**3과목 데이터베이스 구축**

1장 논리 데이터베이스 설계

076 데이터베이스 설계

- 사용자의 요구를 분석하여 데이터베이스의 구조에 맞게 변형한 후 DBMS로 데이터베이스를 구현하여 일반 사용자들이 사용하게 하는 것

무결성 / 일관성 / 회복/ 보안 / 효율성 / 데이터베이스 확장

요구 조건 분석 -> 개념적 설계 -> 논리적 설계 -> 물리적 설계 -> 구현

개념적 설계 – E-R다이어그램 논리적 설계 – 트랜잭션 인터페이스

077 데이터 모델의 개념

- 현실 세계의 정보를 컴퓨터에 표현하기 위해 단순화, 추상화하여 체계적으로 표현

데이터 모델 구성 요소 : 개체, 속성, 관계

데이터 모델 종류 : 개념적 데이터 모델, 논리적 데이터 모델, 물리적 데이터 모델

데이터 모델에 표시할 요소 : 구조, 연산, 제약조건

구조 – 데이터 구조 및 정적 성질 / 연산 – 데이터베이스를 조작하는 기본 도구 / 제약 조건 – 논리적인 제약 조건

078 데이터 모델의 구성 요소 – 개체

- 데이터베이스에 표현하려는 것, 사람이 생각하는 개념이나 정보 단위 같은 현실 세계의 대상체

유형, 무형의 정보 / 영속적 / 유일한 식별자로 구별 가능 / 하나 이상의 관계

업무에서 사용하는 용어로 지정 / 약어 사용 X / 단수 명사 사용 / 개체명 유일 / 의미에 따라 이름 부여

079 데이터 모델의 구성 요소 – 속성

- 데이터 베이스를 구성하는 가장 작은 논리적 단위

데이터 항목 또는 데이터 필드

개체를 구성하는 항목

개체의 특성을 기술

속성의 수 – 디그리 or 차수

기본 속성 / 설계 속성 / 파생 속성

업무에서 사용하는 용어로 지정 / 서술형 지정 X / 약어 사용 X / 개체명은 속성명으로 사용X / 개체에서 유일하게 식별 가능

080 데이터 모델의 구성 요소 – 관계

- 개체와 개체 사이의 논리적인 연결

개체 간의 관계 / 속성 간의 관계

일 대 일(1:1) / 일 대 다(1:n) / 다 대 다(n:m)

종속 관계 / 중복 관계 / 재귀 관계 / 배타 관계

식별 관계 – B 개체의 존재 여부가 A 개체의 존재 여부에 의존적일 때 / 실선

비식별 관계 – B 개체의 존재 여부가 A 개체의 존재 여부와 관계없을 때 / 점선

081 식별자

- 각각의 인스턴스를 유일하게 구분할 수 있는 구분자 모든 개체는 한 개 이상의 식별자를 반드시 가져야 함

대표성 여부 : 주 식별자 / 보조 식별자

스스로 생성 여부 : 내부 식별자 / 외부 식별자

단일 속성 여부 : 단일 식별자 / 복합 식별자

대체 여부 : 원조 식별자 / 대리 식별자

주 식별자 – 개체를 대표하는 유일한 식별자

보조 식별자 – 주 식별자를 대신하여 개체를 식별할 수 있는 속성

유일성 / 최소성 / 불변성 / 존재성 -> 주 식별자의 특성

내부 식별자 – 개체 내에서 스스로 만들어지는 식별자

외부 식별자 – 외부 개체의 식별자를 가져와 사용하는 식별자

단일 식별자 – 주 식별자가 한 가지 속성으로만 구성된 식별자

복합 식별자 – 두 개 이상의 속성으로 구성된 식별자

원조 식별자 – 업무에 의해 만들어지는 가공되지 않은 원래의 식별자 (본질 식별자)

대리 식별자 – 속성이 두 개 이상인 경우 속성들을 하나의 속성으로 묶어 사용하는 식별자 (인조 식별자)

후보 식별자 – 각 인스턴스를 유일하게 식별할 수 있는 속성

082 E-R 모델

- 개념적 데이터 모델의 가장 대표적인 것

개체 - 사각형 / 관계 - 마름모 / 속성 – 타원

083 관계형 데이터 모델

- 2차원적인 표를 이용해서 데이터 상호 관계를 정의하는 DB 구조

대표적 언어 : SQL

1 : 1 / 1: N / N : M

084 관계형 데이터베이스의 구조

개체나 관계를 릴레이션이라는 표로 표현

튜플 – 행, 레코드 기수/대응수

속성 – 열, 데이터 항목/필드, 디그리/차수

도메인 – 같은 타입의 원자값들의 집합

릴레이션의 특징

튜플 사이에는 순서가 없음

릴레이션은 시간에 따라 변함

속성들 가느이 순서는 중요하지 않음

085 관계형 데이터베이스의 제약조건 – 키

키 – 튜플을 찾거나 순서대로 정렬할 때 튜플들을 서로 구분할 수 있는 기준이 되는 속성

후보키 – 튜플을 유일하게 식별하기 위해 사용 / 기본키로 사용할 수 있는 속성들

유일성 / 최소성 만족

기본키 – 후보키 중 주키로 중복죈 값을 가질 수 없음 / 널 값을 가질 수 없음

유일성 / 최소성 만족

대체키 – 후보키가 둘 이상일 때 기본키를 제외한 나머지 후보키

슈퍼키 – 릴레이션을 구성하는 모든 튜플들 중 슈퍼키로 구성된 속성의 잡합과 동일한 값은 나타나지 않음

유일성만 만족

외래키 – 다른 릴레이션의 기본키를 참조하는 속성

086 관계형 데이터베이스의 제약 조건 – 무결성

무결성 – 데이터베이스에 저장된 데이터 값과 그것이 표현하는 현실 세계의 실제값이 일치하는 정확성

개체 무결성 – 기본키를 구성하는 어떤 속성도 널값이나 중복값을 가질 수 없음

도메인 무결성 – 주어진 속성 값이 정의된 도메인에 속한 값이어야 함

참조 무결성 – 외래키 값은 널값이거나 참조 릴레이션의 기본키 값과 동일해야 함

사용자 정의 무결성 – 속성 값들이 사용자가 정의한 제약조건에 만족해야 함

087 관계대수 및 관계해석

관계대수 – 원하는 정보와 정보를 검색하기 위해 어떻게 유도하는가를 기술하는 절차적 언어

순수 관계 연산자 : select, project, join, division

일반 집합 연산자 : union, intersection, difference, Cartesian product

select – 선택 조건을 만족하는 튜플의 부분집합을 구하여 새로운 릴레이션을 만드는 연산

수평 연산, 시그마 사용

project – 속성 리스트에 제시된 속성 값만을 추출하여 새로운 릴레이션을 만드는 연산

수직 연산, 파이 사용

join – 공통 속성을 중심으로 두 개의 릴레이션을 하나로 합쳐 새로운 릴레이션을 만드는 연산

division – r의 속성이 s의 속성값을 모두 가진 튜플에서 s가 가진 속성을 제외한 속성만을 구하는 연산

union – 합집합 intersection – 교집합 difference – 차집합

관계 해석 – 원하는 정보가 무엇이라는 것만 정의하는 비절차적 특성

088 정규화

- 종속성 이론을 이용하여 잘못 설계된 관계형 스키마를 더 작은 속성의 세트로 쪼개어 바람직한 스키마로 만들어 가는 과정

일관성 / 정확성 / 단순성 / 비중복성 / 안전성

삽입 이상 / 삭제 이상 / 갱신 이상

제1정규형 – 릴레이션 속한 모든 도메인이 원자값만으로 되어 있는 정규형

제2정규형 – 기본키가 아닌 모든 속성이 기본키에 대하여 완전 함수적 종속을 만족하는 정규형

제3정규형 – 기본키가 아닌 모든 속성이 기본키에 대해 이행적 종속을 만족하지 않는 정규형

BCNF – 결정자가 모두 후보키인 정규형

제4정규형 – 다치 종속이 성립하는 경우 r이 모든 속성이 a에 함수적 종속 관계를 만족하는 정규형

제5정규형 – 모든 조인 종속이 r의 후보키를 통해서만 성립되는 정규형

089 반정규화

- 정규화된 데이터 모델을 통합, 중복, 분리하는 과정으로 의도적으로 정규화 원칙을 위배하는 행위

테이블 통합 / 테이블 분할 – 수평 분할, 수직 분할 / 중복 테이블 추가 – 집계 테이블 추가, 진행 테이블 추가, 특정 부분만을 포함하는 테이블 추가 / 중복 속성 추가

090 시스템 카탈로그

- 시스템 자체에 관련이 있는 다양한 객체에 관한 정보를 포함하는 시스템 데이터베이스

시스템 카탈로그 = 데이터 사전

시스템 카탈로그에 저장된 정보 -> 메타 데이터

카탈로그의 특징

일반 이용자가 SQL을 이용하여 내용 검색 가능

카탈로그를 갱신하는 것은 허용되지 않음

DBMS가 스스로 생성하고 유지 -> 시스템 자동으로 갱신

문제정리

트랜잭션 인터페이스 설계, 스키마 평가 및 정제 -> 논리적 설계 단계

트랜잭션 모델링 -> 개념적 설계 단계

허가 받은 사용자만이 갱신할 수 있다는 설명 -> 소프트웨어 개발 시 충족시켜야 할 보안 요소인 무결성

도메인 -> 속성들이 가질 수 있는 모든 값들의 집합

개체 타입 -> 같은 개체를 갖는 속성들의 집합

오너-멤버 관계 -> 네트워크(망)형 모델

테이블에 속한 튜플 수 -> 카디널리티 / 속성의 수 -> 차수

E-R모델에서 관계 -> 개체 간의 관계 표현

릴레이션에 포함된 각 튜플 사이에는 순서가 없음

후보키 – 릴레이션을 구성하는 속성들 중에서 튜플을 유일하게 식별하기 위해 사용하는 속성들의 부분집합 / 기본키로 사용할 수 있는 속성들

대체키 – 후보키가 둘 이상일 때 기본키를 제외한 나머지 후보키를 의미

외래키 – 다른 릴레이션의 기본키를 참조하는 속성

수학의 Predicate Calculus에 기반을 두고 관계 데이터베이스를 위해 제안된 것 -> 관계 해석

3NF에서 모든 이행 종속을 제거하지 못한 경우 정보의 중복 문제가 존재

정규형들은 차수가 높아질수록 만족시켜야 할 제약조건이 증가

무결성 종류 – 개체 무결성 / 참조 무결성 / 도메인 무결성 / 사용자 정의 무결성

1NF : 릴레이션에 속한 모든 도메인이 원자값만으로 되어 있는 정규형, 릴레이션의 모든 속성 값이 원자 값으로만 되어 있는 정규형

2NF : 릴레이션 R이 1NF이고 기본키가 아닌 모든 속성이 기본키에 대하여 완전 함수적 종속을 만족하는 정규형

3NF: 릴레이션 R이 2NF이고 기본키가 아닌 모든 속성이 기본키에 대해 이행적 종속을 만족하지 않는 정규형

BCNF : 릴레이션 R에서 결정자가 모두 후보키인 정규형

4NF : 릴레이션 R에 다치 종속 AB가 성립하는 경우 R의 모든 속성이 A에 함수적 종속 관계를 만족하는 정규형

5NF : 릴레이션 R의 모든 조인 종속이 R의 후보키를 통해서만 성립되는 정규형

키워드 : 1. 개체 2. 속성 3. 관계 4. 튜플 5. 도메인 6. 키 7. 무결성 8. 관계대수 9. 정규화 10. 시스템 카탈로그