

Τελική Εξέταση : περιόδου Ιανουαρίου
31/ /2006

Θέμα 1 [25] Θέλουμε να σχεδιάσουμε μια ΒΔ για πόλεις, τις χώρες στις οποίες ανήκουν και προξενεία που υπάρχουν σ' αυτές. Η προς σχεδίαση ΒΔ τηρεί τα ακόλουθα στοιχεία:

- Κάθε πόλη ανήκει σε μία μόνο χώρα. Οι πληροφορίες για την πόλη περιλαμβάνουν το όνομα και τον πληθυσμό της. Τα ονόματα των πόλεων είναι διαφορετικά για μια χώρα, όμως πόλεις σε διαφορετικές χώρες μπορεί να έχουν ίδιο όνομα (και πιθανόν ίδιο πληθυσμό).
- Για κάθε χώρα τηρούμε το όνομά της, που θεωρείται μοναδικό και τον πρωθυπουργό της.
- Για τα προξενεία αποθηκεύουμε πληροφορίες σχετικές με το μοναδικό όνομά τους (π.χ. «Γαλλικό προξενείο στην Αθήνα») και τη διεύθυνσή τους (οδό και αριθμό), που είναι μοναδική σε μια πόλη, αλλά μπορεί να είναι ίδια για προξενεία που βρίσκονται σε διαφορετικές πόλεις.
- Μια πόλη ανήκει σε κάποια χώρα (συσχέτιση ανήκει).
- Τα προξενεία και οι πόλεις στις οποίες βρίσκονται σχετίζονται με τη συσχέτιση τοποθεσία.
- Οι χώρες σχετίζονται με τα προξενεία με τη συσχέτιση εκπροσώπηση. Κάθε χώρα που εκπροσωπείται μέσω ενός προξενείου έχει έναν πρόξενο για τον οποίο διατηρούμε μόνο το όνομά του. Υπάρχει περίπτωση ένα προξενείο να εκπροσωπεί περισσότερες από μία χώρα σε μια άλλη, οπότε σ' αυτό υπάρχουν τότε τόσοι πρόξενοι όσες και οι χώρες που εκπροσωπούνται.
- Τέλος, μια χώρα μπορεί να έχει παραπάνω από ένα προξενεία σε μια πόλη καθένα από τα οποία έχει τον πρόξενό του.

α) [15] Σχεδιάστε ένα μοντέλο οντοτήτων συσχετίσεων

β) [10] Μετατρέψτε το παραπάνω μοντέλο σε σχεσιακό.

Θέμα 2 [5] Για τις επερωτήσεις $Q1$ και $Q2$ στη συνέχεια, έστω ότι R και S είναι δύο σχέσεις με μοναδικά γνώρισματα τα a και b .

$Q1: \pi_a(R) \cap \pi_a(S)$ $Q2: \pi_a(R \cap S)$

Τι από τα ακόλουθα (α)-(δ) ισχύει και γιατί;

- (α) Οι $Q1$ και $Q2$ παράγουν το ίδιο αποτέλεσμα.
- (β) Το αποτέλεσμα της $Q1$ περιέχεται πάντα στο αποτέλεσμα της $Q2$.
- (γ) Το αποτέλεσμα της $Q2$ περιέχεται πάντα στο αποτέλεσμα της $Q1$.
- (δ) Οι $Q1$ και $Q2$ παράγουν διαφορετικά αποτελέσματα.

Θέμα 3 [5] Έστω ότι τα σχήματα δύο σχέσεων είναι $R(a,b)$ και $S(c)$.

$Q1: \text{SELECT } a$
 $\text{FROM } R$
 $\text{WHERE } R.b > \text{ALL}(\text{SELECT } c \text{ FROM } S)$

$Q2: \text{SELECT } a$
 $\text{FROM } R$
 $\text{WHERE } R.b > \text{ANY}(\text{SELECT } c \text{ FROM } S)$

Τι από τα ακόλουθα (α)-(δ) ισχύει και γιατί;

- (α) Οι $Q1$ και $Q2$ παράγουν το ίδιο αποτέλεσμα.
- (β) Το αποτέλεσμα της $Q1$ περιέχεται πάντα στο αποτέλεσμα της $Q2$.
- (γ) Το αποτέλεσμα της $Q2$ περιέχεται πάντα στο αποτέλεσμα της $Q1$.
- (δ) Οι $Q1$ και $Q2$ παράγουν διαφορετικά αποτελέσματα.

Θέμα 4 [5] Μία από τις ακόλουθες τέσσερις εκφράσεις σχεσιακής άλγεβρας δεν είναι ισοδύναμη με τις άλλες τρεις. Βασίζονται όλες στις σχέσεις $R(A,B)$ και $S(B,C)$.

Υποδείξτε ποια δεν είναι ισοδύναμη με τις άλλες και δικαιολογείστε την απάντησή σας.

- (α) $\pi_{AB}(R \bowtie S)$ (β) $R \bowtie \pi_B(S)$ (γ) $R \cap (\pi_A(R) \times \pi_B(S))$ (δ) $\pi_{A,B}(R \bowtie S)$

Θέμα 5 [20] Θεωρείστε το ακόλουθο σχεσιακό σχήμα:

Movie (title, year, genre, duration)

Actor (name, date-of-birth, place-of-birth, gender)

Stars (name, title, year, role)

Rating (title, year, overall, number-of-reviews)

Η σχέση Movie περιγράφει μια ταινία κι έχει γνωρίσματα τον τίτλο της (title), το έτος που γυρίστηκε (year), το είδος στο οποίο ανήκει (genre με τιμές τις drama, adventure, fantasy, comedy, romance και thriller) και τη διάρκειά της σε λεπτά (duration). Η σχέση Actor που αναφέρεται στους ηθοποιούς, έχει γνωρίσματα το όνομά τους (name), την ημερομηνία και τον τόπο γέννησής τους (date, place-of-birth) και το φύλο τους (gender). Η σχέση Stars περιγράφει ποιοι ηθοποιοί (name) παίζουν σε ποιες ταινίες (title), ποια χρονολογία (year) και το ρόλο τους (role). Τέλος η σχέση Ratings δίνει για κάθε ταινία (title), που γυρίστηκε ένα συγκεκριμένο χρόνο (year), ένα βαθμό (rating) που είναι δεκαδικός από 1 έως 10, καθώς και το πλήθος των βαθμολογητών της ταινίας (number-of-reviews).

Διατυπώστε σε σχεσιακή άλγεβρα ερωτήσεις για τα παρακάτω:

- Ταινίες που δεν είναι "drama" ή "romance".
- Ηθοποιούς που παίζουν σε τρεις τουλάχιστον διαφορετικές ταινίες.
- Ηθοποιούς που έχουν παίξει μόνο σε ταινίες με είδος "drama".
- Ηθοποιούς που έχουν παίξει σε όλες τις ταινίες με βαθμό τουλάχιστον 8 και πλήθος βαθμολογητών τουλάχιστον 50.
- Ταινίες με διάρκεια 50 ή 60 λεπτά και βαθμό από 7 έως 8.

Θέμα 6 [10] Έστω το σχεσιακό σχήμα $R(A, B, C, D, X, Y, Z)$ στο οποίο ισχύει το ακόλουθο σύνολο συναρτησιακών εξαρτήσεων $F = \{AB \rightarrow X, AC \rightarrow D, Y \rightarrow C, YZ \rightarrow X, XB \rightarrow D, BD \rightarrow Z\}$. Βρείτε την ελάχιστη κάλυψη για το F .

Θέμα 7 [10] Έστω ότι στην $R(A, B, C, D)$ ισχύει η ΣΕ $B \rightarrow C$. Ποια από τις ακόλουθες ΣΕ αν ισχύει επιπλέον θα σημαίνει ότι η R είναι σε 3NF αλλά όχι σε BCNF: α) $D \rightarrow AB$, β) $AC \rightarrow D$, γ) $CD \rightarrow B$ ή δ) $AD \rightarrow B$.

Θέμα 8 [10] Στη σχέση $R(A, B, C, D, E)$ ισχύει το ακόλουθο σύνολο ΣΕ: $\{A \rightarrow B, A \rightarrow D, BC \rightarrow A\}$. Αν διάσπασουμε την R προκειμένου να τη μετατρέψουμε σε BCNF ξεκινώντας με την εξάρτηση $A \rightarrow B$ ποια ΣΕ δεν μπορεί να ελεγχθεί σε μία από τις σχέσεις που προκύπτουν από τη διάσπαση και γιατί:

- (α) $A \rightarrow B$, (β) $A \rightarrow D$, (γ) $BC \rightarrow A$, (δ) Καμία από τις παραπάνω (δηλαδή, κάθε ΣΕ πάνω στην R μπορεί να ελεγχθεί σε μία από τις σχέσεις που προκύπτουν από τη διάσπαση).

Θέμα 9 [10] Έστω ένα B+ δέντρο τάξης 400 (ίδια και για τους εσωτερικούς κόμβους και για τα φύλλα).

- [5] Ποιος είναι ο μικρότερος και ποιος ο μεγαλύτερος αριθμός κλειδιών που μπορεί να κρατήσει ένα τέτοιο δέντρο;
- [5] Θεωρείστε ότι κάθε κόμβος του δέντρου είναι γεμάτος κατά 70%. Πόσα επίπεδα κόμβων απαιτούνται για $6 \cdot 10^9$ κλειδιά; Τι ποσοστό του συνολικού χώρου (blocks) είναι φύλλα στην περίπτωση αυτή;

Θέμα 10 [15] Θεωρείστε τον ακόλουθο (τροποποιημένο) ορισμό της σειριακοποιησιμότητας: «Ένα πρόγραμμα σύγχρονης εκτέλεσης S είναι σειριακοποιήσιμο αν και μόνο αν παράγει την ίδια κατάσταση της βάσης δεδομένων με ένα σειριακό πρόγραμμα T το οποίο περιλαμβάνει τις ίδιες ακριβώς δοσοληψίες με το S , διατεταγμένες όμως με τη σειρά της πρώτης τους εμφάνισης στο S ».

α) [5] Δείξτε ότι ο ορισμός αυτός είναι αρκετά περιοριστικός δηλαδή, υπάρχουν σειριακοποιήσιμα προγράμματα σύγχρονης εκτέλεσης τα οποία εξαιρούνται από τον ορισμό αυτό.

β) [5] Δείξτε επίσης ότι ο ορισμός αυτός δεν είναι αρκετά αυστηρός δηλαδή, υπάρχουν μη-σειριακοποιήσιμα προγράμματα σύγχρονης εκτέλεσης τα οποία ο ορισμός αυτός χαρακτηρίζει ως σειριακοποιήσιμα

γ) [5] Θεωρείστε το ακόλουθο πρόγραμμα σύγχρονης εκτέλεσης: $S: R_1(X) R_2(X) W_1(Y) R_3(Y) W_1(Z) W_2(Y) C_1 W_3(Z) C_3 C_2$ Δώστε ένα πρόγραμμα σύγχρονης εκτέλεσης των ίδιων δοσοληψιών το οποίο να ακολουθεί το πρωτόκολλο 2-Phase Locking.

$X \quad R_1 \quad t_1 \quad A \quad L \rightarrow S$
 $\quad \quad \quad \quad \quad \quad \quad \quad L \rightarrow 3S$
 $\quad \quad \quad \quad \quad \quad \quad \quad 400 \quad L \rightarrow 2S$

Καλή επιτυχία