

데이터베이스 4장

DB 설계

3가지 설계

- 개념적 설계

개념적 설계 단계의 결과물은 논리적 설계 단계의 입력으로 사용된다.

개념적 설계에서는 높은 추상화 수준의 데이터 모델을 기반으로 정형적인 언어로 데이터구조를 명시한다.

여기에선 엔티티 타입, 관계타입, 애트리뷰트들을 식별하고, 애트리뷰트들의 도메인을 결정하고, 후보 키와 기본 키 애트리뷰트들을 결정한다.

- 논리적 설계

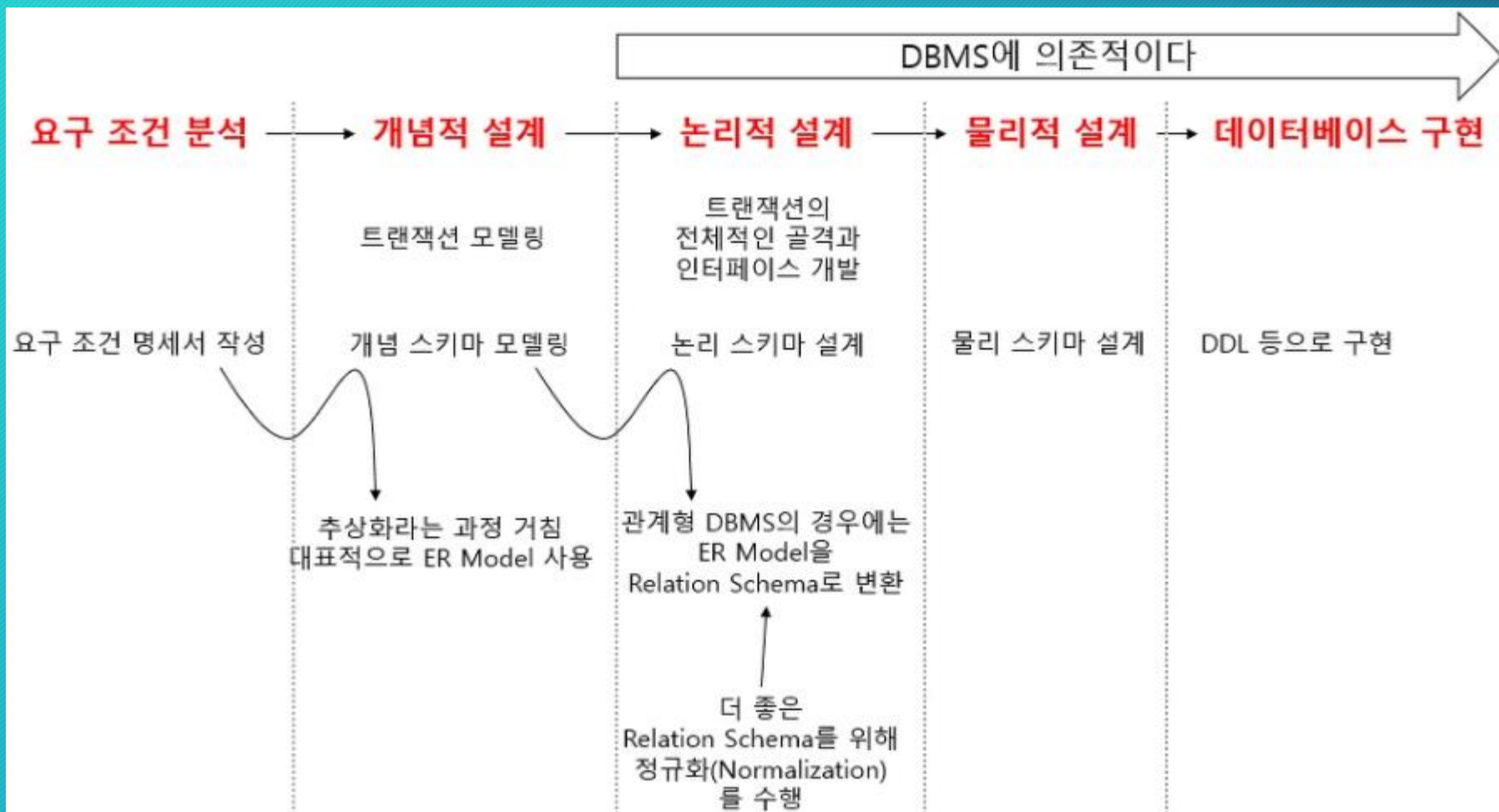
논리적 설계에선 DB관리를 위해 선택한 DBMS의 데이터 모델을 사용하여 논리적 스키마를 생성한다.

- 물리적 설계

성능상의 주요기준은 질의와 갱신이 얼마나 오래걸리는가?

1초당 얼마나 많은 트랜잭션들이 처리될수 있는가? 에 초점이 맞춰져있다.

설계 단계 과정



ER 모델, 엔티티

- ER모델

ER모델은 실세계를 엔티티, 애트리뷰트, 엔티티들 간의 관계로 표현한다.

- 엔티티

하나의 엔티티는 사람, 장소, 사물, 사건 등과 같이 독립적으로 존재하면서 고유하게 식별이 가능한 실세계의 객체이다.

- 엔티티타입

강한 엔티티 타입(정규 엔티티 타입)은 독자적으로 존재하며 엔티티 타입내에서 자신의 키 애트리뷰트를 사용하여 고유하게 엔티티들을 식별할 수 있는 엔티티 타입을 의미한다.

약한 엔티티 타입은 키를 형성하기에 충분한 애트리뷰트들을 갖지 못한 엔티티 타입이다.

애트리뷰트(1)

하나의 엔티티는 연관된 애트리뷰트들의 집합으로 설명된다.

한 애트리뷰트의 도메인은 그 애트리뷰트가 가질 수 있는 모든 가능한 값들의 집합을 의미

- 단순 애트리뷰트

단순 애트리뷰트는 더 이상 다른 애트리뷰트로 나눌 수 없는 애트리뷰트이다.

- 복합 애트리뷰트

복합 애트리뷰트는 두 개 이상의 애트리뷰트로 이루어진 애트리뷰트이다.

- 단일 값 애트리뷰트

단일 값 애트리뷰트는 각 엔티티마다 정확하게 하나의 값을 갖는 애트리뷰트이다.

- 다치 애트리뷰트

다치 애트리뷰트는 각 엔티티마다 여러 개의 값을 가질 수 있는 애트리뷰트이다.

애트리뷰트(2)

- 저장된 애트리뷰트

저장된 애트리뷰트는 다른 애트리뷰트와 독립적으로 존재하는 애트리뷰트(유도와 반대!)

- 유도된 애트리뷰트

다른 애트리뷰트의 값으로부터 얻어진 애트리뷰트, 예를 들어 나이 애트리뷰트는 주민등록번호 애트리뷰트로부터 유도될 수 있는 애트리뷰트이다.

ER모델에서의 표현 그림 (1)

1. 엔티티

- 엔티티는 **실세계의 객체**(차, 자전거 등)와 **추상적인 것**(학과, 과목 등)을 표현한다.

2. 강한 엔티티 타입(정규 엔티티 타입)

- 독자적으로 존재할 수 있는 엔티티 타입이다.
- 직사각형으로 표현한다.



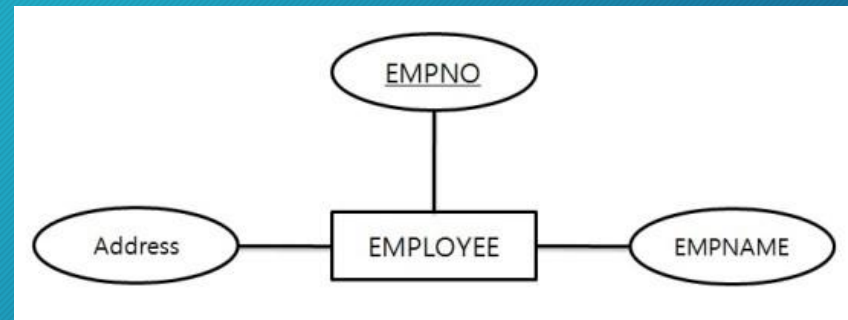
ER모델에서의 표현 그림 (2)

3. 애트리뷰트

- 하나의 엔티티는 연관성이 있는 애트리뷰트의 집합이다.
- 애트리뷰트는 여러가지 종류가 있다.

(1) 단순 애트리뷰트

- 이 애트리뷰트는 더 이상 다른 애트리뷰트로 나눌 수 없다.
- 단순히 하나의 값을 가지는 애트리뷰트이고, 실선 타원으로 나타낸다.



그림에서 **EMPNO**, **Address**, **EMPNAME**은 단순 애트리뷰트이다.

그런데 **Address**는 단순 애트리뷰트가 아닐수도 있다. 주소라는 값은 시, 구, 동, 번지라는 값으로 나뉘질 수 있기 때문이다.

단순 애트리뷰트는 단순히 생각해서 단순한 값을 가지는 애트리뷰트라 생각하자.

ER모델에서의 표현 그림 (3)

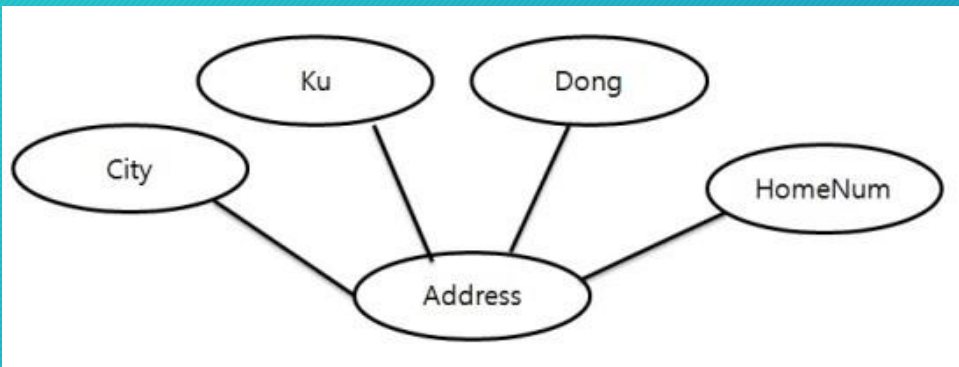
(2) 복합 애트리뷰트

위의 그림에서 이야기한 Address가 복합 애트리뷰트로 생각되어질 수 있다.

Address 는 또 다른 여러가지의 애트리뷰트로 나누어 질 수 있기 때문이다.

예를들어 주소가 서울시 영등포구 여의도동 100번지 라고 한다면 Address의 값은 시, 구, 동, 번지 4가지의 애트리뷰트로 **분할이** 가능해진다.

복합 애트리뷰트의 표현은 다음과 같다.



ER모델에서의 표현 그림 (4)

(3) 단일 값 애트리뷰트

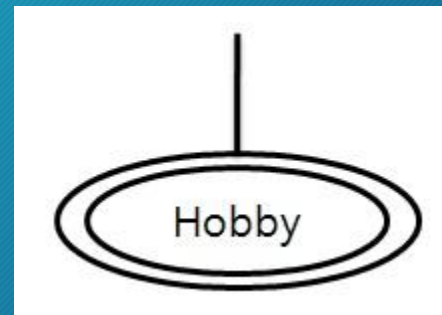
정확하게 하나의 값을 갖는 애트리뷰트이다. 예를 들어 아까 그림에서 기본 키로 지정되어 있는 EMPNO가 단일 값 애트리뷰트이다. 기본 키 값은 중복될 수 없으므로, 유일한 값을 가지기 때문이다. (원자 값을 가지는 애트리뷰트)

(4) 다치 애트리뷰트

각 엔티티마다 여러 개의 값을 가질 수 있는 값이다.

사원의 취미의 경우를 생각해 본다면, 사람마다 취미가 여러개인 사람이 있을수가 있기 때문에, 취미는 다치 애트리뷰트로 분류될 수 있다.

다치 애트리뷰트의 표현은 이중선 타원으로 표현한다.



ER모델에서의 표현 그림 (5)

(5) 저장된 애트리뷰트

다른 애트리뷰트와 독립적으로 존재하는 애트리뷰트이다.

표현 방법은 단순 애트리뷰트와 동일하게 취급된다.

예를 들면 사원의 급여에 대한 정보는 따로 저장되어 있다.

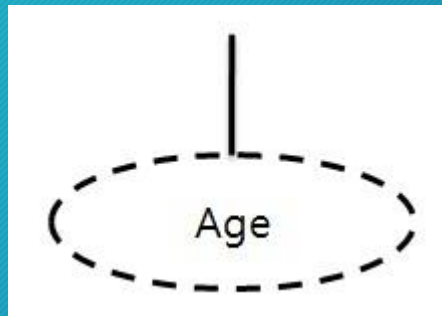
(6) 유도된 애트리뷰트

다른 애트리뷰트의 값으로 부터 얻어진 애트리뷰트이다.

예를 들어 주민등록번호를 통해 이 사람이 몇 살인가를 판단할 수 있다.

그래서 나이라는 정보의 애트리뷰트는 유도된 애트리뷰트의 상태로 있을 수 있다.

표기법은 점선 타원으로 표시한다.



ER모델에서의 관계 표현 (1)

1. 약한 엔티티 타입

엔티티 타입 내의 엔티티들이 자체적으로 갖고 있는 애트리뷰트들의 값에 의해 고유하게 식별이 불가능한 엔티티 타입이다. 다시말해 키를 보유하지 못한 엔티티를 말한다.

약한 엔티티 타입은 다른 엔티티 타입으로 부터 키 애트리뷰트를 가져온다.

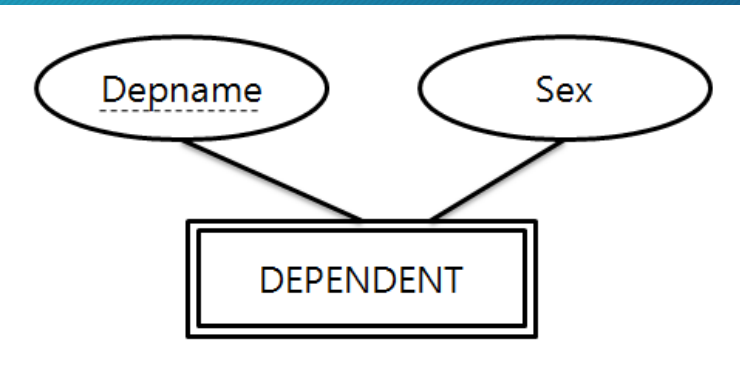
약한 엔티티에 키 애트리뷰트를 제공하는 엔티티는 소유 엔티티 타입, 식별 엔티티 타입이라고 한다.

표기는 이중선 직사각형으로 표시하고, 부분 키는 점선 밑줄을 그어 나타낸다.

부분 키는 전체적으로 구분할 수 없으므로 부모 키 + 자식 키 형태로 해서 구분한다.

옆의 그림에서 DEPENDENT 엔티티는 약한 엔티티 타입이다.

Depname 애트리뷰트는 부분 키 애트리뷰트이다.



ER모델에서의 관계 표현 (2)

2. 관계와 관계타입

어떤 한 엔티티가 다른 엔티티와 연관을 가지는 것을 관계라고 한다.

이러한 관계는 관계타입으로 표시할 수 있다.

관계 타입은 다이아몬드로 표기하고, 서로 연관시키는 엔티티 타입들을 관계 타입에 실선으로 연결한다.

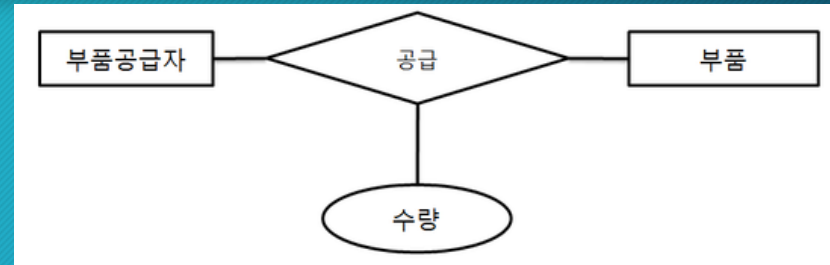


ER모델에서의 관계 표현 (3)

(1) 관계의 애트리뷰트

관계타입은 관계의 특징을 기술하는 애트리뷰트를 가질 수 있다.

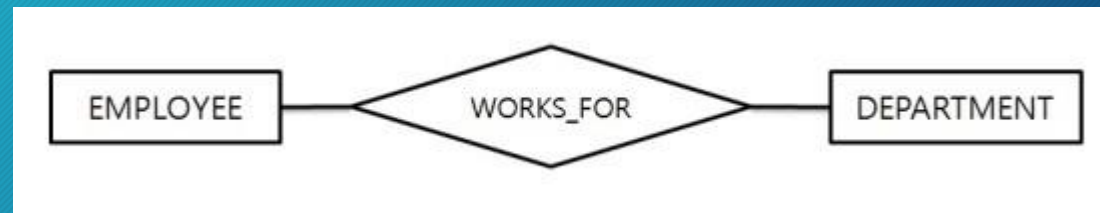
예를 들어서 설명하면, 부품공급자와 부품 관계 사이에는 공급이라는 관계 타입이 존재할 수 있으며, 부품을 공급하는 과정에서 부품을 몇 개를 공급하는가의 정보가 필요할 수 있기 때문에, 수량 애트리뷰트를 관계의 애트리뷰트로 지정한 다.



(2) 차수

차수는 관계로 연결된 엔티티 타입들의 개수를 의미
1진, 2진, 3진 관계는 차수가 1, 2, 3 이다.

옆 사진은 2진관계이다.



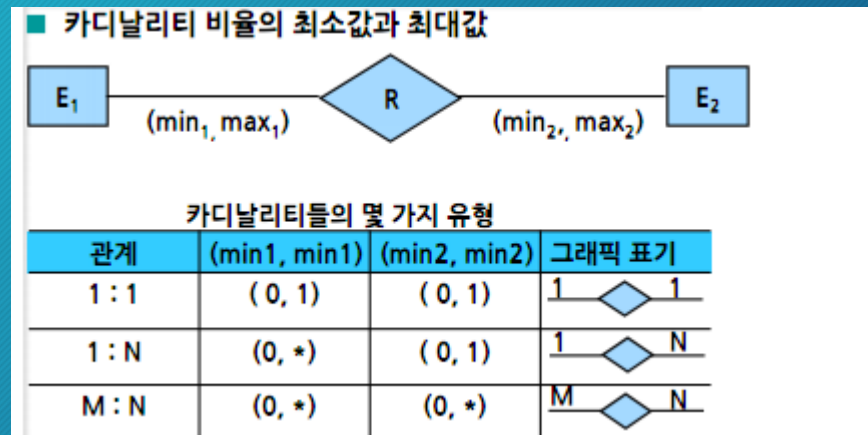
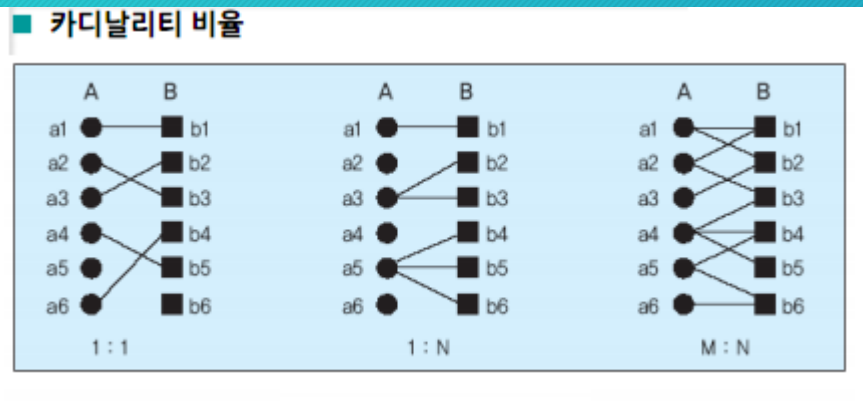
ER모델에서의 관계 표현 (4) -1

3. 카디널리티 비율

몇 개의 엔티티가 다른 엔티티 타입의 몇 개의 엔티티와 연관되는가를 나타낸다.

모델링하는 실세계의 시나리오로 부터 유도된다. 관계는 1:1(일대일), 1:N(일대다), M:N(다대다)로 구분된다.

카디널리티 비율은 최소값과 최대값을 사용하여 좀 더 명확하게 나타낼 수 있다.



오른쪽 그림에서 min은 부분참여(0) 전체참여를(1) 뜻하는것.

ER모델에서의 관계 표현 (4) -2

4. 전체 참여와 부분참여

전체참여 : 어떤 관계에 엔티티 타입 E1의 모든 엔티티들이 관계타입 R에 의해 어떤 엔티티 타입 E2의 어떤 엔티티와 연관된다.

부분참여 : 어떤 관계에 엔티티 타입 E1의 일부 엔티티들이 관계타입 R에 의해 어떤 엔티티 타입 E2의 어떤 엔티티와 연관된다.

★ 약한 엔티티 타입은 항상 관계에 전체 참여를 한다.

※ 전체참여는 간선을 두 줄로 표현하고, 부분참여는 간선을 한 줄로 표현한다.

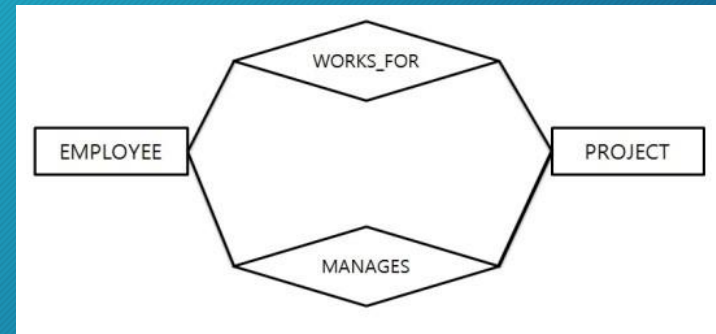


설명을 좀더 하자면, 부서(DEPT)를 관리(MANAGES)하는 사원(EMP)가 있는데 사원은 전부가 부서를 관리 하는게 아니니 부분참여이고 부서는 관리 당하는 부서가 없을테니 (부장이 없는 부서는 없죠.) 전체참여이다.

ER모델에서의 관계 표현 (5)

5. 다중 관계

엔티티 사이에 관계 타입이 여러개 존재하고, 엔티티들이 그 관계에 모두 참여하는 경우, 다중관계로 표현할 수 있다.

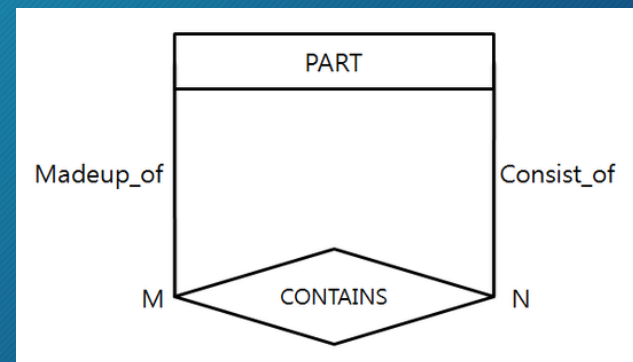


6. 순환적 관계

하나의 엔티티 타입이 동일한 관계 타입에 두 번 이상 참여하는 것이다.

(1) 역할

그림에서 Madeup_of, Consist_of 처럼 이 관계가 어떤건지 명확하게 설명하기위해 사용하는것.


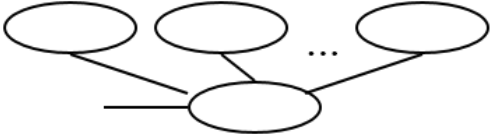
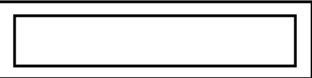

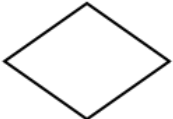
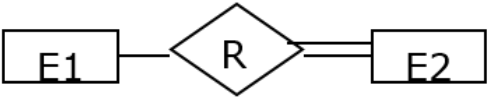
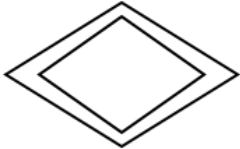
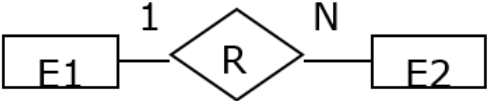

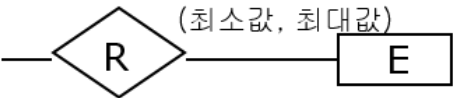




ER모델 작성절차

1. Entity 도출	명사형, 집합형을 이루고 대표성을 갖는 Entity 도출 예) 학생, 학과, 교수
2. Relationship 도출	Entity와 entity 간의 관계 도출 예) 학생은 학과에 소속된다.
3. Attribute 도출	Entity의 상세 속성 도출 예) 학생 = {학번, 이름, 학과, ...}
4. Cardinality 도출	두 Entity 간의 관계에서 참여자의 수를 비율로 표시 예) 1:1, 1:N, N:M
5. 식별자 도출	기본키, 외래키, 수퍼키, 후보키, 대체키 등에 대한 정의
6. 정규화	1차, 2차, 3차, BCNF, 4차, 5차 정규화
7. 통합/분할	Entity type의 성격에 따라 통합, 분할 수행
8. Data 모델 검증	Entity type 속성 관계 등에 대한 적합성 검증

※참조. DB배움터 282쪽 (ER스키마 단계적 완성)

ER모델 표현 정리

Symbol	Meaning	Symbol	Meaning
	엔티티 타입		복합 애트리뷰트
	약한 엔티티 타입		유도된 애트리뷰트
	관계 타입		E1이 R에 부분참여 E2가 R에 전체참여
	식별 관계 타입		R에서 E1:E2의 카디날리티 비율이 1:N
	애트리뷰트		R에서 E의 참여에 대한 구조적 제약조건 (최소값, 최대값)
	키 애트리뷰트		
	다치 애트리뷰트		