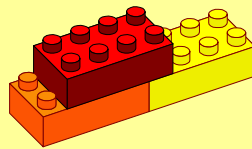


HY-280

«ΘΕΩΡΙΑ ΥΠΟΛΟΓΙΣΜΟΥ»

θεμελιακές έννοιες της επιστήμης του υπολογισμού



ΑΣΚΗΣΕΙΣ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

Γεώργιος Φ.Δ. Γεωργακόπουλος

μέρος Α΄

Εισαγωγή, και η βασική θεωρία των πεπερασμένων αυτομάτων.

(#1) ΟΜΑΛΕΣ ΓΡΑΜΜΑΤΙΚΕΣ.

Το εάν μια γραμματική είναι ομαλή ή όχι, φαίνεται στην *μορφή* της, στον τρόπο με τον οποίο είναι γραμμένη: πρέπει (i) τα παραγωγικά σύμβολα («κεφαλαία») να μετατρέπονται μόνον σε μια λέξη λ από τερματικά σύμβολα («πεζά»), ακολουθούμενη από ακριβώς ένα παραγωγικό σύμβολο, ή κανένα.

Ποιές από τις εξής γραμματικές είναι ομαλές;

- 1) $I \rightarrow \alpha X \mid \emptyset, \quad X \rightarrow \alpha \beta \gamma, \quad \gamma \rightarrow \beta \alpha X X \mid I.$
- 2) $I \rightarrow 01\Lambda \mid \emptyset, \quad X \rightarrow 0I, \quad \gamma \rightarrow 00X \mid \emptyset, \quad Z \rightarrow 1I \mid X$
- 3) $I \rightarrow \beta \alpha M \mid \emptyset, \quad X \rightarrow I \alpha \gamma \alpha \mid \emptyset, \quad \gamma \rightarrow \gamma \gamma Z \mid I, \quad Z \rightarrow \alpha \beta \mid X$
- 4) $I \rightarrow \alpha \alpha \gamma, \quad X \rightarrow \gamma \mid \alpha Z, \quad \alpha \gamma \rightarrow \alpha I \mid X, \quad Z \rightarrow I \mid \alpha \alpha \Lambda.$

ΠΡΟΣ ΠΑΡΑΔΟΣΗ (#1): (4 από (1, 2, 3, 4)).

(#2) ΟΜΑΛΕΣ ΓΛΩΣΣΕΣ ΚΑΙ ΚΛΕΙΣΤΟΤΗΤΑ.

Το εάν μια γλώσσα είναι ομαλή ή όχι, εξαρτάται από το εάν μπορούμε ή όχι να βρούμε μια ομαλή γραμματική που να την παράγει, (ή ένα διάγραμμα μεταβάσεων, ή ένα πεπερασμένο αυτόματο). Στις απλές περιπτώσεις αυτό το κάνουμε «κατ' ευθείαν». Στις πιο σύνθετες χρησιμοποιούμε τους κανόνες κλειστότητας.

Δείξατε με χρήση των κανόνων κλειστότητας ότι οι εξής γλώσσες είναι ομαλές:

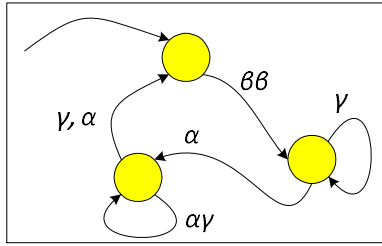
- 1) $HL = \{ \lambda \in \{\alpha, \beta\}^* : \text{η λέξη } \lambda \text{ αν περιέχει το υπόλεκτο «}\beta\beta\text{» τότε περιέχει και το «}\alpha\alpha\alpha\text{»} \}$.
- 2) $HL = \{ \lambda \in \{\alpha, \beta\}^* : \text{η λέξη } \lambda \text{ περιέχει λιγότερα από 5 «}\alpha\text{» και περιττό πλήθος από «}\beta\text{»} \}$.
- 3) Δείξατε ότι εάν η γλώσσα L είναι ομαλή, τότε είναι ομαλή και η γλώσσα που περιέχει όλες και μόνον τις λέξεις της L με άρτιο μήκος (δηλαδή τις λ με $|\lambda| = 2k$, για κάποιο φυσικό αριθμό k).
- 4) Ένα έτος (λ.χ. το 1864) θεωρείται δίσεκτο, εάν διαιρείται δια 4, εκτός εάν διαιρείται δια 100, αλλά και πάλι η εξαίρεση δεν ισχύει εάν διαιρείται δια 400, (λ.χ. το έτος 2000). Η γλώσσα επί του $\Sigma = \{0..9\}$ των «δίσεκτων» ετών είναι ομαλή ή όχι;

ΠΡΟΣ ΠΑΡΑΔΟΣΗ (#2): (2 από (1, 2, 3, 4)).

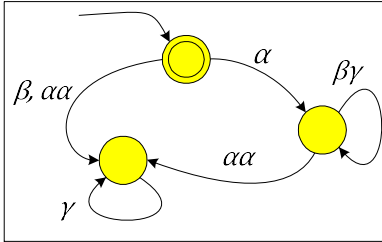
(#3) ΔΙΑΓΡΑΜΜΑΤΑ ΜΕΤΑΒΑΣΕΩΝ:

Σε απλές περιπτώσεις μπορούμε να σχεδιάσουμε ένα διάγραμμα μετάβασης ώστε να δέχεται μια απλή γλώσσα. Σε απλές περιπτώσεις μπορούμε επίσης να αναγνώσουμε ένα διάγραμμα μετάβασης, και να καταλάβουμε τί γλώσσα αποδέχεται.

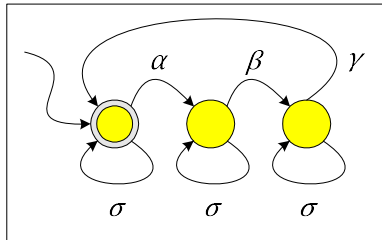
- 1) Σχεδιάστε αιτιοκρατικό αυτόματο (με διάγραμμα μεταβάσεων) που αποδέχεται την εξής γλώσσα, και δώστε μια σύντομη εξήγηση γιατί δουλεύει σωστά:
 $\{ \lambda \in \{\alpha, \beta\}^* : \text{η λέξη } \lambda \text{ δεν περιέχει ούτε το υπόλεκτο «}\alpha\alpha\text{», ούτε το «}\beta\beta\text{»} \}$.
- 2) Σχεδιάστε αιτιοκρατικό αυτόματο (με διάγραμμα μεταβάσεων) που αποδέχεται την εξής γλώσσα, και δώστε μια σύντομη εξήγηση γιατί δουλεύει σωστά:
 $\{ \lambda \in \{\alpha, \beta\}^* : \text{η λέξη } \lambda \text{ περιέχει άρτιο πλήθος από } \alpha \text{ και περιττό από } \beta \}$
- 3) Περιγράψτε, με απλά και σαφή Ελληνικά, το ποιά γλώσσα επί του $\Sigma = \{ \alpha, \beta \}$ αποδέχεται το εξής διάγραμμα μεταβάσεων (αυτόματο):



- 4) Περιγράψτε, με απλά και σαφή Ελληνικά, το ποιά γλώσσα επί του $\Sigma = \{ \alpha, \beta, \gamma \}$ αποδέχεται το εξής διάγραμμα μεταβάσεων (αυτόματο):



- 5) Περιγράψτε, με απλά και σαφή Ελληνικά, το ποιά γλώσσα επί του $\Sigma = \{ \alpha, \beta, \gamma, \sigma \}$ αποδέχεται το εξής διάγραμμα μεταβάσεων (αυτόματο):

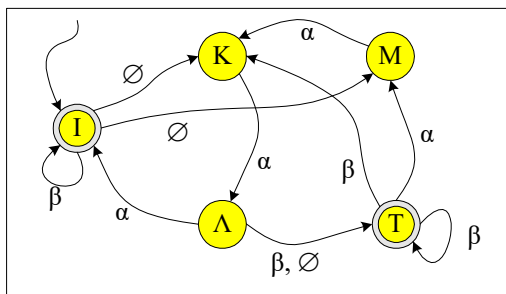


- 6) Έστω ένα διάγραμμα μεταβάσεων M όπου κάθε μετάβαση δεν επιγράφεται με μία λέξη λ , αλλά με ένα ζεύγος λέξεων $\langle \lambda, \mu \rangle$, η ερμηνεία του οποίου είναι: «εάν διαβάσεις λ τότε γράψε μ ». Αν το διάγραμμα διαβάσει μια λέξη d από το Σ^* θα γράψει μια άλλη λέξη του Σ^* . (Ως εάν το M να «μεταφράζει» την d). Έστω $\Gamma(M)$ όλες οι λέξεις που μπορεί να γράψει το M αρχίζοντας από την αφετηριακή και καταλήγοντας (έχοντας διαβάσει όλο το d) σε αποδεκτική κατάσταση. Δείξατε ότι η γλώσσα $\Gamma(M)$ είναι ομαλή.

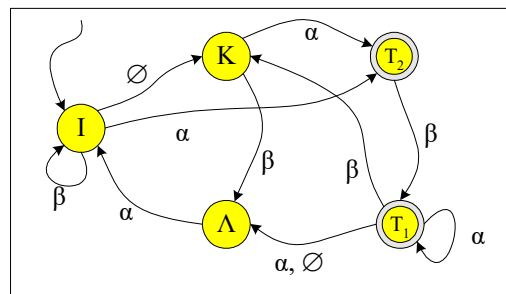
ΠΡΟΣ ΠΑΡΑΔΟΣΗ (#3): (1 από (1, 2)), (2 από (3, 4, 5)). Η 6^η προαιρετικά ως bonus.

(#4) ΔΙΑΓΝΩΣΗ – ΑΠΟΔΟΧΗ:

Για την αποδοχή μιας λέξης από ένα αυτόματο αρκεί ένας αποδεκτικός περίπατος στο αντίστοιχο διάγραμμα μεταβάσεων. Πώς ανακαλύπτουμε κάτι τέτοιο; Σε απλές περιπτώσεις αρκεί η παρατήρηση. Σε πιο σύνθετες θα πρέπει να εφαρμόσουμε την σχετική μέθοδο, (σημειώσεις ενότητα 6^η και 10^η, βλ. και το επόμενο (#5)).



(α)



(β)

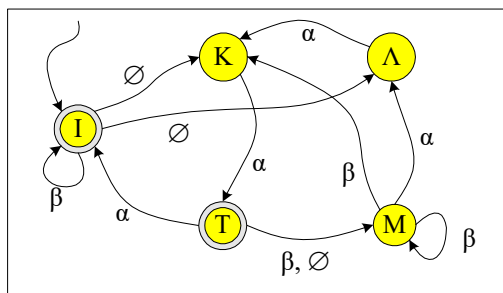
- 1) Ελέγξατε εάν το αυτόματο αριστερά, αποδέχεται τη λέξη $\lambda = \beta \beta \alpha \alpha$.
- 2) Ελέγξατε εάν το αυτόματο δεξιά, αποδέχεται τη λέξη $\lambda = \alpha \beta \alpha \alpha \beta$.

ΠΡΟΣ ΠΑΡΑΔΟΣΗ (#4): (2 από (1, 2)).

(#5) ΔΙΑΓΝΩΣΗ – ΑΠΟΡΡΙΨΗ:

Για να δείξουμε ότι μια λέξη *_δεν_* ανήκει σε μια ομαλή γλώσσα πρέπει να πιστοποιήσουμε ότι δεν υπάρχει καμία διαδρομή από την αφετηριακή κατάσταση σε μια αποδεκτική κατάσταση η οποία σχηματίζει αυτή την λέξη. Αυτό σημαίνει ότι πρέπει να εξετάσουμε – μέ κάποιο βολικό και σωστό τρόπο – όλες τις δυνατές διαδρομές για αυτή την λέξη. Αυτό είναι πολύ εύκολο στα αιτιοκρατικά αυτόματα, αλλά όχι στα μη-αιτιοκρατικά. Για τα δεύτερα ένας τρόπος είναι το να σημειώνουμε σε κάθε βήμα όλες τις καταστάσεις στις οποίες θα μπορούσαμε να είχαμε βρεθεί διαβάζοντας άλλο ένα γράμμα από την λέξη. Αν – στο τέλος – θα μπορούσαμε να είχαμε φθάσει σε έστω μία αποδεκτική κατάσταση τότε – και μόνον τότε – η λέξη ανήκει στη γλώσσα.

- 1) Δείξτε ότι η λέξη $\alpha\alpha\beta\alpha\beta$ δεν γίνεται δεκτή από το παρακάτω αυτόματο.
- 2) Δείξτε ότι η λέξη $\beta\beta\beta\alpha\beta$ δεν γίνεται δεκτή από το παρακάτω αυτόματο.



ΠΡΟΣ ΠΑΡΑΔΟΣΗ (#5): (2 από (1, 2)).

(#6) ΛΗΜΜΑ ΑΝΤΛΗΣΗΣ – ΑΝΩΜΑΛΕΣ ΓΛΩΣΣΕΣ

Για να δείξουμε ότι μια γλώσσα Λ είναι ομαλή αρκεί να βρούμε έστω μία ομαλή γραμματική Γ που να την παράγει ακριβώς. Για να δείξουμε όμως ότι *_δεν_* είναι ομαλή πρέπει να δείξουμε ότι *_καμία_* ομαλή γραμματική Γ δεν την παράγει ακριβώς, ότι δηλαδή η Γ ,

- είτε (i) προσθέτει λέξεις που δεν έχει η Λ ,
- είτε (ii) δεν παράγει λέξεις που έχει η Λ .

Το εργαλείο εδώ για το (i) είναι το λήμμα άντλησης: δείχνουμε ότι (λόγω απείρου πλήθους λέξεων) αν η Λ ήταν ομαλή τότε θα έπρεπε να περιέχει (μεταξύ άλλων) και όλες λέξεις της μορφής $x\mu^{(k)}y$, για κάποιες συγκεκριμένες σταθερές λέξεις $x, \mu, y, \mu \neq \emptyset, k = 0, 1, 2, 3, \dots$ και δείχνουμε ότι (οποιαδήποτε μορφή και εάν είχαν τα x, μ, y), για κάποιο k θα συνέβαινε $x\mu^{(k)}y \notin \Lambda$ – πράγμα που οδηγεί σε άτοπο. Προσέξτε ότι μερικές φορές βολεύει να δείξουμε όχι ότι η Λ είναι ανώμαλη, αλλά ότι ένα κομμάτι της Λ' είναι ανώμαλο – τέτοιο όμως που θα έπρεπε να ήταν ομαλό εάν ήταν και η Λ .

Εξηγήσατε την μη-ομαλότητα των γλωσσών στις εξής περιπτώσεις:

- 1) $L = \{ \lambda \in \{\alpha, \beta\}^* \mid (\lambda^R : \text{η λέξη } \lambda \text{ γραμμένη «ανάποδα» (ή κατοπτρικά)}) \}$
- 2) $L = \{ \lambda \in \{ \times \}^* \mid (\text{πλήθος «}\times\text{» στην } \lambda) = i \cdot j, \text{ όπου } i, j \text{ φυσικοί } \geq 2 \}$
- 3) $L = \{ \lambda \in \{ 0, 1 \}^* \mid (\text{πλήθος «}0\text{» στην } \lambda) = (\text{πλήθος «}1\text{» στην } \lambda)^2 \}$

ΠΡΟΣ ΠΑΡΑΔΟΣΗ (#6): (2 από (1, 2, 3)).

(#7) ΟΜΑΛΕΣ ΕΚΦΡΑΣΕΙΣ:

Η περιγραφή μιας (ομαλής) γλώσσας είναι πολλές φορές προτιμότερο να γίνεται μέσα από μια ομαλή έκφραση, διότι αυτές περιγράφουν με αρκετά φυσικό τρόπο το πώς «γράφεται» μια λέξη της γλώσσας.

- 1) Σχεδιάστε ένα διάγραμμα μεταβάσεων (όχι κατ' ανάγκην αιτιοκρατικό) που δέχεται την γλώσσα από την εξής ομαλή έκφραση:

$$R = (\{ \{ 0, 1 \}^*, \{ \emptyset, 1^* \} \})^*$$

- 2) Δώστε μια ομαλή έκφραση που να παράγει ακριβώς την γλώσσα των «κλασμάτων», λ.χ. '+ 35 / 56' ή '2 / 3' ή '-58 / 9'.
- 3) Δώστε μια ομαλή έκφραση που να παράγει όλες και μόνον τις διευθύνσεις «e-mail» της μορφής που χρησιμοποιείται, π.χ. ugram@csd.uoc.gr.
- 4) Έστω Μ μια ομαλή έκφραση που περιγράφει τα ονόματα μεταβλητών που χρησιμοποιούνται σε μια γλώσσα προγραμματισμού. Μια συνάρτηση ή διαδικασία, με τις παραμέτρους της, έχει την εξής μορφή:

$$\mu_0 (\mu_1 , \mu_2 , \dots , \mu_k)$$

όπου τα μ_i είναι λέξεις της γλώσσας $L(M)$ και $k > 0$. Π.χ. $\text{MAX} (Z , X , 23 , Y)$. Δώστε μια ομαλή έκφραση που χρησιμοποιεί την έκφραση Μ (ως έχει, σκέτο Μ), και που παράγει όλες και μόνον τις λέξεις της παραπάνω μορφής.

- 5) Δώστε μια ομαλή έκφραση που να περιγράφει την γλώσσα επί του $\Sigma = \{ \alpha, \beta \}$ η οποία περιέχει και το υπόλεκτο «βαβαβ», και το υπόλεκτο «αββα».

ΠΡΟΣ ΠΑΡΑΔΟΣΗ (#7): (1 από (1)), (1 από (2, 3)), (1 από 4, 5)).

(#8) ΘΕΩΡΗΜΑ MYHILL-NERODE:

Το θ. Myhill-Nerode, (9^ο κεφάλαιο σημειώσεων), είναι βαρεια σημαντικό θεώρημα για τα αυτόματα, διότι μας επιτρέπει, (αναλύοντας την γλώσσα προς αποδοχή), να εκτιμήσουμε ένα κάτω φράγμα για το πλήθος των καταστάσεων που θα χρειαζόταν ένα σχετικό αυτόματο, (δηλαδή το πόσες καταστάσεις είναι αναγκαίο αυτό να έχει κατ' ελάχιστο).

- 1) Δείξατε με βάση το θ. Myhill-Nerode ότι η εξής γλώσσα δεν είναι ομαλή:

$$L = \{ \lambda \in \{ \alpha, \beta \}^* \text{ και } \text{πλήθος «α» στην } \lambda = \text{πλήθος «β» στην } \lambda \}$$

- 2) Αποδείξτε με το θ. Myhill-Nerode ότι η τομή ομαλών γλωσσών είναι ομαλή, ότι δηλαδή εάν οι γλώσσες Γ_1 και Γ_2 έχουν πεπερασμένο πλήθος κλάσεων ισοδυναμίας $\approx_f$ τότε το ίδιο ισχύει και για την $\Gamma_1 \cap \Gamma_2$.
- 3) Έστω η γλώσσα $\Lambda = \{ \text{οι δυαδικές παραστάσεις φυσικών αριθμών που είναι πολλαπλάσια του 5} \}$. Εξηγήστε γιατί δεν υπάρχει «αυτόματο» με λιγότερες από 5 καταστάσεις που την αποδέχεται.

ΠΡΟΣ ΠΑΡΑΔΟΣΗ (#8): (1 από (1, 2, 3)). Οποιαδήποτε πρόσθετη, δεκτή ως bonus.