**데이터 베이스 프로그래밍 복습하기**

**담당 주제 : (정규화)**

모둠원 학번/이름 : 20816 이시현,20818 이찬휘

<내용 정리하기>

**1. 이상(anomaly) 현상**

-불필요한 데이터 중복으로 인해 릴레이션에 대한 데이터 삽입·수정·삭제 연산을 수행할 때 발생할 수 있는 부작용

**>삽입 이상(insertion anomaly)**

릴레이션에 새 데이터를 삽입하려면 불필요한 데이터도 함께 삽입해야 하는 문제

**>갱신 이상(update anomaly)**

릴레이션의 중복된 투플들 중 일부만 수정하여 데이터가 불일치하게 되는 모순이 발생하는 문제

**>삭제 이상(deletion anomaly)**

릴레이션에서 투플을 삭제하면 꼭 필요한 데이터까지 손실되는 연쇄 삭제 현상이 발생하는 문제

**2. 정규화**

-이상 현상을 제거하면서 데이터베이스를 올바르게 설계해 나가는 과정

-릴레이션을 관련 있는 속성들로만 구성하기 위해 릴레이션을 분해(decomposition)하는 과정

-**함수적 종속성\***을 판단하여 정규화를 수행함

**\*함수 종속**

-속성들 간의 관련성

-X → Y로 표현하며, X가 Y를 함수적으로 결정한다, “Y가 X에 함수적으로 종속되어 있다”와 같은 의미

**>완전 함수 종속**

릴레이션에서 속성 집합 Y가 속성 집합 X에 함수적으로 종속되어 있지만, 속성 집합 X의 전체가 아닌 일부분에는 종속되지 않음을 의미

**>부분 함수 종속**

릴레이션에서 속성 집합 Y가 속성 집합 X의 전체가 아닌 일부분에도 함수적으로 종속됨을 의미

**>이행적 함수 종속**

릴레이션을 구성하는 3개의 속성 집합 X, Y, Z에 대해 함수 종속 관계 X → Y와 Y → Z가 존재하면 논리적으로 X → Z가 성립되는데, 이때 Z가 X에 이행적으로 함수 종속되었다고 함

**3. 정규형(NF; Normal Form)**

-릴레이션이 정규화된 정도를 의미하며, 각 정규형마다 제약조건이 존재

-정규형의 차수가 높아질수록 요구되는 제약조건이 많아지고 엄격해짐

-릴레이션의 특성을 고려해서 적합한 정규형을 선택

**>제1정규형**(**1NF**; First Normal Form)

릴레이션의 모든 속성이 더는 분해되지 않는 원자 값(atomic value)만 가지면 제1정규형을 만족함

**>제2정규형**(**2NF**; Second Normal Form)

릴레이션이 제1정규형에 속하고, 기본키가 아닌 모든 속성이 기본키에 완전 함수 종속되면 제2정규형을 만족함

**>제3정규형**(**3NF**; Third Normal Form)

릴레이션이 제2정규형에 속하고, 기본키가 아닌 모든 속성이 기본키에 이행적 함수 종속이 되지 않으면 제3정규형을 만족함

**>보이스/코드 정규형**(**BCNF**; Boyce/Codd Normal Form/**강한 제3정규형**(strong 3NF))

**-의미**

후보키를 여러 개 가지고 있는 릴레이션에 발생할 수 있는 이상 현상을 해결하기 위해 제3정규형보다 좀 더 엄격한 제약조건을 제시하는 정규형

보이스/코드 정규형에 속하는 모든 릴레이션은 제3정규형에 속하지만, 제3정규형에 속하는 모든 릴레이션이 보이스/코드 정규형에 속하는 것은 아님

**-필요성**

하나의 릴레이션에 여러 개의 후보키가 존재하는 경우, 제3정규형까지 모두 만족해도 이상 현상이 발생할 수 있음

**>제4정규형**

릴레이션이 보이스/코드 정규형을 만족하면서, 함수 종속이 아닌 다치 종속(MVD; Multi Valued Dependency)을 제거하면 제4정규형에 속함

**>제5정규형**

릴레이션이 제4정규형을 만족하면서, 후보키를 통하지 않는 조인 종속(JD; Join Dependency)을 제거하면 제5정규형에 속함

**정규화 과정 정리**

****