

Chapter 01

데이터베이스 시스템

...
오라클로 배우는 데이터베이스 개론과 실습



1. 데이터, 정보, 지식

- 데이터 : 관찰의 결과로 나타난 정량적 혹은 정성적인 실제 값
- 정보 : 데이터에 의미를 부여한 것
- 지식 : 사물이나 현상에 대한 이해



그림 1-1 데이터, 정보, 지식

2. 일상생활의 데이터베이스

- 데이터베이스 : 조직에 필요한 정보를 얻기 위해 논리적으로 연관된 데이터를 모아 구조적으로 통합해 놓은 것

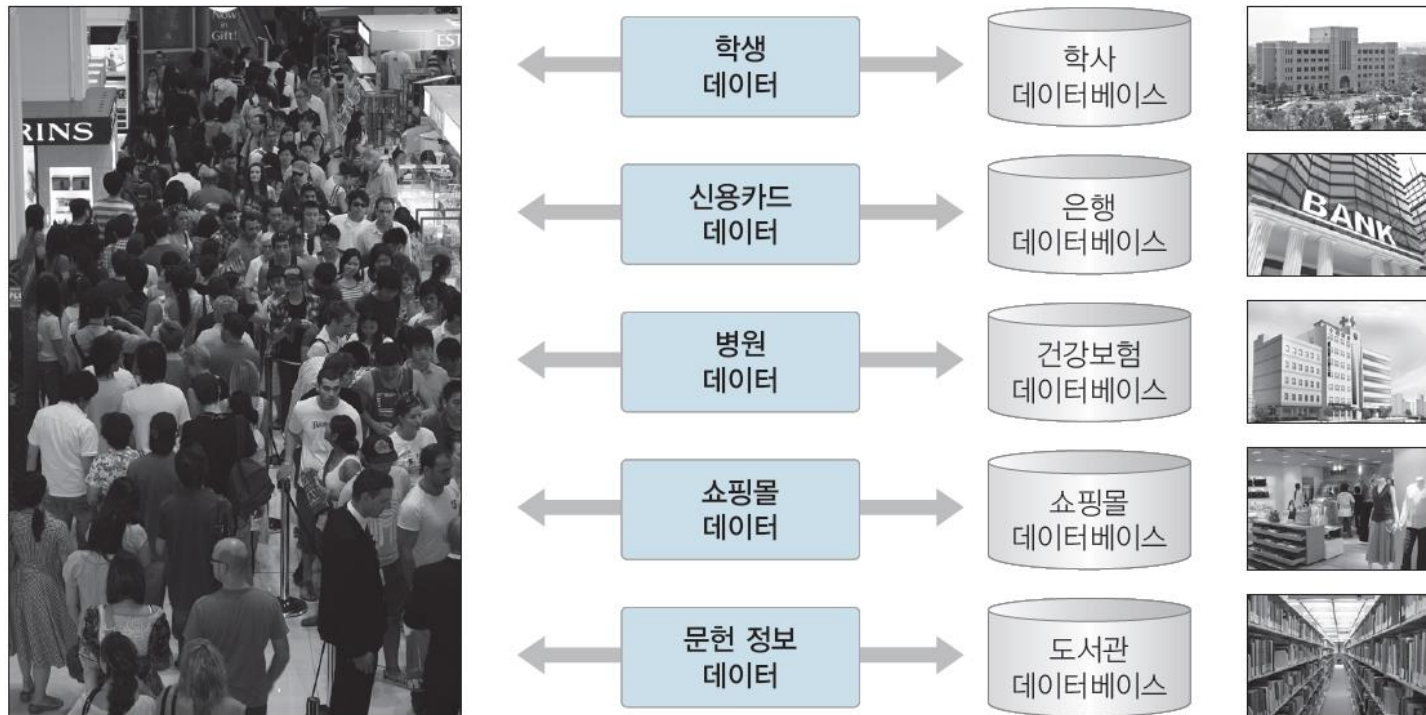


그림 1-2 일상생활에서 생성되는 데이터베이스

2. 일상생활의 데이터베이스

- 데이터베이스 시스템은 데이터의 검색과 변경 작업을 주로 수행함
- 변경이란 시간에 따라 변하는 데이터 값을 데이터베이스에 반영하기 위해 수행하는 삽입, 삭제, 수정 등의 작업을 말함

표 1-2 검색과 변경 빈도에 따른 데이터베이스 유형

유형	검색 빈도	변경 빈도	데이터베이스 예	특징
유형1	적다	적다	공통 데이터베이스	<ul style="list-style-type: none">• 검색이 많지 않아 데이터베이스를 구축할 필요 없음• 보존가치가 있는 경우에 구축
유형2	많다	적다	도서 데이터베이스	<ul style="list-style-type: none">• 사용자 수 보통• 검색은 많지만 데이터에 대한 변경은 적음
유형3	적다	많다	비행기 예약 데이터베이스	<ul style="list-style-type: none">• 예약 변경/취소 등 데이터 변경은 많지만 검색은 적음, 검색은 변경을 위하여 먼저 시도됨• 실시간 검색 및 변경이 중요함
유형4	많다	많다	증권 데이터베이스	<ul style="list-style-type: none">• 사용자 수 많음• 검색도 많고 거래로 인한 변경도 많음

구축이
쉬움



구축이
어려움

3. 데이터베이스의 개념

① 통합된 데이터(integrated data)

데이터를 통합하는 개념으로, 각자 사용하던 데이터의 중복을 최소화하여 중복으로 인한 데이터 불일치 현상을 제거

② 저장된 데이터(stored data)

문서로 보관된 데이터가 아니라 디스크, 테이프 같은 컴퓨터 저장장치에 저장된 데이터를 의미

③ 운영 데이터(operational data)

조직의 목적을 위해 사용되는 데이터를 의미한다. 즉 업무를 위한 검색을 할 목적으로 저장된 데이터

④ 공용 데이터(shared data)

한 사람 또는 한 업무를 위해 사용되는 데이터가 아니라 공동으로 사용되는 데이터를 의미

3. 데이터베이스의 특징

① 실시간 접근성(real time accessibility)

데이터베이스는 실시간으로 서비스된다. 사용자가 데이터를 요청하면 몇 시간이나 몇 일 뒤에 결과를 전송하는 것이 아니라 수 초 내에 결과를 서비스한다.

② 지속적인 변화(continuous change)

데이터베이스에 저장된 내용은 어느 한 순간의 상태를 나타내지만, 데이터 값은 시간에 따라 항상 바뀐다. 데이터베이스는 삽입(insert), 삭제(delete), 수정(update) 등의 작업을 통하여 바뀐 데이터 값을 저장한다.

③ 동시 공유(concurrent sharing)

데이터베이스는 서로 다른 업무 또는 여러 사용자에게 동시에 공유된다. 동시(concurrent)는 병행이라고도 하며, 데이터베이스에 접근하는 프로그램이 여러 개 있다는 의미다.

④ 내용에 따른 참조(reference by content)

데이터베이스에 저장된 데이터는 데이터의 물리적인 위치가 아니라 데이터 값에 따라 참조된다.

4. 데이터베이스 시스템의 구성

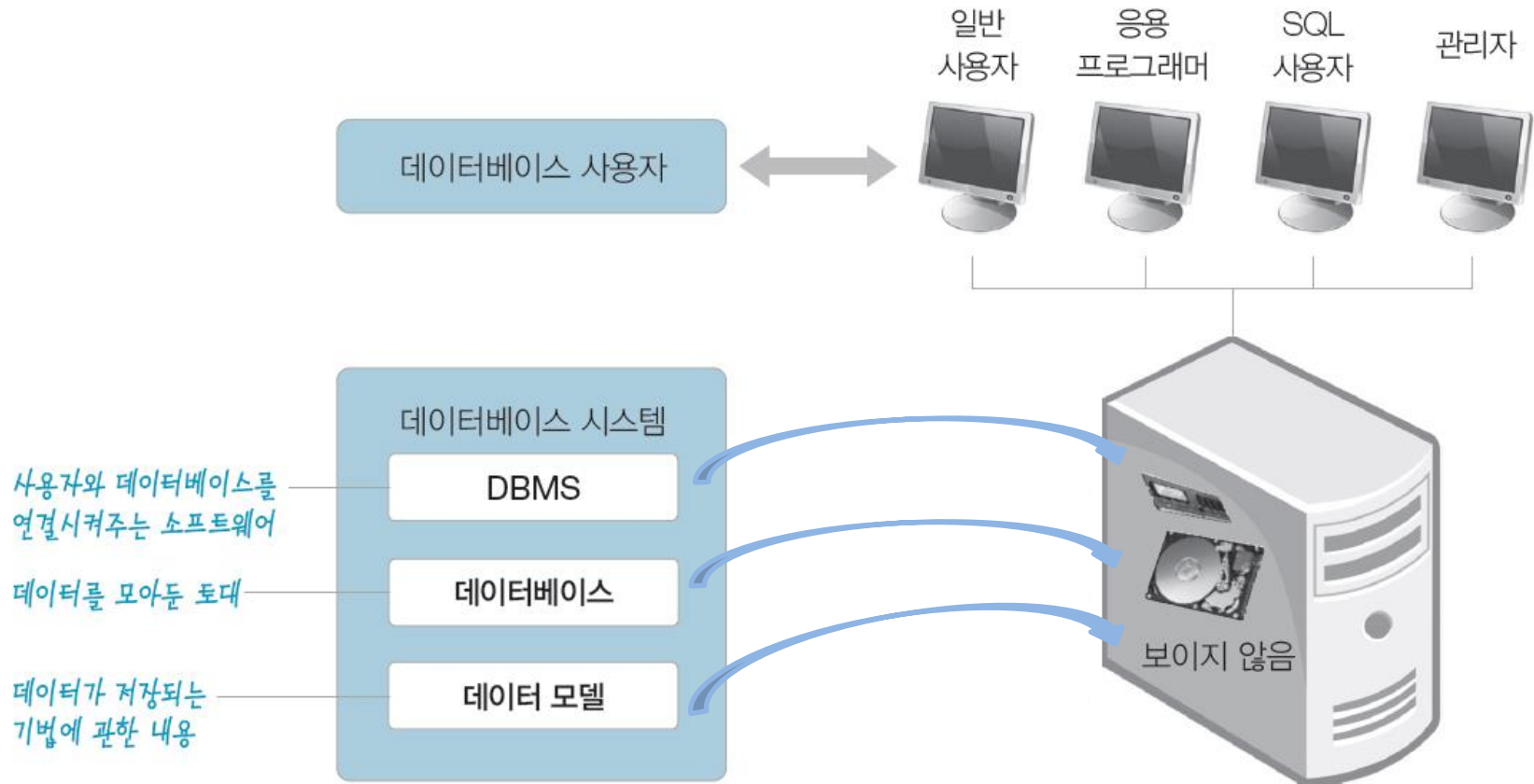


그림 1-5 데이터베이스 시스템의 구성 요소와 물리적인 위치

3. 파일 시스템과 DBMS의 비교

표 1-6 DBMS의 장점

구분	파일 시스템	DBMS
데이터 중복	데이터를 파일 단위로 저장하므로 중복 가능	DBMS를 이용하여 데이터를 공유하기 때문에 중복 가능성 낮음
데이터 일관성	데이터의 중복 저장으로 일관성이 결여됨	중복 제거로 데이터의 일관성이 유지됨
데이터 독립성	데이터 정의와 프로그램의 독립성 유지 불가능	데이터 정의와 프로그램의 독립성 유지 가능
관리 기능	보통	데이터 복구, 보안, 동시성 제어, 데이터 관리 기능 등을 수행
프로그램 개발 생산성	나쁨	짧은 시간에 큰 프로그램을 개발할 수 있음
기타 장점	보통	데이터 무결성 유지, 데이터 표준 준수 용이

04. 데이터베이스 시스템의 구성

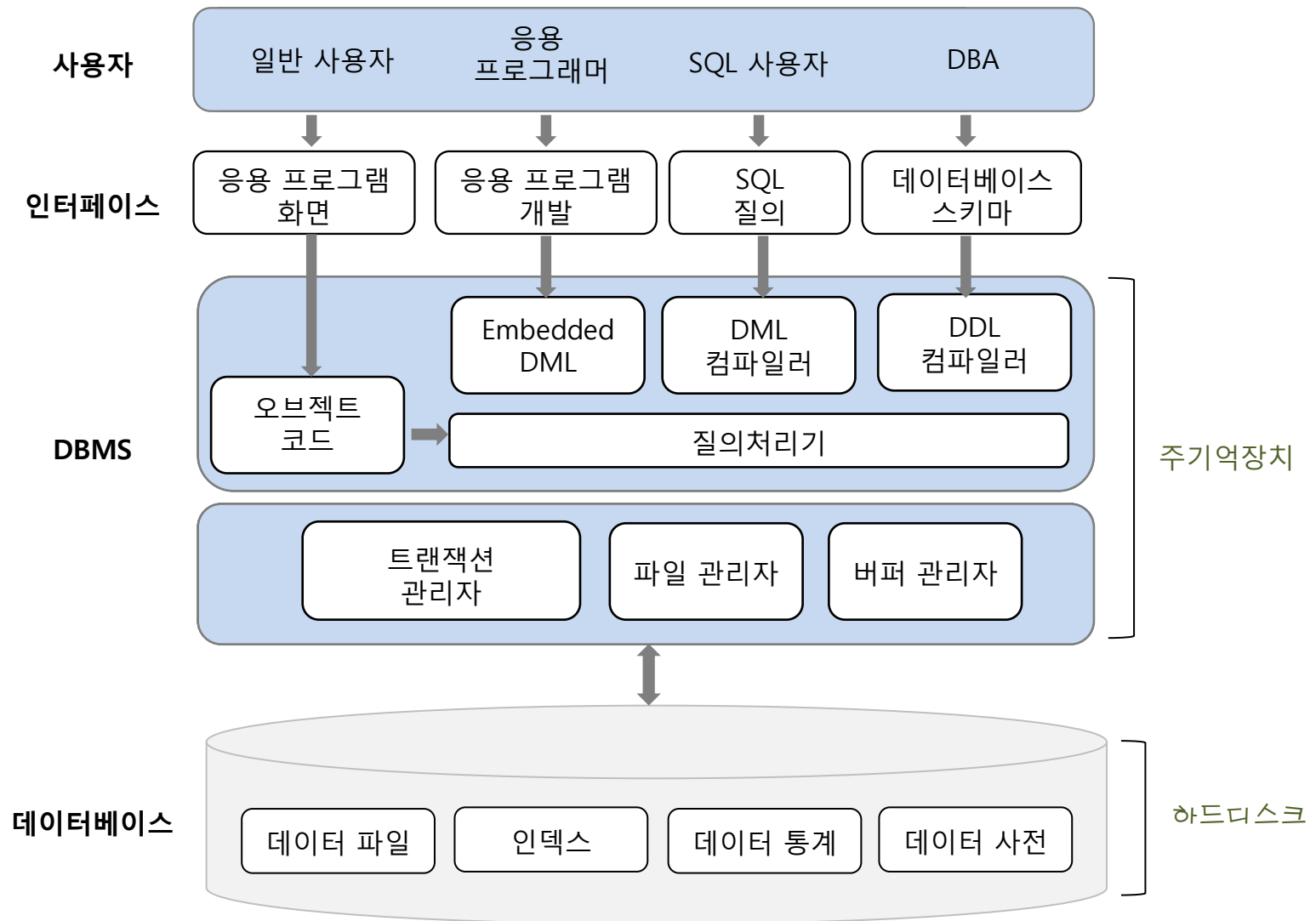


그림 1-22 데이터베이스 시스템의 구성

3. DBMS

표 1-8 DBMS의 기능

데이터 정의(Definition)	데이터의 구조를 정의하고 데이터 구조에 대한 삭제 및 변경 기능을 수행함
데이터 조작(manipulation)	데이터를 조작하는 소프트웨어(응용 프로그램)가 요청하는 데이터의 삽입, 수정, 삭제 작업을 지원함
데이터 추출(Retrieval)	사용자가 조회하는 데이터 혹은 응용 프로그램의 데이터를 추출함
데이터 제어(Control)	데이터베이스 사용자를 생성하고 모니터링하며 접근을 제어함. 백업과 회복, 동시성 제어 등의 기능을 지원함

4. 데이터 모델

① 포인터 사용 : 계층 데이터 모델, 네트워크 데이터 모델

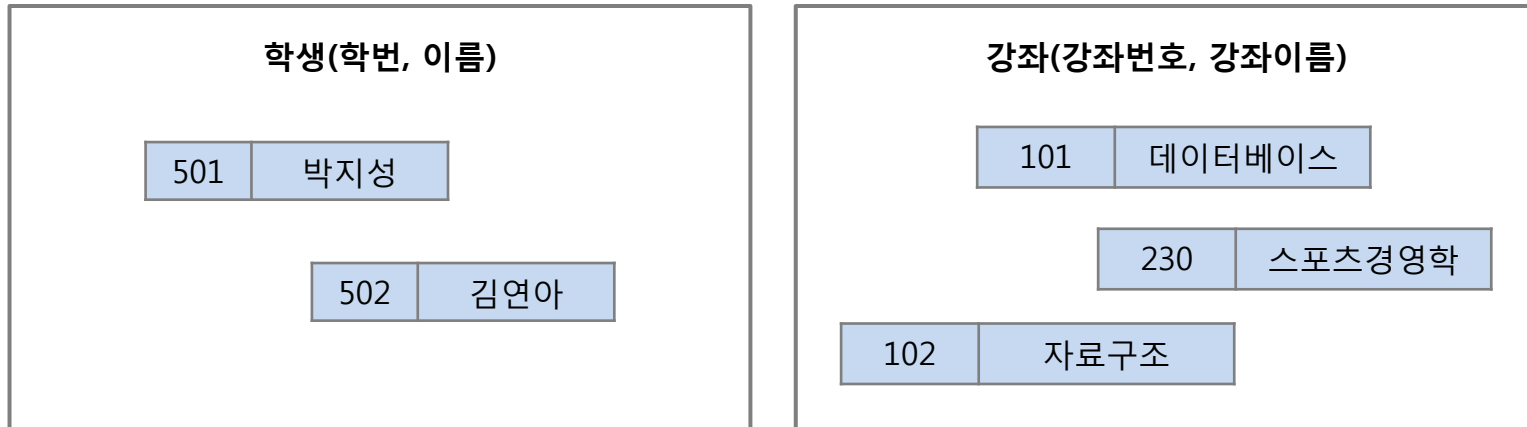


그림 1-23 관계 표현을 위한 예시

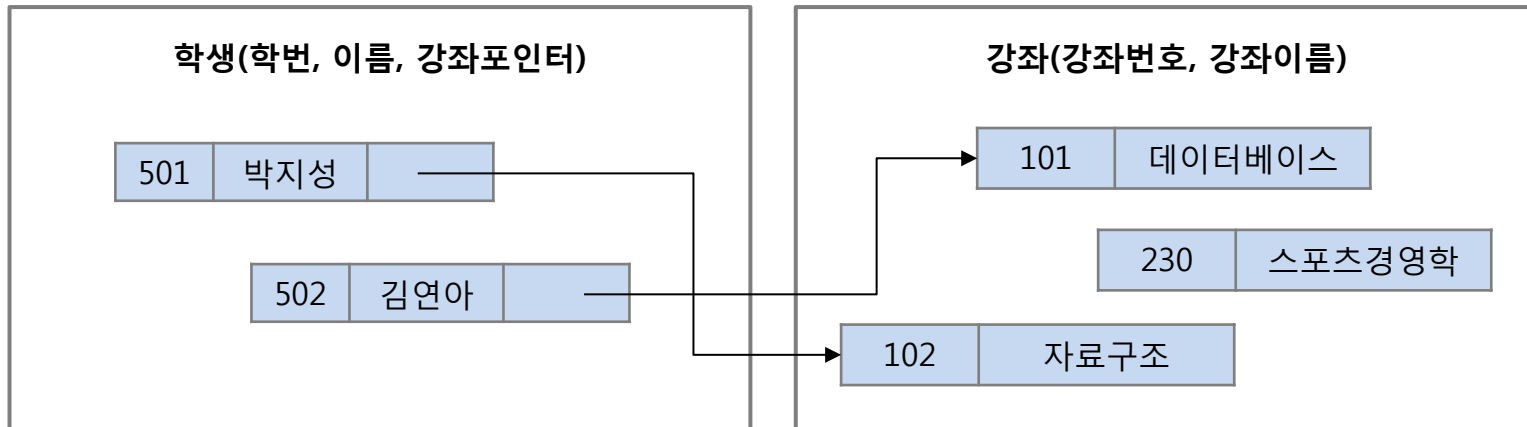


그림 1-24 포인터를 사용하여 관계 표현

4. 데이터 모델

② 속성 값 사용 : 관계 데이터 모델

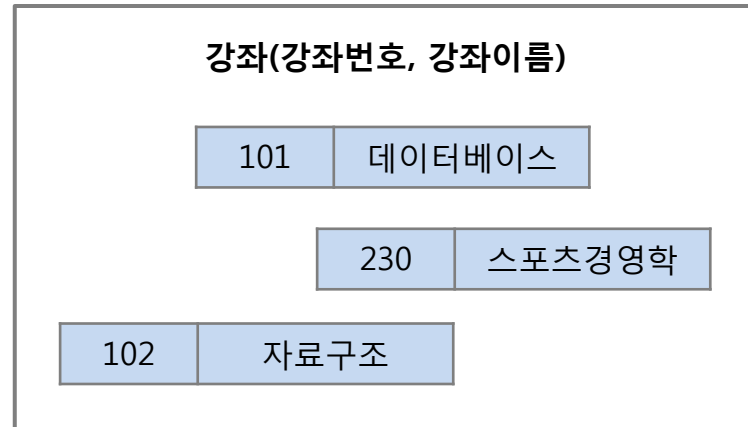
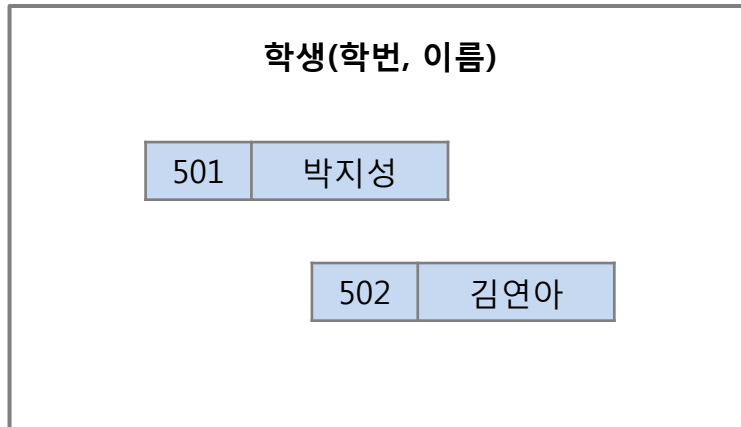


그림 1-23 관계 표현을 위한 예시

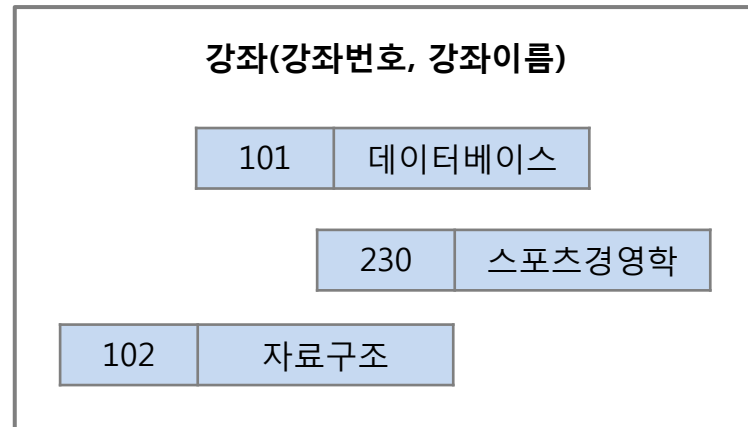
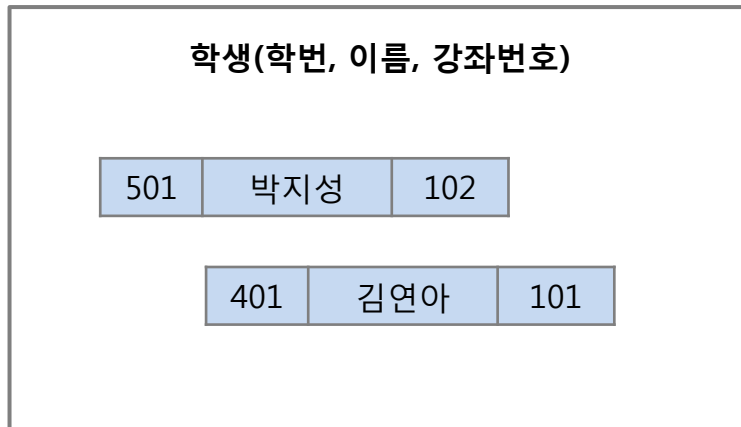


그림 1-25 속성 값을 사용하여 관계 표현

4. 데이터 모델

③ 객체식별자 사용 : 객체 데이터 모델

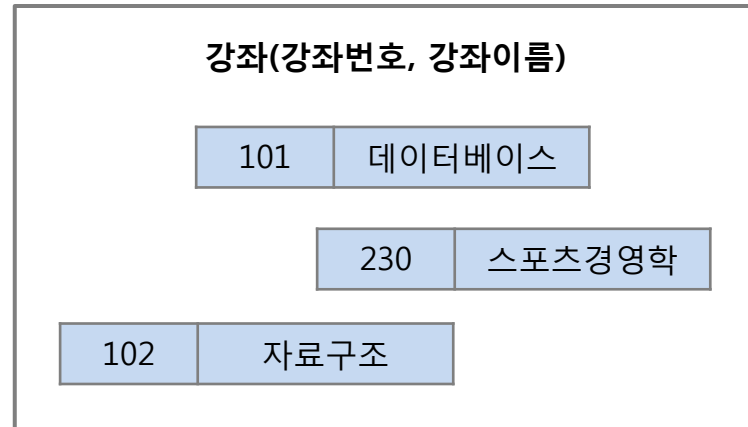
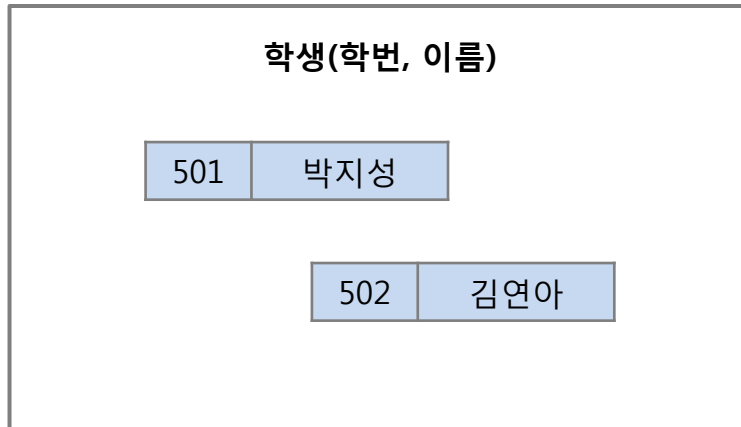


그림 1-23 관계 표현을 위한 예시

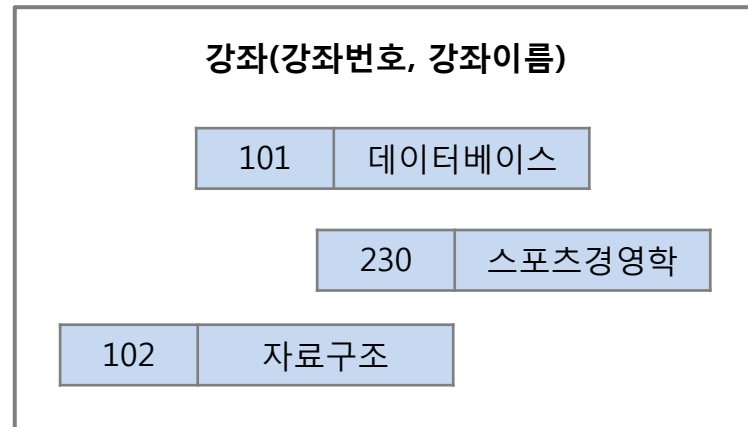
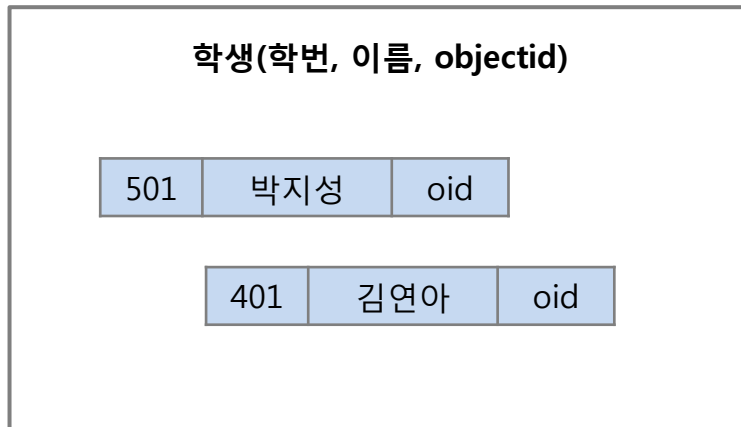
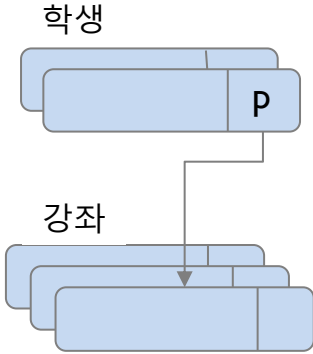
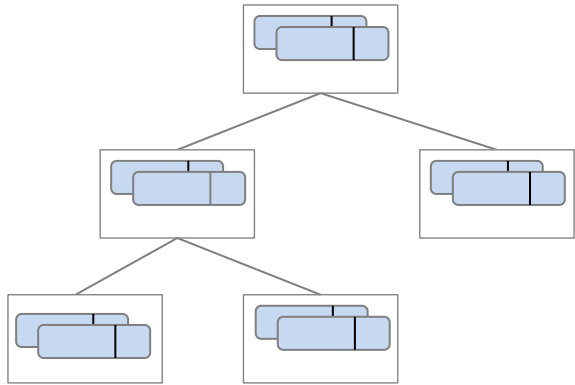
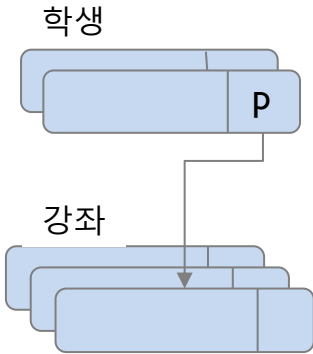
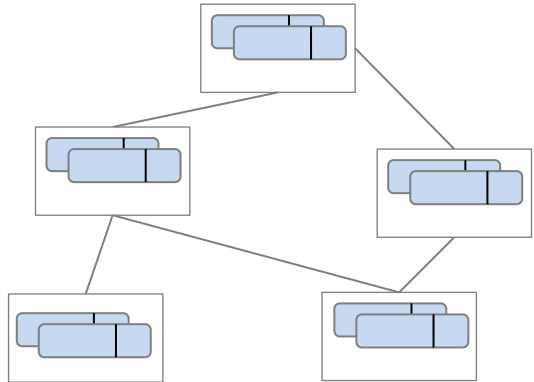


그림 1-26 객체식별자를 사용하여 관계 표현

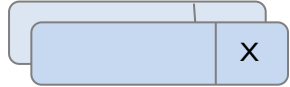
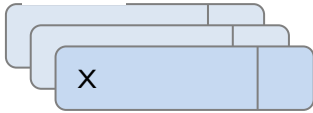
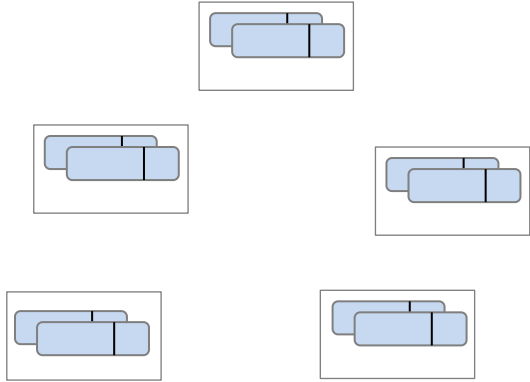
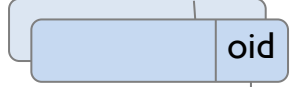

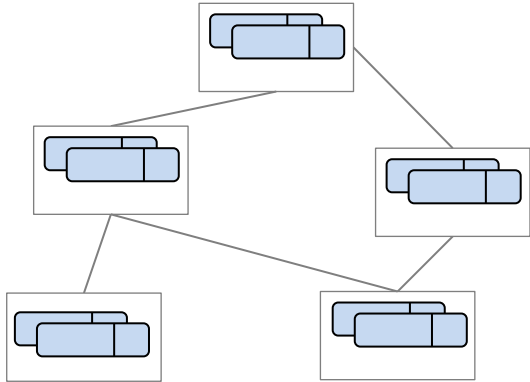
4. 데이터 모델

표 1-9 데이터 모델과 각 모델에서 관계의 표현 방법

데이터 모델	관계의 표현	데이터 구성
계층 데이터 모델 (포인터 사용)	<p>학생</p>  <p>강좌</p>	
네트워크 데이터 모델 (포인터 사용)	<p>학생</p>  <p>강좌</p>	

4. 데이터 모델

표 1-9 데이터 모델과 각 모델에서 관계의 표현 방법

데이터 모델	관계의 표현	데이터 구성
관계 데이터 모델 (속성 값 사용)	<p>학생</p>  <p>강좌</p> 	
객체 데이터 모델 (객체식별자 사용)	<p>학생</p>  <p>강좌</p>  <p>객체 번호 oid</p>	

5.1 3단계 데이터베이스 구조

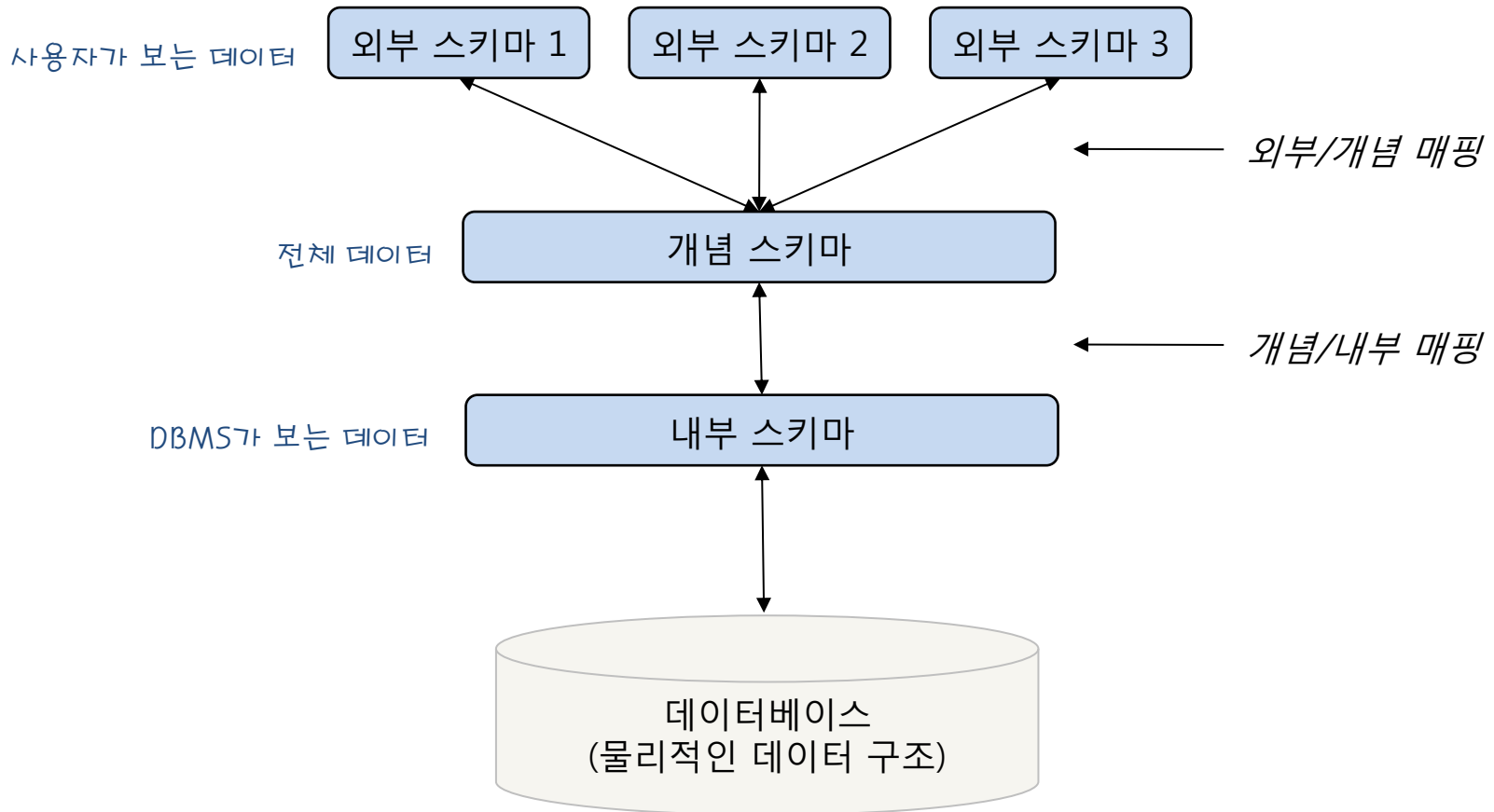


그림 1-27 ANSI의 3단계 데이터베이스 구조

5.1 3단계 데이터베이스 구조

■ 외부 스키마

- 일반 사용자나 응용 프로그래머가 접근하는 계층으로 전체 데이터베이스 중에서 하나의 논리적인 부분을 의미
- 여러 개의 외부 스키마(external schema)가 있을 수 있음
- 서브 스키마(sub schema)라고도 하며, 뷰(view)의 개념임

■ 개념 스키마

- 전체 데이터베이스의 정의를 의미
- 통합 조직별로 하나만 존재하며 DBA가 관리함
- 하나의 데이터베이스에는 하나의 개념 스키마(conceptual schema)가 있음

■ 내부 스키마

- 물리적 저장 장치에 데이터베이스가 실제로 저장되는 방법의 표현
- 내부 스키마(intenal schema)는 하나
- 인덱스, 데이터 레코드의 배치 방법, 데이터 압축 등에 관한 사항이 포함됨

5.1 3단계 데이터베이스 구조

■ 외부/개념 매핑

- 사용자의 외부 스키마와 개념 스키마 간의 매핑(사상)
- 외부 스키마의 데이터가 개념 스키마의 어느 부분에 해당되는지 대응시킴

■ 개념/내부 매핑

- 개념 스키마의 데이터가 내부 스키마의 물리적 장치 어디에 어떤 방법으로 저장되는지 대응시킴

5.1 3단계 데이터베이스 구조

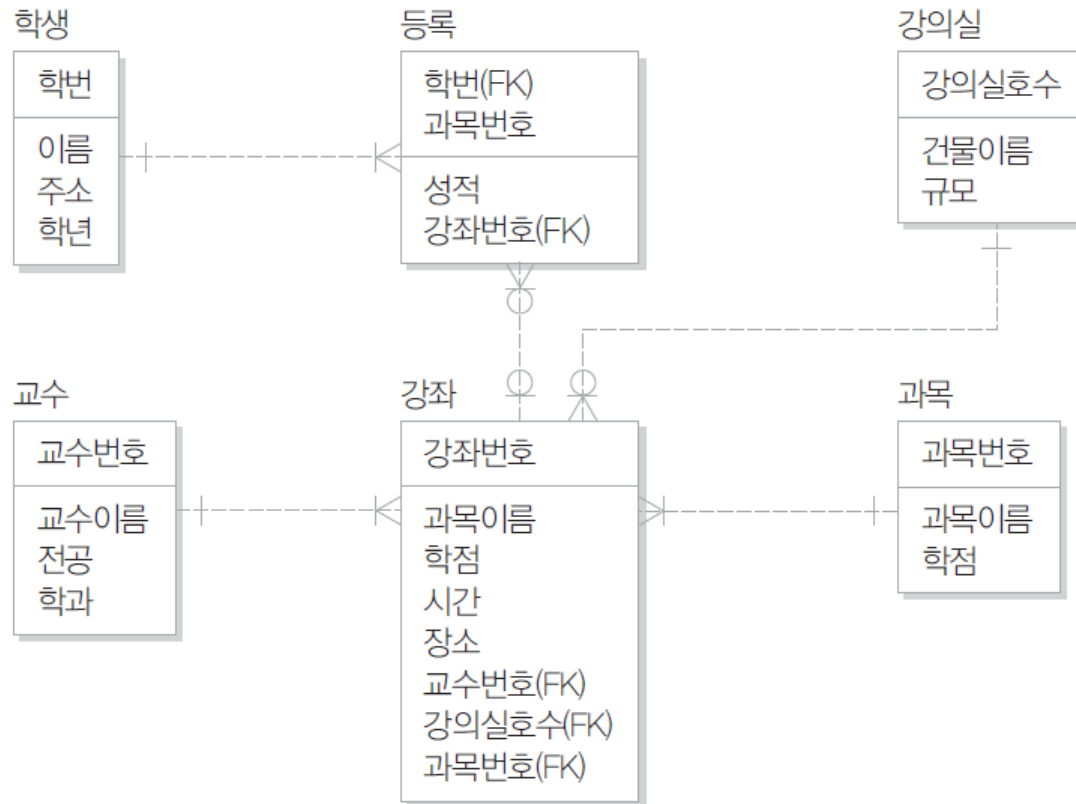


그림 1-28 수강신청 데이터베이스의 개념 스키마

5.1 3단계 데이터베이스 구조

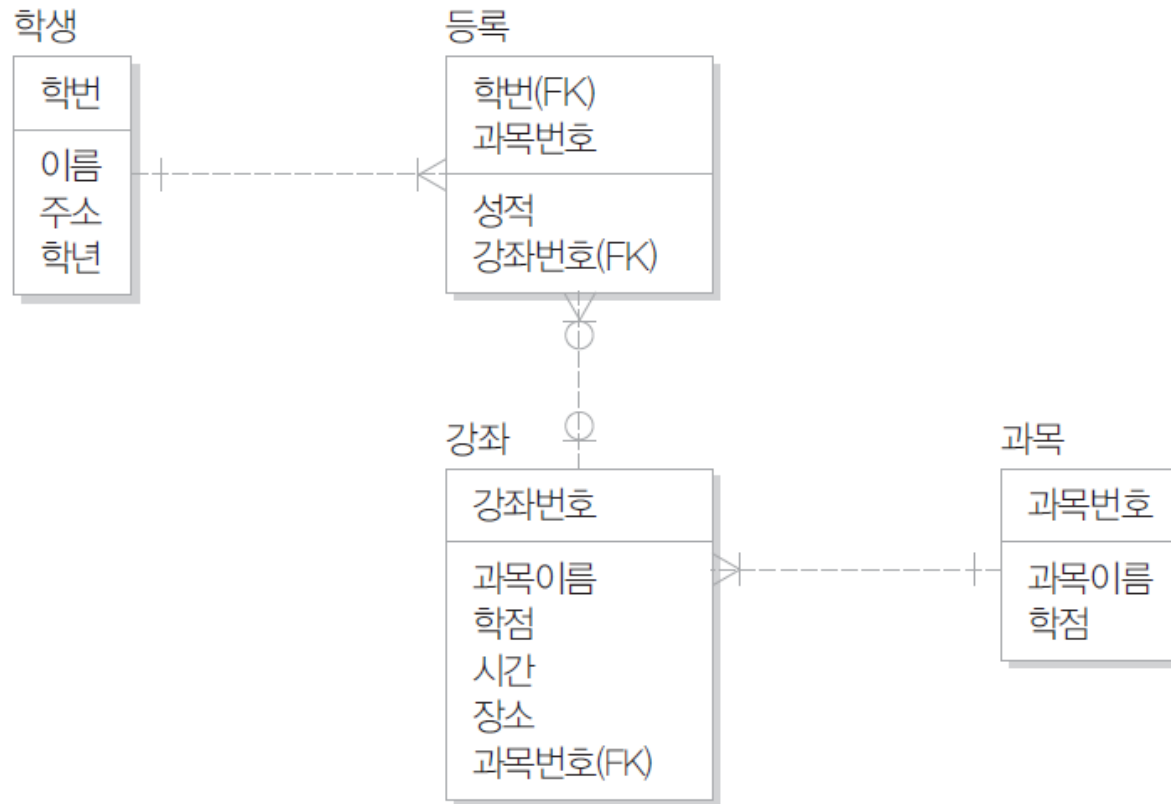


그림 1-29 수강등록 담당 부서에서 필요한 데이터베이스(외부 스키마1)

5.1 3단계 데이터베이스 구조

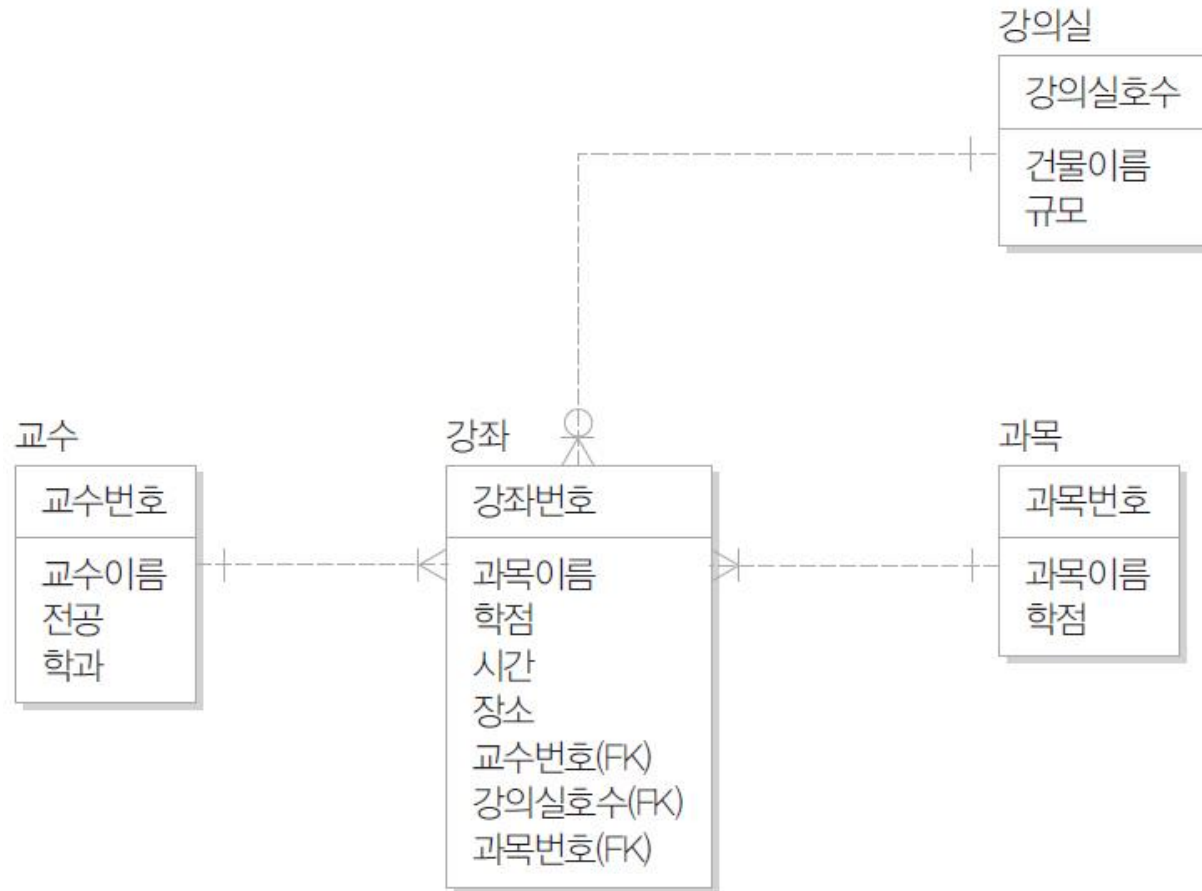


그림 1-30 시간표 담당 부서에서 필요한 데이터베이스(외부 스키마2)

5.1 3단계 데이터베이스 구조

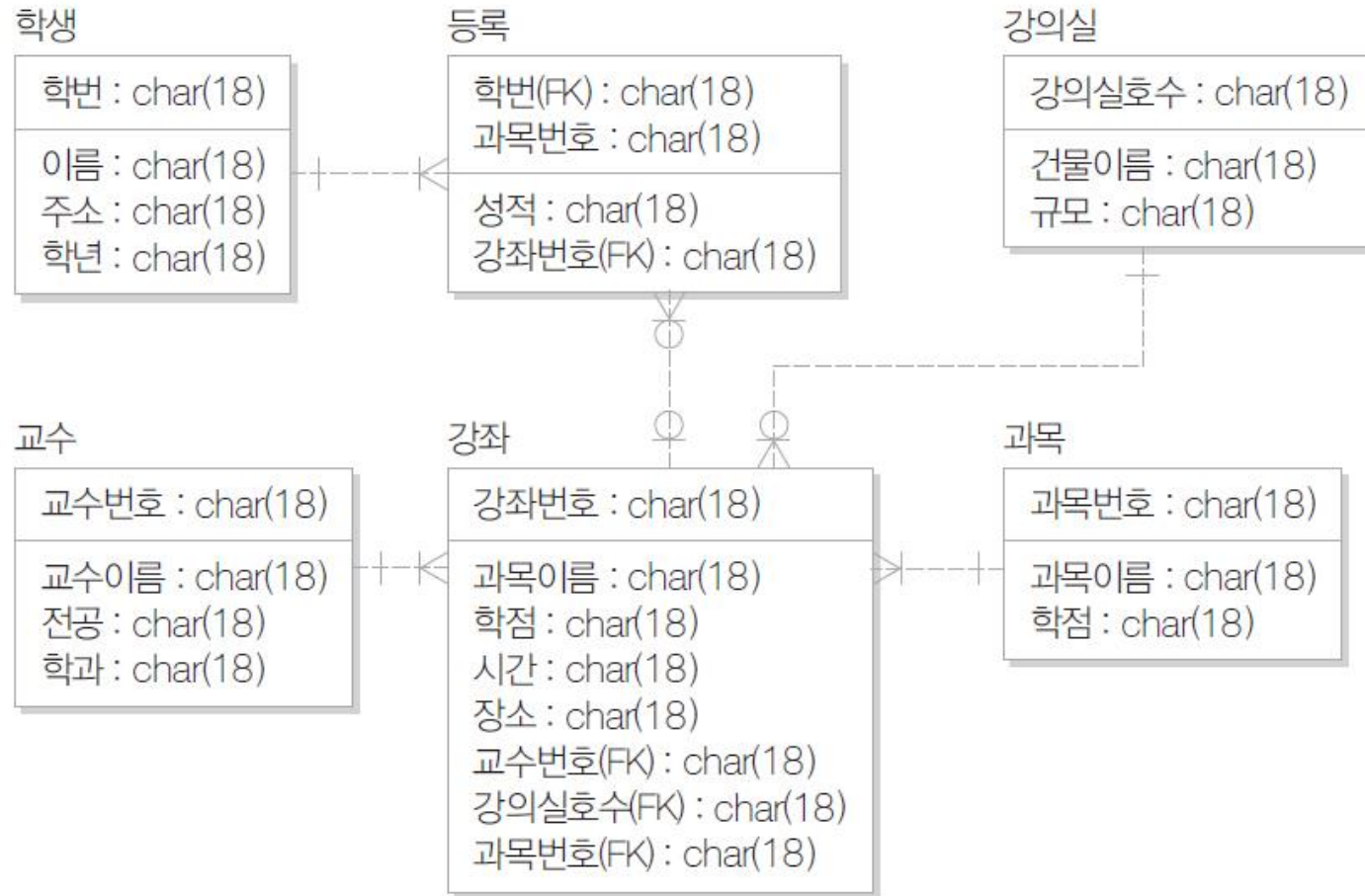


그림 1-31 수강신청 데이터베이스의 내부 스키마

5.1 3단계 데이터베이스 구조

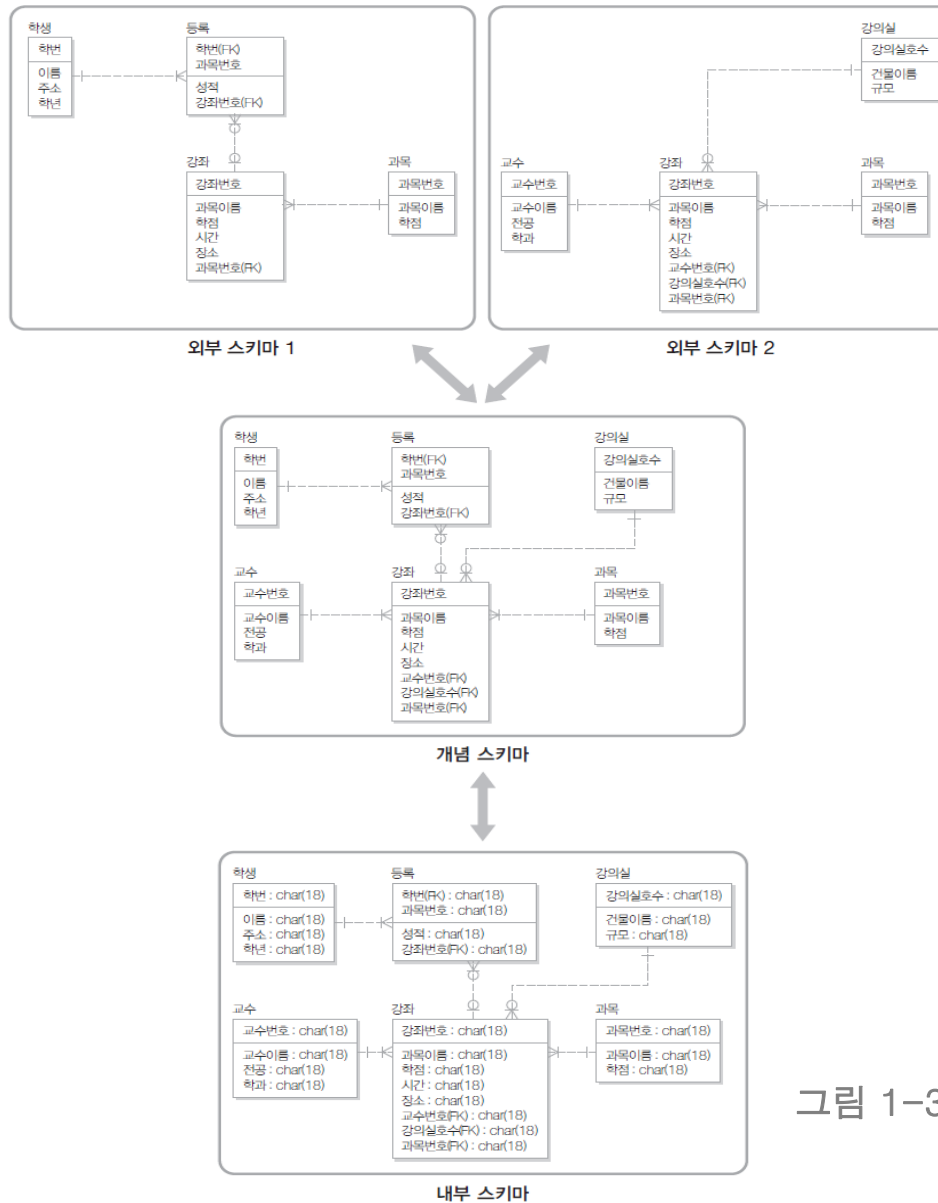


그림 1-32 수강신청 데이터베이스의 3단계 구조

5.2 데이터 독립성

■ 논리적 데이터 독립성(logical data independence)

- 외부 단계(외부 스키마)와 개념 단계(개념 스키마) 사이의 독립성
- 개념 스키마가 변경되어도 외부 스키마에는 영향을 미치지 않도록 지원
- 논리적 구조가 변경되어도 응용 프로그램에는 영향이 없도록 하는 개념
- 개념 스키마의 테이블을 생성하거나 변경하여도 외부 스키마가 직접 다루는 테이블이 아니면 영향이 없음

■ 물리적 데이터 독립성(physical data independence)

- 개념 단계(개념 스키마)와 내부 단계(내부 스키마) 사이의 독립성
- 저장장치 구조 변경과 같이 내부 스키마가 변경되어도 개념 스키마에 영향을 미치지 않도록 지원
- 성능 개선을 위하여 물리적 저장 장치를 재구성할 경우 개념 스키마나 응용 프로그램 같은 외부 스키마에 영향이 없음
- 물리적 독립성은 논리적 독립성보다 구현하기 쉬움

1.1 마당서점의 데이터

bookid	bookname	publisher	price
1	축구의 역사	굿스포츠	7000
2	축구아는 여자	나무수	13000
3	축구의 이해	대한미디어	22000
4	골프 바이블	대한미디어	35000
5	피겨 교본	굿스포츠	8000
6	역도 단계별기술	굿스포츠	6000
7	야구의 추억	이상미디어	20000
8	야구를 부탁해	이상미디어	13000
9	올림픽 이야기	삼성당	7500
10	Olympic Champions	Pearson	13000

그림 3-3 Book 테이블

orderid	custid	bookid	saleprice	orderdate
1	1	1	6000	2014-07-01
2	1	3	21000	2014-07-03
3	2	5	8000	2014-07-03
4	3	6	6000	2014-07-04
5	4	7	20000	2014-07-05
6	1	2	12000	2014-07-07
7	4	8	13000	2014-07-07
8	3	10	12000	2014-07-08
9	2	10	7000	2014-07-09
10	3	8	13000	2014-07-10

그림 3-4 Orders 테이블

custid	name	address	phone
1	박지성	영국 맨체스타	000-5000-0001
2	김연아	대한민국 서울	000-6000-0001
3	장미란	대한민국 강원도	000-7000-0001
4	추신수	미국 클리블랜드	000-8000-0001
5	박세리	대한민국 대전	NULL

그림 3-4 Customer 테이블

1.1 마당서점의 데이터

Book(bookid, bookname, publisher, price)

BOOKID	BOOKNAME	PUBLISHER	PRICE
1	축구역사	굿스포츠	7000

Orders(orderid, custid, bookid, saleprice, orderdate)

ORDERID	CUSTID	BOOKID	SALEPRICE	ORDERDATE
1	1	1	6000	14/07/01

Customer(custid, name, address, phone)

CUSTID	NAME	ADDRESS	PHONE
1	박지성	영국 맨체스터	000-5000-0001

그림 3-6 마당서점의 데이터 구성도