4-44: Θεωρία Υπολογισμού

Λύσεις της 3ης Ομάδας Προτεινομένων Ασκήσεων

Άσκηση 1

Κατασκευάστε πεπερασμένο αυτόματο που αποδέχεται τις συμβολοσειρές από το αλφάβητο $\Sigma = \{a, b\}$ που

- ▷ έχουν μήκος το πολύ 2 ή
- ▷ έχουν μήκος τουλάχιστον 3 και κάθε τριάδα διαδοχικών συμβόλων της συμβολοσειράς περιέχει τουλάχιστον δύο a.

Λύση

Η δυσκολία της άσκησης έγκειται στον έλεγχο και την αποδοχή όσων από τις συμβολοσειρές μήκους τουλάχιστον 3 θα πρέπει να γίνονται αποδεκτές. Ο έλεγχος αυτών των συμβολοσειρών μπορεί να γίνει αποτελεσματικά ως εξής: για κάθε σύμβολο της συμβολοσειράς που διαβάζεται, ελέγχεται ότι στα τελευταία τρία σύμβολα που έχουν διαβαστεί περιέχονται τουλάχιστον δύο a και εάν όχι τότε το αυτόματο οδηγείται σε κατάσταση αδιεξόδου $[X]$. Από αυτό φαίνεται ότι για αυτές τις περιπτώσεις αρκεί το αυτόματο να "θυμάται" τα τελευταία δύο σύμβολα του τμήματος της συμβολοσειράς που έχει δει: για να το επιτύχουμε αυτό θα χρησιμοποιήσουμε καταστάσεις που θα συμβολίζονται ως $[..\sigma_1\sigma_2]$ όπου $\sigma_1, \sigma_2 \in \{a, b\}$ και μεταβάσεις όπως $\delta([..aa], a) = [..aa]$, $\delta([..aa], b) = [..ab]$, κ.λπ.

Για τις συμβολοσειρές μήκους το πολύ 2, εύκολα μπορούμε να φροντίσουμε ώστε αυτές να γίνουν αποδεκτές χρησιμοποιώντας καταστάσεις $[e]$, $[a]$, $[b]$, $[aa]$, $[ab]$, $[ba]$ και $[bb]$, οι οποίες δηλώνουν ότι η είσοδος που έχει διαβαστεί είναι e, a, …, bb αντίστοιχα, και είναι τελικές.

Με βάση τα παραπάνω, έχουμε το εξής αυτόματο: $M = \Big(K,\, \Sigma,\, \delta,\, q_0,\, F\Big)$

όπου

$$\begin{aligned}
\Sigma &= \{a, b\} \\
K &= \{[e],\, [a],\, [b],\, [aa],\, [ab],\, [ba],\, [bb],\, [..aa],\, [..ab],\, [..ba],\, [X]\} \\
q_0 &= [e] \\
F &= \{[e],\, [a],\, [b],\, [aa],\, [ab],\, [ba],\, [bb],\, [..aa],\, [..ab],\, [..ba]\}
\end{aligned}$$

δ :

	a	b
$[e]$	$[a]$	$[b]$
$[a]$	$[aa]$	$[ab]$
$[b]$	$[ba]$	$[bb]$
$[aa]$	$[..aa]$	$[..ab]$
$[ab]$	$[..ba]$	$[X]$
$[ba]$	$[..aa]$	$[X]$
$[bb]$	$[X]$	$[X]$
$[..aa]$	$[..aa]$	$[..ab]$
$[..ab]$	$[..ba]$	$[X]$
$[..ba]$	$[..aa]$	$[X]$
$[X]$	$[X]$	$[X]$

(Παρατηρήστε ότι κάθε φορά που από μια κατάσταση διαφορετική της $[X]$ το αυτόματο μεταβαίνει στην $[X]$, τα τελευταία 3 σύμβολα που έχει διαβάσει το αυτόματο περιέχουν το πολύ 1 a.)

Άσκηση 2

Δοθέντων δύο πεπερασμένων αυτομάτων M_1 και M_2, περιγράψτε πεπερασμένο αυτόματο M το οποίο αποδέχεται τη γλώσσα $L(M_1) \cap L(M_2)$.

Λύση

Έστω $M_1' = (K_1, \Sigma, \delta_1, s_1, F_1)$ και $M_2' = (K_2, \Sigma, \delta_2, s_2, F_2)$ **ντετερμινιστικά** πεπερασμένα αυτόματα ισοδύναμα προς τα δοθέντα αυτόματα M_1 και M_2 αντίστοιχα. (Για λόγους απλότητας της περιγραφής, υποθέτουμε ότι τα αλφάβητα εισόδου των δύο αυτομάτων ταυτίζονται. Εάν αυτό δεν ισχύει, τότε επαυξάνουμε το διάγραμμα καταστάσεων του αυτόματου M_1' ως εξής: συμπεριλαμβάνουμε μία κατάσταση αδιεξόδου και προσθέτουμε για καθένα από τα σύμβολα του αλφαβήτου του M_2' τα οποία δεν ανήκουν στο αλφάβητο του M_1' μία μετάβαση από καθεμία από τις υπόλοιπες κατάστασεις του M_1' προς την κατάσταση αδιεξόδου. Παρατηρήστε ότι το αυτόματο που προκύπτει είναι ντετερμινιστικό, έχει ως αλφάβητο την ένωση των αλφαβήτων των M_1' και M_2', και είναι ισοδύναμο προς το M_1'. Αντίστοιχα επαυξάνουμε και το διάγραμμα καταστάσεων του αυτόματου M_2'.)

Το αυτόματο M που ζητάμε θα πρέπει να αποδέχεται ακριβώς εκείνες τις συμβολοσειρές οι οποίες γίνονται αποδεκτές τόσο από το M_1 (άρα και από το M_1') όσο και από το M_2 (άρα και από το M_2'). Για να το εξασφαλίσουμε αυτό, προσομοιώνουμε την "εν παραλλήλω" λειτουργία των αυτομάτων M_1' και M_2': συγκεκριμένα, για κάθε σύμβολο της συμβολοσειράς εισόδου με τη σειρά, εκτελούμε την αντίστοιχη μετάβαση για το M_1' και για το M_2'. Αυτό επιτυγχάνεται φροντίζοντας ώστε οι καταστάσεις του ζητούμενου αυτόματου M να είναι ζεύγη της μορφής $\langle q, q' \rangle$, όπου $q \in K_1$ και $q' \in K_2$. Τότε, η "εν παραλλήλω" προσομοίωση ενός βήματος της λειτουργίας των M_1' και M_2' για το σύμβολο $\sigma \in \Sigma$, όταν αυτά βρίσκονται στις καταστάσεις q και q' αντίστοιχα, επιτυγχάνεται με τη μετάβαση:

$$\delta(\langle q, q' \rangle,\, \sigma) \;=\; \langle \delta_1(q, \sigma),\, \delta_2(q', \sigma) \rangle.$$

Από τα παραπάνω, φαίνεται ότι η αρχική κατάσταση του αυτόματου M θα είναι η $\langle s_1, s_2 \rangle$, ενώ το σύνολο τελικών καταστάσεων θα είναι

$$F = \{\, \langle q, q' \rangle \mid q \in F_1 \text{ και } q' \in F_2 \,\}$$

ώστε να γίνονται αποδεκτές ακριβώς οι συμβολοσειρές που γίνονται αποδεκτές τόσο από το M_1' όσο και από το M_2'.

Συνοψίζοντας, το ζητούμενο αυτόματο M θα είναι ως εξής: $M = \Big(K,\, \Sigma,\, \delta,\, s,\, F \Big)$

$$\begin{array}{llcl}
\text{όπου} & K & = & \{\, \langle q, q' \rangle \mid q \in K_1 \text{ και } q' \in K_2 \,\} \\
 & \Sigma & : & \text{το (κοινό) αλφάβητο των αυτομάτων } M_1' \text{ και } M_2' \\
 & \delta & : & \forall q \in K_1, \forall q' \in K_2, \forall \sigma \in \Sigma, \quad \delta(\langle q, q' \rangle,\, \sigma) = \langle \delta_1(q, \sigma),\, \delta_2(q', \sigma) \rangle \\
 & s & = & \langle s_1, s_2 \rangle \\
 & F & = & \{\, \langle q, q' \rangle \mid q \in F_1 \text{ και } q' \in F_2 \,\}.
\end{array}$$

Άσκηση 3

Γράψτε μια κανονική έκφραση για τη γλώσσα που περιέχει όλες τις συμβολοσειρές από το αλφάβητο $\Sigma = \{a, b\}$ που έχουν τουλάχιστον δύο συνεχόμενα a και το πολύ μία εμφάνιση δύο συνεχόμενων b.

Λύση

Διακρίνουμε τις εξής περιπτώσεις.

(α) *Συμβολοσειρές που περιέχουν τουλάχιστον ένα aa και δεν περιέχουν bb*: Τότε, κάθε b, εάν δεν βρίσκεται στο τέλος της συμβολοσειράς, θα πρέπει να ακολουθείται από a. Επίσης, το τμήμα της συμβολοσειράς πιο αριστερά από την αριστερότερη εμφάνιση του aa θα πρέπει να έχει εναλλάξ a και b. Συνεπώς, για αυτές τις συμβολοσειρές έχουμε την κανονική έκφραση

$$r_1 \;=\; (\varnothing^* \cup \mathrm{b})\,(\mathrm{a\,b})^*\,\mathrm{a\,a}\,(\mathrm{a} \cup \mathrm{b\,a})^*\,(\varnothing^* \cup \mathrm{b})$$

ή επίσης

$$r_1 \;=\; (\varnothing^* \cup \mathrm{b})\,(\mathrm{a\,b})^*\,\mathrm{a\,a\,a}^*\,(\mathrm{b\,a\,a}^*)^*\,(\varnothing^* \cup \mathrm{b}).$$

(β) *Συμβολοσειρές που περιέχουν τουλάχιστον ένα aa και ακριβώς ένα bb και το bb είναι πιο αριστερά από την αριστερότερη εμφάνιση του aa*: Σε αυτήν την περίπτωση έχουμε ότι αριστερότερα από το (μοναδικό) bb καθώς και μεταξύ του bb και του αριστερότερου aa έχουμε εναλλάξ a και b, οπότε:

$$r_2 \;=\; (\varnothing^* \cup \mathrm{a})\,(\mathrm{b\,a})^*\,\mathrm{b\,b}\,(\mathrm{a\,b})^*\,\mathrm{a\,a}\,(\mathrm{a} \cup \mathrm{b\,a})^*\,(\varnothing^* \cup \mathrm{b})$$

ή επίσης

$$r_2 \;=\; (\varnothing^* \cup \mathrm{a})\,(\mathrm{b\,a})^*\,\mathrm{b\,b}\,(\mathrm{a\,b})^*\,\mathrm{a\,a\,a}^*\,(\mathrm{b\,a\,a}^*)^*\,(\varnothing^* \cup \mathrm{b}).$$

(γ) *Συμβολοσειρές που περιέχουν τουλάχιστον ένα aa και ακριβώς ένα bb και το bb είναι πιο δεξιά από την αριστερότερη εμφάνιση του aa*: Με ανάλογο τρόπο προκύπτει ότι η αντιστοιχη κανονική έκφραση για αυτές τις συμβολοσειρές είναι

$$r_3 \;=\; (\varnothing^* \cup \mathrm{b})\,(\mathrm{a\,b})^*\,\mathrm{a\,a}\,(\mathrm{a} \cup \mathrm{b\,a})^*\,\mathrm{b\,b}\,(\mathrm{a} \cup \mathrm{a\,b})^*$$

ή επίσης

$$\begin{aligned}
& r_3 \;=\; (\varnothing^* \cup \mathrm{b})\,(\mathrm{a\,b})^*\,\mathrm{a\,a}\,(\mathrm{a} \cup \mathrm{b\,a})^*\,\mathrm{b\,b}\,(\mathrm{a}^*\,\mathrm{a\,b})^*\,\mathrm{a}^* \\
ή\quad & r_3 \;=\; (\varnothing^* \cup \mathrm{b})\,(\mathrm{a\,b})^*\,\mathrm{a\,a\,a}^*\,(\mathrm{b\,a\,a}^*)^*\,\mathrm{b\,b}\,(\mathrm{a} \cup \mathrm{a\,b})^* \\
ή\quad & r_3 \;=\; (\varnothing^* \cup \mathrm{b})\,(\mathrm{a\,b})^*\,\mathrm{a\,a\,a}^*\,(\mathrm{b\,a\,a}^*)^*\,\mathrm{b\,b}\,(\mathrm{a}^*\,\mathrm{a\,b})^*\,\mathrm{a}^*.
\end{aligned}$$

Τελικά, καθώς η δοθείσα γλώσσα είναι η ένωση των συνόλων των συμβολοσειρών των περιπτώσεων (α), (β) και (γ), η ζητούμενη κανονική έκφραση είναι $r_1 \cup r_2 \cup r_3$.

Άσκηση 4

Ποια είναι η γλώσσα $\mathcal{L}\Big(\mathrm{c}^*(\mathrm{a} \cup \mathrm{b}^*\mathrm{c}^*)^*\Big)$; Αποδείξτε την απάντησή σας.

Λύση

Θα αποδείξουμε ότι η γλώσσα της συγκεκριμένης άσκησης, έστω L_4, ταυτίζεται με τη γλώσσα $\mathcal{L}\Big((\mathrm{a} \cup \mathrm{b} \cup \mathrm{c})^*\Big)$ όλων των συμβολοσειρών από το αλφάβητο $\{a, b, c\}$. Συγκεκριμένα, θα δείξουμε ότι $L_4 \subseteq \mathcal{L}\Big((\mathrm{a} \cup \mathrm{b} \cup \mathrm{c})^*\Big)$ και $\mathcal{L}\Big((\mathrm{a} \cup \mathrm{b} \cup \mathrm{c})^*\Big) \subseteq L_4$.

Εν πρώτοις, προφανώς $L_4 \subseteq \mathcal{L}\Big((\mathrm{a}\cup\mathrm{b}\cup\mathrm{c})^*\Big)$ καθώς η γλώσσα L_4 ορίζεται στο αλφάβητο $\{a,b,c\}$. Αρκεί λοιπόν να δείξουμε ότι $\mathcal{L}\Big((\mathrm{a}\cup\mathrm{b}\cup\mathrm{c})^*\Big) \subseteq L_4$. Ας θεωρήσουμε μια οποιαδήποτε συμβολοσειρά $w = \sigma_1\sigma_2\cdots\sigma_t$, όπου $\sigma_i \in \{a,b,c\}$, η οποία ανήκει στην $\mathcal{L}\Big((\mathrm{a}\cup\mathrm{b}\cup\mathrm{c})^*\Big)$. Τότε, καθώς τόσο το a, όσο και το b, αλλά και το c περιγράφονται από την κανονική έκφραση $\mathrm{a}\cup\mathrm{b}^*\mathrm{c}^*$, ολόκληρη η συμβολοσειρά w περιγράφεται από την κανονική έκφραση $(\mathrm{a}\cup\mathrm{b}^*\mathrm{c}^*)^*$ οπότε και από την $\mathrm{c}^*(\mathrm{a}\cup\mathrm{b}^*\mathrm{c}^*)^*$ δεδομένου ότι η κανονική έκφραση c^* περιγράφει την κενή συμβολοσειρά. Δηλαδή, $w \in L_4$, και άρα $\mathcal{L}\Big((\mathrm{a}\cup\mathrm{b}\cup\mathrm{c})^*\Big) \subseteq L_4$.

Άσκηση 5

Ποιες από τις παρακάτω προτάσεις ισχύουν; Δικαιολογήστε τις απαντήσεις σας.

(α) Κάθε υποσύνολο μιας κανονικής γλώσσας είναι κανονική γλώσσα.

(β) Εάν η L είναι κανονική, τότε και η γλώσσα $\{uv \mid u \in L,\ v \in \overline{L}\}$ είναι κανονική.

(γ) Εάν η L είναι κανονική, τότε και η γλώσσα $Pref(L) = \{u \mid \exists v : uv \in L\}$ είναι κανονική.

(δ) Εάν η L είναι κανονική, τότε και η γλώσσα $\{u \mid uu^R \in L\}$ είναι κανονική.

(ε) Η γλώσσα που περιέχει όλες τις συμβολοσειρές από το αλφάβητο $\Sigma = \{a,b,c\}$ με ίσο πλήθος από a και c είναι κανονική.

(στ) Η γλώσσα $\{a^ib^j \mid \text{μέγιστος\_κοινός\_διαιρέτης}(i,j) = 1\}$ είναι κανονική.

(ζ) Η γλώσσα $\{u \mid u \in \{a,b,c\}^* \text{ και } u = u^R\}$ είναι κανονική.

(η) Η γλώσσα $\{uvu^R \mid u,v \in \{a,b,c\}^*\}$ είναι κανονική.

Λύση

(α) **Δεν ισχύει.** Παρατηρήστε ότι η γλώσσα $\{a^ib^i \mid i \geq 0\}$, για την οποία γνωρίζουμε ότι δεν είναι κανονική, είναι υποσύνολο της κανονικής γλώσσας που περιγράφεται από την κανονική έκφραση $\mathrm{a}^*\mathrm{b}^*$.

(β) **Ισχύει.** Έστω $L_{(\beta)}$ η γλώσσα $\{uv \mid u \in L,\ v \in \overline{L}\}$. Ο ορισμός της $L_{(\beta)}$ συνεπάγεται ότι αυτή είναι ίση με την παράθεση $L\,\overline{L}$. Άρα, η $L_{(\beta)}$ είναι κανονική καθώς οι κανονικές γλώσσες είναι κλειστές ως προς το συμπλήρωμα και την παράθεση.

(γ) **Ισχύει.** Θα το αποδείξουμε κατασκευάζοντας ένα πεπερασμένο αυτόματο που αποδέχεται τη γλώσσα $Pref(L)$ χρησιμοποιώντας ένα ντετερμινιστικό πεπερασμένο αυτόματο M_L για την L (το οποίο σίγουρα υπάρχει καθώς μας δίδεται ότι η L είναι κανονική).

Από τον ορισμό της $Pref(L)$ προκύπτει ότι το αυτόματο για την $Pref(L)$ θα πρέπει να αποδέχεται όλα τα προθέματα όλων των συμβολοσειρών της L. Δεδομένου ότι για οποιαδήποτε συμβολοσειρά x της L υπάρχει μοναδική διαδρομή στο διάγραμμα καταστάσεων του αυτόματου M_L από την αρχική κατάσταση σε κάποια τελική κατάστση και τα προθέματα της x μας οδηγούν στις καταστάσεις αυτής της διαδρομής, θα επιτύχουμε τον στόχο μας εάν κάνουμε τελικές όλες τις καταστάσεις στις οποίες φθάνουμε από την αρχική κατάσταση και από τις οποίες καταλήγουμε σε κάποια τελική κατάσταση του M_L. Δηλαδή, το ζητούμενο αυτόματο για την $Pref(L)$ έχει το ίδιο σύνολο καταστάσεων, το ίδιο αλφάβητο, τις ίδιες μεταβάσεις και την ίδια αρχική κατάσταση με το M_L, μόνο που οι τελικές του καταστάσεις είναι όλες εκείνες οι καταστάσεις για τις οποίες υπάρχει ακολουθία μεταβάσεων από την αρχική κατάσταση προς αυτές και ακολουθία μεταβάσεων από αυτές σε κάποια τελική κατάσταση του M_L.

(δ) **Ισχύει.** Δουλεύουμε όπως και στο υποερώτημα (γ) αξιοποιώντας ένα ντετερμινιστικό αυτόματο $M_L = (K, \Sigma, q_0, \delta, F)$ για την L ώστε να κατασκευάσουμε ένα αυτόματο που θα αποδέχεται την $L_{(\delta)} = \{u \mid uu^R \in L\}$. Στην προκειμένη περίπτωση, η κατασκευή είναι πιο σύνθετη και βασίζεται στην εξής παρατήρηση: Έστω u μια συμβολοσειρά και έστω q η κατάσταση του M_L στην οποία καταλήγουμε εάν ξεκινήσουμε από την αρχική κατάσταση q_0 του M_L και ακολουθήσουμε μεταβάσεις για τα σύμβολα της u με τη σειρά που εμφανίζονται στην u. Τότε, η u ανήκει στην $L_{(\delta)}$ εάν και μόνον εάν από την q για την συμβολοσειρά u^R καταλήγουμε σε τελική καταστασηση του M_L.

Για να μπορέσουμε να έχουμε στη διάθεσή μας όλες τις πληροφορίες που χρειαζόμαστε, θα υλοποιήσουμε στο ζητούμενο αυτόματο για την $L_{(\delta)}$ "κίνηση κατά την ευθεία φορά" και "κίνηση κατά την αντίστροφη φορά". Συγκεκριμένα, ας θεωρήσουμε μια συμβολοσειρά $u = \sigma_1\sigma_2\cdots\sigma_t$. Για το πρώτο σύμβολο σ_1 της u, θα κινηθούμε κατά την "ευθεία φορά" από την αρχική κατάσταση q_0 του M_L στην κατάσταση $q' = \delta(q_0, \sigma_1)$, ενώ παράλληλα κατά την "αντίστροφη φορά" θα βρούμε το σύνολο, έστω S', όλων των καταστάσεων από τις οποίες για το σ_1 υπάρχει μετάβαση που οδηγεί σε κάποια τελική κατάσταση. Κατόπιν, για το δεύτερο σύμβολο σ_2 της u, θα κινηθούμε κατά την "ευθεία φορά" από την κατάσταση q' του M_L στην κατάσταση $\delta(q', \sigma_2)$, ενώ παράλληλα κατά την "αντίστροφη φορά" θα βρούμε το σύνολο, έστω S'', όλων των καταστάσεων από τις οποίες για το σ_2 υπάρχει μετάβαση που οδηγεί σε κάποια κατάσταση στο S', κ.ο.κ. Η συμβολοσειρά u θα γίνει αποδεκτή εάν, όταν την διαβάσουμε ολόκληρη, η κατάσταση του M_L στην οποία θα βρισκόμαστε κατά την "ευθεία φορά" ανήκει στο σύνολο καταστάσεων που έχουν προκύψει κατά την κίνηση κατά την "αντίστροφη φορά".

Θεωρώντας τη συνάρτηση $back(S, \sigma) = \{q \in K \mid \delta(q, \sigma) \in S\}$, το ζητούμενο αυτόματο είναι $M = (K', \Sigma, q_0', \delta', F')$

$$\begin{array}{llll} \text{όπου} & K' & = & \{[p, S] \mid \forall p \in K \text{ και } \forall S \subseteq K\} \\ & q_0' & = & [q_0, F] \\ & \delta' & : & \delta'([p, S], \sigma) = [\delta(p, \sigma), back(S, \sigma)] \\ & F' & = & \{[p, S] \mid p \in S\} \end{array}$$

(ε) **Δεν ισχύει.** Θα το αποδείξουμε χρησιμοποιώντας το Θεώρημα Άντλησης.
Υποθέτουμε λοιπόν ότι η γλώσσα του συγκεκριμένου υποερωτήματος, έστω $L_{(\varepsilon)}$, είναι κανονική και έστω k η σταθερά του Θεωρήματος Άντλησης για την $L_{(\varepsilon)}$.
Θεωρούμε τη συμβολοσειρά $w = a^k c^k$ η οποία ανήκει στην $L_{(\varepsilon)}$ και έχει μήκος τουλάχιστον k.
Τότε, μπορούμε να χειριστούμε κάθε πιθανό χωρισμό της w υπό τους περιορισμούς του Θεωρήματος Άντλησης για κανονικές γλώσσες (δηλαδή, $w = xyz$ όπου $|xy| \leq k$ και $|y| \geq 1$) με βάση την εξής μία αλλά γενική περίπτωση:

$$x = a^p \qquad y = a^q \qquad z = a^{k-p-q} c^k$$

όπου $0 \leq p \leq k - q$ και $1 \leq q \leq k - p$. Τότε, όμως, για $i = 2$, η συμβολοσειρά xy^iz είναι $a^{k+q} c^k$, η οποία προφανώς δεν ανήκει στη γλώσσα $L_{(\varepsilon)}$. Καταλήξαμε λοιπόν σε άτοπο, και άρα η υπόθεσή μας ότι η $L_{(\varepsilon)}$ είναι κανονική δεν ισχύει.

Ένας εναλλακτικός τρόπος απόδειξης ότι η γλώσσα $L_{(\varepsilon)}$ δεν είναι κανονική βασίζεται σε εις άτοπον απαγωγή με χρήση των ιδιοτήτων κλειστότητας των κανονικών γλωσσών. Υποθέτουμε λοιπόν ότι η $L_{(\varepsilon)}$ είναι κανονική. Τότε και η $L_{(\varepsilon)} \cap \mathcal{L}(\mathrm{a}^*\mathrm{c}^*)$ θα πρέπει να είναι κανονική. Αλλά $L_{(\varepsilon)} \cap \mathcal{L}(\mathrm{a}^*\mathrm{c}^*) = \{a^i c^i \mid i \geq 0\}$, η οποία, όπως και η $\{a^i b^i \mid i \geq 0\}$, δεν είναι κανονική. Άτοπο. Άρα η $L_{(\varepsilon)}$ δεν είναι κανονική.

(στ) **Δεν ισχύει.** Θα το αποδείξουμε χρησιμοποιώντας το Θεώρημα Άντλησης.
Υποθέτουμε λοιπόν ότι η γλώσσα του συγκεκριμένου υποερωτήματος, έστω $L_{(\sigma\tau)}$, είναι κανονική και έστω k η σταθερά του Θεωρήματος Άντλησης για την $L_{(\sigma\tau)}$.
Θεωρούμε τη συμβολοσειρά $w = a^{(k+1)!+1}\, b^{(k+1)!}$ η οποία ανήκει στην $L_{(\sigma\tau)}$ και έχει μήκος τουλάχιστον k (ο συμβολισμός $(k+1)!$ δηλώνει το παραγοντικό του αριθμού $k+1$).
Τότε, μπορούμε να χειριστούμε κάθε πιθανό χωρισμό της w υπό τους περιορισμούς του Θεωρήματος Άντλησης για κανονικές γλώσσες (δηλαδή, $w = xyz$ όπου $|xy| \leq k$ και $|y| \geq 1$) με βάση την εξής μία αλλά γενική περίπτωση:

$$x = a^p \qquad y = a^q \qquad z = a^{(k+1)!+1-p-q}\, b^{(k+1)!}$$

όπου $0 \leq p \leq k-q$ και $1 \leq q \leq k-p$. Τότε, όμως, για $i=2$, η συμβολοσειρά xy^iz είναι $a^{(k+1)!+1+q}\, b^{(k+1)!}$, η οποία προφανώς δεν ανήκει στη γλώσσα $L_{(\sigma\tau)}$, καθώς τα πλήθη $(k+1)!+1+q$ και $(k+1)!$ των a και b έχουν ως κοινό διαιρέτη το $q+1$ (παρατηρήστε ότι $2 \leq q+1 \leq k+1$), και άρα ο μέγιστος κοινός διαιρέτης τους είναι τουλάχιστον ίσος με $q+1 \geq 2$. Καταλήξαμε λοιπόν σε άτοπο, και άρα η υπόθεσή μας ότι η $L_{(\sigma\tau)}$ είναι κανονική δεν ισχύει.

(ζ) **Δεν ισχύει.** Θα το αποδείξουμε χρησιμοποιώντας το Θεώρημα Άντλησης.
Υποθέτουμε λοιπόν ότι η γλώσσα του συγκεκριμένου υποερωτήματος, έστω $L_{(\zeta)}$, είναι κανονική και έστω k η σταθερά του Θεωρήματος Άντλησης για την $L_{(\zeta)}$.
Θεωρούμε τη συμβολοσειρά $w = a^k\, b\, a^k$ η οποία ανήκει στην $L_{(\zeta)}$ και έχει μήκος τουλάχιστον k.
Τότε, μπορούμε να χειριστούμε κάθε πιθανό χωρισμό της w υπό τους περιορισμούς του Θεωρήματος Άντλησης για κανονικές γλώσσες (δηλαδή, $w = xyz$ όπου $|xy| \leq k$ και $|y| \geq 1$) με βάση την εξής μία αλλά γενική περίπτωση:

$$x = a^p \qquad y = a^q \qquad z = a^{k-p-q}\, b\, a^k$$

όπου $0 \leq p \leq k-q$ και $1 \leq q \leq k-p$. Τότε, όμως, για $i=2$, η συμβολοσειρά xy^iz είναι $a^{k+q}\, b\, a^k$, η οποία προφανώς δεν ανήκει στη γλώσσα $L_{(\zeta)}$ καθώς δεν είναι παλίνδρομη. Καταλήξαμε λοιπόν σε άτοπο, και άρα η υπόθεσή μας ότι η $L_{(\zeta)}$ είναι κανονική δεν ισχύει.

(η) **Ισχύει.** Δεν είναι δύσκολο να δειχθεί ότι η γλώσσα του συγκεκριμένου υποερωτήματος, έστω $L_{(\eta)}$, ταυτίζεται με τη γλώσσα $\mathcal{L}\Big((\mathrm{a} \cup \mathrm{b} \cup \mathrm{c})^*\Big)$ όλων των συμβολοσειρών από το αλφάβητο $\{a, b, c\}$ και άρα είναι κανονική. Παρατηρήστε ότι $L_{(\eta)} \subseteq \mathcal{L}\Big((\mathrm{a} \cup \mathrm{b} \cup \mathrm{c})^*\Big)$ ενώ επίσης κάθε συμβολοσειρά w από a, b, c ανήκει στην $L_{(\eta)}$ καθώς μπορεί να γραφεί ως $w = uvu^R$ όπου $u = e$ και $v = w$.

Άσκηση 6

Δείξτε ότι το Θεώρημα Άντλησης για κανονικές γλώσσες δεν μας βοηθάει να δείξουμε ότι η γλώσσα $\{uu^Rv \mid u, v \in \{a,b\}\{a,b\}^*\}$ δεν είναι κανονική.

Λύση

Ας προσπαθήσουμε να εφαρμόσουμε το Θεώρημα Άντλήσης.
Υποθέτουμε ότι η δοθείσα γλώσσα, έστω L_6, είναι κανονική και έστω k η σταθερά του Θεωρήματος Άντλησης για την L_6. Ακολούθως, θα πρέπει να επιλέξουμε μια συμβολοσειρά w που ανήκει στη γλώσσα L_6 και έχει μήκος τουλάχιστον k. Για οποιαδήποτε επιλογή της συμβολοσειράς w, θα προκύψει μία από τις εξής δύο περιπτώσεις:

(α) $|u| = 1$. Έστω $u = \sigma$ όπου $\sigma \in \{a, b\}$.
Τότε $w = \sigma\sigma v$ όπου $v = \sigma_1\sigma_2 \cdots \sigma_t$ για $\sigma_i \in \{a, b\}$ και έστω $t \geq \max\{k-2, 2\}$.
Θεωρήστε τον εξής χωρισμό τής w: $w = xyz$ όπου $x = \sigma\sigma$, $y = \sigma_1$, $z = \sigma_2 \cdots \sigma_t$.
(Ο χωρισμός αυτός είναι συμβατός με τους περιορισμούς $|xy| \leq k$ και $|y| \geq 1$, υποθέτοντας χωρίς βλάβη της γενικότητας ότι $k \geq 3$.)

- ▷ Εάν $i = 0$, τότε $xy^0z = \sigma\sigma\sigma_2 \cdots \sigma_t$, η οποία ανήκει στην L_6 καθώς $xy^0z = u'u'^Rv'$ για $u' = \sigma$ και $v' = \sigma_2 \cdots \sigma_t$ (παρατηρήστε ότι $u', v' \in \{a, b\}\{a, b\}^*$).
- ▷ Εάν $i = 1$, τότε $xy^1z = w$, η οποία εξ ορισμού ανήκει στην L_6.
- ▷ Εάν $i \geq 2$, τότε $xy^iz = \sigma\sigma\underbrace{\sigma_1 \cdots \sigma_1}_{i}\sigma_2 \cdots \sigma_t$ (δηλ., υπάρχουν i επαναλήψεις του σ_1), η οποία επίσης ανήκει στην L_6 για κάθε $i \geq 2$, καθώς $xy^iz = u'u'^Rv'$ για $u' = \sigma$ και $v' = \underbrace{\sigma_1 \cdots \sigma_1}_{i}\sigma_2 \cdots \sigma_t$ (παρατηρήστε ότι $u', v' \in \{a, b\}\{a, b\}^*$).

Άρα, στη συγκεκριμένη περίπτωση, για κάθε τιμή του $i \geq 0$, η συμβολοσειρά xy^iz ανήκει στη γλώσσα L_6.

(β) $|u| \geq 2$. Έστω $u = \sigma_1\sigma_2 \cdots \sigma_t$ όπου $\sigma_i \in \{a, b\}$ και $t \geq 2$.
Τότε, $w = uu^Rv$ για v τέτοιο ώστε $|v| + 2t \geq k$.
Θεωρήστε τον εξής χωρισμό τής w: $w = xyz$ όπου $x = e$, $y = \sigma_1$, $z = \sigma_2 \cdots \sigma_t\sigma_t \cdots \sigma_2\sigma_1 v$ (και ο χωρισμός αυτός είναι συμβατός με τους περιορισμούς $|xy| \leq k$ και $|y| \geq 1$).

- ▷ Εάν $i = 0$, τότε $xy^0z = \sigma_2 \cdots \sigma_t\sigma_t \cdots \sigma_2\sigma_1 v$, η οποία ανήκει στην L_6 καθώς $xy^0z = u'u'^Rv'$ για $u' = \sigma_2 \cdots \sigma_t$ και $v' = \sigma_1 v$ (παρατηρήστε ότι $u', v' \in \{a, b\}\{a, b\}^*$).
- ▷ Εάν $i = 1$, τότε $xy^1z = w$, η οποία εξ ορισμού ανήκει στην L_6.
- ▷ Εάν $i \geq 2$, τότε $xy^iz = \underbrace{\sigma_1 \cdots \sigma_1}_{i}\sigma_2 \cdots \sigma_t\sigma_t \cdots \sigma_2\sigma_1 v$ η οποία επίσης ανήκει στην L_6 για κάθε $i \geq 2$ καθώς $xy^iz = u'u'^Rv'$ για $u' = \sigma_1$ και $v' = \underbrace{\sigma_1 \cdots \sigma_1}_{i-2}\sigma_2 \cdots \sigma_t\sigma_t \cdots \sigma_2\sigma_1 v$ (παρατηρήστε ότι $u', v' \in \{a, b\}\{a, b\}^*$).

Άρα, και σε αυτήν την περίπτωση, για κάθε τιμή του $i \geq 0$, η συμβολοσειρά xy^iz ανήκει στη γλώσσα L_6.

Τελικά, για οποιαδήποτε επιλογή της συμβολοσειράς w, υπάρχει χωρισμός της w σύμφωνα με τους περιορισμούς του Θεωρήματος Άντλησης για τον οποίο δεν υπάρχει $i \geq 0$ τέτοιο ώστε $xy^iz \notin L_6$ (το οποίο θα μας οδηγούσε σε άτοπο).