



**Εθνικό Μετσόβιο Πολυτεχνείο**  
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών  
**Αλγοριθμική Επιστήμη Δεδομένων 2021 – 2022**

Διδάσκοντες: Α. Παγουρτζής, Θ. Σούλιου

**2η Σειρά Ασκήσεων**

**Άσκηση 1.** Έστω ένα σύνολο  $U$ . Μια οικογένεια συναρτήσεων κατακερματισμού  $\mathcal{H} = \{h : U \rightarrow [m]\}$  λέγεται *καθολική* αν

$$\forall x, y \in U, x \neq y : \Pr_{h \in \mathcal{H}} [h(x) = h(y)] \leq \frac{1}{m}$$

(σημείωση: χρησιμοποιούμε τον συμβολισμό  $[m] = \{0, \dots, m-1\}$ )

Ισοδύναμα, για κάθε δύο διαφορετικές τιμές  $x, y \in U$ , υπάρχουν το πολύ  $|\mathcal{H}|/m$  συναρτήσεις  $h \in \mathcal{H}$  για τις οποίες  $h(x) = h(y)$ .

(α) Αποδείξτε ότι για  $a \in [m] \setminus \{0\}, b \in [m]$  η οικογένεια συναρτήσεων  $h_{a,b}(x) = (ax + b) \bmod m$  δεν έχει την ιδιότητα της καθολικότητας για  $U = [m^k], k \geq 2$ .

(β) Αποδείξτε ότι για πρώτο αριθμό  $p > m^k, k \geq 2$  και για  $a \in [p] \setminus \{0\}, b \in [p]$  η οικογένεια συναρτήσεων  $h_{a,b}(x) = ((ax + b) \bmod p) \bmod m$  έχει την ιδιότητα της καθολικότητας για  $U = [m^k]$ .

(γ) Εξακολουθεί να ισχύει η ιδιότητα της καθολικότητας αν στο ερώτημα (β) θέσουμε  $U = [p]$ ;

**Άσκηση 2.** (\*) Εξετάστε την μέθοδο κατακερματισμού ανοιχτής διεθυνσιοδότησης (open addressing) και:

(α) Εξηγήστε γιατί ο μέσος χρόνος επιτυχούς αναζήτησης, μετά από εισαγωγή  $n$  στοιχείων, είναι ίδιος με τον μέσο χρόνο εισαγωγής των στοιχείων στον πίνακα.

(β) Αποδείξτε ότι ο χρόνος αυτός φράσσεται άνω από την ποσότητα  $\frac{1}{\alpha} \ln \frac{1}{1-\alpha}$ , όπου  $\alpha = n/m$  ο παράγοντας φόρτου.

*Υπόδειξη:* ξεκινήστε με μια εκτίμηση για το αναμενόμενο πλήθος δοκιμών κατά την εισαγωγή του  $i$ -οστού στοιχείου υποθέτοντας *uniform hashing*.

**Άσκηση 3.** (α) Περιγράψτε λεπτομερώς, σε μορφή ψευδοκώδικα, τον αλγόριθμο Girvan-Newman του βιβλίου LRU (κεφ. 10.2.3, 10.2.4) για υπολογισμό της τιμής edge betweeness σε ένα γράφημα. Εξηγήστε την πολυπλοκότητα του αλγορίθμου.

(β) Πώς μπορεί να απλοποιηθεί ο αλγόριθμος αν ο γράφος είναι δέντρο; Ποια είναι η πολυπλοκότητα του απλοποιημένου αλγορίθμου;

**Άσκηση 4.** (\*) Να λύσετε τις ασκήσεις **10.4.1** και **10.4.2** από το βιβλίο LRU.

**Άσκηση 5.** Για τον γράφο του Σχήματος 10.16 [LRU], υπολογίστε:

(α) την τιμή της normalized cut, για την διαμέριση  $\{1, 2, 3\}, \{4, 5, 6\}$ .

(β) την τιμή modularity για την παραπάνω διαμέριση, καθώς και για τη διαμέριση  $\{1\}, \{2, 3\}, \{4, 5\}, \{6\}$ .

**Άσκηση 6.** Να λύσετε τις ασκήσεις **8.2.1** και **8.4.1** από το βιβλίο LRU.

**Άσκηση 7.** Να λύσετε τις ασκήσεις **9.2.3** και (\*)**9.3.2** από το βιβλίο LRU.

**Προθεσμία υποβολής και οδηγιές.** Οι απαντήσεις θα πρέπει να υποβληθούν έως τις 27/6/2022, σε ηλεκτρονική μορφή. Οι ασκήσεις με (\*) είναι προαιρετικές.

Για απορίες / διευκρινίσεις: στείλτε μήνυμα στη διεύθυνση [ads@corelab.ntua.gr](mailto:ads@corelab.ntua.gr).