Πανεπιστημιο Ιωαννινων
Τμήμα Πληροφορικής

Βάσεις Δεδομένων
Ακαδημαϊκό Έτος 2003-2004

ΤΕΛΙΚΟ ΔΙΑΓΩΝΙΣΜΑ
ΕΞΕΤΑΣΤΙΚΗ ΙΑΝΟΥΑΡΙΟΥ

Απαντείστε στις ερωτήσεις με *σαφήνεια.*
ΠΡΟΣΟΧΗ!! Απαντήσεις χωρίς εξήγηση δεν θα βαθμολογηθούν.
Καλή Επιτυχία.

Θέμα 1 [Μονάδες 20] Θεωρείστε μια βάση δεδομένων που θα κρατά το πρόγραμμα των αιθουσών για τις εξετάσεις. Κάθε μάθημα μπορεί να έχει ένα ή περισσότερα τμήματα και το όνομα μαθήματος είναι μοναδικό. Θεωρείστε τους ακόλουθους δύο σχεδιασμούς.
Σχεδιασμός Α Υπάρχει μόνο μια οντότητα η Εξέταση με γνωρίσματα Όνομα-Μαθήματος, Τμήμα-Μαθήματος, Αριθμός-Δωματίου και Ώρα.
Σχεδιασμός Β Ορίζουμε επιπρόσθετες οντότητες και συσχετίσεις στη θέση των γνωρισμάτων της οντότητας Εξέταση.

- Μάθημα με γνωρίσματα Όνομα, Τμήμα και Αριθμός-Μαθήματος
- Τμήμα με γνωρίσματα Αριθμός και Αριθμός-Εγγεγραμμένων-Φοιτητών που εξαρτάται ως ασθενής οντότητα από το Μάθημα
- Αίθουσα με γνωρίσματα Αριθμός και Χωρητικότητα

(i) Δώστε το διάγραμμα Οντοτήτων/Συσχετίσεων για τον Σχεδιασμό Β.
(ii) Δώστε τις (μη τετριμμένες) συναρτησιακές εξαρτήσεις που ισχύουν για τον Σχεδιασμό Α. Εξηγείστε.
(iii) Υποθέστε ότι ισχύει ένας νέος κανονισμός που καθορίζει ότι οι εξετάσεις όλων των τμημάτων ενός μαθήματος θα πρέπει να γίνονται στο ίδιο δωμάτιο. Πως αυτό επηρεάζει τις απαντήσεις σας στο (i) και (ii).
(iv) Δώστε το σχεσιακό μοντέλο για το διάγραμμα του ερωτήματος (ii) (χωρίς την αλλαγή του (iii).

Θέμα 2 [Μονάδες 15] Έστω ένα σχήμα σχέσης $R(A, B, C, D)$ όπου το μοναδικό (υποψήφιο) κλειδί είναι το {A, B}. Σε τι κανονική μορφή (3NF, BCNF ή σε καμία) είναι η σχέση αν ισχύει: (Προσοχή: μη ξεχάστε τη συναρτησιακή εξάρτηση που προκύπτει από τον περιορισμό κλειδού).
(i) $C \to D$
(ii) $B \to D$
(iii) $C \to B$
(iv) Θεωρείστε την περίπτωση (iii). Έστω ότι η σχέση διασπάται σε $R_1(B, C)$ και $R_2(A, C, D)$. Η διάσπαση διατηρεί τις εξαρτήσεις; Είναι διάσπαση χωρίς απώλειες;

Θέμα 3 [Μονάδες 35]

(α) Θεωρείστε το παρακάτω σχεσιακό σχήμα για ένα μαγαζί που νοικιάζει αποκριάτικες στολές.

ΠΕΛΑΤΗΣ(<u>Αριθμός_Ταυτότητα</u>, Όνομα)
ΣΤΟΛΗ(<u>Αριθμός_Στολής</u>, Είδος, Μέγεθος, Χρώμα, Τιμή)
ΝΟΙΚΙΑΖΕΙ(<u>Αριθμός_Ταυτότητας</u>, <u>Αριθμός_Στολής</u>, Ημερομηνία_Επιστροφής)

Δώστε σε σχεσιακή άλγεβρα τις παρακάτω ερωτήσεις
(i) Το είδος, χρώμα και ημερομηνία επιστροφής της στολής (στολών) που νοίκιασε ο Βασίλης.
(ii) Έστω ότι τα μεγέθη είναι XL, L, S και M και ότι υπάρχει τουλάχιστον μια στολή για κάθε μέγεθος. Δώστε τα Είδη στολών που υπάρχουν σε όλα τα μεγέθη χωρίς να χρησιμοποιείστε διαίρεση.
(iii) Επαναλάβετε το (ii) χρησιμοποιώντας διαίρεση.
(iv) Όλες τις στολές (τον Αριθμός_Στολής) είδος Κλόουν που είναι διαθέσιμες, δηλαδή δεν έχουν νοικιαστεί.

Δώστε σε SQL τις παρακάτω ερωτήσεις
(v) Το είδος, χρώμα και τιμή της φθηνότερης στολής σε μέγεθος L (αγνοείστε αν έχει ή όχι νοικιαστεί).
(vi) Τον αριθμό (πλήθος) για κάθε είδος στολής, για τις οποίες υπάρχουν τουλάχιστον 5 στολές μεγέθους M με τιμή μικρότερη των 100 (αγνοείστε αν έχει ή όχι νοικιαστεί).

(β) Έστω η σχέση $R(A, B)$ που έχει r πλειάδες και η σχέση $S(A, C)$ που έχει s πλειάδες, όπου r, $s > 0$. Δώστε το μικρότερο και το μεγαλύτερο αριθμό πλειάδων για το $R * S$ (όπου *, φυσική συνένωση) και για το $\pi_A(R) \cup \pi_A(S)$ για κάθε μια από τις παρακάτω περιπτώσεις (εξηγείστε την απάντησή σας):
(i) Όταν δεν ισχύει καμία συνθήκη για το γνωρίσματα.
(ii) Όταν το A είναι κλειδί για την R.
(iii) Όταν το A είναι κλειδί για την R και ξένο κλειδί για την S.

Θέμα 4 [Μονάδες 20] Υποθέστε ότι έχετε 12.000.000 εγγραφές μεγέθους 200 bytes που είναι αποθηκευμένες σε ένα ταξινομημένο αρχείο. Θεωρείστε μη εκτεινόμενη οργάνωση. Το μέγεθος του block είναι 4096 bytes. Κατασκευάζουμε ένα μη πυκνό B+ ευρετήριο όπου τα φύλλα έχουν δείκτες στην πρώτη εγγραφή κάθε block. Θεωρείστε ότι το κλειδί ευρετηρίου είναι 35 bytes, οι δείκτες δέντρου και αρχείου (δεδομένων) είναι 30 bytes.

(α) Πόσες τιμές πρέπει να αποθηκευτούν στα φύλλα αυτού του ευρετηρίου; Ποιός είναι ο ελάχιστος αριθμός φύλλων που πρέπει να έχει ένα τέτοιο ευρετήριο;
(β) Υπολογίστε το ύψος ενός B+ ευρετηρίου του οποίου η ρίζα έχει 4 δείκτες και οι εσωτερικοί κόμβοι είναι γεμάτοι κατά 90%. Πόσο γεμάτα είναι τα φύλλα;
(γ) Θεωρείστε ότι χρησιμοποιούμε το δέντρο του ερωτήματος (β) για να υπολογίσουμε μια συνθήκη διαστήματος (δηλαδή, για να βρούμε όλες τις εγγραφές με κλειδί x όπου $a \leq x \leq b$, για κάποιες τιμές a και b), και ότι τη συνθήκη την ικανοποιούν 860 εγγραφές. Υπολογείστε τον αριθμό των blocks που χρειάζεται να μεταφερθούν και εξηγείστε γιατί.

Θέμα 5 [Μονάδες 10]

(α) Αφού τα ευρετήρια επιταχύνουν την επεξεργασία ερωτήσεων, γιατί να μην διατηρούμε ευρετήρια για κάθε γνώρισμα μιας σχέσης;

(β) Πόσα μη πυκνά ευρετήρια μπορούμε να έχουμε ανά αρχείο δεδομένων και γιατί;

(γ) Θεωρείστε επεκτατό δυναμικό κατακερματισμό. Ο διπλασιασμός του καταλόγου απαιτεί την εξέταση όλων των κάδων με τοπικό βάθος ίσο με το ολικό βάθος ή όχι; Εξηγείστε.