

1. 개체-관계 모델

1) 개체-관계(Entity Relationship: E-R) 정의

- 현실 세계를 개념적으로 표현하기 위해 가장 널리 사용되는 모델임
- 피터 첸(Peter Chen)이 제안한 개념적 데이터 모델
- 개체와 개체 간의 관계를 이용해 현실 세계를 개념적 구조로 표현
- 핵심 요소: 개체, 속성, 관계
- 개체는 자동차, 사람, 집, 사원 등과같이 실제로 존재하는 개체들임
- 관계는 개체와 다른 개체간의 연관관계를 나타내는 것임

2) 개체-관계 다이어그램(E-R diagram)

<E-R 표기법>



(1) 개체

- 현실 세계에서 조직을 운영하는 데 꼭 필요한 사람이나 사물과 같이 구별되는 모든 것
- 저장할 가치가 있는 중요 데이터를 가지고 있는 사람이나 사물, 개념, 사건 등
- 다른 개체와 구별되는 이름을 가지고 있고, 각 개체만의 고유한 특성이나 상태, 즉 속성을 하나 이상 가지고 있음

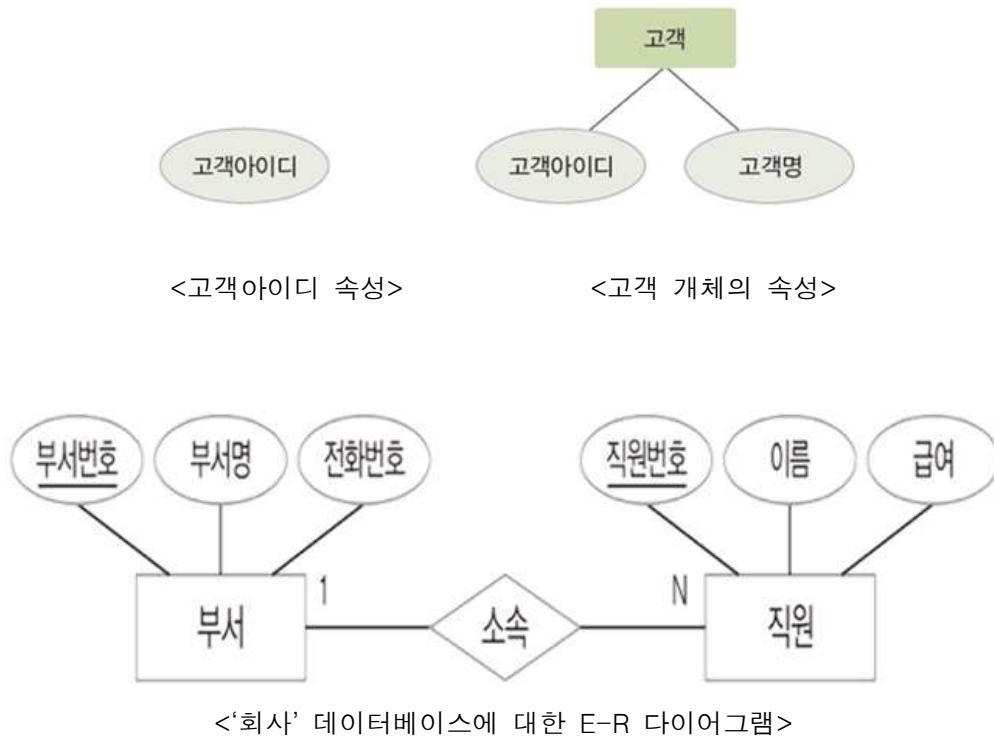
서점에 필요한 개체: 고객, 책

학교에 필요한 개체: 학과, 과목

- 파일 구조의 레코드(record)와 대응됨

(2) 속성

- 개체나 관계가 가지고 있는 고유의 특성
- 의미 있는 데이터의 가장 작은 논리적 단위
- 파일 구조의 필드(field)와 대응됨
- E-R 다이어그램에서 타원으로 표현하고 타원 안에 이름을 표기



2. 관계형 데이터베이스 시스템

1) 관계형 데이터베이스 시스템 개요

- 관계형 데이터 모델을 기반으로 하며 2차원 테이블 형태로 표현되고 데이터베이스는 테이블들의 집합체로 나타냄
- 데이터베이스 구조는 데이터 정의어(Data Definition Language : DDL)를 이용하여 정의됨
- 각 응용은 자기 자신의 뷰(view)를 전체 데이터베이스의 부분 집합으로 구성할 수 있으며 뷰(view)를 구성할 때는 테이블의 일부분 또는 여러 테이블을 조인(join)하여 구성함
- 실제 데이터에 접근할 때는 데이터 조작어(Data Manipulation Language: DML)를 이용하며, DML은 비절차적 질의에 기반을 두고 있음
- 사용자는 검색하려는 데이터 값이나 테이블 이름만 명시하면 됨
- 비절차적이면서 데이터베이스 구축이나 조작 및 검색에 편리한 언어가 SQL(Structured Query Language)임
- IBM에서 처음 개발하였으나 지금은 ISO 국제 표준으로 대부분 사용 데이터베이스 개발 업체들이 표준으로 수용함

2) 관계형 데이터베이스의 기본 구조

키(key)	테이블 내의 유일한 식별자 기능을 수행함
투플(tuple)	테이블의 행(레코드 단위)을 말함
속성(attribute)	열에 해당, 도메인은 속성들의 값의 범위를 의미함



<관계형 데이터베이스 모델의 기본 구조>

3) 관계형 데이터베이스 시스템 구성 요소

(1) 릴레이션(relation)

- 튜플 내의 각 속성값으로 원자 값(값이 더 이상 나누어 질 수 없는 단일값)을 가짐
- 각 속성은 유일한 이름을 가지며 속성의 순서는 큰 의미가 없음
- 테이블에서 모든 튜플은 서로 다른 값을 가지며, 튜플의 순서는 없고 릴레이션은 튜플들의 집합으로 정의됨

(2) 튜플(tuple)

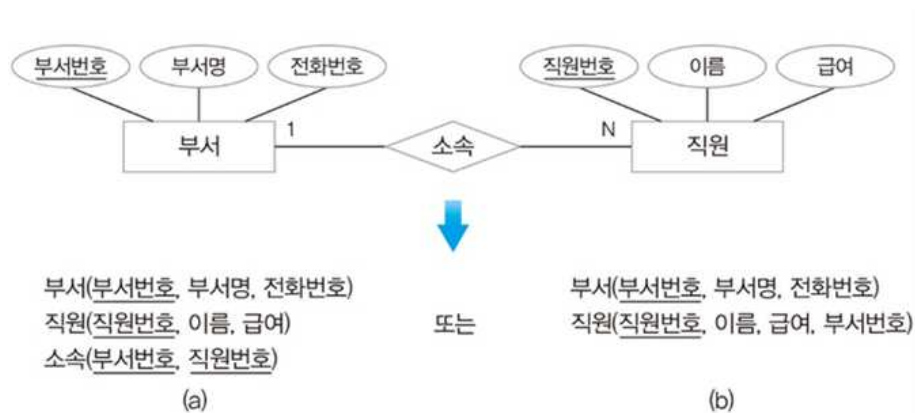
- 데이터의 행을 튜플이라 함
- 파일의 레코드와 유사한 개념으로 행이 튜플에 해당됨
- 서로 다른 속성으로 구성

(3) 속성(attribute)

- 데이터의 가장 작은 논리적 단위
- 속성은 각 개체의 특성을 나타내는 값을 표현한 것으로 단독으로 존재하기 어려움
- item, field, column 등으로도 불림

4) E - R 모델을 관계형 데이터베이스 스키마로 변환하기

1단계	‘부서’나 ‘직원’과 같은 개체 타입은 관계형 데이터 모델의 릴레이션으로 바뀜
2단계	‘소속’과 같은 관계타입은 별도의 릴레이션이 되거나, 대응 관계가 1: N 인 경우 1측의 기본 키를 N측의 외래 키로 사용해야 함
3단계	각 개체에 포함된 속성은 릴레이션의 속성으로 변환해야 함



<E-R 모델을 관계형 데이터베이스 스키마로 변환>

5) My-Sql 접속: 회사 데이터베이스 구축하기

- '부서', '직원' 테이블을 포함하는 '회사' 데이터베이스를 생성한다. 이때 SQL 언어로 데이터베이스를 구축하기 위해서는 편의상 데이터베이스 이름인 '회사'를 영어인 company 로 함

```
CREATE DATABASE company;
```

☞ company 데이터베이스를 만든다.

(1) 데이터베이스 선택하기

- 데이터베이스를 생성했으면 해당 데이터베이스 'company' 를 선택해야 함

```
use company;
```

☞ company 데이터베이스를 선택한다.

(2) 테이블 생성하기

- 데이터베이스 'company'를 구성하는 '부서(department)' 테이블과 '직원(employee)' 테이블을 생성해야 함

```
CREATE TABLE department (
    dept_number INT NOT NULL,
    dept_name CHAR(20),
    tel CHAR(20),
    PRIMARY KEY(dept_number));

CREATE TABLE employee (emp_number INT NOT NULL,
    emp_name CHAR(20),
    salary CHAR(20),
    dept_no INT,
    PRIMARY KEY(emp_number),
    FOREIGN KEY(dept_no) REFERENCES department(dept_number));

show tables;
```

☞ company 데이터베이스를 구성하는 테이블들을 보여준다.

(3) 테이블에 데이터 삽입하기

- Department, employee 각 테이블에 데이터를 삽입함

```
INSERT INTO department VALUES (1, '총무부', '02-1234-5678');  
INSERT INTO department VALUES (2, '자재부', '02-1234-5679');  
INSERT INTO department VALUES (3, '기획부', '02-1234-5680');  
  
INSERT INTO employee VALUES (10, '박종호', 2500000, 2);  
INSERT INTO employee VALUES (11, '성인옥', 2200000, 1);  
INSERT INTO employee VALUES (12, '한병엽', 1800000, 3);  
INSERT INTO employee VALUES (13, '박진수', 3000000, 1);
```

(4) 테이블의 데이터 검색하기

- 삽입된 데이터를 확인하기 위해 SELECT 문장을 수행함

```
SELECT * FROM department;  
SELECT * FROM employee;
```

3. 최근 데이터베이스 활용(빅데이터)

1) 빅데이터 개요

- 기존의 데이터베이스 관리 도구를 이용하여 데이터를 수집, 저장, 관리, 분석할 수 있는 범위를 넘어서는 방대한 양의 데이터 집합으로부터 가치 있는 정보를 빠르고 효율적으로 추출하고 결과를 분석하는 최신기술
- 빅데이터는 기존의 정형화된 데이터뿐만 아니라 다양한 형태의 사진, 동영상, 행동패턴, 위치정보, 메타 정보, 센서 데이터 등 멀티미디어 정보와 비정형 데이터들을 망라함
- 빅데이터 기술의 발전은 다변화된 현대 사회를 더욱 정확하게 예측하여 효율적으로 작동케 함
- 정치, 사회, 경제, 문화, 과학 기술 등 전 영역에 걸쳐서 사회와 인류에게 가치 있는 정보를 제공할 수 있는 가능성을 제시하며 그 중요성이 부각되고 있음

2) 빅데이터의 활용 과정

- 기업의 비즈니스 요구사항을 확인함
- 필요한 다량의 데이터를 검색하여 수집함
- 수집한 데이터를 적절한 형태로 가공함
- 처리된 데이터를 분석하고 시각화하여 이용함

3) 빅데이터의 활용 사례

(1) 한국석유공사의 국내 유가 예측 서비스

- 한국석유공사는 데이터 분석 전문기업과 협력하여 유가예측 시스템을 개발함
- 유가의 단기 미래가격을 예측하는 오피스넷 시스템을 구축하여 국내 1,300여 개의 주유소로부터 하루 6차례씩 수집된 휘발유 가격 정보를 활용함
- 사용자에게 현재의 차량 위치를 중심으로 최저가의 유가 서비스를 제공할 수 있고, 인근 주유소들 간에도 유가정보를 공유하여 건전한 경쟁을 유도하여 유가를 낮출 수 있으며, 이를 통하여 국제 유가에 민감한 국내 물가 안정에도 기여함

(2) 구글의 검색어 분석을 통한 독감 예보 서비스

- 구글은 홈페이지에서 독감이라는 단어를 포함하여 기침, 인플루엔자, 오한 등 독감과 관련된 빅데이터 검색어 분석을 통하여 '구글 독감 동향(Google Flu Trends)'이라는 독감 확산 조기 경보체계를 마련함
- 독감환자의 발병 시간, 분포 및 지역별 확산 정보를 예측하는데, 미국 질병통제 예방센터의 데이터와 비교결과 실제 독감 증세를 보인 환자 수와 매우 밀접한 상관관계가 있는 독감 예보 서비스임이 확인됨