

7장: 데이터베이스

7장: 데이터베이스

- 데이터베이스 개요
 - 데이터와 정보
 - 데이터베이스
- 데이터베이스 시스템
 - 데이터베이스 시스템의 개념
 - 데이터베이스 시스템 구조

데이터베이스 - 이장의 목적

- 정보화 사회에서 **정보**는 중요한 자원이자 의사결정하는 중요한 수단
 - 중요한 정보를 잘 보관하고 있는가?
 - 정보를 어떻게 보관하고 관리하는가?
 - **정보**는 무엇인가?
 - **데이터**는 무엇인가?

차이

- 데이터베이스 시스템을 이용한 데이터베이스의 설계, 구축, 관리
 - 데이터를 효율적으로 관리하는 방법은 무엇인가?
 - 가치 있는 정보는 어떻게 확보할 수 있는가?

7.1 데이터베이스 개요

데이터와 정보

- 정보: 추상적, 관념적
 - 예: '덥다', '춥다', '비가 온다' ...
- 데이터: 구체적이고 객관적인 의미자료, 컴퓨터 안에 저장 가능한 형태
 - 예: '섭씨 영하 10도', '비가 10mm 온다' ...

“ 컴퓨터에 저장 가능한 형태
+ (추상적, 관념적) 의 의미자료 ”

데이터와 정보

□ 데이터는

- 관찰이나 측정을 통해 수집된 사실이나 값
- 가공을 가하지 않은 값을 의미 *→ 데이터 라임으로 문장..*
- 객관적인 실재를 반영하며 기호로 표현
- 예) 고객의 주소, 성별, 이름, 나이, 휴대폰 기종, 휴대폰 활용 회수 등은 구체화하고 숫자나 기호로 표현 가능한 값

□ 데이터 종류

- ▶ 원자료 또는 1차 자료 : 어떤 카드회사에 가입한 회원이 작성한 카드나 작성한 설문지 결과를 입력한 자료
- ▶ 2차 자료 : 다른 목적으로 수집한 자료를 의미하며 집계된 자료라는 의미의 매크로 데이터, 예를 들어, 정부에서 생산하고 배포하는 각종 자료를 여러 목적에 따라 간접 혹은 다른 목적으로 활용하면 2차 자료가 됨

데이터

데이터와 정보

□ 정보는

- 의사결정에 사용, 특정 목적이나 문제 해결에 적용할 지식
- 데이터를 의사 결정에 활용할 수 있도록 처리하고 체계적으로 조직한 의미 있는 결과물
목적에 맞게 활용 할 수 있도록
- 사용자가 목적으로 하는 값으로, 의사결정을 위한 최종 값
- 아침 출근 전

① 데이터 : 예상 기온, 예상 바람, 예상 강수량 →

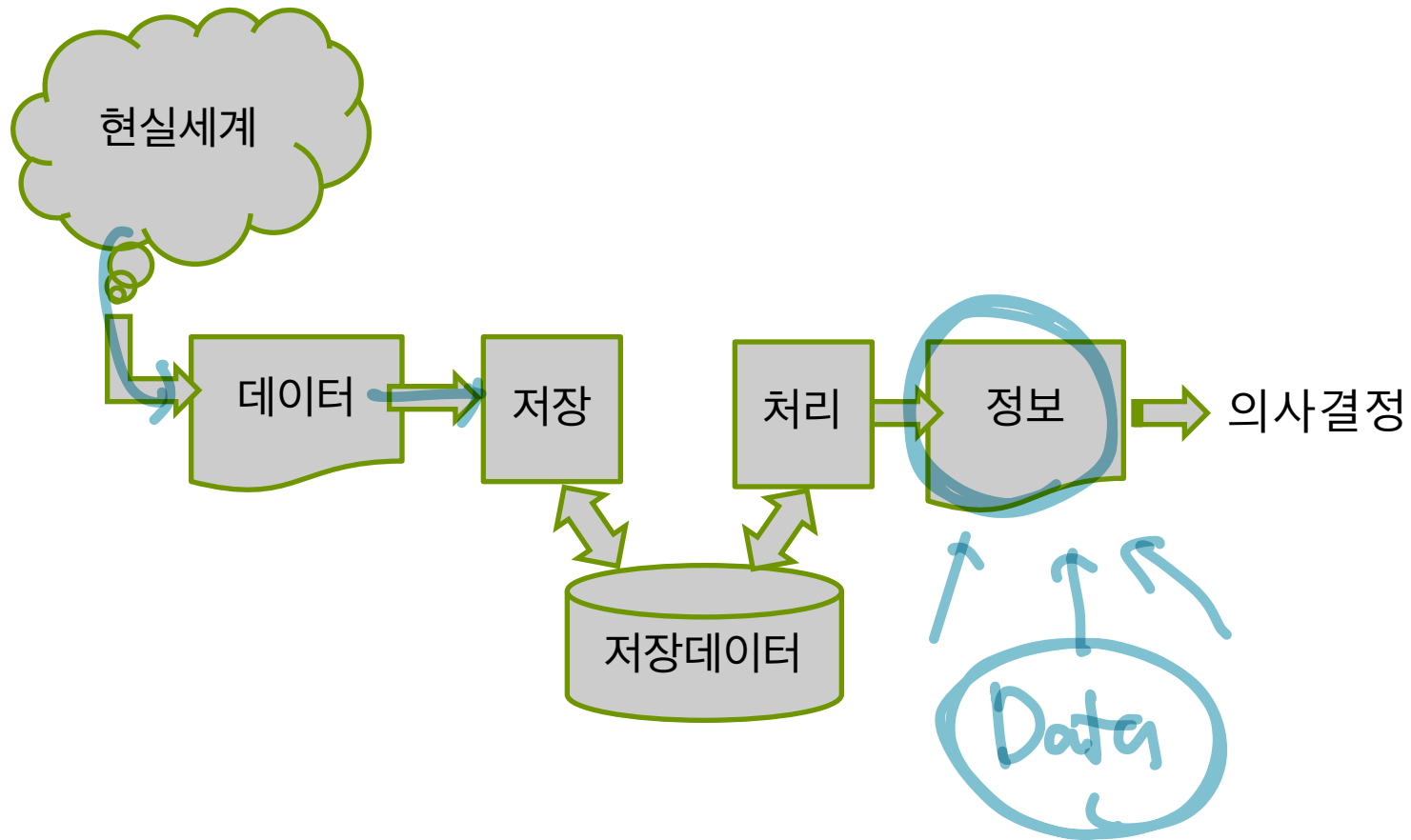
② 정보 : 오늘 예상 날씨 : 감기 걸릴 정도의 나쁜 날씨 →

③ 의사결정 : 방수되는 파카를 가지고 가자

*의사결정이
가능할 수 있다*



데이터와 정보



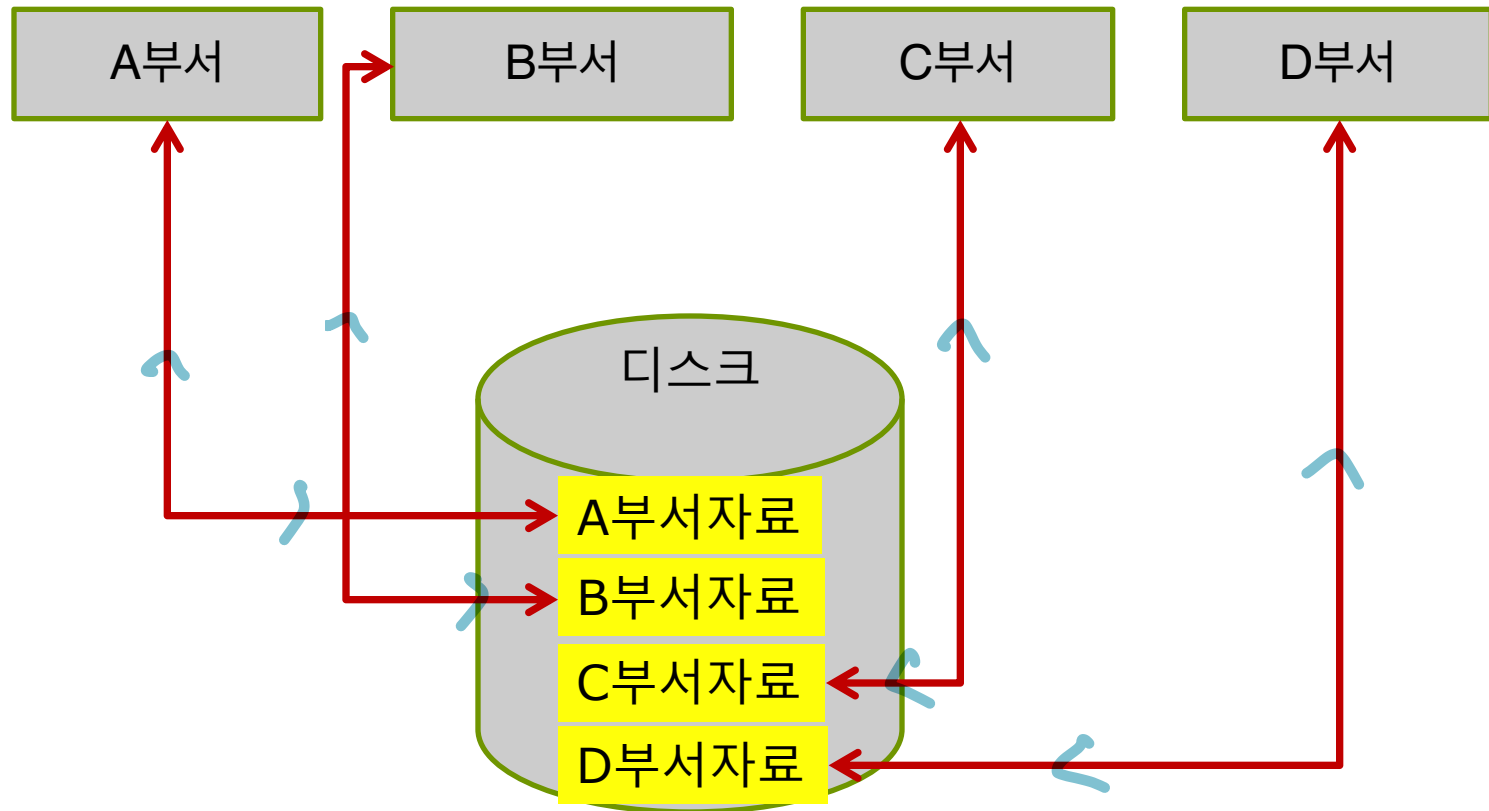
데이터베이스 개념

- 어떤 기관의 각 부서마다 각각 자신들의 파일을 따로 유지하면
- 예를 들어 회원이 100,000명이 있는 아래를 살펴보자.
 - A부서에서는 (주민등록 번호, 이름, 주소, 나이, A 부서에서만 필요한 자료
 - B부서에서는 (주민등록 번호, 이름, 주소, 나이, B 부서에서만 필요한 자료
 - C부서에서는 (주민등록 번호, 이름, 주소, 나이, C 부서에서만 필요한 자료
 - D부서에서는 (주민등록 번호, 이름, 주소, 나이, D 부서에서만 필요한 자료

각 부서

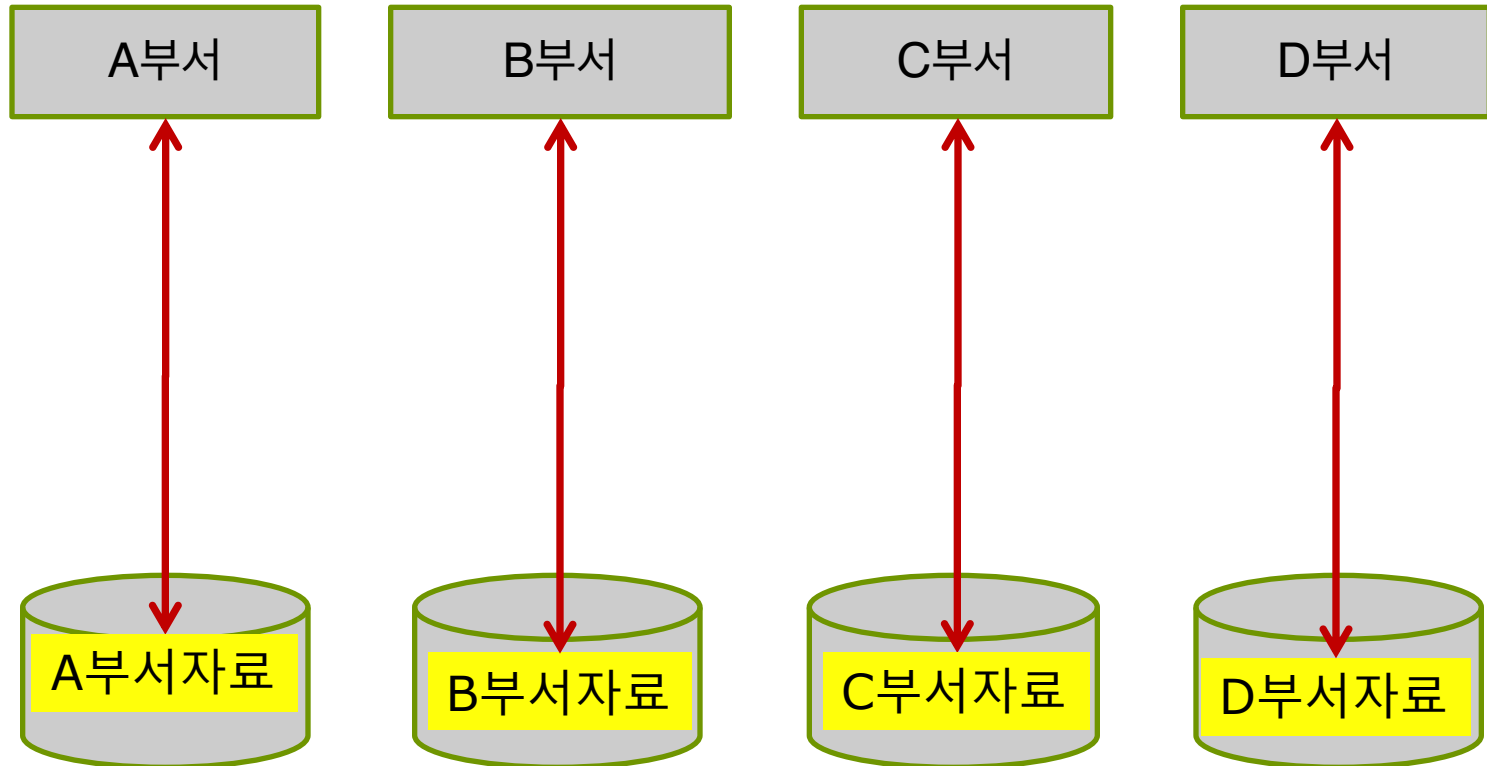
"자료"

부서별 독자적 자료관리시스템 - 1



각 부서마다 자료를 유지/관리하는 시스템

부서별 독자적 자료관리시스템 - 2



각 부서마다 자료를 유지/관리하는 시스템

부서별 독자적 자료관리 문제점

□ 자료 중복

□ 디스크 용량 낭비

용량 × 회원수 × 부서 수

(주민번호, 이름, 주소, 나이)용량(바이트) × 회원수 × 부서 수

□ 자료를 일관성 있게 유지하기 어려움(Data inconsistency 발생)

→ 일관성 있게 유지하기 어려움

create read update delete

▶ 자료의 crud 때 마다 모든 부서 자료를 수정해야 함

□ 자료 불일치 가능성 → 자료의 무결성 보장 못함. ==> 자료를 못 씀

□ 자료의 무결성이 확보되지 못하면

□ "은행 대출부서 : 고객 홍길동이 1달 전에 100억원을 대출해 감

□ 은행 통장관리부서 : 홍길동이라는 고객은 10년 전에 탈퇴 했음

□ 은행 고객관리부 : 홍길동이라는 고객은 20년 전에 돌아가심

□ 홍길동은 돈 빌린 적 없다 주장

□ 이 은행의 자료는 (모두 확인하거나 모두 버리거나) 해야함

□ 중복 자료는 통합관리 해야 함 : 주민등록 번호, 이름, 주소, 나이 통합 관리

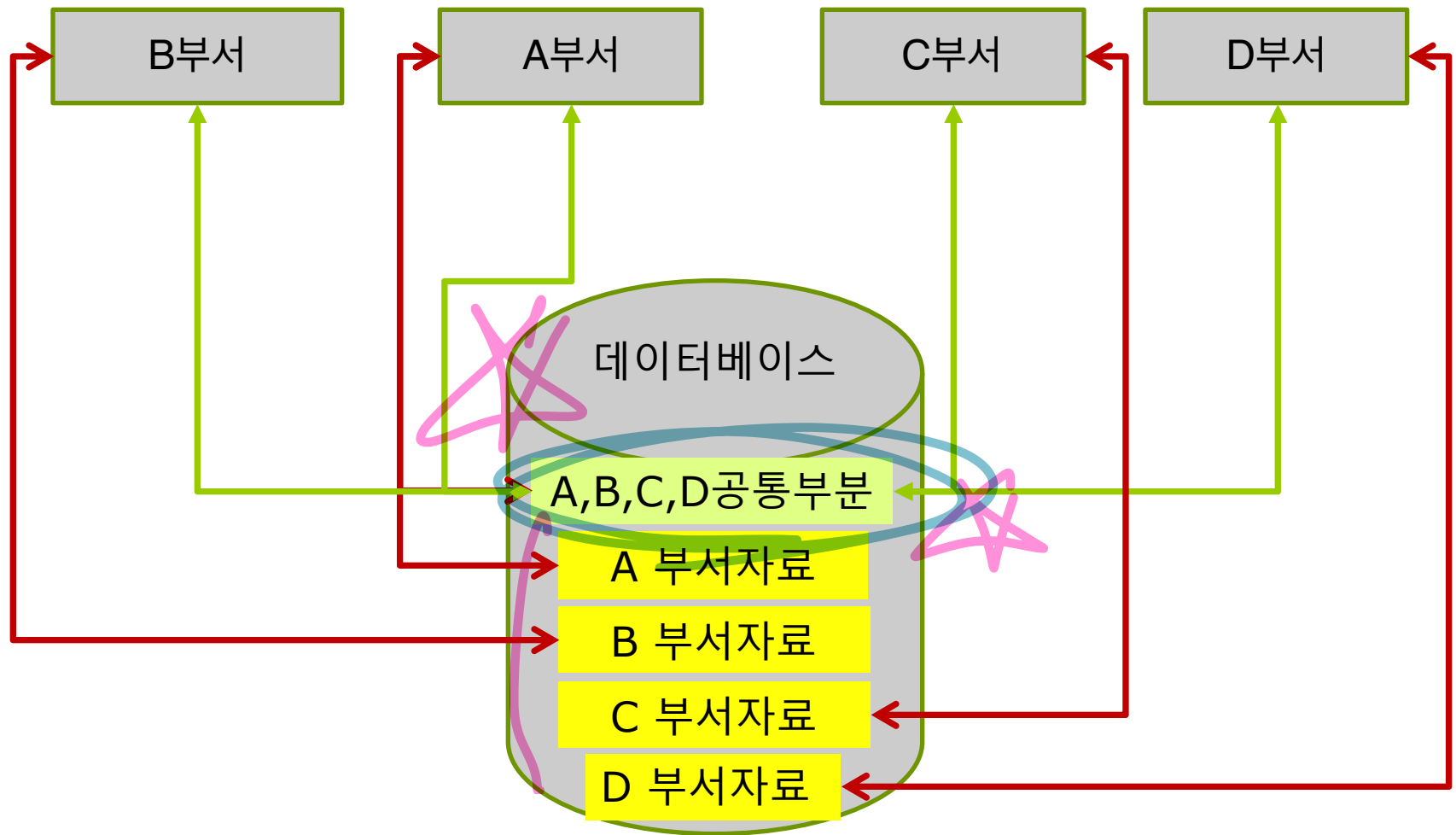
자료의 집중관리 필요

- 불필요한 자료 중복없애기 위해서는
 - 중앙 집중 저장, 관리, 공용 사용 필요
 - 필요한 자료 중복도 최소화
 - ▶ 중복자료는 철저하게 시스템이 기계적으로 관리해야 함.



DBMS 관리

데이터베이스 개념 - 자료중복 배제



데이터베이스 개념-불필요한 중복 배제

데이터베이스 개념

- 데이터의 구조 변경에 따른 응용프로그램 수행문제
 - 물리적인 자료구조 및 파일 시스템 독립인 논리적 자료구조표현
- 논리적으로 연관된 여러 개의 데이터를 데이터의 중복을 최소화하고 데이터를 통합관리
 - (CRUD) 환자 정보.
 - 데이터의 검색과 갱신을 효율적으로 수행
- 데이터가 일관성을 유지하고 더 나아가 자료 무결성 유지
 - 처방 고객 생년월일이 1230년 1월 2일이다 (무결성 의심??)
- 데이터베이스 예 :
 - 병원의 환자 기록 데이터베이스,
 - 은행의 고객 데이터베이스,
 - 학교의 학생 데이터베이스, 도서관의 도서 데이터베이스, 자동차 회사 고객 데이터베이스, 기차나 항공기 예약 현황 데이터베이스

데이터베이스 개념

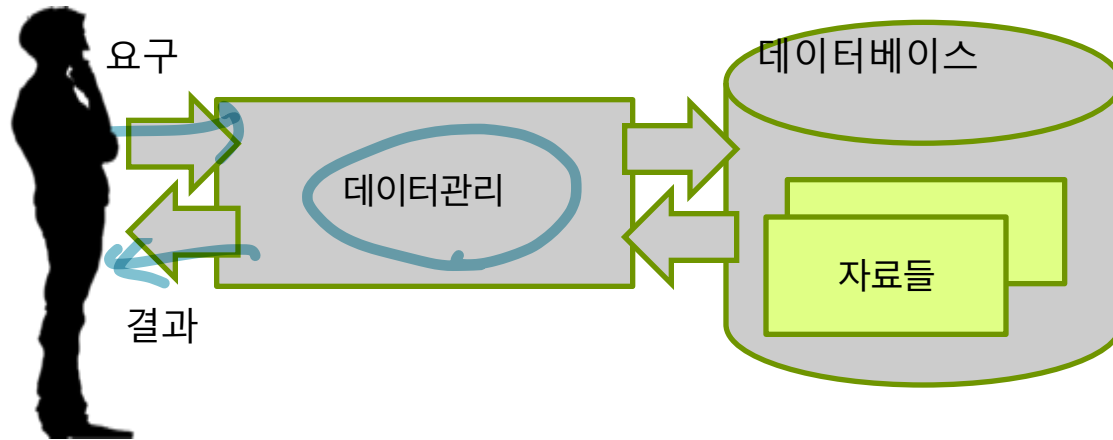
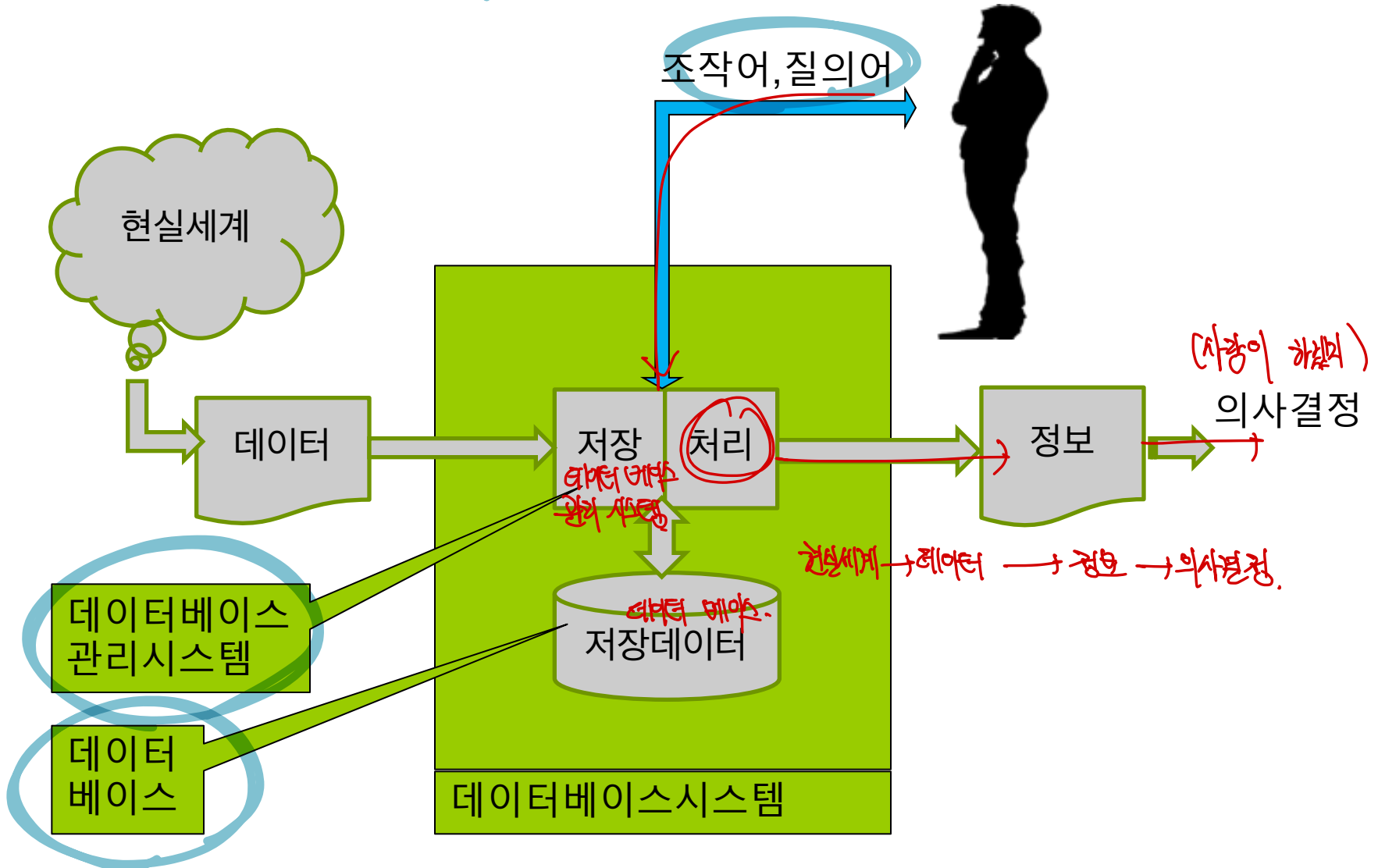


그림7.4 데이터의 보존과 검색 요구 실행 – **textbook** 수정필요

P8 - (데이터베이스 시스템)



데이터베이스관리시스템 - DBMS

(저장, 처리) 등에 사용 됨

- 정보를 정확히 추출하여 유용하게 활용할 수 있도록 잘 정리
- 여러 사용자가 목적에 따라 공유하여 사용
- since 1963년 IDS라는 데이터베이스 관리시스템 개발

데이터 종류

동계원 중목 해

- ① **통합 데이터** : 똑같은 데이터가 중복되지 않음, 논리적으로 통합됨
- ② **공유 데이터** : 여러 응용 프로그램(사용자)이 공동으로 사용과 유지 가능한 데이터, 사용 권한이 다름(권한관리), 네트워크를 통하여 공유
- ③ **운영 데이터** : DB 사용자가 사용하고 유지하는 데이터
 - 실세계 현황을 모델 - 수강생 자료, 개설과목 등
 - 최신의 자료로 계속 crud
- ④ **저장 데이터** : 책상 위의 종이에 적혀있는 자료가 아니라, 반드시 컴퓨터로 접근할 수 있는 저장매체(테이프, 디스크, CD 등)에 저장하되,
 - 저장된 위치나 주소가 아닌 값으로 참조하는 내용에 의한 참조,

권한

"
주제에 의한 참조 x
값으로 참조하는 내용에 의한 참조"

데이터베이스 장점

- 데이터베이스 장점
 - 데이터 중복의 최소화
 - 여러 사용자에게 의한 데이터 공유.
 - 데이터의 일관성
 - 데이터의 무결성
 - 정확한 최신의 정보 이용 가능
 - 데이터 처리속도 증가
 - 백업 및 회복.

원본

부본

8.2 (데이터베이스 시스템)

데이터베이스 시스템의 개념

□ 데이터베이스 시스템은

□ 컴퓨터 화일시스템에 저장한 데이터(DB)

□ 데이터를 저장하고 저장된 데이터를 관리하여 사용자에게 정보를 생성해 주는 시스템(DBMS) → Data Base Management System.

□ 데이터를 저장할 DB를 생성, 삭제, 갱신, 조회(데이터언어)하고 데이터를 DB에 CRUD (데이터언어)

DB를 (생성, 삭제, 갱신, 조회) 하는

□ 데이터베이스 시스템의 구성

□ 데이터베이스, 데이터베이스 관리 시스템, 데이터언어, 사용자

= DB를 만들

데이터베이스 시스템

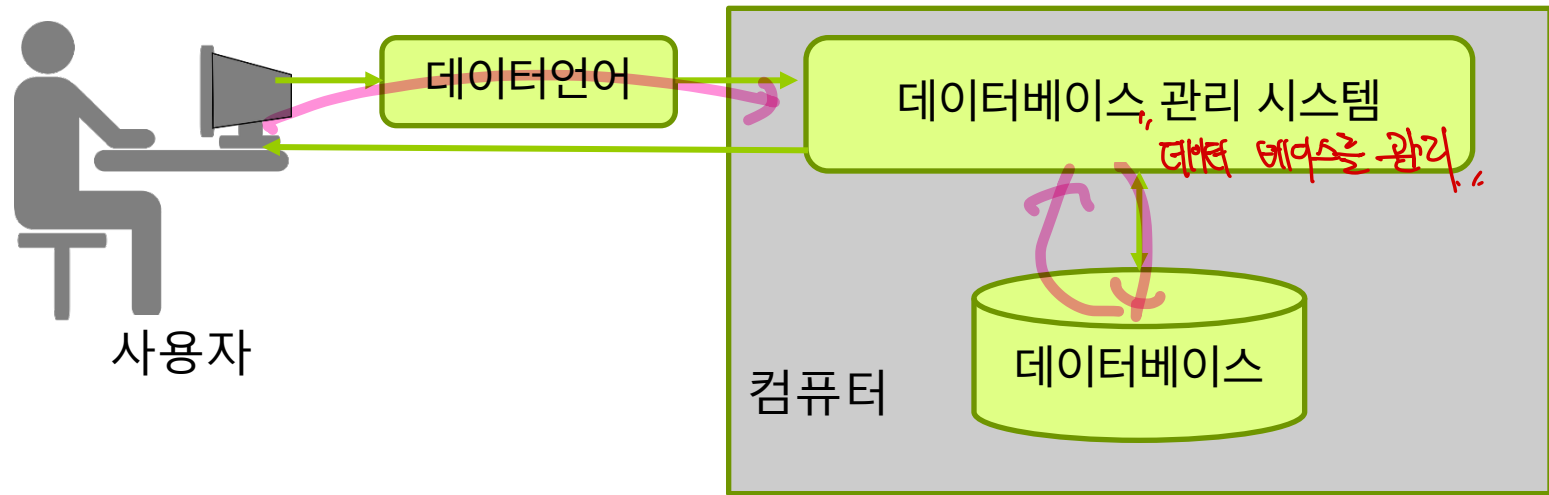


그림7.5 데이터베이스 시스템

데이터베이스 시스템

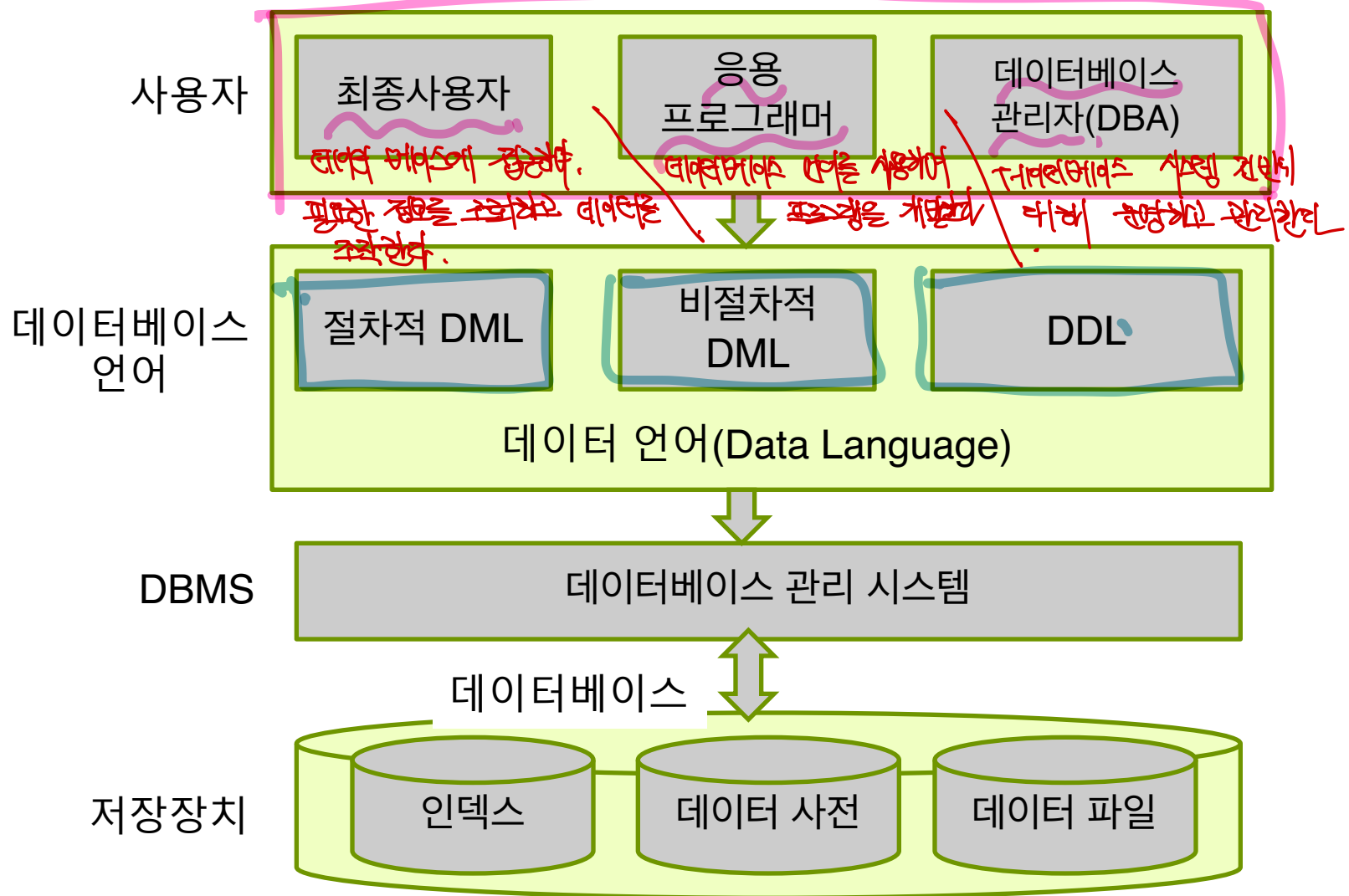


그림 7.7 데이터베이스 시스템의 구성

데이터베이스 사용자

- 데이터베이스 시스템 사용자는 데이터베이스에 접근하는 모든 사람
- 목적에 따라 크게
 - 데이터베이스 관리자,
 - 응용 프로그래머,
 - 최종 사용자

데이터베이스 관리자(DBA, DB Administrator)

데이터베이스 관리자

- 데이터베이스 시스템을 총괄 운영하고 관리
- 데이터베이스를 설치, 구축, 갱신, 관리, 분석
- 데이터베이스 시스템의 고장이나 문제의 원인을 파악하여 복구하고,
- 사용자에게 교육과 기술을 지원
- 데이터베이스를 주기적으로 백업하여 데이터의 손실을 예방
- 자료를 수집하여 정리하고, 입력된 자료를 검토하여 문제점을 보완
- 데이터베이스에 접근하여 사용자 등록, 데이터의 접근 범위, 읽기/쓰기 권한, 사용자 로그인과 암호 관리
- 데이터 정의 언어와 제어 언어를 주로 사용
- 데이터베이스가 큰 조직에서는 데이터베이스 설계자가 데이터베이스 설계만 따로 담당

응용프로그래머

데이터를 관리

→ 데이터베이스의 저장 구조를 정의하고 설계

- 데이터베이스 관리 프로그램을 사용하여 데이터베이스의 저장 구조를 정의하고 설계
- 학교 전산화 응용프로그래머
- 학생데이터베이스, 교실데이터베이스, 성적데이터베이스 등 테이블 구조 만들기
- 최종 사용자의 요구에 따른 인터페이스와 질의 처리에 적절한 응용 프로그램을 개발
- 데이터베이스 프로그램을 개발하는 개발자

일반 사용자

- 최종 사용자로서 응용 프로그래머가 개발한 프로그램으로 데이터베이스에 접근하여 자신이 원하는 자료를 사용
- 조직의 업무를 담당하거나 홈페이지 등에서 상품을 조회하거나 구매하기도 한다.

데이터베이스 언어

- 관리자용 언어 사용권한 필요
- 데이터 조작 언어, 데이터 정의 언어, 데이터 제어 언어로 구성
 - **데이터 조작 언어 (DML)**: 데이터 검색, 등록, 갱신, 삭제 ^{CRUD}
 - **데이터 정의 언어 (DDL)**: 데이터 구조를 생성, 갱신, 삭제 ^{DDL}
 - **데이터 제어 언어**: 접근을 제어하는 언어 ^{ACL}
- **SQL** (Structured Query Language) 가장 많이 사용

데이터베이스 관리 시스템(DBMS)

- 사용자 및 응용 프로그램과 데이터베이스 사이의 인터페이스
 - 사용자 및 응용 프로그램이 데이터베이스 관리 시스템(DBMS)을 통하여 데이터베이스를 이용
- SQL을 사용하여 자료구조 및 자료를 CRUD
 - Data Definition
 - Data Manipulation
- 제어기능 : 데이터 사용 권한 제어, 보안, 병행 수행제어.
 - 같은 레코드에 여러 사용자가 (동시 접근 금지) : 어떤 사용자가 한 데이터를 사용 중일 때는 다른 사용자가 사용할 수 없도록 한다.
 - 데이터 보호를 위해 패스워드 관리, 장애에 대비한 마스터 파일 이중화 등

실제로 컴퓨터에 설치하여 사용하는 소프트웨어.

자료구조에 데이터베이스를 구축하고, 자료를 일관성 있고, 효율하게 저장하고

관찰하여, DB언어를 통하여 방문 사용자의 자료 생성 조작 삭제 행위를 처리

상용 DBMS

- 
- SQL 서버
 - DB2
 - 인포믹스,
 - 오라클
 - 사이베이스
 - MySQL 등이 대표적

데이터 모델

- 실세계 데이터를 구조화한 자료 모델.
- 계층 모델, 네트워크 모델, 관계 모델, 객체-지향 모델, 멀티미디어 모델 등 다양.
- ① □ 관계형 데이터 모델 – 가장 많이 사용
 - 데이터를 테이블로 표현하고 집합으로 취급
 - 테이블을 관계(relation)라고 함
 - Relational DB(관계형DB)는 테이블로 자료를 나타내는 DB
 - 데이터를 조작하기 위해 집합 연산을 테이블에 적용하여 실행
 - ▶ relational algebra- set theory 기반 algebra
 - 그림 7.16, 7.17

ADB → 테이블로 자료를 나타냄

관계형모델

학생 테이블 ← 테이블(=릴레이션, =엔티티, =관계)

필드(열, 속성)

필드명

학번	이름	학과명	성별	주소
20051201	김철수	컴퓨터	남	대구
20050315	이영희	수학	여	포항
20040527	홍길동	과학	남	구미
20031209	이영희	컴퓨터	여	경주
20040307	박희동	수학	남	대구

레코드
(행, 튜플)

= 학생 데이터

그림 7.16 학생 관련 테이블

필드: 열, 속성

필드명

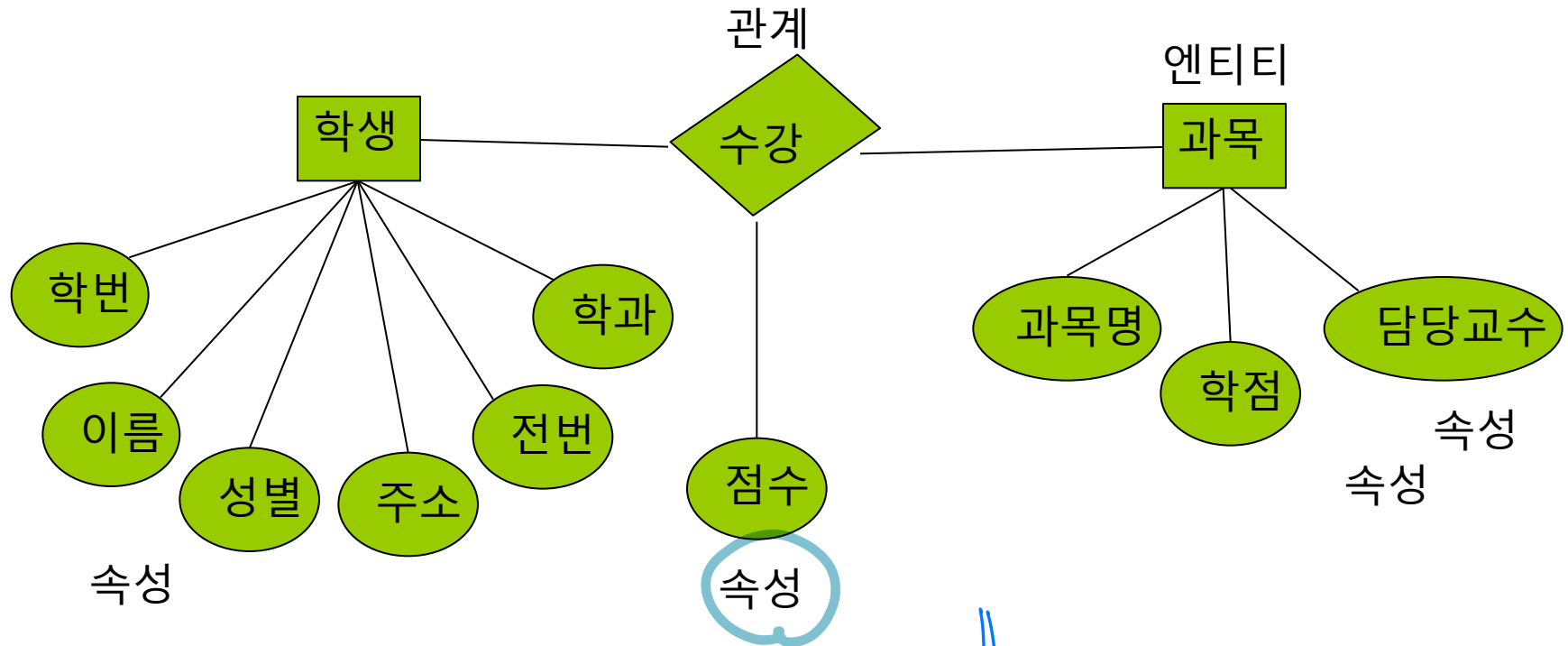
레코드: 행, 튜플

테이블: 릴레이션, 엔티티, 관계

= 현재 사제 관련 학생 데이터

관계형 data 모델(Entity-Relationship-diagram)

- 실세계의 엔티티와 그 사이의 관계를 모델
- 엔티티 : 현실 세계 관리 대상 개체 (학생, 학과, 과목, 성적, 교수, 수강 등)



↓ "ER diagram" 대용

관계형 데이터모델(ER diagram 대응)

□ E-R diagram → Relation(table)로 변경

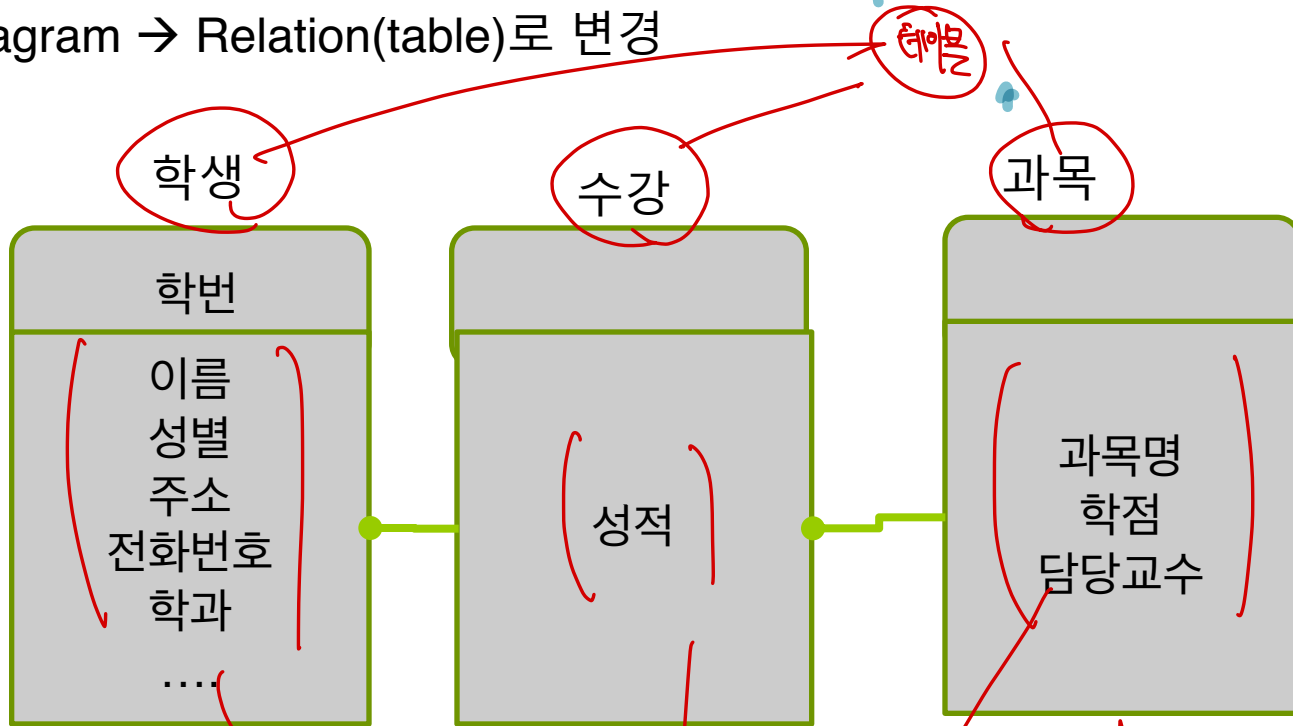


그림 7.17 관계 데이터베이스 모델

= ER diagram

데이터베이스(ER-diagram)

- E-R diagram → Relation(table)로 변경 (테이블의 키 결정)

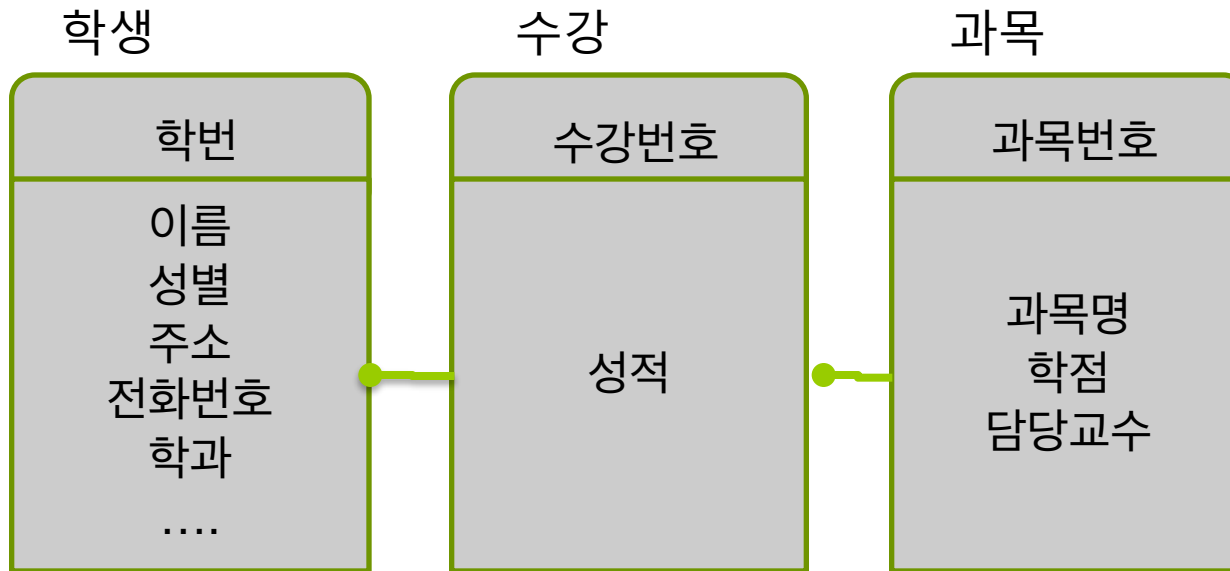


그림 8.8 관계 데이터베이스 모델

이제 공부

관계형 데이터 모델 – Relational DB

- E-R diagram에 대응하는 RDB의 Relation(table) 들



그림 7.17 관계 데이터베이스 모델

관계형 데이터 모델

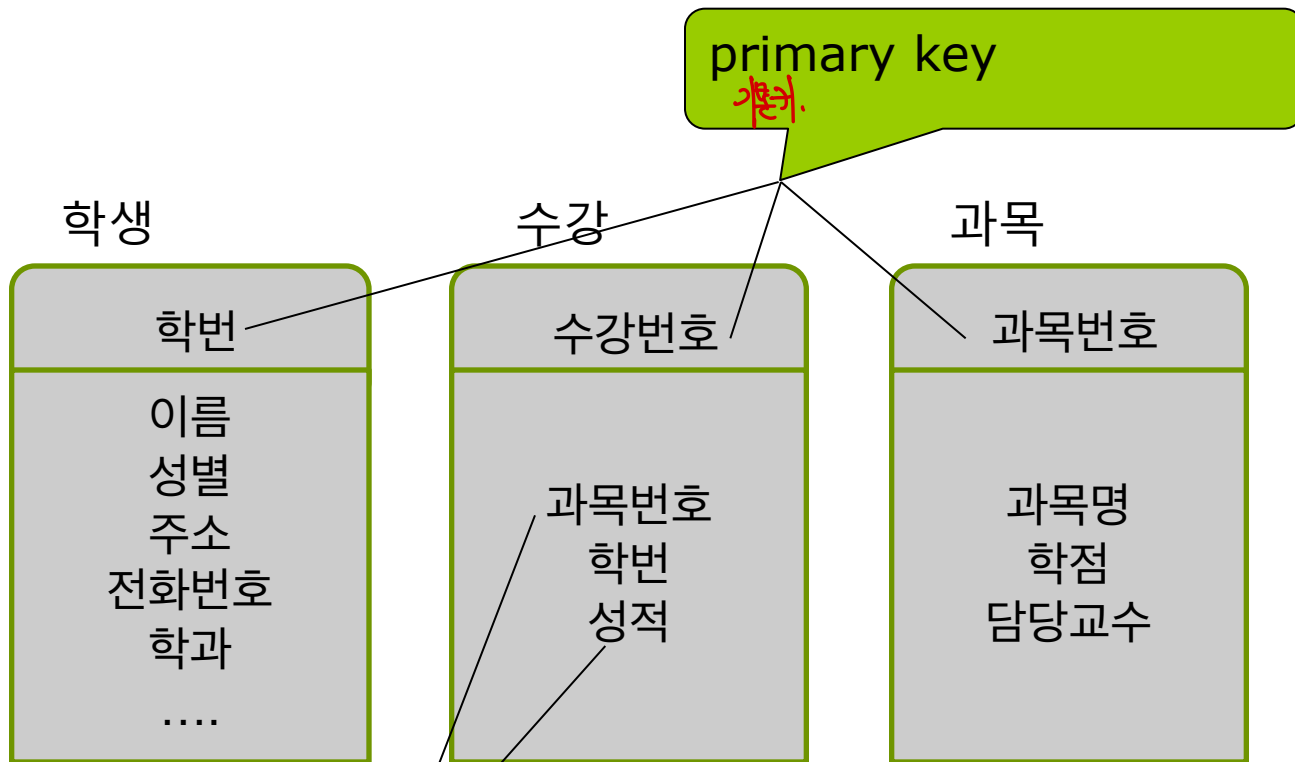


그림 8.8 관계 데이터베이스 모델

Foreign key
외래키

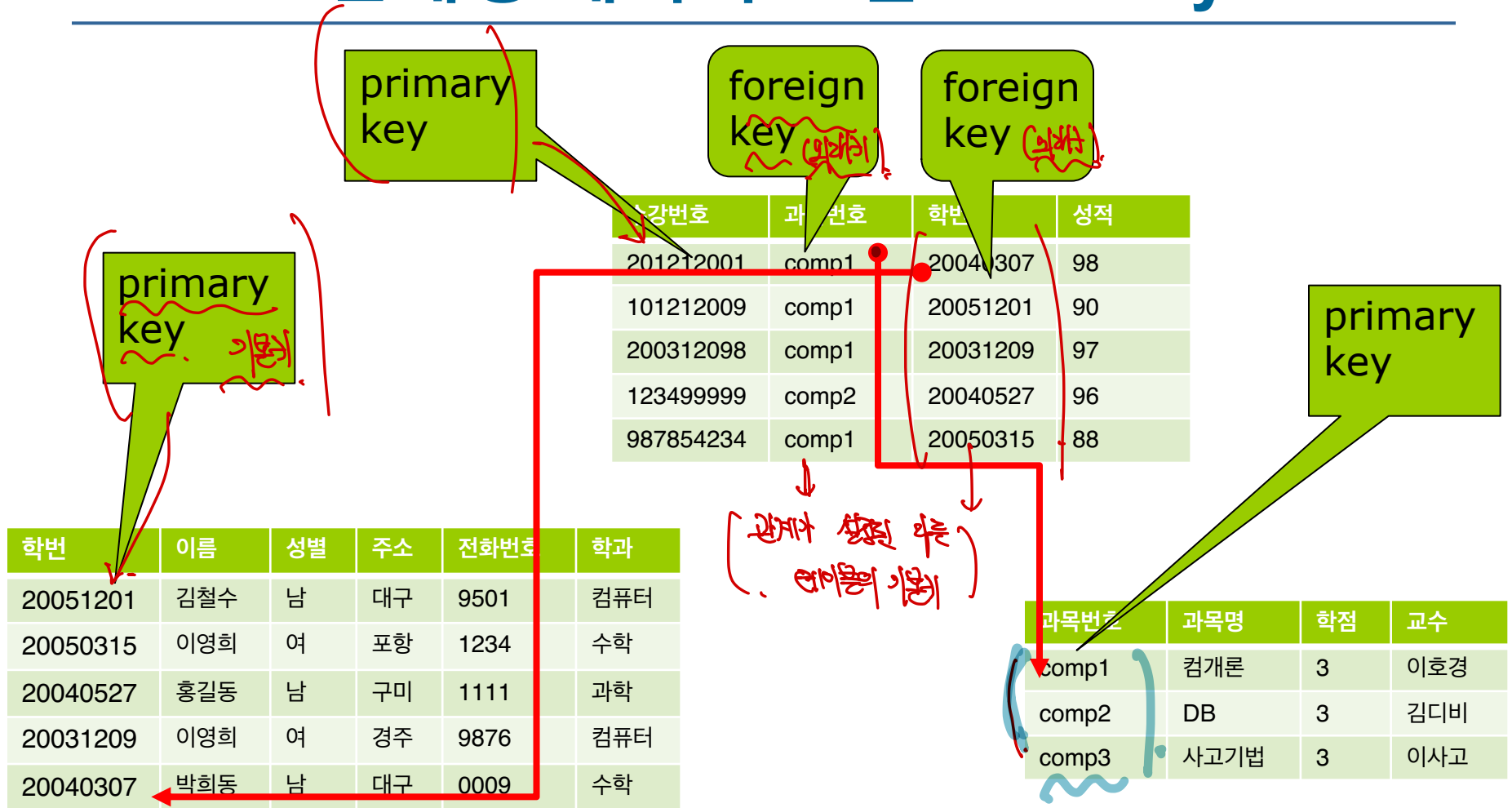
관계형 데이터 모델 – snap shot

수강번호	과목번호	학번	성적
201212001	comp1	20040307	98
101212009	comp1	20051201	90
200312098	comp1	20031209	97
123499999	comp2	20040527	96
987854234	comp1	20050315	88

학번	이름	성별	주소	전화번호	학과
20051201	김철수	남	대구	9501	컴퓨터
20050315	이영희	여	포항	1234	수학
20040527	홍길동	남	구미	1111	과학
20031209	이영희	여	경주	9876	컴퓨터
20040307	박희동	남	대구	0009	수학

과목번호	과목명	학점	교수
comp1	컴개론	3	이호경
comp2	DB	3	김디비
comp3	사고기법	3	이사고

관계형 데이터 모델 - DB Key



□ DB의 논리적인 자료 구조

관계형 데이터 모델

□ 데이터베이스는 테이블, 필드, 레코드, 도메인, 키의 요소로 구성

□ **테이블** : 관계(릴레이션)를 엔티티라고도 하며, 열과 행으로 구성한다.

↓ □ **필드** : 열을 말하는 것으로 속성이라 하며, 데이터 값을 저장하는 단위이다. 학번열, 이름열, 학과면열, 성별열, 주소열을 필드명이라 한다.

→ □ **레코드** : 행을 말하는 것으로 ~~투플~~ 튜플이라 하며, 각 행의 데이터 내용이다.

□ **도메인** : 필드가 가질 수 있는 값들의 집합이다

필드의 원수 또는 값들의 집합

□ **키** : 주어진 관계에서 모든 인스턴스 가운데 유일성을 갖는 여러 개의 속성 집합이다. 키로 데이터를 구별하기 때문에 식별자라 한다.

키의 유형에는

후보 키 (한 테이블에서 데이터 레코드를 고유하게 식별할 수 있는 필드),

기본 키 (후보키 중에서 대표로 선택한 키),

수퍼 키 (복합 키라고도 하며, 두 개 이상의 필드로 구성된 기본 키),

외래 키 (참조 키라고도 하며, 관계가 설정된 다른 테이블의 기본 키)

의 유형이 있다.

관계형 데이터 모델-key

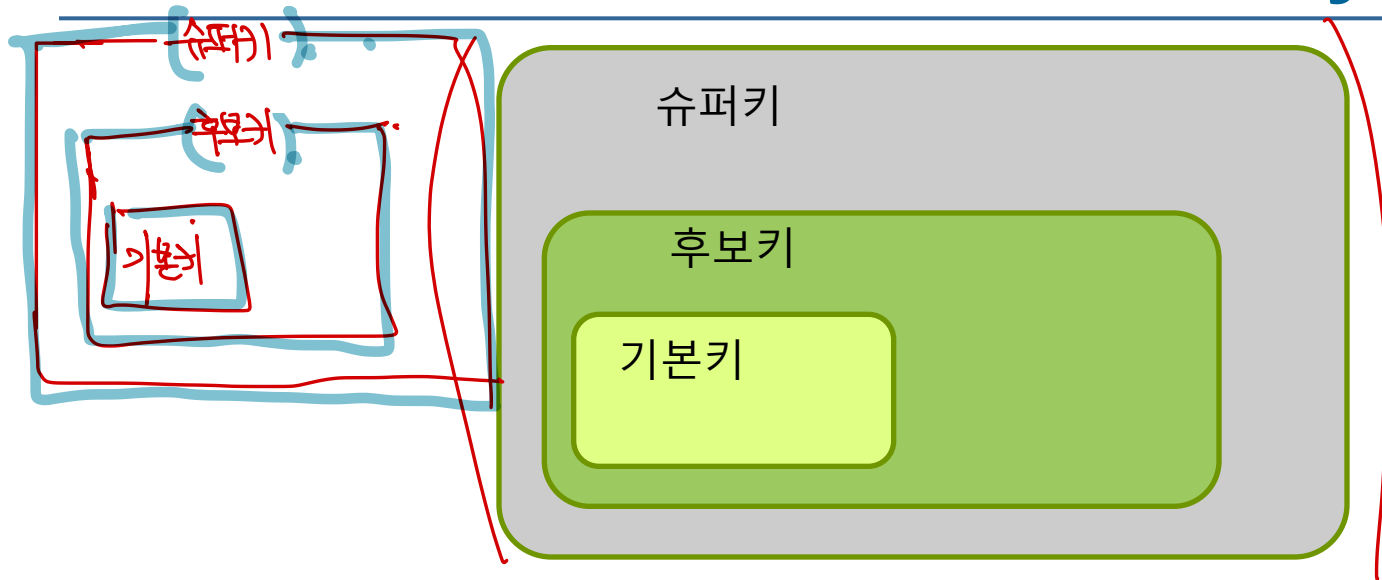
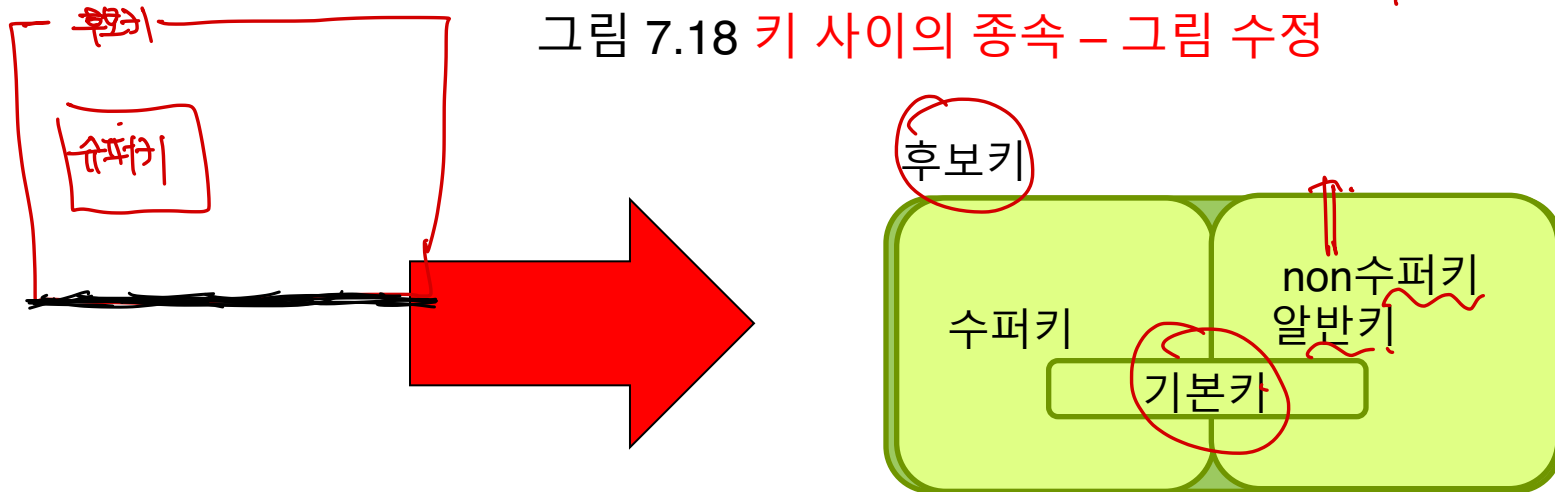


그림 7.18 키 사이의 종속 - 그림 수정



(관계형 데이터 베이스)

RDB.

□ 관계형 데이터 모델을 지원하고 관계연산을 지원하는 데이터베이스 시스템

□ DB의 물리적 구조 관점 (internal schema) - 물리적인 저장장치 관점에서의 저장구조 - DB개발자

□ DB의 논리적인 자료 모델 (conceptual schema) - 통합 DB 관점

중복X

→ 에 맞게 데이터를 적어야 함

수강번호	과목번호	학번	성적
201212001	comp1	20040307	98
101212009	comp1	20051201	90
200312098	comp1	20031209	97
123499999	comp2	20040527	96
987854234	comp1	20050315	88

과목번호	과목명	학점	교수
comp1	컴개론	3	이호경
comp2	DB	3	김디비
comp3	사고기법	3	이사고

학번	이름	성별	주소	전화번호	학과
20051201	김철수	남	대구	9501	컴퓨터
20050315	이영희	여	포항	1234	수학
20040527	홍길동	남	구미	1111	과학
20031209	이영희	여	경주	9876	컴퓨터
20040307	박희동	남	대구	0009	수학

DB 사용자 관점 (External Schema)

- DB 사용자마다 관점이 다르다 사용자 관점별 자료 모델 구조
- A,B,C,D 부서의 운용자료가 다르다. (A,B,C,D 부서의 external schema)
- A,B,C, D 부서의 모든 자료를 모은 DB의 구조 conceptual schema

수강번호	과목번호	학번	성적
201212001	comp1	20040307	98
101212009	comp1	20051201	90
200312098	comp1	20031209	97
123499999	comp2	20040527	96
987854234	comp1	20050315	88

과목번호	과목명	학점	교수
comp1	컴개론	3	이호경
comp2	DB	3	김디비
comp3	사고기법	3	이사고

학번	이름	성별	주소	전화번호	학과
20051201	김철수	남	대구	9501	CS
20050315	이영희	여	포항	1234	MAT
20040527	홍길동	남	구미	1111	SCI
20031209	이영희	여	경주	9876	CS
20040307	박희동	남	대구	0009	MAT

학과코드	학과명	건물
CS	컴퓨터학과	IT5
MAT	수학과	MA1
SCI	과학과	SC1

- 입학생을 등록하는 부서의 external schema

(입학생을 등록하는 부서의 external schema)

RDB – 관계연산 SQL -DDL

□ CREATE TABLE 학생 {

학번 varchar(10) not NULL

이름 varchar(10) not NULL

학과 varchar(8) not NULL

.....

PRIMARY KEY(학번)

}

주키 = 기본키 (학번)

학번	이름	학과	성별	주소	특산물	효능
20051201	김철수	컴퓨터	남	대구	사과	피부
20050315	이영희	수학	여	포항	과메기	뇌
20040527	홍길동	과학	남	구미	방울토마토	건강
20031209	이영희	컴퓨터	여	경주	배	폐
20040307	박희동	수학	남	대구	사과	피부

RDB : (SQL-DDL)

Data Define Language.

```
□ ALTER TABLE 학생 {  
    ..... 속성 변경  
}
```

```
DROP TABLE 학생 CSACADE I RESTRICT;
```

관계형 데이터베이스 연산

- 사용자 요구결과 테이블에서 새로운 테이블을 생성 - DDL
- SELECT, PROJECT, JOIN 등의 연산을 이용 - DML
- 학생과 학과 정보를 저장하고 있는 관계 데이터베이스에서 예를 들어보자.
 - 학생 정보와 학과 정보를 저장하고 있는 두 개의 테이블.
 - 학생 테이블에서 각행은 한 학생에 대한 데이터 집합이고,
 - 각 열은 학생의 여러 속성 값이다.

↓ ↓ ↓ 학생의 속성
→

RDB 연산

테이블 이름 : 학생

학사정보

학번	이름	학년	학과
20161345	김미진	2	체육학과
20162482	최은경	2	심리학과
20173344	여진규	1	컴퓨터과학과

→ 학생 정보
↓ 추가.
학생 정보.

테이블 이름 : 학과

수업정보

학과번호	학과	정원
D1	심리학과	50명
D2	체육학과	10명
D3	컴퓨터과학과	30명

→ 수업정보.
학과 정보.

데이터베이스 모델

- SELECT 연산은 테이블에서 지정된 행을 추출

```
SELECT '학과' from 학과 where '정원' > 25명
```

select '학과' from. 학과. where '정원' > 25

- '학과' 테이블에서 '정원' 이 25명 이상인 '학과' 이름을 찾아 그 행을 모두 출력 하라
- 결과

학과번호	<i>학과</i>	정원
D1	심리학과	50명
D3	컴퓨터과학과	30명

(관계형데이터베이스 SQL-DML)

- (select, insert, update, delete)
- 테이블의 (엔트리 조회 삽입 갱신 삭제)

studentTable

학번	이름	학과	성별	주소	특산물	효능
20051201	김철수	컴퓨터	남	대구	사과	피부
20050315	이영희	수학	여	포항	과메기	뇌
20040527	홍길동	과학	남	구미	방울토마토	건강
20031209	이영희	컴퓨터	여	경주	배	폐
20040307	박희동	수학	남	대구	사과	피부

- SELECT 학번 주소 (FROM studentTable WHERE 성별 = '남')

학번	주소
20051201	대구
20040527	구미
20040307	대구

관계형데이터베이스 SQL DML

- (select, insert, update, delete)
 (생성, 추가, 갱신, 삭제)
- 테이블의 엔트리 조회 삽입 갱신 삭제
- INSERT INTO studentTable (학번, 이름, 학과, 성별, 주소, 특산물, 효능)
VALUES ('20180001', '최유경', '컴퓨터', '여', '대구', '사과', '피부')
→ 도메인에 맞게 추가

studentTable

학번	이름	학과	성별	주소	특산물	효능
20051201	김철수	컴퓨터	남	대구	사과	피부
20050315	이영희	수학	여	포항	과메기	뇌
20040527	홍길동	과학	남	구미	방울토마토	건강
20031209	이영희	컴퓨터	여	경주	배	폐
20040307	박희동	수학	남	대구	사과	피부
20180001	최유경	컴퓨터	여	대구	사과	피부

7. 데이터베이스 모델

- PROJECT 연산은 테이블에서 특정 행을 추출

PROJECT '학생' over '학번', '이름'

- '학생' 테이블에서 '학번' 과 '이름' 찾아 그 행을 모두 출력하라
- 결과



학번	이름
20161345	김미진
20162482	최은경
20173344	여진규

project '학생' over '학번', '이름'

데이터베이스 – JOIN연산

- JOIN 연산은 공통 행을 기준으로 두 개 테이블을 합친다. 이 때 반드시 공통된 열이 있어야 한다.
만약 두 테이블에 공통된 열이 있더라도 일치하는 값이 없으면 연산은 제대로 수행이 되지 않을 수도 있다.

JOIN '학과' and '학생' over '학과'

- '학과' 테이블과 '학생' 테이블에서 '학과' 를 공통으로 하여, 두 테이블을 합치고 출력하라
- 연산 및 결과-next page

Join '학과' and '학생' over '학과'

Join 연산 및 결과

학과	학번	이름	학년	학과
	20161345	김미진	2	체육학과
	20162482	최은경	2	심리학과
	20173344	여진규	1	컴퓨터과학과

학생	학과번호	학과	정원
	D1	심리학과	50명
	D2	체육학과	10명
	D3	컴퓨터과학과	30명

학생 ← 학과, '학과' 면 같.

JOIN '학과' and '학생' over '학과'

학과	학과번호	정원	학번	이름	학년
심리학과	D1	50	20162482	최은경	2
체육학과	D2	10	20161345	김미진	2
컴퓨터과학과	D3	30	20173344	여진규	1

이원

Join연산 활용

학과	학번	이름	학년	학과
	20161345	김미진	2	체육학과
	20162482	최은경	2	심리학과
	20173344	여진규	1	컴퓨터과학과

학생	학과번호	학과	정원
	D1	심리학과	50명
	D2	체육학과	10명
	D3	컴퓨터과학과	30명

JOIN '학과' and '학생' over '학과'

학과	학과번호	정원	학번	이름	학년
심리학과	D1	50	20162482	최은경	2
체육학과	D2	10	20161345	김미진	2
컴퓨터과학과	D3	30	20173344	여진규	1

(정원이 40명을 초과하는 학과의 학생의 학년을 나열하라!)

Select '학년' from (JOIN '학과' and '학생' over '학과')
where '정원' > 40

관계형데이터베이스 SQL-DCL

- Data Control Language
- COMMIT: 자료변경 확정
- ROLLBACK: 자료변경 취소
- GRANT: 자료조작 권한 할당, REVOKE: 자료조작 권한 회수

생각해보기-테이블 구조 설계

- 어떤 학생이 어떤 교수가 강의하는 어떤 과목을 수강하며 몇 점을 받았는지 관리하는 테이블을 만들고자 한다.

학번	이름	학과	성별	주소	과목	수강	학점	교수	성적
20051201	김철수	컴퓨터	남	대구	컴개론	comp1	3	이호경	98
20050315	이영희	수학	여	포항	컴개론	comp2	3	김길동	90
20040527	홍길동	과학	남	구미	컴개론	comp1	3	이호경	97
20031209	이영희	컴퓨터	여	경주	컴개론	comp2	3	이호경	96
20040307	박희동	수학	남	대구	컴개론	comp1	3	이호경	88

두 테이블의 문제점은 무엇인가?

학번	이름	학과	성별	주소	특산물	효능
20051201	김철수	컴퓨터	남	(대구)	사과	피부
20050315	이영희	수학	여	포항	과메기	뇌
20040527	홍길동	과학	남	구미	방울토마토	건강
20031209	이영희	컴퓨터	여	경주	배	폐
20040307	박희동	수학	남	(대구)	사과	피부

↓
여는게 좋은 거라.

불필요한 자료의 중복

□ 불필요한 자료 중복

- 대구의 특산물이 사과에서 배로 변화하면 테이블의 끝까지 찾아서 수정 필요함(경북대 출신 모든 학생을 관리하고 있다면 수억 개의 튜플 또는 레코드를 수정해야 함)

□ 자료의 불일치

- 자료가 입력이나 수정단계에서 실수로 인해 불일치가 발생하면 어떻게 수정할 것인가
- comp1을 강의하는 사람이 이호경인가? 김호경인가?

학번	이름	학과	성별	주소	과목	수강	학점	교수	성적
20051201	김철수	컴퓨터	남	대구	컴개론	comp1	3	이호경	98
20050315	이영희	수학	여	포항	컴개론	comp2	3	김길동	90
20040527	홍길동	과학	남	구미	컴개론	comp1	3	이호경	97
20031209	이영희	컴퓨터	여	경주	컴개론	comp2	3	이호경	96
20040307	박희동	수학	남	대구	컴개론	comp1	3	이호경	88

학번	이름	학과	성별	주소
20051201	김철수	컴퓨터	남	대구
20050315	이영희	수학	여	포항
20040527	홍길동	과학	남	구미
20031209	이영희	컴퓨터	여	경주
20040307	박희동	수학	남	대구

과목	수강	학점	교수
컴개론	comp1	3	이호경
컴개론	comp2	3	김길동

학번	수강	학점	성적
20051201	comp1	3	98
20050315	comp2	3	90
20040527	comp1	3	97
20031209	comp2	3	96
20040307	comp1	3	88

특성이 맞게 분리

학번	이름	학과	성별	주소	특산물	효능
20051201	김철수	컴퓨터	남	대구	사과	피부
20050315	이영희	수학	여	포항	과메기	뇌
20040527	홍길동	과학	남	구미	방울토마토	건강
20031209	이영희	컴퓨터	여	경주	배	폐
20040307	박희동	수학	남	대구	사과	피부

학번	이름	학과	성별	주소	특산물	특산물	효능
20051201	김철수	컴퓨터	남	대구	사과	사과	피부
20050315	이영희	수학	여	포항	과메기	과메기	뇌
20040527	홍길동	과학	남	구미	방울토마토	방울토마토	건강
20031209	이영희	컴퓨터	여	경주	배	배	폐
20040307	박희동	수학	남				