

데이터베이스론

2020년도 국가공무원 5급(기술) 공개경쟁채용 제2차시험

응시번호 :

성명 :

제 1 문. 정규화에 대한 다음 물음에 답하시오.

(총 20점)

- 1) 암스트롱이 제안한 3가지 기본 공리(axiom)는 재귀(reflexive) 규칙, 증가(augmentation) 규칙, 이행(transitive) 규칙이다. 이 3가지 공리만을 이용하여 가이행 규칙(pseudotransitivity rule), $X \rightarrow Y$ 와 $WY \rightarrow Z$ 이면 $WX \rightarrow Z$ 임을 증명하시오. (5점)
- 2) 릴레이션 $R(A, B, C, D, E, F)$ 과 함수종속 집합 $F = \{A \rightarrow B, A \rightarrow C, BC \rightarrow E, CD \rightarrow F, B \rightarrow E\}$ 가 주어졌을 때, $AB \rightarrow CD$ 가 F 의 폐포(F^+)에 속하는지 여부를 판별하고, 이유를 설명하시오. (5점)
- 3) 릴레이션 $R(A, B, C)$ 과 함수종속 집합 $F = \{C \rightarrow A, AB \rightarrow C\}$ 가 주어졌을 때, 주어진 릴레이션 R 의 최대 정규형을 판별하고, 이유를 설명하시오. (단, 릴레이션 R 은 제1정규형을 만족한다고 가정한다) (5점)
- 4) 3)에서 주어진 릴레이션 R 과 함수종속 집합 F 에 대하여 갱신 이상(update anomaly)이 발생하지 않도록 정규화하시오. 그리고 분해된 릴레이션들이 종속성 보존(dependency preservation)이 되는지를 판별하고, 이유를 설명하시오. (5점)

제 2 문. 기업 프로젝트를 관리하는 데이터베이스가 다음과 같이 직원, 부서, 프로젝트, 업무의 릴레이션으로 구성되어 있다. 직원 릴레이션은 기업 내의 직원에 대한 정보를, 부서 릴레이션은 기업 내의 부서에 대한 정보를, 프로젝트 릴레이션은 프로젝트에 대한 정보를, 그리고 업무 릴레이션은 직원이 프로젝트에 참여한 정보를 나타낸다. 한 직원은 여러 개의 프로젝트에 참여할 수 있고 한 프로젝트에 여러 직원이 참여할 수 있을 때, 다음 물음에 답하시오. (단, 각 릴레이션의 기본키는 밑줄로 표시되어 있다) (총 20점)

직원(직원번호, 직원이름, 전화번호, 주소, 직급, 소속부서번호)

부서(부서번호, 부서이름, 부서장직원번호)

프로젝트(과제번호, 과제명, 부서번호)

업무(직원번호, 과제번호, 과제수행시간)

- 1) 주어진 릴레이션의 의미와 그들 사이의 관계를 살펴보고, 각 릴레이션에서 외래키(foreign key)를 모두 찾으시오. (5점)
- 2) 'IoT융합부'에서 일하는 직원의 직원이름, 직급, 소속부서번호를 찾는 관계대수식을 작성하시오. (단, 관계대수식은 프로젝션(projection), 선택(selection), 세타조인(theta join)만을 이용하여 작성한다) (5점)
- 3) 과제명이 'A'인 프로젝트에 참여하는 직원이름을 찾는 관계대수식을 세타조인, 선택, 프로젝션 순서로 수행하도록 작성하시오. 그리고 작성된 관계대수식에 대하여 중간 생성파일의 크기를 줄이기 위한 최적화된 관계대수식을 구하시오. (단, 관계대수식은 선택, 프로젝션, 세타조인만을 이용하여 작성한다) (10점)

제 3 문. 어떤 회사의 판매정보를 관리하는 데이터베이스가 다음과 같이 네 개의 테이블(table)과 하나의 뷰(view)로 구성되어 있다. 아래에 제시된 SQL 문이 각각 실행될 때 “성공” 또는 “실패” 여부와 그 이유를 설명하시오. (총 20점)

데이터베이스 스키마		
<pre> create table Customers (custnum varchar2(4) NOT NULL, cname varchar2(10), phone varchar2(15), primary key (custnum)) ; create table Products (prodnum varchar2(6) NOT NULL, pname varchar2(30), unitprice number(8), stock number(4), primary key (prodnum)) ; create table Orders (ordnum number(7) NOT NULL, custnum varchar2(4), orddate date, amount number(12), primary key (ordnum), foreign key (custnum) references Customers (custnum) on delete set NULL) ; create table Order_details (ordnum number(7) NOT NULL, prodnum varchar2(6) NOT NULL, quantity number(9), primary key (ordnum, prodnum), foreign key (ordnum) references Orders (ordnum) on delete cascade, foreign key (prodnum) references Products (prodnum)) ; create view Products_5000_view as select * from Products where unitprice >= 5000 ; </pre>		

판매 데이터베이스			
Customers			
custnum	cname	phone	
c1	김우진	010-0001-1111	
c2	장주권	010-0002-2222	
c3	김장명	010-0003-3333	
c4	박송준	010-0004-4444	
Products			
prodnum	pname	unitprice	stock
p1	B5	1,800	50
p2	60매	2,000	50
p3	90매	2,500	50
p4	삼색볼펜	3,000	100
Orders			
ordnum	custnum	orddate	amount
1801	c1	2018-06-10	3,600
1802	c2	2018-06-11	13,600
1805	c3	2018-06-15	15,000
1809	c4	2018-06-25	6,800
Order_details			
ordnum	prodnum	quantity	
1801	p1	2	
1802	p1	2	
1802	p2	5	
1805	p4	2	
1809	p1	1	

- 1) delete from Products where prodnum = 'p4' ; (5점)
- 2) insert into Orders values
(1809, 'c10', to_date('2018-06-25', 'yyyy-mm-dd'), 16500) ; (5점)
- 3) update Products
set prodnum = 'p11' where prodnum = 'p3' ; (5점)
- 4) insert into Products_5000_view values
('p20', '캠퍼스노트 유선', 2000, 50) ; (5점)

제 4 문. 다음은 날씨와 바람 속성 값에 따라 특정 게임의 진행여부(게임여부)를 보여주는 테이블이다. 어떤 조건이 주어졌을 때, 해당 게임의 진행여부를 결정하기 위한 결정 트리(decision tree)를 생성하고자 한다. 생성 과정에서 각 노드의 불순도(impurity)는 지니(Gini) 계수로 계산하고, 속성의 중요도는 정보 이득(information gain)으로 판정한다. 다음 물음에 답하시오. (단, 지니 계수는 $1 - \sum_{i=1}^c p(\frac{i}{t})^2$ 이고, 이 때 c 는 클래스 개수, $p(\frac{i}{t})$ 는 노드 t 에 존재하는 데이터 중에서 클래스 i 에 속하는 비율을 의미한다)

날씨	바람	게임여부
맑음	없음	no
맑음	있음	no
흐림	없음	yes
비	없음	yes

- 결정 트리의 루트 노드에 위치할 속성을 결정하고, 결정 과정을 설명하시오. (20점)
- 결정 트리를 완성하고 이를 이용하여 새로운 조건(날씨 = 흐림, 바람 = 있음)이 주어졌을 때 해당 게임의 진행여부를 결정하시오. (5점)

제 5 문. 트랜잭션 고립성 수준(isolation level)에 대한 다음 물음에 답하시오. (총 15점)

- 트랜잭션들이 수행되는 동안 반복할 수 없는 읽기(non-repeatable read)는 어떤 경우에 발생하는지 예를 들어 설명하시오. (5점)
- 2단계 로킹 규약(2-Phase Locking Protocol: 2PL)에서 팬텀 읽기(phantom read) 문제를 방지하기 위하여 2PL 기법을 어떻게 수정해야 하는지 설명하시오. (5점)
- 로킹 기법에 기반하여 트랜잭션들을 병행 수행하고자 한다. 높은 수준의 고립성 수준을 사용하는 경우 발생하는 단점을 설명하시오. (5점)