CS 235: Introduction to Databases
Svetlozar Nestorov
Lecture Notes #2

The Big Picture
- Stages of building DB application: data tier
- Real-world domain.
  - understand client needs.
- Design data model:
  - using entity-relationship (E/R) model.
- Database data model:
  - using relational model.
- Create schema in DBMS, load data.
  - using SQL, loader.

Last Time
- Entity-Relationship Model

Outline
- More design issues:
  - Subclasses,
  - Keys,
  - Weak entity sets.
- Exercise problems

Subclasses
- Subclass:
  - special case
  - fewer entities
  - more properties.
- Example: Ales are a kind of beer.
  - In addition to the properties (= attributes and relationships) of beers, there is a color attribute for ales.

E/R Subclasses
- isa triangles indicate the subclass relation.

Different Subclass Viewpoints
- **E/R viewpoint:** An entity has a component in each entity set to which it logically belongs.
  - Its properties are the union of the properties of these E.S.
- **Object-oriented viewpoint:** An object (entity) belongs to exactly one class.
  - It inherits properties of its superclasses.

Subclasses Example

**Multiple Inheritance**
- Theoretically, an E.S. could be a subclass of several other entity sets.

**Problems**
- How should conflicts be resolved?
- Example: manf means grower for wines, bottler for beers. What does manf mean for "grape beers"?
- Need ad-hoc notation to resolve meanings.
- In practice, we shall assume a tree of entity sets connected by isa, with all "isas" pointing from child to parent.

**Keys**
- A key is a set of attributes whose values can belong to at most one entity.
  - The value of a key is unique.
- In E/R model, every E.S. must have a key.
  - It could have more than one key, but one set of attributes is the designated key.
- In E/R diagrams, you should underline all attributes of the designated key.

**Example**
- Suppose name is key for Beers.
  - Beer name is also key for ales.
    - In general, key at root is key for all.
Example: A Multiattribute Key

- What is the key?

Example: Email Addresses

- Email address = user name + host name, e.g., evtimov@cs.uchicago.edu.
- Email address corresponds to a user name on a particular host.
- Once on a host, you only need user name, e.g., evtimov
- Key for an email = the user name at the host (which is unique for that host only) + the IP address of the host (which is unique globally).

Example: Chain of Weakness

- Consider IP addresses consisting of a primary domain (e.g., edu), subdomain (e.g., uchicago), and host (e.g., cs).

Weak Entity Sets

- Sometimes an E.S. E's key comes not (completely) from its own attributes, but from the keys of one or more E.S.'s to which E is linked by a supporting many-one relationship.
- Called a weak E.S.
- Represented by putting double rectangle around E and a double diamond around each supporting relationship.
- Many-one-ness of supporting relationship (includes 1-1) essential.
  - With many-many, we wouldn't know which entity provided the key value.

Email Addresses

- Design issue: Under what circumstances could we simply make user-name and host-name be attributes of email, and dispense with the weak E.S.?
All Connecting Entity Sets Are Weak

In this special case, where bar and beer determine a price, we can omit price from the key, and remove the double diamond from ThePrice.

Better: price is an attribute of BBP.

Design Principles

- Setting: client has (possibly vague) idea of what he/she wants. You must design a database that represents these thoughts and only these thoughts.
- Avoid redundancy.
  - Wastes space and encourages inconsistency.
  - Intuition: something is redundant if it could be hidden from view, and you could still figure out what it is from the other data.
- Avoid intermediate concepts.

Good and Bad Design

Bad?

- Question: Why is it OK to have Beers with just its key as attribute? Why not make set of beers an attribute of manufacturers?
Exercise Problem 2

- E/R diagrams

Exercise Problem 3

- Multiway relationships