Mỗi câu hỏi có thể có nhiều đáp án đúng, học viên điền những đáp án đó vào bảng trà lời bên dưới. Sáu câu hỏi có yêu cầu giải thích

1. Những phát biểu nào sau đây là đúng:
   1. Tuple, row. record là thuật ngữ tương đương trong mô hình CDSL quan hệ
   2. Một quan hệ có thể có nhiều khóa chính (primary key)
   3. Một quan hệ phải có ít nhất một khóa
   4. Một quan hệ phái có ít nhất một khóa ngoại (foreign key)
2. Một quan hệ có thể thuộc dạng chuẩn 3 (3NF) mà không cần phải thuộc dạng chuẩn 2 (2NF ). Phát biểu này là:
   1. Đúng
   2. Sai
3. Cho lược đồ quan hệ R(A1,A2,A3); P1, P2 chứa kết quả của các phép truy xuãt dữ liệu như mô tả dưới đây:

P1 πA2A3 (σA2=”a”(R))

P2 SELECT A2,A3 FROM R WHERE A2="a".

Những phát biểu nào sau đây là đúng:

* 1. |P1| = |P2|
  2. |P1| > |P2|
  3. |P1| < |P2|
  4. Tất cả các câu trên đều chưa chính xác

1. Cho quan hệ T(id,x,y,z) dùng để lưu tọa độ các điểm trong không gian 3 chiều dùng CSDL quan hệ như sau: điểm P(px,py,pz) sẽ được lưu trữ thành một hàng tP{P\_id,px,py,pz} trong bảng T. Phát biểu nào sau đây là đúng
   1. Để giúp cho việc thao tác với tập dữ liệu không gian Pi, i=1 .. N được hiệu quả về sau, người ta nên đánh chỉ mục lên mỗi thuộc tính T.x, T.y, T.z và T.id
   2. Có thể dùng một cây B\*-tree để đánh chỉ mục trực tiếp lên 3 cột {T.x, T.y, T.z}
   3. Có thể dùng một cây B\*-tree để đánh chỉ mục trực tiếp lên 4 cột {T\_id, T.x, T.y, T.z}
   4. Có thể dùng một cây R-tree để đánh chỉ mục trực tiếp lên 3 cột {T.x, T.y, T.z}
   5. Có thể dùng một cây R-tree để đánh chỉ mục trực tiếp lên 4 cột {T\_id, T.x, T.y, T.z}
2. Trong cách tiếp cận dùng chung file (shared file approach) cho các ứng dụng sau đây thì phát biểu nào sau đây là sai:
   1. Vấn đề cấu trúc dữ liệu không uyển chuyển (regrid data structure) xuất hiện và phải cần giải quyết
   2. Vấn đề dư thừa dữ liệu (data redundancy) đã được loại bỏ hoàn toàn
   3. Vấn đề không nhất quán dữ liệu (data inconsistency) đã đưọc loại bỏ hoàn toàn
   4. Vấn đề phụ thuộc dữ liệu có tính chất vật lý (physical data dependency) đã loại bỏ hoàn toàn
3. Những ràng buộc toàn vẹn được định nghĩa cho specialization là
   1. Participation và completeness
   2. Participation và disjointness
   3. Key constraint và disjointness
   4. Cardinality và participation
   5. Completeness và disjointness
4. Cho lược đồ quan hệ LOTS/PROPERTY\_ID#, COUNTY\_NAME, LOT#, AREA} và 3 phụ thuộc hàm

FD1: PROPERTY\_ID { COUNTY\_NAME, LOT#, AREA};

FD2: {COUNTY\_NAME, LOT# } { PROPERTY\_ID#, AREA};

FD3: AREA COUNTY\_NAME

Chúng ta có LOTS thuộc những dạng chuẩn nào sau đây

* 1. BCNF
  2. 3NF
  3. 2NF
  4. 1NF

1. Thiết kế ERD sau đây có thể gây fan trap



* 1. Đúng
  2. Sai

1. Những đối tượng nào sau đây trong ERD/E-ERD sẽ được chuyển thành một quan hệ trong quá trình ánh xạ theo quy tắc mà các anh chị đã được học
   1. Thực thể yếu
   2. Mối liên kết 3 ngôi one-one-one
   3. Mối liên kết 2 ngôi one-to-many với phía “many” là phía bắt buộc
   4. Thuộc tính đa trị
2. Dư thừa dữ liệu trong CSDL xảy ra khi dữ liệu giống nhau được lưu trữ nhiều lần ở vị trí khác nhau trong CSDL. Phát biểu “Sự dư thừa dữ liệu có điều khiển là khi CSDL bảo đảm được rằng những bản sao của cùng một dữ liệu là nhất quán với nhau”
   1. Sai
   2. Đúng
3. Phát biểu SQL là ngôn ngữ phi thủ tục (non-procedural language) là
   1. Đúng
   2. Sai
4. Giả sử A, B là 2 đối tượng trong một CSDL có k đối tượng không gian được đánh chỉ mục (indexing) sử dụng cây R-tree. Phát biểu “Nếu A là NN (nearest neighbour” của B thì B sẽ là một trong các k-NN của A là
   1. Sai
   2. Đúng

Hình sau dùng cho câu 13 và 14

Hình (a)

 Hình (b)

1. Với đối tượng hỏi là Q1 thì những node nào của cây R-tree sẽ được truy xuất để trả lời cho câu hỏi : “Tìm các đối tượng có chứa dữ liệu Q1 ?”
   1. Root, R1, R3
   2. Root, R2, R6
   3. Root, R1, R4
   4. Các câu trên đều sai
2. Với đối tượng hỏi là Q1 thì những node nào của cây R-tree sẽ được truy xuất để trả lời cho câu hỏi : “Tìm các đối tượng có giao với Q1 ?”
   1. Root, R1, R3, R5, R13
   2. Root, R2, R6, R7, R17
   3. Root, R1, R3, R4, R12
   4. Các câu trên đều sai
3. Phát biểu “Các hàm gộp (MIN, MAX, AVERAGE, COUNT, SUM) chỉ thao tác trên các giá trị khác NULL” là:
   1. Sai vì hàm COUNT(\*) thao tác trên cả giá trị NULL
   2. Đúng vì các giá trị NULL sẽ bị bỏ qua đối với các hàm gộp
   3. Sai vì hàm SUM sẽ coi các giá trị NULL bằng với số 0
   4. Các câu A và C đúng
   5. Các câu trên đều sai
4. Trong một câu lệnh SELECT, các hàm gộp (MIN, MAX, AVERAGE, COUNT, SUM) chỉ có thể xuất hiện trong danh sách các giá trị được trả về trong phần SELECT hoặc trong các mệnh đề HAVING và ORDER BY. Phát biểu này là
   1. Đúng
   2. Sai
5. Tên của một thuộc tính phải là duy nhất trong
   1. Một CSDL
   2. Một quan hệ (relation)
   3. Câu a và b sai
   4. Câu a và b đúng
6. Những kiểu ràng buộc nào sau đây không phải là một kiểu ràng buộc trong mô hình dữ liệu quan hệ (relational data model)
   1. Entity integrity constraint
   2. Domain constraint
   3. Referential integrity
   4. Participation constraint
   5. Key constraint
7. Phát biểu "Một khóa ngoại (foreign key) có thể có chứa 4 thuộc tính" là:
   1. Sai vì khóa ngoại chỉ có thể chứa 1 thuộc tính
   2. Đúng
8. Một câu lệnh SELECT của SQL bắt buộc phải có ít nhất tà:
   1. Các mệnh đề SELECT, FROM, WHERE
   2. Các mệnh đề SELECT, WHERE
   3. Các mệnh đề SELECT, FROM
   4. Mệnh đề SELECT
   5. Không câu nào trong các câu trên đúng
9. Bậc của một kiểu quan hệ (the degree of a relationship type) được định nghĩa là:
   1. The number of relationship instances
   2. The number of relationship attributes
   3. The number of participating entity types
   4. Các câu a và b đều đúng.
10. Bậc của một quan hệ (the degree of arelation) được định nghĩa là:
    1. The number of its relationship instances
    2. The number of its relation schema attributes
    3. The number of its relalionship attributes
    4. Câu B và C đều đúng
11. Cho quan hệ R( A,B,C,D) và một tập các phụ thuộc hàm

F={AB CE, B D, D A}. Những phụ thuộc hàm nào sau đây không thể suy ra được từ F;

* 1. BC DE
  2. AD CE
  3. B E
  4. AB A
  5. Tất cả các phụ thuộc hàm ở trên
  6. Không có phụ thuộc hàm nào ở trên

1. Phát biểu "Mọi quan hệ có hai thuộc tính đều thuộc dạng chuẩn BCNF là:
   1. Đúng
   2. Sai
2. Giả sử kích thước các nodes (node size) và B\*-tree là như nhau và các cây này dùng để đánh chỉ số trên cùng một cột của một bảng trong một CSDL thì B\*-tree luôn có chiều cao nhỏ hơn hoặc bằng chiều cao của B-tree. Phát biểu này là:
   1. Sai
   2. Đúng
3. Một quan hệ luôn luôn thuộc dạng chuẩn 2 (2NF) khi:
   1. Chỉ có một khóa và khóa này chỉ có một thuộc tính
   2. Tồn tại những phụ thuộc hàm bắc cầu
   3. Chỉ có một khóa
   4. Tất cả những khóa dự tuyển là những khóa chỉ có một thuộc tính
   5. Tất cả những phụ thuộc hàm đều thuộc dạng không đầy đủ (partial)
4. Cho T là một cây R-tree lưu trữ các MBR (minimum bounding rectangles) của các đối tượng (dữ liệu không gian 3 chiều phức tạp. Giả sử Q(x,y,z) là một điểm trong không gian dữ liệu của T. Để xác định xem đối tượng dữ liệu nào trong T thực sự chứa Q thì:
   1. Chỉ phải duyệt cây T theo 1 nhánh
   2. Phải duyệt cây T theo 2 nhánh
   3. Có thể chỉ phải kiểm tra gốc (root) của cây
   4. Chưa thể xác định phải duyệt cây T theo bao nhiêu nhánh
   5. Tất cả các câu trên đều sai
5. Cho quan hệ R(A, B, C) và ba tác vụ sau đây
6. σA=c(R(A,B,C));
7. {t|R(t) AND t.A = c}
8. {xyz|R(xyz) AND x =c AND z=c}
   1. Các ba tác vụ là tương đương nhau trên R
   2. Các tác vụ (1) và (2) là tương đương nhau trên R
   3. Các tác vụ (2) và (3) là tương đương nhau trên R
   4. Các tác vụ (1) và (3) là không tương đương nhau trên R
   5. Tất cả các câu trên đều sai
9. Tập các thuộc tính K của một lược đồ quan hệ R được gọi là một khóa của R nếu với bất kì hai hàng khác nhau t1, t2 trong bất kì trạng thái nào của R chúng ta luôn có ràng buộc sau đây được thỏa: t1[K] ≠ t1[K]
   1. Phát biểu trên là đúng
   2. Phát biểu trên là sai
   3. Tập K tạo thành một khóa dự tuyển
   4. Tập K không tạo thành một khóa dự tuyển
10. Những ngôn ngữ thao tác dữ liệu DMLs (data manipulation language) là những ngôn ngữ phi thủ tục (non-procedural language).
    1. Phát biểu trên là sai
    2. Phát biểu trên là đúng
    3. Chỉ có DML cho mô hình dữ liệu mạng (network data model) là ngôn ngữ thủ tục
    4. Chỉ có DML cho mô hình dữ liệu phân cấp (hierarchical data model) là ngôn ngữ thủ tục
11. Mọi thuộc tính được liệt kê trong mệnh đề ORDER BY của một câu lệnh SELECT phải xuất hiện trong mệnh đề SELECT.
    1. Phát biểu trên là đúng khi thay cụm từ “ORDER BY" trong phát biểu bảng cụm từ "GROUP BY”
    2. Phát biểu trên là đúng
    3. Phát biểu trên là sai
    4. Câu A và C đúng
12. Khóa chính của một quan hệ có thể được nhận dạng bằng cách kiểm tra trạng thái mở rộng (extension state) của một quan hệ.
    1. Phát biểu trên là sai
    2. Phát biểu trên là đúng và điều này là một cách tiếp cận thực tế khi thiết kế các CSDL
    3. Phát biểu trên chỉ đúng cho một số quan hệ
    4. Câu B và C đúng
13. Cho lược đồ quan hệ R. Những phát biểu nào sau đây là SAI:
    1. σcond1(σcond2(R)) = σcond2(σcond1(R))
    2. σ<cond1 > (σcond2(R)) = σ<cond1 AND cond2 > (R))
    3. σ<cond1 > (σ<cond2>(R)) ≤ |σ<cond1>(R))|
    4. Tất cả các câu trên đều sai
14. Cho hai lược đồ quan hệ R và S. Hãy chọn những phát biểu ĐÚNG sau đây. Kí hiệu x là tích Đề cát, \* là phép toán nhân):
    1. |R x S| = |R| \* |S|
    2. |R x S| < |R| \* |S|
    3. Bậc (R x S) = Bậc (R) \* Bậc (S)
    4. Bậc (R x S) = Bậc (R) + Bậc (S)
15. Các khái niệm cơ bản về bảo mật CSDL cho rằng:
    1. *Identification và authentication* là hai công việc giống nhau
    2. Nguyên lý no write-down trong MAC cũng tồn tại trong RBAC
    3. Dùng dấu vân tay (fingerprint) là một kỹ thuật sử dụng trong việc authentication
    4. Tất cả các câu trên đều đúng
16. Chức năng nào sau đây KHÔNG PHẢI là của một DBA
    1. Security and authentication
    2. Database design
    3. Transaction management
    4. Database tunining
    5. Back up
    6. Performance monitoring
17. Mô hình Bell – LaPudala
    1. Quan tâm đến tính toàn vẹn (integrity) SAI
    2. Quan tâm đến tính chất chống thoái thác (anti repudiation)
    3. Quan tâm đến tính sẵn sàng (availability) SAI
    4. Quan tâm đến tính mật (confidential)
    5. Quan tâm đến tính mật và toàn vẹn dữ liệu SAI
18. Cho 2 thao tác T1 và T2 được thực thi như bên dưới và giả sử các giá trị ban đầu là X=100, Y=50, M =10, N=20. Sau khi 2 giao tác T1 & T2 kết thúc và giao tác T1 không thành công trong câu lệnh sau cùng read-item(Y) thì:
    1. X=100, M=20
    2. X=110, N=20
    3. X=90, M=10
    4. X=100, M=10
    5. Tất cả các câu trên đều sai

|  |  |
| --- | --- |
| T1 | T2 |
| read\_item(X);  X:=X-N;  write\_item(X);  read\_item(Y); | read\_item(X);  X:=X+M;  write\_item(X); |

1. Trong bảo mật CSDL với 2 kĩ thuật điều khiển truy xuất là DAC và MAC
   1. Hệ thống sử dụng MAC không gặp phải vấn đề Trojan Horses ĐÚNG do nó kết hợp BLB và BP
   2. Vấn đề covert channel không tồn tại trong các hệ thống sử dụng MAC SAI do nó chỉ là disavantage.
   3. DAC thích hợp cho các ứng dụng trong môi trường web
   4. Tất cả các câu trên đều sai
2. Những phát biểu nào sau đây là SAI
   1. Khóa ngoại không được NULL
   2. Mọi thành phần của khóa chính không được NULL
   3. Mỗi quan hệ chỉ có thể có nhiều nhất một khóa ngoại
   4. Mỗi thành phần nào đó của khóa chính cũng có thể là khóa ngoại tham khảo đến khóa chính của quan hệ khác
3. Chức năng nào sau đây không thuộc về mô hình dữ liệu ý niệm (conceptual) Mức ERD
   1. Enable the database to be implemented in a DBMS
   2. Enable the database user to understand the structure of database
   3. Acts as a communication tool between the end user and designers
   4. Represents the types of data stored in the database
   5. Represents the relationships between data
4. Những yếu tố nào sau đây đảm bảo tính nhất quán của cơ sở dữ liệu:
   1. Khóa ngoại và khóa chính
   2. Các khóa dự tuyển và khóa chính
   3. Hệ quản trị CSDL ( =DBMS)
   4. Nhà cung cấp hệ quản trị CSDL
   5. DBMS
5. A database extension is the state of the database
   1. Đúng
   2. Sai
6. Hai giao tác T1 và T2 sau đây thỏa mãn các điều kiện của giao thức 2PL
   1. Đúng, giải thích:

…

* 1. Sai, giải thích

|  |  |
| --- | --- |
| T1 | T2 |
| read\_lock(Y);  read\_item(Y);  unlock(Y);    write\_lock(X);  read\_item(X);  X:=X+Y  write\_item(X);  unlock(X); | read\_lock(X);  read\_item(X);  unlock(X);  write\_lock(Y);  read\_item(Y);  Y:=X+Y  write\_item(Y);  unlock(Y); |

1. Với ngôn ngữ SQL với các DBMS thì:
   1. Một phát biểu SQL đơn (a single SQL, statement) luôn được đảm bảo thỏa mãn tính chất atomic
   2. Khi isolation level (tính cô lập) mang giá trị UNCOMMITED thì vấn đề dirty read có thể xảy ra
   3. Vấn đề deadlock trong các giao lúc giao tác phải được giải quyết khi viết các câu lệnh SQL trong giao tác
   4. Tất cả các câu trên đều đúng
2. Một kênh bí mật (covert channel):
   1. Là một đường liên lạc không được thiết kế cho mục đích liên lạc trong hệ thống
   2. Là một Trojan hoặc virus máy tính
   3. Chỉ tồn tại trong các hệ thống bảo mật đóng (close security systems)
   4. Có thể là một hệ thống file (file system)
3. Với dữ liệu lớn (big data) thi phát biểu nào sau đây là SAI
   1. Các DBMS thương mại hiện nay chưa thể quản lý và xử lý câu truy vấn hiệu quá
   2. Dữ liệu phải được quản lý trên hạ tầng điện toán đám mây
   3. Lượng dữ liệu cần lưu trữ nhiều
   4. Tốc độ dữ liệu dược tạo ra cao
   5. Dữ liệu lưu trữ da dạng
4. "Physical database design" có thể gồm những việc làm sau đây:
   1. Thiết kế lại ERD cho CSDL
   2. Thiết kế chỉ mục mức vật lý
   3. "Denormalization" nhằm tăng tốc độ xử lý cho các câu truy vấn
   4. Tất cả các câu trên đều sai
5. Các cấu trúc chỉ mục nhiều chiều dùng để
   1. Đánh chỉ mục cho các dữ liệu nhiều chiều
   2. Tăng tốc độ xử lý câu truy vấn với dữ liệu nhiều chiều
   3. Tránh việc phải truy xuất tuần tự toàn bộ CSDL khi xứ lý câu truy vấn
   4. Tất cả các câu trên đều sai
6. Những phát biểu nào sau đây là đúng:
   1. Hệ quản trị CSDL bao gồm CSDL và các ứng dụng
   2. CSDL bao gồm hệ quản trị
   3. Hệ cơ sở dữ liệu bao gồm hệ quản trị CSDL và CSDL
   4. Tất cả các câu trên đều sai