



ΗΥ - 280 Θεωρία Υπολογισμού

Διάλεξη 5

Κωνσταντίνος Βάρσος

Τμήμα Επιστήμης Υπολογιστών – Πανεπιστήμιο Κρήτης

Χειμερινό εξάμηνο 2025 - 26

Γλώσσες και Αυτόματα

Τι είδαμε στην Διάλεξη 4

- Ισοδυναμία DFA και NFA.
- Αλγόριθμος μετατροπής NF σε DFA.
- Ιδιότητες κανονικών γλωσσών: ένωση, συμπλήρωμα και τομή.
- Κανονικές εκφράσεις.

Κανονικές Εκφράσεις

π.χ., Βρείτε κανονική έκφραση για την γλώσσα ...

$$A = \{w \in \{0, 1\}^* \mid w \text{ περιέχει } 10 \text{ και τελειώνει σε } 1\}$$

$$(0 \cup 1)^* 10 (0 \cup 1)^* 1$$

$$B = \{w \in \{\omega, \theta\}^* \mid w \text{ περιέχει ζυγό αριθμό } \omega\}$$

$$(\theta^* \omega \theta^* \omega \theta^*)^* \text{ ή } (\theta^* \omega \theta^* \omega)^* \theta^* \text{ ή } (\theta \cup \omega \theta^* \omega)^*$$

$$C = \{w \in \{a, b\}^* \mid w \text{ τελειώνει με } b \text{ και δεν περιέχει } aa\}$$

$$(b \cup ab) \cup (b \cup ab)^* \text{ ή } (b \cup ab)^+$$

Κανονικές Εκφράσεις

π.χ., $\Sigma = \{0, 1\}$

Κανονική έκφραση	Γλώσσα
01	
$01 \cup 1$	
$(01 + \varepsilon)1$	
$(0 + 10)^*(\varepsilon + 1)$	
0^*10^*	
$0^* \cup 0^*10^*$	
$((0 + 1)(0 + 1)(0 + 1))^*$	
$1((0 \cup 1)1)^*$	
$((0 \cup 1)1)^* \cup (0 \cup 1)(1(0 \cup 1))^*$	

Κανονικές Εκφράσεις

π.χ., $\Sigma = \{0, 1\}$

Κανονική έκφραση	Γλώσσα
01	$\{01\}$
$01 \cup 1$	$\{01, 1\}$
$(01 + \varepsilon)1$	$\{011, 1\}$
$(0 + 10)^*(\varepsilon + 1)$	$\{\varepsilon, 1, 0, 10, 01, 101, 00, 010, 1010, 001, \dots\}$
0^*10^*	$\{\omega \mid \omega \text{ έχει μοναδικό } 1\}$
$0^* \cup 0^*10^*$	$\{\omega \mid \omega \text{ έχει το πολύ ένα } 1\}$
$((0 + 1)(0 + 1)(0 + 1))^*$	$\{\omega \mid \omega \text{ πολλαπλάσιο του } 3\}$
$1((0 \cup 1)1)^*$	$\{\omega \mid \omega \text{ έχει } 1 \text{ σε κάθε περιττή θέση} \wedge \omega \text{ είναι περιττός}\}$
$((0 \cup 1)1)^* \cup (0 \cup 1)(1(0 \cup 1))^*$	$\{\omega \mid \omega \text{ έχει } 1 \text{ σε κάθε άρτια θέση}\}$

Κανονικές Εκφράσεις

Μερικές βασικές αλγεβρικές ιδιότητες των Regex,

- $R_1 \cup (R_2 \cup R_3) = (R_1 \cup R_2) \cup R_3$
- $R_1(R_2 R_3) = (R_1 R_2) R_3$
- $R_1(R_2 \cup R_3) = R_1 R_2 \cup R_1 R_3, (R_1 \cup R_2) R_3 = R_1 R_3 \cup R_2 R_3$
- $R_1 \cup R_2 = R_2 \cup R_1$
- $(R^*)^* = R^*$
- $R\varepsilon = \varepsilon R = R$
- ! $R \cup \emptyset = R, R\emptyset = \emptyset, R\emptyset^* = R$

Κανονικές Εκφράσεις

Μερικές βασικές αλγεβρικές ιδιότητες των Regex,

- $R_1 \cup (R_2 \cup R_3) = (R_1 \cup R_2) \cup R_3$
- $R_1(R_2 R_3) = (R_1 R_2) R_3$
- $R_1(R_2 \cup R_3) = R_1 R_2 \cup R_1 R_3, (R_1 \cup R_2) R_3 = R_1 R_3 \cup R_2 R_3$
- $R_1 \cup R_2 = R_2 \cup R_1$
- $(R^*)^* = R^*$
- $R\varepsilon = \varepsilon R = R$
- ! $R \cup \emptyset = R, R\emptyset = \emptyset, R\emptyset^* = R$

■ Ισχύει $R_1 R_2 = R_2 R_1$;

Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

Θεώρημα. [Kleene'56] Μια γλώσσα A είναι κανονική αν $A = \mathcal{L}(R)$ για κάποια κανονική έκφραση R .

Απόδειξη. [αντίστροφο] (**Κατασκευή**) Δοθείσας R θα κατασκευάσουμε NFA N , τ.ω. $\mathcal{L}(R) = \mathcal{L}(N)$.

$R = \emptyset$



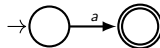
$R_1 \cup R_2$ όπως κλειστότητα NFA

$R = \varepsilon$



$R_1 R_2$ όπως κλειστότητα NFA

$R = a$

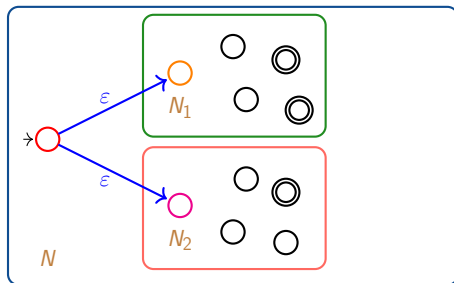


R_1^* όπως κλειστότητα NFA

Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

Απόδειξη. (συνέχεια)

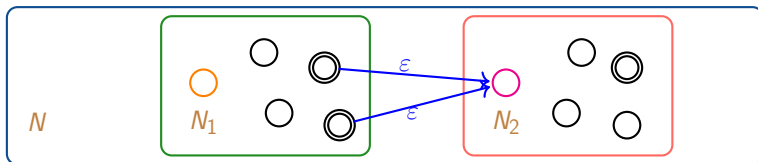
$R = R_1 \cup R_2$, όπου R_1 και R_2 είναι Regex. Έστω N_1 και N_2 τα NFA για τις R_1 και R_2 , αντίστοιχα. Τότε το NFA N για την R γίνεται



Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

Απόδειξη. (συνέχεια)

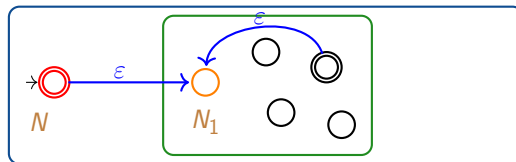
$R = R_1 R_2$, όπου R_1 και R_2 είναι Regex. Έστω N_1 και N_2 τα NFA για τις R_1 και R_2 , αντίστοιχα. Τότε το NFA N για την R γίνεται



Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

Απόδειξη. (συνέχεια)

$R = R_1^*$, όπου R_1 είναι Regex. Έστω N_1 NFA για την R_1 . Τότε το NFA N για την R γίνεται

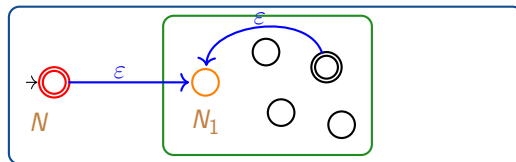


□

Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

Απόδειξη. (συνέχεια)

$R = R_1^*$, όπου R_1 είναι Regex. Έστω N_1 NFA για την R_1 . Τότε το NFA N για την R γίνεται



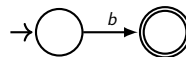
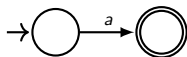
□

- Η προηγούμενη μέθοδος κατασκευάζει NFA από Regex με επαγωγικό τρόπο.
- Η κλάση των γλωσσών που περιγράφεται από Regex είναι υποσύνολο των κανονικών γλωσσών.

Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

π.χ., Μετατρέψτε την κανονική έκφραση $(ab \cup a)^*$ σε NFA.

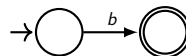
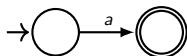
Ξεκινάμε από τα σύμβολα του αλφαβήτου,



Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

π.χ., Μετατρέψτε την κανονική έκφραση $(ab \cup a)^*$ σε NFA.

Ξεκινάμε από τα σύμβολα του αλφαβήτου,



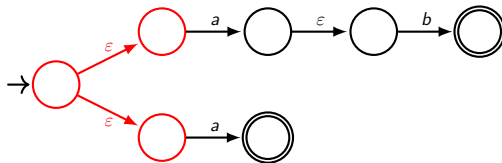
Σταδιακά κατασκευάζουμε τα πιο σύνθετα μέρη της έκφρασης. Για την ab χρησιμοποιούμε τον μηχανισμό της παράθεσης



Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

π.χ., (συνέχεια)

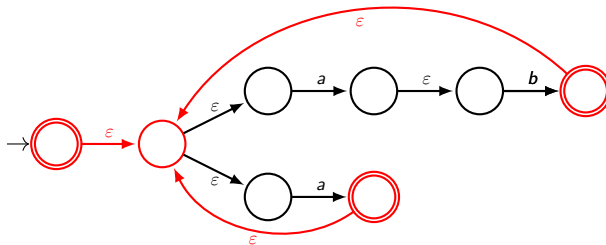
Έπειτα για την $ab \cup a$, χρησιμοποιούμε την ένωση



Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

π.χ., (συνέχεια)

Για την $(ab \cup a)^*$, καταλήγουμε στο



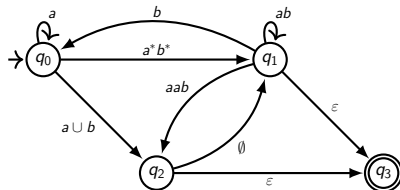
Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

- Μένει να δείξουμε ότι για κάθε κανονική γλώσσα υπάρχει κανονική έκφραση.
- Θα χρησιμοποιήσουμε μια νέα μορφή αυτομάτου, το γενικευμένο μη-ντετερμινιστικό πεπερασμένο αυτόματο (GNFA).

Ορισμός. Ένα γενικευμένο μη-ντετερμινιστικό πεπερασμένο αυτόματο GNFA είναι ένα NFA το οποίο επιτρέπει κανονικές εκφράσεις ως μεταβάσεις.

Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

π.χ., πεπερασμένο αυτόματο με κανονικές εκφράσεις ως ετικέτες,



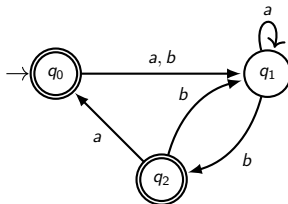
Τα GNFA που μας ενδιαφέρουν έχουν **και** τα παρακάτω χαρακτηριστικά:

- Μία αρχική, μια (διαφορετική) τερματική κατάσταση.
- Πλήρες γράφημα, εκτός από αρχική κ τελική κατάσταση.
- Καμία εισερχόμενη στην αρχική, καμία εξερχόμενη στην τελική.

Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

Μετατροπή NFA σε GNFA.

π.χ., Έστω NFA N ,

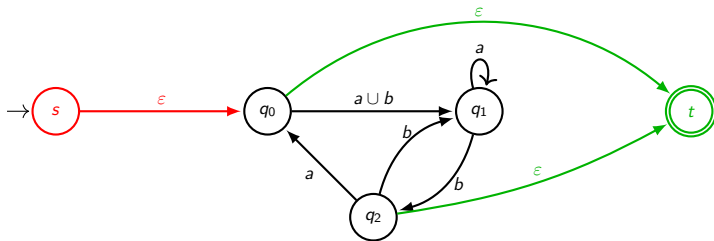


■ Στην μετάβαση με ετικέτα a, b γράφουμε την ισοδύναμη κανονική έκφραση $a \cup b$.

Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

■ Προσθέτουμε μια καινούργια αρχική κατάσταση, s , και μια καινούργια τελική κατάσταση, t .

■ Προσθέτουμε ϵ -μετάβαση από την s στην αρχική κατάσταση του N και ϵ -μετάβαση από τις τελικές καταστάσεις του N στην t .

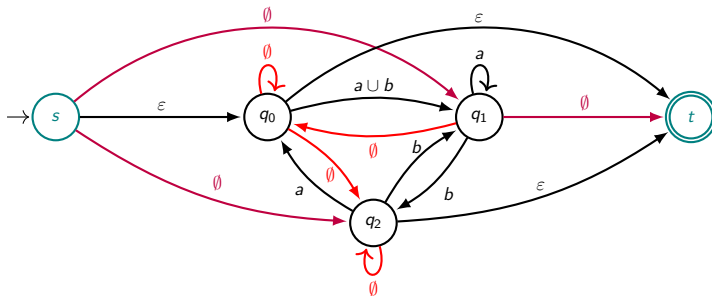


■ Τα δύο αυτά βήματα ονομάζονται και κανονικοποίηση του αυτομάτου.

Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

■ Προσθέτουμε $\emptyset$ -μεταβάσεις από την s προς κάθε κατάσταση $K \cup \{t\}$, και από κάθε κατάσταση $K \cup \{s\}$ προς την t .

■ Προσθέτουμε $\emptyset$ -μεταβάσεις από κάθε κατάσταση $K \setminus \{s, t\}$ προς κάθε κατάσταση $K \setminus \{s, t\}$, αν η μετάβαση δεν υπάρχει ήδη.



Υπάρχει και η μετάβαση $s \rightarrow t$, με ετικέτα $\emptyset$, δεν σημειώνεται για πρακτικούς λόγους.

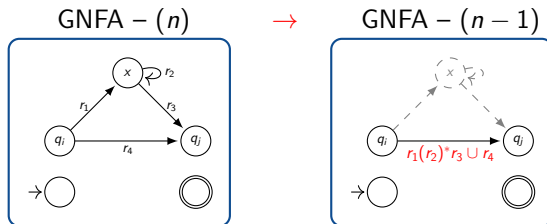
Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

Λήμμα. Κάθε GNFA έχει μια ισοδύναμη κανονική έκφραση R .

Απόδειξη. [Με μαθηματική επαγωγή στον αριθμό των καταστάσεων του GNFA]

Για $n = 2$: $\rightarrow \circ \xrightarrow{r} \odot$ και $R = r$.

Για $n > 2$: Υποθέτουμε ότι ισχύει για $n - 1$ καταστάσεις, θα δείξουμε ότι ισχύει και για n καταστάσεις.



Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

Απόδειξη (συνέχεια).

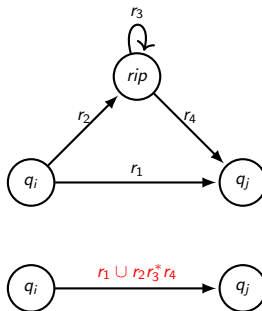
- Διαλέγουμε οποιαδήποτε κατάσταση x , πλην της αρχικής και της τελικής.
- Αφαιρούμε την κατάσταση x .
- Επαναδρομολογούμε όλα τα διερχόμενα από την x μονοπάτια.
- Επανακαθορίζουμε τις μεταβάσεις για κάθε ζεύγος καταστάσεων q_i, q_j .



- Συνεπώς, υπάρχει ισοδυναμία μεταξύ DFAs και κανονικών εκφράσεων.

Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

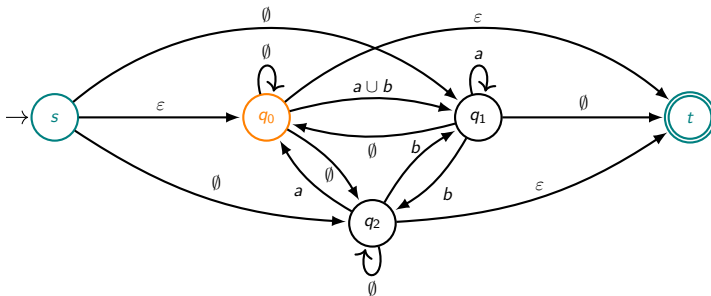
- Αφαιρούμε καταστάσεις από το $K \setminus \{s, t\}$, μια κάθε φορά.
- Επαναδρομολογούμε τις μεταβάσεις με κατάλληλες ετικέτες, όπως φαίνεται στο σχήμα



Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

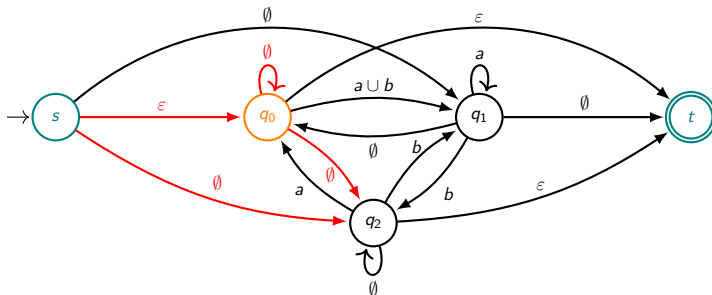
π.χ., (συνέχεια)

■ Ας διαγράψουμε την κατάσταση q_0 το GNFA θα γίνει



Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

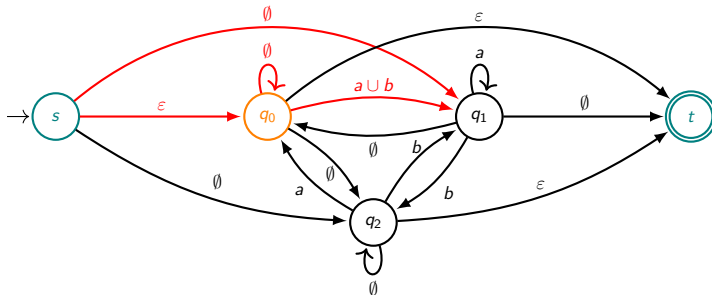
π.χ., (συνέχεια) ας διαγράψουμε την κατάσταση q_0



■ Η s συνδέεται μέσω της $\emptyset \cup \epsilon \emptyset^* \emptyset = \emptyset$ με την q_2 .

Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

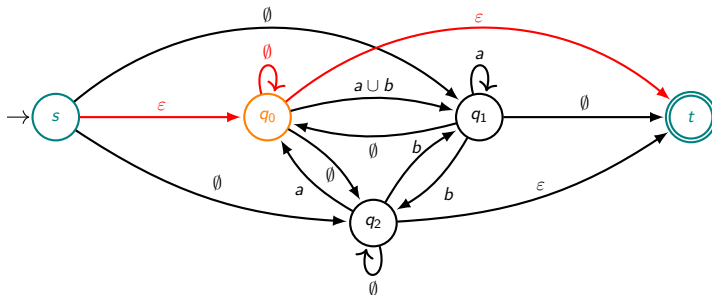
π.χ., (συνέχεια) ας διαγράψουμε την κατάσταση q_0



■ Η s συνδέεται μέσω της $\emptyset \cup \epsilon \emptyset^*(a \cup b) = a \cup b$ με την q_1 .

Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

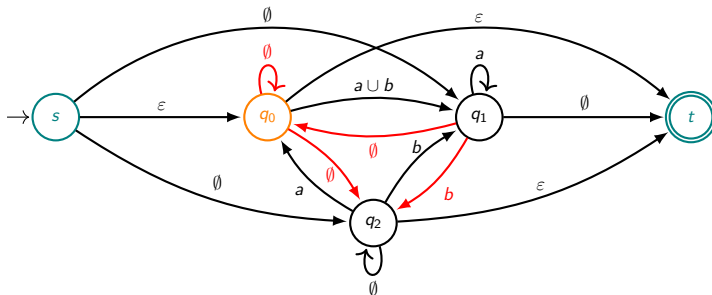
π.χ., (συνέχεια) ας διαγράψουμε την κατάσταση q_0



■ Η s συνδέεται, μέσω της $\emptyset \cup \epsilon\emptyset^*\epsilon = \epsilon$ με την t .

Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

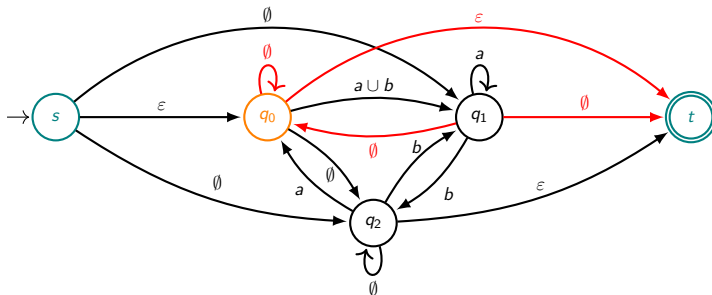
π.χ., (συνέχεια) ας διαγράψουμε την κατάσταση q_0



■ Η q_1 έχει βρόγχο a και συνδέεται μέσω της $b \cup \emptyset\emptyset^*\emptyset = b$ με την q_2 .

Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

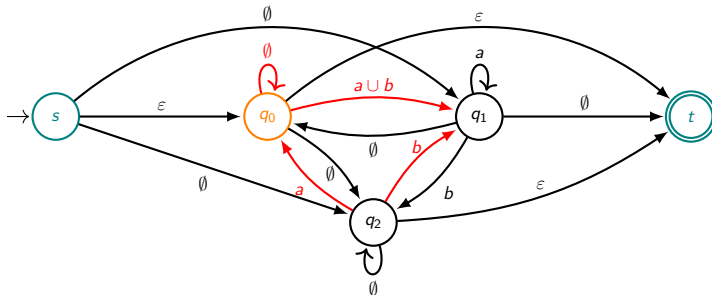
π.χ., (συνέχεια) ας διαγράψουμε την κατάσταση q_0



■ Η q_1 συνδέεται μέσω της $\emptyset \cup \emptyset\emptyset^*\epsilon = \emptyset$ με την t .

Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

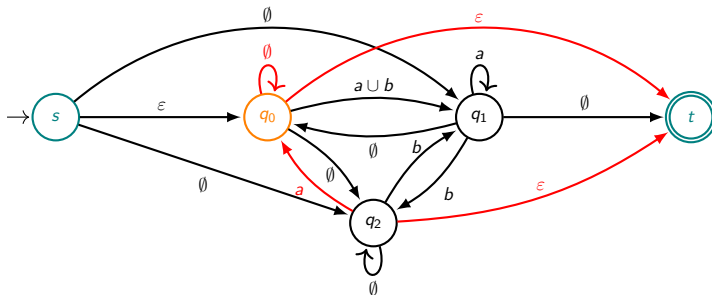
π.χ., (συνέχεια) ας διαγράψουμε την κατάσταση q_0



■ Η q_2 έχει βρόγχο $\emptyset$ και συνδέεται μέσω της $b \cup a\emptyset^*(a \cup b) = b \cup a(a \cup b)$ με την q_1 .

Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

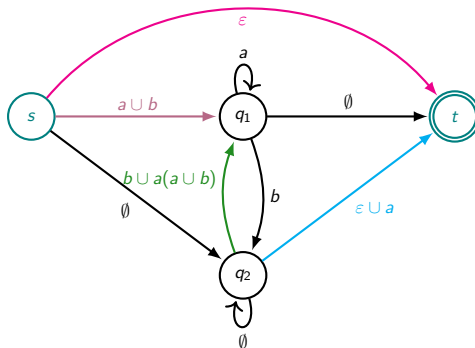
π.χ., (συνέχεια) ας διαγράψουμε την κατάσταση q_0



■ Η q_2 συνδέεται μέσω της $\epsilon \cup a\emptyset^*\epsilon = \epsilon \cup a$ με την t .

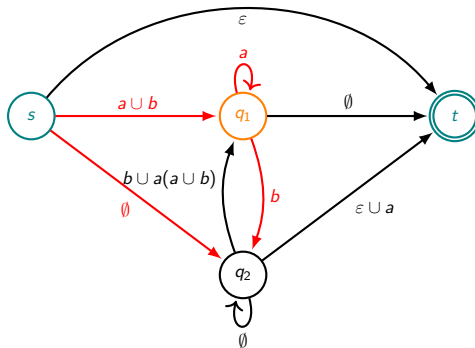
Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

π.χ., (συνέχεια) αφού διαγράψουμε την κατάσταση q_0 και επαναδρομολογήσουμε τις μεταβάσεις, το GNFA θα γίνει



Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

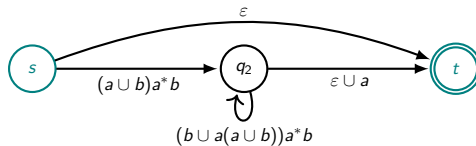
π.χ., (συνέχεια) ας διαγράψουμε την κατάσταση q_1



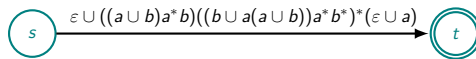
■ Η s συνδέεται μέσω της $\emptyset \cup (a \cup b)a^*b$ με την q_2 . Η q_2 τώρα θα έχει βρόγχο $(b \cup a(a \cup b))a^*b$ και συνδέεται μέσω της $(\varepsilon \cup a) \cup (b \cup a(a \cup b))a^*\emptyset = \varepsilon \cup a$ με την t .

Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

π.χ., (συνέχεια) αφού διαγράψουμε την κατάσταση q_1 και επαναδρομολογήσουμε τις μεταβάσεις, το GNFA θα γίνει,

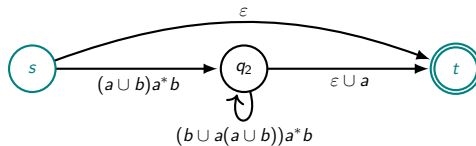


Τέλος, διαγράφοντας την κατάσταση q_2 παίρνουμε,

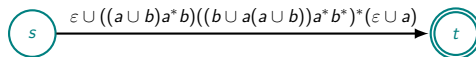


Πεπερασμένα αυτόματα και Κανονικές Εκφράσεις

π.χ., (συνέχεια) αφού διαγράψουμε την κατάσταση q_1 και επαναδρομολογήσουμε τις μεταβάσεις, το GNFA θα γίνει,



Τέλος, διαγράφοντας την κατάσταση q_2 παίρνουμε,



Συνεπώς, η κανονική έκφραση που αντιστοιχεί στο δοθέν NFA είναι

$$\varepsilon \cup ((a \cup b)a^*b)((b \cup a(a \cup b))a^*b^*)^*(\varepsilon \cup a).$$

□

Μη Κανονικές Γλώσσες

- Για να δείξουμε ότι μια γλώσσα είναι κανονική χρειαζόμαστε ένα DFA.
- Για να δείξουμε ότι μια γλώσσα είναι **δεν** είναι κανονική χρειαζόμαστε απόδειξη.
- Είναι ανεπαρκές το επιχείρημα ότι επειδή δεν μπορούμε να βρούμε κάποιο DFA, τότε και η γλώσσα δεν είναι κανονική. (εκτός αν αποδείξουμε ότι δεν υπάρχει)

Μη Κανονικές Γλώσσες

Λήμμα Άντλησης. Για κάθε κανονική γλώσσα A , υπάρχει αριθμός p τ.ω.

αν $s \in A$ και $|s| \geq p$, τότε $s = xyz$ όπου

- $xy^iz \in A \ \forall i \geq 0$
- $y \neq \varepsilon, [|y| \neq 0]$
- $|xy| \leq p$
- Η διαμέριση xyz μπορεί να είναι διαφορετική για κάθε $s \in A$.
- Αν A κανονική γλώσσα, τότε κάθε 'μεγάλη' συμβολοσειρά στην A μπορεί να 'αντληθεί' και το αποτέλεσμα να παραμείνει στην A .

Μη Κανονικές Γλώσσες

Απόδειξη.

Έστω DFA M με $|K| = p$ που δέχεται την A . Έστω $s \in A$ τ.ω. $|s| \geq p$. Τα πρώτα p βήματα του υπολογισμού της M για τα πρώτα p σύμβολα $(s_1, \dots, s_p)$ της s είναι

$$(q_0, s_1 \dots s_p) \vdash_M (q_1, s_2 \dots s_p) \vdash_M \dots \vdash_M (q_p, \varepsilon)$$

Τουλάχιστον 2 από τις καταστάσεις $q_0, \dots, q_p$ είναι **ίδιες**.

[Αρχή Περιστερώνας: Το M έχει p καταστάσεις και διατρέχει παραπάνω από p βήματα, άρα θα επισκεφθεί μια κατάσταση τουλάχιστον 2 φορές.]

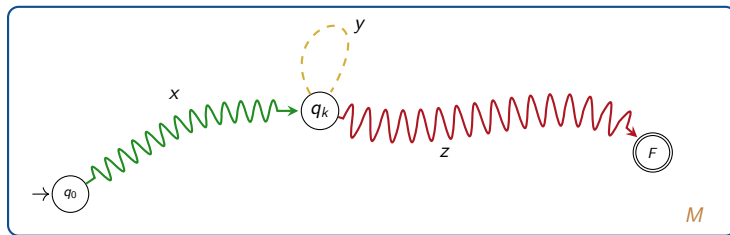
Έστω $s_k = s_j$ για $k < j \leq p$. Έστω τα παρακάτω τμήματα του w

- $x = s_1 \dots s_k$, φέρνει το M στην κατάσταση q_k ,
- $y = s_{k+1} \dots s_j$
- $z = s_{j+1} \dots s_{|w|}$

$s = xyz$, $y \neq \varepsilon$, $|xy| \leq p$.

Μη Κανονικές Γλώσσες

Απόδειξη (συνέχεια).



□

Το M δέχεται **όλες** τις συμβολοσειρές xy^iz .

Μη Κανονικές Γλώσσες

Ένας δυνατός τρόπος απόδειξης ότι μια γλώσσα A **δεν** είναι κανονική.

- Αρχικά υποθέτουμε ότι είναι κανονική.
- Άρα το Θεώρημα Άντλησης θα πρέπει να ισχύει.
 - Διαλέγουμε $s \in A$, τ.ω. $|s| \geq p$.
 - Διαλέγουμε y , τ.ω. $|y| \neq \emptyset$.
 - Διαλέγουμε x , τ.ω. $|xy| \leq p$.
 - Ότι έχει απομείνει από την s το αναθέτουμε στην z .
- Διαλέγουμε k ακέραιο, τ.ω. $xy^kz \notin A$.
- Καταλήγουμε σε άτοπο.

Μη Κανονικές Γλώσσες

π.χ., Έστω $A = \{0^k 1^k \mid k \geq 0\}$, θ.δ.ο η A δεν είναι κανονική.

Απόδειξη. (1ος τρόπος)

Το Λήμμα Άντλησης δίνει το p . Έστω $s = 0^p 1^p \in A$.

Από το Λήμμα Άντλησης μπορούμε να διαμερίσουμε την $s = xyz$ ικανοποιώντας τις 3 συνθήκες.

Από τις συνθήκες $|xy| \leq p$ και $y \neq \emptyset$ έχουμε ότι η y αποτελείται από 0, έστω 0^k , ενώ όλα οι 1 βρίσκονται στην z και τέλος x^{p-k} .

Συνεπώς, στη $xyyz = 0^{p-k} 0^k 0^k 1^p = 0^{p+k} 1^p$, τα 0 αυξάνονται και οι 1 παραμένουν σταθεροί σε πλήθος, επομένως $xyyz \notin A$, άτοπο. Άρα, η A μη-κανονική. $\square$

Μη Κανονικές Γλώσσες

π.χ., Έστω $A = \{0^k 1^k \mid k \geq 0\}$, θ.δ.ο η A δεν είναι κανονική.

Απόδειξη. (2ος τρόπος)

Το Λήμμα Άντλησης δίνει το p . Έστω $s = 0^q 1^q \in A$, για $q < p \leq 2q$.

Από το Λήμμα Άντλησης μπορούμε να διαμερίσουμε την $s = xyz$ ικανοποιώντας τις 3 συνθήκες.

Από τις συνθήκες $|xy| \leq p$ και $y \neq \emptyset$ τώρα η y μπορεί να αποτελείται και από 0, και από 1.

Εξετάζουμε τις περιπτώσεις: $y = \{0^k, 1^k, 0^m 1^\ell\}$ και αντλούμε.

Σε κάθε περίπτωση καταλήγουμε ότι $xyyz \notin A$, άτοπο. Άρα, η A μη-κανονική. □

Μη Κανονικές Γλώσσες

π.χ., Έστω $B = \{ww \mid w \in \Sigma^*\}$ όπου $\Sigma^* = \{0, 1\}^*$. Θ.δ.ο η B δεν είναι κανονική.

Απόδειξη. (Μέσω εις άτοπο)

Το Λήμμα Άντλησης δίνει το p . Πρέπει να διαλέξουμε κατάλληλο $s \in B$.

- $s = 0^p 0^p \in B$. Δεν είναι καλή επιλογή, γιατί μπορεί να αντληθεί και να παραμείνει μέσα στο B . [γιατί ;]
- $s = 0^p 1 0^p 1 \in B$. Δεν μπορεί να αντληθεί και να ικανοποιούνται οι 3 προϋποθέσεις του λήμματος. [γιατί ;]
- $xyz \notin B$, άρα άτοπο.

Άρα η B δεν είναι κανονική. □

Μη Κανονικές Γλώσσες

π.χ., Έστω $C = \{w \mid \text{η } w \text{ έχει ίδιο αριθμό 0 και 1}\}$. Θ.δ.ο η C δεν είναι κανονική.

Θα χρησιμοποιήσουμε ιδιότητες κλειστότητας.

Απόδειξη. (Μέσω εις άτοπο)

- Έστω ότι η C είναι κανονική.
- Η 0^*1^* είναι κανονική [γιατί ;], άρα και η $C \cap 0^*1^*$ είναι κανονική [από κλειστότητα τομής].
- Αλλά ξέρουμε ότι η $D = C \cap 0^*1^*$ είναι μη-κανονική [γιατί ;], άρα άτοπο.
- Άρα C μη-κανονική.

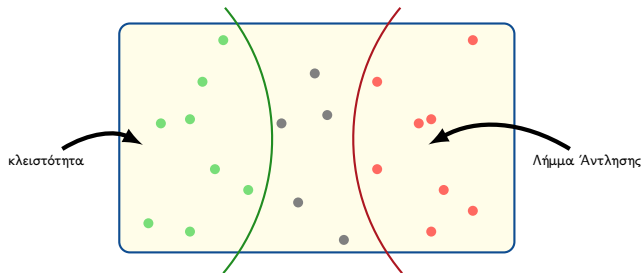


Κανονικές Γλώσσες

■. Είδαμε ότι υπάρχουν κανονικές γλώσσες

π.χ., $\{w \in \{0,1\}^* \mid \text{ο δεκαδικός αριθμός που αναπαρίσταται ως } w \text{ διαιρείται με το } 6\}$

και μη-κανονικές γλώσσες, π.χ., $\{0^k 1^k \mid k \geq 0\}$.



■ Μπορούμε να βρούμε κάποιον χαρακτηρισμό ;