

데이터베이스론

문 1. 관계형 데이터베이스에서 참조 무결성 제약조건을 만족하도록 외래키를 생성하고자 한다. 이와 관련된 SQL 명령문의 제약조건에 대한 설명으로 옳지 않은 것은?

- ① 'on delete set null'은 참조되는 테이블의 행이 삭제되면, 참조하는 테이블의 행에 있는 외래키 열에 null을 저장한다.
- ② 'on delete cascade'는 참조되는 테이블의 행이 삭제되면, 참조하는 테이블의 행에 있는 외래키 열을 삭제한다.
- ③ 'on delete set default'는 참조되는 테이블의 행이 삭제되면, 참조하는 테이블의 행에 있는 외래키 열에 사전에 정의된 default 값을 저장한다.
- ④ 'on delete no action'은 참조되는 테이블의 행을 삭제하려고 할 때, 참조하는 테이블의 행이 존재할 경우 삭제 명령이 수행되지 못하도록 한다.

문 2. 다음 과제 테이블은 학년별 동아리에 가입한 학생 수와 제출한 과제 수를 저장하고 있다. '학생이 10명 이상 가입한 동아리에 대하여 동아리와 제출한 총 과제 수를 출력하시오'를 수행하기 위한 SQL문으로 옳은 것은?

과제			
학년	동아리	학생수	과제수
1	A	20	20
1	B	10	20
1	C	5	8
2	A	15	10
2	B	15	20
3	A	5	15
3	B	12	20
3	C	8	15

- ① SELECT 동아리, SUM(과제수)
FROM 과제
GROUP BY 동아리
HAVING SUM(학생수) >= 10 ;
- ② SELECT 동아리, SUM(과제수)
FROM 과제
WHERE SUM(학생수) >= 10 ;
- ③ SELECT 동아리, SUM(과제수)
FROM 과제
HAVING SUM(학생수) >= 10 ;
- ④ SELECT 동아리, SUM(과제수)
FROM 과제
WHERE 학생수 >= 10
GROUP BY 동아리 ;

문 3. 다음 관계형 데이터베이스의 세 가지 기능적 요소에 대한 설명에서 ㉠ ~ ㉣에 들어갈 용어를 바르게 연결한 것은?

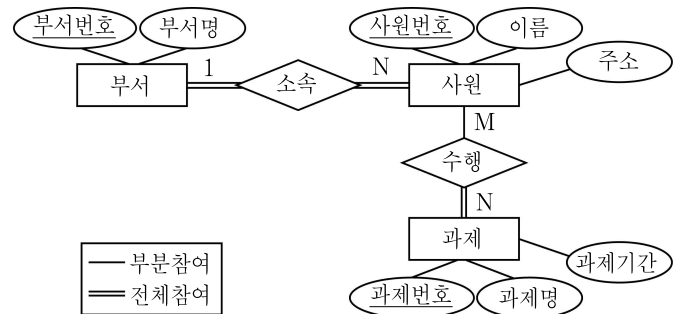
- (㉠)는(은) SQL에서 삽입, 삭제, 갱신과 같은 데이터 변경문을 실행할 때 미리 명시된 조건을 만족하는 경우 특정한 동작을 자동으로 수행할 수 있도록 한다.
○ (㉡)는(은) 데이터베이스 내에 존재하는 작업 순서가 정해진 수행 단위로서 DBMS에서 컴파일된 후 실행된다.
○ (㉢)는(은) 데이터베이스에서 데이터를 신속하게 탐색할 수 있도록 만든 데이터 구조이다.

- | | | |
|-------|--------------|---------------------------|
| ㉠ | ㉡ | ㉢ |
| ① 인덱스 | 트리거(trigger) | 주장(assertion) |
| ② 주장 | 인덱스 | 저장 프로시저(stored procedure) |
| ③ 주장 | 인덱스 | 트리거 |
| ④ 트리거 | 저장 프로시저 | 인덱스 |

문 4. 데이터베이스 보안과 관련한 설명으로 옳지 않은 것은?

- ① SQL 삽입(injection) 공격은 공격자가 악의적으로 만든 SQL 명령을 응용 프로그램이 수행하도록 하는 것이다.
- ② 데이터베이스 관리자(DBA)가 각 사용자에게 데이터베이스에 대한 접근 권한을 부여하거나 취소할 때 grant 명령을 사용한다.
- ③ 데이터베이스에 대한 권한은 역할(role)에도 부여할 수 있다.
- ④ 전자 서명(digital signature)은 공개키 암호화 기법의 특성을 이용하여 인증을 수행한다.

문 5. 다음 ERD(Entity-Relationship Diagram)에 대한 설명으로 옳지 않은 것은? (단, 속성 이름에 대한 밑줄은 기본키이다)



- ① 주어진 ERD를 릴레이션으로 사상(mapping)하면, 4개의 릴레이션이 생성된다.
- ② 모든 사원은 적어도 하나 이상의 과제를 수행하여야 한다.
- ③ 모든 부서는 적어도 1명 이상의 사원이 존재하여야 하며, 사원도 반드시 하나의 부서에 소속되어야 한다.
- ④ 사원은 2개의 부서에 동시에 소속될 수 없다.

문 6. 다음 함수종속성 집합 FD의 최소커버(minimal cover) FD_{min}는?

$$FD = \{Y \rightarrow X, Z \rightarrow XYW\}$$

- ① FD_{min} = {Y → X, Z → X, Z → Y, Z → W}
- ② FD_{min} = {Y → X, Z → Y, Z → W}
- ③ FD_{min} = {Y → X, Z → X, Z → W}
- ④ FD_{min} = {Y → X, Z → X, Z → Y}

문 7. 다음 인덱스 기법에 대한 설명으로 옳은 것만을 모두 고르면?

- ㄱ. B-트리 전체를 노드 내의 키 값에 따라 순차 검색하기 위해서는 트리의 각 노드를 전위 순회(preorder traversal)한다.
- ㄴ. 밀집(dense) 인덱스는 희소(sparse) 인덱스에 비해 액세스 시간은 빠르지만 더 많은 공간을 필요로 한다.
- ㄷ. B-트리에서 오버플로가 발생하여 리프 노드가 분할될 때, 중간 키 값이 부모 노드뿐만 아니라 새로 분할된 노드에도 저장된다.
- ㄹ. 1,000명의 사원 정보가 저장된 '사원' 테이블의 '부서' 필드에 대하여, 30명이 소속된 '총무과' 사원에 대한 비트맵 인덱스를 구성할 경우 1,000비트가 필요하다.

- ① ㄱ, ㄷ
② ㄱ, ㄹ
③ ㄴ, ㄷ
④ ㄴ, ㄹ

문 8. 다음은 시스템 고장이 발생할 때 트랜잭션 T₁, T₂, T₃, T₄, T₅를 복구하기 위해 고장 전에 마지막으로 기록된 DBMS 로그이다. 이 로그를 사용하여 즉시 갱신 회복 기법의 undo-redo 알고리즘을 수행할 때, 회복과정에 대한 설명으로 옳지 않은 것은? (단, <checkpoint>는 검사점 기록 로그 레코드이며, <T₁, A, 200, 400>은 'T₁이 데이터 항목 A의 현재 값 200을 400으로 갱신한다'를 의미하는 로그 레코드이다. 표현되지 않은 로그의 다른 속성은 고려하지 않는다)

로그 번호	로그 레코드
1	<T ₁ start>
2	<T ₁ , A, 200, 400>
3	<T ₁ commit>
4	<T ₂ start>
5	<T ₂ , B, 900, 1900>
6	<checkpoint>
7	<T ₃ start>
8	<T ₃ , A, 400, 700>
9	<checkpoint>
10	<T ₅ start>
11	<T ₅ , C, 100, 5000>
12	<T ₃ commit>
13	<T ₄ start>
14	<T ₄ , D, 700, 1300>
15	<T ₅ commit>
16	<T ₄ , E, 1300, 1500>
--시스템 고장발생--	

- ① T₁은 어떠한 undo와 redo 연산도 수행하지 않는다.
② 데이터 항목 A와 B의 값은 각각 700과 900으로 갱신된다.
③ 데이터 항목 C와 D의 값은 각각 5000과 700으로 갱신된다.
④ redo 연산은 T₅, T₃ 순서로 수행된다.

문 9. <보기 1>에서 공급업체와 부품, 카탈로그 테이블을 생성하는 SQL문을 수행한 후 튜플을 삽입하여 세 테이블의 상태가 다음과 같을 때, <보기 2>의 SQL문을 수행한 결과로 옳은 것은?

—<보기 1>—

```
CREATE TABLE 공급업체 (
  업체번호 INT NOT NULL,
  업체명 VARCHAR(20),
  PRIMARY KEY(업체번호));
CREATE TABLE 부품 (
  부품번호 INT NOT NULL,
  부품명 VARCHAR(20),
  색상 VARCHAR(20),
  PRIMARY KEY(부품번호));
CREATE TABLE 카탈로그 (
  업체번호 INT NOT NULL,
  부품번호 INT NOT NULL,
  가격 INT,
  PRIMARY KEY(업체번호, 부품번호),
  FOREIGN KEY(업체번호) REFERENCES 공급업체(업체번호),
  FOREIGN KEY(부품번호) REFERENCES 부품(부품번호));
```

공급업체	
업체번호	업체명
1	공급A
2	공급B
3	공급C

부품		
부품번호	부품명	색상
1	부품1	빨강
2	부품2	파랑
3	부품3	노랑

카탈로그		
업체번호	부품번호	가격
1	3	10000
2	1	20000
3	2	30000

—<보기 2>—

```
SELECT 업체번호, 업체명
FROM 공급업체
WHERE NOT EXISTS (
  SELECT 부품.부품번호
  FROM 부품
  WHERE 부품.색상='빨강' AND EXISTS (
    SELECT *
    FROM 카탈로그
    WHERE 카탈로그.부품번호=부품.부품번호
    AND 카탈로그.업체번호=공급업체.업체번호));
```

- ①

업체번호	업체명
2	공급B
- ②

업체번호	업체명
2	공급B
- ③

업체번호	업체명
1	공급A
3	공급C
- ④

업체번호	업체명
1	공급A
2	공급B
3	공급C

문 10. 다음 SQL문에 대하여 질의 최적화를 수행하고자 한다. 이에 대한 설명으로 옳지 않은 것은?

```
SELECT 학생.이름, 교수.이름
FROM 학생, 학과, 교수
WHERE 학생.학과코드 = 학과.학과코드
      AND 교수.교수코드 = 학과.학과장코드
      AND 학과.학과코드 = 'CS' ;
```

- ① 카탈로그는 해당 테이블에 대한 통계 정보를 저장하고 있어 비용 계산에 활용될 수 있지만, 통계 정보를 실시간으로 갱신하려면 부하가 커서 주기적으로 갱신하기도 한다.
- ② 조인연산 $\text{학생} \bowtie_{\text{학과코드}=\text{학과코드}} \text{학과}$ 는 기본적으로 중첩 루프(nested loop)를 이용하여 구현하지만, '학과.학과코드'에 인덱스가 구축되어 있다면 비용을 감소시킬 수 있다.
- ③ 경험적 질의 최적화 기법에서는 주어진 SQL문의 WHERE절에서 조인연산보다 학과.학과코드 = 'CS'를 먼저 수행하도록 한다.
- ④ $|R|$ 을 릴레이션 R 의 튜플의 수로 정의할 때, $R \bowtie_{A=B} S$ 에 대하여 A 가 R 의 기본키이면 $|(R \bowtie_{A=B} S)| \leq |S|$ 이고 조인 선택률(j_s)은 $j_s > 1/|R|$ 이다.

문 11. 두 릴레이션 T_1 , T_2 에 관계 대수 연산 ㉠ ~ ㉣을 수행한 결과 릴레이션이 각각 R_1 , R_2 , R_3 과 같다. 연산을 바르게 연결한 것은?

T_1		T_2	
A	B	B	C
aa	10	20	xx
bb	20	30	yy
cc	30	40	zz
dd	20		

$R_1 \leftarrow T_1 \text{ ㉠ } T_2$			$R_2 \leftarrow T_1 \text{ ㉡ } T_2$			$R_3 \leftarrow T_1 \text{ ㉢ } T_2$		
A	B	C	A	B	C	A	B	C
aa	10		aa	10		bb	20	xx
bb	20		bb	20	xx	dd	20	xx
cc	30		cc	30	yy	cc	30	yy
dd	20		dd	20	xx		40	zz
	20	xx		40	zz			
	30	yy						
	40	zz						

㉠

㉡

㉢

- ① 완전 외부조인($\bowtie$) 외부 합집합($\cup$) 왼쪽 외부조인($\bowtie$)
- ② 완전 외부조인($\bowtie$) 외부 합집합($\cup$) 오른쪽 외부조인($\bowtie$)
- ③ 외부 합집합($\cup$) 완전 외부조인($\bowtie$) 왼쪽 외부조인($\bowtie$)
- ④ 외부 합집합($\cup$) 완전 외부조인($\bowtie$) 오른쪽 외부조인($\bowtie$)

문 12. SQL에서 뷰(view)의 역할에 대한 설명으로 옳지 않은 것은?

- ① 기본 테이블(base table)들만으로 작성된 질의를 간소화시킬 수 있다.
- ② 사용자의 접근권한에 따라 동일한 기본 테이블의 속성들을 선택적으로 제공할 수 있다.
- ③ 기본 테이블들만으로 작성된 질의 처리 성능을 향상시키기 위해 개발되었다.
- ④ 기본 테이블들의 물리적 구조를 변경시키지 않고 사용자가 원하는 새로운 가상 테이블을 생성시킬 수 있다.

문 13. 데이터베이스와 데이터베이스 관리시스템(DBMS)에 대한 설명으로 옳지 않은 것은?

- ① 데이터 종속성(data dependency)을 유지하여 데이터와 이를 이용하는 프로그램이 밀접하게 연결되어 동작하도록 도와준다.
- ② 데이터베이스 사용자들에게 공용 데이터에 대한 다양한 관점을 제공해 준다.
- ③ 데이터베이스 시스템의 자기 기술성은 데이터베이스 구조와 제약조건에 대한 정의를 가지고 있음을 의미한다.
- ④ 분산 데이터베이스 시스템에 포함된 각 지역의 DBMS는 지역 자치성(local autonomy)을 가질 수 있다.

문 14. 다음 트랜잭션 스케줄 S_1 , S_2 , S_3 , S_4 중 회복 불가능한 스케줄은? (단, $r_i(X)$ 와 $w_i(X)$ 는 각각 트랜잭션 T_i 의 데이터 항목 X 에 대한 읽기연산과 쓰기연산이고, c_i 와 a_i 는 각각 T_i 의 완료연산과 철회 연산이다)

- ① S_1 : $r_1(X)$; $r_2(X)$; $r_1(Y)$; $w_2(X)$; $w_1(Y)$; a_1 ; $r_2(Y)$; c_2 ;
- ② S_2 : $r_1(X)$; $w_1(X)$; $r_2(X)$; $r_1(Y)$; $w_2(X)$; c_2 ; a_1 ;
- ③ S_3 : $r_1(X)$; $w_1(X)$; $r_2(Y)$; $w_2(Y)$; $w_2(X)$; c_2 ; c_1 ;
- ④ S_4 : $r_1(X)$; $w_1(X)$; $r_2(X)$; $r_1(Y)$; $w_2(X)$; $w_1(Y)$; a_1 ; a_2 ;

문 15. 릴레이션 스키마 $R(A, B, C, D)$ 는 다음 함수적 종속성 집합 FD를 만족한다. R 을 BCNF 정규화하였을 때 분해된 두 릴레이션 스키마 R_1 과 R_2 로 옳은 것은?

$$FD = \{A \rightarrow BCD, B \rightarrow ACD, D \rightarrow C\}$$

- ① $R_1(B, C, D)$, $R_2(A, B)$
- ② $R_1(A, C, D)$, $R_2(B, C, D)$
- ③ $R_1(A, B, D)$, $R_2(C, D)$
- ④ $R_1(A, B, C)$, $R_2(C, D)$

문 16. 다음과 같은 두 트랜잭션 T_1 과 T_2 에 대하여 현재 T_1 은 $x\text{-lock}(X)$ 를, T_2 는 $s\text{-lock}(Y)$ 를 성공적으로 수행한 상태이다. 이후 두 트랜잭션이 각각 순서대로 마지막 연산까지 수행하는 과정에 대한 설명으로 옳은 것은? (단, $\text{read}(X)$ 와 $\text{write}(X)$ 는 각각 트랜잭션의 데이터 항목 X 에 대한 읽기연산과 쓰기연산이고, $x\text{-lock}(X)$ 와 $s\text{-lock}(X)$ 는 각각 X 에 대한 배타로크를 거는 연산과 공유로크를 거는 연산이며, $\text{unlock}(X)$ 는 X 에 대한 로크를 해제하는 연산이다)

T_1	T_2
$x\text{-lock}(X)$	$s\text{-lock}(Y)$
$\text{read}(X)$	$\text{read}(Y)$
$X = X + 500$	$x\text{-lock}(X)$
$\text{write}(X)$	$\text{unlock}(Y)$
$x\text{-lock}(Y)$	$\text{read}(X)$
$\text{unlock}(X)$	$X = X + Y$
$Y = Y + 500$	$\text{unlock}(X)$
$\text{unlock}(Y)$	

- ① T_1 은 성공적으로 완료되지만 T_2 는 교착 상태에 빠진다.
- ② T_1 은 교착 상태에 빠지지만 T_2 는 성공적으로 완료된다.
- ③ T_1 과 T_2 가 모두 성공적으로 완료된다.
- ④ T_1 과 T_2 가 모두 교착 상태에 빠진다.

문 17. 빅데이터(big data)에 대한 설명으로 옳지 않은 것은?

- ① 디지털 환경에서 생성되는 데이터로 규모가 방대하고, 정형, 반정형, 비정형 등 다양한 형태의 데이터를 포함한다.
- ② NoSQL 시스템은 반구조적이고 자기 기술적인 데이터를 허용하므로 대개는 스키마를 요구하지 않는다.
- ③ NoSQL의 키-값(key-value) 데이터 모델은 키와 값의 쌍으로 저장하며, 값은 이미지나 동영상 등 다양한 형태의 데이터가 될 수 있다.
- ④ 빅데이터 분석 과정에서 추출된 정보를 시각화하는 기술로 Hadoop의 맵 리듀스(MapReduce)를 사용한다.

문 18. XML(eXtensible Markup Language)에 대한 설명으로 옳은 것은?

- ① DTD(Document Type Definition) 문서는 XML 문서를 작성하기 위한 구문 규칙과 동일한 방식으로 작성된다.
- ② DTD(Document Type Definition) 문서는 XML 스키마 문서와 달리 네임스페이스(namespace)를 지원한다.
- ③ XQuery 질의문은 FOR, LET, WHERE, RETURN 절로 구성되고 둘 이상의 변수에 대한 조인연산을 작성할 수 있다.
- ④ XML 스키마는 XML 문서의 요소 및 속성을 데이터베이스 테이블에 맵핑(mapping)할 때 문자열 데이터를 제외한 다른 데이터형(data type)을 지정할 수 없다.

문 19. 3개의 트랜잭션 T_1 , T_2 , T_3 에 대하여, 다음과 같은 스케줄의 뷰(view)/충돌(conflict) 직렬 가능성(serializability)을 설명한 것으로 옳은 것은?

시간	T_1	T_2	T_3
	$\text{read}(X)$		
	$\text{write}(X)$	$\text{write}(X)$	
			$\text{write}(X)$

- ① 뷰 직렬 가능하며, 동시에 충돌 직렬 가능하다.
- ② 뷰 직렬 가능하지 않으며, 충돌 직렬 가능하지 않다.
- ③ 뷰 직렬 가능하지만, 충돌 직렬 가능하지 않다.
- ④ 뷰 직렬 가능하지 않지만, 충돌 직렬 가능하다.

문 20. 관계형 데이터베이스 역정규화(denormalization)에 대한 설명으로 옳지 않은 것만을 모두 고르면?

- ㄱ. 릴레이션들은 역정규화한 후 정규형 수준이 높아진다.
- ㄴ. 데이터베이스의 데이터 무결성을 강화할 목적으로 개발되었다.
- ㄷ. 릴레이션들의 데이터 중복을 줄임으로써 데이터베이스의 크기를 감소시킨다.
- ㄹ. 둘 이상의 릴레이션들에 대하여 조인한 결과를 빈번하게 이용하는 경우, 역정규화 함으로써 질의응답 시간이 단축될 수 있다.

- ① ㄱ, ㄴ, ㄷ
- ② ㄱ, ㄴ, ㄹ
- ③ ㄴ, ㄷ, ㄹ
- ④ ㄱ, ㄴ, ㄷ, ㄹ