Πανεπιστήμιο Ιωαννίνων
Πρόγραμμα Προπτυχιακών Σπουδών

Τμήμα Πληροφορικής
Εαρινό Εξάμηνο 2010-2011

4-44: Θεωρία Υπολογισμού

Λύσεις της 1ης Ομάδας Προτεινομένων Ασκήσεων

Άσκηση 1

Γράψτε κανονικές εκφράσεις για καθεμία από τις παρακάτω γλώσσες ορισμένες στο αλφάβητο $\{a, b\}$.

α) Το σύνολο των συμβολοσειρών που περιέχουν (τουλάχιστον ένα) aa.

β) Το σύνολο των συμβολοσειρών που δεν περιέχουν aa.

γ) Το σύνολο των συμβολοσειρών που δεν περιέχουν ούτε aa ούτε bb.

δ) Το σύνολο των συμβολοσειρών με το πολύ ένα ζεύγος διαδοχικών a και το πολύ ένα ζεύγος διαδοχικών b.

ε) Το σύνολο των συμβολοσειρών που δεν περιέχουν το bab.

Λύση

α) Η κανονική έκφραση είναι: $(\mathtt{a} \cup \mathtt{b})^*\mathtt{aa}(\mathtt{a} \cup \mathtt{b})^*$.

β) Σε περίπτωση που η ιδιότητα των συμβολοσειρών (όπως στη συγκεκριμένη περίπτωση) διατυπώνεται με μια αρνητική πρόταση, προσπαθούμε να βρούμε μια ισοδύναμη καταφατική πρόταση. Καθώς σε μια συμβολοσειρά που δεν περιέχει aa, κάθε a που δεν είναι το τελευταίο σύμβολο της συμβολοσειράς θα πρέπει να ακολουθείται από b, καταλήγουμε στην εξής κανονική έκφραση: $\mathtt{b}^*(\mathtt{abb}^*)^*(\varnothing^* \cup \mathtt{a})$.
Άλλες επίσης σωστές λύσεις είναι: $(\mathtt{b} \cup \mathtt{ab})^*(\varnothing^* \cup \mathtt{a})$ και $(\varnothing^* \cup \mathtt{a})(\mathtt{b} \cup \mathtt{ba})^*$.

γ) Σε αυτήν την περίπτωση, οι συμβολοσειρές θα πρέπει να έχουν a και b εναλλάξ. Δύο σωστές λύσεις είναι: $(\varnothing^* \cup \mathtt{b})(\mathtt{ab})^*(\varnothing^* \cup \mathtt{a})$ και $(\varnothing^* \cup \mathtt{a})(\mathtt{ba})^*(\varnothing^* \cup \mathtt{b})$.

δ) Διακρίνουμε περιπτώσεις.

1. *Η συμβολοσειρά δεν περιέχει ούτε aa ούτε bb:*
Τότε από το υποερώτημα (γ) έχουμε: $r_1 = (\varnothing^* \cup \mathtt{b})(\mathtt{ab})^*(\varnothing^* \cup \mathtt{a})$.
2. *Η συμβολοσειρά περιέχει ένα aa αλλά όχι bb:*
Τότε: $r_2 = (\varnothing^* \cup \mathtt{b})(\mathtt{ab})^*\mathtt{aa}(\mathtt{ba})^*(\varnothing^* \cup \mathtt{b})$.
3. *Η συμβολοσειρά περιέχει ένα bb αλλά όχι aa:*
Τότε όπως στην περίπτωση 2: $r_3 = (\varnothing^* \cup \mathtt{a})(\mathtt{ba})^*\mathtt{bb}(\mathtt{ab})^*(\varnothing^* \cup \mathtt{a})$.
4. *Η συμβολοσειρά περιέχει ένα aa και ένα bb και το aa είναι πριν το bb:*
Τότε: $r_4 = (\varnothing^* \cup \mathtt{b})(\mathtt{ab})^*\mathtt{aa}(\mathtt{ba})^*\mathtt{bb}(\mathtt{ab})^*(\varnothing^* \cup \mathtt{a})$.
5. *Η συμβολοσειρά περιέχει ένα aa και ένα bb και το aa είναι μετά το bb:*
Τότε: $r_5 = (\varnothing^* \cup \mathtt{a})(\mathtt{ba})^*\mathtt{bb}(\mathtt{ab})^*\mathtt{aa}(\mathtt{ba})^*(\varnothing^* \cup \mathtt{b})$.

Τελικά, η κανονική έκφραση για τη δοθείσα γλώσσα είναι η $r = r_1 \cup r_2 \cup r_3 \cup r_4 \cup r_5$.

ε) Σε αυτήν την περίπτωση, κάθε ba που δεν βρίσκεται στο τέλος της συμβολοσειράς θα πρέπει να ακολουθείται από a, ενώ επίσης το πρόθεμα της συμβολοσειράς έως και

πριν την 1η εμφάνιση του *ba* και οι υποσυμβολοσειρές μεταξύ διαδοχικών εμφανίσεων του *ba* δεν θα πρέπει να περιέχουν *ba* οπότε περιγράφονται από την κανονική έκφραση $\texttt{a}^*\texttt{b}^*$. Άρα μια λύση είναι: $\texttt{a}^*\texttt{b}^*(\texttt{baaa}^*\texttt{b}^*)^*(\emptyset^* \cup \texttt{ba})$.
Εξίσου σωστές λύσεις είναι και οι: $\texttt{a}^*\texttt{b}^*(\texttt{baaa}^*\texttt{b}^*)^*(\emptyset^* \cup \texttt{a})$ και $\texttt{a}^*\texttt{b}^*(\texttt{baaa}^*\texttt{b}^*)^*\texttt{a}^*$.

Άσκηση 2

Περιγράψτε τις γλώσσες που περιγράφονται από τις ακόλουθες κανονικές εκφράσεις.

α) $(\texttt{bb} \cup \texttt{a})^*$

β) $(\texttt{bb} \cup \texttt{aa})^*$

γ) $(\texttt{b} \cup \texttt{ab} \cup \texttt{aab})^*(\emptyset^* \cup \texttt{a} \cup \texttt{aa})$

δ) $(\texttt{aa} \cup \texttt{bb} \cup (\texttt{ab} \cup \texttt{ba})(\texttt{aa} \cup \texttt{bb})^*(\texttt{ab} \cup \texttt{ba}))^*$

Λύση

Για την λύση των υποερωτημάτων (α) και (β), ορίζουμε τις μέγιστες υποσυμβολοσειρές μιας συμβολοσειράς w ως εξής: κάθε υποσυμβολοσειρά y της w η οποία αποτελείται από παράθεση του ιδίου συμβόλου και είναι τέτοια ώστε εάν $w = xyz$ τότε

▷ $x = e$ ή το δεξιότερο σύμβολο της x είναι διαφορετικό από τα σύμβολα της y

και

▷ $z = e$ ή το αριστερότερο σύμβολο της z είναι διαφορετικό από τα σύμβολα της y

είναι μια *μέγιστη υποσυμβολοσειρά* της w. Για παράδειγμα, (όλες) οι μέγιστες υποσυμβολοσειρές της συμβολοσειράς *abbbaacaaaa* είναι: *a*, *bbb*, *aa*, *c* και *aaaa*· από αυτές οι *a*, *aa* και *aaaa* είναι οι μέγιστες υποσυμβολοσειρές από *a*.

α) Από την κανονική έκφραση συμπεραίνουμε ότι οι συμβολοσειρές που περιγράφονται προκύπτουν από παράθεση υποσυμβολοσειρών που είναι *a* ή *bb*. Βλέπουμε λοιπόν ότι τα *b* εμφανίζονται σε ζεύγη διαδοχικών *b*, από όπου συμπεραίνουμε ότι η κανονική έκφραση περιγράφει τη γλώσσα των συμβολοσειρών από *a* και *b* στις οποίες καθεμία από τις μέγιστες υποσυμβολοσειρές από *b* είναι αρτίου μήκους.

β) Αντίστοιχα προς το υποερώτημα (α), η κανονική έκφραση περιγράφει τη γλώσσα των συμβολοσειρών από *a* και *b* στις οποίες καθεμία από τις μέγιστες υποσυμβολοσειρές από *b* και καθεμία από τις μέγιστες υποσυμβολοσειρές από *a* είναι αρτίου μήκους.

γ) Η συγκεκριμένη κανονική έκφραση αποτελεί γενίκευση της κανονικής έκφρασης $(\texttt{b} \cup \texttt{ab})^*(\emptyset^* \cup \texttt{a})$ που, όπως έχουμε δει, περιγράφει όλες τις συμβολοσειρές από *a* και *b* που δεν περιέχουν το *aa* (υποερώτημα (β) της Άσκησης 1). Συνεπώς, η κανονική έκφραση του υποερωτήματος (γ) περιγράφει τη γλώσσα των συμβολοσειρών από *a* και *b* που δεν περιέχουν το *aaa*.

δ) Η συγκεκριμένη κανονική έκφραση περιγράφει τις συμβολοσειρές από *a* και *b* που έχουν άρτιο πλήθος *a* και άρτιο πλήθος *b*.
Κατά πρώτον, παρατηρούμε ότι κάθε συμβολοσειρά u που περιγράφεται από την κανονική έκφραση έχει άρτιο πλήθος *a* και άρτιο πλήθος *b*: η u προκύπτει από παραθέσεις συμβολοσειρών που περιγράφονται από τις κανονικές εκφράσεις $\texttt{aa}$, $\texttt{bb}$ και $(\texttt{ab} \cup \texttt{ba})(\texttt{aa} \cup \texttt{bb})^*(\texttt{ab} \cup \texttt{ba})$, καθεμία από τις οποίες περιγράφει συμβολοσειρές με άρτιο πλήθος *a* και άρτιο πλήθος *b*.
Τέλος, πρέπει να δείξουμε ότι κάθε συμβολοσειρά w με άρτιο πλήθος *a* και άρτιο πλή-

θος b περιγράφεται από την κανονική έκφραση του υποερωτήματος. Αποδεικνύουμε ότι αυτό πράγματι ισχύει ως εξής: Χωρίζουμε την w παίρνοντας δύο-δύο τα σύμβολά της (π.χ. η συμβολοσειρά $abbbab$ χωρίζεται σε ab, bb, ab). Κάποια από αυτά τα ζεύγη ενδέχεται να είναι ab ή ba· ας ονομάσουμε αυτά τα ζεύγη *ανόμοια*. Καθώς η w έχει άρτιο πλήθος a, το πλήθος των ανόμοιων ζευγών είναι άρτιο. Τότε μπορεί να δει κάποιος ότι η υποσυμβολοσειρά της w από την αρχή έως (και χωρίς να συμπεριλαμβάνει) το πρώτο ανόμοιο ζεύγος περιγράφεται από την $(\mathtt{aa} \cup \mathtt{bb})^*$, η υποσυμβολοσειρά από το πρώτο ανόμοιο ζεύγος (συμπεριλαμβανομένου) έως το δεύτερο ανόμοιο ζεύγος (συμπεριλαμβανομένου) περιγράφεται από την $(\mathtt{ab} \cup \mathtt{ba})(\mathtt{aa} \cup \mathtt{bb})^*(\mathtt{ab} \cup \mathtt{ba})$, η υποσυμβολοσειρά από εκεί έως το τρίτο ανόμοιο ζεύγος (μη συμπεριλαμβανομένου) περιγράφεται από την $(\mathtt{aa} \cup \mathtt{bb})^*$, η υποσυμβολοσειρά από το τρίτο ανόμοιο ζεύγος (συμπεριλαμβανομένου) έως το τέταρτο ανόμοιο ζεύγος (συμπεριλαμβανομένου) περιγράφεται από την $(\mathtt{ab} \cup \mathtt{ba})(\mathtt{aa} \cup \mathtt{bb})^*(\mathtt{ab} \cup \mathtt{ba})$, και ούτω καθεξής, ενώ η υποσυμβολοσειρά από το τελευταίο ανόμοιο ζεύγος (μη συμπεριλαμβανομένου) έως το τέλος της w περιγράφεται από την $(\mathtt{aa} \cup \mathtt{bb})^*$. Δηλαδή, η w περιγράφεται από τη δοθείσα κανονική έκφραση.

Άσκηση 3

Αποδείξτε τις ακόλουθες ταυτότητες που ισχύουν για κανονικές εκφράσεις.

α) $r \cup s \;=\; s \cup r$

β) $(rs)t \;=\; r(st)$

γ) $r(s \cup t) \;=\; rs \cup rt$

δ) $(\varnothing^* \cup r)r^* \;=\; r^*$

ε) $(r^*)^* \;=\; r^*$

Λύση

Τα υποερωτήματα (α)-(γ) προκύπτουν άμεσα με εφαρμογή του ορισμού των κανονικών εκφράσεων δεδομένου ότι για σύνολα R, S και T ισχύει ότι $R \cup S = S \cup R$, $(RS)T = R(ST)$ και $R(S \cup T) = RS \cup RT$.

δ) Από το υποερώτημα (γ) έχουμε ότι $(\varnothing^* \cup r)r^* = \varnothing^* r^* \cup rr^* = r^* \cup rr^*$ το οποίο είναι ίσο με r^* καθώς $\mathcal{L}(rr^*) \subseteq \mathcal{L}(r^*)$.
Σημειώστε ότι η $\mathcal{L}(rr^*)$ είναι η γλώσσα των συμβολοσειρών που προκύπτουν από παραθέσεις μίας ή περισσοτέρων συμβολοσειρών που περιγράφονται από την κανονική έκφραση r, ενώ η $\mathcal{L}(r^*)$ είναι η γλώσσα των συμβολοσειρών που προκύπτουν από παραθέσεις οσωνδήποτε (ακόμη και μηδέν) συμβολοσειρών που περιγράφονται από την κανονική έκφραση r.

ε) Θα δείξουμε ότι $\mathcal{L}\big((r^*)^*\big) = \mathcal{L}(r^*)$. Προφανώς, $\mathcal{L}(r^*) \subseteq \mathcal{L}\big((r^*)^*\big)$: εάν θέσουμε $s = r^*$, τότε η προηγούμενη σχέση γράφεται $\mathcal{L}(s) \subseteq \mathcal{L}(s^*)$ που ισχύει για κάθε s.
Για να δείξουμε ότι $\mathcal{L}\big((r^*)^*\big) \subseteq \mathcal{L}(r^*)$, θεωρούμε μια συμβολοσειρά $w \in \mathcal{L}\big((r^*)^*\big)$. Τότε όμως η w είναι παράθεση υποσυμβολοσειρών $u_1, u_2, \ldots, u_k$ (δηλαδή, $w = u_1 u_2 \cdots u_k$) καθεμία από τις οποίες περιγράφεται από την κανονική έκφραση r^*. Συνεπώς, για κάθε $i = 1, \ldots, k$, ισχύει ότι $u_i = x_{i,1} x_{i,2} \cdots x_{i,\ell_i}$,[1] $\ell_i \geq 0$, όπου κάθε υποσυμβολοσειρά $x_{i,j}$

[1] Σημειώνεται ότι δεν είναι απαραίτητο τα $\ell_1, \ell_2, \ldots, \ell_k$ να είναι ίσα μεταξύ τους.

περιγράφεται από την κανονική έκφραση r. Τότε, συνολικά

$$w = x_{1,1}\, x_{1,2} \cdots x_{1,\ell_1}\, x_{2,1}\, x_{2,2} \cdots x_{2,\ell_2} \quad \cdots \quad x_{k,1}\, x_{k,2} \cdots x_{k,\ell_k}\,,$$

δηλαδή, η w είναι παράθεση συμβολοσειρών που περιγράφονται από την κανονική έκφραση r. Άρα, $w \in \mathcal{L}(r^*)$. Καθώς τα προηγούμενα ισχύουν για οποιαδήποτε συμβολοσειρά w της $\mathcal{L}\big((r^*)^*\big)$ συμπεραίνουμε ότι $\mathcal{L}\big((r^*)^*\big) \subseteq \mathcal{L}(r^*)$.
Οι σχέσεις $\mathcal{L}(r^*) \subseteq \mathcal{L}\big((r^*)^*\big)$ και $\mathcal{L}\big((r^*)^*\big) \subseteq \mathcal{L}(r^*)$ συνεπάγονται ότι $\mathcal{L}\big((r^*)^*\big) = \mathcal{L}(r^*)$.

Άσκηση 4

Ποιες από τις παρακάτω ισότητες για κανονικές εκφράσεις ισχύουν και γιατί;

α) $(rs \cup r)^* r \;=\; r(sr \cup r)^*$

β) $s(rs \cup s)^* r \;=\; rr^* s(rr^* s)^*$

Λύση

α) Δουλεύοντας με "μικρά" παραδείγματα διαπιστώνουμε ότι η ισότητα φαίνεται να ισχύει. Με λίγη προσοχή παρατηρούμε ότι ισχύει η εξής ισότητα

$$(rs \cup r)^i r \;=\; r\,(sr \cup r)^i \qquad \forall i \geq 0 \tag{1}$$

όπου ο συμβολισμός t^i για μια κανονική έκφραση t δηλώνει μια κανονική έκφραση που περιγράφει συμβολοσειρές που προκύπτουν από την παράθεση i συμβολοσειρών καθεμία από τις οποίες περιγράφεται από την t. Η απόδειξη της ισότητας (1) βασίζεται σε επαγωγή στο i.
Βάση: Για $i = 0$, έχουμε ότι $(rs \cup r)^i r = (rs \cup r)^0 r = r$ το οποίο είναι ίσο με $r = r(sr \cup r)^0$. Άρα η βάση ισχύει.
Επαγ. Υπόθεση: Έστω $(rs \cup r)^k r = r(sr \cup r)^k$ για $k \geq 0$.
Επαγ. Βήμα: Θα δείξουμε ότι $(rs \cup r)^{k+1} r \;=\; r(sr \cup r)^{k+1}$.

$$\begin{aligned}
(rs \cup r)^{k+1} r &= (rs \cup r)\,(rs \cup r)^k r \\
&= (rs \cup r) r\,(sr \cup r)^k && \text{(από επαγ. υπόθεση)} \\
&= (rsr \cup rr)\,(sr \cup r)^k \\
&= r(sr \cup r)\,(sr \cup r)^k \\
&= r\,(sr \cup r)^{k+1}.
\end{aligned}$$

Η απόδειξη του επαγωγικού βήματος ολοκληρώθηκε και άρα η ισότητα (1) ισχύει για κάθε $i \geq 0$. Η ισχύς της ισότητας (1) συνεπάγεται την ισχύ της ισότητας του υποερωτήματος (α) καθώς κάθε συμβολοσειρά που περιγράφεται από την κανονική έκφραση $(rs \cup r)^* r$ θα περιγράφεται από την $(rs \cup r)^i r$ για κάποιο $i \geq 0$, οπότε και από την $r(sr \cup r)^i$ λόγω της (1) και άρα και από την $r(sr \cup r)^*$. Όμοια και για συμβολοσειρές που περιγράφονται από την κανονική έκφραση $r(sr \cup r)^*$.

β) Η ισότητα δεν ισχύει. Ως αντιπαράδειγμα θεωρείστε τις κανονικές εκφράσεις $r = \mathtt{a}$ και $s = \mathtt{b}$. Τότε η κανονική έκφραση $s(rs \cup s)^* r$ γίνεται $\mathtt{b(ab \cup b)^*a}$ και περιγράφει συμβολοσειρές που αρχίζουν με b και τελειώνουν με a (και επίσης δεν περιέχουν aa), ενώ η $rr^* s(rr^* s)^*$ γίνεται $\mathtt{aa^*b(aa^*b)^*}$ και περιγράφει συμβολοσειρές που αρχίζουν με a και τελειώνουν με b (και επίσης δεν περιέχουν bb). Δηλαδή, τα δύο σύνολα συμβολοσειρών διαφέρουν.