The Entity-Relationship Model
ER Model - Part 1: Basics

By Michael Hahsler

Based on slides for CS145 Introduction to Databases (Stanford)
Introduction to Database Design
Database Design

• Database design: Why do we need it?
  • Agree on structure of the database before deciding on a particular implementation

• Consider issues such as:
  • What entities to model
  • How entities are related
  • What constraints exist in the domain
  • How to achieve good designs

• Several formalisms exist
  • We discuss one flavor of ER diagrams
Database Design Process

1. Requirements analysis

• What is going to be stored?
• How is it going to be used?
• What are we going to do with the data?
• Who should access the data?
2. Conceptual Design

- A **high-level description** of the database
- Sufficiently **precise** that technical people can understand it
- But, **not so precise** that non-technical people cannot **participate**
Database Design Process

3. Implementation:

- Logical Database Design
- Physical Database Design
- Security Design
ER is a visual syntax for DB design which is precise enough for technical points, but abstracted enough for non-technical people.
Impact of the ER model

• The ER model is one of the most cited articles in Computer Science
  • “The Entity-Relationship model – toward a unified view of data”
    Peter Chen, 1976

• Used by companies big and small
1. ER Basics: Entities & Relations
Entities and Entity Sets

• **Entities & entity types** are the primitive units of the ER model

  • **Entities** are the individual objects (instances), which are members of entity types
  • **Entity type** are the *classes* or *types* of objects in our model
  • Example: Person is an entity type while Michael is an entity.
  • *We use entity types in ER models*
Entities and Entity Types

- An entity type has **attributes** represented by ovals attached to an entity type

Shapes are **important**. Colors used here are **not**.
Entities vs. Entity Sets

Example:

- **Product**
  - **Name**: Xbox
  - **Category**: Gaming Console
  - **Price**: $250

- **Product**
  - **Name**: My Little Pony Doll
  - **Category**: Toy
  - **Price**: $25

Entities are **not** explicitly represented in ER diagrams!
Keys

A **key** is a **minimal** set of attributes that uniquely identifies an entity.

Denote elements of the primary key by underlining.

Here, \{name, category\} is **not** a key (it is not *minimal*).

*If it were, what would it mean?*

The ER model forces us to designate a single **primary key**, though there may be multiple candidate keys. Often, we introduce an **artificial key** attribute (also called a **synthetic or surrogate key**).
Entity Types Define Relations

<table>
<thead>
<tr>
<th>name</th>
<th>category</th>
<th>price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gizmo</td>
<td>Electronics</td>
<td>$9.99</td>
</tr>
<tr>
<td>GizmoLite</td>
<td>Electronics</td>
<td>$7.50</td>
</tr>
<tr>
<td>Gadget</td>
<td>Toys</td>
<td>$5.50</td>
</tr>
</tbody>
</table>
The R in ER: Relationships

• A relationship type is between two entity types

How to read a relationship in both directions:
1. A product is made by a company
2. A company makes a product
What is a Relationship?

• A mathematical definition (called “Relation” in Math):
  
  • Let $A$, $B$ be sets
    • $A=\{1,2,3\}$, $B=\{a,b,c,d\}$
What is a Relationship?

• **A mathematical definition:**

  • Let A, B be sets
    • $A=\{1,2,3\}, \ B=\{a,b,c,d\}$

  • $A \times B$ (the **cross-product**) is the set of all pairs $(a,b)$
    • $A \times B = \{(1,a), (1,b), (1,c), (1,d), (2,a), (2,b), (2,c), (2,d), (3,a), (3,b), (3,c), (3,d)\}$
What is a Relationship?

• A mathematical definition:
  
  • Let $A$, $B$ be sets
    • $A = \{1, 2, 3\}$, $B = \{a, b, c, d\}$
  
  • $A \times B$ (the cross-product) is the set of all pairs $(a, b)$
    • $A \times B = \{(1,a), (1,b), (1,c), (1,d), (2,a), (2,b), (2,c), (2,d), (3,a), (3,b), (3,c), (3,d)\}$
  
  • We define a relationship (relation) to be a subset of $A \times B$
    $R = \{(1,a), (2,c), (2,d), (3,b)\}$
What is a Relationship?

• **A mathematical definition:**
  • Let A, B be sets
  • A x B (the *cross-product*) is the set of all pairs
  • A relationship (relation) is a subset of A x B

• **Example:**Makes is a relationship. It is a *subset* of Product × Company:
What is a Relationship?

A relationship between entity sets **P** and **C** is a *subset of all possible pairs of entities in P and C*, with tuples uniquely identified by **P** and **C**’s keys.
A **relationship** between entity sets **P** and **C** is a **subset of all possible pairs of entities in P and C**, with tuples uniquely identified by **P** and **C**’s keys.
What is a Relationship?

<table>
<thead>
<tr>
<th>Company</th>
<th>Product</th>
<th>price</th>
<th>C.name</th>
<th>P.name</th>
<th>P.category</th>
<th>P.price</th>
</tr>
</thead>
<tbody>
<tr>
<td>GizmoWorks</td>
<td>GizmoLite</td>
<td>$7.50</td>
<td>GazmoWorks</td>
<td>GizmoLite</td>
<td>Electronics</td>
<td>$7.50</td>
</tr>
<tr>
<td>GizmoWorks</td>
<td>Gadget</td>
<td>$5.50</td>
<td>GazmoWorks</td>
<td>Gadget</td>
<td>Toys</td>
<td>$5.50</td>
</tr>
<tr>
<td>GadgetCorp</td>
<td>GizmoLite</td>
<td>$7.50</td>
<td>GadgetCorp</td>
<td>GizmoLite</td>
<td>Electronics</td>
<td>$7.50</td>
</tr>
<tr>
<td>GadgetCorp</td>
<td>Gadget</td>
<td>$5.50</td>
<td>GadgetCorp</td>
<td>Gadget</td>
<td>Toys</td>
<td>$5.50</td>
</tr>
</tbody>
</table>

A **relationship** between entity sets **P and C** is a *subset of all possible pairs of entities in P and C*, with tuples uniquely identified by **P and C’s keys**.
What is a Relationship?

• There can only be one relationship for every unique combination of entities

• This also means that the relationship is uniquely determined by the keys of its entities

• Example: the key for Makes (to right) is {Product.name, Company.name}

This follows from our mathematical definition of a relationship (it is a set)
Relationships and Attributes

Relationships may have attributes as well.

For example: “since” records when company started making a product.

Note: For each product/company pair there is automatically only a single since value since there can only be one unique product/company pair in makes.
Q: What does this say?

A: A person can only buy a specific product once per day (date)

Modeling something as a relationship makes it unique. **What if this is not appropriate?**
Decision: Relationship vs. Entity?

What about this way?

Now we can have multiple purchases per product, person pair!

We can always use a new entity instead of a relationship. For example, to permit multiple instances of each entity combination!
Note on Relationships vs. Relation

**ER Model:** How do Entity types relate to each other

**Math:** A Relation (= a subset of the cross product)

**Relational Algebra:** A table with data (a set)

Relations are used to implement entity types and certain relationship types!
How to Create an ER Diagram

“Rules of thumb" for mapping natural language descriptions into ER diagrams:

<table>
<thead>
<tr>
<th>English grammar structure</th>
<th>ER structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common noun</td>
<td>Entity type</td>
</tr>
<tr>
<td>Proper noun</td>
<td>Entity</td>
</tr>
<tr>
<td>Verb</td>
<td>Relationship type</td>
</tr>
<tr>
<td>Adjective</td>
<td>Attribute for entity</td>
</tr>
<tr>
<td>Adverb</td>
<td>Attribute for relationship</td>
</tr>
</tbody>
</table>
Example: How to Create an ER Diagram

Here is what the person in charge said:

“Our company is called PowerSeller and we sell health products on Ebay. Our products are made by different manufacturers. Products belong to different product categories (e.g., supplements, cosmetics, etc.) and each product is sold at a fixed price. We use customer IDs for our customers, and we know for all of them the shipping address and name, and for most we also know a phone number.”
Example: How to Create an ER Diagram

Here is what the person in charge said:

“Our company is called PowerSeller and we sell health products on Ebay. Our products are made by different manufacturers. Products belong to different product categories (e.g., supplements, cosmetics, etc.) and each product is sold at a fixed price. We use customer IDs for our customers, and we know for all of them the shipping address and name, and for most we also know a phone number.”
Example: How to Create an ER Diagram

*Identify entity types and attributes (nouns):*
- company, PowerSeller
- health product, product, product category, price, manufacturer
- customerID, customer, shipping address, name, phone number

*Identify relationship types (verbs):*
- know
- belongs to
- sell
- make
Example: How to Create an ER Diagram

*Identify Entity types (bold) and attributes:*

- company, PowerSeller
- health product, **product**, product category, price, **manufacturer**
- customer ID, **customer**, shipping address, name, phone number

*Identify relationship types:*

- know
- belongs to
- sell, buy
- make
Product

- name
- category
- price

make

Manufacturer

name

Customer

- phone
- address
- name
- ID

buy
Exercise: Draw an ER diagram for football

Teams play each other in Games. Each pair of teams can play each other multiple times.

Players belong to Teams (assume no trades / changes).

A Game is made up of Plays that result in a yardage gain/loss, and potentially a touchdown.

A Play will contain either a Pass from one player to another, or a Run by one player.