ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ

ΑΝΟΙΚΤΑ ΑΚΑΔΗΜΑΙΚΑ ΜΑΘΗΜΑΤΑ

Θεωρία Υπολογισμού

Ενότητα 12: Κανονικότητα ή μη των γλωσσών

Επ. Καθ. Π. Κατσαρός
Τμήμα Πληροφορικής

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Άδειες Χρήσης

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «Ανοικτά Ακαδημαϊκά Μαθήματα στο Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης» έχει χρηματοδοτήσει μόνο την αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

1. Αποδεικνύοντας την κανονικότητα των γλωσσών

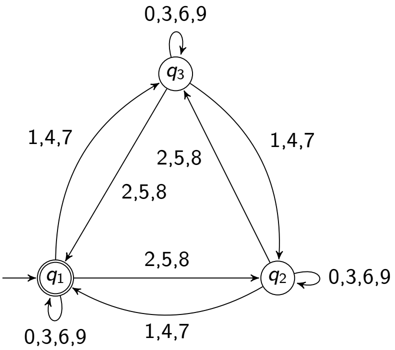
2. Θεώρημα της «άντλησης»
   - Απόδειξη θεωρήματος της «άντλησης»
   - Εφαρμογή του θεωρήματος της «άντλησης»

Αποδεικνύοντας την κανονικότητα (1/3)

- Για ν.δ.ό μια γλώσσα είναι κανονική (ΚΓ) αρκεί ν.δ.ό
  - γίνεται δεκτή από ένα ΠΑ.
  - ή περιγράφεται από μια κανονική έκφραση (ΚΕ)
- Παράδειγμα: Ν.δ.ό το σύνολο $L \subseteq \Sigma^*$ των μη αρνητικών ακεραίων που διαιρούνται με 2 ή 3, ($\Sigma = \{0, .., 9\}$) είναι ΚΓ.
  Απόδειξη.
  - Το σύνολο των μη αρνητικών ακεραίων $L_1 = 0 \cup \{1, .., 9\}\Sigma^*$ περιγράφεται από μία ΚΕ άρα είναι κανονική γλώσσα.
  - Η γλώσσα των μη αρνητικών ακεραίων που διαιρούνται με 2, $L_2 = L_1 \cap \Sigma^*\{0, 2, 4, 6, 8\}$ είναι κανονική (τομή δύο ΚΓ).
  - Έστω L_4 (οι αριθμοί που διαιρούνται με 3), τότε $L_3 = L_1 \cap L_4$ αυτοί που είναι και ακέραιοι.
    - Βοήθεια: Ένας αριθμός διαιρείται με 3 αν και μόνο αν το άθροισμα των ψηφίων των διαιρείται με 3.
    - Η L_4 περιγράφεται από ΠΑ (επόμενη διαφάνεια) που θυμάται το άθροισμα modulo 3 των ψηφίων της συμβολοσειράς
- Τελικά $L = L_2 \cup L_3$ κανονική (ένωση δύο ΚΓ).

Αποδεικνύοντας την κανονικότητα (2/3)

- Το ΠΑ για την L_4



Αποδεικνύοντας την κανονικότητα (3/3)

- Δύο διαισθητικές **ιδιότητες** που ισχύουν για τις κανονικές γλώσσες *και όχι για κάποιες* μη κανονικές:
  - Καθώς μια συμβολοσειρά διαβάζεται το **ποσό της μνήμης** που απαιτείται ώστε να αποφαστιστεί αν ανήκει στη γλώσσα πρέπει να είναι *φραγμένο* και *μη εξαρτώμενο από τη συμβολοσειρά*.
    - Παράδειγμα: Η $\{a^n b^n : n \geq 0\}$ δεν είναι κανονική αφού δεν μπορούμε να φτιάξουμε ένα ΠΑ που να θυμάται πόσα είναι τα a, όταν φτάσει στο σύνορο ανάμεσα στα a και b.
  - Στις κανονικές γλώσσες με άπειρο αριθμό συμβολοσειρών οι συμβολοσειρές έχουν **απλή επαναληπτική δομή** που προκύπτει από την Kleene star στην αντίστοιχη ΚΕ ή από κύκλο στο διάγραμμα καταστάσεων μεταβάσεων ενός ΠΑ.
    - Παράδειγμα: Η $\{a^n : n \geq 1 \text{ είναι πρώτος}\}$ δεν είναι κανονική, καθώς δεν υπάρχει απλή περιοδικότητα στο σύνολο των πρώτων αριθμών.

Θεώρημα της «άντλησης»

Θεώρημα 1 (Θεώρημα της «άντλησης»(pumping lemma»

Έστω L μία κανονική γλώσσα. Υπάρχει ακέραιος $n \geq 1$ τέτοιος ώστε κάθε συμβολοσειρά $w \in L$ με $|w| \geq n$ να μπορεί να γραφτεί διαφορετικά ως $w = xyz$ έτσι ώστε $y \neq e, |xy| \leq n$ και $xy^iz \in L$ για κάθε $i \geq 0$.

- Πολύ σημαντικό θεώρημα για ν.δ.ό μία γλώσσα δεν είναι κανονική.

Απόδειξη θεωρήματος της «άντλησης»

Απόδειξη.

- Εφόσον η L είναι κανονική, γίνεται δεκτή από ένα ΠΑ, έστω M με n καταστάσεις με q_0 την αρχική
  - Για μια w, $|w| \geq n$, και τα πρώτα n σύμβολά της $w_1 w_2..w_n$:
    $$(q_0, w_1 w_2..w_n) \vdash_M (q_1, w_2..w_n) \vdash_M .. \vdash_M (q_n, e)$$
  - Αφού το M έχει μόνο n καταστάσεις και υπάρχουν συνολικά $n+1$ καταστάσεις στον παραπάνω υπολογισμό, από αρχή του περιστεριώνα υπάρχουν $i, j, 0 \leq i \leq j \leq n$, τέτοια ώστε $q_i = q_j$.
    - Δηλαδή η $y = w_i w_{i+1}..w_j$ οδηγεί το M από την q_i ξανά στην q_i και $y \neq e$ αφού $i < j$
    - Τότε ακόμα και αν η y αφαιρεθεί ή προστεθεί επαναληπτικά στην w αμέσως μετά το j-οστό της σύμβολο, πάντα το M θα δέχεται την w.
    - Δηλαδή το M δέχεται τη $xy^i z \in L, \forall i \geq 0$, όπου $x = w_1..w_i$ και $x = w_{j+1}..w_m$

Εφαρμογή του θεωρήματος της «άντλησης» (1/2)

Παράδειγμα 1: Η $\{a^i b^i : i \geq 0\}$ δεν είναι κανονική γιατί αν ήταν κανονική θα μπορούσε να εφαρμοστεί το θεώρημα της άντλησης για κάποιο ακέραιο έστω n.

- Ας θεωρήσουμε τη συμβολοσειρά $w = a^n b^n \in L$.
- Σύμφωνα με το θεώρημα μπορεί να γραφτεί διαφορετικά ως $w = xyz$ έτσι ώστε $|xy| \leq n$ και $y \neq e$, δηλαδή $y = aj$ για κάποιο $j > 0$.
- Τότε όμως $xz = a^{n-j} b^n \notin L$, που έρχεται σε αντίθεση με το θεώρημα.
- Άρα η γλώσσα δεν είναι κανονική.

Εφαρμογή του θεωρήματος της «άντλησης» (2/2)

Παράδειγμα 2: Η γλώσσα $L = \{a^n : n \geq 1 \text{ είναι πρώτος}\}$ δεν είναι κανονική.

- Ας υποθέσουμε ότι είναι, και έστω x, y και z έτσι ώστε $x = a^p, y = a^q$ και $z = a^r$, όπου $p, r \geq 0$ και $q > 0$.
- Σύμφωνα με το θεώρημα, $xy^nz \in L$ για κάθε $n \geq 0$,
- δηλ. $p + nq + r$ είναι πρώτος για κάθε $n \geq 0$.
- Αυτό είναι αδύνατο γιατί έστω $n = p + 2q + r + 2$. Τότε $p + nq + r = (q + 1)(p + 2q + r)$, που είναι γινόμενο φυσικών καθένας από τους οποίους είναι μεγαλύτερος του 1.

Τέλος ενότητας

Επεξεργασία: Εμμανουέλα Στάχτιαρη
Θεσσαλονίκη, 24/07/2014