

【데이터베이스론】

1. 다음은 데이터베이스 관리 시스템(DBMS)의 주요 구성요소에 대한 설명이다. 옳은 것을 모두 고른 것은?

- ㉠ 데이터정의어(DDL) 컴파일러는 DDL로 명세된 스키마 정의를 내부 형태로 처리하여 시스템 카탈로그에 저장한다.
- ㉡ 질의처리기는 질의문을 파싱, 분석, 컴파일하여 런타임 처리기를 통해 실행한다.
- ㉢ 데이터조작어(DML) 컴파일러는 DML 명령어를 추출 및 파싱하여 DB의 내용에 대한 정확성 및 안정성 유지 기능을 수행한다.
- ㉣ 트랜잭션 관리자는 무결성 제약조건을 준수하여 데이터 접근 권한 등의 통제수행, 동시성제어, 회복작업 등을 수행한다.

- ① ㉠㉡㉢ ② ㉡㉣ ③ ㉠㉡㉣ ④ ㉢㉣

2. 다음의 15개 숫자 데이터에서 이진 검색 알고리즘을 활용하여 숫자 2를 검색하고자 한다. 이에 대한 설명으로 가장 적절한 것은?

(정렬 전)

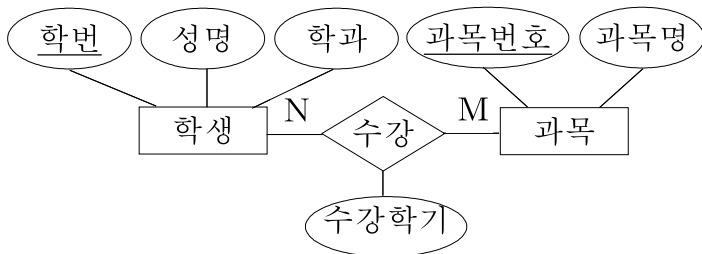
11	3	1	6	10	5	14	12	4	15	7	13	8	2	9
----	---	---	---	----	---	----	----	---	----	---	----	---	---	---

(정렬 후)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----

- ① 정렬 후 데이터를 기준으로 2를 찾는 비교횟수는 4회이다.
- ② 정렬 후 데이터를 기준으로 비교횟수를 거듭할 때마다 검색 대상이 되는 데이터의 수가 약 반(1/2)으로 줄어든다.
- ③ 피보나치 수열에 따라 다음에 비교할 대상을 선정하여 검색한다.
- ④ 15개의 숫자 데이터들은 정렬되어 있지 않아도 항상 검색이 가능하다.

3. 다음 ERD(Entity Relationship Diagram)를 관계형 데이터베이스 스키마로 가장 적절하게 변환한 것은? (단, 밑줄()은 기본키를 의미한다.)



- ① 학생(학번, 성명, 학과), 과목(과목번호, 과목명), 수강(학번, 과목번호)
- ② 학생(학번, 성명, 학과), 과목(과목번호, 과목명), 수강(학번, 과목번호)
- ③ 학생(학번, 성명, 학과), 과목(과목번호, 과목명), 수강(학번, 과목번호, 수강학기)
- ④ 학생(학번, 성명, 학과), 과목(과목번호, 과목명), 수강(학번, 과목번호, 수강학기)

4. 다음은 무결성의 종류에 대한 설명이다. (가)~(라)에 해당되는 것으로 가장 알맞게 짝지은 것은?

(가)	릴레이션은 참조할 수 없는 외래키 값을 가질 수 없다.
(나)	속성 값은 그 속성의 도메인 범위 내에 존재해야 한다.
(다)	기본키에 속해 있는 속성은 널(NULL) 값을 가질 수 없다.
(라)	한 릴레이션에 같은 키값을 가진 튜플들은 허용하지 않는다.

- | | | | |
|-----------|---------|---------|---------|
| (가) | (나) | (다) | (라) |
| ① 개체 무결성 | 키 무결성 | 참조 무결성 | 도메인 무결성 |
| ② 참조 무결성 | 도메인 무결성 | 개체 무결성 | 키 무결성 |
| ③ 도메인 무결성 | 키 무결성 | 참조 무결성 | 개체 무결성 |
| ④ 참조 무결성 | 개체 무결성 | 도메인 무결성 | 키 무결성 |

5. 다음의 두 릴레이션 R과 S에 대한 조인(Join) 결과로 옳지 않은 것은?

R			S		
A	B	C	B	C	D
2	나	b	나	b	2
3	다	c	나	c	3
4	라	d	라	d	4
1	가	a	가	a	1

① R ⋈ S (Natural Join)

A	B	C	D
2	나	b	2
4	라	d	4
1	가	a	1

② R ⋈ S (Left Outer Join)

A	B	C	B	C	D
2	나	b	나	b	2
3	다	c	NULL	NULL	NULL
4	라	d	라	d	4
1	가	a	가	a	1

③ R ⋈ S (Right Outer Join)

A	B	C	B	C	D
2	나	b	나	b	2
NULL	NULL	NULL	나	c	3
4	라	d	라	d	4
1	가	a	가	a	1

④ R ⋈ S (Full Outer Join)

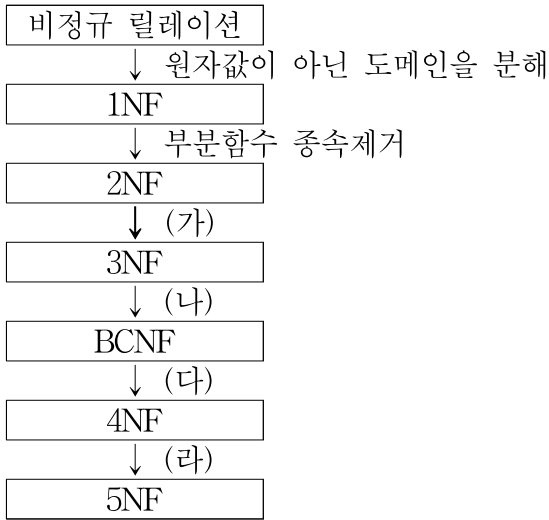
A	B	C	B	C	D
3	다	c	NULL	NULL	NULL
NULL	NULL	NULL	나	c	3

6. 다음의 릴레이션 tbl_salary(index, name, address, age, salary)에 대해, 여러 관계 대수식을 수행하려고 한다. 관계 대수식의 결과로 생성되는 릴레이션의 차수(Degree)와 카디널리티(Cardinality)로 옳지 않은 것은?

index	name	address	age	salary
1	차	서울	35	350
2	경	대전	27	270
3	강	경남	31	320
4	김	경기	28	270
5	박	경기	29	290
6	최	부산	20	230
7	오	전북	23	250

- ① 대수식: $\sigma_{\text{address}='경기'}(\text{tbl_salary}) \cup \sigma_{\text{address}='서울'}(\text{tbl_salary})$
(결과) 차수: 5, 카디널리티: 3
- ② 대수식: $\sigma_{\text{salary} \geq 300}(\text{tbl_salary}) - \sigma_{\text{address}='경기'}(\text{tbl_salary})$
(결과) 차수: 5, 카디널리티: 2
- ③ 대수식: $\Pi_{\text{name}, \text{address}}(\sigma_{\text{age} > 29}(\text{tbl_salary}) \cup \sigma_{\text{age} \leq 23}(\text{tbl_salary}))$
(결과) 차수: 3, 카디널리티: 4
- ④ 대수식: $\Pi_{\text{name}, \text{age}, \text{salary}}(\sigma_{\text{age} > 28}(\text{tbl_salary}))$
(결과) 차수: 3, 카디널리티: 3

7. 다음은 정규화 과정을 도식화한 것이다. 정규화 과정 중 (가)~(라)에 해당되는 과정을 <보기>에서 골라 알맞게 나열한 것은?



- <보 기>
- ㉠ 후보키를 통하지 않은 조인종속 제거
 ㉡ 결정자가 후보키가 아닌 함수종속 제거
 ㉢ 이행함수 종속제거
 ㉣ 함수종속이 아닌 다치종속 제거

- (가) (나) (다) (라)
 ① ㉢ ㉡ ㉠
 ② ㉡ ㉠ ㉢
 ③ ㉡ ㉢ ㉠
 ④ ㉢ ㉣ ㉡

8. 다음의 릴레이션 R에 존재하는 함수 종속성을 나타낸 것으로 옳은 것은?

R

A	B	C	D	E
1	c	20	a	1
1	b	30	b	3
2	c	20	a	2
2	c	20	a	3
2	a	30	c	2
3	b	20	b	2
3	b	30	b	1

- ① B→DE ② A→BC ③ AB→E ④ CD→B

9. 다음의 tbl_student 테이블을 이용하여 아래의 SQL을 수행하였을 때 예상되는 결과로 옳은 것은?

id	name	grade	subject	score
2017001	Ryu	2	math	60
2017002	Cho	1	kor	80
2019006	Kim	1	kor	55
2018002	Yang	3	eng	85
2018004	Park	2	math	45
2016003	Choi	3	eng	55
2016006	Kang	3	eng	60

SELECT count(*) FROM tbl_student
GROUP BY subject HAVING count(*) > 2;

- ① 0 ② 1 ③ 2 ④ 3

10. 다음의 SQL에서 명령어 짝의 사용이 적절하지 않은 것을 모두 고른 것은?

- ㉠ UPDATE ... / AS ...
 ㉡ CREATE VIEW ... / SET ...
 ㉢ SELECT.../FROM.../WHERE.../GROUP BY... /HAVING ...

- ① ㉠ ② ㉠㉡ ③ ㉡㉢ ④ ㉠㉢

11. 다음의 뷰에 대한 설명으로 적절하지 않은 것을 모두 고른 것은?

- ㉠ 관계형 데이터베이스에서 뷰는 사용자의 전체 외부 스키마 대신에 하나의 가상 릴레이션(Virtual Relation)을 의미한다.
 ㉡ 뷰에 대한 검색은 기본 테이블과 동일하기 때문에 삽입, 삭제, 갱신의 경우 제약을 받지 않는다.
 ㉢ 뷰가 기본키를 포함한 행과 열의 부분 집합으로 구성되어도 갱신 연산은 수행이 불가능하다.
 ㉣ 데이터베이스의 구조가 바뀌어도 뷰를 정의하는 질의만 변경하면 되므로, 기존의 응용 프로그램을 다시 작성할 필요성을 줄일 수 있다.
 ㉤ 뷰는 독자적으로 인덱스를 가질 수 있으며, 뷰의 정의를 쉽게 변경할 수 있다.

- ① ㉠㉡㉤ ② ㉡㉢㉤ ③ ㉠㉢㉣ ④ ㉢㉣㉤

12. 물리적 데이터베이스 설계 시 고려해야 할 사항으로 가장 적절하지 않은 것은?

- ① 저장 레코드 형식 설계 및 저장 레코드 집중화
 ② 수행될 질의와 트랜잭션의 예상 빈도
 ③ 질의문의 조건 절에서 자주 사용하는 애틀리뷰트에 대한 인덱스 생성
 ④ 클러스터 인덱스를 여러 개의 인덱스로 정의

13. 데이터베이스 설계 단계의 튜닝과정에서 고려해야 할 사항에 대한 설명으로 가장 적절하지 않은 것은?

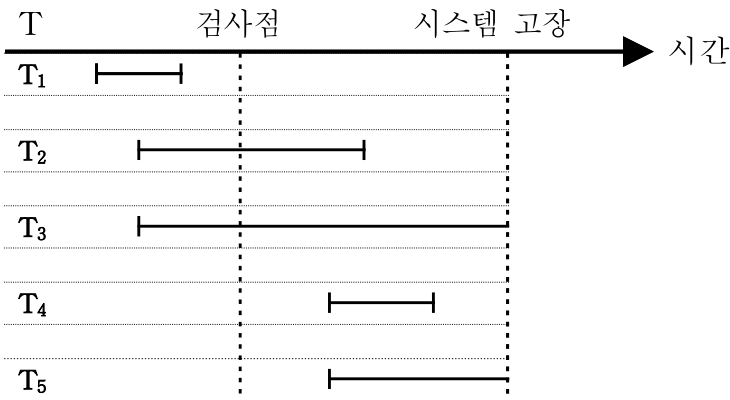
- ① 성능 향상을 위해 릴레이션을 수직 또는 수평으로 분할할 수 있다.
 ② 성능 향상을 위해 롤백 세그먼트를 하나로 통합하여 구성할 수 있다.
 ③ 자주 함께 조인되는 테이블들을 역정규화하여 조인 경로를 줄일 수 있다.
 ④ 분산 데이터베이스를 적용했을 경우, 원격 데이터베이스를 이용할 때 성능 저하가 예상된다면 스냅샷을 이용한 복제 테이블 생성을 고려할 수 있다.

14. 다음의 RAID(Redundant Array of Inexpensive Disks)에 대한 설명으로 적절한 것을 모두 고른 것은?

- ㉠ RAID는 디스크 시스템의 성능을 향상시키기 위해 여러 개의 디스크를 병렬로 연결하여 하나의 고성능 논리적 디스크처럼 동작한다.
 ㉡ RAID에서 디스크 시스템의 성능 향상은 데이터를 다중 디스크에 투명하게 분산시키는 미러링(Mirroring) 기법을 통해 구현한다.
 ㉢ RAID 0(Stripping)은 데이터의 빠른 입출력을 위해 데이터를 여러 드라이브에 분산 저장하므로 읽기 및 쓰기 성능이 뛰어나지만 복구를 위한 추가적인 정보를 기록하지 않아 장애가 발생하면 데이터는 모두 손실된다.
 ㉣ RAID 3(Pararell Transfer with Parity)은 RAID 0의 빠른 속도와 RAID 1의 안정적인 복구 기능을 합쳐 놓은 방식이다.
 ㉤ RAID 5(Independent Data Disks with Distributed Parity Blocks)는 패리티 정보를 모든 드라이브에 나눠 기록하기 때문에 패리티를 담당하는 디스크가 병목현상을 일으키지 않아 다중 처리 시스템에 보다 적합하다.

- ① ㉠㉡㉢ ② ㉠㉢㉤ ③ ㉠㉡㉣ ④ ㉢㉣㉤

15. 다음의 검사점(Checkpoint)이 있는 로그(Log)에서 시스템 고장이 발생하는 경우, 트랜잭션 T에 대한 Undo-list와 Redo-list를 바르게 나열한 것은?



- | | | |
|---|------------------|------------------|
| | Undo-list | Redo-list |
| ① | {T3, T5} | {T2, T4} |
| ② | {T3, T4, T5} | {T1, T2} |
| ③ | {T1, T3, T5} | {T2, T4} |
| ④ | {T3, T5} | {T1, T2, T4} |

16. 트랜잭션의 병행제어를 위해 사용하는 로킹(Locking)기법에 대한 설명 중 가장 적절하지 않은 것은?

- ① 데이터 A에 대하여 트랜잭션 T의 배타 잠금(Lock)이 설정되어 있다면 그 잠금이 해제될 때까지는 다른 어떤 트랜잭션도 데이터 A에 대하여 잠금을 설정하지 못하고 대기해야 한다.
- ② 특정 스케줄에 참여하고 있는 모든 트랜잭션들이 2단계 로킹 기법(2PL: 2Phase Locking)을 준수한다면, 그 스케줄은 직렬 가능하므로 항상 옳다고 간주할 수 있다.
- ③ 엄격한 2PL 기법의 경우 엄격한 관리를 통하여 교착상태(Deadlock)를 막을 수 있다.
- ④ 2PL은 크게 확장단계(Growing Phase)와 수축단계(Shrinking Phase)로 구성된다.

17. 테이블 student에 대한 권한을 부여하는 ㉠과 회수하는 ㉡을 SQL로 바르게 표현한 것은?

- ㉠ 테이블 student에 대한 SELECT, INSERT 권한을 Kim과 Lee에게 부여한다.

㉡ 테이블 student에 대한 SELECT, INSERT 권한을 Lee로 부터 회수한다.

- ① ㉠ GRANT SELECT, INSERT ON student TO Kim, Lee;
㉡ REVOKE SELECT, INSERT ON student FROM Lee;
- ② ㉠ AUTHORIZE SELECT, INSERT ON student TO Kim, Lee;
㉡ CANCEL SELECT, INSERT ON student FROM Lee;
- ③ ㉠ AUTHORIZE student TO Kim, Lee ON SELECT, INSERT;
㉡ CANCEL student TO Lee ON SELECT, INSERT;
- ④ ㉠ GRANT student TO Kim, Lee ON SELECT, INSERT;
㉡ REVOKE student TO Lee ON SELECT, INSERT;

18. 분산 데이터베이스 시스템의 장점에 대한 설명으로 가장 적절하지 않은 것은?

- ① 데이터 처리량을 증가시켜야 할 경우 중앙 집중 시스템은 보다 큰 용량의 시스템으로 대체하는 것이 일반적이지만 분산 시스템에서는 새로운 노드를 추가함으로써 가능하게 할 수 있다.
- ② 어느 한 사이트의 장애나 개별적인 통신 장애가 있을 경우 성능이 저하되기는 하지만 그 기능을 계속 수행할 수 있다.
- ③ 중앙집중형 데이터베이스 시스템보다 구현이 간단하기 때문에 소프트웨어 개발비용을 더 많이 줄일 수 있다.
- ④ 임의의 질의어가 여러 사이트에 있는 데이터를 이용한다면, 여러 개의 부 질의로 분할하여 병렬로 실행할 수 있다.

19. 다음은 데이터마이닝의 응용 프로그램에 대한 설명이다. (가)~(라)에 해당되는 것으로 알맞게 짝지은 것은?

(가)	데이터를 서로 다른 부류 혹은 클래스로 나누어주는 모델을 학습하는 과정
(나)	유사한 레코드들은 같은 그룹에 배치하고 서로 다른 레코드들은 다른 그룹에 배치시키고, 그룹들 간에는 서로 겹치지 않음
(다)	연속성을 갖는 변수의 미래 값을 예측하는 방법
(라)	인간이 경험으로부터 학습해 가는 두뇌의 신경망 활동에 착안한 기법

- | | | | | |
|---|---------------------|---------------------|------------------|------------------|
| | (가) | (나) | (다) | (라) |
| ① | classification rule | clustering | regression | neural networks |
| ② | clustering | classification rule | association rule | regression |
| ③ | classification rule | association rule | regression | neural networks |
| ④ | clustering | classification rule | neural networks | association rule |

20. 다음과 같은 과목수강 트랜잭션에 대해 연관규칙 탐색을 수행한 경우, {OS, PL}→{DB}에 대한 지지도(Support)와 신뢰도(Confidence)는 각각 얼마인가?

TID	OS	PL	CA	DB
1	1	0	0	1
2	0	1	1	1
3	1	1	1	1
4	1	1	1	1
5	1	1	1	0
6	1	1	0	1

- ① 지지도 0.67, 신뢰도 0.5
- ② 지지도 0.75, 신뢰도 0.5
- ③ 지지도 0.5, 신뢰도 0.75
- ④ 지지도 0.5, 신뢰도 0.67