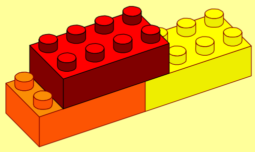
ΗΥ-280

«ΘΕΩΡΙΑ ΥΠΟΛΟΓΙΣΜΟΥ»

θεμελιακές έννοιες της επιστήμης του υπολογισμού



ΑΣΚΗΣΕΙΣ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

Γεώργιος Φρ. Γεωργακόπουλος

μέρος Α΄
Εισαγωγή, και η βασική θεωρία των πεπερασμένων αυτομάτων.

Α′ ΜΕΡΟΣ – ΑΣΚΗΣΕΙΣ

ΠΕΠΕΡΑΣΜΕΝΑ ΑΥΤΟΜΑΤΑ & ΟΜΑΛΕΣ ΓΡΑΜΜΑΤΙΚΕΣ

Τα ερωτήματα είναι απλά και με λυμμένα παραδείγματα. Προσπαθείστε να δώσετε τα περισσότερα που μπορείτε. Βαθμολογικά θα μετρήσουν τα εξής:

- το πόση «ύλη» καλύψατε, (με στόχο πάνω από το 70%, και ένα τουλάχιστον υπο-ερώτημα από κάθε ερώτημα),
- η ποιότητα των απαντήσεων (ζητείται σαφήνεια, ακρίβεια, εξηγήσεις, και «σωστή γλώσσα»/ορολογία).

Η διορία παράδοσης είναι: στο μάθημα της 23ης / 11ου / 2015

(#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.

ΕΡΩΤΗΜΑ-ΠΑΡΑΔΕΙΓΜΑ: Ποιές από τις εξής γραμματικές είναι ομαλές;

Ι → α β Κ ǀ ∅ Κ → Λ ǀ Μ Λ → β β Κ ǀ ∅ Μ → γ Μ ǀ Λ	Ι → α β Κ Κ → Λ ǀ Κ Λ → β γ Κ Μ → γ Μ ǀ Λ	Ι → α β Κ ǀ ∅ αΚ → Λ ǀ Μ Λ → β β Κ ǀ ∅ Μ → α Λ ǀ Ι	Ι → α β Λ ǀ ∅ Κ → α β γ ǀ ∅ Λ → γ γ Κ ǀ Λ Μ → Μ γ β ǀ Λ
(α)	(β)	(γ)	(δ)

ΕΝΔΕΙΚΤΙΚΕΣ ΑΠΑΝΤΗΣΕΙΣ:

(α) Είναι ομαλή: τηρούνται οι κανόνες μορφής: αριστερά του «→» έχουμε =1 παραγωγικό σύμβολο, και δεξιά μια λέξη (λ.χ. 'αβ') και ακριβώς 1 ή 0 παραγωγικά σύμβολα. Προσέξτε ότι ένας κανόνας «Χ → ∅» περιέχει (τρόπος του λέγειν) την κενή λέξη και κανένα (επόμενο) παραγωγικό σύμβολο.

(β) Είναι ομαλή. Προσέξτε ότι δεν μας πειράζει που δεν περιέχει απαλειφές της μορφής Χ → ∅ - απλά η γραμματική δεν παράγει καμμία λέξη – (δηλαδή δεν παράγει *ούτε κάν την κενή λέξη*, δηλαδή *παράγει την κενή γλώσσα*).

(γ) Δεν είναι ομαλή. Ο 2ος κανόνας δεν έχει ένα (και ακριβώς) παραγωγικό σύμβολο αριστερά του «→».

(δ) Δεν είναι ομαλή. Ο 4ος κανόνας έχει τα τερματικά σύμβολα μετά το παραγωγικό. Θα πρέπει να είναι παντού πριν από αυτό (ή παντού μετά, αλλά όχι τα δύο ανάμεικτα).

■

ΑΣΚΗΣΗ ΠΡΟΣ ΠΑΡΑΔΟΣΗ (#1)

(1) Ποιές από τις εξής γραμματικές είναι ομαλές;

Ι → β γ Λ ǀ ∅ Κ → Μ Λ → α α Κ ǀ ∅ Μ → α Ι ǀ Λ	Ι → β β Μ ǀ ∅ Κ → Ι β β γ ǀ ∅ Λ → γ γ Κ ǀ Λ Μ → α β ǀ Κ	Ι → β γ Λ αΚ → Λ ǀ Μ Λ → α Ι ǀ Κ Μ → β β Λ ǀ Ι	Ι → Κ ǀ ∅ Κ → α β γ Κ Λ → γ Κ γ
(1.1)	(1.2)	(1.3)	(1.4)

(#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. Στις απλές περιπτώσεις αυτό το κάνουμε «κατ' ευθείαν». Στις πιο σύνθετες χρησιμοποιούμε τους κανόνες κλειστότητας.

ΕΡΩΤΗΜΑ-ΠΑΡΑΔΕΙΓΜΑ: Δείξτε ότι οι εξής γλώσσες Λ (επί του Σ = {α, β, γ } αν δεν δίδεται κάτι άλλο) είναι ομαλές:

(α) Λ = «όλες οι λέξεις που περιέχουν ως υπόλεκτο το ***ββαβα*** »

(β) Λ = «όλες οι λέξεις που _δεν_ περιέχουν το υπόλεκτο ***βββ*** »

(γ) Λ = «όλες οι λέξεις που _αν_ περιέχουν το ***αβγ*** τότε περιέχουν και το ***βγα*** »

(δ) Λ = «οι λέξεις που έχουν την μορφή μιας σειράς φυσικών δεκαδικών αριθμών (≥ 1), χωρισμένων με παύλες και εγκλεισμένων σε αγκύλες, π.χ. **[23 – 45 – 6 – 12]** »

(ε) Λ = «οι δεκαδικοί γραμμένοι σε 'επιστημονική' γραφή, λ.χ. –0.314159E+2 »

Έστω ότι μια γλώσσα L επί του αλφαβήτου Σ είναι ομαλή, και $M \subseteq \Sigma^*$. Δείξτε ότι και οι εξής γλώσσες θα είναι ομαλές:

(στ) *Πρόθεμα*(L) ≡ { λ: λ ∈ L, και υπάρχει τέλος τ ∈ Σ* για τη λ ώστε λ τ ∈ L }

(ζ) *Maximum*(L) ≡ { λ: λ ∈ L, και για κάθε επέκταση τ ∈ Σ* της λ (τ ≠ ∅), λ τ ∉ L }

ΕΝΔΕΙΚΤΙΚΕΣ ΑΠΑΝΤΗΣΕΙΣ:

(α) Η πιο απλή και θεμελιακή περίπτωση: Λ = Σ* || {ββαβα} || Σ*, δηλαδή η παράθεση τριών προφανώς ομαλών γλωσσών.

(β) Η χρησιμότητα της «άρνησης» ή «συμπληρώματος»: παίρνουμε την γλώσσα που έχει το υπόλεκτο «βββ» και εξ αυτής το συμπλήρωμα: Λ = Σ* – (Σ* {βββ} Σ*).

(γ) Αυτό είναι πιο πονηρό: απαιτεί απλά στοιχεία προτασιακού λογισμού: έχουμε την συνθήκη (P ⇒ Q), όπου P = «περιέχει αβγ», Q = «περιέχει βγα». Επειδή η έκφραση (P ⇒ Q) ισοδυναμεί με την έκφραση (¬P ∨ Q) έχουμε ισοδυνάμως την ένωση [(*δεν* «έχει αβγ») ***ή*** «έχει βγα»]. Έστω οι γλώσσες

$$\Lambda_{\alpha\beta\gamma} = (\Sigma^* \{\alpha\beta\gamma\} \Sigma^*),\ \Lambda_{\beta\gamma\alpha} = (\Sigma^* \{\beta\gamma\alpha\} \Sigma^*),\ \text{και}\ \Lambda = (\Sigma^* - \Lambda_{\alpha\beta\gamma}) \cup \Lambda_{\beta\gamma\alpha}.$$

Οι $\Lambda_{\alpha\beta\gamma}$, $\Lambda_{\beta\gamma\alpha}$ είναι εμφανώς ομαλές και το ζητούμενο για την προκύπτει από την κλειστότητα ως προς συμπλήρωμα «Σ* –», και την ένωση.

Στα παραπάνω εφαρμόζουμε κυρίως τους κανόνες κλειστότητας ως προς τομή, ένωση, συμπλήρωμα, για να αναλύσουμε τις ιδιότητες που έχουν οι λέξεις της γλώσσας. Έχουμε όμως στη διάθεσή μας και τους «γραμματικούς» κανόνες κλειστότητας ως προς παράθεση, επανάληψη και κατοπτρισμό.

(δ) Η γλώσσα Φ των φυσικών δεκαδικών είναι (εύκολα) ομαλή, και μια σειρά σαν την παραπάνω αρχίζει με [Φ (*παράθεση*) συνεχίζει με μια αόριστη *επανάληψη* (– Φ)* και τελειώνει με] . Η όλη γλώσσα είναι ομαλή από την κλειστότητα ως προς παράθεση και αόριστη επανάληψη.

(ε) Οι αριθμοί αυτοί αποτελούνται από 7 μέρη: ένα προαιρετικό πρόσημο, ένα ακέραιο μέρος, μια τελεία, ένα δεκαδικό μέρος, τον εκθέτη «Ε», ένα προαιρετικό πρόσημο, και έναν φυσικό αριθμό. Όλα αυτά εκφράζονται εύκολα με ομαλές γλώσσες και ο όλος αριθμός δίδεται από την παράθεσή τους. Τα «προαιρετικά» μέρη μπορούν να εκφραστούν από την κενή λέξη (και σε επίπεδο αυτομάτου) από κενές μεταβάσεις.

Τέλος, από την τροποποίηση ενός διαγράμματος μεταβάσεων, μπορούμε να πάρουμε και άλλες ομαλές γλώσσες. Τα δύο τελευταία ερωτήματα απαντώνται με αυτόν τον τρόπο.

(στ) Έστω Α ένα αυτόματο που αποδέχεται την Λ, και Τ το σύνολο των αποδεκτικών καταστάσεων του. Ελέγχουμε κάθε κατάσταση Κ για το εάν υπάρχει διαδρομή από αυτήν προς κάποια αποδεκτική κατάσταση εντός του Τ, και εάν ναι την σημειώνουμε. Στο τέλος χαρακτηρίζουμε κάθε σημειωμένη κατάσταση ως αποδεκτική.

(ζ) Έστω Α ένα αυτόματο που αποδέχεται την Λ, και Τ το σύνολο των αποδεκτικών καταστάσεων του. Ελέγχουμε εάν από μια κατάσταση Κ του Τ υπάρχει διαδρομή που οδηγεί σε κατάσταση του Τ (σε άλλη ή ακόμα και στον εαυτό της) και αν ναι τότε την σημειώνουμε. Στο τέλος χαρακτηρίζουμε τις σημειωμένες καταστάσεις ως απορριπτικές.

■

ΑΣΚΗΣΗ ΠΡΟΣ ΠΑΡΑΔΟΣΗ (#2)

Δείξτε ότι οι εξής γλώσσες είναι ομαλές.

(2.1) L = { λ ∈ {α, β}* : η λέξη λ δεν περιέχει το υπόλεκτο βαβ, εκτός, ίσως, αν περιέχει το ααα }.

(2.2) L = { λ ∈ {α, β}* : η λέξη λ περιέχει άρτιο πλήθος από α και περιττό από β }.

(2.3) *Κατάληξη*(L) ≡ { λ: λ ∈ L, και υπάρχει αρχή α ∈ Σ* για τη λ ώστε α λ ∈ L }

(2.4) *L-προ-M* ≡ { λ: λ ∈ L, και υπάρχει τ ∈ M ώστε λ τ ∈ L }

(#3) ΔΙΑΓΡΑΜΜΑΤΑ ΜΕΤΑΒΑΣΕΩΝ:

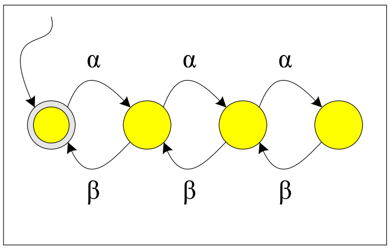
Σε απλές περιπτώσεις μπορούμε να σχεδιάσουμε ένα διάγραμμα μετάβασης ώστε να δέχεται μια απλή γλώσσα. Σε απλές περιπτώσεις μπορούμε επίσης να αναγνώσουμε ένα διάγραμμα μετάβασης, και να καταλάβουμε τί γλώσσα αποδέχεται.

ΕΡΩΤΗΜΑ-ΠΑΡΑΔΕΙΓΜΑ:

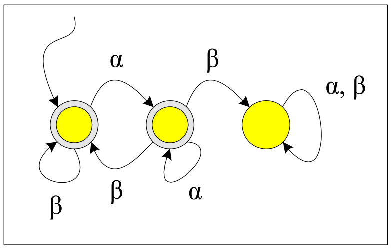
(α) Σχεδιάστε (μέσω διαγραμμάτων μετάβασης) αιτιοκρατικό αυτόματα τέτοια ώστε να δέχονται τις εξής γλώσσες: { λ ∈ {α, β}* : κάθε α στην λέξη λ εμφανίζεται μόνον μετά και πρίν από ένα β }

(β) Περιγράψτε (σε απλά και σαφή Ελληνικά) το ποιές γλώσσες επί του Σ = { α, β } αποδέχονται τα παρακάτω αυτόματα:

(α)



(β)



ΕΝΔΕΙΚΤΙΚΕΣ ΑΠΑΝΤΗΣΕΙΣ:

(α) Οι εμφανίσεις των ‘α’ και ‘β’ ανοίγουν και κλείνουν σαν παρενθέσεις, και δεν έχουμε πάνω από τρία διαδοχικά ‘α’.

(β) Η μετάβαση από την $2^{η}$ κατάσταση στην $3^{η}$ οδηγεί σε απόρριψη, αλλά είναι περιττή. Χωρίς αυτήν και την $3^{η}$ κατάσταση το αυτόματο αποδέχεται κάθε λέξη, αφού σε κατάσταση διαβάζει οποιοδήποτε σύμβολο, και όλες οι καταστάσεις του είναι αποδεκτικές. Το αυτόματο αποδέχεται λοιπόν κάθε λέξη – απλά δεν χρησιμοποιούμε ποτέ την μετάβαση από την $2^{η}$ κατάσταση στην $3^{η}$.

ΑΣΚΗΣΗ ΠΡΟΣ ΠΑΡΑΔΟΣΗ (#3)

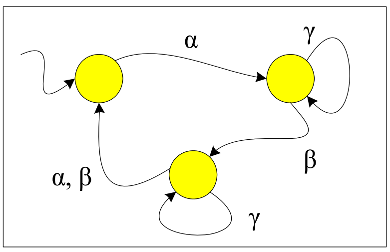
(3.1) Σχεδιάστε (μέσω διαγραμμάτων μετάβασης) αιτιοκρατικά αυτόματα τέτοια ώστε να δέχονται τις εξής γλώσσες:

(3.1.1) { λ ∈ {α, β}* : η λέξη λ δεν περιέχει ούτε την (υπο)λέξη ααα ούτε την ββ }.

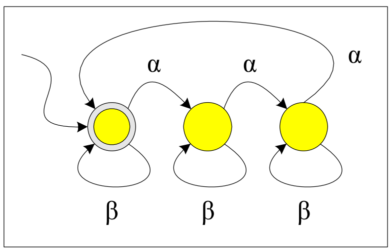
(3.1.2) { λ ∈ {α, β}* : η λέξη λ περιέχει άρτιο πλήθος από α και περιττό από β }

(3.2) Περιγράψτε (σε απλά και σαφή Ελληνικά) το ποιές γλώσσες επί του Σ = { α, β } αποδέχονται τα παρακάτω αυτόματα:

(1)

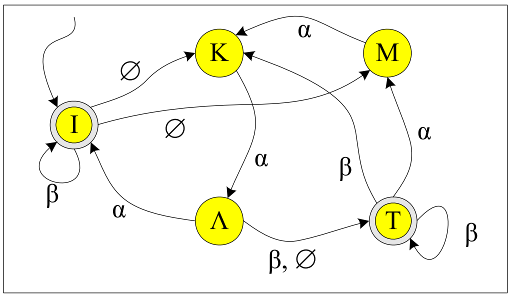


(2)

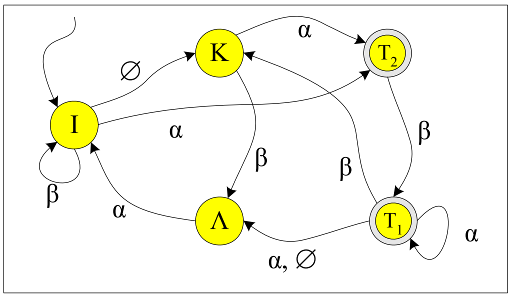


(#4) Διάγνωση – αποδοχή:

Για την αποδοχή μιας λέξης από ένα αυτόματο αρκεί ένας αποδκετικός περίπατος στο αντίστοιχο διάγραμμα μεταβάσεων. Πώς ανακαλύπτουμε κάτι τέτοιο; Σε απλές περιπτώσεις αρκεί η παρατήρηση. Σε πιο σύνθετες θα πρέπει να εφαρμόσουμε την μέθοδο της επόμενης ενότητας (5).



(α)



(β)

ΕΡΩΤΗΜΑ-ΠΑΡΑΔΕΙΓΜΑ:

(α) Ελέγξατε εάν το αυτόματο αριστερά, αποδέχεται τη λέξη λ = ***α α α β β***.

(β) Ελέγξατε εάν το αυτόματο δεξιά, αποδέχεται τη λέξη λ = ***β α α β α***.

ΕΝΔΕΙΚΤΙΚΕΣ ΑΠΑΝΤΗΣΕΙΣ:

(α) Αρκεί η διαδρομή I → M → K → Λ → I → I → I.

(β) Αρκεί η διαδρομή I → K → Λ → I → K → T2 → T1 → T1.

ΑΣΚΗΣΗ ΠΡΟΣ ΠΑΡΑΔΟΣΗ (#4)

(4.1) Ελέγξατε εάν το αυτόματο αριστερά, αποδέχεται τη λέξη λ = ***β β α α α***.

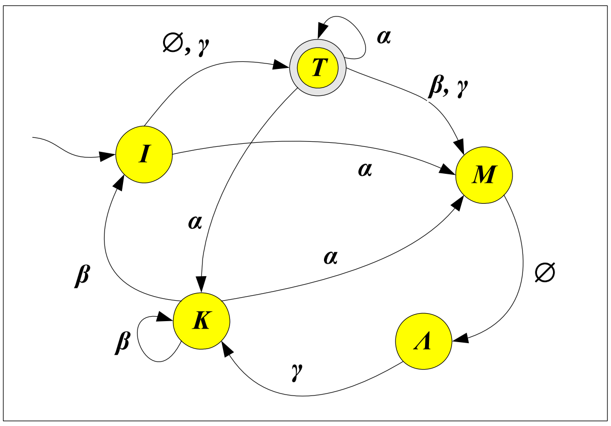
(4.2) Ελέγξατε εάν το αυτόματο δεξιά, αποδέχεται τη λέξη λ = ***α β α α β***.

(#5) ΔΙΑΓΝΩΣΗ – ΑΠΟΡΡΙΨΗ:

Για να δείξουμε ότι μια λέξη _δεν_ ανήκει σε μια ομαλή γλώσσα πρέπει να πιστοποιήσουμε ότι δεν υπάρχει *καμμία 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.

ΕΡΩΤΗΜΑ-ΠΑΡΑΔΕΙΓΜΑ:

Δείξατε ότι το εξής (μη-αιτιοκρατικό) αυτόματο _δεν_ αποδέχεται τη λέξη: ***α α β α β*** .



ΕΝΔΕΙΚΤΙΚΕΣ ΑΠΑΝΤΗΣΕΙΣ:

Για να γίνει αποδεκτή η λέξη *ααβαβ* αρκεί να υπάρχει *έστω ένας* περίπατος από την αφετηριακή κατάσταση I σε μια αποδεκτική κατάσταση (εδώ την T) που να σχηματίζει αυτή τη λέξη. Πρέπει λοιπόν να δείξουμε ότι δεν υπάρχει κανένας τέτοιος περίπατος – άρα θα πρέπει να ελέγξουμε μεθοδικά όλες τις δυνατές διαδρομές που θα μπορούσε να ακολουθήσει το αυτόματο. Προς τούτο υπάρχει απλός και ταχύς συστηματικός τρόπος:

- Σχηματίζουμε έναν πίνακα με τόσες στήλες όσες οι καταστάσεις του αυτομάτου, και τόσες γραμμές όσες μία για την αρχή και από μία για κάθε σύμβολο που θα διαβάσουμε.

- Συμπληρώνουμε την πρώτη γραμμή σημειώνοντας με '✓' τις καστάσεις στις οποίες θα μπορούσε να βρεθεί το αυτόματο αφετηριακά δηλαδή την αφετηριακή και συνεχίζοντας με μία ή περισσότερες κενές μεταβάσεις.
- Συμπληρώνουμε κάθε γραμμή που αντιστοιχεί σε σύμβολο σ, εξετάζοντας το σε ποιές καταστάσεις θα μπορούσε να ήταν το αυτόματο (από την προηγούμενη γραμμή) και σημειώνοντας όσες θα μπορούσε να βρεθεί σε αυτές διαβάζοντας σ – και συνεχίζοντας με μία ή περισσότερες κενές μεταβάσεις.
- Ελέγχουμε εάν στην τελική γραμμή υπάρχει *έστω μία* αποδεκτική κατάσταση.

Στο παράδειγμα θα είχαμε τα εξής:

	Ι	Κ	Λ	Μ	Τ
αρχή	✓				✓
α		✓	✓	✓	✓
α		✓	✓	✓	✓
β	✓	✓	✓	✓	✓
α		✓	✓	✓	✓
γ		✓	✓	✓	

– : Σημειώνουμε την αφετηριακή Ι και την Τ, διότι από Ι πάμε Τ με κενή μετάβαση.

α: Από Ι μέσω 'α' πάμε στην Μ και από εκεί με κενή στην Λ // Από Τ μέσω 'α' στην Τ, και στη Κ. // Σημειώνουμε τις Κ, Λ, Μ, Τ.

α: Από Κ μέσω 'α' στην Μ και από εκεί με κενή/κενές στην Λ // Από Λ μέσω 'α' δεν πάμε πουθενά. // Από Μ μέσω 'α' δεν πάμε πουθενά // Από Τ μέσω 'α' στην Τ και στην Κ Σημειώνουμε τις Κ, Λ, Μ, Τ.

β: Από Κ μέσω 'β' στην Κ και στην Ι, και από εκεί με κενή στην Τ. // Από Λ μέσω 'β' δεν πάμε πουθενά. // Από Μ μέσω 'β' δεν πάμε πουθενά. // Από Τ μέσω 'β' στην Μ και από εκεί με κενή στην Λ. // Σημειώνουμε Ι, Κ, Λ, Μ, Τ (όλες...).

α: Από Ι μέσω 'α' στην Μ και από εκεί με κενή στην Λ. // Από Κ μέσω 'α' στην Μ και από εκεί με κενή την Λ. // Από Λ μέσω 'α' δεν πάμε πουθενά. // Από Μ μέσω 'α' δεν πάμε πουθενά. // Από Τ μέσω 'α' στην Τ και στην Κ. // Σημειώνουμε τις Κ, Λ, Μ, Τ.

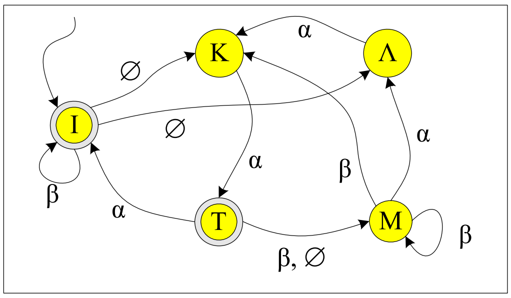
γ: Από Κ μέσω 'γ' δεν πάμε πουθενά. // Από Λ μέσω 'γ' στην Κ. // Από Μ μέσω 'γ' δεν πάμε πουθενά. // Από Τ μέσω 'γ' στην Μ και από εκεί με κενή στην Λ. // Σημειώνουμε τις Κ, Λ, Μ.

Η αποδεκτική κατάσταση Τ δεν περιλαμβάνεται στην σημειωμένες καταστάσεις της τελευταίας γραμμής, άρα όποια διαδρομή και αν ακολουθούσε το αυτόματο δεν θα αποδεχόταν την λέξη ***α α β α γ***.

■

ΑΣΚΗΣΗ ΠΡΟΣ ΠΑΡΑΔΟΣΗ (#5)

(5.1) Δείξτε ότι η λέξη ***ααβαβ*** δεν γίνεται δεκτή από το παρακάτω αυτόματο.
(5.2) Δείξτε ότι η λέξη ***βββαβ*** δεν γίνεται δεκτή από το παρακάτω αυτόματο.



Για να δείξουμε όμως ότι _δεν_ είναι ομαλή πρέπει να δείξουμε ότι _καμμία_ 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 $\boldsymbol{x}\,\boldsymbol{\mu}^{(\kappa)}\,\boldsymbol{y}$, για κάποιες συγκεκριμμένες σταθερές λέξεις $\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{y}$, $\boldsymbol{\mu} \neq \varnothing$, κ = 0, 1 , 2, 3, … και δείχνουμε ότι (*οποιαδήποτε* μορφή και εάν είχαν τα $\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{y}$), για κάποιο κ θα συνέβαινε $\boldsymbol{x}\,\boldsymbol{\mu}^{(\kappa)}\,\boldsymbol{y} \notin \Lambda$ – 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.

ΕΡΩΤΗΜΑ-ΠΑΡΑΔΕΙΓΜΑ: Δείξατε ότι οι εξής γλώσσες δεν είναι ομαλές.
(α) η γλώσσα Λ επί του Σ = { α, β } όπου πλήθος α = πλήθος β.
(β) η γλώσσα Λ επί του Σ = { α, β } όλων των λ για τα οποία $\lambda = \lambda^{(R)}$ (κατοπτρικό).

ΕΝΔΕΙΚΤΙΚΕΣ ΑΠΑΝΤΗΣΕΙΣ:
(α) Η Λ έχει οσοδήποτε μεγάλες λέξεις επομένως, από οποιαδήποτε ομαλή γραμματική και εάν παραγόταν, θα έπρεπε για κάποια συγκεκριμμένα $\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{y}$ όλες οι λέξεις $\boldsymbol{x}\,\boldsymbol{\mu}^{(\kappa)}\,\boldsymbol{y}$, κ ≥ 0, να ανήκουν στην Λ, δηλαδή να έχουν ίσο πλήθος από 'α' και 'β'. Αλλά αυτό είναι δυνατόν! Αρκεί λ.χ. τα x και y να έχουν ίσο πλήθος από 'α' και 'β', και το ίδιο να ισχύει για το 'μ'. Δεν είναι η επανάληψη του μ που μπορεί να κάνει (εδώ) ζημιά στο πλήθος των 'α' και 'β'. Γι' αυτό χρησιμοποιούμε την γλώσσα $\{ \alpha^{(\kappa)} \beta^{(\lambda)} : \kappa, \lambda \geq 0 \}$. Αυτή είναι ομαλή (εύκολα), και η τομή της Λ ' με την Λ είναι η γλώσσα $\{ \alpha^{(\kappa)} \beta^{(\lambda)} : \kappa, \lambda \geq 0, \kappa = \lambda\}$. Αυτή γνωρίζουμε ότι δεν είναι ομαλή – αλλά ας επαναλάβουμε το επιχείρημα:

Η Λ' έχει οσοδήποτε μεγάλες λέξεις, επομένως, από οποιαδήποτε ομαλή γραμματική και εάν παραγόταν, θα έπρεπε για κάποια συγκεκριμμένα $\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{y}$ όλες οι λέξεις $\boldsymbol{x}\,\boldsymbol{\mu}^{(\kappa)}\,\boldsymbol{y}$, κ ≥ 0, να ανήκουν στην Λ, δηλαδή να έχουν ίσο πλήθος από 'α' και 'β', με όλα τα 'α' στην αρχή και όλα τα 'β' στο τέλος. Αλλά αυτό είναι (τώρα) αδύνατον:
(i) αν το μ έχει ίσο πλήθος από 'α' και 'β' τότε η επανάληψη του θα τα ανεμίγνυε και δεν θα είχαμε όλα τα 'α' στην αρχή και όλα τα 'β' στο τέλος.
(ii) αν το μ έχει διαφορετικό πλήθος από 'α' και 'β' τότε η επανάληψη του – έστω και μία φορά – θα κατέστρεφε το ισοζύγιο και δεν θα είχαμε ίσο πλήθος από 'α' και 'β'.
Επομένως η Λ' δεν παράγεται από καμμία ομαλή γραμματική, άρα ούτε και η Λ (διότι τότε και η Λ' θα ήταν ομαλή (από θ. κλειστότητας τομής)).

(β) Έστω Χ η γλώσσα των λέξεων με ένα και μόνον 'β'. Η Χ είναι προφανώς ομαλή. Εάν η Λ ήταν ομαλή θα ήταν και η $\Lambda' = \Lambda \cap X$. Αλλά εάν $\lambda = \lambda^{(R)}$ και η λ έχει ένα ακριβώς 'β', τότε $\lambda = \alpha^{(\kappa)}\beta\alpha^{(\kappa)}$, κ ≥ 0. Γνωρίζουμε ότι αυτή τη γλώσσα δεν είναι ομαλή – αλλά ας επαναλάβουμε το επιχείρημα: Η Λ' έχει οσοδήποτε μεγάλες λέξεις, επομένως, από οποιαδήποτε ομαλή γραμματική και εάν παραγόταν, θα έπρεπε για κάποια συγκεκριμμένα $\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{y}$ <u>όλες</u> οι λέξεις $\boldsymbol{x}\,\boldsymbol{\mu}^{(\kappa)}\,\boldsymbol{y}$, κ ≥ 0, να ανήκουν (*μεταξύ άλλων*) στην Λ, δηλαδή να έχουν ένα 'β' και ίσο πλήθος από 'α' πριν και μετά από το 'β'. Αλλά αυτό είναι αδύνατον.
Έστω ότι η $\lambda = \boldsymbol{x}\,\boldsymbol{\mu}\,\boldsymbol{y}$ ανήκει στην Λ', ότι έχει δηλαδή την μορφή $\lambda = \alpha^{(\kappa)}\beta\alpha^{(\kappa)}$.
(i) αν το μ περιέχει το 'β' τότε η $\boldsymbol{x}\,\boldsymbol{\mu}\,\boldsymbol{\mu}\,\boldsymbol{y}$ θα είχε 2 'β' και η λέξη λ δεν θα ήταν στην Λ'.
(ii) αν το μ δεν περιέχει 'β' τότε ανήκει όλο στο 1ο μέρος πριν το 'β', είτε όλο στο 2ο μέρος, και η επανάληψη του στην λέξη $\boldsymbol{x}\,\boldsymbol{\mu}\,\boldsymbol{\mu}\,\boldsymbol{y}$ θα κατέστρεφε το ισοζύγιο των 'α' ανάμεσα στο 1ο και 2ο μέρος της λέξης λ.
Επομένως η Λ' δεν παράγεται από καμμία ομαλή γραμματική, άρα ούτε και η Λ (διότι τότε και η Λ' θα ήταν ομαλή (από θ. κλειστότητας τομής)).

■

ΑΣΚΗΣΗ ΠΡΟΣ ΠΑΡΑΔΟΣΗ (#6)

(6.1) Το λήμμα άντλησης λέει το εξής:

«εάν η γλώσσα L(A) ενός αιτιοκρατικού αυτομάτου Α έχει άπειρο πλήθος λέξεων, τότε υπάρχουν λέξεις α, μ, τ ∈ Σ*, ώστε κάθε λέξη $α\ μ^{(i)}\ τ$, $i \geq 0$, ανήκει στην L(A).».

Δείξατε ότι το λήμμα μπορεί να ενισχυθεί κατά τον εξής τρόπο:

« ... τότε υπάρχουν λέξεις α, μ, τ ∈ Σ*, ώστε $|α\ τ| \leq |Σ_Κ|$, και κάθε λέξη $α\ μ^{(i)}\ τ$, $i \geq 0$, ανήκει στην L(A).»

(6.2) Δείξατε ότι οι εξής γλώσσες δεν είναι ομαλές. Μπορείτε να χρησιμοποιείστε το (1) είτε το αποδείξατε είτε ή όχι.

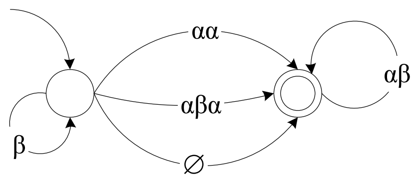
(6.1.1) $L = \{ λ\ λ^R : λ \in \{α, β\}^* \}$ ($λ^R$: η λέξη λ γραμμένη «ανάποδα» (ή κατοπτρικά))

(6.2.2) $L = \{ λ\ λ : λ \in \{α, β\}^* \}$

(#7) ΚΑΝΟΝΙΚΕΣ ΕΚΦΡΑΣΕΙΣ:

ΕΡΩΤΗΜΑ-ΠΑΡΑΔΕΙΓΜΑ: Σχεδιάστε αυτόματο (όχι κατ' ανάγκη αιτιοκρατικό) που δέχεται τη γλώσσα της εξής ομαλής έκφρασης: $R = β^* \{ αα, αβα, \varnothing \} (αβ)^*$

ΕΝΔΕΙΚΤΙΚΕΣ ΑΠΑΝΤΗΣΕΙΣ: Φτιάχνουμε τα επί μέρους αυτόματα, για τις επί μέρους γλώσσες β*, { αα, αβα, ∅ }, (αβ)*, και τα συνδέουμε όπως εξηγείται στην ενότητα περί ομαλών εκφράσεων και κλειστότητας. Στο τέλος πολλές απλοποιήσεις ίσως να είναι εφικτές. Προσέξτε ότι η επανάληψη εκφράζεται με «ανακυκλώσεις», η παράθεση με «αλληλουχία» και η ένωση με «διακλαδώσεις»:



■

ΑΣΚΗΣΗ ΠΡΟΣ ΠΑΡΑΔΟΣΗ (#7)

(7.1) Σχεδιάστε τα αυτόματα (όχι κατ' ανάγκη αιτιοκρατικά) που δέχονται τις γλώσσες των εξής ομαλών εκφράσεων:

(1) $R = (\{ \{α, β\}^*, (\varnothing , γ)^* \})^*$

(2) $R = \{ (βα)^*, (γβ)^+ \} βγ$

(7.2) 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:

$$μ_0\ (\ μ_1\ ,\ μ_2\ ,\ \ldots\ ,\ μ_κ\ )$$

όπου τα $μ_i$ είναι λέξεις της L(M) και κ > 0. Δώστε μια ομαλή έκφραση που να παράγει όλες και μόνον τις λέξεις της παραπάνω μορφής.

(#8) ΘΕΩΡΗΜΑ MYHILL-NERODE:

ΕΡΩΤΗΜΑ-ΠΑΡΑΔΕΙΓΜΑ: «ποιά είναι η κλάση ισοδυναμίας της λέξης λ ως προς την γλώσσα L;»

ΕΝΔΕΙΚΤΙΚΕΣ ΑΠΑΝΤΗΣΕΙΣ: Στην σημειώσεις, στην ενότητα για το θ. Myhill-Nerode, υπάρχει ένα εκτενές παράδειγμα υπολογισμού της κλάσης ισοδυναμίας μιας λέξης λ ως προς μία γλώσσα L, και αρκεί για τα παρακάτω:

ΑΣΚΗΣΗ ΠΡΟΣ ΠΑΡΑΔΟΣΗ (#8)

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

$$L = \{ σ^{(ν)}\ μ\ σ^{(ν)} :\ μ \in Σ^*, σ \in Σ, μ \text{ δεν περιέχει } σ, \text{ και } ν \geq 0 \}.$$

(8.1.2) Η γλώσσα $L = \{ λ \in \{α, β\}^* :$ η λέξη λ δεν περιέχει την (υπο)λέξη *ββα* $\}$ είναι ομαλή γλώσσα.

(1) Ποιές είναι οι κλάσεις ισοδυναμίας της σχέσης $\approx_L$; (Υπόδειξη: εξετάστε το με ποιό τρόπο μπορεί να καταλήγει μια λέξη λ, και το εάν περιέχει ήδη ή όχι την υπολέξη *ββα* .)

(2) Και ως αποτέλεσμα του (1) – ποιό είναι το μικρότερο (σε πλήθος καταστάσεων) αιτιοκρατικό αυτόματο που την αποδέχεται;