

1장(데이터베이스 시스템의 개요) 연습문제 해답

1. 다음 용어들을 설명하시오.

데이터베이스, DBMS, 데이터 모델, 데이터베이스 시스템

- **데이터베이스(DB: database)** : 조직체의 응용 시스템들이 공용으로 사용하는 운영 데이터들이 구조적으로 통합하여 저장된 데이터로 컴퓨터 내부의 하드디스크에 저장된다.
- **데이터베이스 시스템** : 데이터베이스 관리 시스템(DBMS: database management system), 데이터베이스, 데이터모델 이 3가지로 구성된 시스템을 말한다.
- **데이터베이스 관리 시스템(DBMS: database management system)** : 소프트웨어 프로그램으로 데이터베이스를 관리하는 기능을 하며, 사용자에게 편리하고 효율적인 데이터베이스 사용 환경을 제공한다. 컴퓨터 주기억 장치에 상주한다.
- **데이터 모델(data model)** : 데이터베이스 시스템에서 데이터를 저장하는 이론적 방법에만 관한 것으로, 데이터베이스에 데이터가 어떻게 구조화되어 저장되는 지를 결정한다. 데이터 모델은 데이터, 연산자, 무결성 제약조건을 요소로 갖는다.

2. 데이터베이스의 개념 네 가지를 설명하시오.

- 조직체의 모든 구성원이 공용으로 사용하는 **공용 데이터**
- 조직의 업무를 목적으로 사용되는 **운영 데이터**
- 데이터베이스에서 모든 데이터는 데이터의 중복을 최소화하면서 통합되는 **통합 데이터**
- 컴퓨터 디스크와 같은 매체에 저장된 데이터를 지칭하므로 **저장 데이터**

3. 데이터베이스의 특징 네 가지를 설명하시오.

- 데이터베이스는 사용자의 요청에 따라 실시간으로 서비스되므로 **실시간 접근성**을 띈다.
- 데이터베이스가 저장하는 내용은 특정 시점의 상태를 나타내며, 시간에 따라 데이터가 **계속적으로 변한다**.
- 베이스는 조직체의 공용정보이므로 여러 사용자가 동시에 사용하고, 데이터베이스를 접근하는 프로그램이 여러 개 있으므로 **동시에 공유**할 수 있는 데이터이다.
- 데이터베이스에 저장된 데이터는 데이터의 물리적인 위치가 아니라 **데이터 값(내용)에 따라 참조**된다.

4. 파일 시스템과 DBMS의 데이터 접근 방법의 차이를 설명하시오.

파일 시스템에서 데이터는 다양한 언어(c, java 등)로 작성되어 운영체제의 파일에 저장되고, 각각의 응용 프로그램마다 별도의 파일을 가지며 직접 이 저장된 파일에 접근한다. 그러나 DBMS에서의 데이터베이스는 표준화된 형식으로 저장되며 여기에 대한 접근은 모두 응용프로그램에서 SQL언어를 사용해 DBMS에 접근함으로써 이루어진다.

5. 파일 시스템과 DBMS의 장단점을 비교하여 설명하시오.

	파일시스템	DBMS
장점	<ul style="list-style-type: none"> ● 운영체제를 설치할 때 함께 설치되기 때문에 별도의 구입비용 없이 사용가능하다. ● 보통 소규모 프로그램에서 사용해 주기억장치를 적게 사용하고, 속도가 빠르다. 	<ul style="list-style-type: none"> ● DBMS에서는 데이터중복을 최소화한 통합 데이터베이스를 구축한다. <ul style="list-style-type: none"> ○ 프로그램-데이터가 독립적이어서 유지보수 비용 적게 든다. ○ 여러 사용자가 통합된 데이터베이스를 공유하며 동시에 데이터베이스에 접근할 수 있어 데이터 공유에 용이하다. ○ 모든 데이터를 하나의 데이터베이스에 통합하므로 데이터 중복과 불일치 감소한다. ○ 중복제거로 데이터 일관성 유지한다. ○ 데이터베이스를 통한 데이터의 중앙집중화로 보안조치 용이하다. ● 데이터에 대한 모든 접근은 데이터를 처리하는 DBMS가 담당한다. <ul style="list-style-type: none"> ○ 데이터를 응용프로그램으로부터 분리시켜 데이터 독립성이 향상된다. ○ 데이터 표준화 시행이 쉽다. ○ 데이터 무결성이 유지된다. ● 데이터베이스 접근 시 시스템이 다운되었을 경우 이전의 일관된 데이터베이스 상태 복구 가능하다.
단점	<ul style="list-style-type: none"> ● 데이터 정의가 프로그램에 내포되어 있다. <ul style="list-style-type: none"> ○ 프로그램-데이터 독립성이 없어 유지보수 비용 크다. ○ 각 응용 프로그램마다 파일을 갖고 있으며, 응용 프로그램마다 각각 다른 언어로 작성되어 있어 데이터 공유가 어려움 (공유 역시 파일 단위로 수행) ● 프로그램에서 데이터를 접근하고 조작하는 것 이외에 별도의 제어기능이 없다. <ul style="list-style-type: none"> ○ 중복된 데이터 변경을 제어하기 어려워 각 파일에 데이터 중복 저장 가능하다. 그래서 중복된 데이터가 변경될 때 두 파일 모두 수정하지 않으면 데이터 불일치가 발생한다. ○ 다수 사용자들을 위한 동시성 제어가 제공되지 않아 데이터 일관성이 훼손된다. ○ 사용자 권한에 따른 접근제어 어려워 보안조치가 미흡하다. ● 응용프로그램에서 파일 내 데이터 수정 시 시스템이 다운되었을 경우 회복기능이 없어 데이터 일관성 복구 어렵다. ● 데이터의 의미와 데이터간의 상호관계를 나타내는 데이터 모델링 개념이 부족하여 프로그래머의 생산성 낮다.(하나하나 기능 정의) 	<ul style="list-style-type: none"> ● DBMS의 구입비용 및 추가적인 하드웨어 구입비용 등 초기 투자비용이 크다. ● 데이터 관리의 오버헤드와 위험이 존재한다. ● 응용프로그램이 단순하고, 데이터 변경이 자주 일어나지 않고, 다수 사용자의 접근이 필요하지 않을 때는 부적합하다.

6. 데이터 모델을 설명하고 종류를 나열하시오.

데이터 모델은 데이터베이스 시스템에서 데이터를 저장하는 이론적 방법에 관한 것으로, 데이터베이스에 데이터가 어떻게 구조화되어 저장되는 지를 결정한다. 데이터 모델을 구분하는 가장 큰 기준은 데이터들 간에 관계를 표현하는 방법인데 대표적으로 3가지가 있다.

- **포인터를 사용하여 데이터 간의 관계를 표현하는 데이터 모델**

계층 데이터 모델, 네트워크 데이터 모델이 여기에 해당하며, 포인터를 사용하여 데이터가 저장된 포인터 값을 다른 테이블 데이터에 저장하는 방법으로 데이터를 연결한다. 데이터를 직접 찾아가 프로그램 속도는 빠르지만, 실제적으로 포인터 연산 구현이 어려워 응용 프로그램 개발 속도 더디다는 단점이 있다.

- **속성 값 사용하여 데이터 간의 관계를 표현하는 데이터 모델**

관계 데이터 모델이 여기에 해당된다. 테이블의 속성, 속성값을 다른 테이블 속성, 속성값에 직접 저장하는 방법으로 데이터를 찾을 때 이 속성, 속성값을 이용하여 어떤 강좌를 신청하였는지 알아낸다. 포인터를 사용하는 것보다 개념이 쉽고 프로그램 개발이 빠르지만 속도가 느리다는 단점이 있다.

- **객체 식별자를 사용하여 데이터 간의 관계를 표현하는 데이터 모델**

모든 객체는 고유의 식별자를 갖고 있기 때문에, 테이블을 객체 개념으로 보고 객체의 고유 식별자를 다른 테이블에 저장하는 방법으로 구현한다. **객체지향 모델**이 여기에 속한다. 이는 객체지향 언어에서 개념을 도입한 것이다..

7. 다음 데이터베이스 사용자들의 역할을 설명하시오.

일반 사용자, 응용 프로그래머, SQL 사용자, DBA

- **일반 사용자**: 데이터를 다루는 업무를 맡은 사람으로 특별한 지식 없이 응용 프로그래머가 일반 사용자를 위해 개발한 **프로그램**을 이용하여 데이터베이스 접근해서 데이터 사용한다.
- **응용 프로그래머**: 일반 사용자가 사용할 수 있는 프로그램을 개발하는 **프로그래머**로 프로그래밍 언어(C, Java, JSP 등)와 SQL 전용 프로그래밍 언어 등을 사용하여 일반 사용자를 위한 데이터 관리 응용 로직을 개발한다.
- **SQL 사용자**: SQL을 사용하여 업무를 처리하는 IT부서의 업무 담당자로 응용프로그램으로 구현되어 있지 않은 업무를 **SQL**을 사용하여 처리하고, 데이터를 모니터링하는 업무를 수행한다.
- **DBA**: 데이터베이스 운영 조직의 데이터베이스 시스템을 **총괄**하는 사람으로 데이터베이스의 설계, 구현, 유지보수의 전 과정과 데이터베이스 사용자 통제, 보안, 성능 모니터링, 데이터 전체 파악 및 관리 등 제반 업무를 담당한다.

8. 데이터 독립성을 정의하고, 데이터베이스 시스템에서 그 중요성을 설명하시오.

데이터 독립성은 하위 단계의 내용을 추상화하여 상위 단계에서 그 세부 사항을 숨김으로써 하위 단계 내의 변경에 대해 상위 단계와 상호간 간섭이 없도록 하는 것을 의미한다. 데이터 독립성은 물리적 독립성과 논리적 독립성이 있는데 이런 데이터 독립성을 수행함으로써 데이터베이스 시스템의 운영과 관련 응용 프로그램의 유지보수가 용이해진다.

9. ANSI의 3단계 데이터베이스 구조에 대해 설명하시오.

ANSI-SPARC구조는 데이터베이스를 보는 관점을 기준으로 3개의 단계를 분리한 것으로, 데이터베이스를 3-Layer 스키마(외부, 내부, 개념 스키마)로 나누어 데이터베이스 구조를 단순화시켜 데이터 물리적 데이터 독립성과 논리적 데이터 독립성을 유지시키는 것을 목적으로 한다.

첫 번째 층인 **외부 스키마**는 일반 사용자나 응용 프로그래머가 접근하는 계층으로 전체 데이터베이스 중에서 일반 사용자들에게 필요한 부분 스키마에 해당된다. 다음으로 두 번째 층인 **개념 스키마**는 전체 데이터베이스를 일컫는 말로 조직 별 하나만 존재하며 DBA가 관리한다. 개념 스키마는 데이터와의 관계, 제약사항, 무결성에 대한 내용을 포함하되, 저장장치에 독립적으로 기술한다. 마지막으로 **내부 스키마**는 DBMS의 관점에서 데이터베이스가 실제로 하드디스크에 저장되는 방법을 표현한 것으로 조직별 하나만 존재하여 인덱스, 데이터 레코드의 배치방법, 데이터 압축 등에 관한 사항을 포함한다.

DBMS는 세 가지 유형의 스키마 간의 사상을 책임진다. 총 2가지 사상이 있다. 먼저, 외부/개념 사상은 각 외부 스키마는 외부/개념 사상에 의해 개념스키마와 연관된다. 즉, 외부/개념 사상은 외부 스키마를 사용해서 입력된 사용자의 질의를 개념 단계의 스키마를 사용한 질의로 변환하는 과정을 수행한다. **논리적 데이터 독립성**을 지원한다. 두 번째로 개념/내부 사상이다. 개념/내부 사상은 위의 작업 결과를 다시 내부 단계의 스키마로 변환하여 디스크의 데이터베이스에 접근한다. 개념/내부 사상을 통해 DBMS가 개념 스키마 내의 한 논리적 레코드를 구성하는 물리적 저장장치 내의 실제 레코드를 찾을 수 있다. **물리적 데이터 독립성**을 지원한다.

10. 다음 데이터베이스 시스템의 구성도를 보고, 빈 곳에 알맞는 용어를 써 넣으시오.

(위에서부터 좌우로)

SQL 사용자, SQL 질의, DDL 컴파일러, 데이터베이스 관리 시스템, 파일 관리자, 인덱스

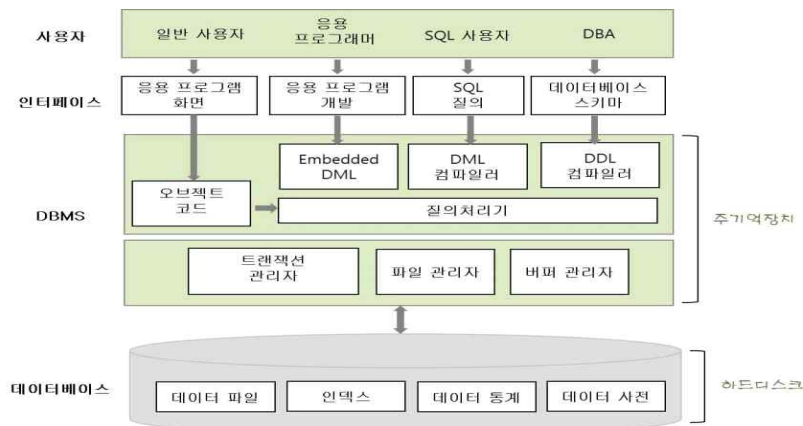


그림 1-22 데이터베이스 시스템의 구성