



데이터베이스 시스템 6장

6.1 설계 과정의 개요

6.1.1 설계 단계

6.1.2 설계 대안

6.2 개체-관계 모델

6.2.1 개체 집합

6.2.2 관계 집합

6.3 복합 속성

6.4 대응 카디널리티

6.5 주 키

6.5.1 개체 집합

6.5.2 관계 집합

6.5.3 약한 개체 집합

6.6 개체 집합에서 중복 속성 제거하기

6.7 E-R 다이어그램을 관계형 스키마로 변환하기

6.7.1 강한 개체 집합의 표현

6.7.2 복합 속성을 지닌 강한 개체 집합의 표현

6.7.3 약한 개체 집합의 표현

6.7.4 관계 집합의 표현

6.7.5 스키마의 중복성

6.7.6 스키마의 결합

6.8 E-R의 확장된 특성

6.8.1 구체화

6.8.2 일반화

6.8.3 속성 상속

6.8.4 구체화/일반화에 대한 제약 조건

6.8.5 통합화

6.8.6 릴레이션 스키마로의 변환

6.8.6.1 일반화의 표현

6.8.6.2 통합화의 표현

6.9 개체-관계 설계 쟁점

6.9.1 E-R 다이어그램 작성 시 흔한 실수

6.9.2 개체 집합 대 속성 사용의 비교

6.9.3 개체 집합 대 관계 집합 사용의 비교

6.9.4 이진 대 비이진 관계 집합

6.10 데이터 모델링을 위한 대안적인 표기법

6.10.2 통합 모델링 언어

6.11.1 기능 요구 사항

6.11.3 스키마의 진화

6.1 설계 과정의 개요

6.1.1 설계 단계

- 데이터베이스 시스템 6장

6.1.2 설계 대안

스키마 설계시 다음 두 가지 중요한 위험성을 피해야 함

1. 중복성

- a. 여러 중복된 데이터들이 있다면 어떻게 업데이트할 것인가?
- b. 한 장소에만 존재하도록 하는 것이 이상적, but 현실은?

2. 불완전성

- a. 모델링하기 어렵거나 불가능하게 하는 경우
- b. 어떤 릴레이션이 존재해야만 다룰 수 있는 경우 → ex) 과목이 있어야만 분반을 다룰 수 있는 경우

6.2 개체-관계 모델

- 데이터베이스의 전체 논리 구조를 나타내는 조직의 스키마를 명시함으로써 데이터베이스 설계를 쉽게 사도록 개발됨
 - 실세계 조직의 의미와 상호작용을 나타내는데 유용
- E-R 데이터 모델은 개체 집합, 관계 집합, 속성을 가짐

6.2.1 개체 집합

- 개체(Entity)는 실세계에서 다른 모든 객체와 구별되는 유무형의 사물
- 개체는 속성(Attribute)의 집합을 가짐
 - 속성은 개체 집합 구성원들이 소유하는 설명 특성
 - 속성 각각은 값을 가짐
- 속성 중 일부 집합은 개체를 고유하게 식별하기도 함
- 개체 집합은(Entity set) 같은 속성을 공유하는 같은 유형의 개체의 집합
 - 확장(Extension)이라는 용어로 개체 집합이 포함하는 개체의 실제 모음을 지칭
- 다른 개체 집합에 중첩될 수도 있음
- E-R 다이어그램에서 사각형으로 표현됨
 - 상단: 개체 집합의 이름
 - 하단: 개체 집합의 모든 속성의 이름

6.2.2 관계 집합

- 관계(Relationship)는 여러 개체들 사이의 연관성
- 관계 집합(Relationship set)은 같은 유형의 관계들의 집합
- 관계 인스턴스(Relationship instance)는 모델링하는 실세계 조직 내의 명명된 개체들 사이에 존재하는 연관성
- E-R 다이어그램에서 다이아몬드로 표시됨
 - 여러 다른 개체 집합을 선으로 연결
 - 두 개 이상의 개체 집합 간의 연관도 표시 가능
- 관계 집합은 2개 이상의 개체 집합 사이의 수학적 릴레이션
 - $\{(e_1, e_2, \dots, e_n) | e_1 \in E_1, e_2 \in E_2, \dots, e_n \in E_n\}$
 - 개체 집합 사이의 연관을 참여한다고 함 = E_1, E_2, \dots, E_n 은 관계 집합 R 에 참여
- 관계에서 개체가 행하는 기능을 역할(Role)이라고 함
 - 보통 일반적으로 개체 집합들이 구별되기에 명시되지는 않음
 - 재귀 관계거나 명확하게 할 필요가 있을 때는 유용
 - 개체 집합과 관계를 연결하는 선에 표기
- 설명 속성
 - 관계 집합의 속성은 점선으로 관계와 연결하고 나뉘지 않는 사각형으로 표시
 - 관계 집합은 여러 설명 속성을 가질 수 있음
- 관계 집합은 한 위치에만 표시될 수 있으나 개체 집합은 둘 이상의 위치에서 반복할 수 없음
- 동일한 개체 집합을 포함하는 하나 이상의 관계 집합이 존재할 수 있음
- 관계 집합은 관계 인스턴스의 집합, 즉 관계에서 하나의 개체는 관계가 있는 다른 하나의 개체와만 관계를 가질 수 있음
 - 하지만 다중값 속성이 있긴함
- n 진 관계 집합 $\Rightarrow n$ 개의 개체 집합이 관련된 관계 집합, 차수는 n

6.3 복합 속성

- 각 속성이 가질 수 있는 값의 집합 \Rightarrow 속성의 도메인, 값 집합
- 단순 속성과 복합 속성: 더 이상 나뉘지 않는 속성 vs 나뉠 수 있는 속성

- 복합 속성을 사용해 모델링을 더 명확하고 연관된 속성을 묶을 수 있고 필요한 속성들만 참고 가능
- 단일값 속성과 다중값 속성
 - 속성이 특정 개체에 대해서 하나의 값만 가지는 경우 ⇒ 단일값 속성
 - 속성이 특정 개체에 대해서 값의 집합을 가지는 경우 ⇒ 다중값 속성
- 유도 속성
 - 다른 관련된 속성이나 개체의 값으로 유도되는 속성
 - 유도 속성의 값은 저장하지 않고 필요할 때 계산하면 됨
- 속성은 개체가 그 속성에 대한 값을 가지지 않을 때 널(null) 값을 가짐
 - 해당 사항 없음, 그 개체에 대한 값이 존재하지 않음을 나타냄

6.4 대응 카디널리티

- 대응 카디널리티란 관계 집합을 통해 다른 개체와 관련될 수 있는 개체의 수를 나타냄
- 일대일, 일대다, 다대일, 다대다
- 화살표나 직선을 이용하여 관계에 대한 카디널리티 제약 표시
- 개체 집합 E의 모든 개체가 관계 집합 R의 적어도 하나의 관계에 참여해야 하는 경우, E의 R에 대한 참여를 전체 참여, 참여하지 않을 수도 있으면 부분 참여라고 함
 - 이중선으로 전체 참여표현
- 관계에 참여하는 각 개체가 나타나는 횟수에 대해서 더 복잡한 제약을 나타낼 수 있음
 - 최솟값(0), 최댓값(*)
 - 최솟값부터 최댓값까지 관계를 가질 수 있음

6.5 주 키

6.5.1 개체 집합

- 개체를 표현하는 속성의 값은 그 개체를 유일하게 구별할 수 있음
 - 키란 각각의 개체를 서로 구별해 주기에 충분한 속성의 집합

6.5.2 관계 집합

- 관계 집합 R을 개체 집합 E1, E2, ..., En이 연관된 관계 집합
- 관계 집합은 관계 인스턴스의 집합이며 각 인스턴스는 참여하는 개체의 의해 고유하게 식별될 수 있음
 - primary-key(Ei) : 개체 집합 Ei의 주 키를 이루는 속성의 집합
 - $\Rightarrow \text{primary-key}(E1) \cup \text{primary-key}(E2) \cup \dots \cup \text{primary-key}(En)$
- 이진 관계 집합의 경우 주 키 선택은 관계 집합의 대응 카디널리티를 고려
 - 다대다 \rightarrow 관계에 속하는 개체 집합의 주 키들의 합집합이기에 둘 다 씀
 - 일대다, 다대일 \rightarrow "다" 쪽의 주 키가 최소 수퍼키이자 주 키
 - 일대일 \rightarrow 한쪽의 주 키가 최소 수퍼 키를 형성
- 비이진 관계의 경우
 - 카디널리티 제약 조건이 없으면 \rightarrow 합집합을 이용한 수퍼 키만 유일한 후보 키이자 주 키
 - 카디널리티 제약 조건이 있으면 \rightarrow 복잡함, 화살표 개수에 따라 다름, 함수 종속으로 해결할 수 있음
- 관계 집합 R에 대한 주 키는 관계 집합 R로부터 들어오는 화살표가 없는 참여 개체 집합 Ei의 주 키의 합집합

6.5.3 약한 개체 집합

- 약한 개체 집합이란 식별 개체 집합이라 하는 다른 개체 집합에 의존하여 존재하는 개체 집합
- 주 키를 약한 개체와 연결하는 대신, 식별 개체의 주 키와 구별자 속성을 함께 사용하여 약한 개체 식별
 - 그 반대는 강한 개체 집합
- 식별 개체와 연관되어야하며 식별 개체 집합에 존재 종속적임
 - 약한 개체 집합(전체 참여) \leftrightarrow 식별 개체 집합, 다대일 관계
 - 식별 관계 집합은 어떠한 설명 속성도 가져서는 안됨
- E-R 다이어그램에서 약한 개체 집합은 이중 사각형으로 표시 구별자는 실선이 아닌 점선, 해당 관계 집합은 이중 다이아몬드로 표시
- 약한 개체 집합은 식별 관계 이외의 관계에도 참여 가능

6.6 개체 집합에서 중복 속성 제거하기

- 어떤 개체 집합도 관계 집합에 의해 중복이 되게 하는 속성을 가지고 있지 않는 것이 좋은 개체-관계 설계임

6.7 E-R 다이어그램을 관계형 스키마로 변환하기

6.7.1 강한 개체 집합의 표현

- E라는 이름의 스키마와 n개의 서로 다른 속성으로 구성 가능
- 주 키는 해당 개체 집합의 주 키를 사용할 수 있음

6.7.2 복합 속성을 지닌 강한 개체 집합의 표현

- 복합 속성을 지니면 분리함
 - name → first_name, middle_name, last_name
- 다중값 속성은 다르게 취급됨
 - 유도 속성 → 명시적으로 표현 X
 - 다중값 속성과 해당 속성이 포함된 테이블의 주 키로 새로운 테이블 구성 가능, but 다중값 속성 테이블이 주 키만 남는 경우 새로운 테이블 생성을 하지 않음

6.7.3 약한 개체 집합의 표현

- 약한 개체 집합의 속성들 + 약한 개체 집합이 의존하는 강한 개체 집합의 주 키 속성들
 - 외래 키 제약

6.7.4 관계 집합의 표현

- 관계 집합에 참여하는 각 개체 집합들의 주 키의 합집합 + 설명 속성
 - 외래 키 제약
- 다대일 관계의 경우 어느 하나의 주 키로만 구성

6.7.5 스키마의 중복성

- 약한 개체 집합과 강한 개체 집합을 연결하는 관계 집합은 약한 개체 집합의 스키마와 중복될 수도 있음

- 중복 제거 필요

6.7.6 스키마의 결합

- 개체 집합 A에서 개체 집합 B로 가는 다대일 관계 집합 AB의 경우 A가 관계 집합 AB에 전체 참여한다면 A와 AB를 결합하여 양쪽 스키마 속성의 합집합으로 구성 가능
- 일대일 관계의 경우 관계 집합의 스키마는 어느 한쪽 개체 집합의 스키마와 결합 가능

6.8 E-R의 확장된 특성

6.8.1 구체화

- 어떤 개체 집합 내에서 하위 그룹을 지정하는 과정을 구체화
 - 복수의 구체화된 개체 집합에 속할 수 있는지에 따라서 중첩 있는 구체화, 중첩 없는 구체화

6.8.2 일반화

- 상위 개체 집합과 하위 개체 집합 사이에 존재하는 포함 관계
- 구체화의 반대
- 실질적으로 구체화와 일반화는 E-R 다이어그램에서 구별하지 않음

6.8.3 속성 상속

- 하위 개체 집합의 모든 계층에 걸쳐 적용
- 관계 집합의 참여도 상속 받음

6.8.4 구체화/일반화에 대한 제약 조건

- 구체화에 대한 중첩 유무와 완전성 제약 조건을 고려해야 함(서로 의존 X)
 - 완전성 제약 조건 : 상위 개체 집합에 있는 개체가 일반화나 구체화에서 반드시 하나 이상의 하위 개체 집합에 속해야 하는지 여부
 - 전체 구체화 or 전체 일반화 → 모든 상위 개체는 어떤 하위 개체 집합에 속함
 - 부분 구체화 or 부분 일반화 → 어떤 상위 개체는 어느 하위 개체 집합에도 속하지 않을 수 있음

6.8.5 통합화

- 관계들 간의 관계를 표현하지 못하는 것이 E-R 모델의 제약점
- 통합화(aggregation)을 이용해서 관계를 상위 개체로 취급할 수 있음

6.8.6 릴레이션 스키마로의 변환

6.8.6.1 일반화의 표현

1. 상위 개체 집합을 위한 스키마를 만들고 각각의 하위 개체 집합을 위해 해당 개체 집합의 속성과 상위 개체 집합의 주 키인 속성을 포함하는 스키마를 생성함
2. 일반화가 전체 중첩 없음 이면 상위 개체 집합을 위한 스키마를 만들지 않고 하위 개체 집합으로 자신의 속성을 나타내는 속성과 상위 개체 집합의 속성을 포함하는 스키마를 생성함
 - a. 단점은 다른 릴레이션이 외래 키로 참조할 수 있는 단일 릴레이션이 존재하지 않음

6.8.6.2 통합화의 표현

- 다른 개체 집합처럼 처리

6.9 개체-관계 설계 쟁점

6.9.1 E-R 다이어그램 작성 시 흔한 실수

- 개체 집합의 주 키를 다른 개체 집합의 속성으로 사용하면 안됨
- 관계 집합의 속성으로 관련 개체 집합의 주 키 속성을 지정하면 안됨
- 다중값 속성이 필요한 상황에서 단일값 속성을 포함하는 관계를 사용하면 안됨
- E-R 다이어그램이 너무 크면 조각으로 분할 필요

6.9.2 개체 집합 대 속성 사용의 비교

- 속성이 아닌 개체로 다루고 싶으면 다른 개체와 관계를 맺어야 하거나 부가적인 정보가 더 필요한 경우는 개체로 다루면 좋음
- 무엇이 속성이 되고, 무엇이 개체 집합을 구성하는지 답은 없고 모델링 하는 구조와 대상이 되는 속성과 관련된 의미에 달려 있음

6.9.3 개체 집합 대 관계 집합 사용의 비교

- 개체 집합 vs 관계 집합 \Rightarrow 명확하지는 않음
- 어떤 정보를 다른 정보와 연결해서 사용한다면 독립적인 개체로 표현해야할 수도 있음
- 두 개체 사이에 일어나는 동작을 기술하는데 관계 집합을 사용

6.9.4 이진 대 비이진 관계 집합

- 이진이 아니어도 이진으로 표현하면 더 잘 표현될 수도 있음
- 비이진 관계를 관계 집합으로 항상 대체 가능함
 - 대체가 불가능할 수도 있음
 - 개체 A, B, C가 관계 R로 3진 관계
 - $E \rightarrow A$ 가는 다대일 관계 R.A, $E \rightarrow B$ 다대일 관계 R.B, $E \rightarrow C$ 다대일 관계 R.C
 - E는 관계에 전체 참여 필수, E에 특수 식별 속성 생성(개체 구분을 위한)
- E-R 관계 모델을 이진 관계 집합만 포함하도록 제한할 수는 있으나 바람직하지는 않음

6.10 데이터 모델링을 위한 대안적인 표기법

6.10.1 대안적인 E-R 표기법

- 까마귀 발로 대응 카디널리티 표시

6.10.2 통합 모델링 언어

- UML
 - 클래스 다이어그램, 유스 케이스 다이어그램, 활동 다이어그램, 구현 다이어그램

6.11 데이터베이스 설계의 다른 관점

6.11.1 기능 요구 사항

- 어떤 종류의 기능을 지원할 것인가?
- 기능을 제공하기 위한 인터페이스 계획

6.11.2 데이터 흐름, 워크플로

- 프로세스에 포함된 데이터와 작업을 정의

6.11.3 스키마의 진화

- 요구사항 변화에 따른 데이터베이스 스키마 변화
- 요구사항 변화에 따른 최소한의 스키마 변경이 좋은 설계
- 영구적인 것과 변경될 것으로 예상되는 부분들을 고려하는 것이 좋음
- 좋은 설계는 현재의 정책을 고려할 뿐만 아닌, 예상되거나 상당한 발생 가능성이 있는 변경으로 인한 수정의 필요성을 피하거나 최소화
- 인간 지향적인 활동 → 요구사항을 이해하기 위해 도메인 전문가와 상호작용 & 데이터와 관련된 모든 사람들의 필요성과 다양한 선호사항 고려

얘기해보면 좋을것

- 기획안부터 어떻게 설계가 이뤄지는지?
- 데이터베이스 설계 시 중요한 포인트?