E-R Diagram
Database Development

• We know how to query a database using SQL
  – A set of tables and their schemas are given
  – Data are properly loaded

• But, how can we develop appropriate tables and their schema for an application?
  – In real applications, data often does not present as tables naturally
  – What are the corresponding data units of tables?
What Is Data in Applications?

• A student information system
  – Objects: students (Ann, Bob, …), courses (354, 459, …), departments (CS, Engineering, …), …

• Objects are related
  – Students taking courses (Ann takes 354, Bob takes 459, …), courses offered by departments (354 and 459 are offered by CS), …

• Generally, an application contains a set of objects and their relationships
Entities

- An entity: an object that exists and is distinguishable from other objects
  - E.g., Ann, Bob, CS, Engineering, 354, 459, …
  - Entities have attributes, e.g., Ann has a phone number and an address

- An entity set: a set of entities of the same type that share the same properties
  - E.g., the set of students, the set of departments, the set of courses, …
## Entity Sets in Relational Databases

<table>
<thead>
<tr>
<th>customer_id</th>
<th>name</th>
<th>street</th>
<th>city</th>
</tr>
</thead>
<tbody>
<tr>
<td>321-12-3123</td>
<td>Jones</td>
<td>Main</td>
<td>Harrison</td>
</tr>
<tr>
<td>019-28-3746</td>
<td>Smith</td>
<td>North</td>
<td>Rye</td>
</tr>
<tr>
<td>677-89-9011</td>
<td>Hayes</td>
<td>Main</td>
<td>Harrison</td>
</tr>
<tr>
<td>555-55-5555</td>
<td>Jackson</td>
<td>Dupont</td>
<td>Woodside</td>
</tr>
<tr>
<td>244-66-8800</td>
<td>Curry</td>
<td>North</td>
<td>Rye</td>
</tr>
<tr>
<td>963-96-3963</td>
<td>Williams</td>
<td>Nassau</td>
<td>Princeton</td>
</tr>
<tr>
<td>335-57-7991</td>
<td>Adams</td>
<td>Spring</td>
<td>Pittsfield</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>loan</th>
<th>amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-17</td>
<td>1000</td>
</tr>
<tr>
<td>L-23</td>
<td>2000</td>
</tr>
<tr>
<td>L-15</td>
<td>1500</td>
</tr>
<tr>
<td>L-14</td>
<td>1500</td>
</tr>
<tr>
<td>L-19</td>
<td>500</td>
</tr>
<tr>
<td>L-11</td>
<td>900</td>
</tr>
<tr>
<td>L-16</td>
<td>1300</td>
</tr>
</tbody>
</table>

*customer* | *loan*
Attributes

• An entity is represented by a set of attributes – the descriptive properties possessed by all members of an entity set

  \[ \text{customer} = (\text{customer	extunderscore id}, \text{customer	extunderscore name}, \text{customer	extunderscore street}, \text{customer	extunderscore city}) \]

  \[ \text{loan} = (\text{loan	extunderscore number}, \text{amount}) \]

• Domain – the set of permitted values for an attribute
Attribute types

• Simple and composite attributes
  – Simple: cannot be divided into subparts
  – Composite: Name = first_name + last_name

• Single-valued and multi-valued attributes
  – Single-valued: each entity has only one value
  – Multi-valued: an entity may have zero, one, or more values, e.g., telephone numbers

• Derived attributes
  – Can be computed from other attributes
  – Example: age, given date_of_birth
Relationships

• A relationship: an association among several entities
  – Ann takes 354, Bob takes 459
  – A set of relationships may share common features: student-taking-courses

• A relationship set: a mathematical relation among $n \geq 2$ entities, each taken from an entity set
  – $\{(e_1, e_2, ..., e_n) : e_1 \in E_1, e_2 \in E_2, ..., e_n \in E_n\}$, where $(e_1, e_2, ..., e_n)$ is a relationship
  – Example: $(Ann, 354) \in std\text{-}take\text{-}crs, (Bob, 459) \in std\text{-}take\text{-}crs$
### Relationship Set borrower

<table>
<thead>
<tr>
<th>Customer ID</th>
<th>First Name</th>
<th>Last Name</th>
<th>Address</th>
<th>City</th>
</tr>
</thead>
<tbody>
<tr>
<td>321-12-3123</td>
<td>Jones</td>
<td>Main</td>
<td>Harrison</td>
<td></td>
</tr>
<tr>
<td>019-28-3746</td>
<td>Smith</td>
<td>North</td>
<td>Rye</td>
<td></td>
</tr>
<tr>
<td>677-89-9011</td>
<td>Hayes</td>
<td>Main</td>
<td>Harrison</td>
<td></td>
</tr>
<tr>
<td>555-55-5555</td>
<td>Jackson</td>
<td>Dupont</td>
<td>Woodside</td>
<td></td>
</tr>
<tr>
<td>244-66-8800</td>
<td>Curry</td>
<td>North</td>
<td>Rye</td>
<td></td>
</tr>
<tr>
<td>963-96-3963</td>
<td>Williams</td>
<td>Nassau</td>
<td>Princeton</td>
<td></td>
</tr>
<tr>
<td>335-57-7991</td>
<td>Adams</td>
<td>Spring</td>
<td>Pittsfield</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Loan ID</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-17</td>
<td>1000</td>
</tr>
<tr>
<td>L-23</td>
<td>2000</td>
</tr>
<tr>
<td>L-15</td>
<td>1500</td>
</tr>
<tr>
<td>L-14</td>
<td>1500</td>
</tr>
<tr>
<td>L-19</td>
<td>500</td>
</tr>
<tr>
<td>L-11</td>
<td>900</td>
</tr>
<tr>
<td>L-16</td>
<td>1300</td>
</tr>
</tbody>
</table>

**customer** → **loan**
Properties of Relationship Sets

- A relationship set can also have properties
Degree of a Relationship Set

- The number of entity sets that participate in a relationship set
  - Relationship sets that involve two entity sets are binary (or of degree two)
  - Most relationship sets in a database system are binary
- Relationship sets may involve more than two entity sets
  - Example: a ternary relationship set between entity sets student, course, and instructor
Mapping Cardinality Constraints

• Express the number of entities to which another entity can be associated via a relationship set
  – Most useful in describing binary relationship sets

• For a binary relationship set the mapping cardinality must be one of the following types
  – One to one, e.g., president – university
  – One to many, e.g., instructor – course
  – Many to one, e.g., course – instructor
  – Many to many, e.g., student – course
Mapping Cardinalities

(a) One to one

(b) One to many
Mapping Cardinalities

(a) Many to one

(b) Many to many
Entity-Relationship (ER) Model

- Elements in a database: data entries
- Data entries represent
  - Entities: data objects, e.g., students, courses, and instructors
  - Relationships among entities: students take courses, instructors teach courses
- ER model: model data using entities and relationships
Object Identity and Keys

• In an application, we need to uniquely identify a natural object, and a natural relationship among multiple objects
  – Student: name, address, phone number
  – Course: name, instructor, time
  – Student-take-course: student-id, course-id

• The identities are modeled as keys
Keys

• A **super key** of an entity set is a set of one or more attributes whose values uniquely determine each entity

• A **candidate key** of an entity set is a minimal super key
  – customer_id is a candidate key of customer
  – account_number is a candidate key of account

• One of the candidate keys is selected to be the primary key
Keys for Relationship Sets

• The combination of primary keys of the participating entity sets forms a super key of a relationship set
  – (customer_id, account_number) is the super key of depositor

• Need to consider the semantics of relationship set in selecting the primary key if more than one candidate key is feasible
Keys and Mapping Cardinality

• One to one relationship set
  – Use a candidate key in either entity set
  – University-president (university, president)

• Many to one relationship set
  – Use a candidate key in the many side entity set
  – Teaching (instructor, courses)

• Many to many relationship set
  – Use a candidate key in each participating entity set
  – Take-course (student, course)
E-R Diagrams

- Rectangles represent entity sets
- Diamonds represent relationship sets
- Lines link attributes to entity sets and entity sets to relationship sets
- Ellipses represent attributes
  - Double ellipses represent multivalued attributes
  - Dashed ellipses denote derived attributes
- Underline indicates primary key attributes
Example

- **Entity set**
  - customer
  - borrower
  - loan

- **Attribute**
  - customer_name
  - customer_street
  - loan_number
  - amount
  - customer_id
  - customer_city

- **Relationship set**
  - customer
  - borrower
  - loan
A More Complicated Example

Composite attribute

Multi-valued attribute

Derived attribute
Relationship Sets with Attributes
Summary

• Model real world data using entities and relationships
• The ER model
• ER diagrams
  – Entities, relationships, attributes
  – Constraints, keys, cardinalities
To-Do-List

• Examine the tables in the TPC data set used in assignment 1. Can you guess for each table whether it models an entity set or a relationship set?