

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ

ΑΣΚΗΣΗ ΔΕΥΤΕΡΗ

Η άσκηση αυτή είναι **ατομική**, δε θα γίνουν δεκτές ασκήσεις από ομάδες φοιτητών.

ΜΑΘΗΜΑ
ΑΚΑΔ. ΕΤΟΣ
ΕΠΙΣΤΡΕΦΕΤΑΙ
ΔΙΔΑΣΚΩΝ

ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ

2014-15

19.1.2015

Ιωάννης Βασιλείου, *Καθηγητής*, Τομέας Τεχνολογίας Πληροφορικής
και Υπολογιστών

Ερώτημα 1.

Θεωρήστε τη σχέση R (A, B, C, D) που περιέχει 1.000.000 εγγραφές, και κάθε σελίδα της σχέσης χωρά 10 εγγραφές. Η R είναι οργανωμένη σε αρχείο σωρού με πυκνά δευτερεύοντα ευρετήρια, και οι εγγραφές της είναι τυχαία ταξινομημένες. Υποθέστε πως το γνώρισμα A είναι υποψήφιο κλειδί της R, με διάστημα τιμών από 0 έως 999.999. Για κάθε ένα από τα παρακάτω αιτήματα, προσδιορίστε το πλήθος των I/O που απαιτούνται για την επεξεργασία του ερωτήματος. Ακολουθούν οι τεχνικές που πρέπει να θεωρήσετε:

- Σάρωση του αρχείου σωρού της R.
- Χρήση ενός ευρετηρίου B+ -δένδρου στο γνώρισμα R.A.
- Χρήση ενός ευρετηρίου κατακερματισμού στο γνώρισμα R.A.

Τα αιτήματα είναι:

1. Βρείτε όλες τις πλειάδες της R ώστε $A < 50$.
2. Βρείτε όλες τις πλειάδες της R ώστε $A = 50$.
3. Βρείτε όλες τις πλειάδες της R. ώστε $A > 50$ και $A < 100$.
4. Βρείτε όλες τις πλειάδες της R ώστε $A \neq 50$.

Ερώτημα 2.

Θεωρήστε μία σχέση R(A,B) οργανωμένη ως κατακερματισμένο αρχείο στο δίσκο. Το κατακερματισμένο αρχείο οργανώνεται σε 1024 blocks (κάδους - buckets) που περιέχουν τις εγγραφές της σχέσης. Για την αποθήκευση μίας εγγραφής (a,b) εφαρμόζουμε πρώτα τη συνάρτηση κατακερματισμού h1 στο πεδίο a λαμβάνοντας X bits. Στη συνέχεια εφαρμόζουμε τη συνάρτηση h2 στο πεδίο b λαμβάνοντας 10-X bits. Τα 10 bits συνολικά ορίζουν τη διεύθυνση του block στο οποίο θα αποθηκευτεί η εγγραφή (a,b).

Θεωρήστε ότι το 20% των ερωτημάτων που αφορούν στη σχέση R είναι της μορφής Q1: SELECT * FROM R WHERE A = a, ενώ το υπόλοιπο 80% είναι της μορφής: Q2: SELECT * FROM R WHERE B = b, όπου a και b είναι σταθερές που δίνονται από τους συντάκτες των ερωτημάτων.

1. Πόσα blocks πρέπει να προσπελαστούν για να απαντηθούν τα ερωτήματα τύπου Q1 και πόσα για τα τύπου Q2; Η απάντησή σας θα είναι συνάρτηση των X bits.
2. Δώστε ένα τύπο που να εκτιμά το μέσο αριθμό blocks που χρειάζεται να προσπελαστούν για την απάντηση ερωτημάτων στη σχέση R.

Ερώτημα 3.

1. Χρησιμοποιείστε τα αξιώματα του Armstrong προκειμένου να αποδείξετε την εγκυρότητα του κανόνα αποσύνθεσης.

$$\text{if } \alpha \rightarrow \beta\gamma, \text{ then } \alpha \rightarrow \beta \text{ and } \alpha \rightarrow \gamma$$

2. Εξετάστε τον παρακάτω προτεινόμενο κανόνα για λειτουργικές εξαρτήσεις: Αν $\alpha \rightarrow \beta$ και $\gamma \rightarrow \beta$, τότε $\alpha \rightarrow \gamma$. Αποδείξτε ότι αυτός ο κανόνας δεν είναι έγκυρος.

Υπόδειξη: Δώστε μια σχέση r που ικανοποιεί το $\alpha \rightarrow \beta$ και το $\gamma \rightarrow \beta$, αλλά δεν ικανοποιεί το $\alpha \rightarrow \gamma$.

Ερώτημα 4.

Δίνεται το σχήμα $R(A, B, C, D, E)$. Δείξτε ότι η παρακάτω αποσύνθεση δεν είναι μία αποσύνθεση χωρίς απώλεια (lossless-join decomposition):

R1 (A, B, C)

R2 (C, D, E).

Υπόδειξη: Δώστε ένα παράδειγμα ενός στιγμιότυπου του R όπου δεν ισχύει η ικανή και αναγκαία συνθήκη για lossless-join decomposition.