

**2ο Σύνολο Ασκήσεων**  
**Ημερομηνία Παράδοσης: στην εξέταση του μαθήματος**

-----  
**Θεματική Ενότητα:** Παράλληλες και Κατανεμημένες Βάσεις Δεδομένων.  
Εξόρυξη Δεδομένων. Βάσεις Διαδικτύου.

**Υλη:** Κεφάλαια 21, 22 και 24 (cow book)

Κεφάλαια 24, 26.2 και 27.1 του βιβλίου των Elmasri & Navathe,  
Τόμος Β'

**Ποσοστό επί του Τελικού Βαθμού:** 15%.

*Οι ασκήσεις θα γίνονται σε ομάδες έως και 2 ατόμων. Αργοπορημένη παράδοση δε γίνεται δεκτή σε καμία περίπτωση.*

-----

**Παράλληλες και Κατανεμημένες Βάσεις Δεδομένων**

1. Θεωρείστε τις σχέσεις  $R(A, B)$ ,  $S(B, C)$  και  $T(C, D)$ . Έστω ότι:

$$S' = S \text{ semijoin } T \quad S'' = S' \text{ semijoin } R''$$

$$R'' = R \text{ semijoin } S' \quad T'' = T \text{ semijoin } S''$$

Οι ακόλουθες εκφράσεις είναι αληθείς; Αποδείξτε τις απαντήσεις σας.

(α)  $R'' = \pi_{AB}(R \text{ join } S \text{ join } T)$

(β) Κάθε εγγραφή στο  $R \text{ join } S'$  συνδέεται (join) με τουλάχιστον μία εγγραφή του  $T$

(γ) Κάθε εγγραφή στο  $T''$  συνδέεται (join) με τουλάχιστον μία εγγραφή στο  $R \text{ join } S$

(δ)  $R \text{ join } S \text{ join } T = R'' \text{ join } S'' \text{ join } T''$

2. Περιγράψτε έναν αλγόριθμο για (α) τον υπολογισμό της συναθροιστικής πράξης count και (β) τον υπολογισμό της συναθροιστικής πράξης count distinct σε έναν παράλληλο σύστημα με  $n$  επεξεργαστές. Για κάθε αλγόριθμο, δώστε μια εκτίμηση του συνολικού χρόνου καθώς και του χρόνου απόκρισης. Ποια είναι η μέγιστη επιτάχυνση και ποια είναι η καλύτερη κλιμάκωση που μπορεί να επιτευχθεί; Εξηγείστε την απάντησή σας.

**Εξόρυξη Δεδομένων**

3. Άσκηση 24.4 του βιβλίου (σελ 1097)

4. Άσκηση 24.7 του βιβλίου (σελ 1098)

**Βάσεις Διαδικτύου**

5. Άσκηση 22.6 του βιβλίου (σελ 1023)

6. Άσκηση 22.7 του βιβλίου (σελ 1024)