Σειρά Προβλημάτων 3 – Λύσεις

Άσκηση 1

Να δώσετε ασυμφραστικές γραμματικές που να παράγουν τις πιο κάτω γλώσσες:

(α) $\{ x \mid x$ *η τιμή της αριθμητικής έκφρασης* $10^{2n} + 10^n + 1, n \geq 1\}$

(β) $\{ a^i b^j c^k d^m \mid i, j, k, m \geq 0$ *και* $i + j \neq k + m \}$

(γ) $\{ w \mid w \in L(a^* b^{2m} a (a^* c)^m), m \geq 0 \}$

Λύση

(α) Οι λέξεις της γλώσσας έχουν τη μορφή 10^k10^k1, $k \geq 0$. Μια γραμματική για τη γλώσσα έχει ως εξής:

S $\rightarrow$ 1X1

X $\rightarrow$ 0X0 | 1

Η γραμματική τοποθετεί το αρχικό και το τελικό 1 και στη συνέχεια τοποθετεί ζευγάρια από 0 δεξιά και αριστερά από το μεσαίο 1.

(β) Μια γραμματική για τη γλώσσα έχει ως εξής:

S $\rightarrow$ aSd | P | T

P $\rightarrow$ bPd | X | B | D

T $\rightarrow$ aTc | X | A | C

X $\rightarrow$ bXc | B | C

A $\rightarrow$ aA | a | B

B $\rightarrow$ bB | b

C $\rightarrow$ cC | c

D $\rightarrow$ dD | d | C

Η γραμματική «κτίζει» λέξεις από έξω προς τα μέσα δημιουργώντας ζευγάρια από τα δύο άκρα της λέξης προς το κέντρο. Τα ζευγάρια αυτά είναι είτε a-d, είτε a-c, είτε b-c είτε b-d τα οποία τοποθετούνται έτσι ώστε να ικανοποιείται η ζητούμενη σειρά των συμβόλων. Για να ολοκληρωθεί η δημιουργία της λέξης όμως θα πρέπει να εισαχθούν επιπρόσθετα σύμβολα είτε στο τμήμα με τα a και b είτε στο τμήμα με τα c και d, έτσι ώστε να διασφαλιστεί ότι η λέξη έχει τη μορφή $a^ib^jc^kd^m$ και $i + j \neq k + m$.

(γ) Μια γραμματική για τη γλώσσα έχει ως εξής:

S $\rightarrow$ aS | X

X $\rightarrow$ bb X T | a

T $\rightarrow$ aT | c

Κατ' αρχάς, τοποθετούμε τα αρχικά *a*. Στη συνέχεια τοποθετούμε ταυτόχρονα αριστερά και δεξιά τη λέξη bb και μια λέξη της μορφής a*c, αντίστοιχα. Αυτό μπορεί να επαναληφθεί 0 ή περισσότερες φορές και στο τέλος τοποθετούμε το μεσαίο a, σύμφωνα με τον ορισμό της γλώσσας.

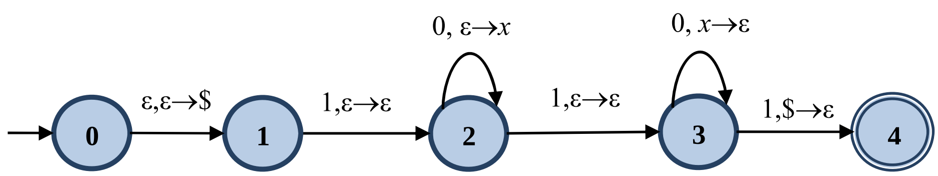
<u>Άσκηση 2</u>

Να κατασκευάσετε αυτόματα στοίβας για τις γλώσσες της Άσκησης 1. *(Να κτίσετε τα αυτόματα κατευθείαν και όχι μέσω μετατροπής των ασυμφραστικών γραμματικών από την Άσκηση 1 σε αυτόματα.)*

<u>Λύση</u>

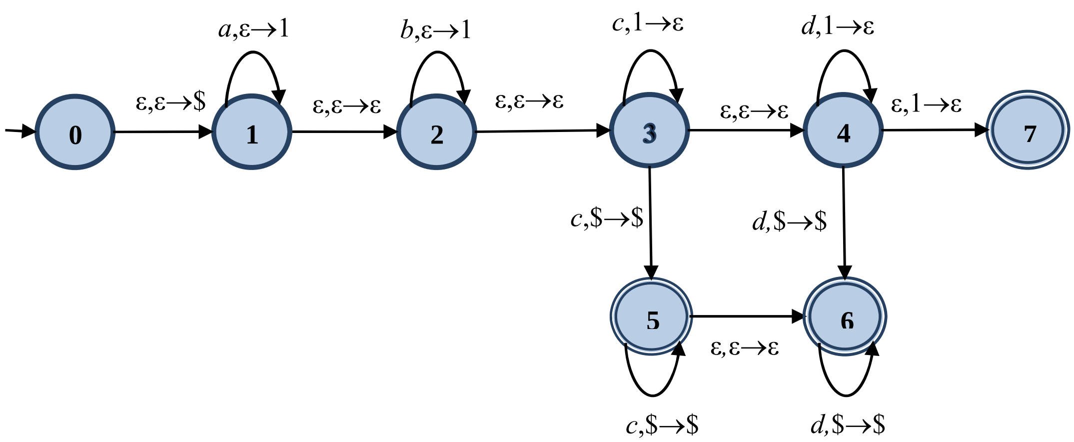
(α) { x | x η τιμή της αριθμητικής έκφρασης $10^{2n} + 10^n + 1$, n ≥ 1}

Ακολουθεί το ζητούμενο αυτόματο. Από την αρχική κατάσταση το αυτόματο γράφει \$. Στη συνέχεια διαβάζει ένα 1 και μηδέν ή περισσότερα 0. Για κάθε 0 που διαβάζει τοποθετεί ένα *x* στη στοίβα. Συνεχίζει διαβάζοντας ένα 1 και στη συνέχεια μηδέν ή περισσότερα 0. Κάθε φορά που διαβάζει ένα 0 στην κατάσταση 3 αφαιρεί ένα x από τη στοίβα. Αποδέχεται αν ακολουθήσει ένα 1 και ταυτόχρονα αδειάσει η στοίβα.



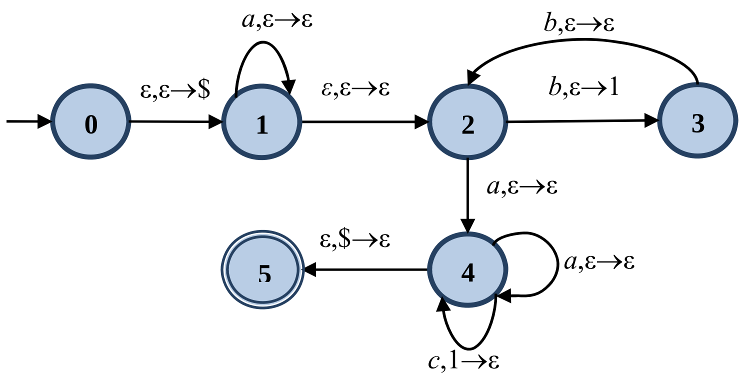
(β) { $a^ib^jc^kd^m$ | i, j, k, m ≥ 0 και i + j ≠ k + m }

Ακολουθεί το ζητούμενο αυτόματο. Από την αρχική κατάσταση το αυτόματο γράφει \$. Στη συνέχεια διαβάζει 0 ή περισσότερα *a* και ακολούθως 0 ή περισσότερα *b*. Για κάθε τέτοιο σύμβολο τοποθετεί ένα 1 μέσα στη στοίβα. Συνεχίζει διαβάζοντας 0 ή περισσότερα *c* και στην συνέχεια 0 ή περισσότερα *d*. Κάθε φορά αφαιρεί ένα 1 από τη στοίβα. Αποδέχεται αν ολοκληρωθεί η ανάγνωση της εισόδου χωρίς να αδειάσει η στοίβα (μετάβαση 4→7) ή αν αδειάσει η στοίβα προτού εξαντληθούν τα c (μετάβαση 3→5) ή και τα d (μετάβαση 4→6).



(γ) { w |w ∈ L($a^*b^{2m}a(a^*c)^m$), m ≥ 0 }

Ακολουθεί το ζητούμενο αυτόματο. Από την αρχική κατάσταση το αυτόματο γράφει \$. Στη συνέχεια διαβάζει μια ακολουθία από a. Με τη συμπλήρωση της ανάγνωσης των a διαβάζει δυάδες από b και για κάθε δυάδα τοποθετεί ένα 1 στη στοίβα. Από την κατάσταση 3 είναι σε θέση να διαβάσει ένα a και στη συνέχεια ακολουθίες της μορφής a*c για κάθε μια από τις οποίες αφαιρεί ένα 1 από τη στοίβα. Όταν αδειάσει η στοίβα, το αυτόματο προχωρεί στην κατάσταση 5 από όπου θα αποδεχθεί τη λέξη.



Άσκηση 3

Θεωρήστε τη γραμματική G = (V, S, P, formula), όπου V = {formula, proposition}, S = { ¬, ∨, (,), True, False } και R οι πιο κάτω κανόνες.

$$\textit{formula} \rightarrow \textit{proposition} \;\; | \;\; \neg\, \textit{formula} \;|\;\; \textit{formula} \vee \textit{formula} \;\; | \;\; (\textit{formula})$$

$$\textit{proposition} \rightarrow \textit{True} \;|\; \textit{False}$$

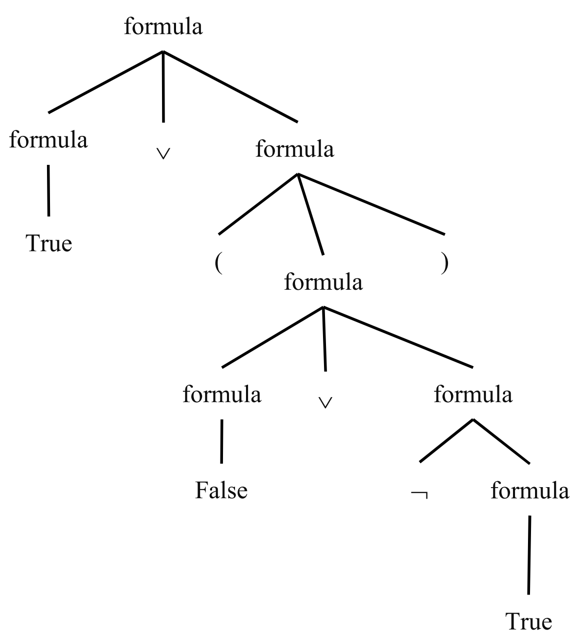
(α) Να κατασκευάσετε παραγωγές και τα αντίστοιχα συντακτικά δέντρα για τις λέξεις:

(i) True ∨(False ∨ ¬True)

(ii) (True ∨ ¬ False) ∨ ¬(True ∨(¬True ∨False) ∨ ¬ True)

Λύση

(i) Συντακτικό Δέντρο:



(ii) Παραλείπεται.

(β) Να εντοπίσετε "πρόταση" φ που να παράγεται από τη γραμματική G μέσω δύο διαφορετικών συντακτικών δέντρων, T_1 και T_2 τέτοια ώστε, αν η φ ερμηνευθεί με βάση το

T_1 τότε θα έχει αποτίμηση True, και αν ερμηνευθεί σύμφωνα με το δέντρο T_2 τότε θα έχει αποτίμηση False.

Λύση

Θεωρούμε τη λέξη $\neg$True $\vee$ True. Η λέξη αυτή θα μπορούσε να ερμηνευθεί τόσο ως την πρόταση ($\neg$True) $\vee$ True (η οποία έχει αποτίμηση True) όσο και την πρόταση $\neg$(True $\vee$ True) (η οποία έχει αποτίμηση False). Οι σχετικές παραγωγές φαίνονται πιο κάτω.

1. formula $\Rightarrow$ formula $\vee$ formula $\Rightarrow$ $\neg$formula $\vee$ formula
   $\Rightarrow$ $\neg$True $\vee$ formula $\Rightarrow$ $\neg$True $\vee$ True
2. formula $\Rightarrow$ $\neg$formula $\Rightarrow$ $\neg$ formula $\vee$ formula
   $\Rightarrow$ $\neg$True $\vee$ formula $\Rightarrow$ $\neg$True $\vee$ True

Επομένως η γραμματική είναι πολύτροπη.

(γ) Να προτείνετε μια καινούρια γραμματική που να παράγει την ίδια γλώσσα με τη G αλλά να είναι μονότροπη. Εξηγήστε σύντομα την απάντησή σας.

Λύση

formula	$\rightarrow$	formula $\vee$ formula | nformula
nformula	$\rightarrow$	$\neg$nformula | pformula
pformula	$\rightarrow$	proposition | (formula)
proposition	$\rightarrow$	True | False

Η πιο πάνω γραμματική εξαναγκάζει, κάθε φορά που θέλουμε να εφαρμόσουμε την άρνηση σε μια σύνθετη πρόταση (π.χ. πρόταση που να περιέχει διάζευξη) να χρησιμοποιούνται παρενθέσεις. Με αυτό τον τρόπο επιτρέπει την παραγωγή 1 από το σκέλος (β) αλλά όχι την παραγωγή 2.

Άσκηση 4

Να δείξετε ότι οι πιο κάτω γλώσσες δεν είναι ασυμφραστικές χρησιμοποιώντας το Λήμμα της Άντλησης για Ασυμφραστικές Γλώσσες.

(α) $\Lambda_1 = \{1010^2 10^3 \ldots 10^{n-1} 10^n 1 \mid n \geq 1\}$

Λύση

Υποθέτουμε για να φτάσουμε σε αντίφαση ότι η Λ_1 είναι ασυμφραστική. Τότε, σύμφωνα με το Λήμμα της Άντλησης, υπάρχει p, το μήκος άντλησης της γλώσσας, τέτοιο ώστε κάθε λέξη της γλώσσας με μήκος μεγαλύτερο από p να ικανοποιεί την ιδιότητα που περιγράφεται στο λήμμα.

Ας επιλέξουμε τη λέξη $w = 1010^2 10^3 \ldots 10^{p-1} 10^p 1$.

Τότε, σύμφωνα με το λήμμα, $w = uvxyz$ έτσι ώστε η υπολέξη vxy περιέχει το πολύ p σύμβολα ($|vxy| \leq p$), τουλάχιστον μία από τις v και y να είναι μη κενή ($|vy| > 0$) και οποιαδήποτε ταυτόχρονη επανάληψη των υπολέξεων v και y να διατηρεί την προκύπτουσα λέξη εντός της γλώσσας ($uv^ixy^iz \in \Lambda_1$, $i \geq 0$).

Διακρίνουμε τις πιο κάτω περιπτώσεις.

- Αν τα *v* και y περιέχουν μόνο 0, τότε αν αντλήσουμε τα τμήματα v και y δύο ή περισσότερες φορές, η λέξη που θα προκύψει δεν θα ανήκει στη γλώσσα αφού το πλήθος των 0 σε κάποια από τα τμήματα της λέξης θα αλλάξει και με αυτό τον τρόπο θα διαταραχθεί η δομή της λέξης.
- Αν τα v και y περιέχουν μόνο 1, τότε αν αντλήσουμε τα δύο τμήματα δύο ή περισσότερες φορές, η λέξη που θα προκύψει δεν θα ανήκει στη γλώσσα αφού τα 1 δεν μπορούν να εμφανίζονται σε ομάδες με περισσότερα από ένα 1.
- Τέλος αν τα v και y περιέχουν και 0 και 1, τότε αν αντλήσουμε τα δύο τμήματα δύο ή περισσότερες φορές, η λέξη που θα προκύψει δεν θα ανήκει στη γλώσσα αφού δύο υπολέξεις της μορφής (0*10*)* θα προστεθούν στη λέξη, αλλοιώνοντας τη δομή της και παραβιάζοντας την ιδιότητα ότι κάθε ομάδα 0* στη λέξη περιέχει ένα περισσότερο 0 από την αμέσως προηγούμενη ομάδα (εκτός από την πρώτη 0* ομάδα) και ένα λιγότερο 0 από την επόμενη (εκτός από την τελευταία 0* ομάδα).

Αυτό μας οδηγεί σε αντίφαση και επομένως η υπόθεσή μας ότι η γλώσσα $Λ_1$ είναι ασυμφραστική ήταν εσφαλμένη.

Συμπέρασμα: Η $Λ_1$ είναι μη ασυμφραστική.

(β) $Λ_2 = \{ b^{2m}a(a^nc)^m \mid m,n \geq 0 \}$

<u>Λύση</u>

Υποθέτουμε για να φτάσουμε σε αντίφαση ότι η $Λ_2$ είναι ασυμφραστική. Τότε, σύμφωνα με το Λήμμα της Άντλησης, υπάρχει *p*, το μήκος άντλησης της γλώσσας, τέτοιο ώστε κάθε λέξη της γλώσσας με μήκος μεγαλύτερο από *p* να ικανοποιεί την ιδιότητα που περιγράφεται στο λήμμα.

Ας επιλέξουμε τη λέξη $w = b^{2p}a(a^pc)^p$ και ας ονομάσουμε τα τμήματα της λέξης ως Α, Β, Γ, όπου $Α=b^{2p}$, Β=a, $Γ = (a^pc)^p$.

Τότε, σύμφωνα με το λήμμα, $w = uvxyz$ έτσι ώστε η υπολέξη *vxy* περιέχει το πολύ *p* σύμβολα ($|vxy| \leq p$), τουλάχιστον μία από τις *v* και *y* να είναι μη κενή ($|vy|> 0$) και οποιαδήποτε ταυτόχρονη επανάληψη των υπολέξεων *v* και *y* να διατηρεί την προκύπτουσα λέξη εντός της γλώσσας ($uv^ixy^iz \in Λ_2, i \geq 0$).

Αφού $|vxy| \leq p$, τότε η λέξη αυτή δεν μπορεί να εκτείνεται σε περισσότερα από δύο τμήματα της λέξης. Διακρίνουμε τις πιο κάτω περιπτώσεις.

- Αν η *v*xy εκτείνεται μόνο στο τμήμα Α, τότε τα v και y θα αποτελούνται από b. Επομένως, αν αντλήσουμε τα τμήματα v και y, η λέξη που θα προκύψει δεν θα ανήκει στη γλώσσα: $uv^2xy^2z = b^{2p+\mu+\nu}a(a^pc)^p \notin Λ_2$, για μ = |v|, λ = |y|.
- Αν η *v*xy εκτείνεται στα τμήματα Α και Β, τότε αν αντλήσουμε τα τμήματα v και y δύο ή περισσότερες φορές, η λέξη που θα προκύψει δεν θα ανήκει στη γλώσσα αφού είτε θα χαλάσει η δομή της λέξης με επανάληψη του συμβόλου a είτε θα αυξηθούν τα b με αποτέλεσμα να διαταραχθεί η σχέση ανάμεσα στα b^2 και τα a^pc.
- Συμμετρικά, αν η *v*xy εκτείνεται στα τμήματα Β και Γ, τότε αν αντλήσουμε τα τμήματα v και y δύο ή περισσότερες φορές, η λέξη που θα προκύψει δεν θα ανήκει στη γλώσσα αφού είτε θα χαλάσει η δομή της λέξης με επανάληψη του συμβόλου a είτε θα επηρεαστούν μέχρι και δύο εμφανίσεις της συμβολοσειράς a^pc με αποτέλεσμα η λέξη που θα προκύψει να μην ανήκει στη γλώσσα.
- Τέλος, αν η *v*xy εκτείνεται μόνο στο τμήμα Γ, τότε αν αντλήσουμε τα τμήματα v και y, δύο ή περισσότερες φορές η λέξη που θα προκύψει δεν θα ανήκει στη γλώσσα αφού

θα επηρεαστούν μέχρι και δύο εμφανίσεις της συμβολοσειράς a^pc με αποτέλεσμα η λέξη να μην περιέχει p επαναλήψεις της συμβολοσειράς a^pc.

Αυτό μας οδηγεί σε αντίφαση και επομένως η υπόθεσή μας ότι η γλώσσα $Λ_2$ είναι ασυμφραστική ήταν εσφαλμένη.

Συμπέρασμα: Η $Λ_2$ είναι μη ασυμφραστική.

(γ) $Λ_3 = \{ a^{n^2+2n} \mid n \geq 0 \}$

Λύση

Υποθέτουμε για να φτάσουμε σε αντίφαση ότι η $Λ_3$ είναι ασυμφραστική. Τότε, σύμφωνα με το Λήμμα της Άντλησης, υπάρχει p, το μήκος άντλησης της γλώσσας, τέτοιο ώστε κάθε λέξη της γλώσσας με μήκος μεγαλύτερο από p να ικανοποιεί την ιδιότητα που περιγράφεται στο λήμμα.

Ας επιλέξουμε τη λέξη $w = a^{p^2+2p}$.

Τότε, σύμφωνα με το λήμμα, $w = uvxyz$ έτσι ώστε η υπολέξη vxy περιέχει το πολύ p σύμβολα ($|vxy| \leq p$), τουλάχιστον μία από τις v και y να είναι μη κενή ($|vy| > 0$) και οποιαδήποτε ταυτόχρονη επανάληψη των υπολέξεων v και y να διατηρεί την προκύπτουσα λέξη εντός της γλώσσας ($uv^ixy^iz \in Λ_3$, $i \geq 0$).

Αφού $|vy| \leq p$, τότε πρέπει να ισχύει ότι $v = a^λ$, $y = a^μ$ και $λ+μ \leq p$.

Επίσης, από το λήμμα, πρέπει να ισχύει ότι $uv^2xy^2z \in Λ_3$.

Έχουμε ότι $uv^2xy^2z = a^{p^2+2p+λ+μ}$.

Για να ανήκει η λέξη στη γλώσσα πρέπει $p^2 + 2p + λ + μ = q^2 + 2q$ για κάποιο ακέραιο q. Αλλά

$p^2 + 2p + λ + μ > p^2 + 2p$ και

$p^2 + 2p + λ + μ < (p + 1)^2 + 2(p + 1) = p^2 + 4p + 3$ για τον λόγο ότι $λ+μ \leq p$.

Αφού η ποσότητα $p^2 + 2p + λ + μ$ βρίσκεται ανάμεσα στις δύο τιμές, και δεν ισούται με καμιά από αυτές, είναι αδύνατο να υπάρχει q τέτοιο ώστε $p^2 + 2p + μ = q^2 + 2q$. Αυτό μας οδηγεί στο συμπέρασμα ότι $uv^2xy^2z \notin Λ_3$.

Αυτό μας οδηγεί σε αντίφαση και επομένως η υπόθεσή μας ότι η γλώσσα $Λ_3$ είναι ασυμφραστική ήταν εσφαλμένη.

Συμπέρασμα: Η $Λ_3$ είναι μη ασυμφραστική.

Άσκηση 5

(α) Έστω μια ασυμφραστική γλώσσα Α και μια κανονική γλώσσα Β. Να δείξετε ότι η γλώσσα ΑΒ είναι ασυμφραστική. Να αποφασίσετε κατά πόσο και η γλώσσα ΒΑ είναι ασυμφραστική.

Λύση

(α) Η απόδειξη είναι κατασκευαστική. Συγκεκριμένα θα δείξουμε ότι αν υπάρχει DFA που αναγνωρίζει τη γλώσσα Α και PDA που αναγνωρίζει τη γλώσσα Β τότε υπάρχει PDA που αναγνωρίζει τη γλώσσα ΑΒ.

Ας υποθέσουμε ότι τα αυτόματα $M_1 = (Q_1, \Sigma, \delta_1, q_1, F_1)$ και $M_2 = (Q_2, \Sigma, \Gamma, \delta_2, q_2, F_2)$ αναγνωρίζουν τις γλώσσες A και B αντίστοιχα. Κατασκευάζουμε το αυτόματο $M = (Q, \Sigma, \Gamma$ $\delta, q_0, F)$ που αναγνωρίζει τη γλώσσα AB ως εξής:

- $Q = Q_1 \cup Q_2$
- Σ: το αλφάβητο είναι το ίδιο με αυτό των M_1 και M_2.
- Για κάθε $r \in Q$, $a \in \Sigma$ και $b \in \Gamma$, θέτουμε

$$\delta(s,a,b) = \begin{cases} \{(\delta_1(s,a),\varepsilon)\}, & \alpha\nu\ s \in Q_1 - F_1\ \kappa\alpha\iota\ b = \varepsilon \\ \{(\delta_1(s,a),\varepsilon),(q_2,\varepsilon)\}, & \alpha\nu\ s \in F_1\ \kappa\alpha\iota\ b = \varepsilon \\ \delta_2(s,a,b), & \alpha\nu\ s \in Q_2 \\ \emptyset, & \delta\iota\alpha\varphi o\rho\varepsilon\tau\iota\kappa\acute{\alpha} \end{cases}$$

- $q_0 = q_1$
- $F = F_2$

Με λόγια, το αυτόματο που έχουμε κατασκευάσει συναρμόζει το αυτόματο M_2 μετά από το αυτόματο M_1 τοποθετώντας ε-μεταβάσεις από τις τελικές καταστάσεις του M_1 προς την αρχική κατάσταση του M_2. Επιπρόσθετα οι μεταβάσεις στο αυτόματο M_1 διατυπώνονται σε ορολογία PDA αγνοώντας την ύπαρξη της στοίβας. Επομένως, το αυτόματο χρησιμοποιεί τη στοίβα αποκλειστικά κατά την επεξεργασία του δεύτερου αυτόματου, που είναι το μόνο αυτόματο που την χρειάζεται. Το αυτόματο αποδέχεται μια λέξη εφόσον η λέξη μπορεί να χωριστεί σε δύο τμήματα όπου το πρώτο τμήμα γίνεται αποδεκτό από το αυτόματο M_1 και το δεύτερο τμήμα γίνεται αποδεκτό από το αυτόματο M_2 όπως είναι και το ζητούμενο της άσκησης.

Μπορούμε να δείξουμε ότι και η γλώσσα BA είναι ασυμφραστική αντιστρέφοντας την πιο πάνω κατασκευή και τοποθετώντας το αυτόματο M_1 μετά από το αυτόματο M_2.

(β) Η πλήρης σύμμειξη δύο γλωσσών A και B είναι η γλώσσα

Σύμμειξη(A,B) $= \{ w \mid w = a_1b_1 \ldots a_kb_k,$ *όπου* $a_1 \ldots a_k \in A,\ b_1 \ldots b_k \in B$ *και κάθε* $a_i, b_i \in \Sigma\}$

Να δείξετε ότι η κλάση των ασυμφραστικών γλωσσών δεν είναι κλειστή ως προς την πλήρη σύμμειξη.

<u>Λύση</u>

Μπορείτε να βρείτε τη λύση της άσκησης στη σελίδα 154 στο βιβλίο του μαθήματος (Michael Sipser, *Εισαγωγή στην θεωρία υπολογισμού*) ως λύση της Άσκησης 2.38 του βιβλίου.