Introduction to Data Management
CSE 344

Lecture 15: E/R Diagrams

Announcements

- Web quiz due Tuesday (tomorrow), 11 pm
- Homework 4 due Thursday, 11 pm
- Midterm: next Monday, November 3, in class
  - Everything through datalog/ra/rc. XML will be on the final (in sections this week, quiz+hw due after midterm)
  - Review Sunday, Nov. 2 at 2 pm, EEB 037
  - Open book or not open book? That is the question.
- CSE talk tomorrow, 3:30, atrium. Hadi Partovi, code.org
- Today: E/R diagrams (4.1-4.6)

Today: E/R Diagrams

Motivating scenario: your boss asks you to build a DBMS describing:

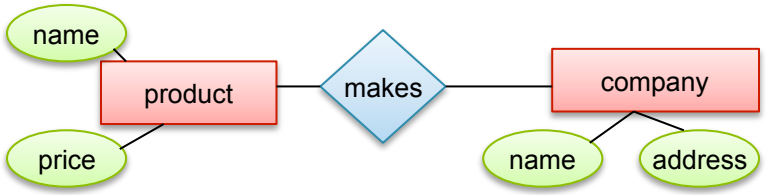
- Companies. Each company has:
  - A name, an address, and a CEO
  - A list of employees, with ssn, name, and address
- Products manufactured by these companies
  - Each product has a name and a price
  - The same product may be manufactured by several companies
- Buyers of these products
  - Each buyer has an ssn, name, and address
  - Some employees may be buyers too

Database Design

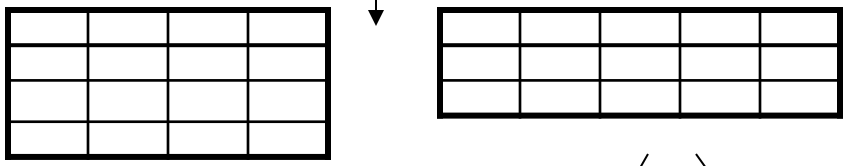
- Why do we need it?
  - Need a way to model real world entities in terms of relations
  - Not easy to go from real-world entities to a database schema
- Consider issues such as:
  - What entities to model
  - How entities are related
  - What constraints exist in the domain
  - How to achieve ***good*** designs
- Several formalisms exists
  - We discuss E/R diagrams

Database Design Process

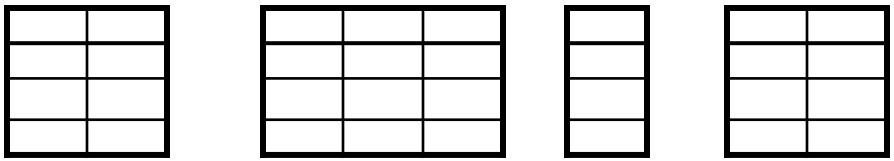
Conceptual Model:



Relational Model:
Tables + constraints
And also functional dep.

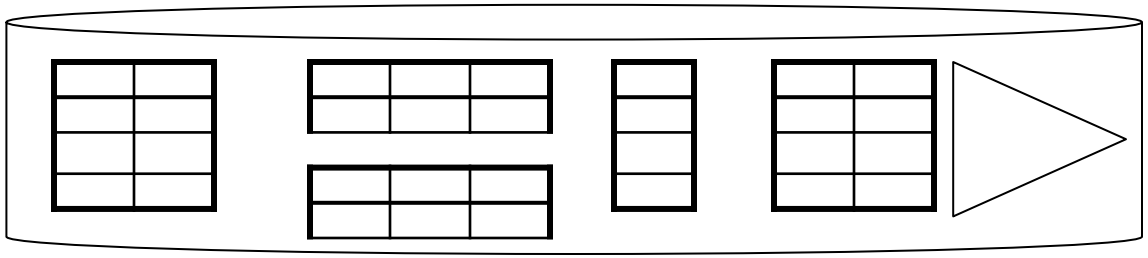


Normalization:
Eliminates anomalies



Conceptual Schema

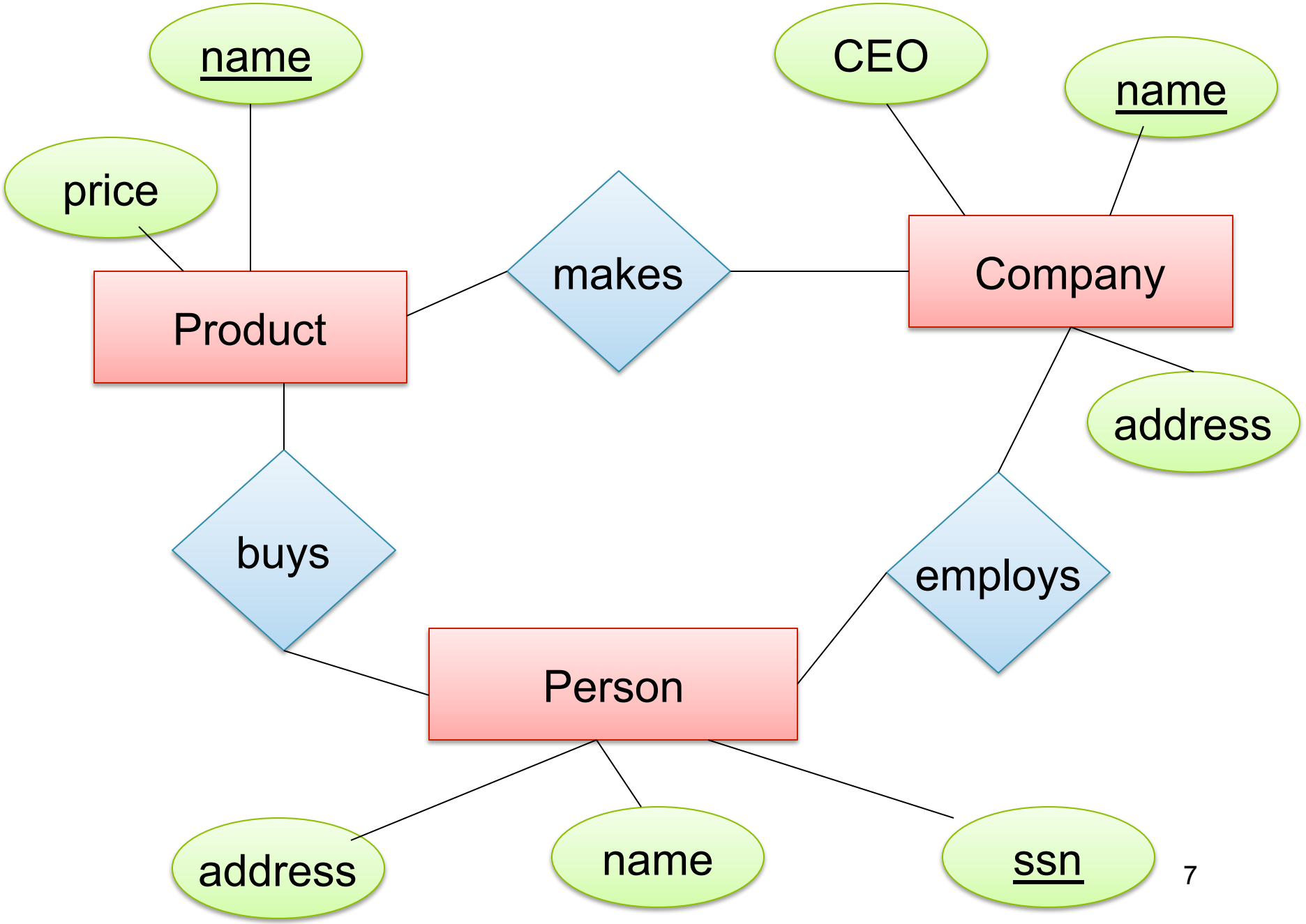
Physical storage details



Physical Schema

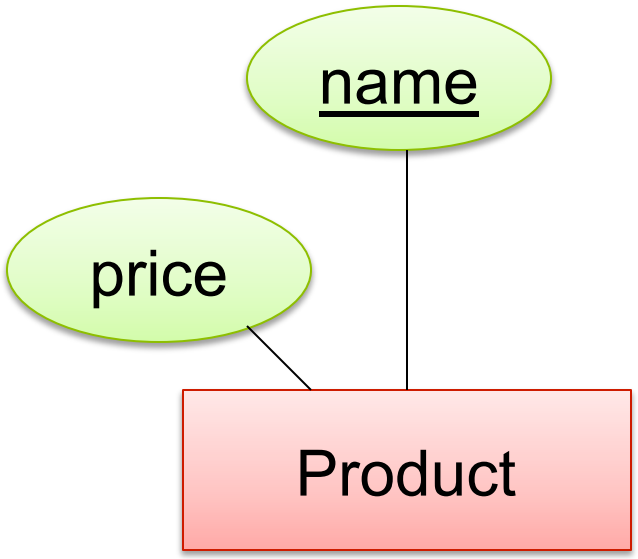
Entity / Relationship Diagrams

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


name
price
Product
makes
CEO
name
Company
address
buys
employs
Person
address
name
ssn

Keys in E/R Diagrams

- Every entity set must have a key

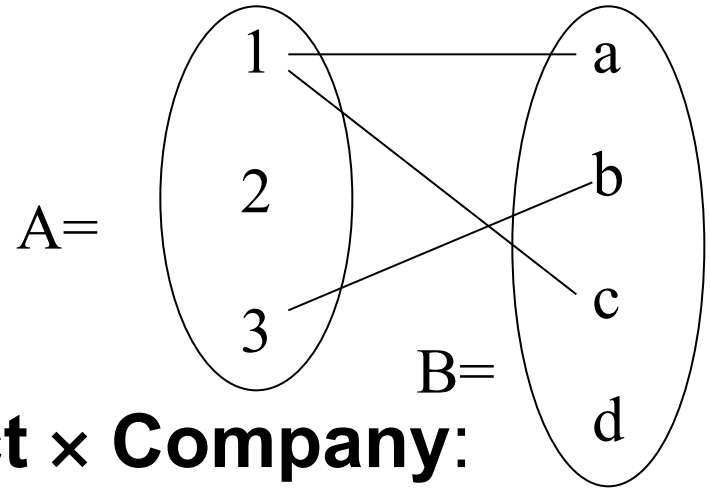


What is a Relation ?

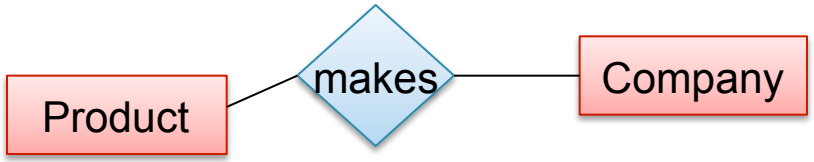
- A mathematical definition:
  - if A, B are sets, then a relation R is a subset of A × B
- A={1,2,3}, B={a,b,c,d},

  A × B = {(1,a),(1,b), . . ., (3,d)}

  R = {(1,a), (1,c), (3,b)}

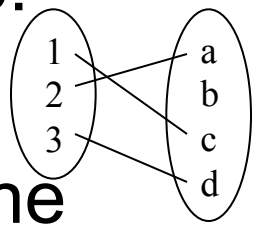


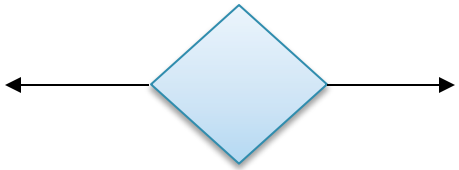
- **makes** is a subset of **Product** × **Company**:



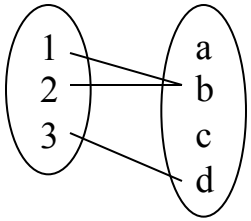
Multiplicity of E/R Relations

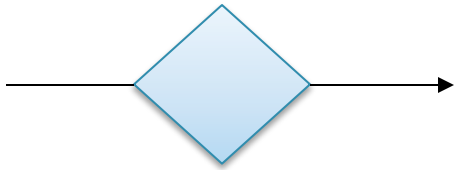
- one-one:



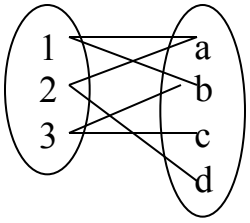


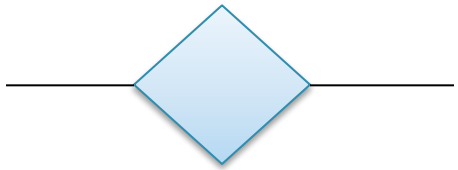
- many-one

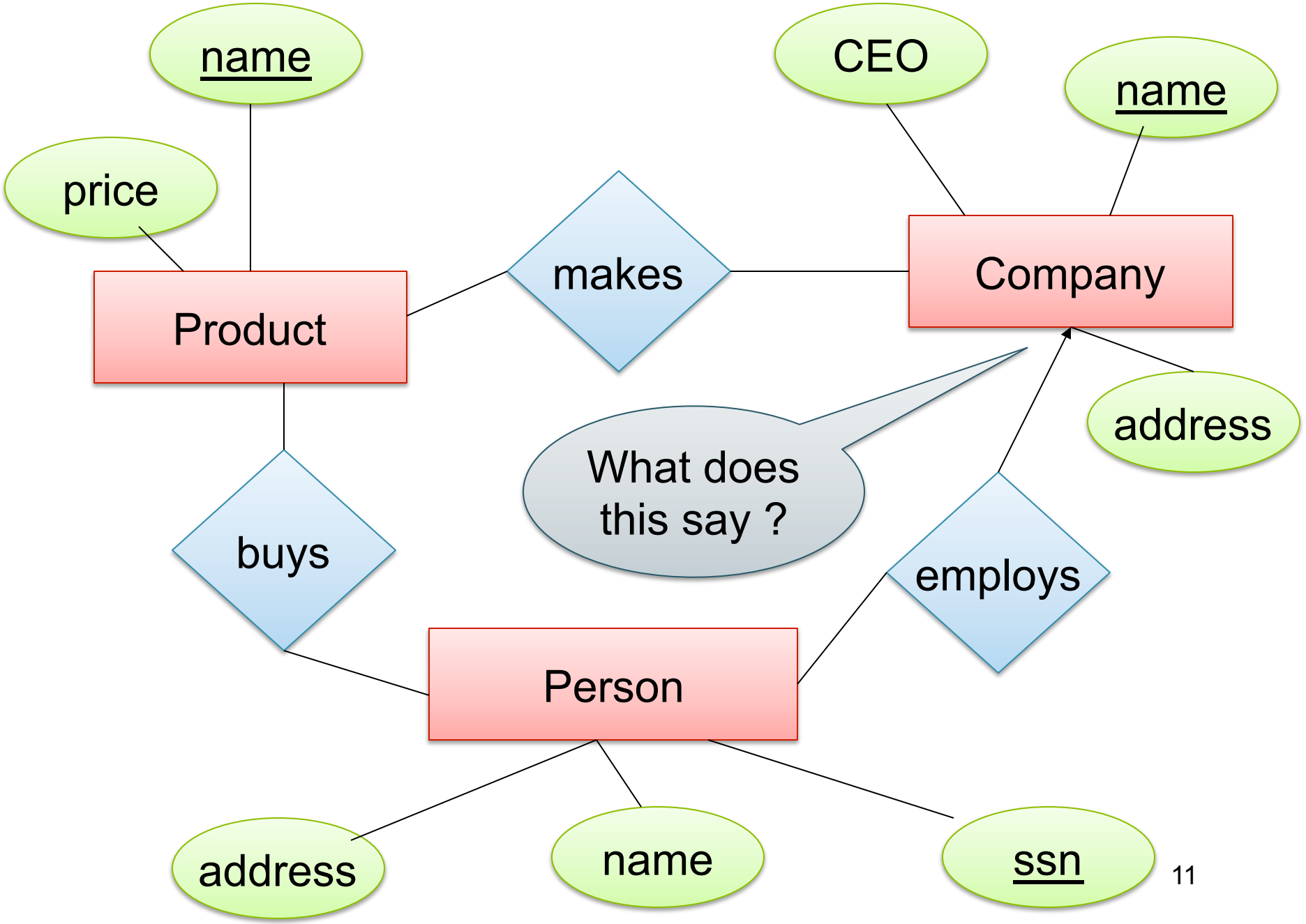




- many-many

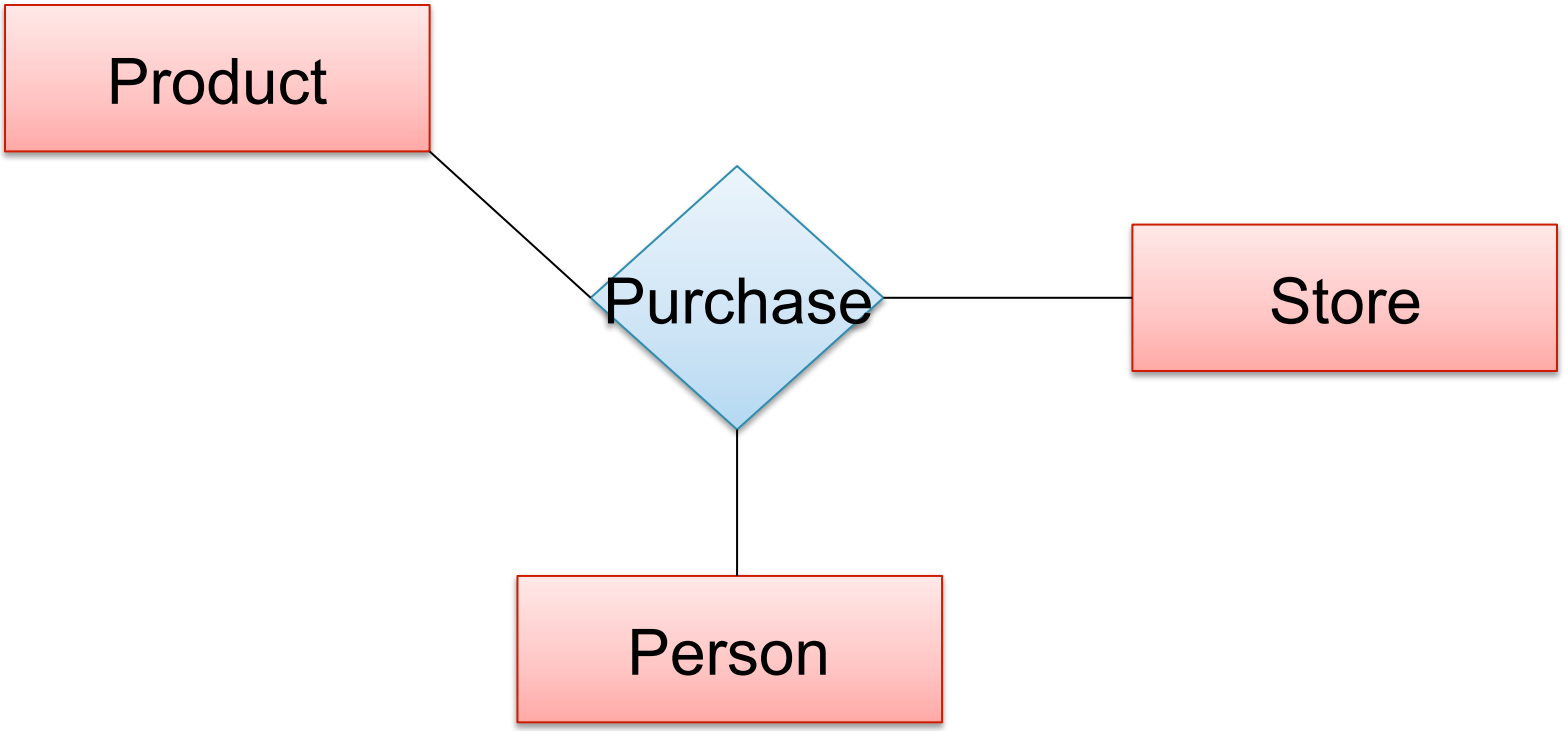





name
price
Product
makes
CEO
name
Company
address
What does this say ?
buys
employs
Person
address
name
ssn

Multi-way Relationships

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

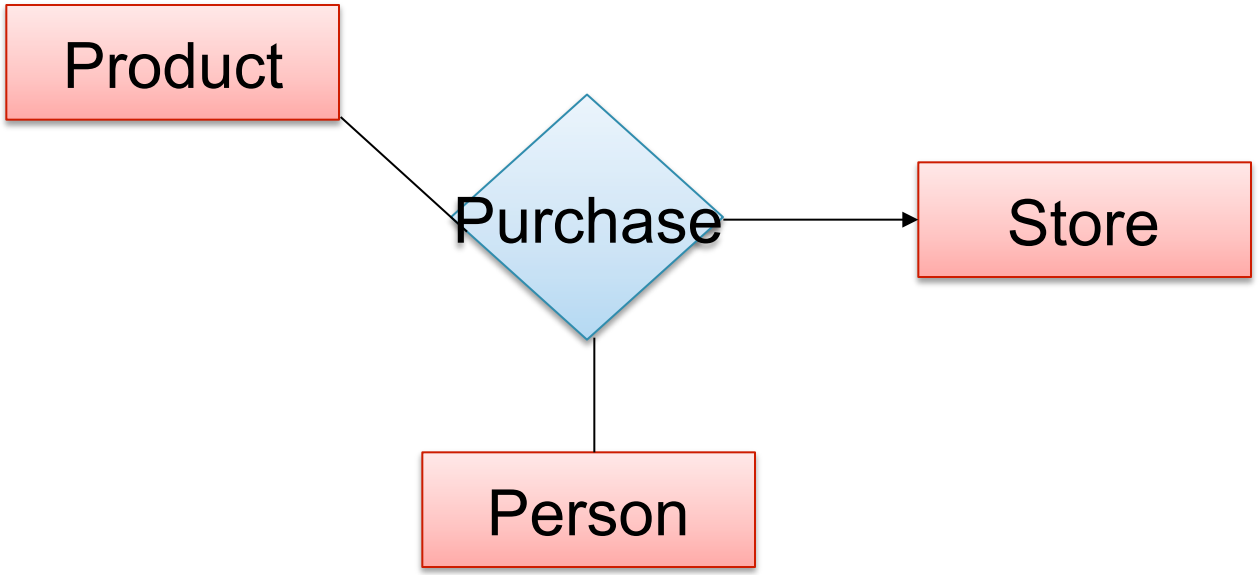


Can still model as a mathematical set (Q. how ?)

A. As a set of triples $\subseteq$ Person $\times$ Product $\times$ Store

Arrows in Multiway Relationships

Q: What does the arrow mean ?

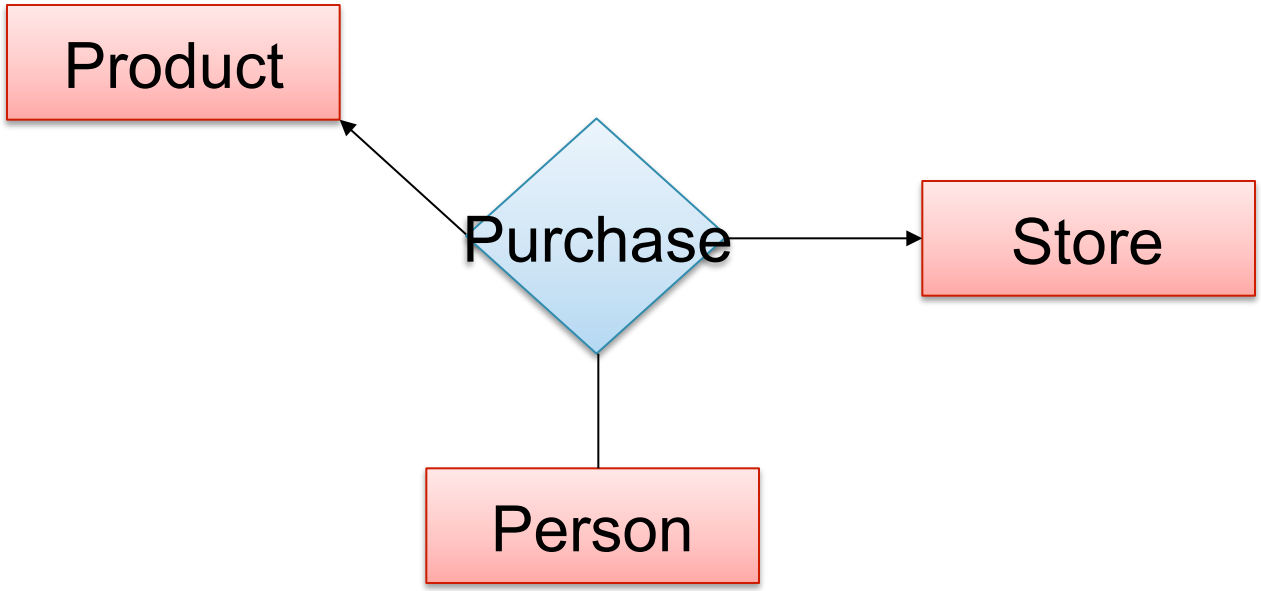


A: A given person buys a given product from at most one store

[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

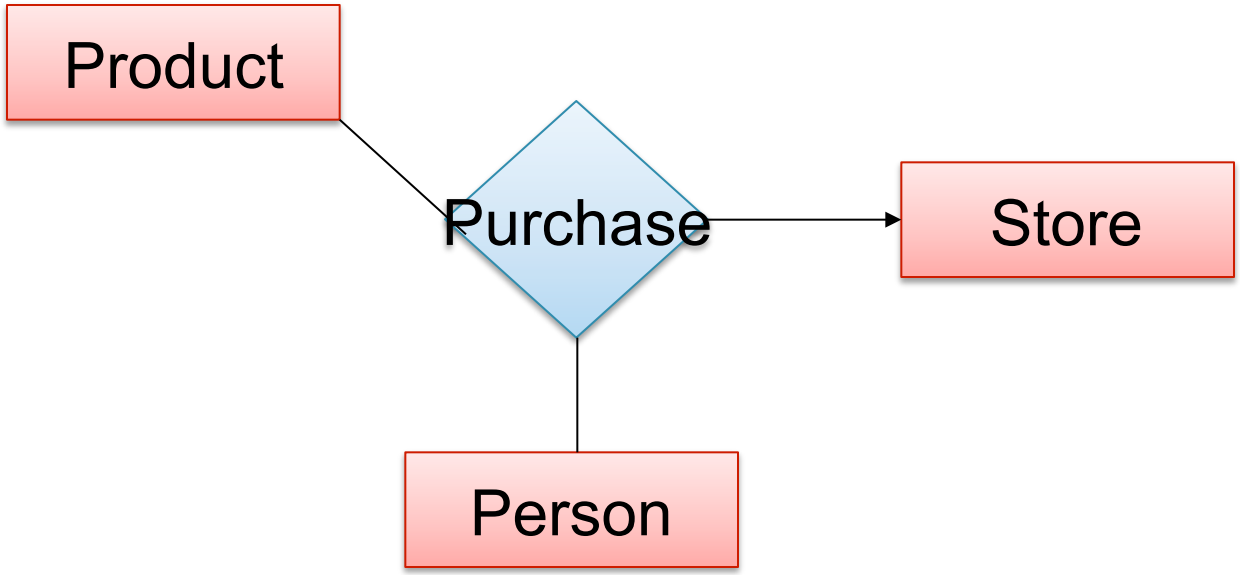
Q: What does the arrow mean ?



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

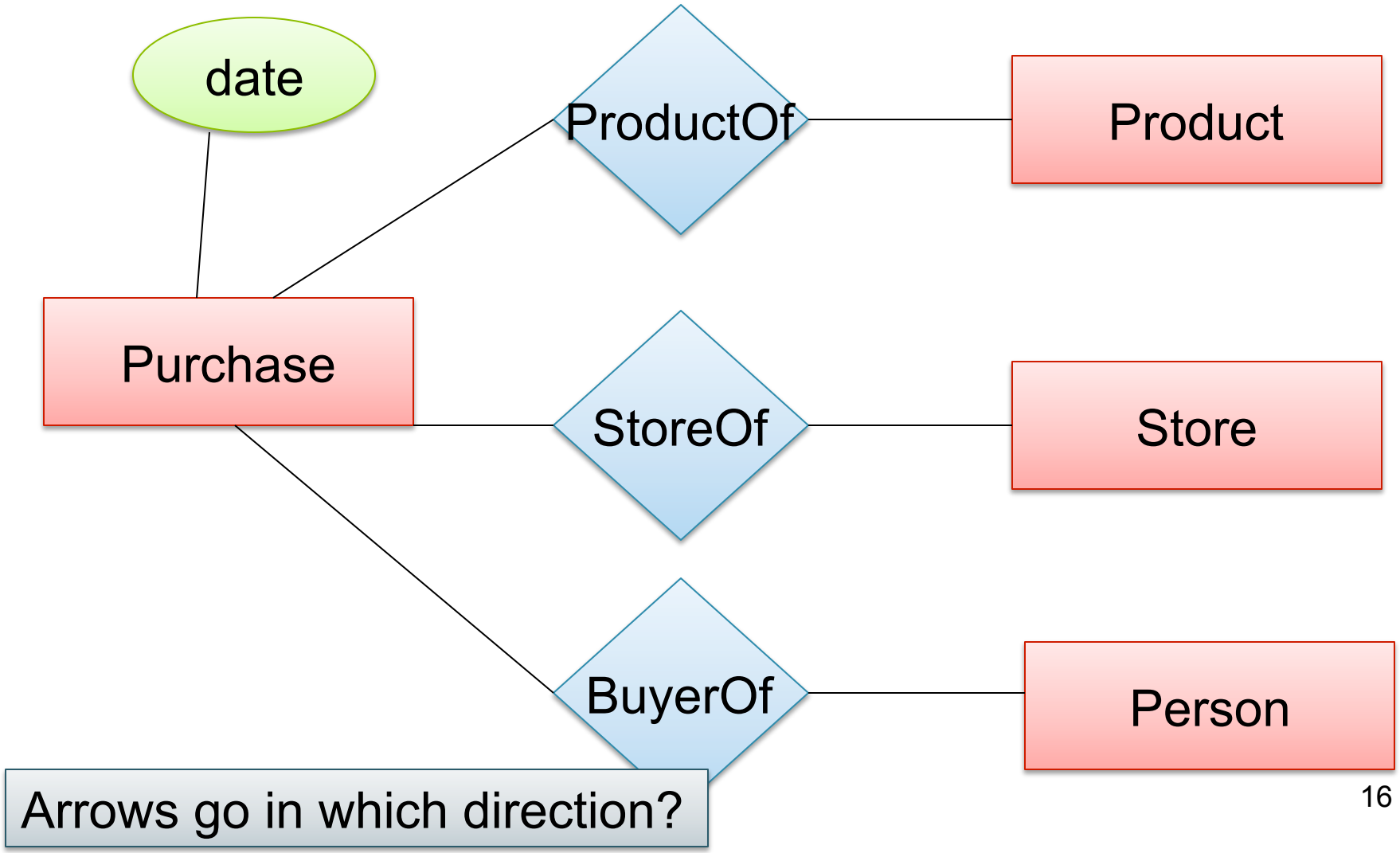
Arrows in Multiway Relationships

Q: How do we say that every person shops at at most one store?

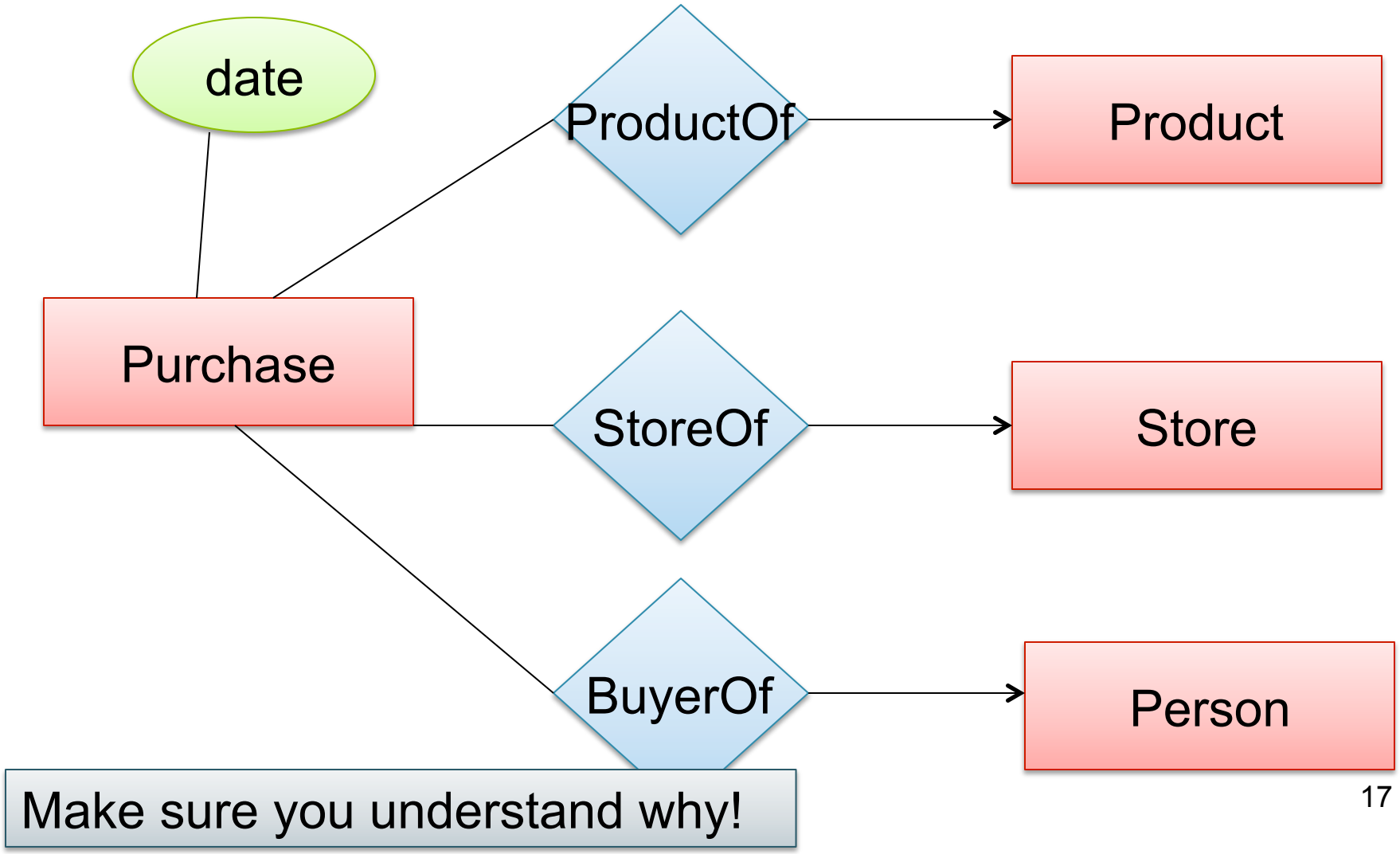


A: Cannot. This is the best approximation.
(Why only approximation ?)

Converting Multi-way Relationships to Binary

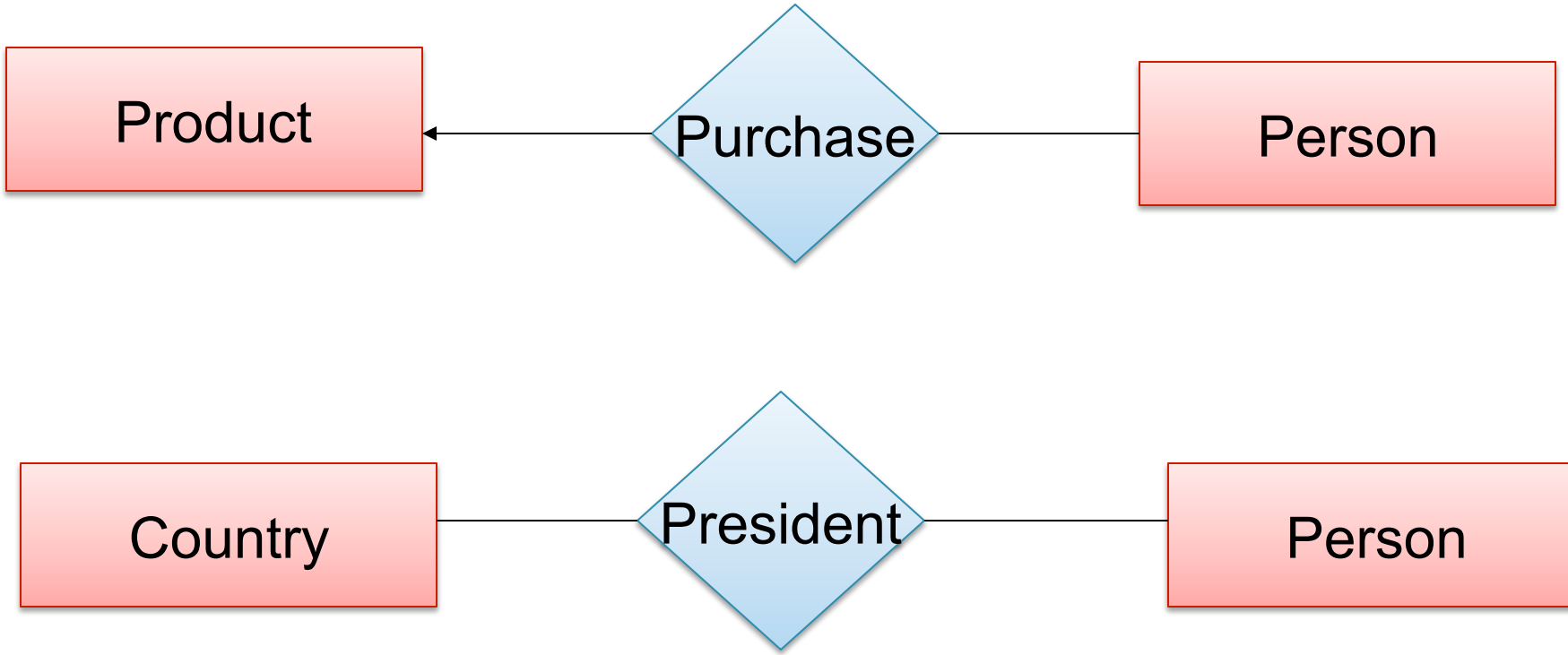


Converting Multi-way Relationships to Binary



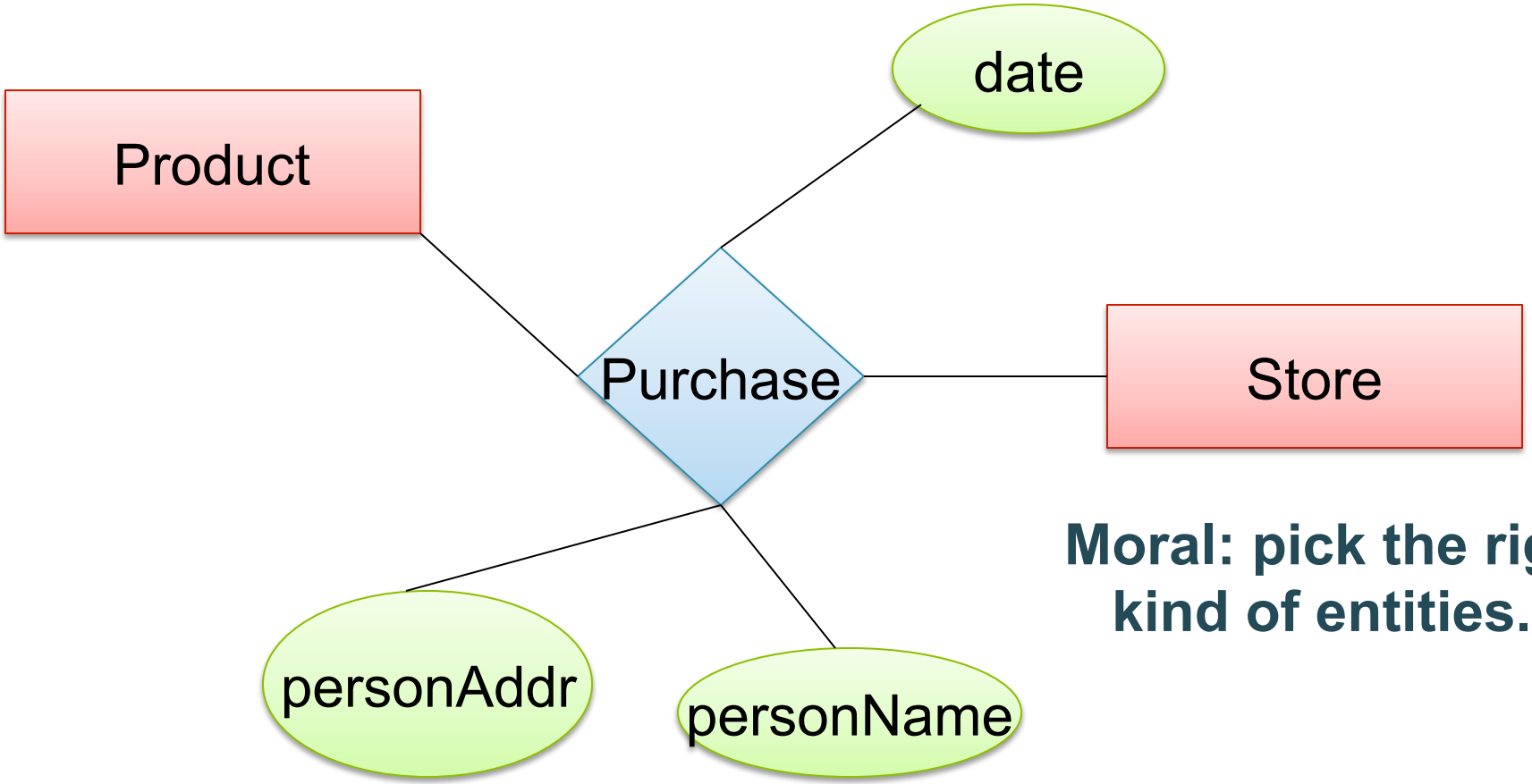
3. Design Principles

What's wrong?



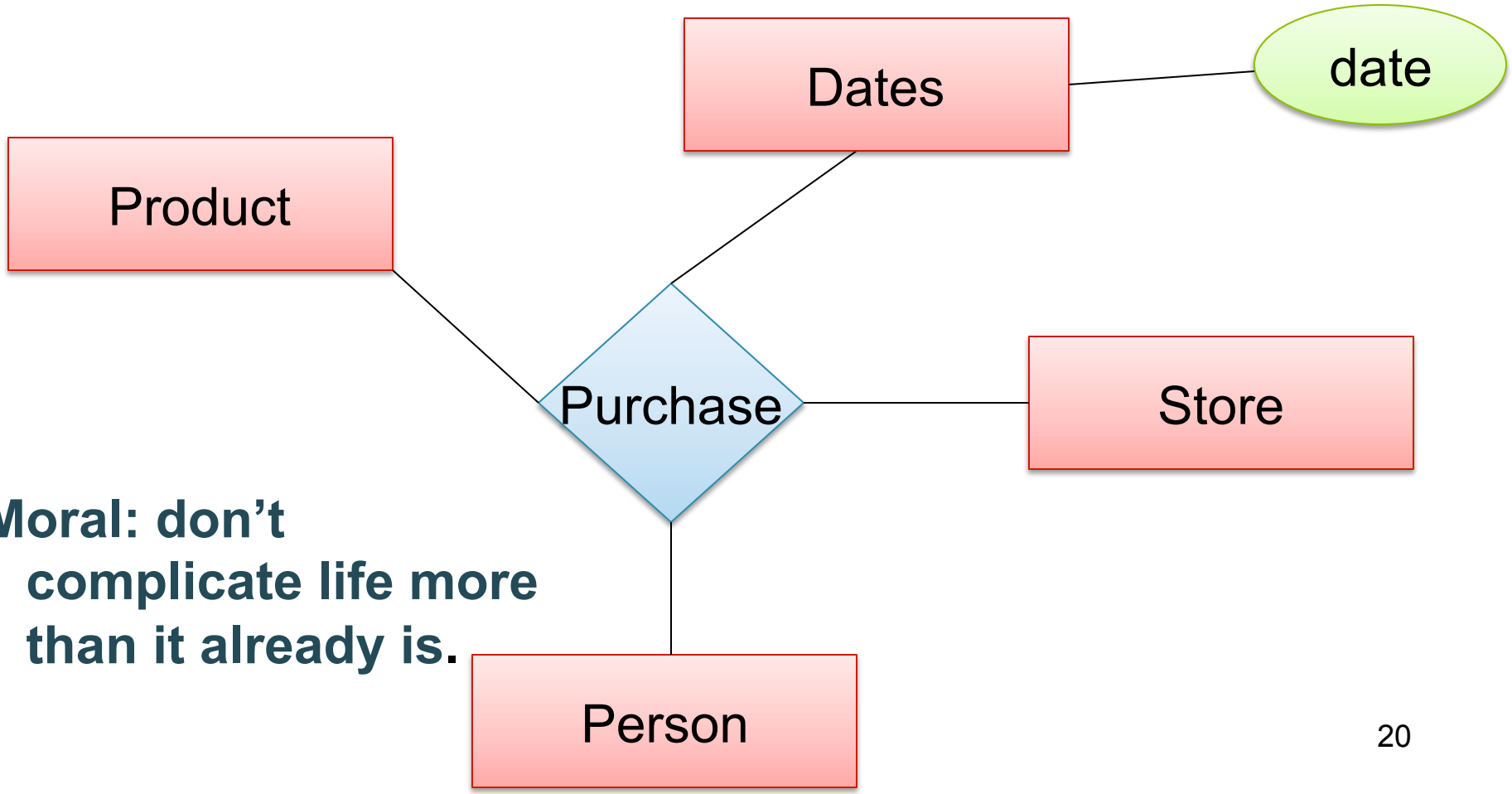
Moral: be faithful to the specifications of the app!

Design Principles: What's Wrong?



Moral: pick the right kind of entities.

Design Principles: What's Wrong?

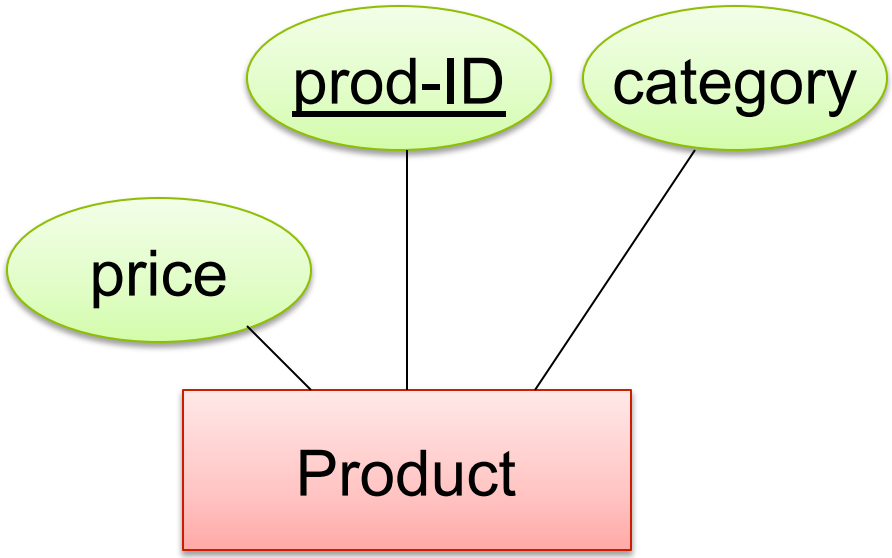


Moral: don't complicate life more than it already is.

From E/R Diagrams to Relational Schema

- Entity set → relation
- Relationship → relation

Entity Set to Relation



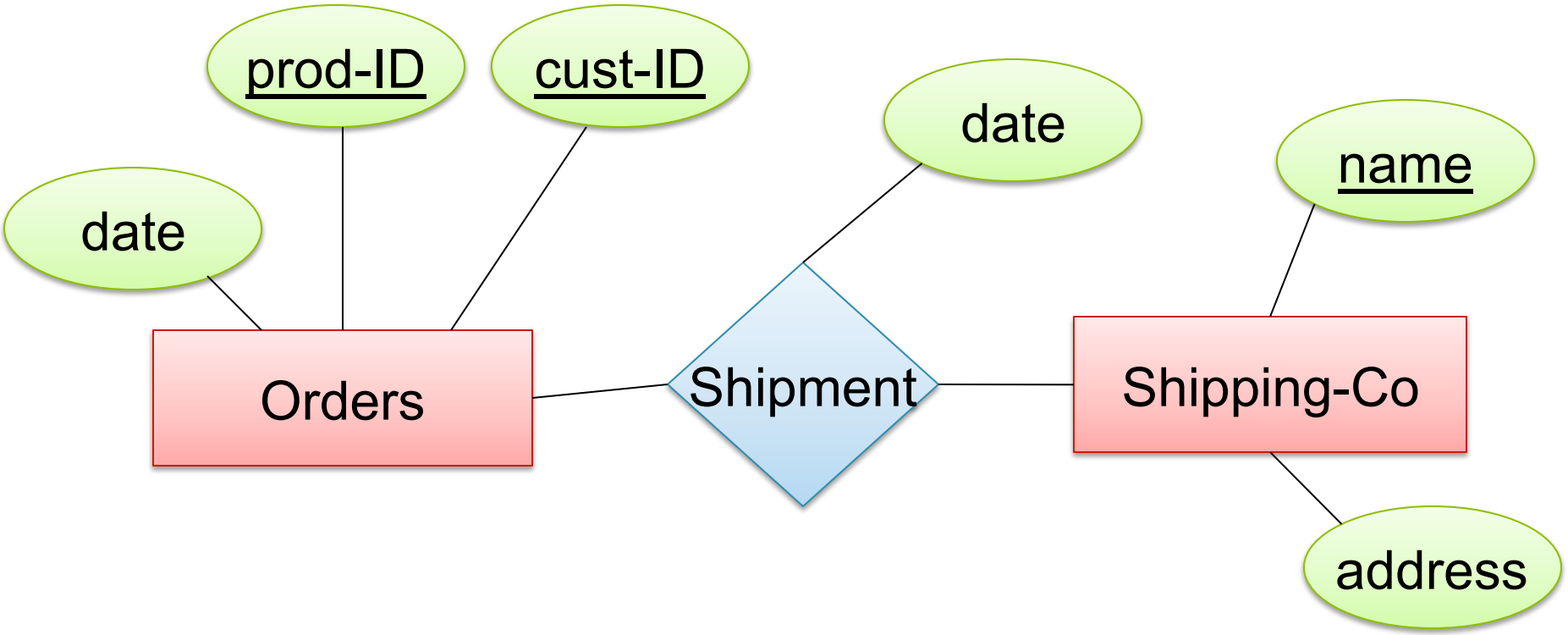
Product(prod-ID, category, price)

prod-ID	category	price
Gizmo55	Camera	99.99
Pokemn19	Toy	29.99

Create Table (SQL)

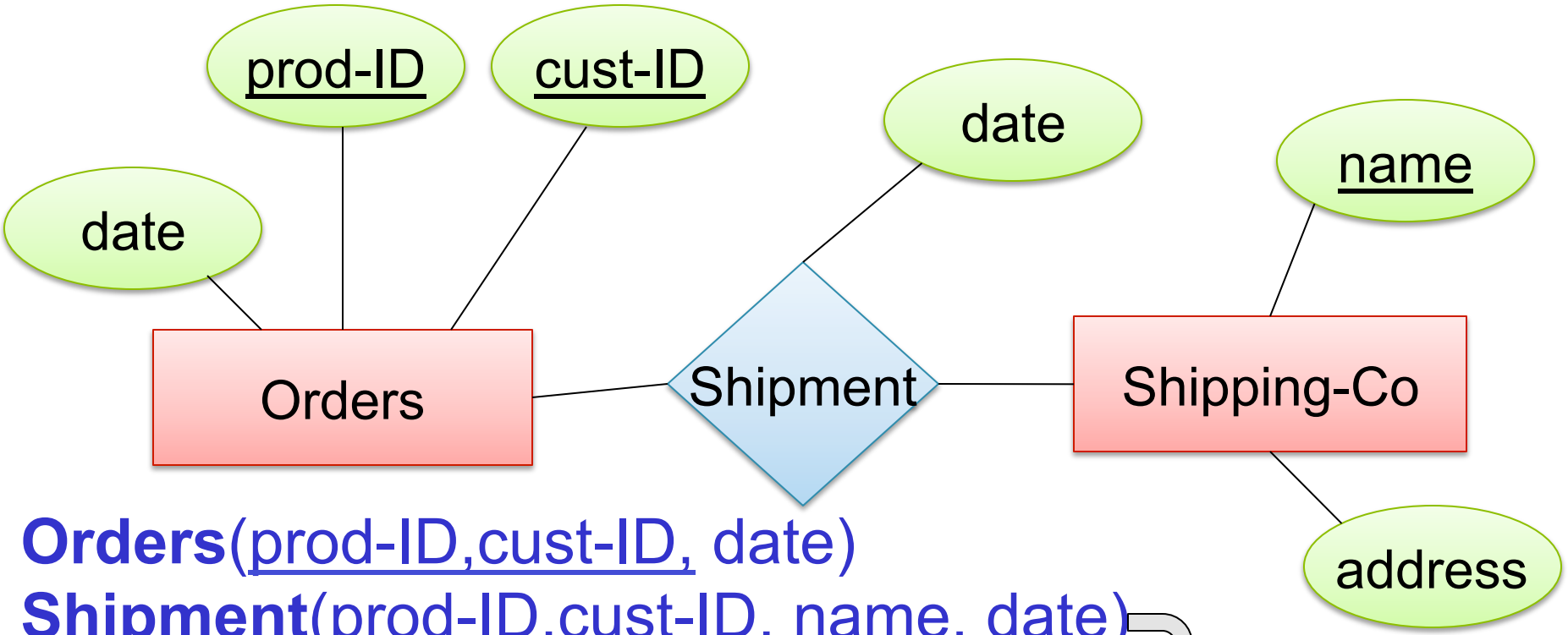
```
CREATE TABLE Product (
        prod-ID CHAR(30) PRIMARY KEY,
        category VARCHAR(20),
         price double)
```

N-N Relationships to Relations



Represent that in relations!

N-N Relationships to Relations



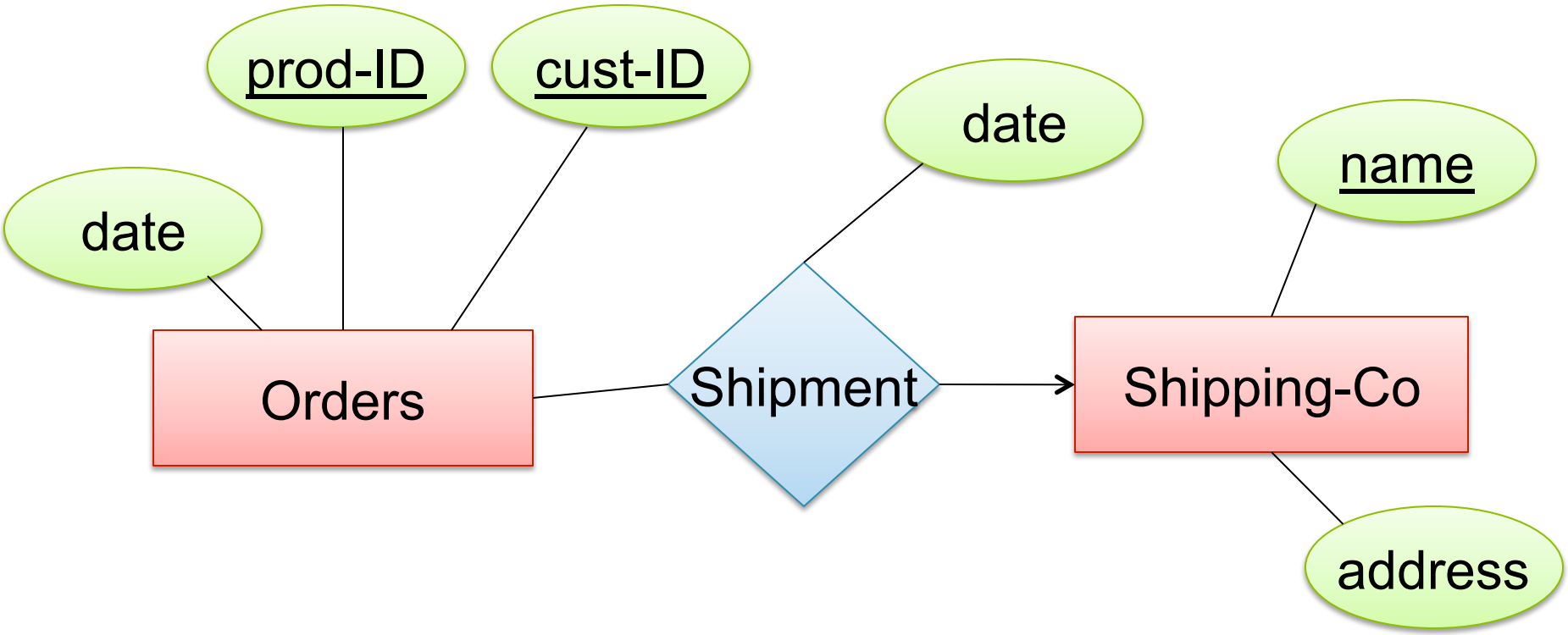
Orders(prod-ID,cust-ID, date)
Shipment(prod-ID,cust-ID, name, date)
Shipping-Co(name, address)

prod-ID	cust-ID	name	date
Gizmo55	Joe12	UPS	4/10/2011
Gizmo55	Joe12	FEDEX	4/9/2011

Create Table (SQL)

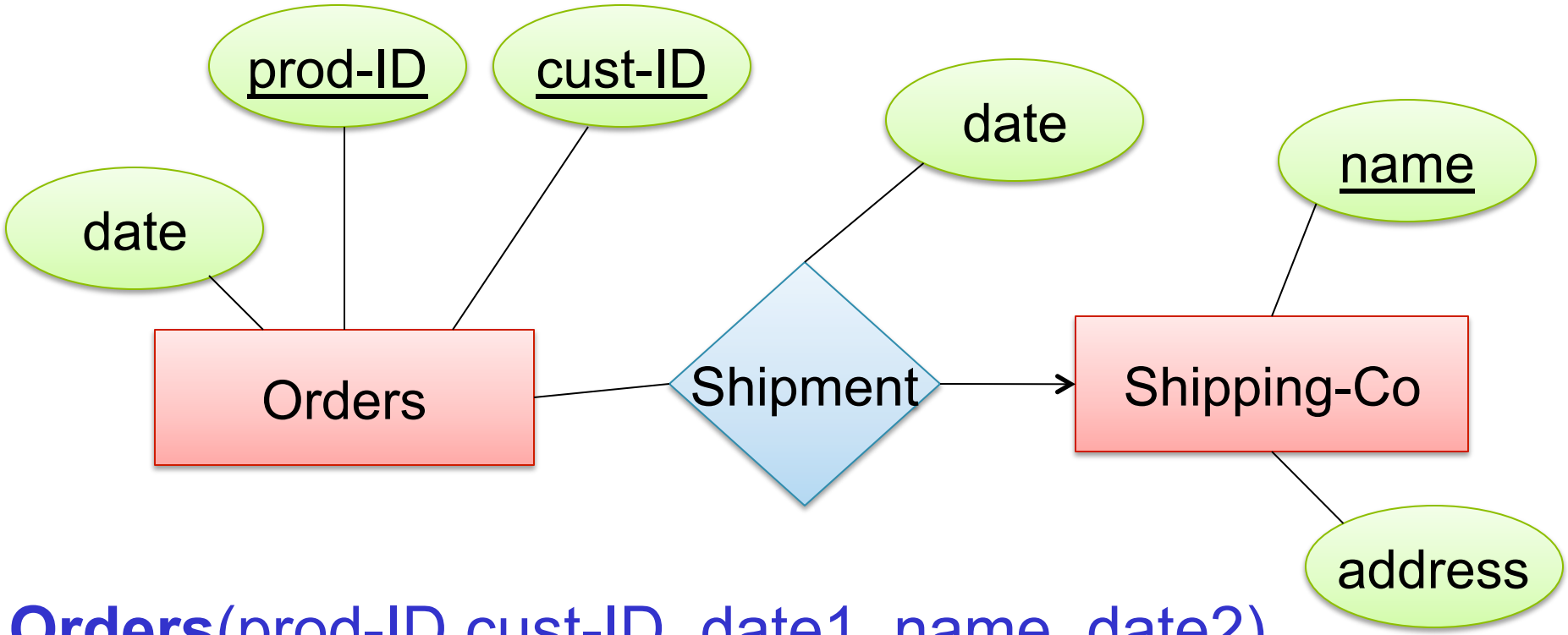
```
CREATE TABLE Shipment(
        name CHAR(30)
                REFERENCES Shipping-Co,
         prod-ID CHAR(30),
         cust-ID VARCHAR(20),
        date DATETIME,
PRIMARY KEY (name, prod-ID, cust-ID),
FOREIGN KEY (prod-ID, cust-ID)
          REFERENCES  Orders
)
```

N-1 Relationships to Relations



Represent this in relations!

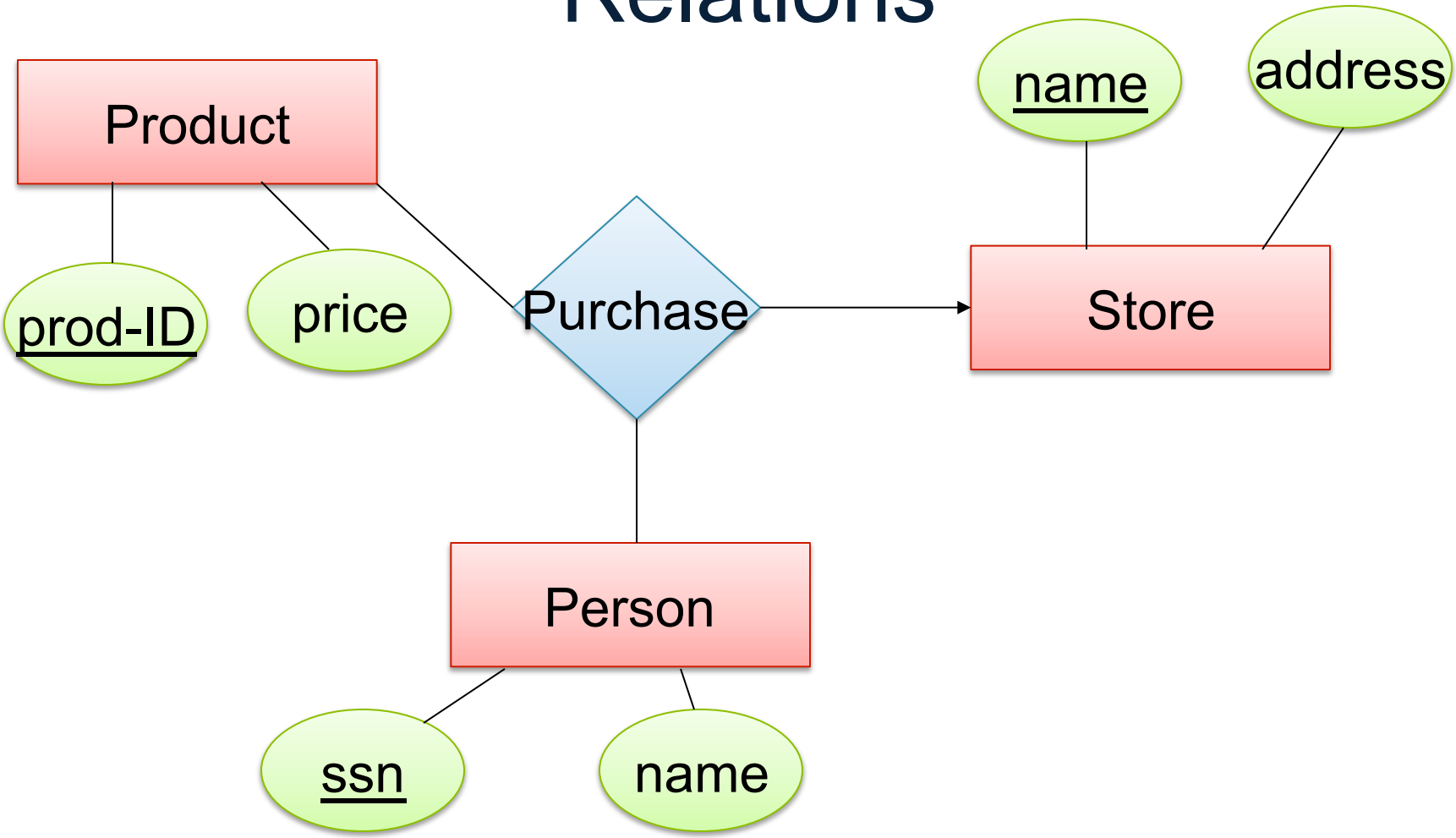
N-1 Relationships to Relations



Orders(prod-ID,cust-ID, date1, name, date2)
Shipping-Co(name, address)

Remember: no separate relations for many-one relationship

Multi-way Relationships to Relations

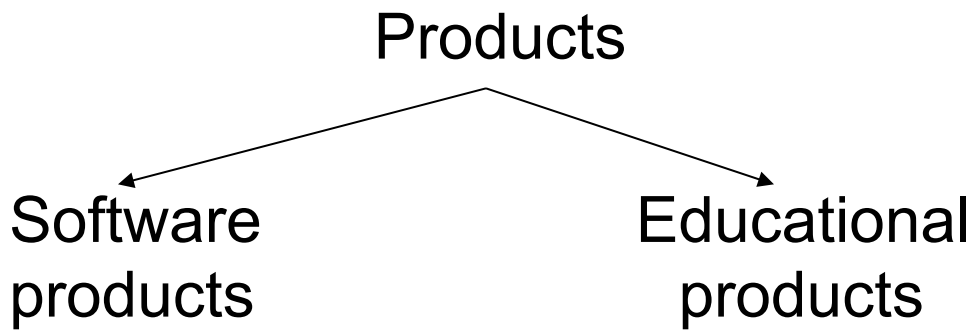


Purchase(prod-ID, cust-ssn, store-name)

Modeling Subclasses

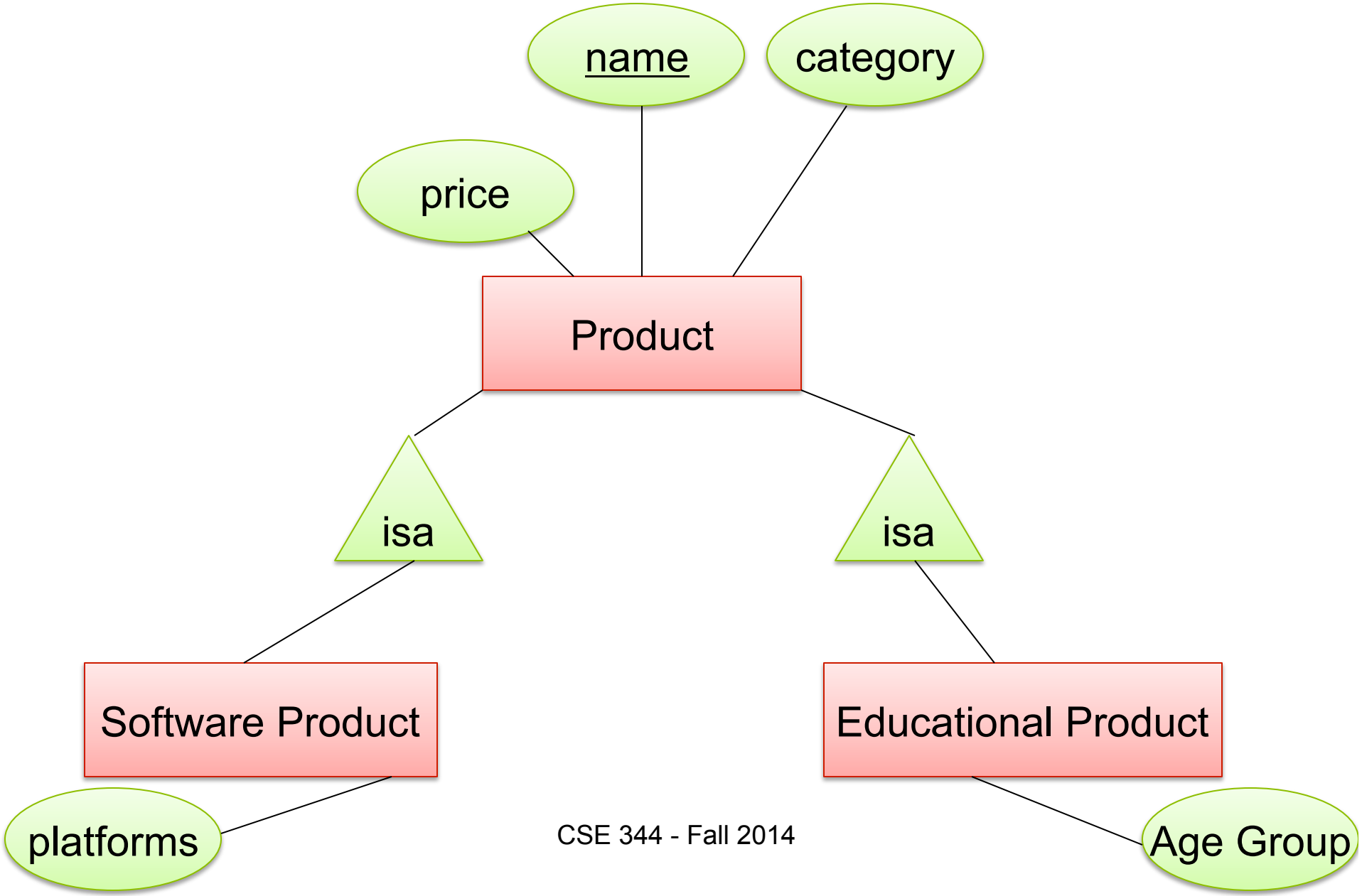
Some objects in a class may be special

- define a new class
- better: define a *subclass*



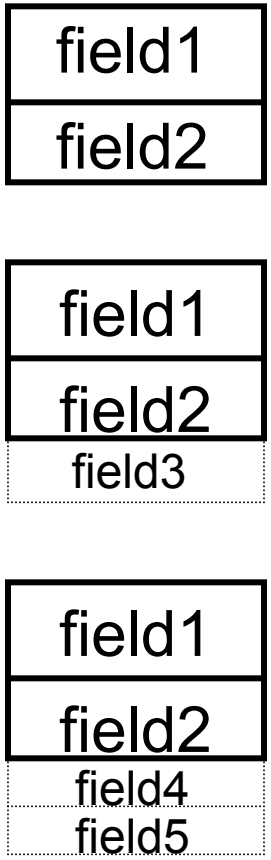
So --- we define subclasses in E/R

Subclasses

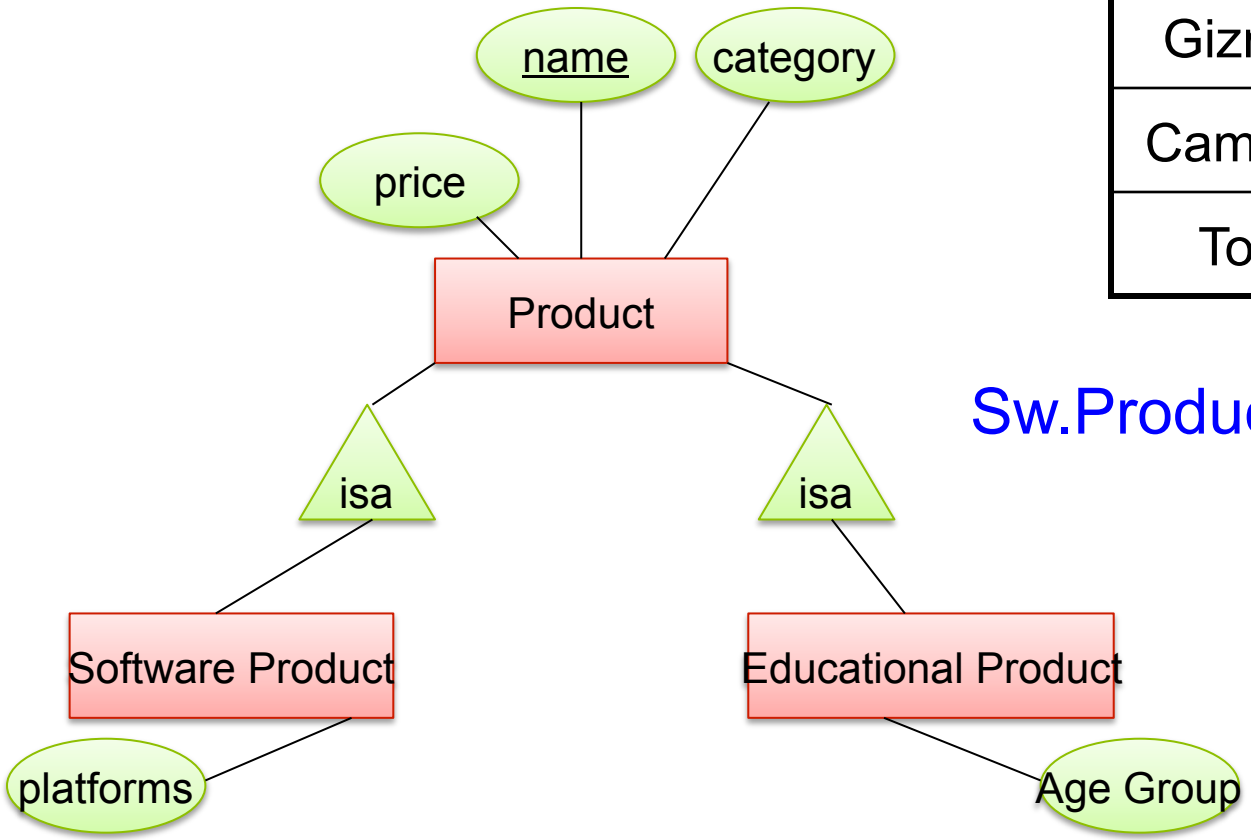

name
category
price
Product
isa
isa
Software Product
Educational Product
platforms
Age Group

Understanding Subclasses

- Think in terms of records:
  - Product
  - SoftwareProduct
  - EducationalProduct



Subclasses to Relations



Product

Name	Price	Category
Gizmo	99	gadget
Camera	49	photo
Toy	39	gadget

Sw.Product

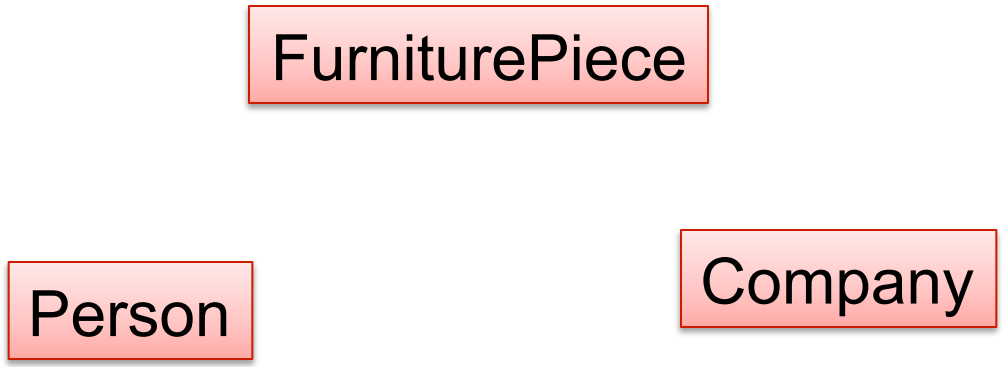
Name	platforms
Gizmo	unix

Ed.Product

Name	Age Group
Gizmo	toddler
Toy	retired

Other ways to convert are possible

Modeling UnionTypes With Subclasses

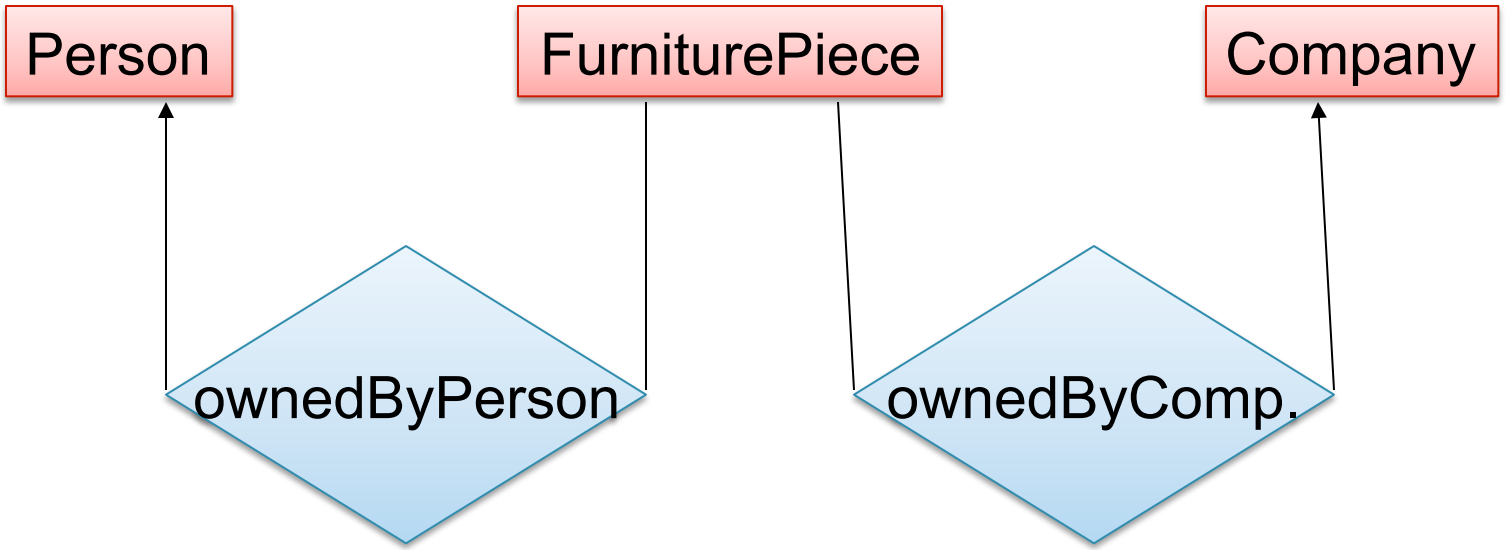


Say: each piece of furniture is owned either by a person or by a company

Modeling Union Types with Subclasses

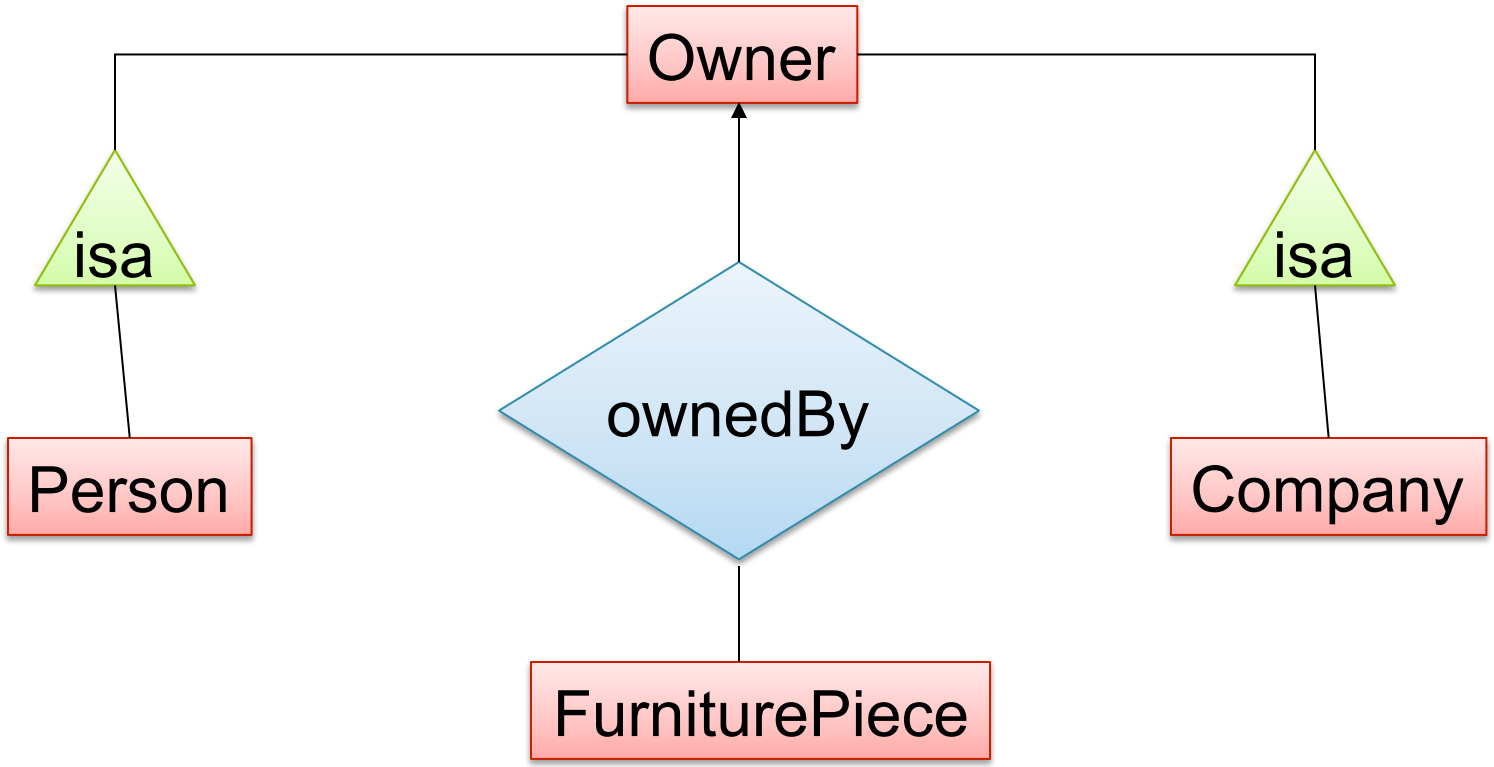
Say: each piece of furniture is owned either by a person or by a company

Solution 1. Acceptable but imperfect (What's wrong ?)



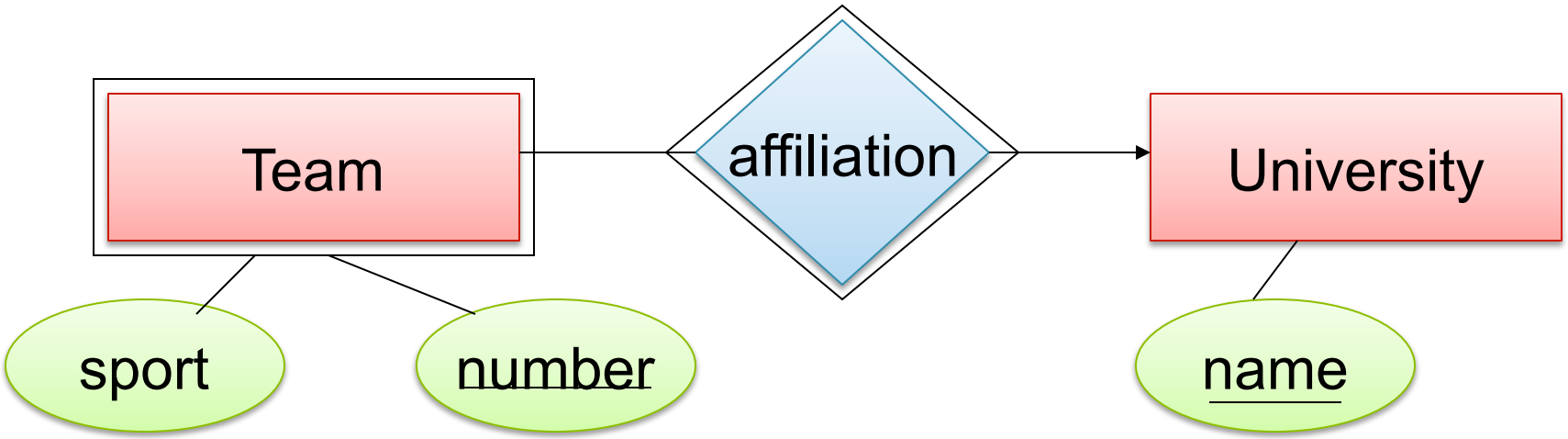
Modeling Union Types with Subclasses

Solution 2: better, more laborious



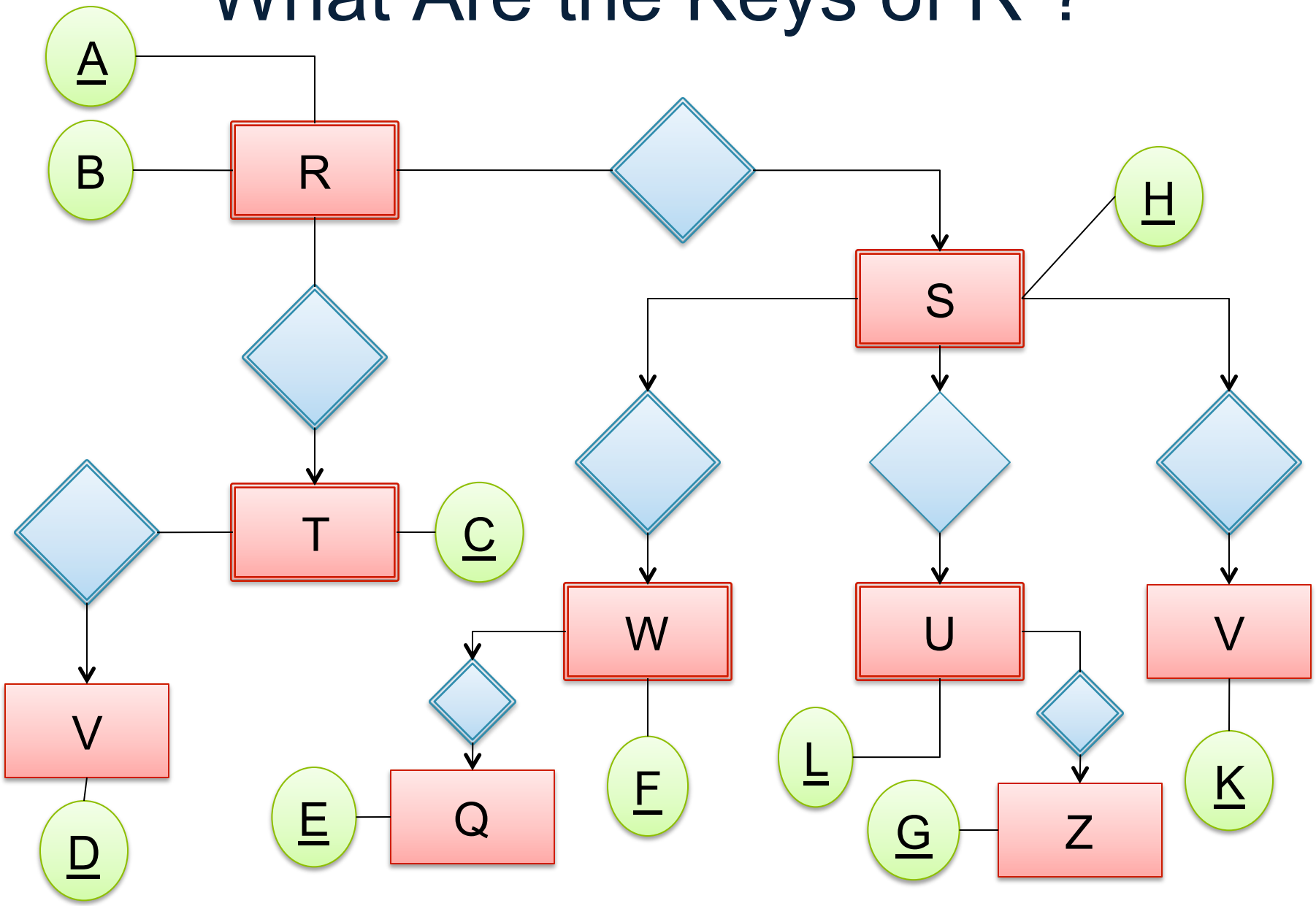
Weak Entity Sets

Entity sets are weak when their key comes from other classes to which they are related.



Team(sport, number, universityName)
University(name)

What Are the Keys of R ?



Where we are

We now have tools for creating models and tables

Next steps are to figure out how get the right tables and relations given possible choices

Next few lectures:

- Constraints and data integrity
- Schema normalization
- Views