Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών & Μηχανικών Υπολογιστών
Τομέας Τεχνολογίας Πληροφορικής και Υπολογιστών
Εργαστήριο Λογικής και Επιστήμης Υπολογισμών

Θεμελιώδη Θέματα Επιστήμης Υπολογιστών

Λυμένες Ασκήσεις

Επιμέλεια: Αντώνης Αντωνόπουλος

30 Νοεμβρίου 2018

Περιεχόμενα

1 Αυτόματα & Κανονικές Εκφράσεις

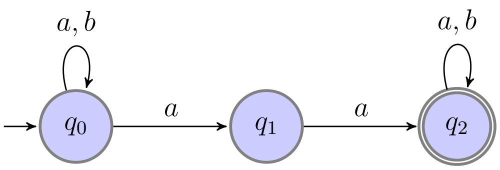
Άσκηση 1.1

Κατασκευάστε NFA, DFA και κανονική έκφραση που να αναγνωρίζει την γλώσσα:

$$L_1 = \{x \in \{a,b\}^* \mid \text{η } x \text{ περιέχει το } 'aa'\}$$

Λύση

Για ευκολία, μπορούμε να κατασκευάσουμε πρώτα το NFA που αποδέχεται την L_1. Θέλουμε το NFA να αποδέχεται όταν υπάρχει το aa μέσα στην συμβολοσειρά που έχουμε ως είσοδο. Οπότε, κατασκευάζουμε τον "σκελετό" του NFA που αποτελείται από τις καταστάσεις q_0, q_1, q_2, όπου μεταβαίνουμε από την q_0 στην q_1 και από την q_1 στην q_2 με a. Η q_2 είναι τελική, αφού όταν φτάσουμε σε αυτήν έχουμε δει το aa.

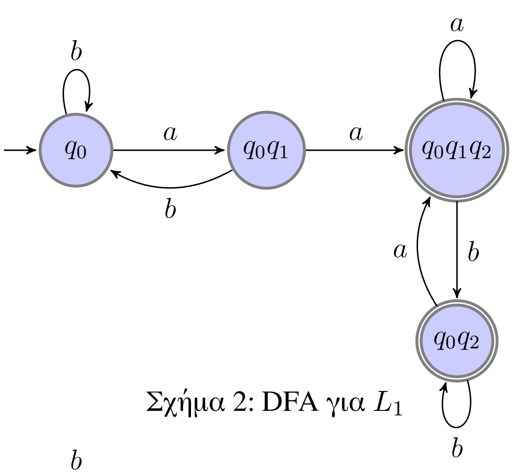


Σχήμα 1: NFA για L_1

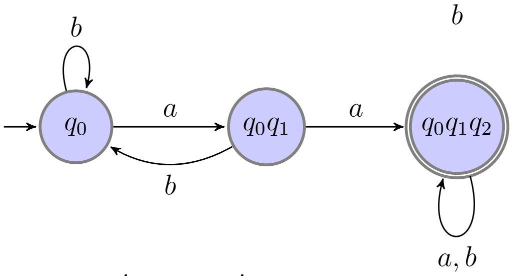
Μετά, πρέπει να προτυποποιήσουμε το γεγονός ότι το aa θα είναι υποσυμβολοσειρά της εισόδου, οπότε προσθέτουμε self-loops στις q_0 και q_2, ώστε να παραμένει το αυτόματο στις καταστάσεις αυτές, μέχρι και αφού βρει το aa (waiting loops). Παρατηρούμε ότι το NFA αποδέχεται αν *υπάρχει* η συμβολοσειρά aa μέσα στην συμβολοσειρά της εισόδου, όπως επίσης και αν η είσοδος είναι μόνο το aa.

Όπως φαίνεται άμεσα από το NFA, η κανονική έκφραση για την L_1 είναι η $(a+b)^*aa(a+b)^*$.

Για να κατασκευάσουμε το DFA για την L_1, θα μετατρέψουμε το NFA σε DFA σύμφωνα με τον αλγόριθμο της απόδειξης του θ. Rabin-Scott.



Σχήμα 2: DFA για L_1



Σχήμα 3: Ελάχιστο DFA για L_1

- Από την κατάσταση q_0 του NFA, με a μεταβαίνουμε στις καταστάσεις q_0 και q_1. Άρα στο DFA θα μεταβούμε στην q_0q_1. Με b μεταβαίνουμε πάλι στην q_0 (self-loop), οπότε θα κάνουμε και στο DFA το ίδιο.

- Από την κατάσταση q_0q_1 του DFA, με a η q_0 του NFA οδηγεί στις q_0 και q_1, και η q_1 του NFA οδηγεί στην q_2. Άρα, στο DFA, η q_0q_1 θα μεταβαίνει στην $q_0q_1q_2$. Με b η q_0 του NFA οδηγεί στον εαυτό της, και η q_1 του NFA δεν οδηγεί πουθενά. Άρα, στο DFA η q_0q_1 θα οδηγεί στην q_0.

- Από την κατάσταση $q_0q_1q_2$ του DFA, με a η q_0 οδηγεί στις q_0 και q_1 στο NFA, η q_1 οδηγεί στην q_2, και η q_2 στον εαυτό της. Άρα, στο DFA, η $q_0q_1q_2$ θα οδηγεί στον εαυτό της. Με b, η q_0 του NFA οδηγεί στον εαυτό της, η q_1 δεν οδηγεί πουθενά, και η q_2 οδηγεί πάλι στον εαυτό της. Άρα, στο DFA η $q_0q_1q_2$ θα οδηγεί στην q_0q_2.

- Από την κατάσταση q_0q_2 του DFA, με a η q_0 του NFA οδηγεί στις q_0 και q_1, και η q_2 στον εαυτό της, άρα η q_0q_2 στο DFA θα οδηγεί στην $q_0q_1q_2$. Με b, οι q_0 και q_2 του NFA οδηγούν στον εαυτό τους, άρα και η q_0q_2 του DFA θα οδηγεί στον εαυτό της.

Τελικές καταστάσεις στο DFA θα είναι αυτές που περιέχουν τουλάχιστον μία τελική του NFA. Άρα οι τελικές καταστάσεις θα είναι οι $q_0q_1q_2$ και q_0q_2, αφού μόνο η q_2 είναι τελική στο NFA. Το DFA φαίνεται στο σχήμα 2.

Παρατηρήστε ότι στο DFA που κατασκευάσαμε, οι (τελικές) καταστάσεις $q_0q_1q_2$ και q_0q_2 δεν έχουν ουσιαστική διαφορά για την λειτουργία του αυτομάτου, αφού αν βρεθούμε στην $q_0q_1q_2$, δεν υπάρχει τρόπος να μεταβούμε πίσω σε κατάσταση απόρριψης. Οπότε, μπορούμε να συγχωνεύσουμε τις δύο αυτές καταστάσεις σε μία. Προφανώς, αν ελαχιστοποιούσαμε αυτό το DFA χρησιμοποιώντας την γνωστή μέθοδο/αλγόριθμο, οι καταστάσεις $q_0q_1q_2$ και q_0q_2 θα ήταν μη-διακρίσιμες, άρα θα συγχωνεύονταν. Το τελικό DFA φαίνεται στο σχήμα 3. $\square$

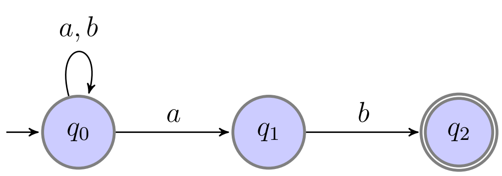
Προσοχή: Η μετατροπή του NFA σε DFA, στην χειρότερη περίπτωση, οδηγεί σε εκθετική αύξηση του πλήθους των καταστάσεων. Κατά την μετατροπή, δεν ασχολούμαστε με μη-προσβάσιμες καταστάσεις (*καταστάσεις στις οποίες δεν φτάνουμε ποτέ από την αρχική*), αλλά στην χειρότερη περίπτωση όλες οι καταστάσεις του δυναμοσυνόλου θα είναι προσβάσιμες, οπότε ένα NFA με n καταστάσεις μετατρέπεται σε DFA με το πολύ 2^n καταστάσεις.

Άσκηση 1.2

Κατασκευάστε NFA, DFA και κανονική έκφραση που να αναγνωρίζει την γλώσσα:

$$L_2 = \{x \in \{a,b\}^* \mid \text{η } x \text{ τελειώνει σε } 'ab'\}$$

Λύση

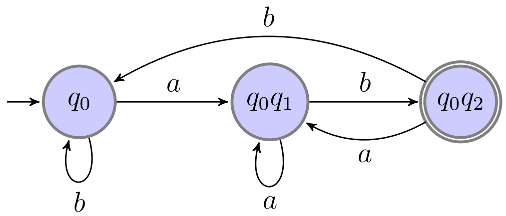


Σχήμα 4: NFA για L_2

Όπως και στο προηγούμενο παράδειγμα, θα κατασκευάσουμε τον "σκελετό" του NFA, δηλαδή τις καταστάσεις q_0, q_1, q_2, με τελική την q_2, που αποδέχονται το $'ab'$. Όμως, αυτό το NFA θέλουμε να αποδέχεται τις συμβολοσειρές που τελειώνουν σε ab, οπότε θα προσθέσουμε waiting loop μόνο στην q_0, ώστε το NFA μας να διαβάζει την συμβολοσειρά εισόδου και να περιμένει μέχρι να "δει" το ab στο τέλος της συμβολοσειράς.
Αν η συμβολοσειρά εισόδου τελειώνει σε ab, θα *υπάρχει* ένα μονοπάτι στο δέντρο υπολογισμού του NFA που να οδηγεί στην q_2. Αν δεν τελειώνει σε ab ή το περιέχει σε άλλο σημείο, *κανένα* μονοπάτι του δέντρου υπολογισμού δεν θα οδηγεί στην q_2.

Όπως φαίνεται από το NFA που κατασκευάσαμε (σχήμα 4), μια κανονική έκφραση για την γλώσσα L_2 είναι η $(a+b)^*ab$.

Θα κατασκευάσουμε DFA όπως στην προηγούμενη Άσκηση 1.1:



Σχήμα 5: DFA για L_2

- Από την κατάσταση q_0 το NFA με a μεταβαίνει στις q_0 και q_1. Άρα το DFA θα μεταβαίνει στην q_0q_1. Με b το NFA μεταβαίνει πάλι στην q_0 (self-loop), άρα το DFA θα κάνει το ίδιο.

- Από την κατάσταση q_0q_1 του DFA, με a η q_0 του NFA μεταβαίνει στις q_0, q_1, και η q_1 δεν μεταβαίνει πουθενά, άρα στο DFA η q_0q_1 θα μεταβαίνει στον εαυτό της. Με b, η q_0 του NFA μεταβαίνει στον εαυτό της, και η q_1 στην q_2, άρα στο DFA η q_0q_1 θα μεταβαίνει στην q_0q_2.

- Από την κατάσταση q_0q_2 του DFA, με a η q_0 του NFA μεταβαίνει στις q_0, q_1, και η q_2 δεν μεταβαίνει πουθενά, άρα στο DFA η q_0q_2 μεταβαίνει στην q_0q_1. Με b, η q_0 του NFA μεταβαίνει στον εαυτό της, και η q_2 δεν μεταβαίνει πουθενά, άρα στο DFA η q_0q_2 θα μεταβαίνει πίσω στην q_0.

Τελικές καταστάσεις στο DFA θα είναι αυτές που περιέχουν τουλάχιστον μία τελική του NFA, άρα η q_0q_2 θα είναι η μόνη τελική στο DFA, που φαίνεται στο σχήμα 5. □

Άσκηση 1.3

Κατασκευάστε DFA που να αποδέχεται την γλώσσα:

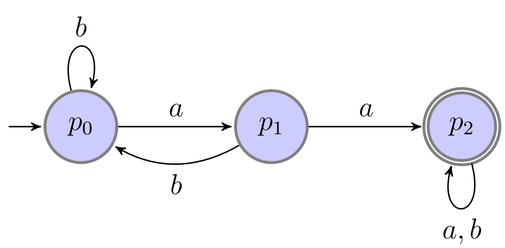
$$L_3 = \{x \in \{a,b\}^* \mid \text{η } x \text{ περιέχει το } 'aa' \text{ ή τελειώνει σε } 'ab'\}$$

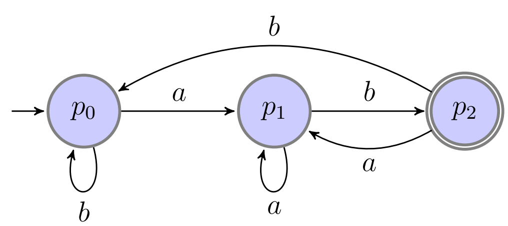
Λύση

Το ζητούμενο DFA αναγνωρίζει την ένωση των γλωσσών των Ασκήσεων 1.1 και 1.2, αφού μία συμβολοσειρά γίνεται αποδεκτή αν την αποδέχεται το αυτόματο του σχήματος 3 ή το αυτόματο του σχήματος 5. Υπάρχουν 2 τρόποι για να βρούμε αυτόματο για την ένωση δύο κανονικών γλωσσών.

1ος τρόπος επίλυσης

Θα συνδυάσουμε αυτά τα δύο αυτόματα κατασκευάζοντας το γινόμενό τους (*δείτε το ένθετο στην επόμενη σελίδα*), που θα πρέπει να αποδέχεται αν τουλάχιστον ένα από τα δύο αυτόματα αποδέχεται. Για ευκολία, θα μετονομάσουμε τις καταστάσεις των δύο αυτομάτων σε p_0, p_1 και p_2. Παρατηρήστε ότι δεν υπάρχει κίνδυνος σύγχυσης στις καταστάσεις, αφού στο γινόμενο κάθε "συντεταγμένη" του διατεταγμένου ζεύγους θα αναφέρεται σε ένα συγκεκριμένο αυτόματο. Θέλουμε, λοιπόν, να κατασκευάσουμε το γινόμενο των δύο ακόλουθων αυτομάτων (σχήματα 3, 5 με αλλαγμένα τα ονόματα των καταστάσεων):





Η αρχική κατάσταση του γινομένου θα είναι η p_0p_0 (*θυμηθείτε ότι γράφουμε ως p_ip_j το διατεταγμένο ζεύγος (p_i, p_j) του γινομένου*).

- Από την κατάσταση p_0p_0 του γινομένου, με a η p_0 του πρώτου DFA οδηγεί στην p_1 και η p_0 του δεύτερου στην p_1, άρα η p_0p_0 οδηγεί στην p_1p_1. Με b, η p_0 του πρώτου και του δεύτερου DFA έχουν self-loops, άρα στο γινόμενο η p_0p_0 θα έχει self-loop.

- Από την κατάσταση p_1p_1 του γινομένου, με a η p_1 του πρώτου DFA οδηγεί στην p_2 και η p_1 του δεύτερου στον εαυτό της, άρα η p_1p_1 οδηγεί στην p_2p_1. Με b, η p_1 του πρώτου οδηγεί στην p_0 και η p_1 του δεύτερου στην p_2, οπότε η p_1p_1 οδηγεί στην p_0p_2.

- Από την κατάσταση p_2p_1 του γινομένου, με a η p_2 του πρώτου DFA και η p_1 του δεύτερου οδηγούν στον εαυτό τους, οπότε η p_2p_1 θα έχει self-loop. Με b, η p_2 του πρώτου οδηγεί στον εαυτό της και η p_1 του δεύτερου στην p_2, οπότε η p_2p_1 οδηγεί στην p_2p_2.

- Από την κατάσταση p_0p_2 του γινομένου, με a η p_0 του πρώτου DFA οδηγεί στην p_1 και η p_2 του δεύτερου οδηγεί στην p_1, άρα η p_0p_2 οδηγεί στην p_1p_1. Με b, η p_0 του πρώτου έχει self-loop, και η p_2 του δεύτερου οδηγεί στην p_0, οπότε η p_0p_2 μεταβαίνει στην p_0p_0.

- Από την κατάσταση p_2p_2 του γινομένου, με a η p_2 του πρώτου DFA οδηγεί στον εαυτό της και η p_2 του δεύτερου στην p_1, άρα η p_2p_2 οδηγεί στην p_2p_1. Με b, η p_2 του πρώτου έχει self-loop και η p_2 του δεύτερου μεταβαίνει στην p_0, οπότε η p_2p_2 οδηγεί στην p_2p_0.

- Από την κατάσταση p_2p_0 του γινομένου, με a η p_2 του πρώτου DFA οδηγεί στον εαυτό της και η p_0 του δεύτερου στην p_1, άρα η p_2p_0 οδηγεί στην p_2p_1. Με b, η p_2 του πρώτου και η p_0 του δεύτερου έχουν self-loops, οπότε η p_2p_0 θα έχει και αυτή self-loop.

Τώρα, πρέπει να ορίσουμε το σύνολο των αποδεκτών καταστάσεων. Αφού θέλουμε το γινόμενο να αποδέχεται όταν αποδέχεται τουλάχιστον ένα από τα δύο DFAs, αποδεκτές καταστάσεις θα είναι αυτές που περιέχουν αποδεκτή κατάσταση σε μία τουλάχιστον από τις δύο συντεταγμένες του διατεταγμένου ζεύγους, άρα στην δική μας περίπτωση οι p_0p_2, p_2p_1, p_2p_0 και p_2p_2.

Γινόμενο Αυτομάτων

Μπορούμε να κατασκευάσουμε το γινόμενο δύο (ή περισσότερων) ντετερμινιστικών αυτομάτων, που θα προσομοιώνει παράλληλα την λειτουργία τους. Ουσιαστικά, το γινόμενο αυτομάτων είναι ένα DFA που “τρέχει” ταυτόχρονα τα DFAs του γινομένου, και αποδέχεται αν ισχύει μία λογική συνθήκη που εξαρτάται από την αποδοχή τους.

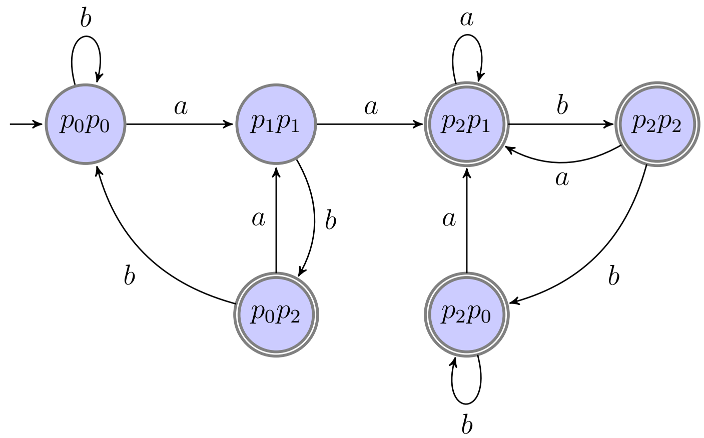
Τυπικά:
Έστω M_1, M_2 DFAs. Το γινόμενό τους $M = M_1 \times M_2$ είναι ένα DFA $M = (Q, \Sigma, \delta, q_0q_0, F)$:

- $Q = Q_1 \times Q_2$ (*γράφουμε* q_iq_j *την* (q_i, q_j)) *Διατεταγμένο ζεύγος*: $q_iq_j \in Q$ αν $q_i \in Q_1$ και $q_j \in Q_2$.
- Κοινό αλφάβητο Σ
- Αρχική κατάσταση q_0q_0
- $\delta(q_iq_j, a) = (\delta_1(q_i, a), \delta_2(q_j, a))$
- Σύνολο τελικών καταστάσεων $F \subseteq Q$.

Τι θα ορίσουμε ως σύνολο τελικών καταστάσεων; Εξαρτάται από το πότε θέλουμε να αποδέχεται το γινόμενο:

- Αν, πχ, θέλουμε το γινόμενο να αποδέχεται όταν αποδέχεται *και* το M_1 *και* το M_2 τότε ορίζουμε ως F:
$$F = \{q_iq_j \mid q_i \in F_1 \text{ και } q_j \in F_2\}$$
$$= F_1 \times F_2$$
- Αν θέλουμε το γινόμενο να αποδέχεται όταν αποδέχεται είτε το M_1 *ή* το M_2, ορίζουμε:
$$F = \{q_iq_j \mid q_i \in F_1 \text{ ή } q_j \in F_2\}$$
$$= (F_1 \times Q_2) \cup (Q_1 \times F_2)$$
- Αν θέλουμε να αποδέχεται όταν αποδέχεται το M_1 *και όχι* το M_2, ορίζουμε:
$$F = \{q_iq_j \mid q_i \in F_1 \text{ και } q_j \notin F_2\}$$
$$= F_1 \times (Q_2 \setminus F_2)$$

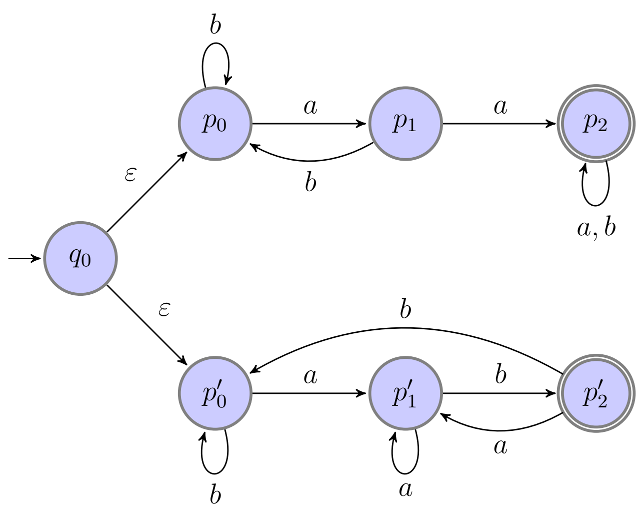
Το γινόμενο που κατασκευάσαμε φαίνεται στο παρακάτω σχήμα:



Σχήμα 6: Γινόμενο DFA για την ένωση των L_1 και L_2

Παρατηρήστε ότι αυτό το DFA *δεν* είναι ελάχιστο. Θα βρούμε το ελάχιστο στην επόμενη ενότητα.

2ος τρόπος επίλυσης
Αυτόματο DFA για την ένωση δύο γλωσσών για τις οποίες έχουμε ήδη DFA (ή NFA, ή και NFA$_\varepsilon$) μπορεί να προκύψει και με τον εξής τρόπο: δημιουργούμε μία νέα αρχική κατάσταση την οποία ενώνουμε με ε-κινήσεις με τις αρχικές καταστάσεις των δύο γνωστών FA. Έτσι, κατασκευάζουμε το εξής NFA$_\varepsilon$:



Στη συνέχεια μετατρέπουμε σε DFA με κάποιον από τους γνωστούς τρόπους (συστήνεται να προτιμάτε αυτόν της απευθείας μετατροπής σε DFA, σελ. 37 των διαφανειών). Η αρχική κατάσταση του DFA θα είναι η (q_0, p_0, p'_0), δηλαδή το ε-κλείσιμο της q_0. Στη συνέχεια, οι μεταβάσεις ορίζονται όπως στη μετατροπή από NFA σε DFA καθώς δεν υπάρχουν ε-κινήσεις μετά από τις υπόλοιπες καταστάσεις. $\square$

Προσοχή: Μερικές φορές η ένωση αυτομάτων μπορεί να γίνει απλά ταυτίζοντας τις αρχικές καταστάσεις τους (αυτό δίνει NFA, αν τα αρχικά αυτόματα είναι NFA) – έτσι μπορεί να πάρουμε ένα απλούστερο αυτόματο. Το αποτέλεσμα όμως δεν είναι πάντοτε σωστό: είναι δυνατόν η εκτέλεση σε ένα τέτοιο αυτόματο να περάσει αρχικά από κάποιες καταστάσεις του ενός αυτομάτου, να επιστρέψει στην κοινή αρχική και να συνεχίσει σε καταστάσεις του άλλου, γεγονός που μπορεί να έχει 'παρενέργειες'. Χρησιμοποιήστε αυτή την 'συντόμευση' με ιδιαίτερη προσοχή.

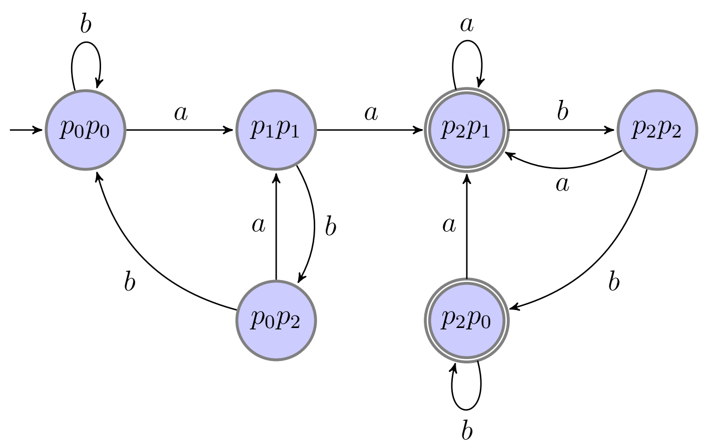
Άσκηση 1.4

Κατασκευάστε DFA που να αποδέχεται την γλώσσα:

$$L_4 = \{x \in \{a, b\}^* \mid \text{η } x \text{ περιέχει το } 'aa' \text{ και δεν τελειώνει σε } 'ab'\}$$

Λύση

Το ζητούμενο αυτόματο αναγνωρίζει την *διαφορά* των γλωσσών των Ασκήσεων 1.1 και 1.2, δηλαδή τις συμβολοσειρές που αποδέχεται το αυτόματο του σχήματος 3, αλλά απορρίπτει το αυτόματο του σχήματος 5, οπότε είναι και πάλι το γινόμενό τους, αλλά με τις αντίστοιχες τελικές καταστάσεις. Έχουμε βρει το γινόμενό τους στην προηγούμενη άσκηση (σχήμα 6). Οι τελικές καταστάσεις τώρα θα είναι αυτές που έχουν κατάσταση αποδοχής για το πρώτο αυτόματο (δηλαδή την p_2), αλλά κατάσταση απόρριψης για το δεύτερο (δηλαδή τις p_0, p_1), άρα θα είναι οι p_2p_0 και p_2p_1:



Σχήμα 7: Γινόμενο DFA για την διαφορά των L_1 και L_2

□

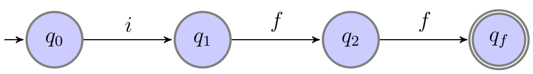
Άσκηση 1.5

Κατασκευάστε DFA και κανονική έκφραση για την γλώσσα:

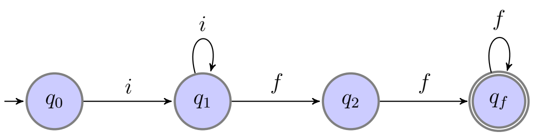
$$L_5 = \{w \in \{f, i\} \mid \text{η } w \text{ περιέχει το } 'iff' \text{ αλλά όχι το } 'fi'\}$$

Λύση

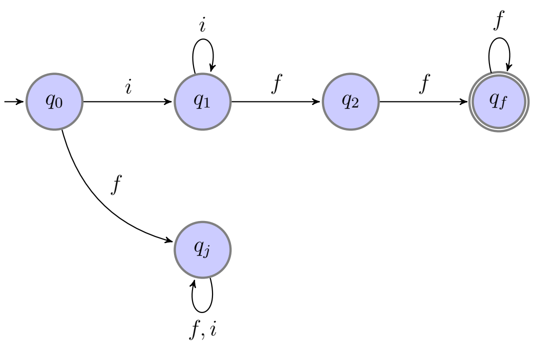
Για την κατασκευή του DFA θα μπορούσαμε να κατασκευάσουμε τα δύο απλά αυτόματα (ένα που να αποδέχεται το iff και ένα που να αποδέχεται το fi), και μετά να κατασκευάσουμε το γινόμενό τους, με αποδεκτές καταστάσεις αυτές στις οποίες αποδέχεται το πρώτο και $\delta\epsilon\nu$ αποδέχεται το δεύτερο. Εναλλακτικά, εδώ, θα προσπαθήσουμε να κατασκευάσουμε το DFA κατευθείαν:



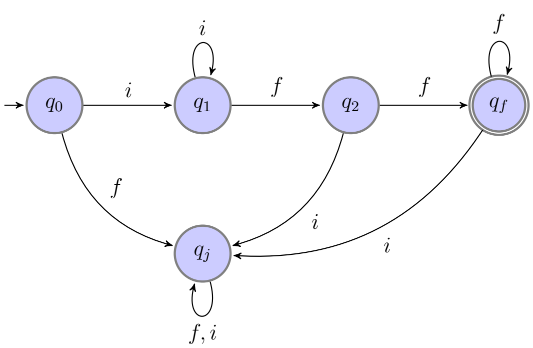
Σχήμα 8α: DFA για L_5



Σχήμα 8β: DFA για L_5



Σχήμα 8γ: DFA για L_5



Σχήμα 8δ: DFA για L_5

- Αρχικά, όπως φαίνεται στο σχήμα 8α, κατασκευάζουμε τον "σκελετό" του DFA, δηλαδή 4 καταστάσεις q_0, q_1, q_2, q_f που να αποδέχονται την συμβολοσειρά iff.

- Σίγουρα μπορούμε να αποδεχόμαστε συμβολοσειρές της μορφής ii^*fff^*, αφού δεν περιέχουν το fi, οπότε προσθέτουμε self-loops στις καταστάσεις q_1 και q_f, με τα σύμβολα i και f αντίστοιχα (σχήμα 8β). Προσοχή, πρέπει να υπάρχει τουλάχιστον ένα i στην αρχή της κάθε αποδεκτής συμβολοσειράς, και τουλάχιστον δύο f στο τέλος της, οπότε οι μεταβάσεις $q_0 \xrightarrow{i} q_1$ και $q_1 \xrightarrow{f} q_2 \xrightarrow{f} q_f$ δεν μπορούν να συγχωνευτούν.

- Τώρα, πρέπει να προσέξουμε την $\mu\eta$ αποδοχή του fi. Σίγουρα, αν το DFA βρίσκεται στην q_0 και διαβάσει σύμβολο f, πρέπει να απορρίψει, μιας και συμβολοσειρές που αρχίζουν από f δεν έχουν ελπίδα να γίνουν αποδεκτές, αφού ο μόνος τρόπος να υπάρξει το iff στην συμβολοσειρά είναι μετά από κάποιο f, άρα θα εμφανιστεί το fi. Οπότε, από την q_0 με f πρέπει να μεταβούμε σε κατάσταση απόρριψης (junk state, q_j), όπως φαίνεται στο σχήμα 8γ.

- Το ίδιο ισχύει και για τις καταστάσεις q_2, q_f. Αν το επόμενο σύμβολο είναι το i, θα προηγείται ένα f (αφού με f μεταβαίνει στις q_2, q_f), οπότε θα εμφανιστεί το fi, και το DFA θα πρέπει να απορρίπτει. Άρα, από τις q_2, q_f με i μεταβαίνει στην q_j (σχήμα 8δ).

- Παρατηρούμε ότι έχουμε συμπληρώσει όλες τις δυνατές μεταβάσεις, οπότε η κατασκευή του DFA ολοκληρώθηκε. Το τελικό DFA φαίνεται στο σχήμα 8δ.

Πως μπορούμε να βρούμε μια κανονική έκφραση για την γλώσσα L_5; *Πρέπει να εντοπίσουμε όλα τα δυνατά μονοπάτια του DFA που οδηγούν σε κατάσταση αποδοχής*: στο αυτόματο του σχήματος 8δ, το μόνο μονοπάτι είναι το $q_0 \to q_1 \to q_2 \to q_f$, οπότε μια κανονική έκφραση για την L_5 είναι η ii^*fff^*, ή ισοδύναμα i^+ff^+.

Προσπαθήστε να δείτε ότι το DFA που βρήκαμε είναι το ελάχιστο DFA που αναγνωρίζει την γλώσσα L_5. Θα το αποδείξουμε στην επόμενη ενότητα. □

2 Ελαχιστοποίηση Αυτομάτων

Άσκηση 2.1

Ελαχιστοποιήστε το αυτόματο του σχήματος 6.

Λύση

Κατασκευάζουμε τον πίνακα των καταστάσεων (για ευκολία, κυκλώνουμε τις τελικές καταστάσεις). Παρατηρήστε ότι ασχολούμαστε μόνο με το μέρος του πίνακα κάτω από την κύρια διαγώνιο, αφού η σχέση της διακρισιμότητας των καταστάσεων είναι συμμετρική και ανακλαστική, οπότε δεν έχει νόημα να συγκρίνουμε ξανά δύο καταστάσεις ή μια κατάσταση με τον εαυτό της.

p_1p_1					
p_0p_2	X_0	X_0			
p_2p_1	X_0	X_0			
p_2p_0	X_0	X_0			
p_2p_2	X_0	X_0			
	p_0p_0	p_1p_1	p_0p_2	p_2p_1	p_2p_0

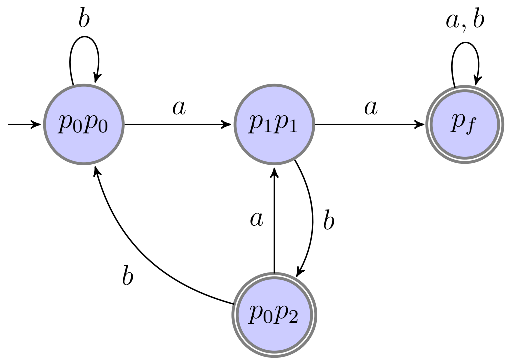
Αρχικά, βάζουμε X_0 στις καταστάσεις που διακρίνονται επειδή η μία είναι τελική και η άλλη δεν είναι.

Στην επόμενη φάση, βάζουμε X_1 στις καταστάσεις που με ένα σύμβολο (a ή b) οδηγούν σε διακρίσιμες καταστάσεις, δηλαδή σε ένα ζεύγος καταστάσεων που ήδη έχουμε διακρίνει με X_0:

- Έχουμε ότι $p_0p_0 \xrightarrow{a} p_1p_1$ και $p_1p_1 \xrightarrow{a} p_2p_1$, και οι p_1p_1 και p_2p_1 είναι διακρίσιμες (αφού η μία είναι τελική και η άλλη δεν είναι), οπότε τις διακρίνουμε με X_1.
- Ομοίως, $p_0p_2 \xrightarrow{b} p_0p_0$ και $p_2p_2 \xrightarrow{b} p_2p_0$, άρα p_0p_2 και p_2p_2 διακρίσιμες.
- $p_0p_2 \xrightarrow{a} p_1p_1$ και $p_2p_0 \xrightarrow{a} p_2p_1$, άρα p_0p_2 και p_2p_0 διακρίσιμες.
- $p_0p_2 \xrightarrow{b} p_0p_0$ και $p_2p_1 \xrightarrow{b} p_2p_2$, άρα p_0p_2 και p_2p_1 διακρίσιμες.

p_1p_1	X_1				
p_0p_2	X_0	X_0			
p_2p_1	X_0	X_0	X_1		
p_2p_0	X_0	X_0	X_1		
p_2p_2	X_0	X_0	X_1		
	p_0p_0	p_1p_1	p_0p_2	p_2p_1	p_2p_0

Δεν μπορούμε να διακρίνουμε άλλες καταστάσεις με κάποιο σύμβολο, οπότε καταλήγουμε στον παραπάνω πίνακα, όπου οι καταστάσεις p_2p_1, p_2p_2 και p_2p_0 συγχωνεύονται σε μία, έστω p_f. Το ελάχιστο αυτόματο φαίνεται παρακάτω:



Σχήμα 9: Ελάχιστο DFA για την ένωση των L_1 και L_2

□

Άσκηση 2.2

Ελαχιστοποιήστε το αυτόματο του σχήματος 7.

Λύση

Όπως και στην προηγούμενη άσκηση, κατασκευάζουμε τον πίνακα των καταστάσεων.
Αρχικά, βάζουμε X_0 στις καταστάσεις που διακρίνονται επειδή η μία είναι τελική και η άλλη δεν είναι.
Στην επόμενη φάση, βάζουμε X_1 στις καταστάσεις που με ένα σύμβολο (a ή b) οδηγούν σε διακρίσιμες καταστάσεις, δηλαδή σε ένα ζεύγος καταστάσεων που ήδη έχουμε διακρίνει με X_0:

- Έχουμε ότι $p_0p_0 \xrightarrow{a} p_1p_1$ και $p_1p_1 \xrightarrow{a} p_2p_1$, και οι p_1p_1 και p_2p_1 είναι διακρίσιμες (αφού η μία είναι τελική και η άλλη δεν είναι), οπότε τις διακρίνουμε με X_1.

- Ομοίως, $p_0p_2 \xrightarrow{a} p_1p_1$ και $p_1p_1 \xrightarrow{a} p_2p_1$, άρα p_0p_2 και p_1p_1 διακρίσιμες.

- $p_0p_0 \xrightarrow{a} p_1p_1$ και $p_2p_2 \xrightarrow{a} p_2p_1$, άρα p_0p_0 και p_2p_2 διακρίσιμες.

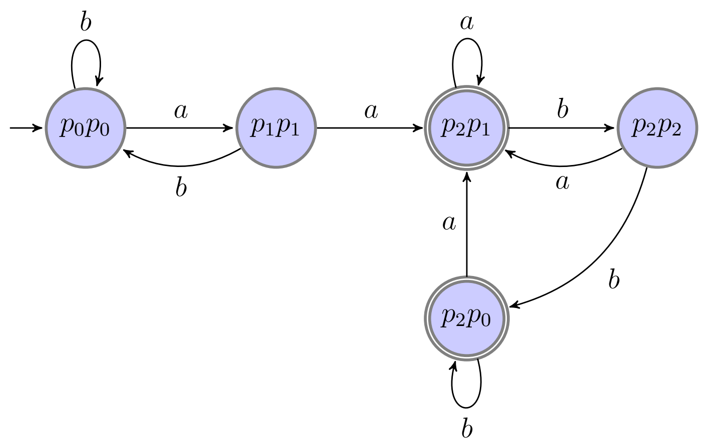
- $p_2p_2 \xrightarrow{b} p_2p_0$ και $p_1p_1 \xrightarrow{b} p_0p_2$, άρα p_2p_2 και p_1p_1 διακρίσιμες.

	p_0p_0	p_1p_1	p_0p_2	p_2p_1	p_2p_0
p_1p_1					
p_0p_2					
p_2p_1	X_0	X_0	X_0		
p_2p_0	X_0	X_0	X_0		
p_2p_2				X_0	X_0

	p_0p_0	p_1p_1	p_0p_2	p_2p_1	p_2p_0
p_1p_1	X_1				
p_0p_2		X_1			
p_2p_1	X_0	X_0	X_0		
p_2p_0	X_0	X_0	X_0	X_1	
p_2p_2	X_1	X_1	X_1	X_0	X_0

- $p_2p_2 \xrightarrow{a} p_2p_1$ και $p_0p_2 \xrightarrow{a} p_1p_1$, άρα p_2p_2 και p_0p_2 διακρίσιμες.
- $p_2p_1 \xrightarrow{b} p_2p_2$ και $p_2p_0 \xrightarrow{b} p_2p_0$, άρα p_2p_1 και p_2p_0 διακρίσιμες.

Παρατηρούμε ότι δεν μπορούμε να διακρίνουμε τις p_0p_2 και p_0p_0, αφού με a οδηγούν στην ίδια κατάσταση, και με b η p_0p_0 οδηγεί στον εαυτό της και η p_0p_2 στην p_0p_0, οπότε οι καταστάσεις αυτές συγχωνεύονται. Το ελάχιστο DFA φαίνεται στο παρακάτω σχήμα:



Σχήμα 10: Ελάχιστο DFA για την διαφορά των L_1 και L_2

□

Παρατηρήστε ότι παρόλο που στις δύο προηγούμενες ασκήσεις ελαχιστοποιήσαμε το γινόμενο των αυτομάτων με μόνη διαφορά τις τελικές καταστάσεις, τα ελάχιστα DFAs είναι πολύ διαφορετικά μεταξύ τους.

Άσκηση 2.3

Ελαχιστοποιήστε το αυτόματο του σχήματος 8δ.

Λύση

Κατασκευάζουμε τον πίνακα των καταστάσεων του αυτομάτου.
Αρχικά, βάζουμε X_0 στις καταστάσεις που διακρίνονται επειδή η μία είναι τελική και η άλλη δεν είναι.
Στην επόμενη φάση, βάζουμε X_1 στις καταστάσεις που με ένα σύμβολο (i ή f) οδηγούν σε διακρίσιμες καταστάσεις, δηλαδή σε ένα ζεύγος καταστάσεων που ήδη έχουμε διακρίνει με X_0:

q_1				
q_2				
q_f	X_0	X_0	X_0	
q_j				X_0
	q_0	q_1	q_2	q_f

- Έχουμε ότι $q_2 \xrightarrow{f} q_f$ και $q_j \xrightarrow{f} q_j$, και οι q_f και q_j είναι διακρίσιμες (αφού η μία είναι τελική και η άλλη δεν είναι), οπότε τις διακρίνουμε με X_1.

- Ομοίως, $q_2 \xrightarrow{f} q_f$ και $q_1 \xrightarrow{f} q_2$, άρα q_1 και q_2 διακρίσιμες.

- $q_0 \xrightarrow{f} q_j$ και $q_2 \xrightarrow{f} q_f$, άρα q_0 και q_2 διακρίσιμες.

  - Έχουμε ότι $q_0 \xrightarrow{f} q_j$ και $q_1 \xrightarrow{f} q_2$, και έχουμε ήδη διακρίνει τις q_2 και q_j με X_1, άρα διακρίνουμε τις q_0 και q_1 με X_2.

  - Ομοίως, $q_1 \xrightarrow{f} q_2$ και $q_j \xrightarrow{f} q_j$, και έχουμε ήδη διακρίνει τις q_2 και q_j με X_1, άρα διακρίνουμε τις q_1 και q_j με X_2.

  - Τέλος, $q_0 \xrightarrow{i} q_1$ και $q_j \xrightarrow{i} q_j$, και έχουμε ήδη διακρίνει τις q_1 και q_j με X_2, άρα διακρίνουμε τις q_0 και q_j με X_3.

q_1	X_2			
q_2	X_1	X_1		
q_f	X_0	X_0	X_0	
q_j	X_3	X_2	X_1	X_0
	q_0	q_1	q_2	q_f

Παρατηρούμε ότι όλος ο πίνακας έχει συμπληρωθεί, οπότε το αυτόματο του σχήματος 8δ είναι ελάχιστο. □

3 Pumping Lemma

Υπενθύμιση: Pumping Lemma
Αν η L είναι κανονική, τότε:
$\exists n \in \mathbb{N}$, τέτοιο ώστε:
$\forall z \in L, |z| \geq n$:
$\exists u, v, w \in \Sigma^*$, τέτοια ώστε:
$z = uvw$ και $|uv| \leq n$ και $|v| \geq 1$ και $\forall i \in \mathbb{N} : uv^iw \in L$.

Άσκηση 3.1

Δείξτε ότι η γλώσσα $L_1 = \{0^{k^2} \mid k > 0\}$ δεν είναι κανονική.

Λύση
Έστω, προς απαγωγή σε άτοπο, πως η L_1 είναι κανονική, άρα ισχύει το pumping lemma, που σημαίνει ότι υπάρχει ένα $n \in \mathbb{N}$ για το οποίο για κάθε $z \in L_1, |z| \geq n$, να υπάρχουν $u, v, w \in \Sigma^*$ ("σπάσιμο" της z) τέτοια ώστε: $z = uvw, |uv| \leq n, |v| \geq 1$ και για κάθε $i \in \mathbb{N} : uv^iw \in L_1$.

Έστω λοιπόν το $n \in \mathbb{N}$. Επιλέγουμε $z = 0^{n^2}$, που σίγουρα υπάρχει, αφού η L_1 είναι άπειρη γλώσσα, και $|z| = n^2 \geq n$. Το pumping lemma μας βεβαιώνει ότι υπάρχουν $u, v, w \in \Sigma^*$, τέτοια ώστε $z = 0^{n^2} = uvw$, και αφού $|v| \geq 1$, $v = 0^m$, για $1 \leq m \leq n$.

Αφού έχουμε για κάθε $i \in \mathbb{N} : uv^iw \in L_1$, επιλέγουμε $i = 2$: Τότε $uv^2w \in L$, και $|uv^2w| = |uvvw| = n^2 + m$, αλλά το $n^2 + m$ *δεν* είναι τέλειο τετράγωνο, αφού $n^2 < n^2 + m \leq n^2 + n = n(n+1) < (n+1)^2$, άρα καταλήξαμε σε άτοπο, ενώ είχαμε υποθέσει ότι η L_1 είναι κανονική, άρα η L_1 *δεν* είναι κανονική. □

Μεθοδολογία

- Θέλουμε να αποδείξουμε ότι η γλώσσα L *δεν* είναι κανονική.
- Υποθέτουμε ότι είναι κανονική, για να καταλήξουμε σε άτοπο, και εφαρμόζουμε το pumping lemma (PL).
- Το PL μας βεβαιώνει ότι υπάρχει ένα $n \in \mathbb{N}$ για το οποίο να ισχύει το συμπέρασμά του.
- Επιλέγουμε $z \in L$, με $|z| \geq n$.
- Το PL μας βεβαιώνει ότι υπάρχει "σπάσιμο" της z στα u, v, w, που όμως ικανοποιούν τα $|uv| \leq n, |v| \geq 1$.
- Αφού $\forall i \in \mathbb{N} : uv^iw \in L$, διαλέγουμε ένα i για το οποίο η συμβολοσειρά uv^iw να μην ανήκει στην L, και καταλήγουμε σε άτοπο.
- Το άτοπο του προηγούμενου βήματος μας οδηγεί στο συμπέρασμα ότι η L *δεν* είναι κανονική.

Άσκηση 3.2

Δείξτε ότι η γλώσσα $L_2 = \{0^p \mid p \text{ πρώτος}\}$ δεν είναι κανονική.

Λύση
Έστω, προς απαγωγή σε άτοπο, πως η L_2 είναι κανονική. Έστω $z = 0^p \in L_2$, για κάποιον πρώτο p, με $|z| \geq n$ (σίγουρα υπάρχει τέτοιο p, αφού οι πρώτοι αριθμοί είναι άπειροι σε πλήθος). Τότε, σύμφωνα με το pumping lemma, θα υπάρχουν u, v, w τέτοια ώστε $z = 0^p = uvw$, $|uv| \leq n$, $|v| \geq 1$, και $uv^iw \in L_2$, για κάθε $i \in \mathbb{N}$.
Έστω $|u| + |w| = k$ και $|v| = \ell \geq 1$. Τότε, αφού $\forall i \in \mathbb{N} :\ uv^iw \in L_2$, τότε και $0^{k+i\ell} \in L_2$ για κάθε $i \in \mathbb{N}$. Για $i = 0$: $0^k \in L_2 \Rightarrow k$ πρώτος. Αν θέσουμε $i = k$, τότε $0^{k+k\ell} = 0^{k(1+\ell)} \in L_2$, άρα $k(1+\ell)$ πρώτος. Όμως $\ell \geq 1 \Rightarrow 1 + \ell \geq 2$, άρα το $1 + \ell$ είναι μη-τετριμμένος παράγοντας του $k(1+\ell)$, άρα δεν είναι πρώτος. Άτοπο. □

4 Τυπικές Γραμματικές

Άσκηση 4.1

Βρείτε γραμματική που να αναγνωρίζει την γλώσσα του αυτομάτου του σχήματος 8δ.

Λύση

Μπορούμε να κατασκευάσουμε δεξιογραμμική κανονική γραμματική, ακολουθώντας τις μεταβάσεις του αυτομάτου. Ονομάζουμε με μη-τερματικά σύμβολα τις καταστάσεις του DFA: $S \equiv q_0$, $A \equiv q_1$, $B \equiv q_2$, $C \equiv q_f$, $D \equiv q_j$. Οι κανόνες θα είναι οι εξής:

$S \to iA|fD$
$A \to iA|fB$
$B \to iD|fC$
$C \to fC|iD|\varepsilon$ (το ε προστίθεται γιατί η q_f είναι τελική κατάσταση.)
$D \to iD|fD$

Μπορούμε να έχουμε γραμματική με λιγότερους κανόνες; Αρχικά, παρατηρούμε ότι αφού το μη-τερματικό D αντιπροσωπεύει την junk state, οπότε το D, όπως φαίνεται και από τους κανόνες, δεν πρόκειται ποτέ να καταλήξει σε τελική συμβολοσειρά (δηλαδή μόνο με τερματικά σύμβολα). Άρα, οι κανόνες που οδηγούν στην D μπορούν να παραλειφθούν. Επίσης, μπορούμε να συγχωνεύσουμε τους κανόνες του A και του B ως εξής:

$$S \to iS|iffC \qquad C \to fC|\varepsilon$$

Η τελική γραμματική θα είναι η $G = (\{S, C\}, \{i, f\}, \{S \to iS|iffC, C \to fC|\varepsilon\}, S)$. □

Άσκηση 4.2

Βρείτε γραμματική που να αναγνωρίζει την γλώσσα:

$$L_2 = \{x \in \{0,1\}^* \mid x = x^R\}$$

(*όπου* x^R *η ανάστροφη συμβολοσειρά, π.χ. αν* $x =$ abc, *τότε* $x^R =$ cba)

Λύση

Για να βρούμε μια τέτοια γραμματική, πρέπει να σκεφτούμε αναδρομικά: μια συμβολοσειρά μήκους k, $k \geq 3$, θα ανήκει στην γλώσσα αν μπορεί να γραφεί ως $0y0$ ή $1y1$, για κάποιο $y \in L_2$ μήκους $k-2$. Άρα, χρειαζόμαστε τους κανόνες $S \to 0S0$ και $S \to 1S1$ για να παράγουμε τις συμβολοσειρές της γλώσσας αναδρομικά. Επίσης, προσθέτουμε τον κανόνα $S \to \varepsilon$ για να καταλήγουμε σε τελική συμβολοσειρά, αλλά τώρα η γλώσσα μας περιορίζεται στις συμβολοσειρές άρτιου μήκους. Για να συμπεριλάβουμε και αυτές με περιττό μήκος, προσθέτουμε τους κανόνες $S \to 0$, $S \to 1$. Η γραμματική θα είναι η:

$$G = (\{S\}, \{0,1\}, \{S \to \varepsilon|0|1|0S0|1S1\}, S)$$

□

Άσκηση 4.3

Έστω η γραμματική $G = (\{S\}, \{0,1\}, P, S)$ με σύνολο κανόνων P:

$$S \to 0S11|\varepsilon$$

Ποια γλώσσα αναγνωρίζει η G;

Λύση

Παρατηρούμε ότι εφαρμόζοντας συνεχώς τον κανόνα παραγωγής:

$$S \to 0S11 \to 00S1111 \to 000S111111 \to \cdots \to 0^n S 1^{2n} \to 0^n \varepsilon 1^{2n} \to 0^n 1^{2n}$$

Επίσης υπάρχει και ο κανόνας $S \to \varepsilon$, άρα $\varepsilon \in L(G)$. Η γλώσσα που αναγνωρίζεται είναι η:

$$L(G) = \{0^n 1^{2n} \mid n \geq 0\}$$

□

5 Αυτόματα για μη-κανονικές γλώσσες

Άσκηση 5.1

Περιγράψτε αυτόματο LBA για την context-sensitive γλώσσα $L = \{a^n b^n a^n \mid n \in \mathbb{N}\}$.

Λύση

Θα περιγράψουμε ένα 2-PDA (PDA με 2 στοίβες) που δεν χρησιμοποιεί περισσότερα σύμβολα στις 2 στοίβες συνολικά από όσα υπάρχουν στην είσοδο.

> Έχει αποδειχθεί ότι μία μηχανή Turing μπορεί να προσομοιωθεί πλήρως από ένα αυτόματο PDA με δύο στοίβες (μπορείτε να σκεφτείτε πώς;). Επομένως, αν ο χώρος που θα χρησιμοποιηθεί είναι το πολύ μια σταθερά επί το μέγεθος της εισόδου τότε το 2-PDA ουσιαστικά είναι ισοδύναμο με LBA.

Η περιγραφή του 2-PDA σε γλώσσα υψηλού επιπέδου είναι η εξής:

- Όσο διαβάζεις a βάζε (push) 0 στην 1η στοίβα.
- Όσο διαβάζεις b βάζε (push) 1 στην 2η στοίβα.
- Όσο διαβάζεις a βγάζε (pop) ένα 0 από την 1η στοίβα και ένα 1 από τη 2η στοίβα (αν κάποια από τις 2 στοίβες είναι κενή προτού βγάλεις το αντίστοιχο 0 ή 1 απόρριψε).
- Αν η είσοδος τελείωσε και οι 2 στοίβες είναι κενές αποδέξου.

Είναι εύκολο να αποδειχθεί ότι το αυτόματο αποδέχεται αν και μόνο αν η συμβολοσειρά εισόδου ανήκει στη γλώσσα L.

Σημείωση: προφανώς μπορούν να χρησιμοποιηθούν και διαφορετικά σύμβολα στις στοίβες, π.χ. a και b αντίστοιχα. □