



ΗΥ - 280 Θεωρία Υπολογισμού

Διάλεξη 6

Κωνσταντίνος Βάρσος

Τμήμα Επιστήμης Υπολογιστών – Πανεπιστήμιο Κρήτης

Χειμερινό εξάμηνο 2025 - 26

Γλώσσες και Αυτόματα

Τι είδαμε στην Διάλεξη 5

- Κανονικές εκφράσεις.
- Μετατροπή κανονικών εκφράσεων σε πεπερασμένα αυτόματα.
- Μη-κανονικές γλώσσες
- Λήμμα Άντλησης για κανονικές γλώσσες.

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

- $xy^iz \in A \quad \forall i \geq 0$

- $y \neq \varepsilon, [|y| \neq 0]$

- $|xy| \leq p$

- 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

Συνέπεια συνθήκης: $|xy| \leq p$

■ η υποσυμβολοσειρά xy βρίσκεται στα πρώτα p σύμβολα της s .

■ η υποσυμβολοσειρά y που θα 'αντλήσουμε' εμφανίζεται μέσα σε ένα οριοθετημένο πρόθεμα της s .

■ δεν αντλούμε από οπουδήποτε, αλλά από το πρώιμο μέρος που καθορίζεται από τον αριθμό των καταστάσεων p στο NFA που αναγνωρίζει την A .

Μη Κανονικές Γλώσσες

Συνέπεια συνθήκης: $y \neq \emptyset$

■ διαφορετικά, η 'άντληση' του y δεν θα είχε κανένα αποτέλεσμα και το λήμμα θα ήταν άνευ νοήματος.

■ η y αντιστοιχεί στον βρόχο στη διαδρομή υπολογισμού στο NFA.

Συνέπεια συνθήκης: $xy^iz \in A, \forall i \geq 0$

■ σε ένα NFA, η επανάληψη του ίδιου βρόχου καταστάσεων δεν αλλάζει την αποδοχή. Το αυτόματο τερματίζει με το ίδιο μοτίβο z καταστάσεων κάθε φορά, επομένως η γλώσσα είναι κλειστή υπό την 'άντληση'.

Μη Κανονικές Γλώσσες

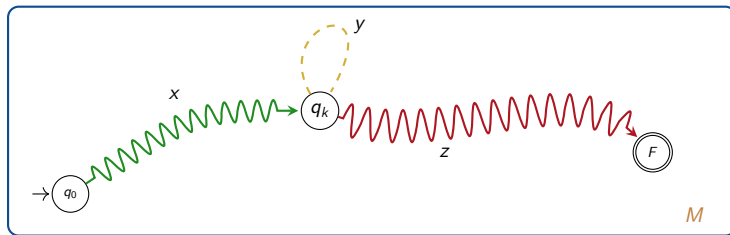
Γιατί αντλούμε μόνο το y ;

■ Επειδή ο πεπερασμένος έλεγχος σε ένα DFA μπορεί να προσπελάσει μόνο ένα μονοπάτι κάθε φορά.

■ Όταν διαβάζετε μια συμβολοσειρά s , το αυτόματο ακολουθεί μια μοναδική ντετερμινιστική διαδρομή. Εάν μια κατάσταση επαναλαμβάνεται, ορίζει έναν βρόχο.

■ Μπορούμε να αντλήσουμε και από άλλες υποσυμβολοσειρές, π.χ. $x_1y_1x_2y_2 \dots x_my_mz$, αλλά τότε η γλώσσα δεν είναι κανονική.

Μη Κανονικές Γλώσσες



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.
- Αν μπορείτε να βρείτε μια συμβολοσειρά $s \in A$ τ.ω. καμία διαμέριση $s = xyz$ να μην ικανοποιεί και τις τρεις συνθήκες, τότε η A δεν είναι κανονική.
- Δηλαδή το λήμμα χρησιμοποιείται μέσω αντιθετοαναστροφής (contraposition, η αντιθετοαναστροφή της συνεπαγωγής $P \rightarrow Q$ είναι η $\neg Q \rightarrow \neg P$.)

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 όλοι 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.

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. □

Κανονικές Γλώσσες

π.χ., Έστω A η κανονική γλώσσα όλων των συμβολοσειρών στο αλφάβητο $\Sigma = \{0, 1\}$ τ.ω.

$$\text{Max}(A) = \{w \in A \mid \text{για κάθε επέκταση } u \in \Sigma^+ \text{ της } w, \ wu \notin A\}.$$

Δείξτε ότι η $\text{Max}(A)$ κανονική.

Κανονικές Γλώσσες

π.χ., Έστω A η κανονική γλώσσα όλων των συμβολοσειρών στο αλφάβητο $\Sigma = \{0, 1\}$ τ.ω.

$$\text{Max}(A) = \{w \in A \mid \text{για κάθε επέκταση } u \in \Sigma^+ \text{ της } w, \ wu \notin A\}.$$

Δείξτε ότι η $\text{Max}(A)$ κανονική.

Απόδειξη. Έστω M αυτόματο που αναγνωρίζει την A , με σύνολο τερματικών καταστάσεων F .

Διαλέγουμε μια κατάσταση $p \in F$. Αν υπάρχει διαδρομή που οδηγεί πάλι στο σύνολο F , σημειώνουμε αυτή την κατάσταση. [γιατί ;]

Μεταβάλλουμε σε μη-τερματικές καταστάσεις όλες τις σημειωμένες καταστάσεις. [γιατί ;]



Μη Κανονικές Γλώσσες

Προσοχή στη χρήση του Λήμματος Άντλησης

■ Το Λήμμα Άντλησης είναι **αναγκαία** αλλά **όχι** και ικανή συνθήκη για να είναι μια γλώσσα κανονική.

■ Το Λήμμα Άντλησης δηλώνει ότι όλες οι κανονικές γλώσσες ικανοποιούν τις 3 συνθήκες.

■ Υπάρχουν μη-κανονικές γλώσσες που ικανοποιούν τις συνθήκες του! **π.χ.**,

$$L = \left\{ uvnwxy \mid \begin{array}{l} u, y \in \Sigma^*, \\ v, w, x \in \Sigma \\ \wedge (v = w \vee v = x \vee x = w) \end{array} \right\} \cup \left\{ w \in \Sigma^* \mid \begin{array}{l} \text{ακριβώς } 1/7 \text{ των συμβόλων της} \\ w \text{ είναι } 3 \end{array} \right\}$$

με $\Sigma = \{0, 1, 2, 3\}$ [πηγή Wikipedia].

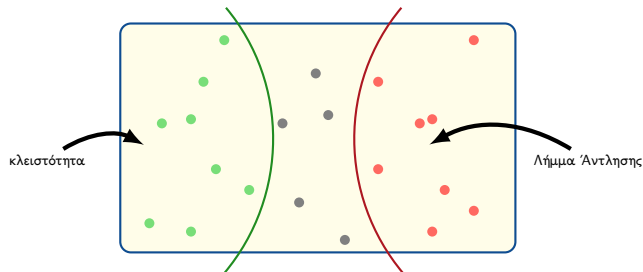
■ Επομένως, χρησιμεύει μόνο για απόδειξη μη-κανονικότητας.

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 ;

Χαρακτηρισμός Myhill-Nerode

■ Πάνω στο Σ^* μπορούμε να ορίσουμε τις εξής σχέσεις ισοδυναμίας:

■ Έστω γλώσσα $A \subseteq \Sigma^*$, ορίζουμε την σχέση R_A τ.ω.

$$xR_A y \text{ ανν } \forall z \in \Sigma^*, (xz \in A \iff yz \in A).$$

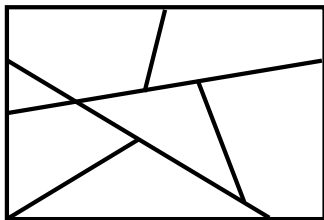
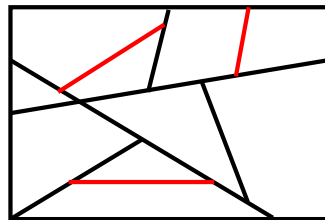
■ Έστω DFA M ορίζουμε την σχέση R_M τ.ω.

$$xR_M y \text{ ανν } \delta(q_0, x) = \delta(q_0, y).$$

■ Δείκτης μιας σχέσης ισοδυναμίας R λέγεται ο πληθάριθμος του συνόλου των κλάσεων ισοδυναμίας της R .

Χαρακτηρισμός Myhill-Nerode

Εκλέπτυνση Ισοδυναμίας. Έστω δύο σχέσεις ισοδυναμίας R και E πάνω στο Ω . Λέμε ότι E είναι εκλέπτυνση (refinement) της R αν κάθε κλάση του E είναι υποσύνολο μιας κλάσης του R .


 Ω/G

 Ω/E

Χαρακτηρισμός Myhill-Nerode

■ Μια σχέση R ορισμένη πάνω στο Σ^* λέγεται **δεξιά αναλλοίωτη** αν $\forall x, y, z \in \Sigma^*$ ισχύει $xRy \Rightarrow xzRyz$.

Πρόταση. Η R_M είναι δεξιά αναλλοίωτη.

Απόδειξη. $\delta(q_0, xz) = \delta(\delta(q_0, x), z) = \delta(\delta(q_0, y), z) = \delta(q_0, yz)$. □

■ Επομένως, $\forall x, y, z \in \Sigma^*$, $xR_M y \Rightarrow xzR_M yz$.

Χαρακτηρισμός Myhill-Nerode

■ Έστω γλώσσα $A \subseteq \Sigma^*$, ορίζουμε την σχέση R_A τ.ω. $xR_A y$ ανν $\forall z \in \Sigma^*$ να ισχύει $xz \in A \Leftrightarrow yz \in A$.

Πρόταση. Η R_A είναι σχέση ισοδυναμίας.

Απόδειξη.

- i) $\forall z \in \Sigma^*$ ισχύει $xz \in A \Leftrightarrow xz \in A$ [$xR_A x$].
- ii) Αν $xz \in A \Leftrightarrow yz \in A$ ανν $yz \in A \Leftrightarrow xz \in A$, $\forall z \in \Sigma^*$ [$xR_A y \Leftrightarrow yR_A x$].
- iii) Αν $xz \in A \Leftrightarrow yz \in A$ και $yz \in A \Leftrightarrow tz \in A$, $\forall z \in \Sigma^*$, τότε $xz \in A \Leftrightarrow tz \in A$, $\forall z \in \Sigma^*$ [$(xR_A y) \wedge (yR_A t) \Leftrightarrow xR_A t$]. □

■ Για κάθε συμβολοσειρά $x \in \Sigma^*$, το σύνολο

$$[x]_{R_A} = \{y \in \Sigma^* \mid xR_A y\},$$

λέγεται η κλάση ισοδυναμίας του x ως προς την σχέση R_A .

Χαρακτηρισμός Myhill-Nerode

■ Έστω γλώσσα $A \subseteq \Sigma^*$, ορίζουμε την σχέση R_A τ.ω. $xR_A y$ ανν $\forall z \in \Sigma^*, (xz \in A \iff yz \in A)$.

π.χ., Έστω η γλώσσα $A = \{w \in \{0,1\}^* \mid w \text{ τελειώνει με } 10\}$.

Οι συμβολοσειρές 01011 και 100 δεν ανήκουν ίδια κλάση ισοδυναμίας καθώς για $z = 0$, $010110 \in A$, ενώ $1000 \notin A$. [άρα είναι διακεκριμένες]

Οι συμβολοσειρές 0 και 100 ανήκουν ίδια κλάση ισοδυναμίας καθώς για $\forall z$, $0z \in A \iff 100z \in A$. [άρα δεν είναι διακεκριμένες]

Οι συμβολοσειρές 1 και 101 ανήκουν ίδια κλάση ισοδυναμίας καθώς για $\forall z$, $1z \in A \iff 101z \in A$. [άρα δεν είναι διακεκριμένες]



Χαρακτηρισμός Myhill-Nerode

■ Έστω DFA M ορίζουμε την σχέση R_M τ.ω. $xR_M y$ ανν $\delta(q_0, x) = \delta(q_0, y)$.

Πρόταση. Η R_M είναι σχέση ισοδυναμίας.

Απόδειξη.

- i) $\delta(q, x) = \delta(q, x)$ [$xR_M x$].
- ii) Αν $\delta(q, x) = \delta(q, y)$, ανν $\delta(q, y) = \delta(q, x)$ [$xR_M y \Leftrightarrow yR_M x$].
- iii) Αν $\delta(q, x) = \delta(q, y)$ και $\delta(q, y) = \delta(q, z)$, τότε $\delta(q, x) = \delta(q, z)$ [$(xR_M y) \wedge (yR_M z) \Leftrightarrow xR_M z$].

□

■ Για κάθε συμβολοσειρά $x \in \Sigma^*$, το σύνολο

$$[x]_{R_M} = \{y \in \Sigma^* \mid xR_M y\},$$

λέγεται η κλάση ισοδυναμίας του x ως προς την σχέση R_M .

Χαρακτηρισμός Myhill-Nerode

Διαισθητικά,

- Δύο προθέματα x, y είναι ισοδύναμα κατά R_A αν παραθέσω οποιοδήποτε z , το αν $xz \in A$ είναι ακριβώς το ίδιο με το αν το $yz \in A$, δηλαδή x, y 'επεκτείνονται ομοίως'.
- Μια γλώσσα $A \subseteq \Sigma^*$ είναι κανονική ακριβώς όταν ο χώρος όλων των προθεμάτων που 'επεκτείνονται ομοίως' διαιρείται σε πεπερασμένα σύνολα.
- Αν υπάρχουν πεπερασμένες τέτοιες κλάσεις, τότε ένα DFA με μια κατάσταση για κάθε κλάση μπορεί να αναγνωρίσει την A .
- Αν υπάρχουν άπειρες κλάσεις, κανένα πεπερασμένο αυτόματο δεν αρκεί.

Χαρακτηρισμός Myhill-Nerode

Θεώρημα [Myhill-Nerode]. Για μια γλώσσα A τα ακόλουθα είναι ισοδύναμα:

- 1 Η γλώσσα A είναι κανονική.
- 2 Η γλώσσα A είναι ένωση κάποιων κλάσεων ισοδυναμίας μιας δεξιάς αναλλοίωτης σχέσης ισοδυναμίας R με πεπερασμένο δείκτη.
- 3 Η σχέση R_A έχει πεπερασμένο δείκτη.

Χαρακτηρισμός Myhill-Nerode

Απόδειξη. $[1 \Rightarrow 2]$

■ Έστω DFA $M = (K, \Sigma, \delta, q_0, F)$ που αναγνωρίζει την A .

■ Η R_M είναι δεξιά αναλλοίωτη.

■ Οι κλάσεις ισοδυναμίας της R_M είναι όσες και οι καταστάσεις του M , δηλαδή η R_M είναι πεπερασμένου δείκτη.

■ Η γλώσσα που γίνεται αποδεκτή από το M είναι ένωση των κλάσεων ισοδυναμίας που αντιστοιχούν σε τελικές καταστάσεις του αυτομάτου.

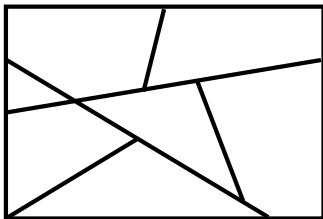
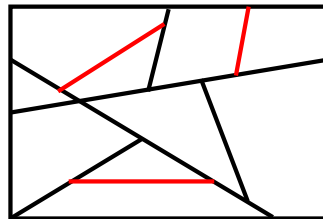
Χαρακτηρισμός Myhill-Nerode

Απόδειξη. $[2 \Rightarrow 3]$ Έστω η ισοδυναμία E που ικανοποιεί την ιδιότητα που δίνεται στο 2. Θα δείξουμε την συνεπαγωγή:

Αν για $x, y \in \Sigma^*$ ισχύει xEy , τότε $xR_A y$

- **Απόδειξη της συνεπαγωγής:** $xEy \Rightarrow xR_A y$.
 - Έστω $z \in \Sigma^*$.
 - Αν $xz \in A$ και xEy , τότε $xzEyz$, επειδή η E είναι δεξιά αναλλοίωτη. [είτε αμφότερα τα xz, yz ανήκουν στην A είτε κανένα.]
 - Αν $xz \in A$ και $xzEyz$, τότε $yz \in A$. [αφού $xz \in A$, όλη η E -κλάση του xz περιέχεται στην A , άρα $yz \in A$]
 - Συνεπώς, από τον ορισμό της R_A παίρνουμε $xR_A y$.
- Άρα, κάθε κλάση ισοδυναμίας της E περιέχεται σε μια κλάση ισοδυναμίας της R_A .
- Δηλαδή, η σχέση E αποτελεί **εκλέπτυνση** της R_A .
- Εφόσον, η E είναι πεπερασμένου δείκτη, και η R_A είναι πεπερασμένου δείκτη.

Χαρακτηρισμός Myhill-Nerode


 Σ^*/R_A

 Σ^*/E

Χαρακτηρισμός Myhill-Nerode

Απόδειξη. $[3 \Rightarrow 1]$ Υποθέτουμε ότι η R_A έχει πεπερασμένο δείκτη και δείχνουμε ότι η A αναγνωρίζεται από κάποιο DFA $M = (K, \Sigma, \delta, q_0, F)$, άρα είναι κανονική.

Ορισμός του M :

- Σύνολο καταστάσεων K : Όλες οι R_A -κλάσεις, οι οποίες είναι πεπερασμένες από την υπόθεση. $[K = \{[x]_{R_A} \mid x \in \Sigma^*\}]$
- Αρχική κατάσταση: Η κλάση $[\varepsilon]_{R_A}$ της κενής συμβολοσειράς.
- Συνάρτηση μετάβασης δ : Ορίζεται ως $\delta([x]_{R_A}, \alpha) = [x\alpha]_{R_A}$, όπου $\alpha \in \Sigma$.
- Τελικές καταστάσεις: $F = \{[x]_{R_A} \mid x \in A\}$.
- Το M αποδέχεται την A αφού $\delta(q_0, x) = [x]_{R_A}$ και επομένως $x \in L(M)$ ανν $[x]_{R_A} \in F$. $\square$

Χαρακτηρισμός Myhill-Nerode

Παρατηρήσεις.

- Για κάθε κανονική γλώσσα A , η R_A είναι πεπερασμένου δείκτη.
- Η R_A είναι δεξιά αναλλοίωτη.
- Αν $[x]_R = [y]_R$, δηλαδή xRy , τότε $[x\alpha]_R = [y\alpha]_R, \forall \alpha \in \Sigma$.
- Συνέπεια του Θεωρήματος είναι ότι μπορούμε να κατασκευάσουμε μοναδικό DFA με τις ελάχιστες δυνατές καταστάσεις.

Χαρακτηρισμός Myhill-Nerode

π.χ., Δείξτε ότι η γλώσσα $B = \{a^n b^n \mid n \geq 0\}$ δεν είναι κανονική.

■ Το $S_1 = \{a^n \mid n \geq 0\}$, στο $\Sigma = \{a, b\}$, είναι άπειρο (άπειρα αριθμήσιμο).

■ Θα δείξουμε ότι οι συμβολοσειρές του είναι ανά δύο διακεκριμένες ως προς την B . Έστω τώρα δύο τυχαίες συμβολοσειρές της B , η a^k και η a^m , όπου k, m θετικοί ακέραιοι με $k \neq m$.

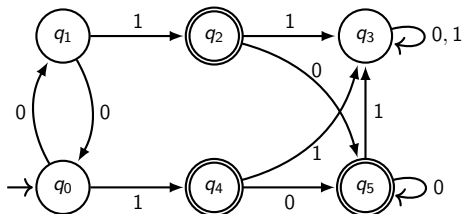
■ Διαλέγουμε μια συμβολοσειρά b^k και την παραθέτουμε στις a^k και a^m . Επομένως, παίρνουμε τις συμβολοσειρές $a^k b^k$ και $a^m b^k$.

■ Συνεπώς, $a^k b^k \in B$ και $a^m b^k \notin B$. Επομένως, οι a^k και a^m είναι διακεκριμένες ως προς την B , άρα δεν ανήκουν στην ίδια κλάση. Καθώς η a^k και η a^m είναι τυχαίες συμβολοσειρές της B , η B θα πρέπει να έχει άπειρες το πλήθος κλάσεις.

■ Άρα, από το Θεωρήμα Myhill-Nerode η B δεν είναι κανονική. $\square$

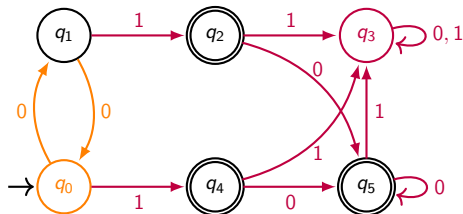
Ελαχιστοποίηση DFA

π.χ., Έστω το αυτόματο M που δίνεται στο σχήμα, το οποίο αποδέχεται την γλώσσα $A = 0^*10^*$.



Ελαχιστοποίηση DFA

π.χ., Έστω το αυτόματο M που δίνεται στο σχήμα, το οποίο αποδέχεται την γλώσσα $A = 0^*10^*$.



Στην $L(M)$ υπάρχουν 6 κλάσεις ισοδυναμίας,

$C_{q_0} = (00)^*$, $C_{q_1} = (00)^*0$, $C_{q_2} = (00)^*01$, $C_{q_3} = 0^*10^*1(0+1)^*$, $C_{q_4} = (00)^*1$, $C_{q_5} = 0^*100^*$,

οπότε $A = C_{q_2} \cup C_{q_4} \cup C_{q_5}$.

Ελαχιστοποίηση DFA

π.χ., Έστω το αυτόματο M που δίνεται στο σχήμα, το οποίο αποδέχεται την γλώσσα $A = 0^*10^*$.

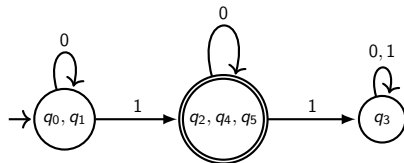
Στην $L(M)$ υπάρχουν 6 κλάσεις ισοδυναμίας,

$$C_{q_0} = (00)^*, C_{q_1} = (00)^*0, C_{q_2} = (00)^*01, C_{q_3} = 0^*10^*1(0+1)^*, C_{q_4} = (00)^*1, C_{q_5} = 0^*100^*,$$

$$\text{οπότε } A = C_{q_2} \cup C_{q_4} \cup C_{q_5}.$$

Παρατηρούμε τα εξής,

- C_{q_0}, C_{q_1} , κλάσεις με 0^*
- $C_{q_2}, C_{q_4}, C_{q_5}$, κλάσεις με μοναδικό 1
- C_{q_3} , κλάσεις με πολλαπλά 1 $\square$



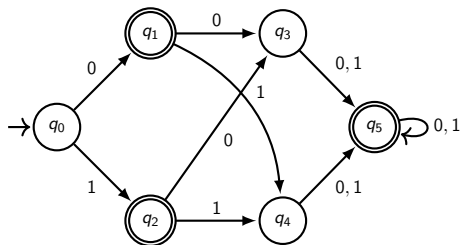
Ελαχιστοποίηση DFA

Αλγόριθμος. Έστω DFA $M = (K, \Sigma, \delta, q_0, F)$ χωρίς μη-προσβάσιμες καταστάσεις.

- 1 Δημιουργούμε πίνακα ζευγών $\{p, q\}$, όπου $p, q \in K$ [$p \neq q$]. Όλα τα στοιχεία του πίνακα είναι αρχικά κενά.
- 2 Μαρκάρουμε το ζεύγος $\{p, q\}$ αν $p \in F$ και $q \notin F$, ή το αντίστροφο.
- 3 Επαναλαμβάνουμε το 2. μέχρι να διατρέξουμε ολόκληρο τον πίνακα χωρίς να μαρκαριστεί κάποιο νέο ζεύγος:
 - Αν το $\{p, q\}$ είναι αμαρκάριστο και υπάρχει ένα σύμβολο $a \in \Sigma$ τ.ω. $\{\delta(p, a), \delta(q, a)\}$ είναι μαρκαρισμένο, τότε μαρκάρουμε το $\{p, q\}$.
- 4 Μετά την ολοκλήρωση, $p \approx q$ αν και μόνο αν το $\{p, q\}$ δεν είναι μαρκαρισμένο.

Ελαχιστοποίηση DFA

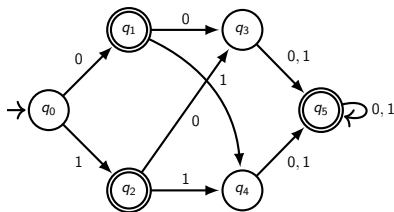
π.χ., Έστω το DFA $M = (K, \Sigma, \delta, q_0, F)$,



Λέμε ότι $p \approx q$ αν $\forall x \in \Sigma^*, \delta(p, x) \in F \Leftrightarrow \delta(q, x) \in F$.

Ελαχιστοποίηση DFA

π.χ., (συνέχεια)

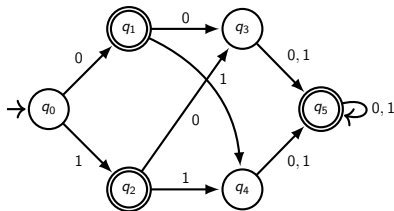


1. Δημιουργούμε πίνακα ζευγών $\{p, q\}$, όπου $p, q \in K$. Όλα τα στοιχεία του πίνακα είναι αρχικά κενά.

	q_0				
☐	q_1				
☐	☐	q_2			
☐	☐	☐	q_3		
☐	☐	☐	☐	q_4	
☐	☐	☐	☐	☐	q_5

Ελαχιστοποίηση DFA

π.χ., (συνέχεια)

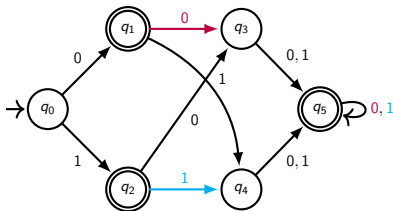


2. Μαρκάρουμε το ζεύγος $\{p, q\}$ αν $p \in F$ και $q \notin F$, ή το αντίστροφο.

q_0				
☒	☐	☐	☐	q_1
☒	☐	☐	☐	q_2
☐	☒	☒	☐	q_3
☐	☒	☒	☐	q_4
☒	☐	☐	☒	☒ q_5

Ελαχιστοποίηση DFA

π.χ., (συνέχεια)



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

	q_0			
☒	q_1			
☒		q_2		
	☒	☒	q_3	
	☒	☒		q_4
☒			☒	☒ q_5

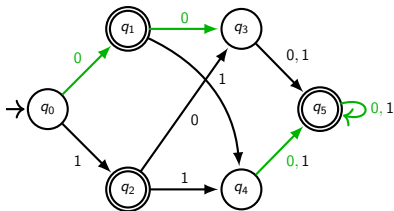
βήμα νο 1

	q_0			
☒	q_1			
☒		q_2		
	☒	☒	q_3	
	☒	☒		q_4
☒	☒	☒	☒	☒ q_5

βήμα νο 2

Ελαχιστοποίηση DFA

π.χ., (συνέχεια)



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

q_0				
✓				q_1
✓				q_2
	✓	✓		q_3
	✓	✓		q_4
✓			✓	✓

βήμα νο 1

q_0				
✓				q_1
✓				q_2
	✓	✓		q_3
	✓	✓		q_4
✓	✓	✓	✓	✓

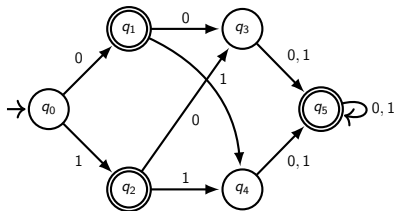
βήμα νο 2

q_0				
✓				q_1
✓				q_2
✓	✓	✓		q_3
✓	✓	✓		q_4
✓	✓	✓	✓	✓

βήμα νο 3

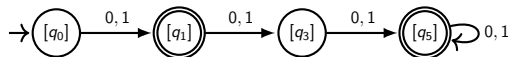
Ελαχιστοποίηση DFA

Άρα DFA $M = (K, \Sigma, \delta, q_0, F)$, λέμε ότι $p \approx q$ αν $\forall x \in \Sigma^*, \delta(p, x) \in F \Leftrightarrow \delta(q, x) \in F$.

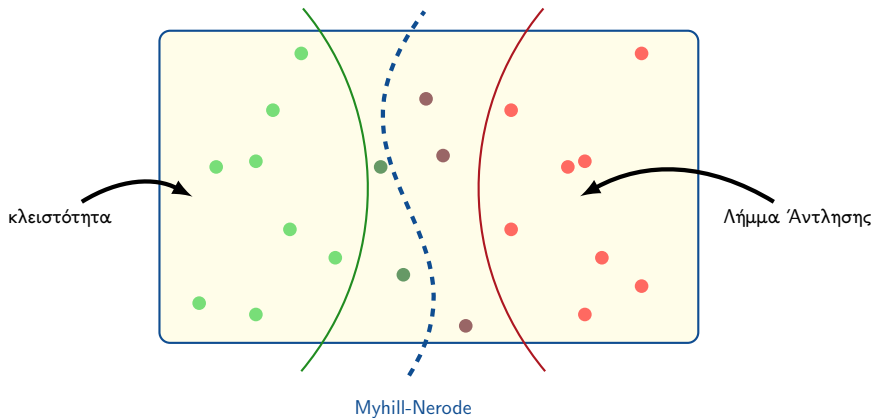


3. Από τον τελευταίο πίνακα παίρνουμε ότι $q_1 \approx q_2$ και $q_3 \approx q_4$.

Επομένως το ελαχιστοποιημένο DFA είναι,



Χαρακτηρισμός Myhill-Nerode



Λήμμα άντλησης vs Myhill-Nerode

ΛΑ. Είναι συνέπεια της ιδέας επανάληψης κατά την εκτέλεση στο DFA.

ΛΑ. Κάθε συμβολοσειρά μήκους $\geq p$ έχει δύο προθέματα που οδηγούν στην ίδια κατάσταση, τότε υπάρχει βρόγχος, τότε μπορούμε να αντλήσουμε (p καταστάσεις $\rightarrow$ loop $\rightarrow$ άντληση). Αυτή είναι μια ιδιότητα που έχουν όλες οι κανονικές γλώσσες.

MN. Ανεξαρτήτως DFA τα προθέματα χωρίζονται σε πεπερασμένες ή άπειρες κλάσεις ισοδυναμίας.

■ Από το MN παίρνουμε το ΛΑ, αν έχει πεπερασμένες κλάσεις, τότε η τρέχουσα θέση του αυτόματου επαναλαμβάνεται (p καταστάσεις $\rightarrow$ loop $\rightarrow$ άντληση).

■ Το αντίστροφο δεν ισχύει. Το ΛΑ είναι αναγκαία συνθήκη, το MN είναι ικανή και αναγκαία συνθήκη (δίνει χαρακτηρισμό).