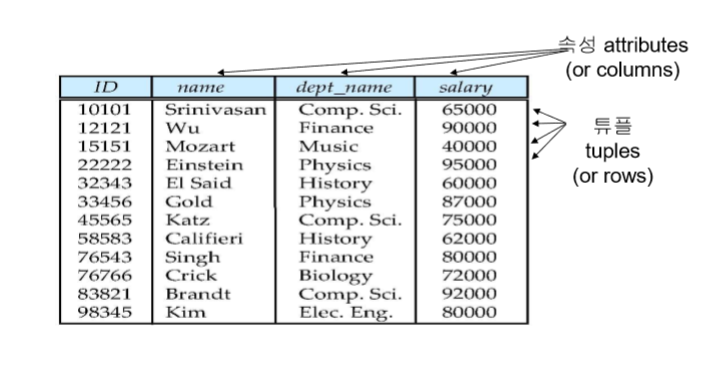
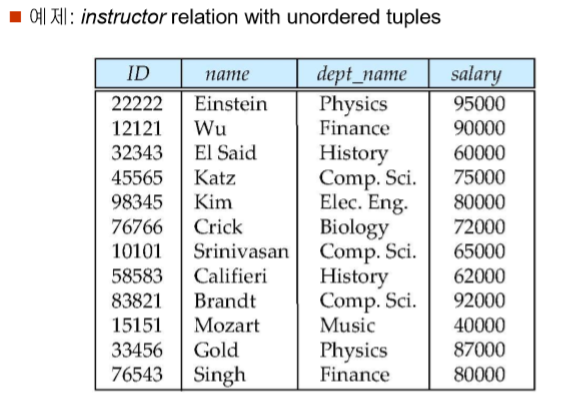
[빅데이터이론] 2019.03.13(수)

관계형 데이터 모델



속성 타입(attribute types)

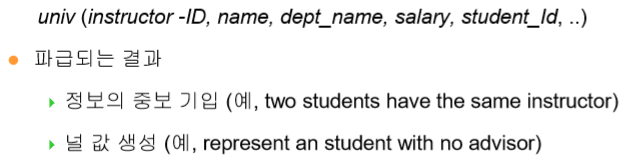
* 속성의 도메인(domain): 속성 값으로 허용되는 값의 집합
* 속성 값은 일반적으로 **원자 값(atomic value)의 성질을 만족**해야 한다.
* 널(null) 값은 모든 도메인에 포함된다고 가정.
  + 값이 아예 없는 것

릴레이션 스키마와 인스턴스

* A1, A2, …, An : 속성
* R = (A1, A2, …, An: 릴레이션 스키마
  + Example:  
    instructor = (ID, name, dept\_name, salary)
* 개체 집합 D1, D2, …. Dn 이 주어지면 릴레이션 r 은 D1 x D2 x … x Dn 의 부분 집합이다.
* 따라서, 릴레이션은 aiDi인 n-튜플(a1, a2, …, an)의 집합이다.
* 튜플들의 순서에는 의미가 없다. 즉, 튜플은 임의의 순서로 저장될 수 있다.

데이터베이스

* 다수의 릴레이션(테이블)으로 구성됨.
* 전체 데이터베이스 정보는 부분으로 나눠 짐.
* 잘못된 예

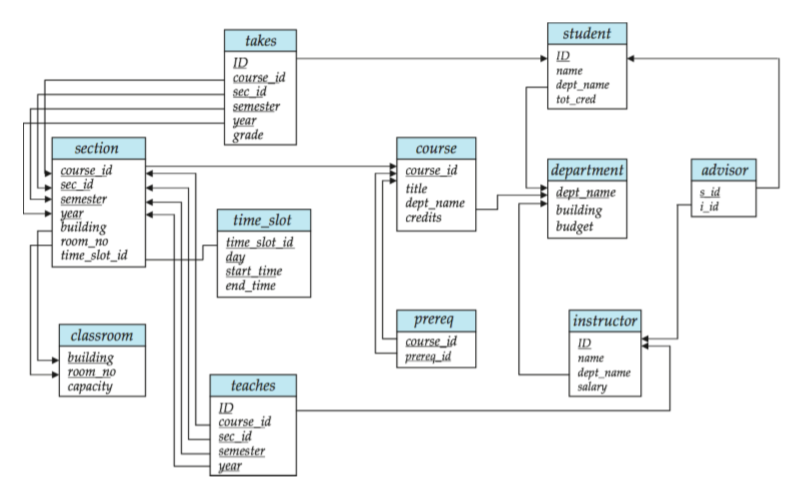


* + 차이점이 생기는데, 그 차이점이 무엇이냐?  
    🡪 데이터 수정 시, 비용이 많이 들어서 나눠서 쓰는 것이다.

키(Key)

* K ⊆ R 이라고 가정
* K의 값이 릴레이션 r(R)의 고유한 튜플을 구분하는데 충분하다면 K는 R의 수퍼 키이다.
  + 예제:  
    {ID} and {ID,name} are both superkeys of instructor.
* 후보 키(candidate key)
  + 수퍼 키 K가 최소의 조건을 만족시키면 후보 키임
  + 예제:  
    {ID} is a candidate key for Instructor
* 주 키(primary key) 🡪 인덱스를 만들기 위해서 사용됨.
  + 데이터베이스 설계자에 의해 선택한 후보 키
* **외래 키(Foreign key)** 🡪 다른 테이블과 관계를 맺어줄 때
  + 릴레이션 r1은 자신의 속성들 가운데 다른 릴레이션 r2의 주 키를 포함할 수 있다. 이러한 속성을 r2를 참조하는 외래 키라 함.
  + **제약조건(constraint**): 릴레이션 r1내의 외래 키에 해당하는 속성 값은 반드시 릴레이션 r2내에 존재하여야 함. //참조 무결성

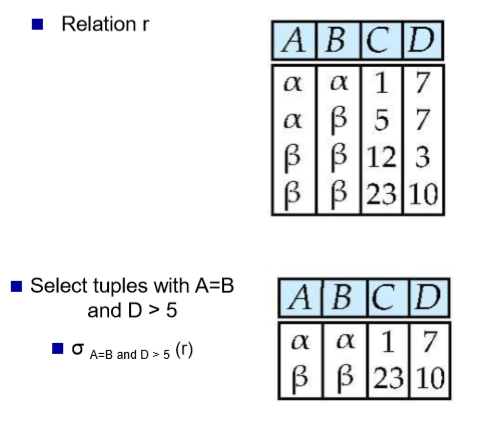
Schema Diagram for University Database



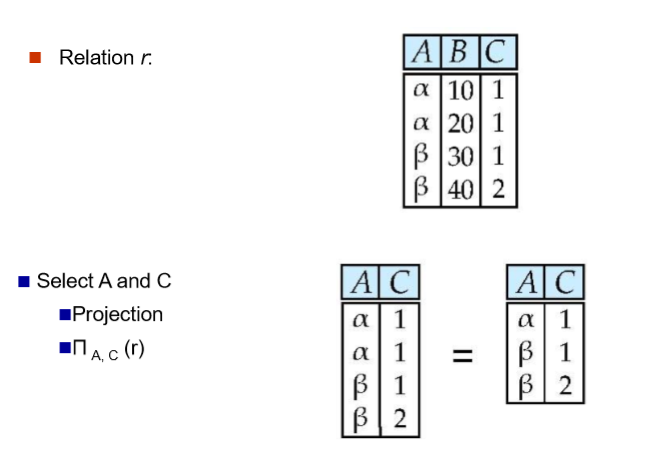
관계형 질의어(Relational Query Languages)

* 사용자가 데이터베이스로부터 정보를 요청한 때 사용하는 언어
* 언어의 분류
  + 절차식(procedural language)
  + 비절차식(nonprocedural language)
* 순수 언어(pure language)
  + 관계 대수 🡪 절차식 언어에 해당
    - 릴레이션에서 원하는 데이터를 추출하기 위해 수학의 대수와 같은 연산을 이용하여 질의하는 방법을 시굴하는 언어
  + 관계 해석 🡪 비절차식 언어에 해당
    - 어떤 데이터를 찾는지만 명시하는 선언적 언어
  + SQL 질의언어는 기본적으로 관계 해석을 기반으로 하지만 DBMS 내부에서는 관계대수에 기반을 둔 연산을 수행하여 최정 결과 집합을 반환함.

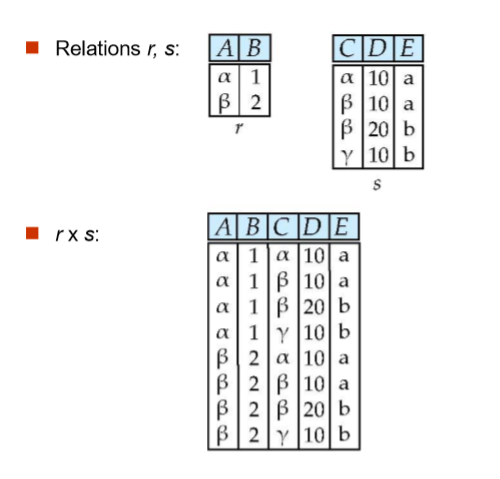
선택연산(Selection of tuples)



추출연산(Selection of Columns)

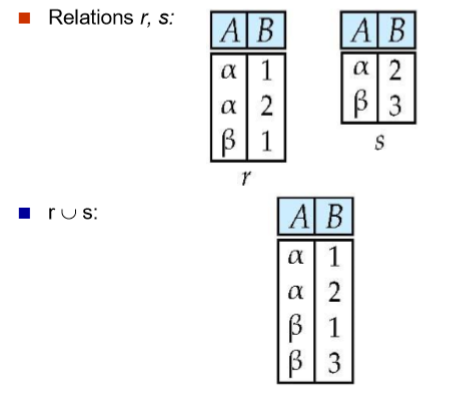


두 릴레이션의 카티션 곱

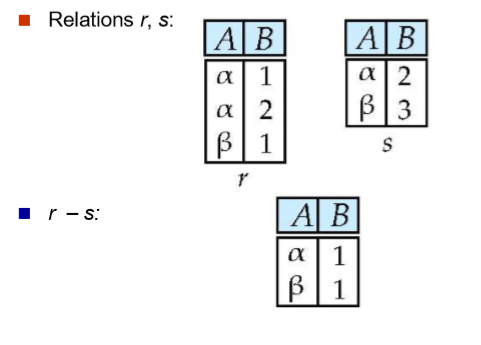


* 공통 속성을 뺌.

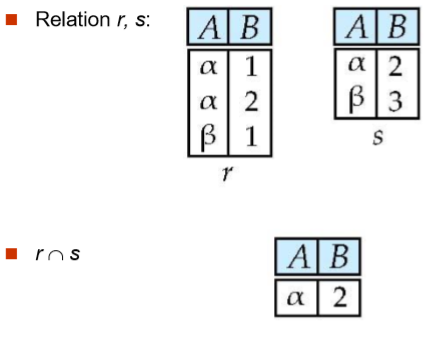
합집합(Union of two relations)



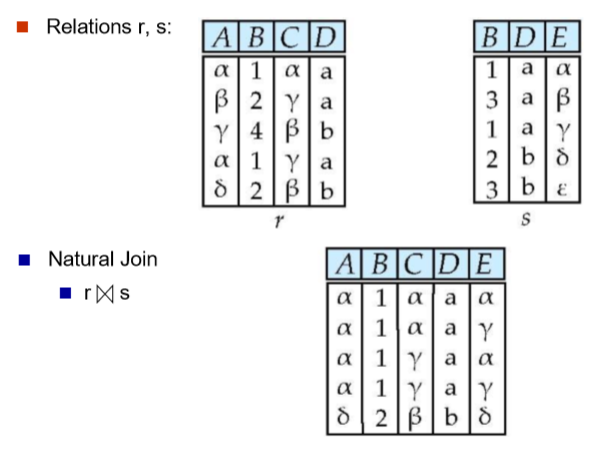
차집합(Set difference of two relations)



공통집합(Set Intersectoin of two relations)



자연 조인 예제



참고

