CS145 Lecture Notes #2
Database Design Using E/R Model

Steps in Building a Database

1. Understand real-world domain being captured
2. Specify it using a database design model
3. Translate specification to the data model of DBMS
4. Create DBMS schema and load data

E/R Basics

Entity-relationship (E/R) model is a data model:
- Historically very popular
- Primarily a design model—not implemented by any major DBMS

Diagrams to represent designs:
- *Entity*: a "thing", like a record or an object
- *Entity set*: a collection of similar things, like a set of records or a class
  ↝ represented as a rectangle
- *Relationship*: an association among two or more entities
- *Relationship set*[*]: a set of relationships of the same type; an association among two or more entity sets
  ↝ represented as a diamond
- *Attributes*: properties of entities or relationships, like record fields
  ↝ represented as ovals

Example: students take courses

---

[*]Ullman/Widom textbook uses slightly different terminology: it uses *relationship* to mean what we call a relationship set, and it uses *relationship set* to refer to the table representation of a relationship.

Keys of Entity Sets

A *key* is a set of attributes whose values can belong to at most one entity in an entity set—like an entity "ID"

Each entity set must have a key

$\rightsquigarrow$ represented by underlining all attributes in the key

Example: {SID} is a key of Students; {dept, number} is a key of Courses

- An entity set may have multiple candidate keys (e.g., {SSN} is also a key of Students); just pick one key
- A key may include multiple attributes (e.g., {dept, number} is a key of Courses)

Attributes on Relationships

Example: students take courses and receive grades

Key attributes identify entities; what identify relationships?

- In a relationship set, each relationship is uniquely identified by the entities it connects
- $\rightsquigarrow$ There can be at most one Take relationship between Bart and CS145
- $\rightsquigarrow$ But there could be another TA relationship between Bart and CS145

Tricky: what if Bart took CS145 twice and got different grades?

Multiplicity of Relationships

- Many-many: each entity in E is related to 0 or more entities in F and vice-versa (e.g., Students-Courses)

- Many-one: each entity in E is related to 0 or 1 entity in F, but each entity in F is related to 0 or more in E (e.g., Courses-Instructors)

- One-one: each entity in E is related to 0 or 1 entity in F and vice-versa (e.g., Students-OracleAccounts)

- "One" is represented by an arrow
- "Exactly one" is represented by a rounded arrow

Arity of Relationships

Usually *binary* relationship sets suffice; but in some cases three or more entity sets must be connected by one *multiway* relationship set

Example: relationship among students, courses, and TA's

- First try two binary relationship sets: Students Take Courses; TAs Assist Courses

  Connection between Students and TAs is *only* via Courses; works in CS145, because each TA is a TA of all students
- But what if each student is assigned to one of the TA's for the course? A ternary relationship set:

How about multiplicity?

Formal meaning of arrow pointing to entity set E participating in a multiway relationship set: pick one entity from each other entity set, together they must be related to 0 or 1 entity in E

Roles

Sometimes an entity set participate more than once in a relationship set (e.g., students are married as husband and wife)
⤳ label edges with *roles* to distinguish

In a *symmetric* relationship set, the roles played by the entity set are the same (e.g., students are roommates of each other)

Weak Entity Sets

Sometimes the key of an entity set E comes not (completely) from its own attributes, but from the keys of other (one or more) entity sets to which E is linked by many-one relationship sets
- E is called a *weak entity set*
- Use of many-one relationship set (includes one-one) essential: with many-many, we wouldn't know which entity provides the key value!
- ⤳ Represented by: a double rectangle around E, and a double diamond around each relationship set through which E is linked to another entity set that provides part of E's key

Example: rooms in buildings

"Chain-of-weakness" example: seats in rooms in buildings

Using Weak Entity Sets to Model Multiway Relationships

Trick: replace an n-ary relationship set with a weak entity set (called a *connecting entity set*) and n binary relationship sets
Example: Students-Courses-TAs

ISA Relationships

Similar to the idea of "subclass" in object-oriented terms: subclass = special case; fewer entities; more properties

⤳ represented as a triangle (direction is important)

Example: graduate students are just students, but they also have advisors and offices