

데이터베이스론

문 1. 데이터베이스 관리 시스템(DBMS)에 대한 설명으로 옳지 않은 것은?

- ① 데이터 중복을 효율적으로 제어할 수 있다.
- ② 여러 사용자의 데이터 공유를 허용하지 않는다.
- ③ 사용자마다 데이터에 대한 적절한 권한을 부여하여 허가 받지 않은 접근을 통제할 수 있다.
- ④ 데이터베이스 시스템의 하드웨어나 소프트웨어의 고장에 대비하여 백업과 회복 기능을 제공한다.

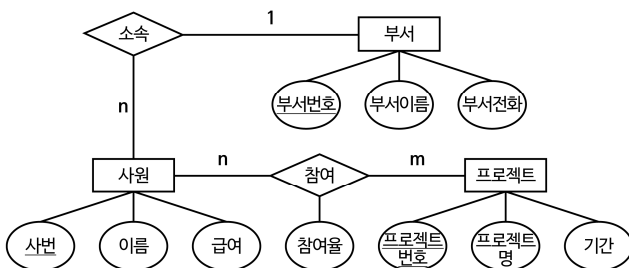
문 2. 외래키(foreign key)에 대한 설명으로 옳은 것은?

- ① 외래키는 중복값을 허용하지 않는다.
- ② 한 릴레이션의 외래키는 같은 릴레이션의 기본키를 참조할 수 있다.
- ③ 한 릴레이션의 외래키는 같은 릴레이션의 기본키의 일부가 될 수 없다.
- ④ 외래키의 도메인과 그 외래키가 참조하는 기본키의 도메인은 서로 다를 수 있다.

문 3. 인덱스에 대한 설명으로 옳지 않은 것은?

- ① B⁺-트리는 다단계 인덱스 구조이다.
- ② 기본키가 아닌 속성(필드)에도 인덱스를 생성할 수 있다.
- ③ 삽입과 삭제 연산이 빈번하게 발생하면 인덱스 유지를 위한 부하가 커진다.
- ④ 탐색하고자 하는 속성이 물리적으로 정렬되어 있지 않은 경우, 클러스터링 인덱스가 효율적이다.

문 4. 다음 E-R 다이어그램을 관계 데이터 모델 스키마로 변환할 때, 도출된 릴레이션으로 옳지 않은 것은?



- ① 사원(사번, 이름, 급여, 부서번호)
- ② 부서(부서번호, 부서이름, 부서전화)
- ③ 참여(사번, 프로젝트번호, 참여율)
- ④ 프로젝트(프로젝트번호, 프로젝트명, 기간, 사번)

문 5. 다음 상점 테이블과 상품 테이블을 생성하는 SQL 구문에 대한 설명으로 옳지 않은 것은?

```
CREATE TABLE 상점(
    ID          INTEGER,
    이름        VARCHAR(200),
    업종코드    INTEGER,
    주소        VARCHAR(200),
    PRIMARY KEY(ID));

CREATE TABLE 상품(
    ID          INTEGER,
    이름        VARCHAR(200),
    상점_ID     INTEGER,
    수량        INTEGER,
    단가        INTEGER,
    PRIMARY KEY(ID),
    FOREIGN KEY(상점_ID) REFERENCES 상점(ID));
```

- ① 상점 테이블의 ID 속성에 동일한 값을 갖는 두 개 이상의 튜플이 존재할 수 없다.
- ② 상점 테이블과 상품 테이블에 대해 자연 조인(natural join) 연산을 수행할 수 있다.
- ③ 상품 테이블의 상점_ID 속성에 NULL 값이 아닌 상점 테이블의 ID 속성에 존재하지 않는 값을 저장할 수 있다.
- ④ 상점 테이블의 이름 속성과 주소 속성에 길이가 200 미만인 문자열을 입력하더라도, 두 속성에서 저장 공간의 낭비가 발생하지 않는다.

문 6. 다음 함수 종속 관계가 있는 릴레이션 ‘학생(학번, 과목번호, 이름, 학과이름, 학과전화번호, 성적)’에 대한 설명으로 옳지 않은 것은? (단, 학생 릴레이션은 이미 제1정규형을 만족한다)

- 학번 → {이름, 학과이름, 학과전화번호}
- 학과이름 → 학과전화번호
- {학번, 과목번호} → 성적

- ① 함수적 종속관계 $X \rightarrow Y$ 가 성립하더라도, $Y \rightarrow X$ 가 항상 성립하는 것은 아니다.
- ② 학생 릴레이션을 분해하더라도 무손실분해(lossless decomposition)이어야 한다.
- ③ 학생 릴레이션이 제2정규형을 만족하도록 분해하면, (학번, 이름, 학과이름, 학과전화번호), (학번, 과목번호, 성적)이다.
- ④ 학생 릴레이션이 제3정규형을 만족하도록 분해하면, (학번, 이름), (학과이름, 학과전화번호), (학번, 과목번호, 성적)이다.

- 문 7. 은행계좌 X와 Y에 대한 트랜잭션 T1과 T2가 수행되는 다음 스케줄에 대한 설명으로 옳지 않은 것은? (단, 두 트랜잭션 수행 전 X계좌 잔액은 20 이상의 양수이고, Y계좌 잔액은 50 이상의 양수이다)

시간	T1	T2
	read(X) X := X + 50 write(X)	
		read(Y) Y := Y + 20 write(Y)
	read(Y) Y := Y - 50 write(Y)	
		read(X) X := X - 20 write(X)

- ① 사이클을 포함한 스케줄이다.
 ② 2단계 로킹 규약(two-phase locking protocol)으로 구현할 수 있다.
 ③ 두 트랜잭션 수행 전과 수행 완료 후 X계좌와 Y계좌 잔액의 합(X + Y)은 같다.
 ④ 두 트랜잭션 수행 완료 후 X계좌 잔액이 180이면, 두 트랜잭션 수행 전 X계좌 잔액은 150이다.

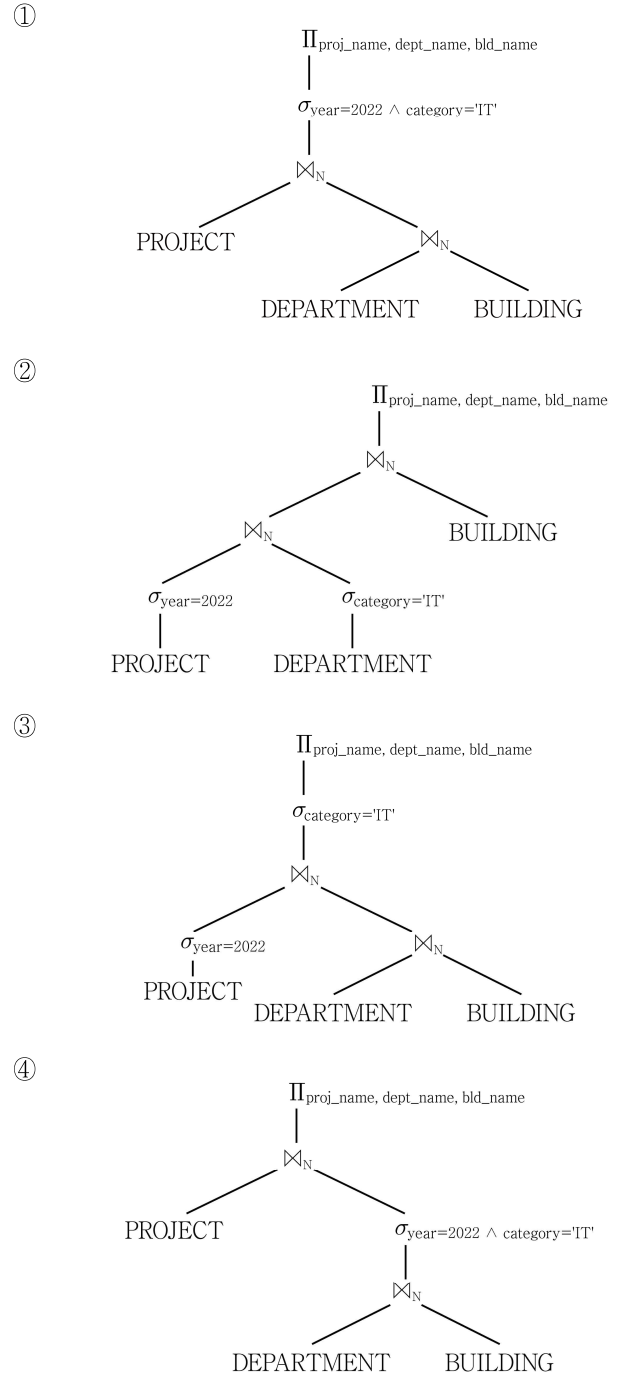
- 문 8. 200개의 트랜잭션에 대한 장바구니 분석을 통해 다음 조건을 만족하는 ‘기저귀→맥주’ 연관 규칙(association rule)을 도출하였다. 이 규칙의 지지도(support)와 신뢰도(confidence)를 순서대로 나열한 것은?

- 맥주는 50개의 트랜잭션에서 구매되었다.
- 기저귀는 40개의 트랜잭션에서 구매되었다.
- 기저귀와 맥주가 동시에 구매된 트랜잭션 수는 20개이다.

	지지도(%)	신뢰도(%)
①	10	50
②	40	45
③	45	40
④	50	10

- 문 9. 다음 릴레이션에 대하여 질의를 수행하고자 한다. 질의 수행 계획(query evaluation plan)의 의미가 다른 것은?

PROJECT (proj_id, proj_name, dept_id, year)
 DEPARTMENT (dept_id, dept_name, bld_id, category)
 BUILDING (bld_id, bld_name, location, capacity)



- 문 10. 릴레이션 R(B, C, G, H, I, S)이 함수 종속성 집합 F를 만족할 때, R의 슈퍼키(superkey)로 옳은 것은?

F = {B → H, CG → H, CG → I, S → B, S → C}

- ① H
 ② S
 ③ CG
 ④ GS

문 11. 관계 데이터베이스 관리 시스템(RDBMS)과 NoSQL에 대한 설명으로 옳은 것은?

- ① RDBMS는 NoSQL보다 약한 스키마를 요구한다.
- ② NoSQL은 RDBMS보다 엄격한 일관성 모델을 보장한다.
- ③ NoSQL은 RDBMS보다 정형 데이터를 저장하기에 적합하다.
- ④ NoSQL 데이터 모델로 키-값(key-value), 문서 기반(document-based), 그래프 기반(graph-based) 모델이 있다.

문 12. 관계 데이터베이스에서 NULL 값에 대한 설명으로 옳지 않은 것은?

- ① 한 개 이상의 NULL 값을 포함한 산술 연산 결과는 NULL이다.
- ② 릴레이션 튜플의 삽입 연산 수행 시, 결과 릴레이션의 속성은 NULL 값을 가질 수 있다.
- ③ NULL 값과 다른 값에 대해 대소 비교 연산자 ‘<’를 수행한 결과는 참(true) 혹은 거짓(false)이다.
- ④ NULL 값을 가지지 않은 속성들만으로 구성된 두 릴레이션에 대해 외부 조인(outer join) 연산을 수행한 경우, 결과 릴레이션의 속성은 NULL 값을 가질 수 있다.

문 13. 다음 중 데이터베이스 관리자(DBA)의 역할만을 모두 고르면?

- ㄱ. 데이터베이스 스키마 정의
 ㄴ. 저장 구조에 대한 접근 방식 결정
 ㄷ. 응용 프로그램 개발
 ㄹ. 데이터베이스 시스템 모니터링 및 성능 분석

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

문 14. 관계 데이터 모델의 릴레이션에 대한 설명으로 옳지 않은 것은?

- ① 릴레이션 R에 포함된 튜플들 사이에는 순서가 없다.
- ② 릴레이션 R의 모든 튜플은 서로 다른 값을 가지고 있다.
- ③ 릴레이션의 인스턴스는 시간에 따라 변하지 않는 정적인 성질을 가지고 있다.
- ④ 릴레이션 스키마에 정의된 속성의 전체 개수를 릴레이션의 차수(degree)라고 한다.

문 15. 데이터베이스 회복에 대한 설명으로 옳은 것은?

- ① 지연 갱신 회복기법에서는 UNDO 연산이 필요 없다.
- ② REDO 연산을 수행할 때는 로그(log)가 필요 없다.
- ③ 검사점(checkpoint) 연산은 장애가 발생할 때 실행된다.
- ④ 즉시 갱신 회복기법에서는 트랜잭션이 완료한 이후 데이터베이스를 갱신한다.

문 16. 다음 Student, Course, Enrollment 테이블에 대한 SQL 구문들을 수행하였을 때, COUNT(*) 값은? (단, Enrollment 테이블 생성 시 sno 속성과 cno 속성은 각각 Student 테이블과 Course 테이블 기본키의 외래키이며, 두 외래키에 모두 ON DELETE CASCADE 조건이 적용된다)

Student			
sno	sname	year	dept
100	John	4	Com
200	David	3	EE
300	Tom	1	Com
400	Mark	4	Com
500	Simon	2	ME

Course				
cno	cname	credit	dept	professor
C123	Java	3	Com	Andy
C312	Data Structure	3	Com	Joseph
C324	File System	3	Com	Helen
C413	Database	3	Com	Luna
E412	OS	3	EE	Elsa

Enrollment				
sno	cno	grade	midterm	final
200	C123	B	85	80
300	C312	A	90	95
400	C312	A	90	95
500	C312	B	85	80
300	C324	C	75	75
400	C324	A	95	90
100	C413	A	90	95
300	C413	A	95	90
400	C413	B	80	85
100	E412	A	95	95
400	E412	C	65	75

SQL 구문

```
DELETE FROM Student WHERE sname= 'Simon';
DELETE FROM Course WHERE cname= 'OS';
SELECT COUNT(*) FROM Enrollment;
```

- ① 6
- ② 7
- ③ 8
- ④ 9

문 17. 다음 직원 테이블과 부서 테이블로 구성된 데이터베이스에서 '기본 인덱스를 사용하여 여러 개의 레코드를 검색하는 방식'으로 구현되는 SELECT문으로 옳은 것은? (단, 직원 테이블의 부서번호 속성은 부서 테이블 기본키의 외래키이다)

- 직원(직원번호, 이름, 봉급, 부서번호)
○ 부서(부서번호, 부서명, 관리자)

- ① SELECT * FROM 직원 WHERE 부서번호 = 5;
② SELECT * FROM 부서 WHERE 부서번호 > 5;
③ SELECT * FROM 직원 WHERE 직원번호 = 100;
④ SELECT * FROM 직원 WHERE 부서번호 = 5 AND 봉급 > 30000;

문 18. 다음 수강 릴레이션에 대한 설명으로 옳지 않은 것은? (단, 수강 릴레이션에 기본키가 설정되어 있으며, 한 강사는 한 과목만 강의할 수 있고, 한 학생은 서로 다른 한 개 이상의 과목들을 수강할 수 있다)

수강		
학번	강사명	과목명
100	John	데이터베이스
100	Marry	네트워크
200	Jeremy	데이터베이스
300	Iaان	인공지능
300	John	데이터베이스

- ① {강사명, 과목명}은 후보키가 아니다.
② 강사 'James'가 강의할 '빅데이터' 과목을 신설하는 경우 수강 학생이 없이 투플을 삽입할 수 있다.
③ 학번이 100인 학생이 '네트워크' 과목 수강을 취소하여 그 수강 투플이 삭제되는 경우, '네트워크' 과목의 강사명도 함께 사라진다.
④ '데이터베이스' 과목의 강사를 'John'에서 'Tony'로 변경한다면, 'John'이 강의하는 과목을 수강하는 모든 학생의 강사명을 수정해야 한다.

문 19. 다음과 같이 실행되고 있는 트랜잭션 T1 ~ T4의 스케줄과 동등한 (equivalent) 충돌 직렬가능(conflict serializable) 스케줄이 아닌 것은?

시간	T1	T2	T3	T4
			read(Z) write(Z)	
	read(X)	write(Y)		
	write(X)			read(Z) write(Z)
				read(Y) write(Y)

- ① T1 → T3 → T2 → T4
② T2 → T3 → T4 → T1
③ T2 → T1 → T4 → T3
④ T3 → T2 → T4 → T1

문 20. 트랜잭션 T1과 T2가 병행 수행될 때, 정확한 결과를 보장할 수 있는 스케줄만을 모두 고르면? (단, 두 트랜잭션 수행 전 A와 B의 값은 각각 50이다)

ㄱ. 시간	T1	T2	ㄴ. 시간	T1	T2
		read(A)		read(A)	
		A := A×2		A := A+100	
		write(A)			read(A)
		read(B)			A := A×2
		B := B×2			write(A)
		write(B)		write(A)	
	read(A)				read(B)
	A := A+100				B := B×2
	write(A)			read(B)	
	read(B)			B := B+100	
	B := B+100			write(B)	
	write(B)				write(B)

ㄷ. 시간	T1	T2	ㄹ. 시간	T1	T2
	lock(A)			lock(A)	
	read(A)			read(A)	
	A := A+100			A := A+100	
	write(A)			write(A)	
	unlock(A)			lock(B)	
		lock(A)		unlock(A)	
		read(A)			lock(A)
		A := A×2			read(A)
		write(A)			A := A×2
		unlock(A)			write(A)
		lock(B)		read(B)	
		read(B)		B := B+100	
		B := B×2		write(B)	
		write(B)		unlock(B)	
		unlock(B)			lock(B)
	lock(B)				unlock(A)
	read(B)				read(B)
	B := B+100				B := B×2
	write(B)				write(B)
	unlock(B)				unlock(B)

- ① ㄱ, ㄴ
② ㄱ, ㄹ
③ ㄴ, ㄷ
④ ㄷ, ㄹ

문 21. 분산 데이터베이스 시스템에 대한 설명으로 옳지 않은 것은?

- ① 사용자들은 원하는 데이터가 어느 지역에 저장되어 있는지를 명확하게 알 수 있다.
- ② 데이터베이스를 여러 지역에 중복하여 분산 저장하면 가용성(availability)이 높아진다.
- ③ 지역 사이트들도 어느 정도 자치성(autonomy)을 가지고 독립적으로 동작할 수 있다.
- ④ 수평적 단편화(horizontal fragmentation)는 릴레이션 튜플들을 몇 개의 부분집합(단편)으로 분할한 후, 각 부분집합을 서로 다른 사이트에 저장하는 것이다.

문 22. 다음 두 트랜잭션이 테이블 R(A, B)에 대해 병행 수행되고 있다. 트랜잭션 T2에서 “select sum(B) from R;”의 결과가 항상 동일한 값을 출력하도록 보장하는 트랜잭션 고립성 수준(isolation level)은? (단, 테이블 R의 속성 A, B의 도메인은 정수이다)

T1	T2
begin transaction; insert into R values (3, 150); commit;	begin transaction; select sum(B) from R; select sum(B) from R; commit;

- ① 직렬가능(SERIALIZABLE)
- ② 반복 가능한 읽기(REPEATABLE READ)
- ③ 커밋된 데이터 읽기(READ COMMITTED)
- ④ 커밋되지 않은 데이터 읽기(READ UNCOMMITTED)

문 23. 다음 관계 대수식에서 동등 규칙(equivalence rule)이 성립하는 것만을 모두 고르면? (단, E, E_1, E_2 는 관계대수식, $\theta, \theta_1, \theta_2$ 는 술어(predicate) 조건을, L_1, L_2 는 속성을 의미한다)

- ㄱ. $\Pi_{L_1}(\Pi_{L_2}(E)) \equiv \Pi_{L_2}(E)$
- ㄴ. $\sigma_{\theta_1 \wedge \theta_2}(E) \equiv \sigma_{\theta_2}(\sigma_{\theta_1}(E))$
- ㄷ. $\sigma_{\theta}(E_1 - E_2) \equiv \sigma_{\theta}(E_1) - E_2$
- ㄹ. $E_1 \bowtie E_2 \equiv E_2 \bowtie E_1$

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

문 24. 트랜잭션 T1 ~ T4가 다음 연산과 조건에 따라 수행하는 아래 스케줄에 대한 설명으로 옳은 것은?

- S(A): 데이터 A에 읽기 연산(R(A))을 위한 공유 로크(shared lock) 요청
- X(A): 데이터 A에 쓰기 연산(W(A))을 위한 전용 로크(exclusive lock) 요청
- 로크를 요청한 트랜잭션은 요청한 로크가 허용될 때까지 대기한다.
- 트랜잭션이 소유한 로크는 트랜잭션이 완료될 때 해제된다.

시간	T1	T2	T3	T4
	S(A) R(A)			
		X(C) W(C)		
				X(D) W(D) S(B) R(B) X(C)
		X(A)		
			X(C)	
	X(B)			

- ① 모든 트랜잭션 T1 ~ T4가 교착상태에 빠졌다.
- ② 트랜잭션 T3을 철회(abort)하면 교착상태가 해결될 수 있다.
- ③ 교착상태를 해결하기 위해 철회할 트랜잭션으로 T4를 선택하는 것이 가장 좋다.
- ④ 트랜잭션 T2를 철회하면, 트랜잭션 T3은 C에 대한 전용 로크를 획득할 수 있다.

문 25. 다음은 OLAP(Online Analytical Processing)의 기능에 대한 설명이다. (가), (나)에 들어갈 용어를 바르게 연결한 것은?

- (가)은 교차 테이블(cross-tab)에서 데이터를 분석하는 차원을 다양하게 변경하는 기능이다.
- 집계 연산에 대해 사용자가 원하는 크기의 데이터로 보여주기 위해, 미세한 단위부터 큰 단위로 변화시키는 연산을 (나)이라고 한다.

(가)

(나)

- ① 피보팅(pivoting) 롤업(roll-up)
- ② 피보팅(pivoting) 드릴다운(drill-down)
- ③ 슬라이싱(slicing) 롤업(roll-up)
- ④ 슬라이싱(slicing) 드릴다운(drill-down)