CS 235: Introduction to Databases

Svetlozar Nestorov

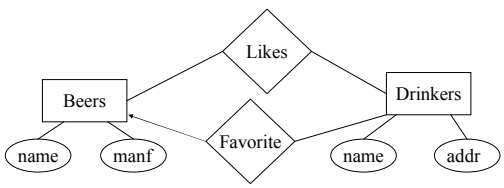
Lecture Notes #2

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

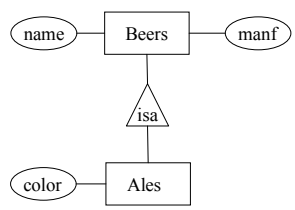
- More design issues
- Subclasses
- Keys

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

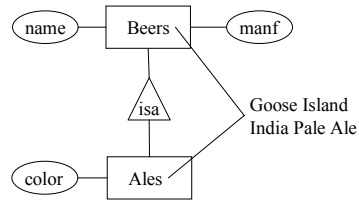
- *isa* triangles indicate the subclass relation.



Different Subclass Viewpoints

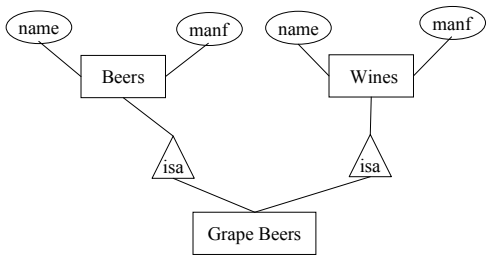
- *E/R viewpoint*: An entity has a *component* in each entity set to which it logically belongs.
  - Its properties are the union of the properties of these E.S.
- *Object-oriented viewpoint*: An object (entity) belongs to exactly one class.
  - It *inherits* properties of its superclasses.

Subclasses Example



Multiple Inheritance

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



Problems

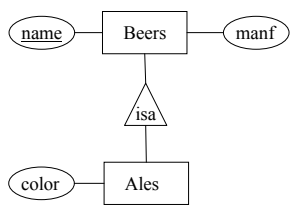
- How should conflicts be resolved?
- Example: *manf* means grower for wines, bottler for beers. What does *manf* mean for "grape beers"?
- Need ad-hoc notation to resolve meanings.
- In practice, we shall assume a tree of entity sets connected by *isa*, with all "isas" pointing from child to parent.

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.

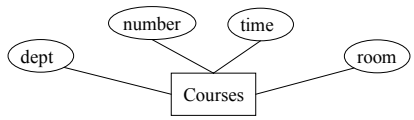
Example

- 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?

Surrogate Keys

- Synthetically generated unique identifiers, often by the database system
- Pros
  - Immutable, performance (integers)
- Cons
  - Disassociations, confusion
- Implementation
  - SEQUENCE (Oracle)
  - AUTO_INCREMENT (MySQL)