



SQL 활용

관계형 데이터 모델



한국기술교육대학교
온라인평생교육원

학습내용

- 관계형 데이터 모델의 구조와 연산
- 관계형 데이터 모델의 제약조건

학습목표

- 관계형 데이터 모델의 구조 및 연산을 사용할 수 있다.
- 관계형 데이터 모델의 제약조건을 설명할 수 있다.

● 관계형 데이터 모델의 구조와 연산

1. 데이터 모델

◆ 모델 $D = \langle S, O, C \rangle$

- Structure : 구조
 - 데이터의 구조
 - 정적 성질, 개체 타입과 이들 간의 관계를 명세함
- Operation : 연산
 - 데이터의 동적 성질
 - 개체 인스턴스를 처리하는 작업에 대한 명세
 - 데이터 조작 기법
- Constraint : 제약조건
 - 데이터의 논리적 제약
 - 구조로부터 파생, 의미적 제약
 - 데이터 조작의 한계를 표현한 규정
- 데이터 모델의 예

예 정수

- 정수(Integer)의 구조
 - ... -2, -1, 0, 1, 2, 3, ...
- 연산
 - 사칙 연산 : +, -, x, /
- 제약조건 : 정수의 세계로 한정함
 - 정수/정수 \Rightarrow 정의할 수 없는 경우가 있음
 - $5/2 \Rightarrow 2.5$: 정수가 아님

● 관계형 데이터 모델의 구조와 연산

2. 관계형 데이터 모델의 구조와 연산

◆ 관계형 데이터 모델을 사용하는 이유

- 모델의 구조가 단순함
- **집합 이론(Set Theory)**이라는 수학적 이론에 기반하여 모델이 강건(Sound)함
- SQL이라는 간단한 비절차적 언어로 사용하기 쉬움

◆ 관계형 데이터 모델

- 구조 : 릴레이션(또는 테이블)
- 연산 : 관계 대수(Relational Algebra)
- 제약조건 : 무결성(Integrity) 제약조건

◆ 관계형 데이터 모델의 구조

- 릴레이션
- 2차원 테이블 형태
- 테이블의 행(Row) ⇔ 튜플
 - 릴레이션의 튜플들의 집합
- 테이블의 열(Column) ⇔ 속성(Attribute)
 - **도메인(Domain)** : 속성이 가질 수 있는 값의 범위

예 학생

| 학번 | 이름 | 주소 | 학년 | 스키마 |
|-----|-----|----|----|-----|
| 100 | 김진현 | 서울 | 4 | 튜플 |
| 101 | 박재영 | 대전 | 1 | |
| 102 | 김홍연 | 제주 | 3 | |

속성

● 관계형 데이터 모델의 구조와 연산

2. 관계형 데이터 모델의 구조와 연산

◆ 릴레이션의 특징

- ① 튜플의 유일성
 - 릴레이션은 튜플의 집합임
 - 집합은 중복을 허용하지 않음
- ② 튜플의 무순서성
 - 릴레이션은 튜플의 집합임
 - 집합에서 원소들 간의 순서는 없음
- ③ 속성(Attribute)의 무순서성
 - 릴레이션 스키마는 속성(Attribute)들의 집합임
- ④ 속성(Attribute)의 원자성(Atomicity)
 - 속성(Attribute)의 값은 원자값임
 - 논리적으로 더 분해할 수 없음

◆ 관계형 데이터 모델의 연산

- 관계 대수(Relational Algebra)
- 관계 대수와 SQL
 - 관계 대수 : 시스템 관점
 - SQL : 사용자 위주

● 관계형 데이터 모델의 구조와 연산

2. 관계형 데이터 모델의 구조와 연산

◆ 관계 대수 연산자의 분류

● 일반 집합 연산자

- 합집합(UNION, \cup)
 - 교집합(INTERSECT, \cap)
 - 차집합(DIFFERENCE, $-$)
 - 카티션 프로덕트(CARTESIAN PRODUCT, \times)
 - 합병호환성
 - \cup , \cap , $-$ 연산의 피연산자(릴레이션)들이 지켜야 할 제약조건
 - ① 차수(Degree : 속성의 수)가 같아야 함
 - ② 대응되는 속성(Attribute) 쌍 별로 타입(또는 도메인)이 같아야 함
 - ③ 대응되는 속성(Attribute) 쌍 별로 의미(Semantic)가 같아야 함
- ①, ② 조건의 시스템에서 확인해줌
 - ③ 조건은 사람이 해야 함

예 몸무게(실수) \cup 키(실수)

- ①, ② 조건은 만족, 그러나 몸무게와 키를 합친 결과는 의미 없음

● 관계형 데이터 모델의 구조와 연산

2. 관계형 데이터 모델의 구조와 연산

◆ 관계 대수 연산자의 분류

- 순수관계 연산자 : 릴레이션이 2차원 구조이기 때문에 유도되는 연산자

- 선택(Select : σ (Sigma))

- A, B가 릴레이션 R의 속성(Attribute)일 때

$$\sigma_{A\theta v}(R) = \{ r \mid r \in R \text{ and } r.A\theta v \}$$

$$\sigma_{A\theta v}(R) = \{ r \mid r \in R \text{ and } r.A\theta r.B \}$$

- θ (Theta) : 비교 연산자

- $\{<, >, \leq, \geq, =, \neq\}$, v : 상수

- 프로젝트(Project : Π (Pi))

- A릴레이션 R(X)에서

$$Y \subseteq X \text{ 이고 } Y = \{B_1, B_2, \dots, B_m\} \text{ 이면,}$$

$$\Pi_Y(R) = \{ \langle r.B_1, \dots, r.B_m \rangle \mid r \in R \}$$

- 프로젝트(Π)의 결과는 선택 조건을 만족하는 릴레이션의 수직적 부분집합(Vertical Subset)

- 폐쇄성질
- 중복된 튜플들은 제거됨

- 관계표현식의 예제

Π 이름, 성적($\sigma_{\text{학번}=300}$ (학생))

| 학생 | 이름 | 학번 | 학과 | 성적 |
|----|-----|-----|----|-----|
| | 김홍연 | 100 | 컴공 | 2.8 |
| | 강성민 | 200 | 컴공 | 3.9 |
| | 빈준길 | 300 | 전자 | 2.3 |
| | 이석주 | 400 | 기계 | 3.5 |

● 관계형 데이터 모델의 구조와 연산

2. 관계형 데이터 모델의 구조와 연산

◆ 관계 대수 연산자의 분류

- 순수관계 연산자 : 릴레이션이 2차원 구조이기 때문에 유도되는 연산자

- 조인(JOIN, \bowtie)

$R(X), S(Y), A \in X, B \in Y$ 에 대하여

$R \bowtie_{A \theta B} S = \{ r \cdot s \mid r \in R \text{ and } s \in S \text{ and } (r.A \theta s.B) \}$

A, B : 조인 속성(Join Attribute)

결과 차수 = R 의 차수 + S 의 차수

- 조인의 예

$\Pi_{\text{교수이름}}(\text{학생} \bowtie_{\text{지도교수번호}=\text{교번}} \text{교수}) = \text{이교수}$

| 학생 | 학생 | | | 교수 | 교수 | |
|----|-----|-----|--------|----|----|------|
| | 학번 | 이름 | 지도교수번호 | | 교번 | 교수이름 |
| | 100 | 개똥이 | 3 | | 1 | 민교수 |
| | 200 | 소똥이 | 1 | | 2 | 박교수 |
| | 300 | 말똥이 | 2 | | 3 | 이교수 |

- 자연 조인(\bowtie_N)

- 일반적으로 아무 말 없이 조인이라고 하면 “자연 조인”을 말함
- 두 테이블에 공통으로 나타나는 속성에 대한 동등 조인으로 간주

● 관계형 데이터 모델의 제약조건

1. 키(Key)

◆ 키(Key)의 개념

- 하나의 테이블 내에서 각 튜플의 유일하게 식별할 수 있는 속성(Attribute)들의 집합
- 실생활의 키
 - 하나의 자물쇠를 열수 있는 열쇠(키)는 오직 하나임
 - 생활의 편리성을 위하여 키를 복제하여 쓰지만 논리적으로는 하나임

◆ 후보키(Candidate Key)

- 한 릴레이션 $R(A_1, \dots, A_n)$ 에 대한 속성의 집합 $K=\{A_i, \dots, A_j\}$ 으로 다음과 같은 성질을 만족함
 - 유일성(Uniqueness) : 서로 다른 두 튜플의 속성집합 K의 값은 같지 않음
 - 최소성(Minimality) : K는 서로 다른 두 튜플을 식별하기 위한 최소한의 속성들로만 이루어져 있음
- 릴레이션의 특징 : 튜플의 유일성
- 후보키 : 튜플의 유일성을 유지시키는 최소 속성 집합

| 학번 | 이름 | 지도교수번호 |
|-----|-----|--------|
| 100 | 개똥이 | 3 |
| 200 | 소똥이 | 1 |
| 300 | 말똥이 | 2 |

- 유일성(Uniqueness)
 - {학번, 이름, 주소}가 같은 튜플은 없음
 - {학번, 이름}이 같은 튜플도 없음
 - {학번}이 같은 튜플도 없음
- 최소성
 - {학번}이 최소로 이루어져 있음

후보키는 {학번}임

● 관계형 데이터 모델의 제약조건

1. 키(Key)

◆ 수퍼키(Super Key)

- 유일성을 만족하는 속성 집합
- 최소성을 만족하지 않아도 됨
- 일반적으로 후보키는 수퍼키의 부분 집합임 \Leftrightarrow 수퍼키는 후보키를 포함함

◆ 기본키(Primary Key)

- 하나의 릴레이션에는 후보키가 여러 개 있을 수 있음
- 여러 개의 후보키 중 DBA가 지정한 하나의 키임

◆ 대체키(Alternative Key)

- 후보키 중 기본키를 제외한 나머지 후보키

| 학번 | 이름 | 주소 | 주민번호 |
|-----|-----|----|-----------|
| 100 | 개똥이 | 천안 | xxx-xxxx1 |
| 200 | 소똥이 | 천안 | xxx-xxxx4 |
| 300 | 말똥이 | 천안 | xxx-xxxx2 |

- 후보키
 - {학번}
 - {주민번호}
- 수퍼키
 - 학번, 이름, 주소, 주민번호}
 - {학번, 이름, 주소}
 - {학번, 이름}
 - {이름, 주소, 주민번호} ...
- 기본키 : {학번}
- 대체키 : {주민번호}

● 관계형 데이터 모델의 제약조건

1. 키(Key)

◆ 외래키

- 한 릴레이션 R1의 튜플과 다른 릴레이션 R2의 하나의 튜플과의 연관 관계를 표시하기 위하여 사용함
- R1의 속성집합 FK의 도메인이 R2의 기본키 일 때, FK를 R1의 외래키라 함
- R1을 참조 릴레이션, R2를 피참조 릴레이션이라고 함
 - R1과 R2가 다른 릴레이션일 필요는 없음

| 사번 | 이름 | 직급 | 관리자 |
|----|----|----|-----|
| 1 | 개 | 사장 | - |
| 2 | 소 | 부장 | 1 |
| 3 | 말 | 과장 | 2 |
| 4 | 닭 | 사원 | 3 |

◆ 개체 무결성

- 의미 : 서로 다른 두 튜플은 같을 수 없음
- 정의 : 기본키 값은 언제나 어느 때고 NULL 값일 수 없음

● 관계형 데이터 모델의 제약조건

1. 키(Key)

◆ 참조 무결성

- 외래키 값은 피참조 릴레이션의 기본키 값이거나 NULL 값임
- 추가 지정을 통해 NULL을 가질 수 없다고 제약을 걸 수 있음

| 학생 | | 외래키 | 기본키 | | 교수 |
|-----|-----|--------|-----|------|----|
| 학번 | 이름 | 지도교수번호 | 교번 | 교수이름 | |
| 100 | 개똥이 | 3 | 1 | 민교수 | |
| 200 | 소똥이 | 1 | 2 | 박교수 | |
| 300 | 말똥이 | 2 | 3 | 이교수 | |
| 400 | 취똥이 | - | | | |
| 500 | 닭똥이 | 9 | | | |

- 취똥이는 신입생이라 아직 지도교수 배정 안됨 ⇨ 아직 모름 ⇨ NULL

| 학생 | | 외래키 | 기본키 | | 교수 |
|-----|-----|--------|-----|------|----|
| 학번 | 이름 | 지도교수번호 | 교번 | 교수이름 | |
| 100 | 개똥이 | 3 | 1 | 민교수 | |
| 200 | 소똥이 | 1 | 2 | 박교수 | |
| 300 | 말똥이 | 2 | 3 | 이교수 | |
| 400 | 취똥이 | - | | | |
| 500 | 닭똥이 | 9 | | | |

- 닭똥이는 지도교수가 9번이라고 하지만 교수테이블에는 9번 교수가 없음
⇨ **참조 무결성 위반**
- 이런 튜플은 존재할 수 없음

● 관계형 데이터 모델의 제약조건

1. 키(Key)

◆ 도메인 무결성

- 속성(Attribute) 값은 해당 속성(Attribute) 도메인에 속한 값들 중 하나이어야 함

예) 대학생의 학년 도메인 : 1, 2, 3, 4

- 9학년 또는 10학년은 존재할 수 없음

- DBMS는 데이터베이스의 상태 변화(삽입, 갱신, 삭제)에도 항상 무결성 제약조건을 검사하고 유지시킴

핵심요약

1. 관계형 데이터 모델의 구조와 연산

■ 데이터 모델

■ 구조

- 데이터의 구조
- 정적 성질, 개체 타입과 이들 간의 관계를 명세

■ 연산

- 데이터의 동적 성질
- 개체 인스턴스를 처리하는 작업에 대한 명세
- 데이터 조작 기법

■ 제약조건

- 데이터의 논리적 제약
- 구조로 부터 파생, 의미적 제약
- 데이터 조작의 한계를 표현한 규정

핵심요약

1. 관계형 데이터 모델의 구조와 연산

■ 관계형 데이터 모델의 구조와 연산

■ 관계형 데이터 모델의 구조

- 릴레이션
- 2차원 테이블 형태
- 테이블의 행(Row) → 튜플
- 테이블의 열(Column) → 속성(Attribute)

■ 일반 집합 연산자

- 합집합(UNION, \cup)
- 교집합(INTERSECT, \cap)
- 차집합(DIFFERENCE, $-$)
- 카티션 프로덕트(CARTESIAN PRODUCT, \times)

■ 순수관계 연산자

- 선택(SELECT, σ)
- 프로젝트(PROJECT, Π)
- 조인(JOIN, \bowtie)
- 디비전(DIVISION, \div)

핵심요약

2. 관계형 데이터 모델의 제약조건

■ 키

- 하나의 테이블 내에서 각 튜플의 유일하게 식별할 수 있는 속성(Attribute)들의 집합
- 하나의 자물쇠를 열수 있는 열쇠(키)는 오직 하나임
- 후보키 (Candidate Key)
 - 한 릴레이션 $R(A_1, \dots, A_n)$ 에 대한 속성의 집합
 - 유일성, 최소성
- 수퍼키(Super Key)
 - 유일성을 만족하는 속성 집합
 - 최소성을 만족하지 않아도 됨
- 기본키(Primary Key)
 - 하나의 릴레이션에는 후보키가 여러 개 있을 수 있음
 - 여러 개의 후보키 중 DBA가 지정한 하나의 키
- 대체키(Alternative Key)
 - 후보키 중 기본키를 제외한 나머지 후보키

핵심요약

2. 관계형 데이터 모델의 제약조건

■ 개체 무결성, 참조 무결성, 도메인 무결성

■ 개체 무결성

- 의미 : 서로 다른 두 튜플은 같을 수 없음
- 정의 : 기본키 값은 언제나 어느 때고 NULL값일 수 없음

■ 참조 무결성

- 외래키 값은 피참조 릴레이션의 기본키 값이거나 NULL 값임
- 추가 지정을 통해 NULL을 가질 수 없다고 제약을 걸을 수 있음

■ 도메인 무결성

- 속성(Attribute) 값은 해당 속성(Attribute) 도메인에 속한 값들 중 하나이어야 함
- DBMS는 데이터베이스의 상태 변화(삽입, 갱신, 삭제)에도 항상 무결성 제약조건을 검사하고 유지시킴