

# Introduction to Data Management

## CSE 344

### Lecture 16: E/R Diagrams and Constraints

# Announcements

- HW6, WQ6 are out
  - Both due 2/27
- One more HW after that

# Midterm

- Stats:
  - Mean: 73
  - Standard deviation: 13
- Check your UW email to access gradescope
- Solutions posted under “exams” on website
- Retrieve paper copies of your exam from CSE front desk until 3/1 (after which they will be recycled)
- Regrade policy:
  - Submit on gradescope by 2/22
  - Clearly state how you were misgraded
  - We will regrade the entire problem after deadline

# Ask the staff...

- TA lectures not very effective
- Taping lectures
- Course grading
  
- Reminder: please sit in the back if you use your laptop to take notes
  - But please not sit in the last 3 rows
  
- Reminder: Keep comments coming via piazza / feedback link on website!

# Meanwhile in data management...

# Welcome to the 2nd half of 344

- Relational data model
    - Instance
    - Schema
    - Query languages
      - SQL, RA, RC, Datalog
  - Query processing
    - Logical & physical plans
    - Indexes
    - Cost estimation
    - Query optimization
  - Non-relational data model
- Conceptual design
    - E/R diagrams
    - Converting to SQL
    - Normalization
- Transactions
    - ACID
    - Transaction Implementation
    - Writing DB applications
  - Parallel query processing
    - MapReduce
    - Spark

# Database Design

What it is:

- Starting from scratch, design the database schema: relation, attributes, keys, foreign keys, constraints etc

Why it's hard

- The database will be in operation for a very long time (years). Updating the schema while in production is very expensive (why?)

# Database Design

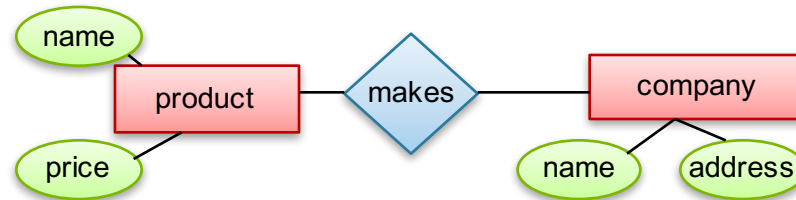
- Consider issues such as:
  - What entities to model
  - How entities are related
  - What constraints exist in the domain
- Several formalisms exists
  - We discuss E/R diagrams
  - UML, model-driven architecture
- Reading: Sec. 4.1-4.6



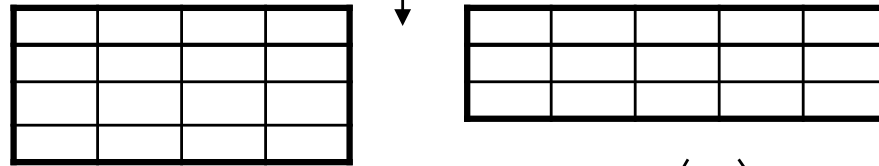


# Database Design Process

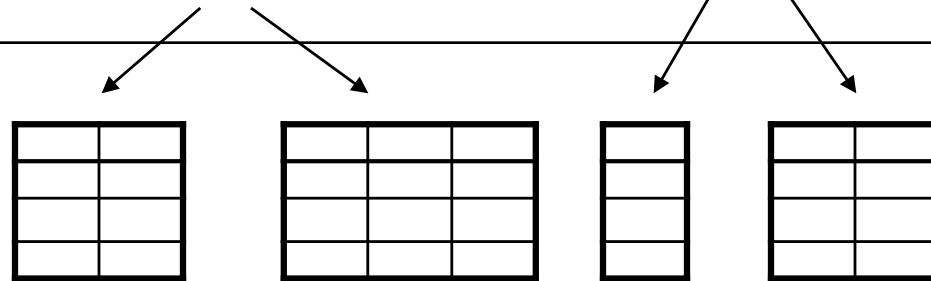
Conceptual Model:



Relational Model:  
Tables + constraints  
And also functional dep.



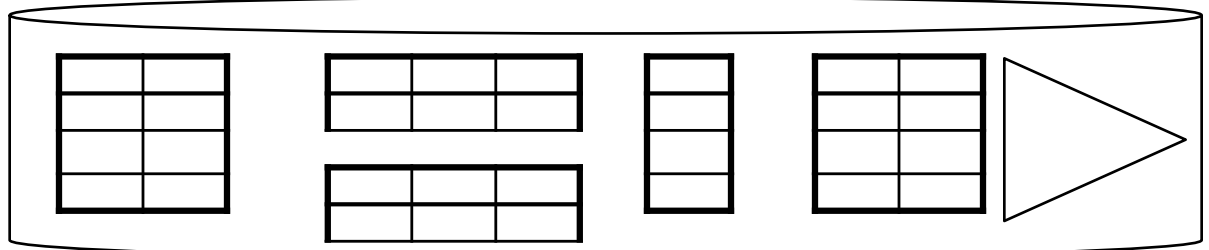
Normalization:  
Eliminates anomalies



Conceptual Schema

Physical storage details

Physical Schema



# Entity / Relationship Diagrams

- Entity set = a class
  - An entity = an object

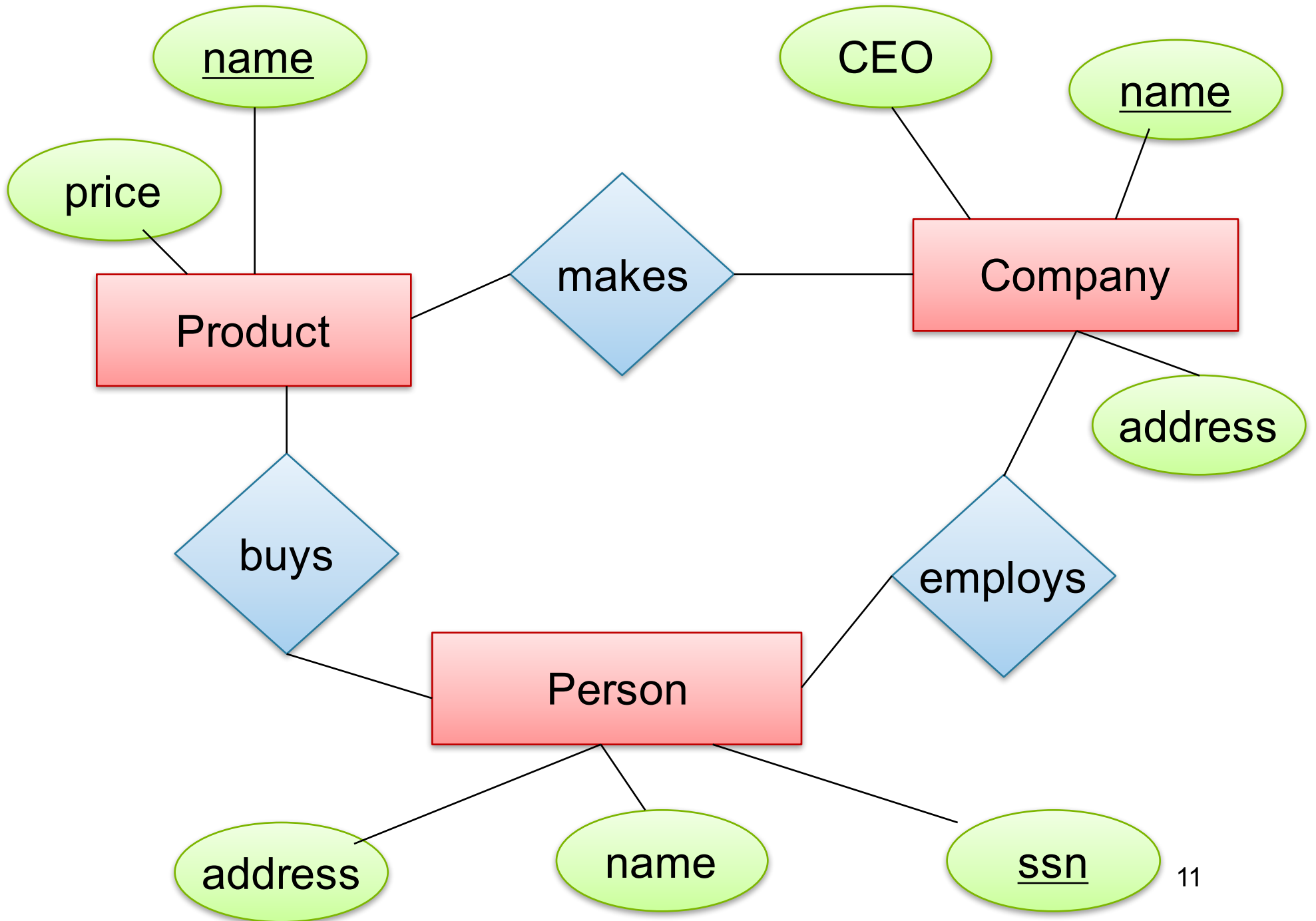


- Attribute



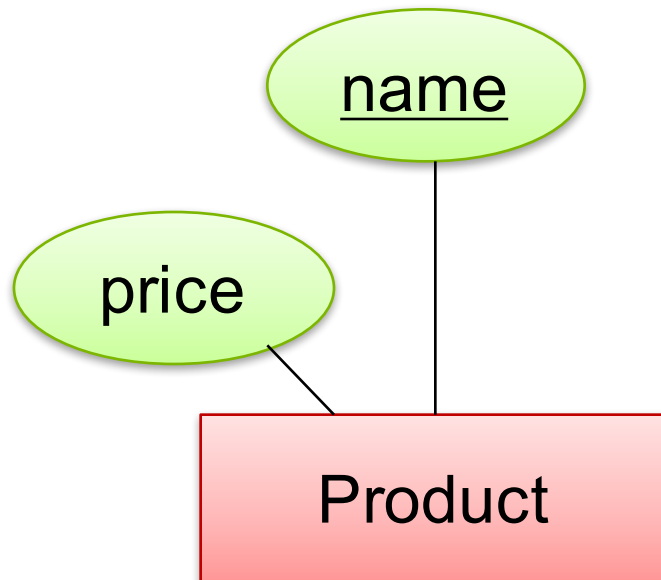
- Relationship





# Keys in E/R Diagrams

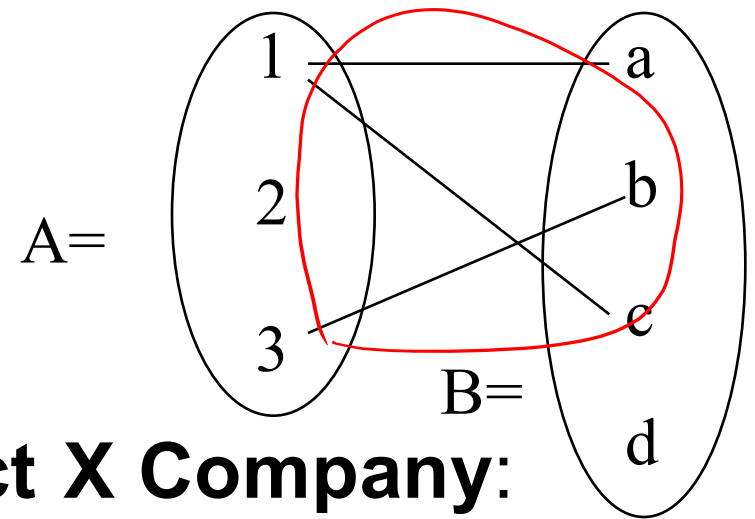
- Every entity set must have a key



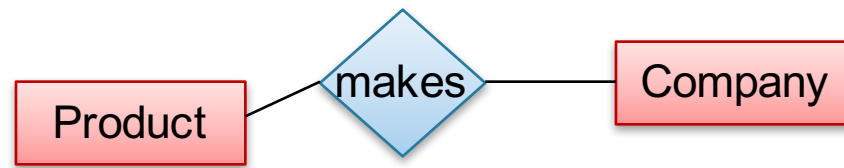
# What is a Relation ?

- A mathematical definition:
  - if A, B are sets, then a relation R is a subset of  $A \times B$

- $A = \{1, 2, 3\}$ ,  $B = \{a, b, c, d\}$ ,  
 $A \times B = \{(1, a), (1, b), \dots, (3, d)\}$   
 $R = \{(1, a), (1, c), (3, b)\}$

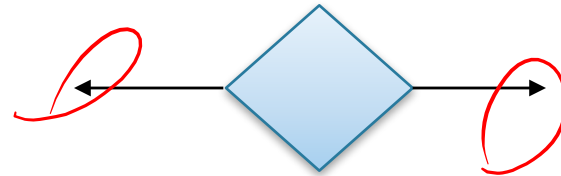
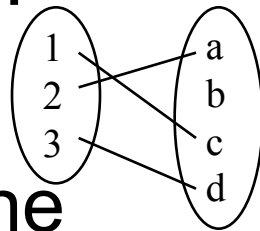


- **makes** is a subset of **Product X Company**:

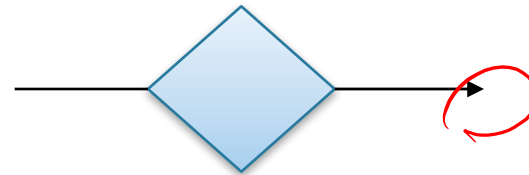
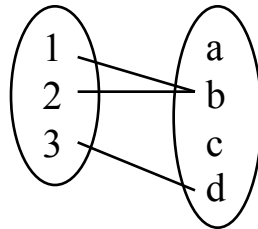


# Multiplicity of E/R Relations

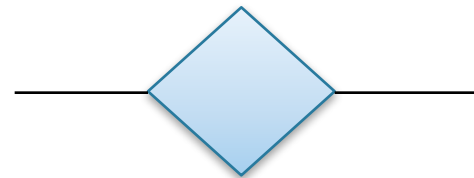
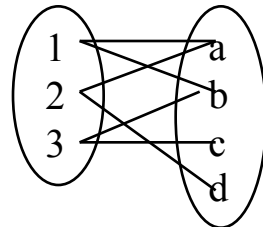
- one-one:

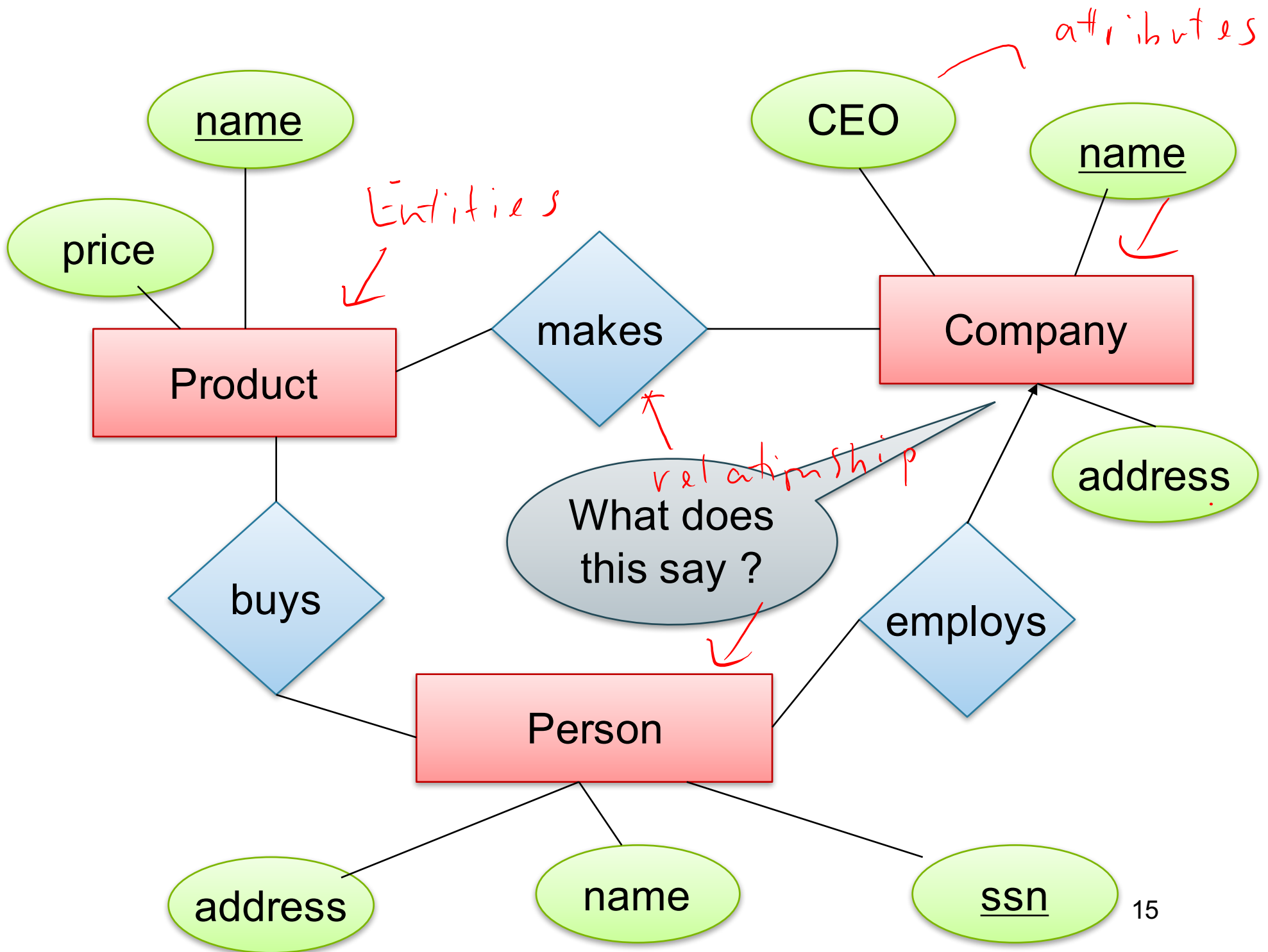


- many-one



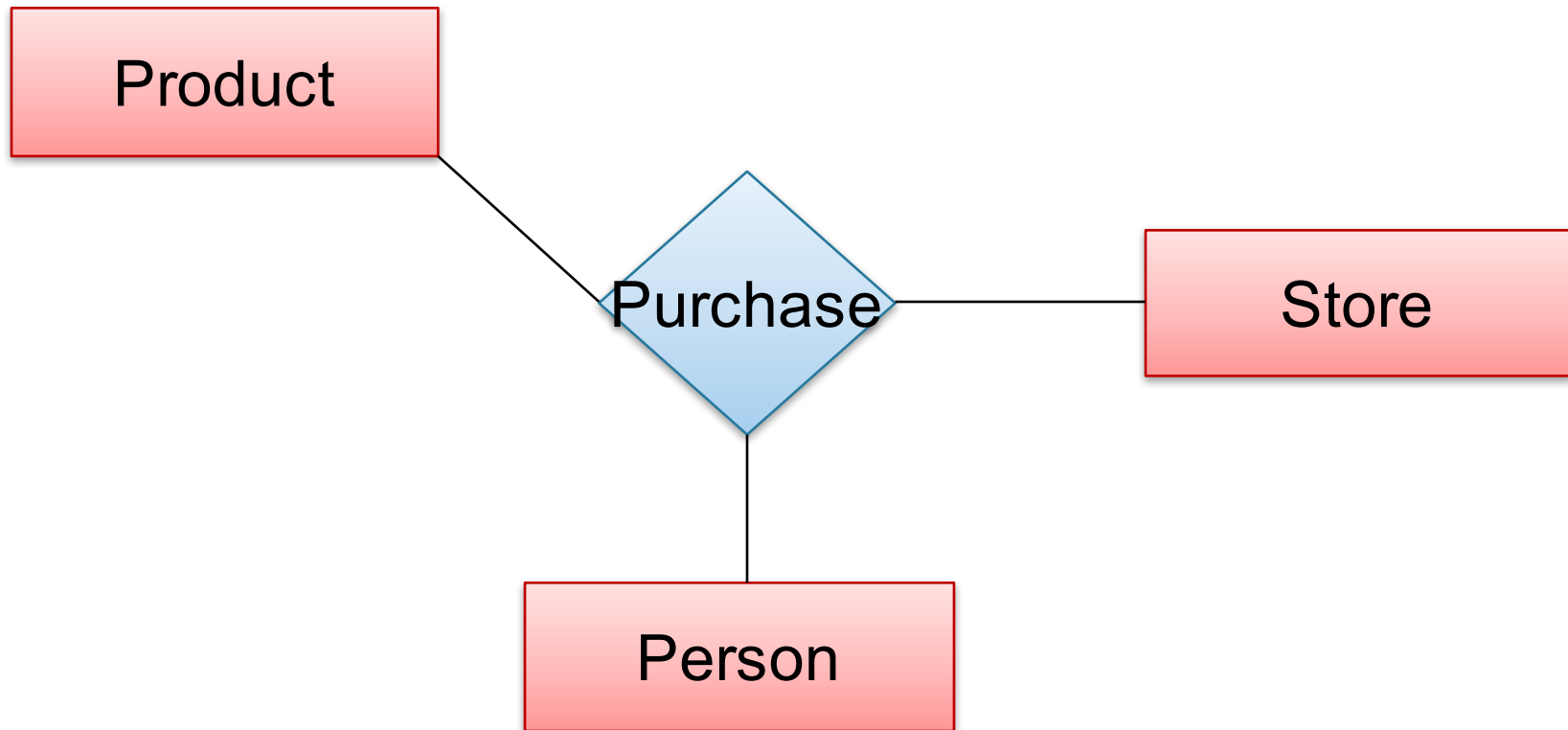
- many-many





# Multi-way Relationships

How do we model a purchase relationship between buyers, products and stores?



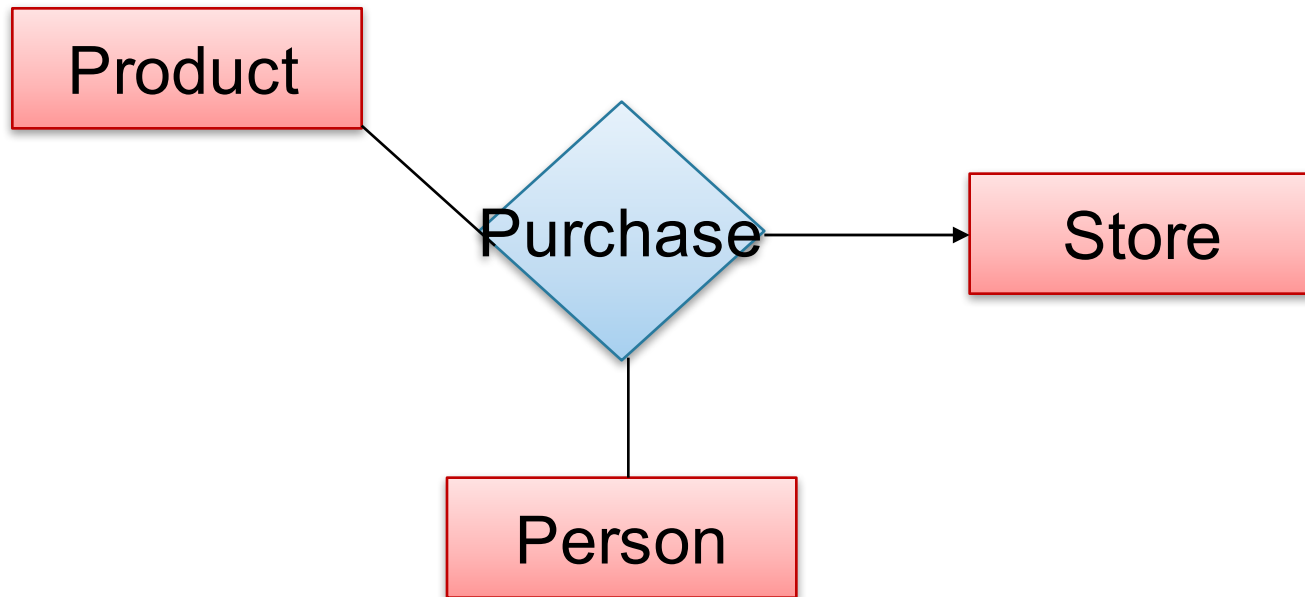
Can still model as a mathematical set (How ?)

As a set of triples  $\subseteq \text{Person} \times \text{Product} \times \text{Store}$



# Arrows in Multiway Relationships

**Q:** What does the arrow mean ?

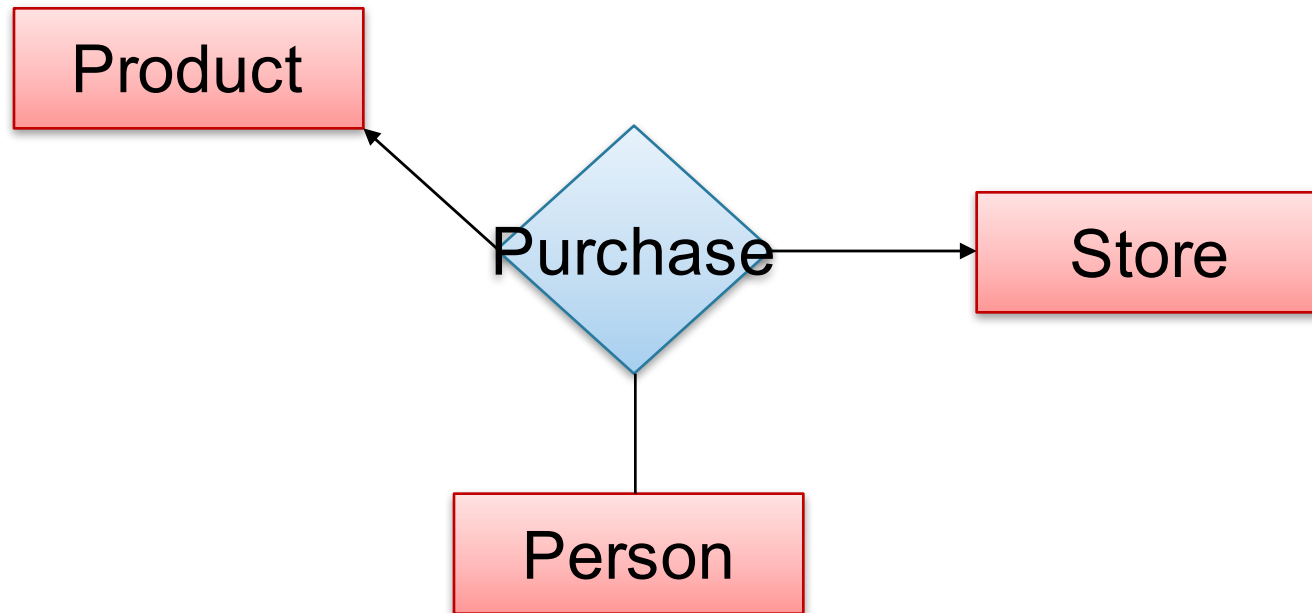


**A:** A given person buys a given product from at most one store

[Fine print: Arrow pointing to E means that if we select one entity from each of the other entity sets in the relationship, those entities are related to at most one entity in E]

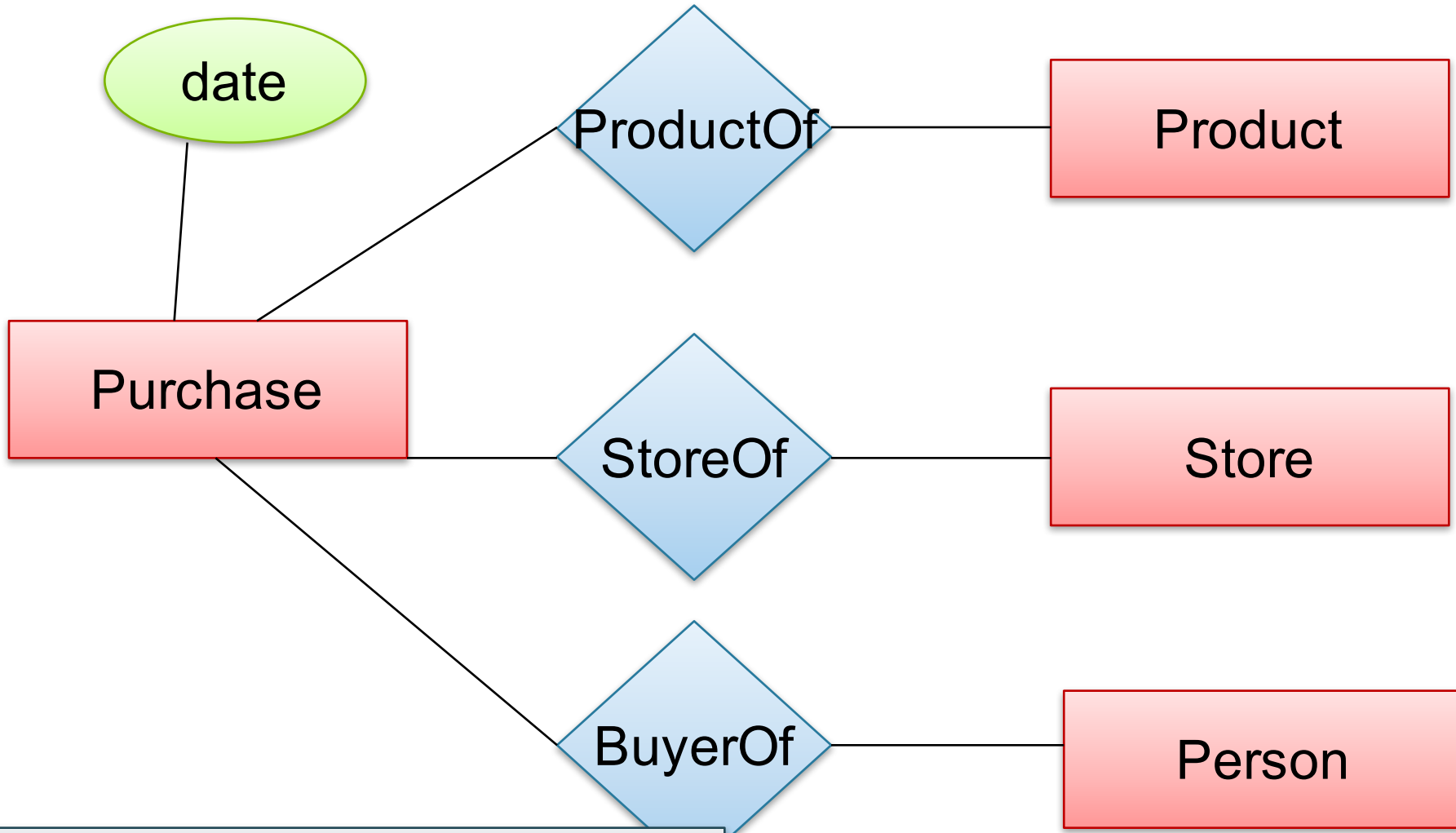
# Arrows in Multiway Relationships

**Q:** What does the arrow mean ?



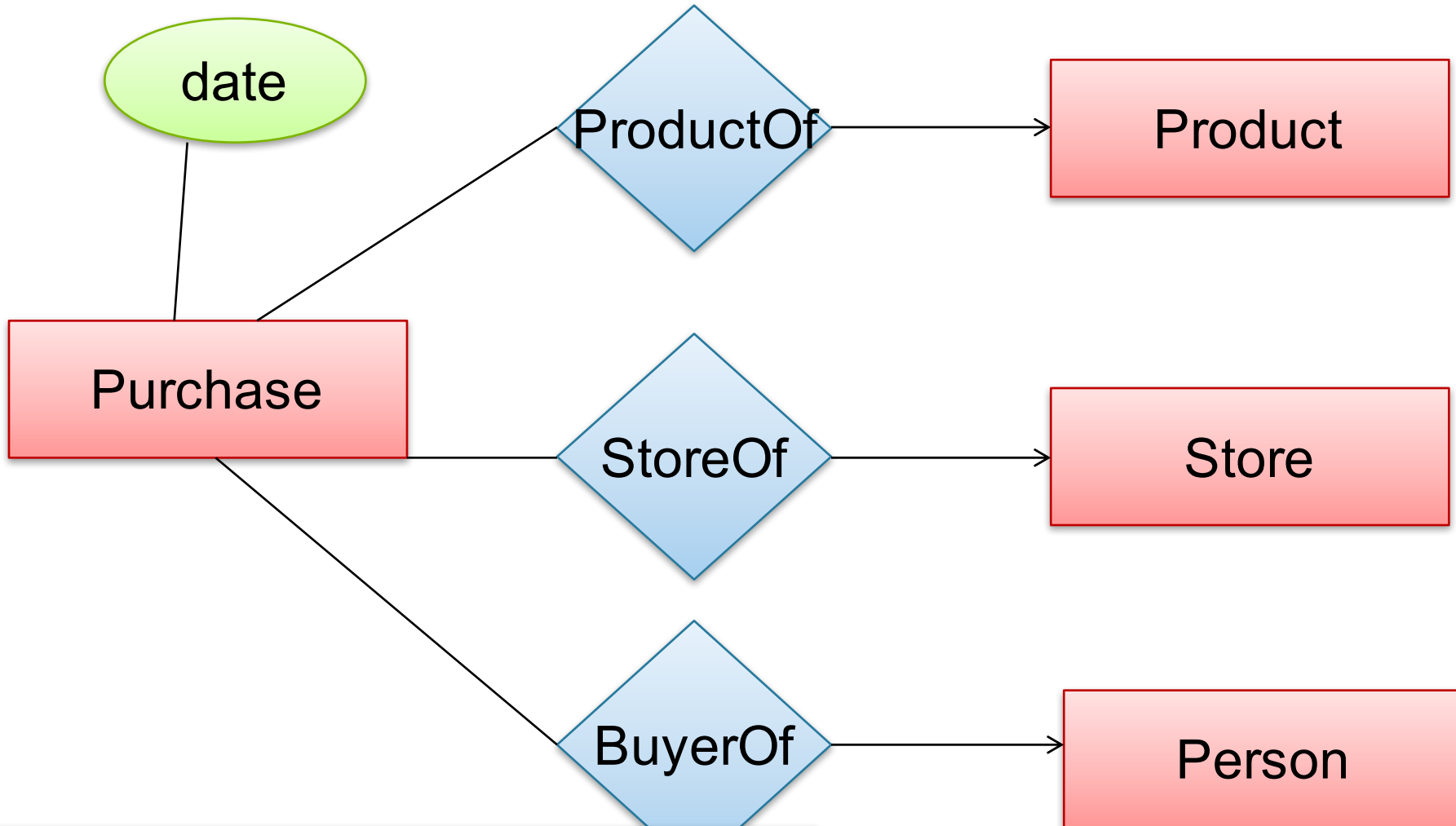
**A:** A given person buys a given product from at most one store  
AND every store sells to every person at most one product

# Converting Multi-way Relationships to Binary



Arrows go in which direction?

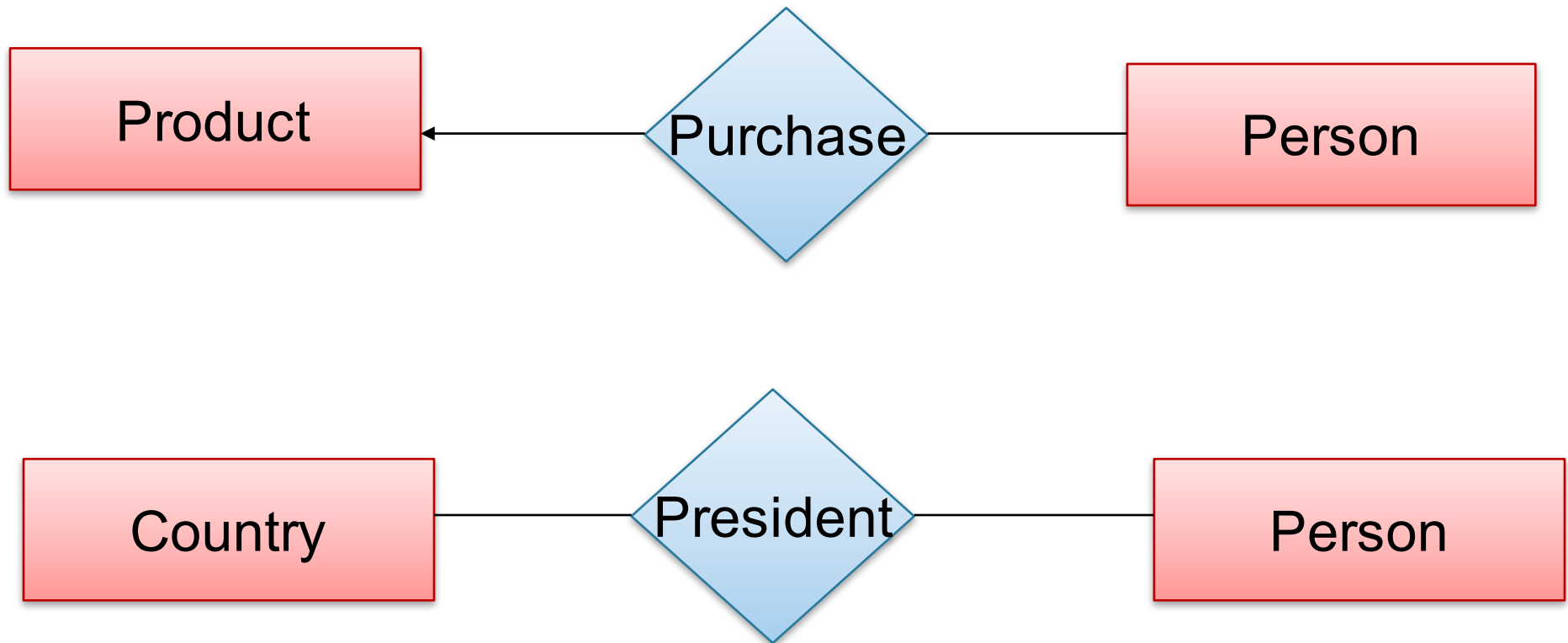
# Converting Multi-way Relationships to Binary



Make sure you understand why!

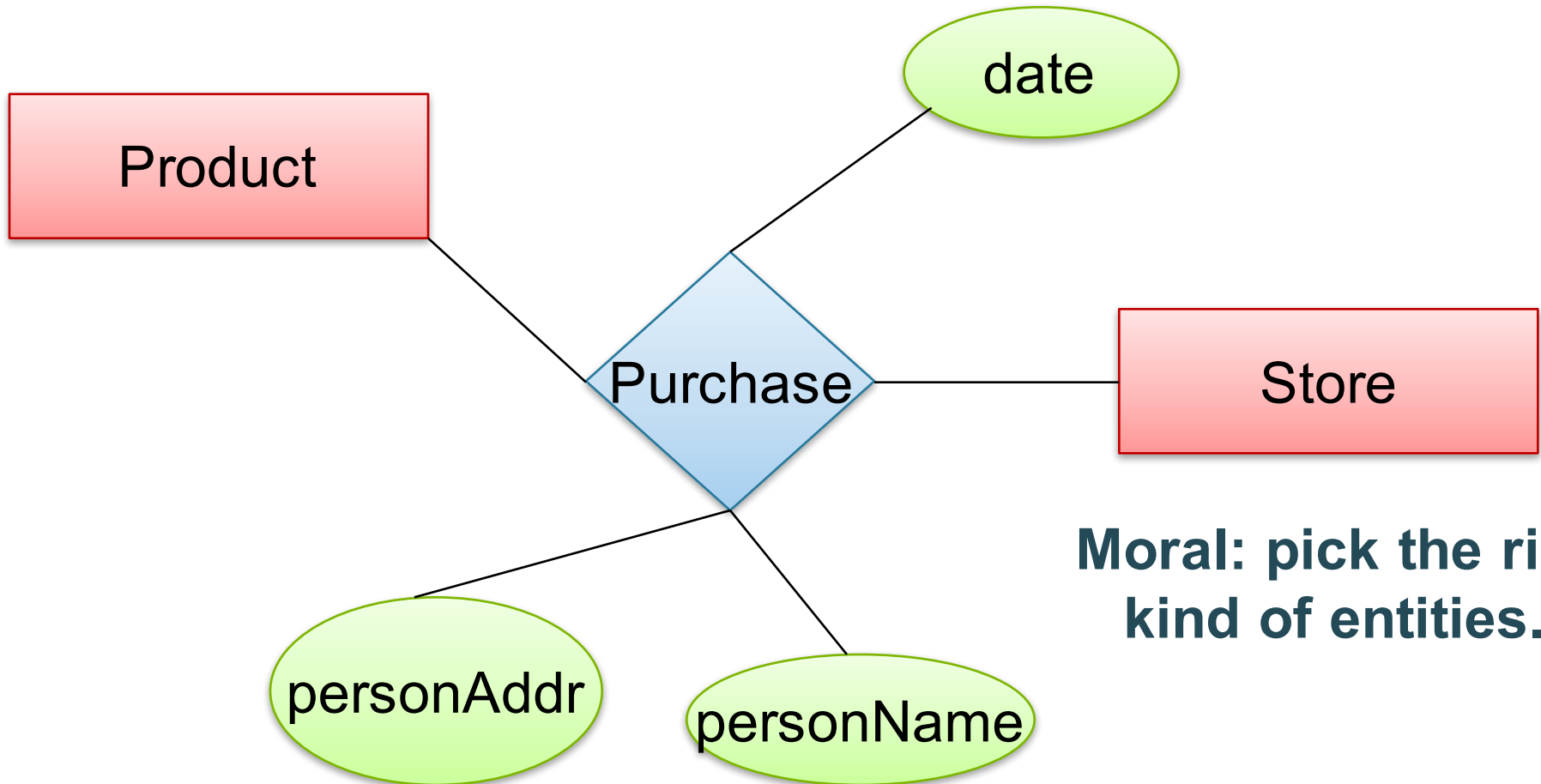
# 3. Design Principles

**What's wrong?**



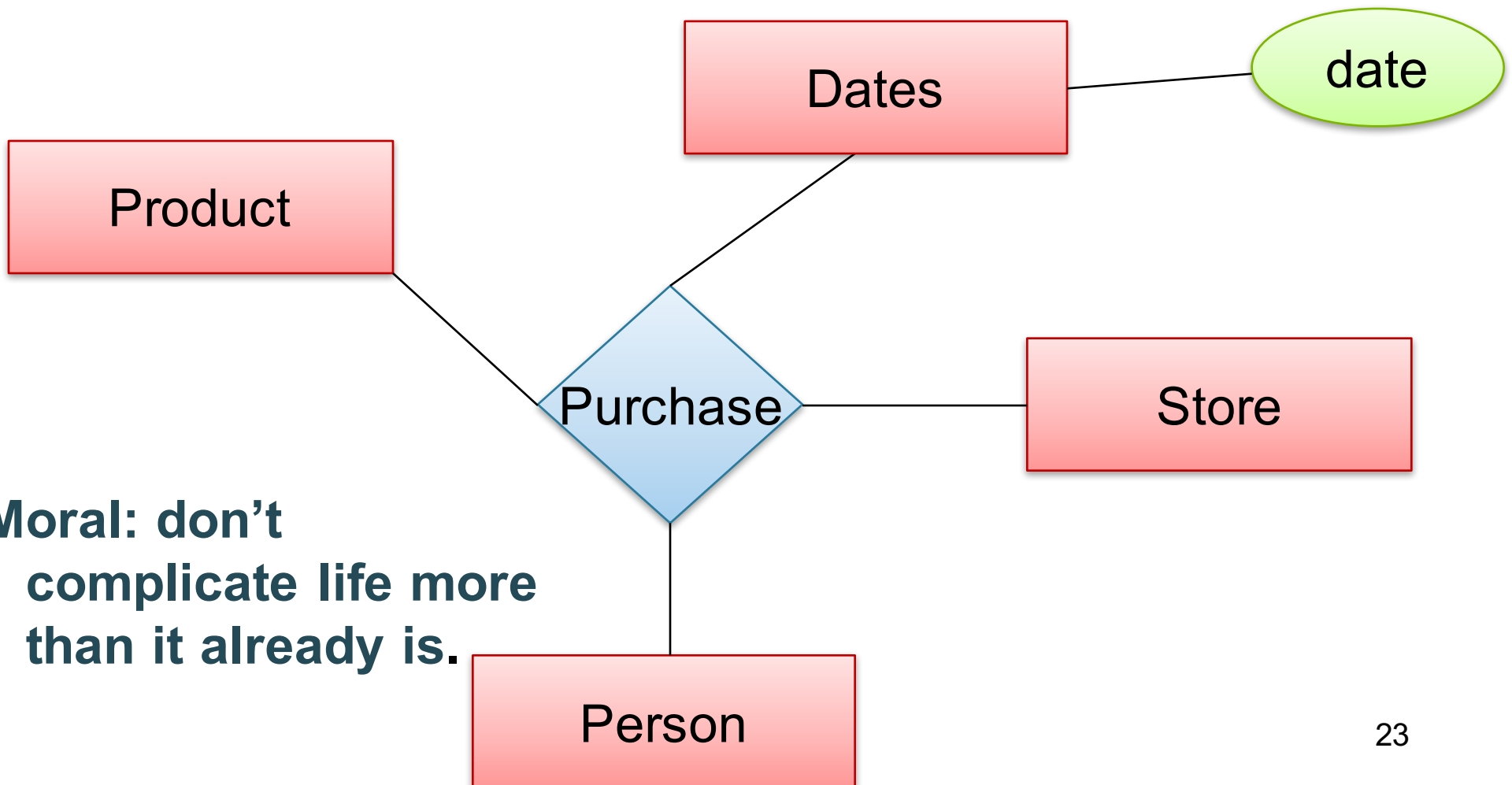
**Moral: Be faithful to the specifications of the application!**

# Design Principles: What's Wrong?



**Moral: pick the right  
kind of entities.**

# Design Principles: What's Wrong?



**Moral: don't  
complicate life more  
than it already is.**

# From E/R Diagrams to Relational Schema

- Entity set  $\rightarrow$  relation
- Relationship  $\rightarrow$  relation