Data Modeling

Problem in Real World

Tables in RDBM

Entity-Relationship (ER) Model

Problem → ER Model → Tables

- *Sort of* an object-oriented approach
  - Support subclasses
  - A graphical representation of the design
    - ER Diagram
  - Easily converted to relational model

ER Diagram

Entity Set and Attributes

- Entity Set is similar to *class* in an OO language
- Attributes are the properties of an entity set
  - Similar to the *class variables* in an OO language
  - Must have simple values like numbers or strings – *cannot be collection or composite type*

Instances of An Entity Set

- Entity – *object* in an OO language
  - (Bud, Anheuser-Busch)
  - (Miller, Miller Brewing)
  - (Bud Lite, Anheuser-Busch)
  - (Joe's Bar, 113 Main St, Full)
  - (Sue's Bar, 20 East St, Beer)
Many-to-Many Relationship

- An entity of either set can be connected to many entities of the other set.

Many-to-One Relationship

- The relationship Favorite between Drinkers and Beers.

One-to-One Relationship

- The relationship Best-seller between Manufactures and Beers.

Instances of a relationship?
Multiway Relationship

Sometimes we need a relationship that connects more than two entity sets. Suppose drinkers will only drink certain beers at certain bars.

A 3-Way Relationship

**Instances of the Preferences Relationship**

<table>
<thead>
<tr>
<th>Bar</th>
<th>Drinker</th>
<th>Beer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joe’s Bar</td>
<td>Ann</td>
<td>Miller</td>
</tr>
<tr>
<td>Sue’s Bar</td>
<td>Ann</td>
<td>Bud</td>
</tr>
<tr>
<td>Sue’s Bar</td>
<td>Ann</td>
<td>Pete’s Ale</td>
</tr>
<tr>
<td>Joe’s Bar</td>
<td>Bob</td>
<td>Bud</td>
</tr>
<tr>
<td>Joe’s Bar</td>
<td>Bob</td>
<td>Miller</td>
</tr>
<tr>
<td>Joe’s Bar</td>
<td>Cal</td>
<td>Miller</td>
</tr>
<tr>
<td>Sue’s Bar</td>
<td>Cal</td>
<td>Bud Late</td>
</tr>
</tbody>
</table>

**“Arrows” in Multi-way Relationships**

What does an arrow mean in a multi-way relationship? Can we add any arrows in the Preferences relationship?

Attributes of Relationships

Sometimes it’s useful to attach an attribute to a relationship.

Can we do without relationship attributes?

Roles

An entity set may appear in the same relationship more than once. Label the edges with names called Roles.
A Different Perspective

<table>
<thead>
<tr>
<th>Husband (Drinkers)</th>
<th>Married</th>
<th>Wife (Drinkers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bob</td>
<td>Sue</td>
<td>Joe</td>
</tr>
</tbody>
</table>

Subclasses

- Subclass
  - Special case
  - More properties
  - No multiple inheritance
- Represented by the isa triangle

Keys

- A key is an attribute or a set of attributes that uniquely identify an entity in an entity set.

Rules about Keys

- Each entity set must have a key
- If there are multiple keys, choose one of them as the **primary key**
- Super class must have all the key attributes

Weak Entity Set

- Entity set $E$ is said to be weak if in order to identify entities of $E$ uniquely, we need to follow one or more many-one relationships from $E$ and include the key of the related entities from the connected entity sets.
Weak Entity Set Example

What's the key for Players??

Representing Weak Entity Sets

The key of a weak entity set consists of its own key attributes and the key attributes of the supporting set.

Referential Integrity

A stronger many-to-one or one-to-one relationship

many-to-one   one-to-one

Design Principles

- Faithfulness
- Avoid redundancy
- Don't use an entity set when an attribute would do
- Limit the use of weak entity set

Avoid Redundancy

Redundancy wastes space, and more importantly, encourages inconsistency.
Entity Set vs. Attributes

- An entity set should satisfy at least one of the following conditions:
  - It is more than the name of something; it has at least one non-key attribute, or
  - It is the "many" in a many-one or many-many relationship.
Don't Overuse Weak Entity Set

- We can usually create unique IDs for entity sets.

Some Common Problems

- Missing arrows
- Not identifying keys
- Two entity sets are connected without a relationship
- Attributes are not of simple type
- Misuse of multiway relationships

Example 1: Receipt Revisited

<table>
<thead>
<tr>
<th>Some Restaurant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Terminal ID: MC2HHRY</td>
</tr>
<tr>
<td>Merchant ID: 4492414532566624</td>
</tr>
<tr>
<td>VISA</td>
</tr>
<tr>
<td>****************1234</td>
</tr>
<tr>
<td><strong>SALE</strong></td>
</tr>
<tr>
<td>inv:00032</td>
</tr>
<tr>
<td>Date: JUN 17, 06  Time: 18:44</td>
</tr>
<tr>
<td>Base: $36.70</td>
</tr>
<tr>
<td>Tip:</td>
</tr>
<tr>
<td>Total:</td>
</tr>
<tr>
<td>Chengyu Sun</td>
</tr>
</tbody>
</table>

Example 2: Product Hierarchy

Example 3: Price That Changes

What if we want to model price that changes??

Price of Intel P3