2+k να είναι τέλειο τετράγωνο. Ο αριθμός p^2+k είναι μεγαλύτερος από το p^2 , αλλά το επόμενο τέλειο τετράγωνο είναι το $(p+1)^2$, για το οποίο ισχύει $(p+1)^2 = p^2 + 2p + 1 > p^2 + k$, εφόσον $k \leq p$.

Επομένως, η συμβολοσειρά xy^2z δεν περιέχεται στην A . Καταλήξαμε σε άτοπο, άρα η A δεν είναι κανονική γλώσσα. $\square$

Σχόλιο. Στην προηγούμενη άσκηση αν παίρναμε $p \geq 0$, τότε για $p = 0$ δεν μπορούμε να αντλήσουμε.

Άσκηση 6. Έστω η γλώσσα $B = \{a^\ell b^m c^n \mid \ell, m, n \geq 0 \text{ και } \ell + m = n\}$. Να δείξετε ότι η B δεν είναι κανονική.

Λύση. Έστω ότι η B είναι κανονική γλώσσα, άρα ισχύει το Λήμμα Άντλησης. Για $p \geq 0$ διαλέγουμε την συμβολοσειρά $s = a^p b^p c^{2p}$. Από το Λήμμα Άντλησης για κάθε διαμέριση του $s = xyz$ θα πρέπει να ισχύει $|xy| \leq p$ και $y \neq \emptyset$. Επομένως, θα πρέπει $y = a^k$ για κάποιο $k > 0$. Ωστόσο, $xygz = a^{p-k} a^k b^p c^{2p} = a^{p+k} b^p c^{2p} \notin B$, καθώς $k > 0$. Άρα, η B δεν είναι κανονική. $\square$

Άσκηση 7. Για κάθε συμβολοσειρά $w = w_1 \cdots w_n$, $w_i \in \Sigma$, ορίζουμε την αντεστραμμένη συμβολοσειρά $w^R = w_n \cdots w_1$, να είναι η w με αντεστραμμένη τη σειρά των γραμμάτων του. Ορίζουμε τη γλώσσα

$$A = \{w \in \{0,1\}^* \mid w = w^R\},$$

να απαρτίζεται από όλες εκείνες τις συμβολοσειρές που διαβάζονται το ίδιο από αριστερά και από τα δεξιά. Να δείξετε ότι η A δεν είναι κανονική.

Λύση. Έστω ότι είναι και έστω p ο φυσικός αριθμός του οποίου η ύπαρξη προκύπτει από το Λήμμα Άντλησης. Ορίζουμε τη συμβολοσειρά $s = 1^p 0 1^p$ που είναι στην A . Από τις προϋποθέσεις του Λήμματος Άντλησης γράφουμε s ως $s = xyz$, με μη κενό y και $|xy| \leq p$, οπότε και οι συμβολοσειρές x, y έχουν μέσα μόνο 1, ώστε $xy^iz \in A$, για $i \geq 0$. Ωστόσο, η συμβολοσειρά $xz \notin A$ αφού αφαιρώντας το y αφαιρέσαμε κάποια 1 από την αριστερή ομάδα 1 αλλά όχι από τη δεξιά ομάδα, οπότε έχουμε καταλήξει σε άτοπο, άρα η A δεν είναι κανονική. $\square$

Άσκηση 8. Με χρήση του Θεωρήματος Myhill-Nerode να δείξετε ότι η γλώσσα $L = \{w \in \{0, 1\}^* \mid ww\}$ δεν είναι κανονική.¹

Λύση. Θεωρήστε το σύνολο των συμβολοσειρών $S = \{0^n \mid n \in \mathbb{N}\}$, θα δείξουμε ότι τα στοιχεία του είναι ανά δύο διαφορετικά ως προς την γλώσσα L . Παίρνουμε δύο οποιαδήποτε στοιχεία του S , έστω 0^k και 0^m , όπου k, m θετικοί ακέραιοι με $k \neq m$. Διαλέγουμε το 10^k1 τη συμβολοσειρά με την οποία επεκτείνουμε τις 0^k και 0^m . Τότε η 0^k10^k1 θα ανήκει στην L , όχι όμως και η 0^m10^k1 . Άρα οι 0^k και 0^m είναι διαφορετικές, μάλιστα για κάθε διαφορετική τιμή του $n \in \mathbb{N}$, η συμβολοσειρά a^n ανήκει σε διαφορετική κλάση ισοδυναμίας. Άρα, υπάρχουν άπειρες κλάσεις ισοδυναμίας, κάτι που σημαίνει ότι η γλώσσα δεν είναι κανονική. $\square$

Άσκηση 9. Με χρήση του Θεωρήματος Myhill-Nerode να δείξετε ότι η γλώσσα $A = \{a^n b^{n^2} \mid n \geq 0\}$ δεν είναι κανονική.

Λύση. Έστω ότι η A είναι κανονική, τότε υπάρχουν μόνο πεπερασμένες κλάσεις ισοδυναμίας. Σύμφωνα με το Θεώρημα Myhill-Nerode, δύο συμβολοσειρές x και y είναι ισοδύναμες αν για κάθε επέκταση z , η ακολουθία xz ανήκει στη A αν και μόνο αν η yz ανήκει στη γλώσσα A . Θεωρούμε τις συμβολοσειρές της μορφής $x_n = a^n$ για διαφορετικές τιμές του n . Θα δείξουμε ότι για κάθε διαφορετικό n , η ακολουθία a^n ανήκει σε διαφορετική κλάση ισοδυναμίας. Για δύο διαφορετικά n και m , με $n \neq m$, και ότι οι ακολουθίες a^n και a^m ανήκουν στην ίδια κλάση ισοδυναμίας. Αυτό σημαίνει ότι για κάθε z , η $a^n z \in A$ αν και μόνο αν η $a^m z \in A$. Ας πάρουμε το $z = b^{n^2}$.

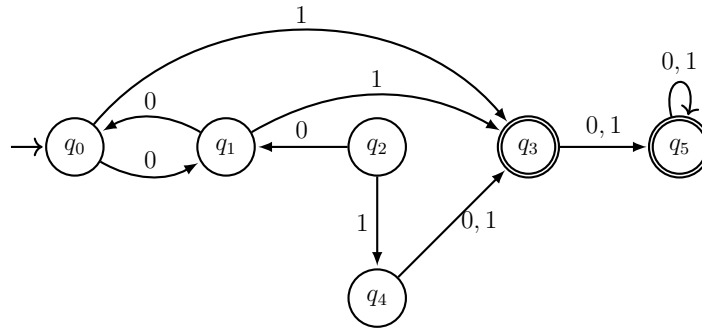
- Για $a^n z = a^n b^{n^2}$, ανήκει στη γλώσσα, αφού η ακολουθία είναι της μορφής $a^n b^{n^2}$.
- Για $a^m z = a^m b^{n^2}$, δεν ανήκει στη γλώσσα, καθώς το πλήθος των b δεν είναι m^2 .

Άρα οι ακολουθίες a^n και a^m δεν μπορούν να είναι ισοδύναμες.

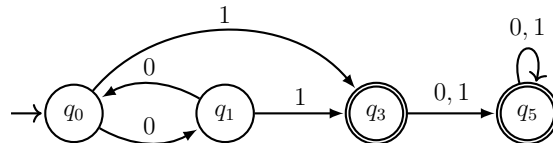
Συνεπώς, για κάθε διαφορετική τιμή του n , η ακολουθία a^n ανήκει σε διαφορετική κλάση ισοδυναμίας. Άρα, υπάρχουν άπειρες κλάσεις ισοδυναμίας, κάτι που σημαίνει ότι η γλώσσα δεν είναι κανονική. $\square$

Άσκηση 10. Ελαχιστοποιήστε το παρακάτω DFA.

¹Στη Διάλεξη 5, αποδείξαμε την μη-κανονικότητα της L χρήση του λήμματος άντλησης για κανονικές γλώσσες.



Λύση. Αρχικά, εξαλείφουμε τις μη-προσβάσιμες καταστάσεις, αυτές είναι οι q_2 , q_4 . Άρα το αυτόματο θα γίνει:



Για το αυτόματο χωρίς μη-προσβάσιμες καταστάσεις, δημιουργούμε πίνακα ζευγών $\{p, q\}$, όπου $p, q \in K$. Όλα τα στοιχεία του πίνακα είναι αρχικώς κενά.

q_0		
	q_1	
		q_3
		q_5

Μαρκάρουμε τα ζεύγη $\{p, q\}$ αν $p \in F$ και $q \notin F$, ή το αντίστροφο.

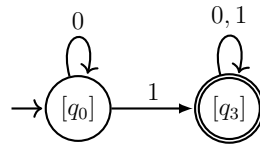
Βήμα 1:

q_0	
☐	q_1
☒	☒ q_3
☒	☒ q_5

Βήμα 2: Μαρκάρουμε τα ζεύγη $\{p, q\}$ αν $\delta(p, a) \in F$ και $\delta(q, a) \notin F$, ή το αντίστροφο, για κάποιο $a \in \Sigma$. Παρατηρούμε ότι για $a = 1$, όλες οι καταστάσεις μεταβαίνουν σε τερματικές, άρα για $a = 1$ δεν μπορούμε να μαρκάρουμε κάποιο ζεύγος. Για $a = 0$, οι q_0, q_1 μεταβαίνουν σε μη-τερματικές, ενώ οι q_3, q_5 μεταβαίνουν σε τερματικές. Όμως, τα ζεύγη $\{q_0, q_1\} \times \{q_3, q_5\}$ τα έχουμε μαρκάρει στο προηγούμενο βήμα.

Η επαναληπτική διαδικασία ολοκληρώνεται καθώς δεν μπορούμε να μαρκάρουμε επιπλέον ζεύγη καταστάσεων.

Από τον τελευταίο πίνακα παίρνουμε ότι $q_0 \approx q_1$ και $q_3 \approx q_5$. Συνεπώς, έχουμε τις κλάσεις $[q_0]$ και $[q_3]$. Συνεπώς, το ελαχιστοποιημένο αυτόματο θα έχει την μορφή,



□