

13

강

컴퓨터과학 개론

# 데이터베이스 (1)

컴퓨터과학과 이관용 교수



KOREA NATIONAL OPEN UNIVERSITY





# 학습목차

- 1 데이터베이스 개념
- 2 데이터베이스 시스템
- 3 데이터 모델링
- 4 개체-관계 모델

01

# 데이터베이스 개념

# 데이터 처리

## ■ 데이터

- 관찰이나 측정을 통해 현실 세계에서 단순히 수집된 사실이나 값
- 적절한 처리를 거쳐야만 정보로서 가치를 가짐

## ■ 정보처리 시스템

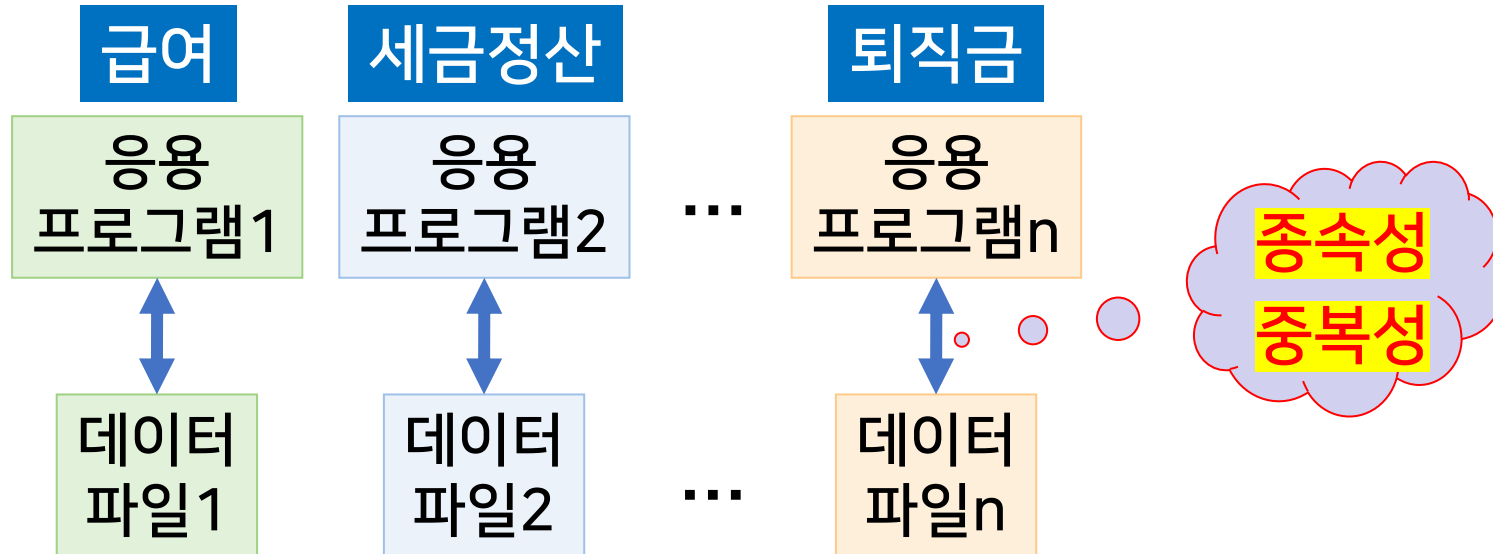
- 한 기관에서 데이터를 수집, 조직, 저장하고 정보를 생성, 분배하는 시스템
- 여러 응용 기술 중에서 실세계의 방대한 데이터를 효과적으로 저장하고 운영하기 위한 기술이 필요

→ “**데이터베이스**”

# 파일 처리 시스템

## ■ 파일 단위의 데이터 저장 및 처리 시스템

- 파일 → 정보 표현에서 1차원적인 저장 시스템
- 각 응용 프로그램이 특정 응용을 위해 필요한 파일을 독립적으로 소유/관리
  - 각 응용 프로그램은 파일의 논리적 구조와 물리적 구조가 대응되도록 구현



# 파일 처리 시스템의 문제점

## ■ 데이터 종속성

- 응용 프로그램과 데이터 사이에 1:1 상호 의존 관계가 존재
  - 파일 구성 요소, 접근 방식 등이 변경되면 해당 응용 프로그램도 함께 변경

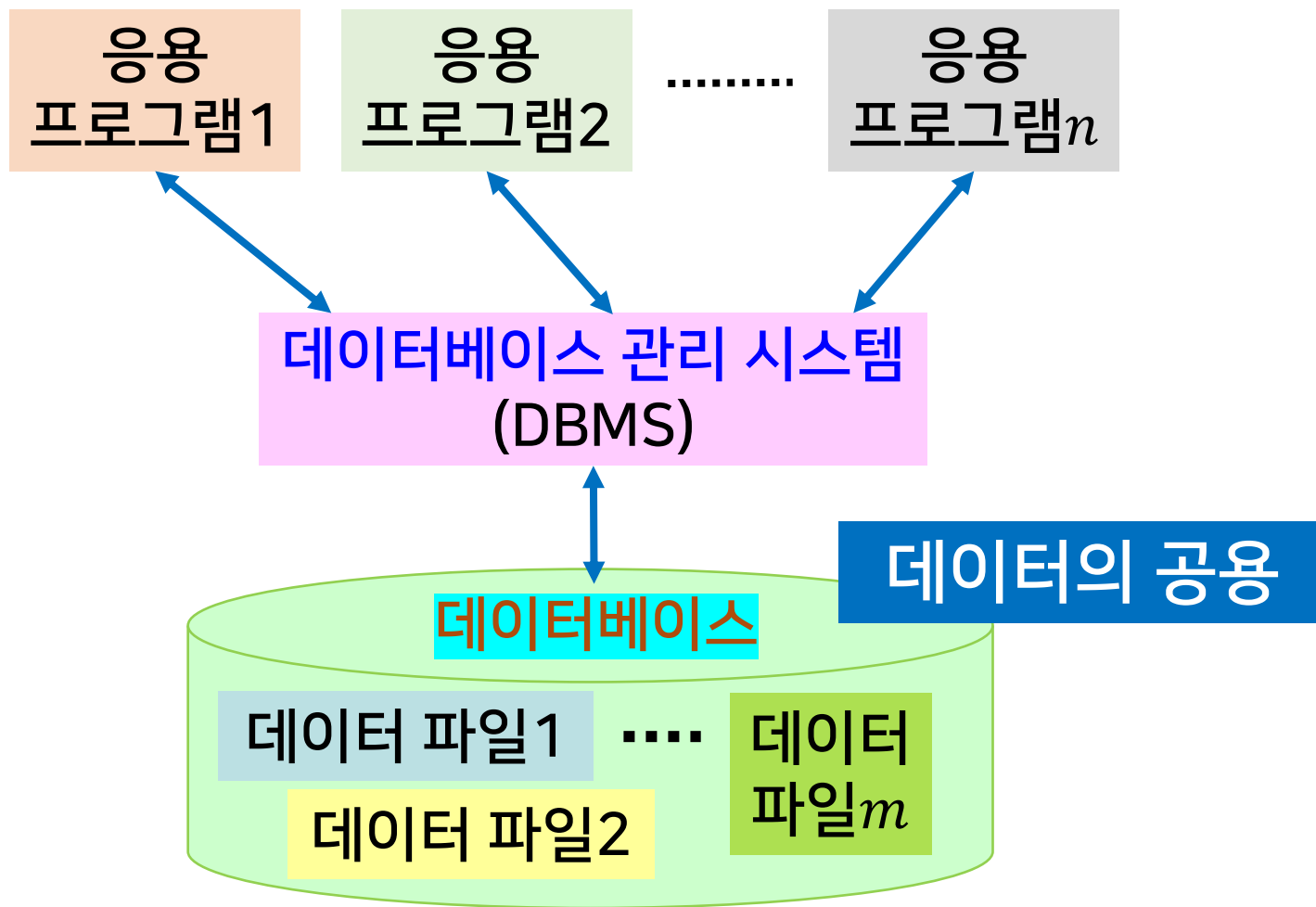
## ■ 데이터 중복성

- 한 시스템에 동일한 데이터가 여러 개 존재
  - 문제점 → 일관성, 보안성, 경제성, 무결성의 확보 및 유지 곤란

데이터 공유 불가

“데이터베이스” 기술 등장

# 데이터베이스 개념



# 데이터베이스 정의

## ■ 한 조직의 여러 응용 시스템이 공유해서 사용하기 위한 통합, 저장, 운영 데이터의 집합

- 공유 shared → 여러 응용 프로그램이 공동으로 소유하고 사용하는 데이터
- 통합 integrated → 중복된 데이터를 배제하여 각 데이터의 일관성 유지
  - 중복의 완전한 배제가 아닌 시스템의 성능 향상을 위해서는  
“제한적으로 통제된 범위 내에서의 최소한의 중복”은 허용
- 저장 stored → 컴퓨터가 접근 가능한 저장매체에 저장된 데이터
- 운영 operational → 어떤 조직의 목적과 유용성 측면에서  
반드시 유지해야 할 데이터
  - 단순한 입출력 데이터 및 처리 과정에서 일시적으로 필요한 데이터는 제외



# 데이터베이스 특징

## ■ 실시간 접근성

- DB에 수시로 접근하는 사용자의 요구를 즉시 처리하여 응답을 제공

## ■ 계속적인 변화

- 삽입/삭제/갱신 등의 연산을 통해 새로운 데이터로 내용을 지속적으로 변화시킴 → 현실 세계의 상태를 정확히 반영한 데이터를 유지

## ■ 동시 공유

- 서로 다른 목적을 가진 여러 사용자가 동시에 원하는 데이터에 접근

## ■ 내용에 의한 참조

- 데이터가 저장된 위치나 주소가 아닌 내용/값에 따라 데이터를 참조

# 파일 처리 방식에 대해 향상된 특징

## ■ 데이터베이스의 자기 기술성

- DB 시스템은 DB 자체뿐만 아니라 **DB에 대한 정의/설명**까지 포함
  - DB에 속하는 각 파일의 구조, 각 항목의 타입과 저장 형식, 데이터의 제약 조건 등이 시스템 카탈로그에 저장

## ■ 프로그램-데이터 독립성

- 데이터 파일 구조에 대한 정보가 응용 프로그램으로부터 분리되어 관리

## ■ 데이터 추상화

- 사용자에게 데이터 저장 방식에 대한 상세 정보보다 데이터에 대한 개념적 표현을 제공 → 보다 용이한 데이터 접근 가능

# 파일 처리 방식에 대해 향상된 특징

## ■ 다중 뷰 제공

- DB에 대한 여러 사용자의 서로 다른 관점의 데이터 요구에 따라 각자의 필요한 부분만을 선별적으로 추출해서 볼 수 있는 기능 제공

## ■ 데이터 공유

- 여러 사용자가 동시에 DB에 접근할 수 있는 기능 제공

## ■ 다수 사용자의 트랜잭션 처리

- 동시성 제어(병행 제어) 기능을 통해 다수 사용자가 동일 데이터를 동시에 변경하는 경우에도 데이터의 일관성 보장

# 데이터베이스의 장단점

■ 데이터의 일관성

현실 세계의 어느 한 사실을 나타내는 두 개 이상의 데이터 간의 일치 여부

■ 데이터의 무결성

DB에 들어 있는 데이터 값과 그것이 표현하는 현실 세계의 실제 값이 일치하는 정확성

■ 데이터의 보안

■ 백업과 회복

■ 표준화

■ 응용 프로그램 개발 시간 단축

■ 융통성

■ 최신 정보의 가용성

■ 규모의 경제성

■ 운영비의 증대

■ 복잡한 자료 처리

■ 백업과 회복의 오버헤드

■ 시스템의 취약성

02

# 데이터베이스 시스템

# 데이터베이스 시스템

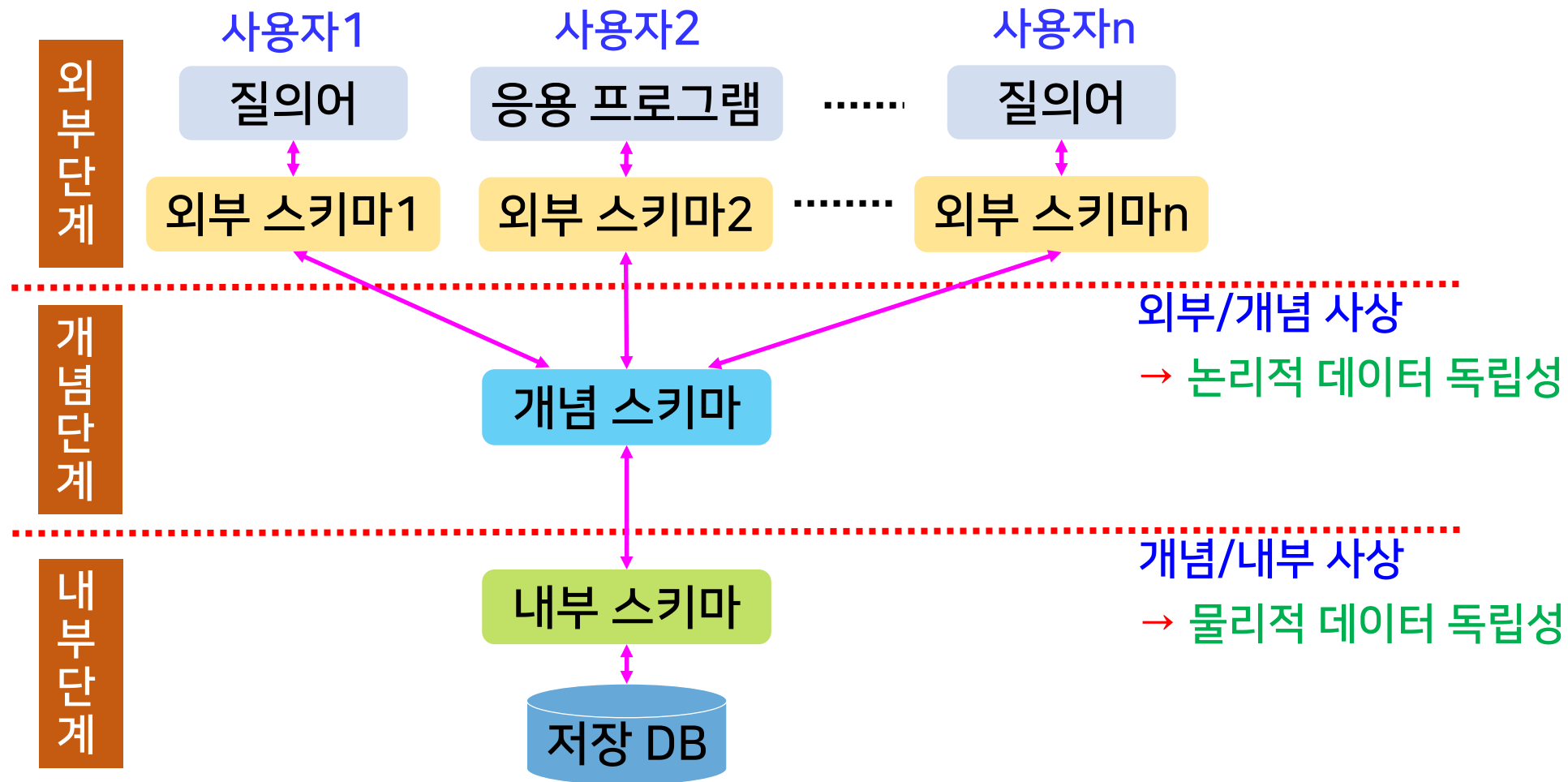
- 데이터를 데이터베이스에 저장하고 관리해서 필요한 정보를 생성하는 컴퓨터 중심의 시스템

- 구성요소

- 데이터베이스
- 데이터베이스 관리 시스템
- 데이터 언어
- 데이터베이스 사용자
- 데이터베이스 관리자
- 데이터베이스 기계

# 데이터베이스 구조

## 3단계(외부-개념-내부) 구조 → “ANSI/SPARC 구조”



# 데이터베이스 구조

## ■ 외부 단계, 뷰 단계

- **개인적 입장** → 각 사용자가 바라보는 개인적인 수준의 DB에 관한 것

## ■ 개념 단계, 논리적 단계

- **범 기관적 입장** → DB의 전체 구조를 추상화하는 단계
  - 물리적 저장구조의 세부 사항을 은폐하고,  
어떤 데이터가 저장되었는지와 데이터 간에 존재하는 관계를 기술

## ■ 내부 단계, 물리적 단계

- **저장 장치 입장** → 데이터가 실제로 어떻게 저장되는가를 기술



# 데이터베이스 구조

## ■ 스키마 schema

- DB 구조에 대한 정의와 제약 조건의 명세를 기술한 것
  - 3가지 관점 → 외부 스키마, 개념 스키마, 내부 스키마

## ■ 외부 스키마

- 개별 사용자/응용 프로그래머가 관심을 두는 DB 일부분만 기술
  - “서브 스키마” → 전체 DB의 한 논리적 부분만을 표현함
- 개별 사용자마다 이에 대응하는 외부 스키마가 존재
  - 하나의 DB에는 여러 개의 외부 스키마가 존재

# 데이터베이스 구조

## ■ 개념 스키마

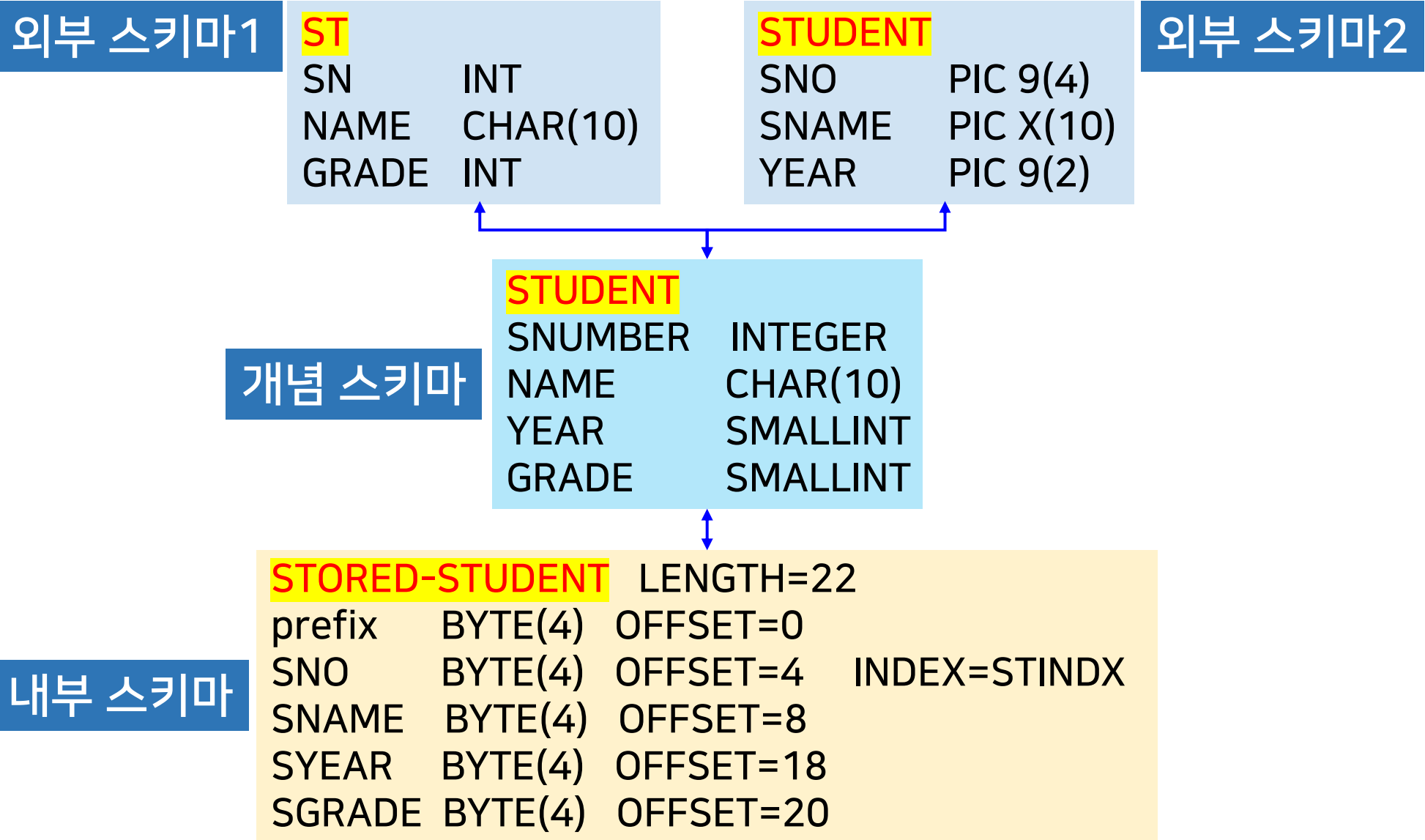
- 모든 사용자나 응용 프로그래머가 필요로 하는 전체적이고 통합된 데이터베이스의 구조를 기술
  - 오직 하나만 존재
- 모든 데이터 개체들에 대한 정의, DB 접근 권한, 보안 정책, 무결성 규칙 등에 대한 명세를 포함
- 개념 스키마로부터 모든 외부 스키마가 생성되고 지원됨

# 데이터베이스 구조

## ■ 내부 스키마

- 물리적인 데이터의 구조를 정의
- 저장장치의 입장에서 데이터가 실제로 저장되는 방법을 기술
  - 저장 레코드 형식, 인덱스 유무, 저장 필드의 표현 방법, 저장 레코드의 물리적 순서 등을 기술
  - 물리적 레코드(페이지, 블록)와 저장장치의 특성(실린더, 트랙)은 고려하지 않음
- “저장 스키마” → 개념 스키마에 대한 저장 구조를 정의

# 데이터베이스 구조



# 데이터베이스 구조\_각 단계의 사상

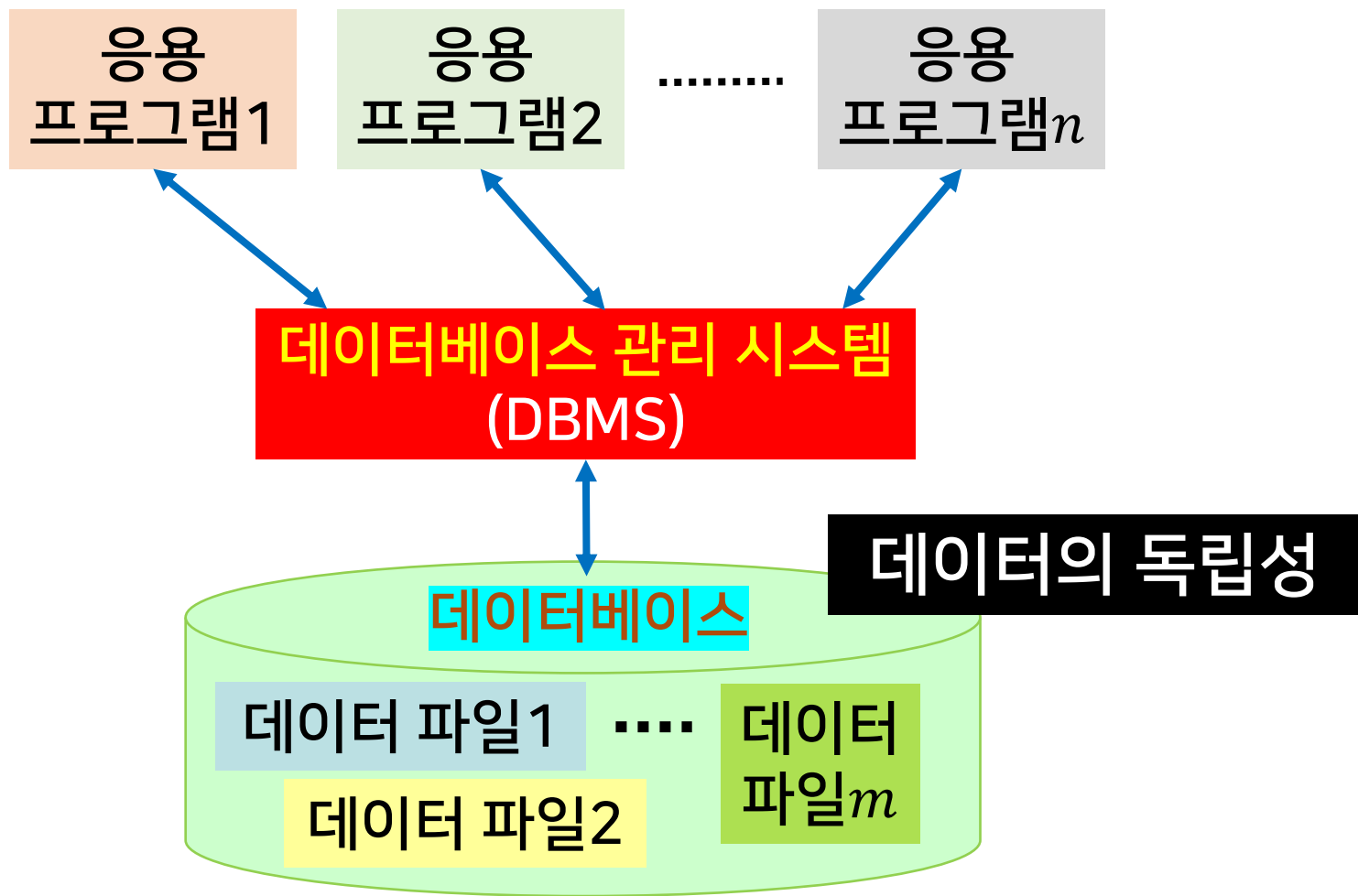
## ■ 외부/개념 사상 → 논리적 데이터 독립성

- 특정 외부 스키마와 개념 스키마 사이의 대응 관계를 정의
  - 응용 프로그램에 영향을 주지 않고 DB의 논리적 구조의 변경이 가능
  - 하나의 논리적 구조로 많은 응용 프로그램이 요구하는 다양한 형태의 논리적 구조를 제공

## ■ 개념/내부 사상 → 물리적 데이터 독립성

- 개념 스키마와 내부 스키마 사이의 대응 관계를 정의
  - 논리적 구조에 영향을 주지 않고 실제 데이터에 대한 저장 양식의 변경이 가능
  - 하나의 논리적 구조로 여러 가지 상이한 물리적 구조의 지원이 가능

# 데이터베이스 관리 시스템



# 데이터베이스 관리 시스템

## ■ DBMS, DataBase Management System

- 응용 프로그램이 데이터베이스를 공유할 수 있도록 관리해주는 소프트웨어 시스템
  - 사용자와 데이터베이스 사이에 위치하여, DB의 구성, 접근 방법, 관리 유지 등에 관한 모든 책임과 권한을 갖고 모든 기능을 통합적으로 수행하여 사용자의 요구에 맞는 정보를 생성해 주는 소프트웨어
- 목적
  - 응용 프로그램이 데이터에 종속되지 않는 데이터의 독립성 제공

# 데이터베이스 관리 시스템

## ■ 개념적인 수준의 작업 과정 (관계형 데이터베이스 관리 시스템에서)

- 사용자는 SQL과 같은 언어를 이용하여 데이터에 대한 접근 요구



- DBMS가 요구를 받아 분석



- 외부스키마, 대응하는 외부와 개념 스키마 접속, 개념 스키마, 개념과 내부 스키마 접속, 그리고 기억장소 구조 정의 순으로 차례대로 검토



- 저장된 데이터베이스에 필수적인 연산 수행



# DBMS의 필수 기능

## ■ 정의 definition

- 물리적으로 구현된 하나의 DB 구조로부터 여러 사용자들의 다양한 요구에 부응할 수 있도록 가장 적합한 DB 구조를 정의하는 기능

## ■ 조작 manipulation

- 사용자와 DB 사이의 상호작용을 위한 수단 → 사용자가 연산 지원 도구를 통해 DB에 체계적으로 접근하고 조작할 수 있는 기능

## ■ 제어 control

- 공용 목적으로 관리되는 DB의 내용을 정확하고 안전하게 유지시키는 기능 → 데이터 무결성 유지, 보안 유지 및 권한 검사, 동시성 처리 등

# 데이터 언어

## ■ DBMS와의 통신 수단

- 기능/목적에 따라 → 데이터 정의어, 데이터 제어어, 데이터 조작어
  - 별도 언어로 존재하기 보다는 하나의 통합된 언어 속에서 기능적으로만 구분

## ■ 데이터 정의어 **DDL**, Data Definition Language

- DB의 구조, 데이터 형식, 처리 방식 등을 정의하는 언어
  - 개념 스키마를 정의하기 위해서 주로 사용
  - 데이터베이스 설계자 또는 관리자가 주로 사용
  - “뷰 정의어” → 외부 스키마를 정의하기 위한 언어
  - “기억장소 정의어” → 내부 스키마를 정의하기 위한 언어

# 데이터 언어

## ■ 데이터 제어어 **DCL**, Data Control Language

- 여러 사용자가 데이터를 공유해서 정확하고 안전하게 사용하도록 데이터 제어를 정의하고 기술하는 언어
  - 데이터베이스 관리자가 사용해서 데이터 보안, 무결성, 데이터 회복, 동시성 제어 등과 관련된 명령어들을 통해서 데이터를 관리

# 데이터 언어

## ■ 데이터 조작어 **DML**, Data Manipulation Language

- DB에 대한 검색, 수정, 삽입, 삭제 등의 조작을 위한 언어
- 절차적 데이터 조작어
  - 필요한 데이터를 **어떻게** *how* 구하는 지를 구체적으로 명시해야 하는 조작어
  - 응용 프로그램 속에 삽입되어 사용 → 한 번에 하나의 레코드를 검색/처리
  - “one-record-at-a-time 데이터 조작어”, “저수준 데이터 조작어”
- 비절차적 데이터 조작어
  - 어떻게 구하는 지는 명시하지 않고, **어떤** *what* 데이터가 필요한지만을 명시
  - “선언적 언어”, “고수준 데이터 조작어”
  - 질의어(→ SQL), 데이터 부속어

# 데이터베이스 사용자

## ■ 데이터베이스에 접근하는 사람의 총칭

- 일반 사용자

- SQL과 같은 질의어를 통해 데이터 조작(검색, 삽입, 수정, 삭제)을 목적으로 단순히 DB에 접근하는 사람

- 응용 프로그래머

- 일반 호스트 언어와 데이터 부속어를 통해 DB에 접근하는 사람

- 데이터베이스 관리자

- 데이터베이스 시스템의 별도의 한 구성요소로 취급

# 데이터베이스 관리자

## ■ DBA, DataBase Administrator

- DDL과 DCL를 통해 DB를 정의하고 제어할 목적으로 접근하여 관리하는 사람
  - 데이터를 여러 사람이 공유할 수 있도록 관리하고 제어하는 사람
  - 조직의 모든 전산 업무, DBMS, 관련 HW/SW 등에 대한 상당한 지식이 필요
- DB에 대한 접근 권한 설정, 제작과 갱신, 보전과 관리, 운영 및 통제, 성능 측정 및 변경 요구에 대한 응답 등의 의무를 가짐

# 데이터베이스 기계

## ■ “데이터베이스 컴퓨터”

- DB 관리 기능을 효율적으로 수행할 수 있도록  
특화되어 설계된 하드웨어/소프트웨어
  - 후위 처리기, 지능형 저장장치, 내용에 의한 참조 메모리,  
병렬 처리, 데이터베이스 연산기 등을 포함

03

# 데이터 모델링



# 데이터 모델링

## ■ 현실 세계의 데이터를 데이터 모델 상의 데이터베이스 구조로 변환하는 과정

- 실세계의 일부분을 DB 시스템이 지원하는 데이터 모델의 형태로 데이터베이스를 표현하는 과정
- 데이터 모델?
  - 데이터 타입, 데이터의 연산, 데이터의 의미 및 일관성 제약 조건 등의 DB 구조를 추상화하여 체계적으로 명시하기 위한 개념의 집합
    - 데이터 모델링의 결과를 표현하는 도구
  - 지원하는 개념의 타입에 따른 구분 → 개념적 모델, 논리적 모델, 물리적 모델

# 데이터 모델

## ■ 개념적 모델

- 현실 세계에서 주요 데이터를 추출하여  
개념 세계의 데이터로 변환하는 과정에서 사용되는 모델
- 현실 세계를 개념 세계로 추상화했을 때  
어떤 요소로 이루어져 있는지를 표현
  - 사용자가 데이터를 인식하는 방법과 밀접한 개념 제공
- 대표적 모델 → 개체-관계 모델(“E-R 모델”)

# 데이터 모델

## ■ 논리적 모델, “구현 모델”

- 개념 세계의 데이터를 데이터베이스에 저장할 구조로 표현하는 과정에서 사용되는 모델
  - 개념적 모델과 물리적 모델의 중간에 위치한 모델
  - 데이터 구성에 대한 세부적인 사항은 숨기고 사용자가 이해할 수 있는 정도의 데이터 저장에 대한 개념 제공
- 종류
  - 계층형 모델, 네트워크형 모델, 관계형 모델, 객체지향형 모델, 객체관계형 모델

# 데이터 모델

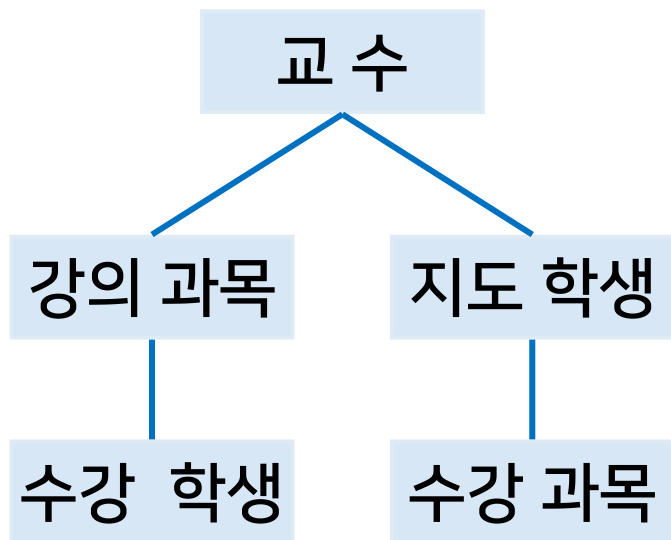
## ■ 물리적 모델

- 사용자 수준보다 더 하위 수준의 저장장치 입장에서 데이터가 어떻게 저장되어 있는지에 대한 세부적인 사항을 기술하는 개념 제공
  - 레코드 형식, 레코드 순서, 접근 경로 등의 정보 표현/제공

# 데이터 모델\_논리적 모델

## ■ 계층형 모델

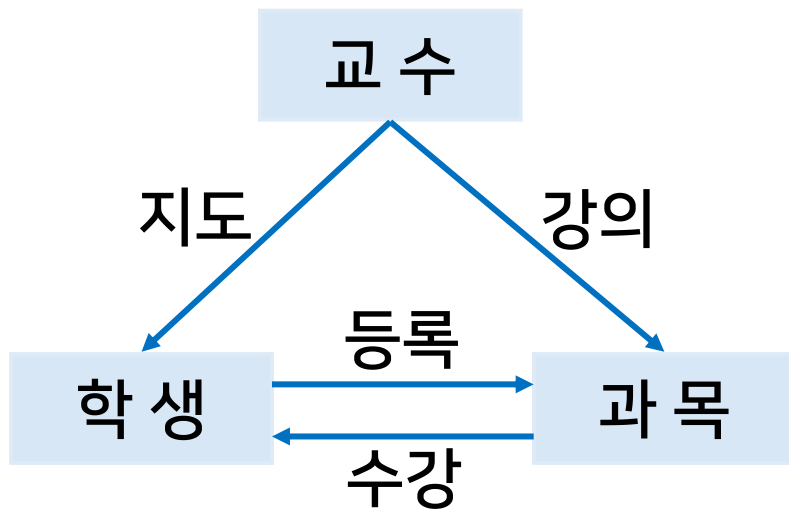
- DB 구조를 트리 형태로 표현
  - 두 개체 사이의 관계는 오직 1:n 관계 → 부모-자식 관계 (“부모 개체”, “자식 개체”)



# 데이터 모델 - 논리적 모델

## 네트워크형 모델, “망형 모델”

- 그래프 형태로 표현
  - 링크로 표현된 데이터 간의 관계 → 오너-멤버 관계



# 데이터 모델\_논리적 모델

## ■ 관계형 모델 → 14강

## ■ 객체지향형 모델

- 데이터와 절차를 일체화된 단위로 다루는 객체지향의 개념을 잘 모아서 정의해 놓은 집합
  - 실세계에 존재하는 개념적 객체를 중심으로 모델링하는 방식

## ■ 객체관계형 모델

- 관계형 모델과 객체지향형 모델의 장점을 결합한 진보된 형태의 모델
  - 관계형 시스템에 새로운 객체 저장 능력을 추가한 형태

04

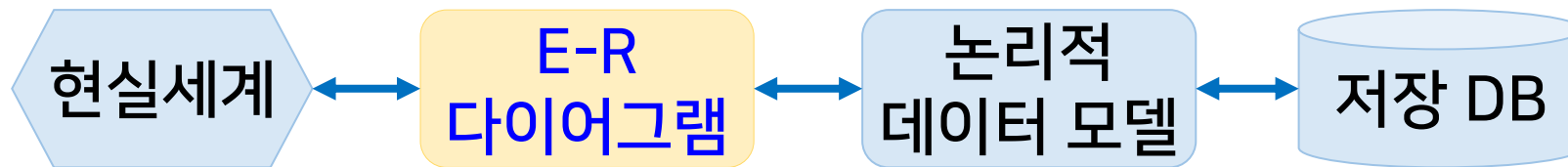
## 개체-관계 모델



# 개체-관계 모델

## ■ E-R Entity-Relationship 모델

- 개체 타입과 이들 간의 관계 타입을 이용해서  
실세계를 사람이 이해할 수 있도록 개념적으로 표현하는 방법
- 개체-관계 다이어그램(ERD) → E-R 모델을 그래프 방식으로 표현한 것

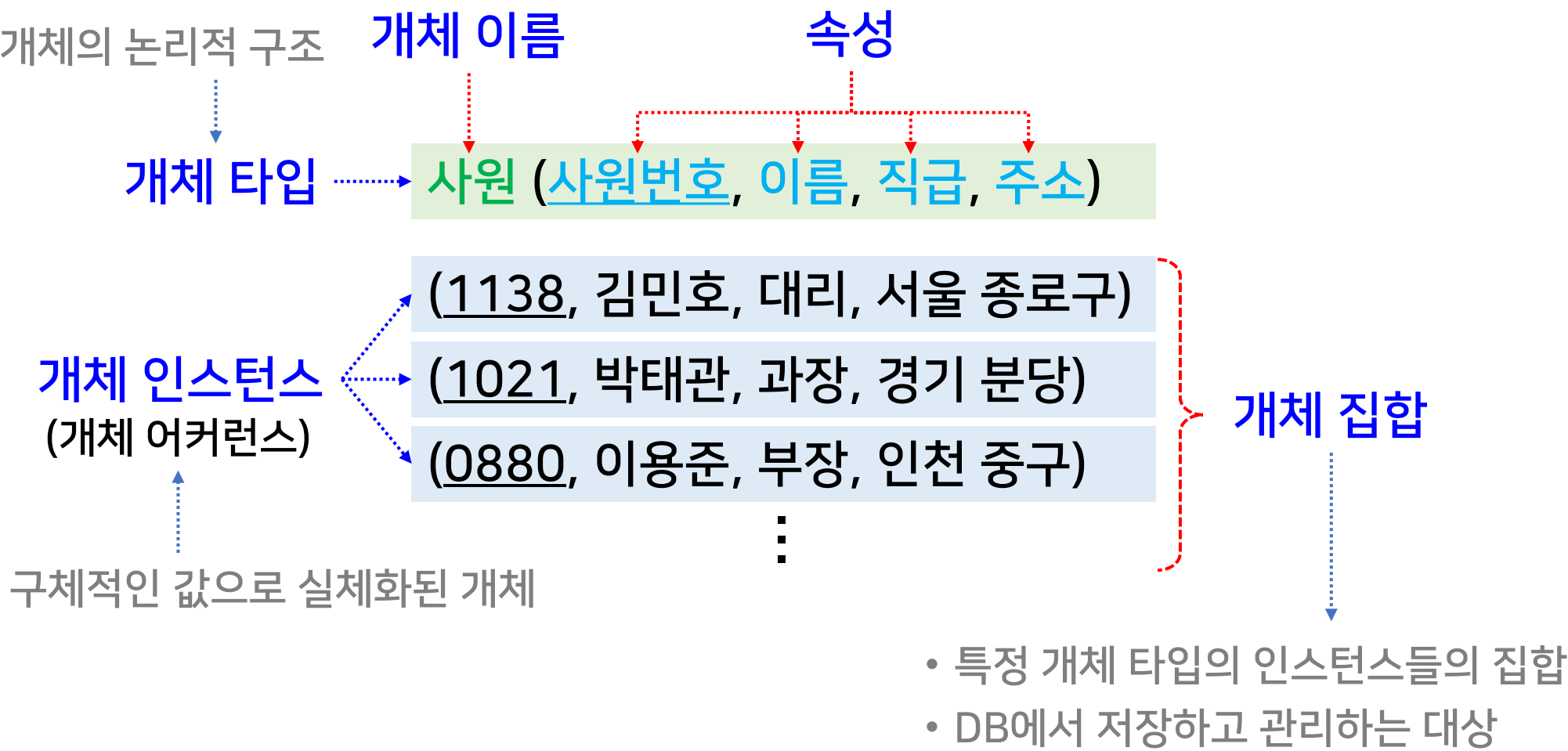


# 개체와 속성

## ■ 개체

- 데이터로 표현하려는 물리적/개념적으로 존재하는 실세계의 유무형의 모든 것
  - 컴퓨터 파일구조에서는 하나의 레코드에 대응
- 하나 이상의 속성으로 구성
  - 속성 → 데이터의 가장 작은 논리적 단위
    - 각 개체의 특성이나 상태를 표현
    - 컴퓨터 파일구조에서 필드에 대응
  - 종류 → 단일값/다중값 속성, 단순/복합 속성, 유도 속성 등

# 개체와 속성



# 관계

## ■ 개체 집합 사이의 대응성(사상)을 의미

- 개체 집합 간의 관계 정보를 표현하는 수단

→ 집합 간에 어떤 관계를 미리 정의해 놓고  
주어진 값을 정의된 관계에 따라 해석하면 유용한 의미 표현이 가능

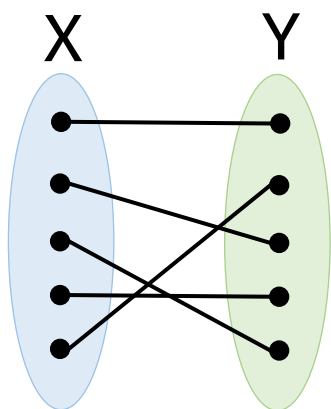


→ " '이관용' 교수가 '홍길동' 학생을 지도한다. "

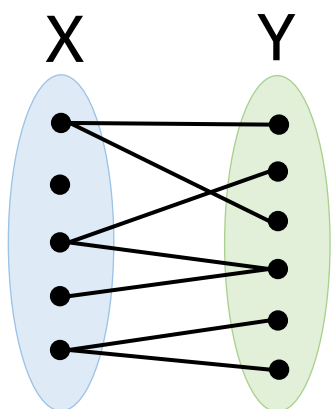
# 관계

## 개체 간의 관계

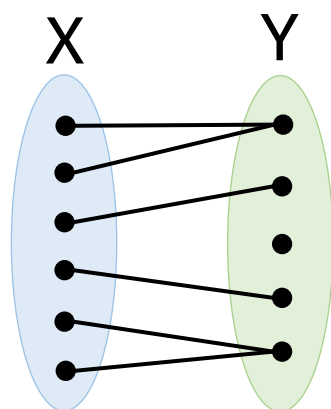
- 사상 원소의 수 mapping cardinality에 따른 분류



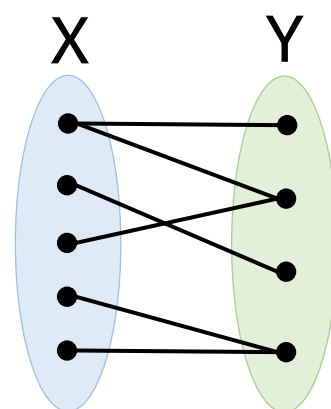
일대일 (1:1)



일대다 (1:n)



다대일 (n:1)



다대다 (m:n)

# 관계

## ■ 두 개체 집합 X와 Y의 관계에서

### ■ “개체 Y가 개체 X에 종속되어 있다”

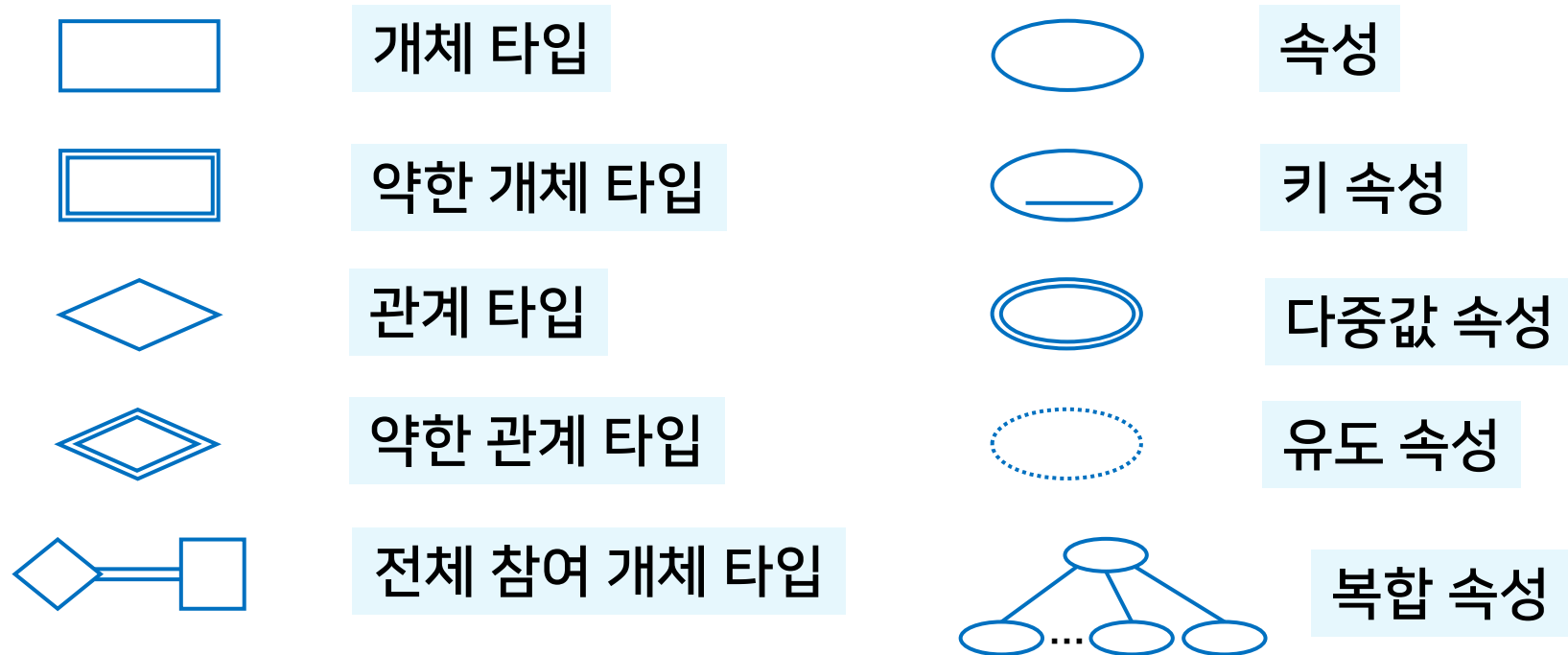
- 개체 X가 존재하는 경우에만 개체 Y가 존재할 수 있고, 개체 X가 삭제되면 개체 Y도 함께 삭제되는 경우
- 개체 X → “오너 개체”, 개체 Y → “약한 개체”

### ■ “전체 참여한다” (“필수적으로 참여한다”)

- 개체 X의 모든 인스턴스가 관계에 반드시 참여하는 경우  
↔ “부분 참여한다”, “선택적으로 참여한다”

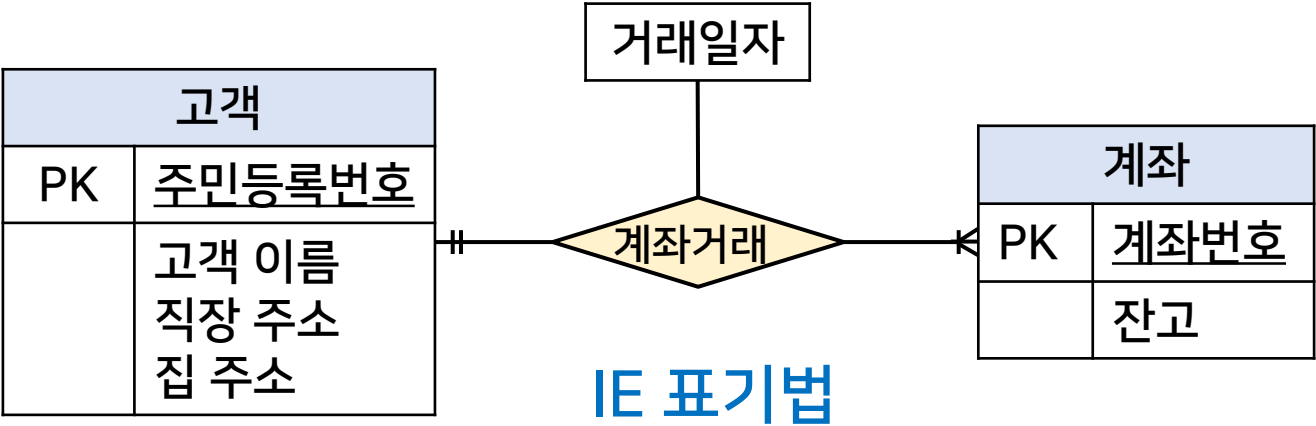
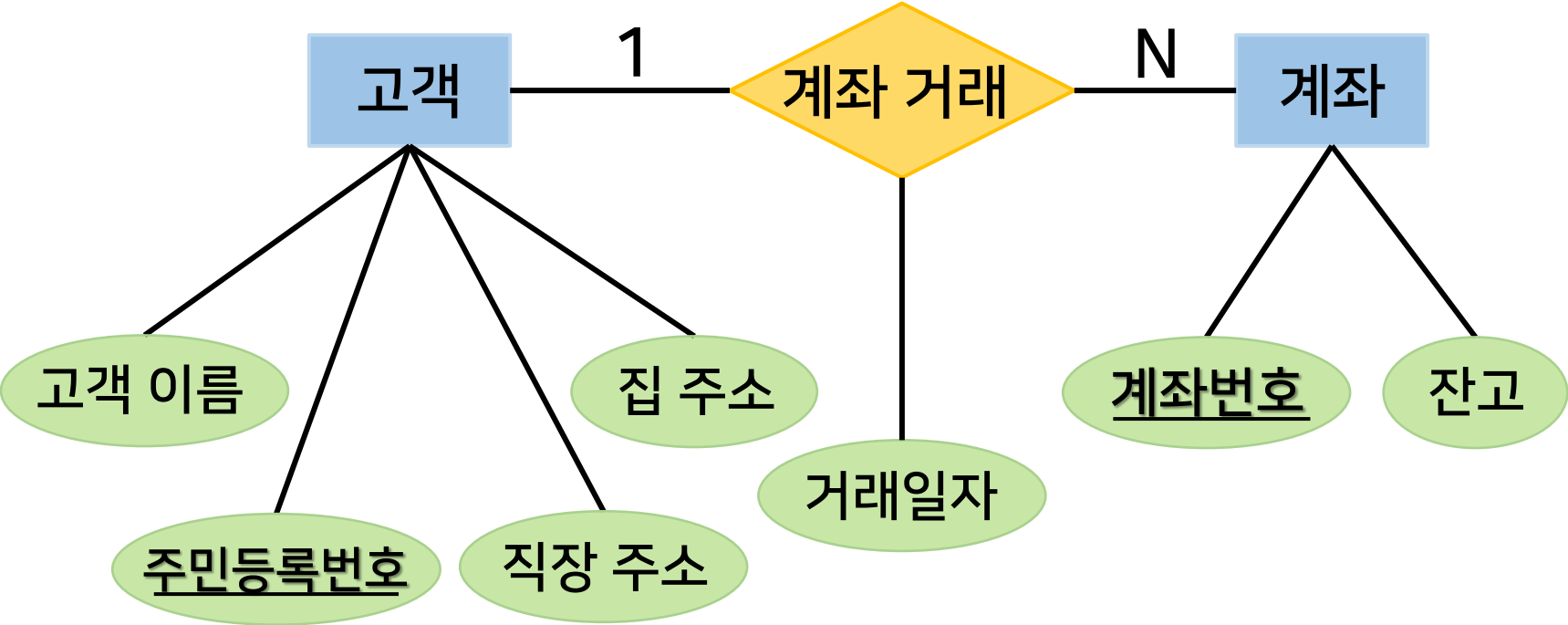
# E-R 다이어그램

## ERD에서 사용하는 기호와 의미



R에서 E1:E2의 사상 원소 비율이 1:N

# E-R 다이어그램







## 1. 데이터베이스 개념

- 정의 → 공유, 통합, 저장, 운영 데이터의 집합
- 특징 → 실시간접근성, 계속적인변화, 동시공유, 내용에의한참조

## 2. 데이터베이스 시스템

- 구성요소 → DB, DBMS, 데이터 언어, 데이터베이스 사용자, DBA, 데이터베이스 기계
- 3단계 데이터베이스 구조 → 외부단계, 개념단계, 내부단계
- 스키마 → 외부스키마, 개념스키마, 내부스키마
- 데이터언어 → DDL, DCL, DML (절차적, 비절차적)

## 3. 데이터 모델링

- 개념적모델, 논리적모델(구현 모델), 물리적모델
- 논리적모델의종류 → 계층형, 네트워크형, 관계형, 객체지향형, 객체 관계형

## 4. 개체-관계 모델

- 개체와개체간의관계를개념적으로표현, E-R 다이어그램

14

강

다음시간 안내

## 데이터베이스 (2)

