

| | |
|-----|--|
| 1 | |
| 2 | |
| 3 | |
| ΣΥΝ | |

ΟΝΟΜΑ
Αρ. Μητρώου

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ

ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ **ΕΞΕΤΑΣΗ 2007-2008 (Μαρτίου)**

ΣΗΜΕΙΩΣΗ: *Οι απαντήσεις δίνονται στα κενά μεταξύ ερωτήσεων !*
Λιδάσκων: **I. Βασιλείου**

ΘΕΜΑ 1.- ΣΧΕΣΙΑΚΟ ΜΟΝΤΕΛΟ ΚΑΙ ΓΛΩΣΣΕΣ [30]

Θεωρήστε την παρακάτω βάση δεδομένων

ΠΙΛΟΤΟΣ(Όνομα, Πόλη, Ώρες-Πτήσης)
ΠΕΤΑΕΙ(Όνομα, Αεροσκάφος, Μισθός)
ΑΝΗΚΕΙ(Αεροσκάφος, Αεροπορική-Εταιρεία, Πόλη-Εγγραφής)

(α) Γράψτε στη γλώσσα SQL τα ερωτήματα (queries). [10]

- i.- Βρες τα ονόματα των πιλότων που πετούνε αεροσκάφη με πόλη εγγραφής την ίδια με αυτή που μένουν.
- ii.- Βρες τα ονόματα των πιλότων που έχουν μεγαλύτερο μισθό από κάθε πιλότο της εταιρείας BLA

(β) [10] Γράψτε σε σχεσιακή άλγεβρα τα ερωτήματα (queries).

- i.- Βρες τα ονόματα των πιλότων που πετούνε αεροσκάφη με πόλη εγγραφής την ίδια με αυτή που μένουν.
 - ii.- Βρες τα ονόματα των πιλότων που έχουν μεγαλύτερο μισθό από 200000 και πετούνε το αεροσκάφος ALB.
-

(γ) [10] Θεωρείστε την παραπάνω Βάση Δεδομένων:

- i. Δώστε ένα παράδειγμα μίας χρήσιμης όψης / προβολής (VIEW) στη σχέση ΠΙΛΟΤΟΣ (χρησιμοποιώντας την SQL), στην οποία επιτρέπεται να γίνουν μεταβολές (ενημερώσεις, διαγραφές, εισαγωγές).
 - ii. Δώστε ένα παράδειγμα μίας χρήσιμης όψης στη σχέση ΠΙΛΟΤΟΣ, στην οποία ΔΕΝ επιτρέπεται να γίνουν ενημερώσεις σε κάποιο γνώρισμα.
 - iii.
-

ΘΕΜΑ 2 – ΓΕΝΙΚΕΣ ΕΡΩΤΗΣΕΙΣ -ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ-ΦΥΣΙΚΟΣ ΣΧΕΔΙΑΣΜΟΣ [35]

(α) [15] Θεωρείστε το Σχήμα R με τα γνωρίσματα (attributes): A, B, C, D, E, G, H και τις λειτουργικές εξαρτήσεις (functional dependencies) που ισχύουν στο R :

$$F = \{ AB \rightarrow C, AC \rightarrow B, AD \rightarrow E, B \rightarrow D, BC \rightarrow A, E \rightarrow G \}$$

Για κάθε ένα από τα παρακάτω σύνολα γνωρισμάτων

- (i) $\{A, B, C\}$
- (ii) $\{A, B, C, E, G\}$
- (iii) $\{A, C, E, H\}$

να κάνετε **τρία πράγματα**:

- (1) Υπολογίστε το σύνολο των εξαρτήσεων που ισχύουν για το συγκεκριμένο σύνολο γνωρισμάτων και διατυπώστε μια (ελάχιστη) κανονική κάλυψη του
- (2) Αναγνωρίστε αν το σύνολο γνωρισμάτων ικανοποιεί την κανονική μορφή BCNF,
- (3) αν το σύνολο δεν είναι σε BCNF, τότε επιδιώξτε μια αποσύνθεση σε ένα σύνολο από BCNF σχήματα που διατηρούν τις εξαρτήσεις.

Υπενθυμίζεται ο ορισμός της κανονικής μορφής BCNF (εάν και μόνο εάν για κάθε $X \rightarrow Y$ που ισχύει τότε ένα εκ των δύο συμβαίνει: είτε το $X \rightarrow Y$ είναι τετριμμένη λειτουργική εξάρτηση ή το X είναι υπερκλειδί)

(β) [15] 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 (Heap) 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 την τεχνική που θα απαιτούσε τις λιγότερες λειτουργίες I/O για την επεξεργασία του ερωτήματος. Ακολουθούν οι τεχνικές που πρέπει να θεωρήσετε:

- Σάρωση του αρχείου σωρού R .
- Χρήση ενός ευρετηρίου B+δέντρου στο γνώρισμα $R.a$.
- Χρήση ενός ευρετηρίου κατακερματισμού στο γνώρισμα $R.a$.

Τα αιτήματα είναι:

- Βρείτε όλες τις πλειάδες της R .
- Βρείτε όλες τις πλειάδες της R ώστε $a < 50$.
- Βρείτε όλες τις πλειάδες της R ώστε $a = 50$.
- Βρείτε όλες τις πλειάδες της R ώστε $a > 50$ και $a < 100$.

(γ) [10] Θεωρείστε την παρακάτω παραλλαγή ενός κλασσικού παραδείγματος βάσης δεδομένων

Emp(eid, did, sal, hobby)
Dept(did, dname, floor, phone)
Finance(did, budget, sales, expenses)

Η σχέση Finance κρατά οικονομικές πληροφορίες ανά τμήμα.

Εστω το ερώτημα (query):

```
SELECT D.dname, F.budget
FROM   Emp E, Dept D, Finance F
WHERE  E.did = D.did AND D.did = F.did AND D.floor = 1 AND E.sal >= 60000
      AND E.hobby = "tennis"
```

- i. Παρουσιάστε ένα δένδρο παράστασης (relational algebra tree) ή μια παράσταση στη σχεσιακή άλγεβρα (relational algebra expression) για το παραπάνω ερώτημα.
- ii. Παρουσιάστε ένα (μετασχηματισμένο) ισοδύναμο δένδρο παράστασης (relational algebra tree) ή μια παράσταση στη σχεσιακή άλγεβρα (relational algebra expression) που κάποιος **στοιχειώδης** optimizer σε ένα DBMS θα κατασκεύαζε.

ΘΕΜΑ 3. -- ΜΟΝΤΕΛΛΟΠΟΙΗΣΗ [35]

Θέλετε να δημιουργήσετε μια βάση δεδομένων για την Βιομηχανία του ΚΙΝΗΜΑΤΟΓΡΑΦΟΥ. Θα πρέπει να αποθηκευτούν πληροφορίες για:

- τις ΤΑΝΙΕΣ που δημιουργούνται (π.χ., Τίτλος, Έτος Παραγωγής, Είδος Ταινίας, Διάρκεια σε λεπτά, κλπ.)
- τους ΗΘΟΠΟΙΟΥΣ (π.χ., όνομα, διεύθυνση, έτος γέννησης,, κλπ)
- τις ΕΤΑΙΡΕΙΕΣ ή ΣΤΟΥΝΤΙΟ στις οποίες ανήκουν οι ταινίες (π.χ., Όνομα, Διεύθυνση, Πρόεδρος, κλπ.)

- i. *Δώστε το διάγραμμα οντοτήτων-συσχετίσεων για τη βάση αυτή. Ποια είναι τα κλειδιά για τη κάθε οντότητα?*
- ii. *Επεκτείνοντας τις πληροφορίες για τη Βάση Δεδομένων, θεωρείστε την έννοια του ΣΥΜΒΟΛΑΙΟΥ μιας εταιρείας με έναν ηθοποιό για μια ταινία . Πως θα το παρουσιάσετε στο παραπάνω διάγραμμα? Σαν μια οντότητα ή μια συσχέτιση (μεταξύ ποιών οντοτήτων)? Πως γίνεται τώρα το διάγραμμα?*
- iii. *Προσπαθήσετε να επεκτείνετε ακόμη περισσότερο τη Βάση Δεδομένων εισάγοντας και νέες (λογικές και χρήσιμες) έννοιες. Για παράδειγμα, πως παρίστανται οι «συνέχειες» (sequel) των ταινιών (π.χ., Άρχοντας των Δακτυλιδιών 1 μετά το 2, κλπ.) Πως μπορούμε να χρησιμοποιήσουμε ISA ιεραρχίες (π.χ., είδη ταινιών), κλπ.*
- iv. *Δώστε την αντίστοιχη περιγραφή (schema) ολόκληρης της Βάσης Δεδομένων) στο σχεσιακό μοντέλο.*

Να γραφούν οι οποιοσδήποτε παραδοχές που κάνετε για τα παραπάνω (π.χ., “μια ταινία ανήκει το πολύ σε ένα στούντιο ”, κλπ.)
