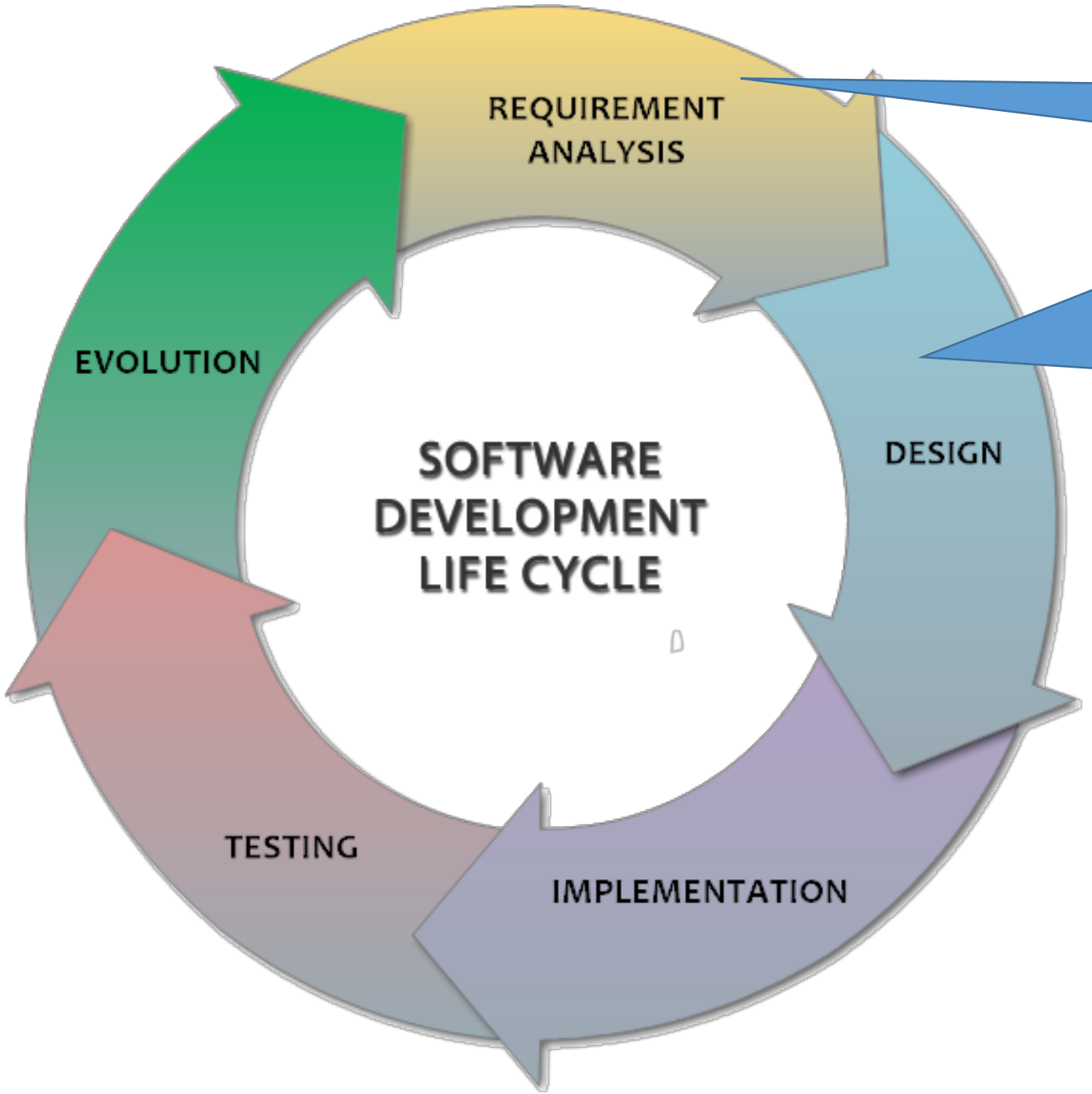
COMP 430
Intro. to Database Systems

Entity-Relationship Diagrams

http://jcsites.juniata.edu/faculty/rhodes/dbms/eermodel

Slides use ideas from Chris Ré.



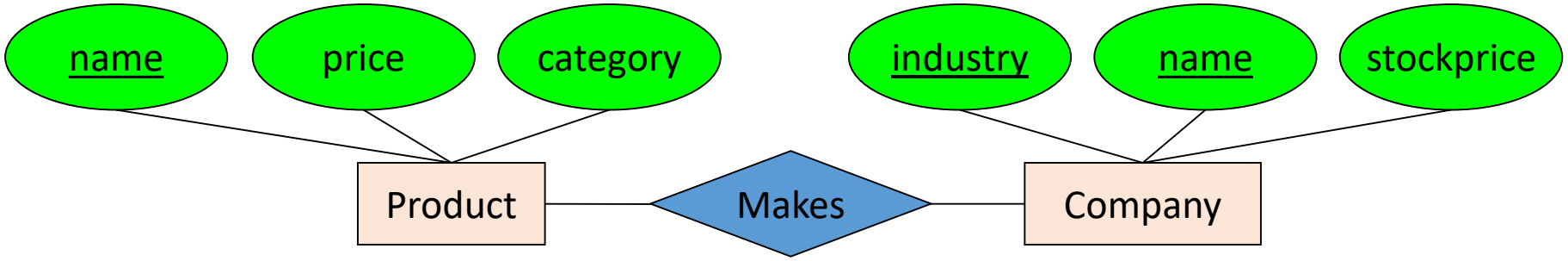
Req. Analysis outlines:
Data & what we want to do with it

Goal: Organize data to support requirements.

Use two complementary techniques:

- Entity-relationship design (now)
- Normalization (later)

Entity-Relationship Diagrams



"The Entity-Relationship model – toward a unified view of data"
Peter Chen, 1976

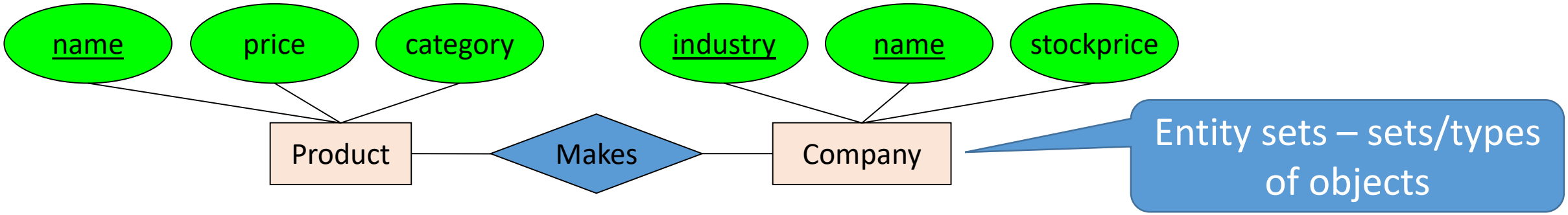
Precise enough for technical decisions

Abstracted enough to omit some implementation details

Abstracted enough for non-technical people

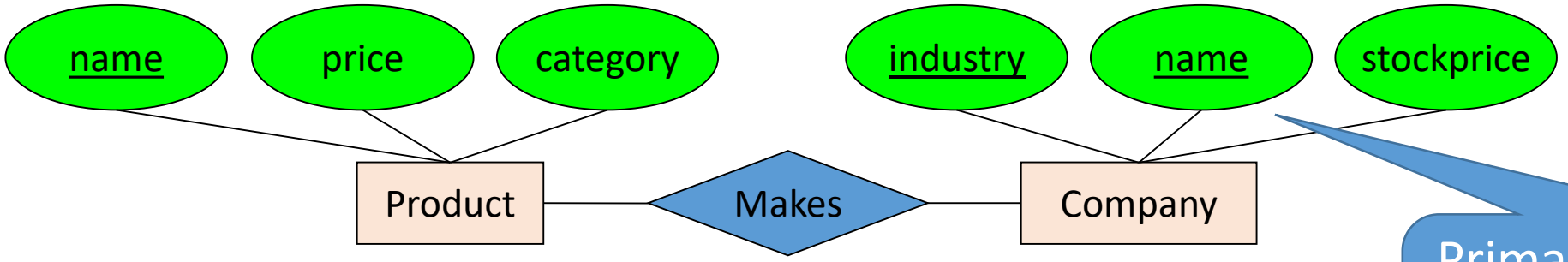
Many variations

For the course, use the given notation.



Product

Company



Attributes of entity sets

Primary key – minimal set of attributes that uniquely identifies entity

http://jcsites.juniata.edu/faculty/rhodes/dbms/eermodel

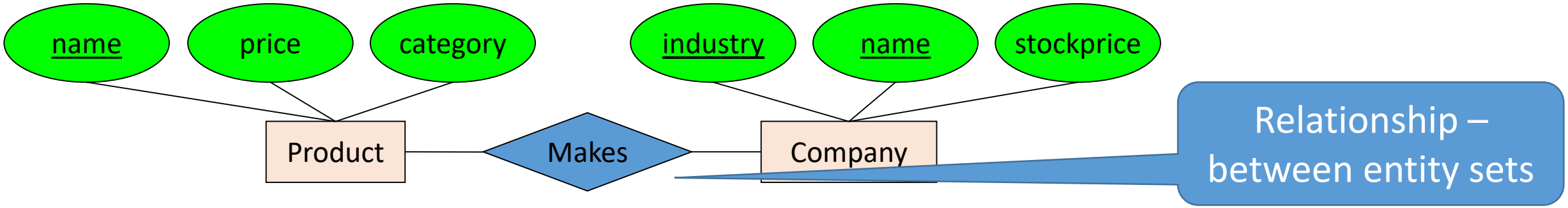
Product

name	price	category
iPad	$200	Tablet
Galaxy Tab	$200	Tablet
Galaxy Note	$500	Smartphone
iPhone	$600	Smartphone
Let It Be	$10	CD

Company

industry	name	stockprice
Music	Apple	
Technology	Apple	$100
Technology	Samsung	$45

Convenient to draw in tabular form. Implemented as tables.



Sets A, B:

A={1,2,3}, B={a,b,c,d}

Cross-product A × B:

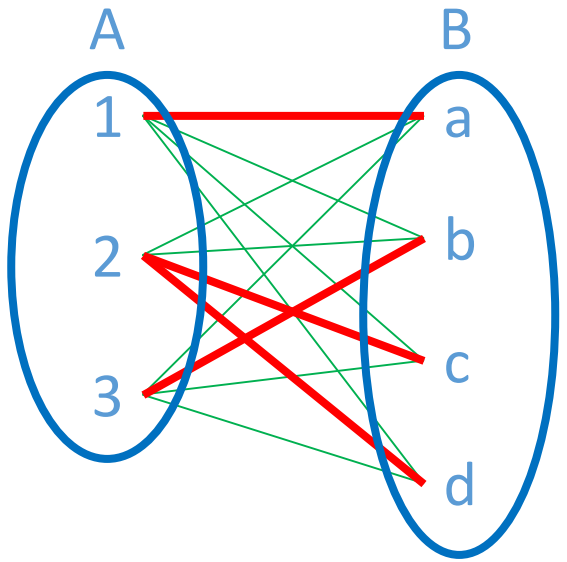
Set of all (a∈A, b∈B) pairs

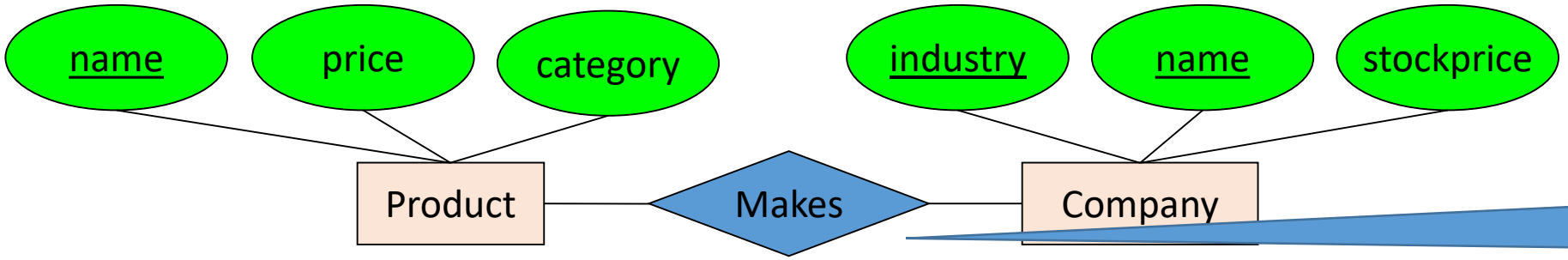
A × B = {(1,a), (1,b), (1,c), (1,d), (2,a), (2,b), (2,c), (2,d), (3,a), (3,b), (3,c), (3,d)}

(Binary) Relationship R:

Some subset of A × B

R = {(1,a), (2,c), (2,d), (3,b)}





Relationship – between entity sets

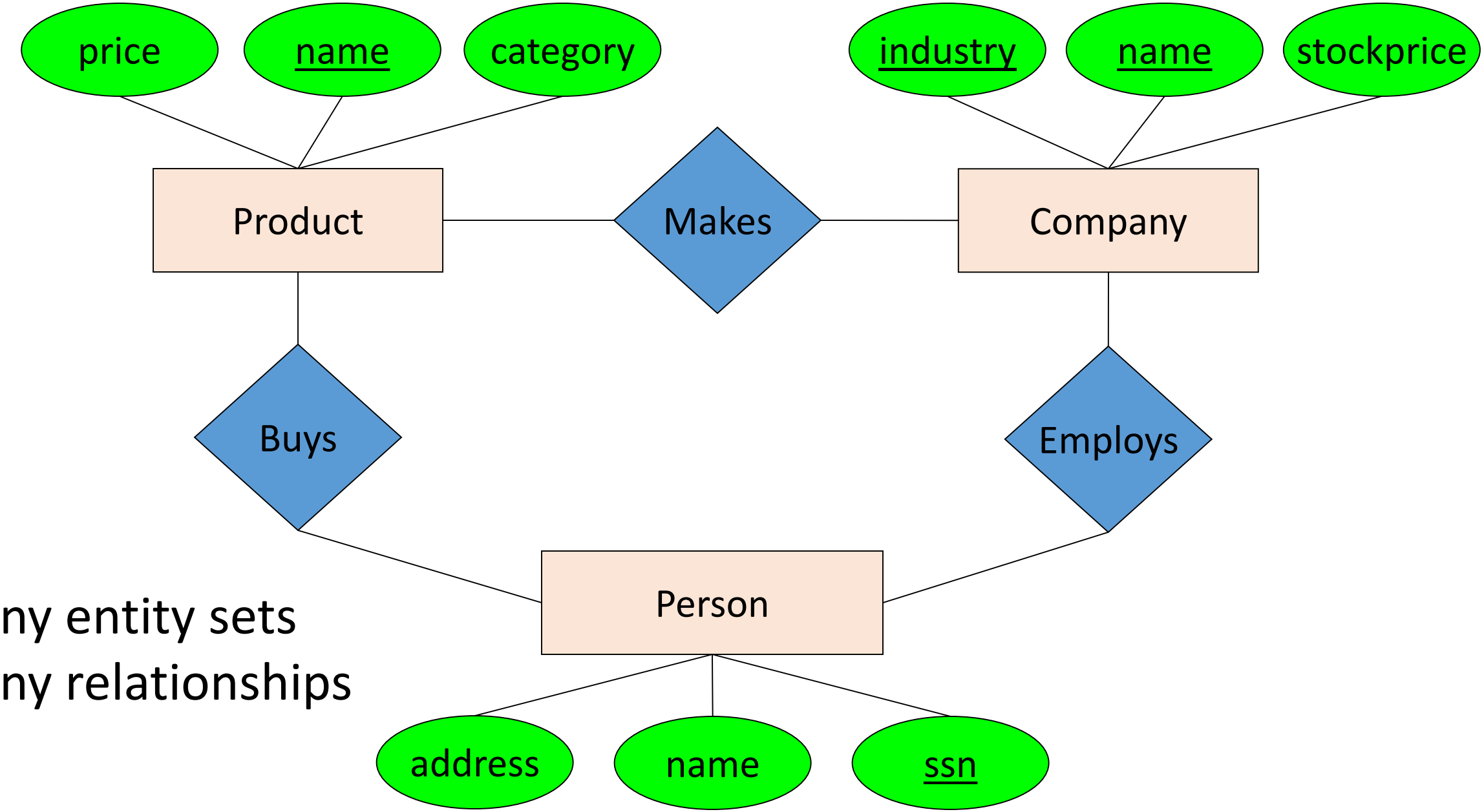
Company c × Product p

c.industry	c.name	p.name	…
Music	Apple	iPad	…
Music	Apple	Galaxy Tab	…
Music	Apple	Galaxy Note	…
Music	Apple	iPhone	…
Music	Apple	Let It Be	…
Technology	Apple	iPad	…
Technology	Apple	Galaxy Tab	…
…	…	…	…
Technology	Samsung	iPhone	…
Technology	Samsung	Let It Be	…

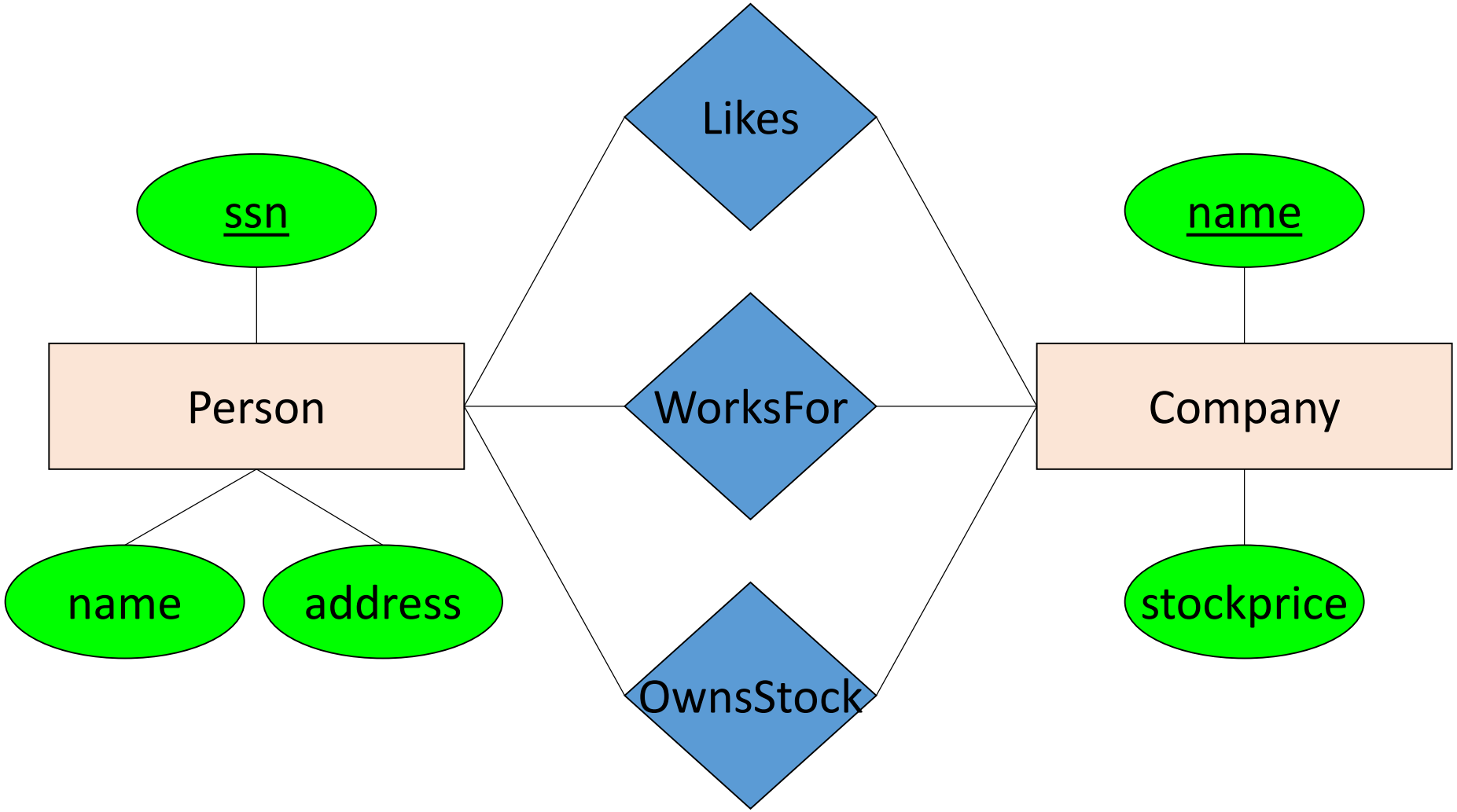
Makes

c.industry	c.name	p.name	…
Music	Apple	Let It Be	…
Technology	Apple	iPad	…
Technology	Apple	iPhone	…
Technology	Samsung	Galaxy Tab	…
Technology	Samsung	Galaxy Note	…

Convenient to draw in tabular form. Implementation as tables discussed later.

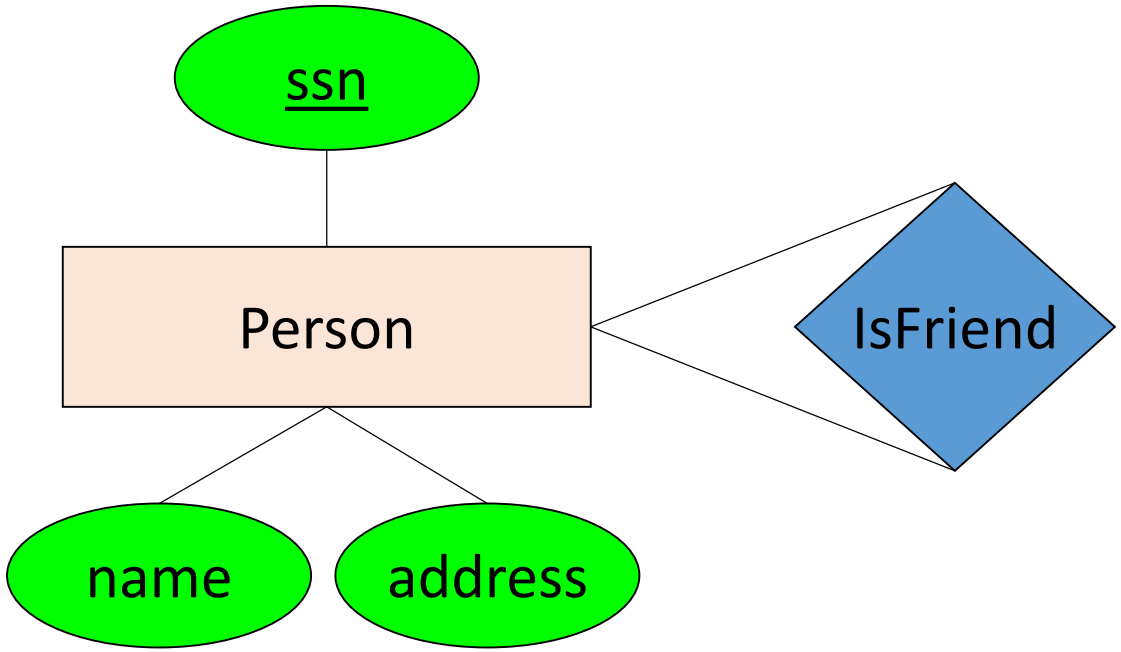


Many entity sets
Many relationships



Can have multiple relationships between same entity sets.

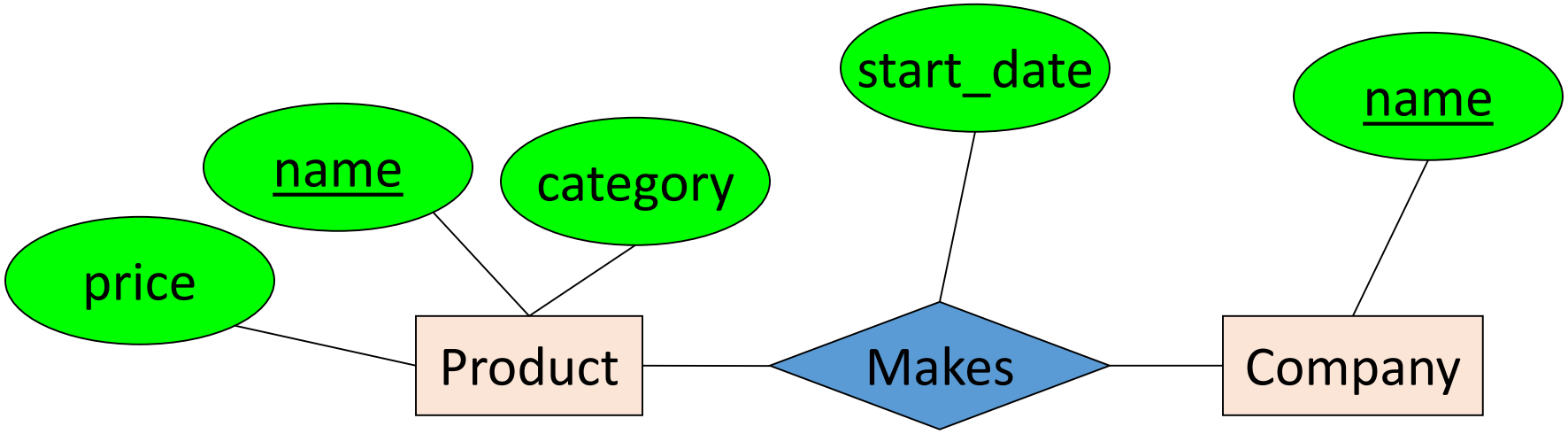
Each is a subset of **Person** × **Company** with primary key (ssn, name).



Can have relationship between single entity set.

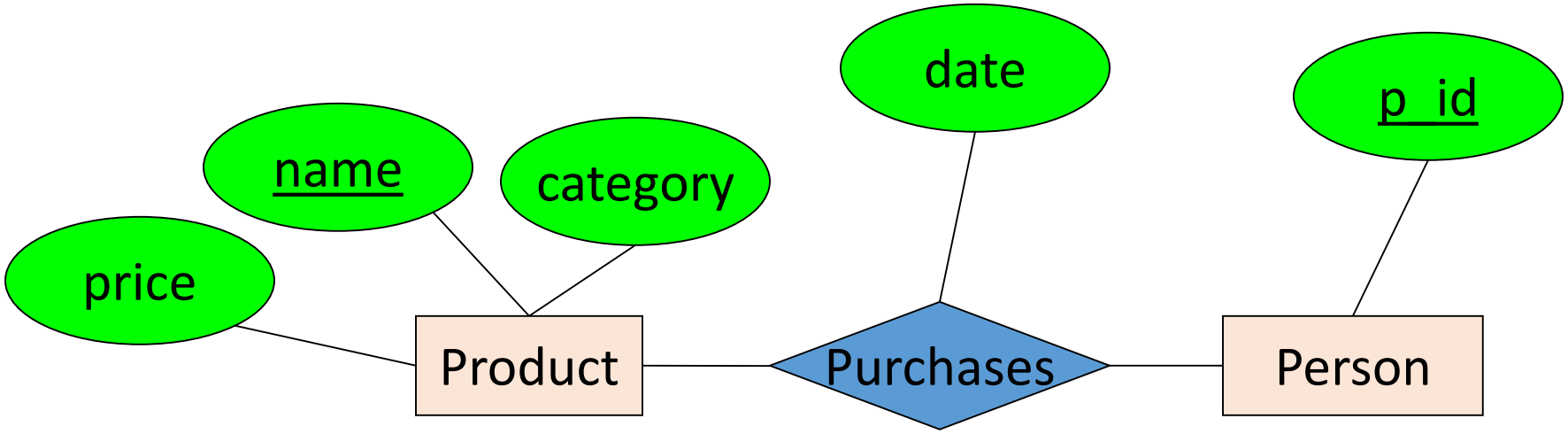
A subset of **Person** × **Person** with primary key (ssn, ssn).

Relationships can have (non-key) attributes



Relationship attributes are implicitly unique per (p.name, c.name).

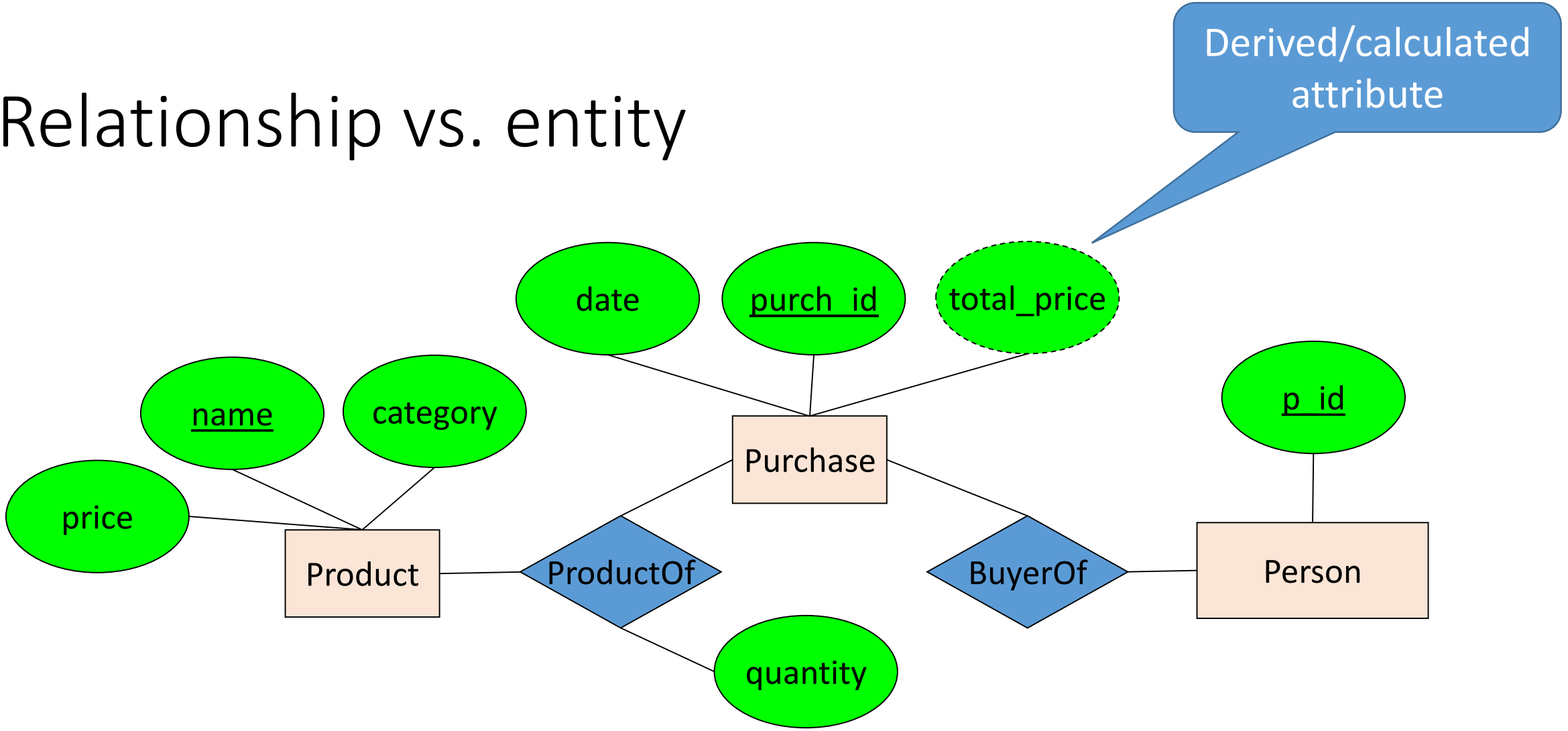
Relationship vs. entity



Relationship attributes are implicitly unique per (name, p_id).

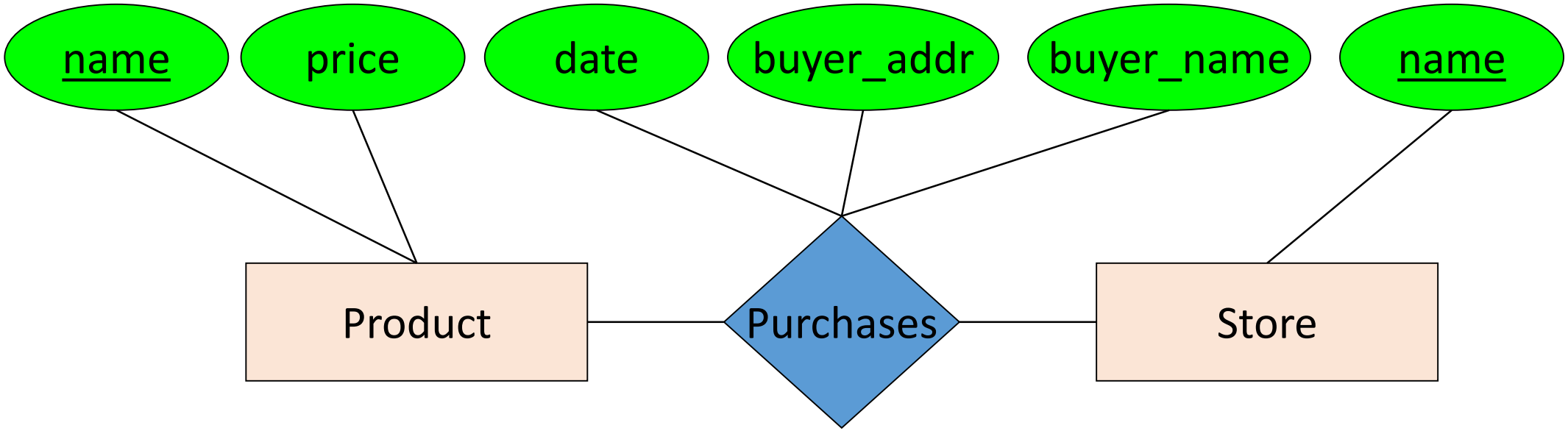
What if we don't want uniqueness?

Relationship vs. entity



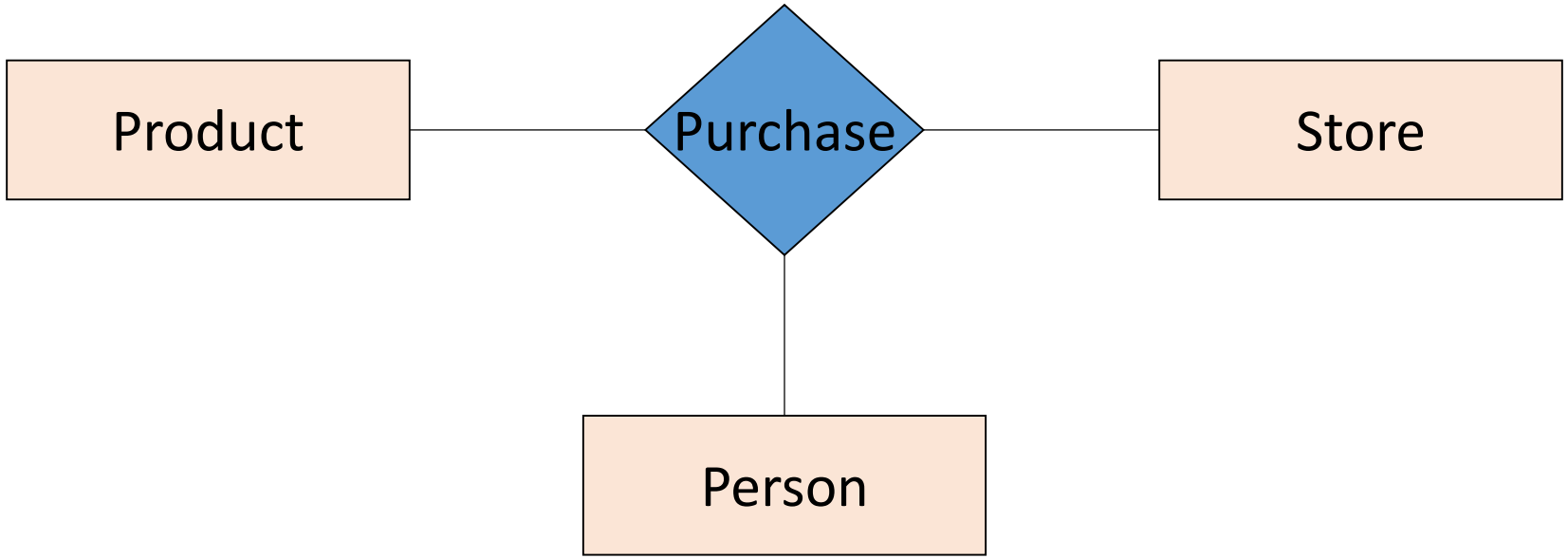
Representing as an entity allows multiple **Purchase**s for each **Product**-**Person** combination.

What's wrong?


name
price
date
buyer_addr
buyer_name
name
Product
Purchases
Store

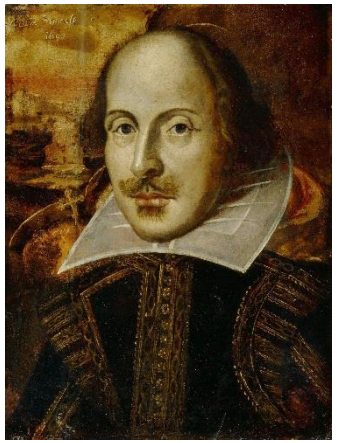
N-ary relationships

Attributes omitted for brevity

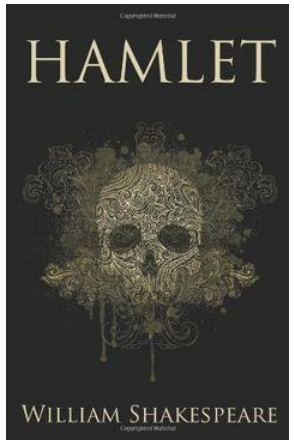


Purchase is a subset of **Person** × **Product** × **Store**.

Activity – Draw ER diagram



Authors have IDs and names. They write books.

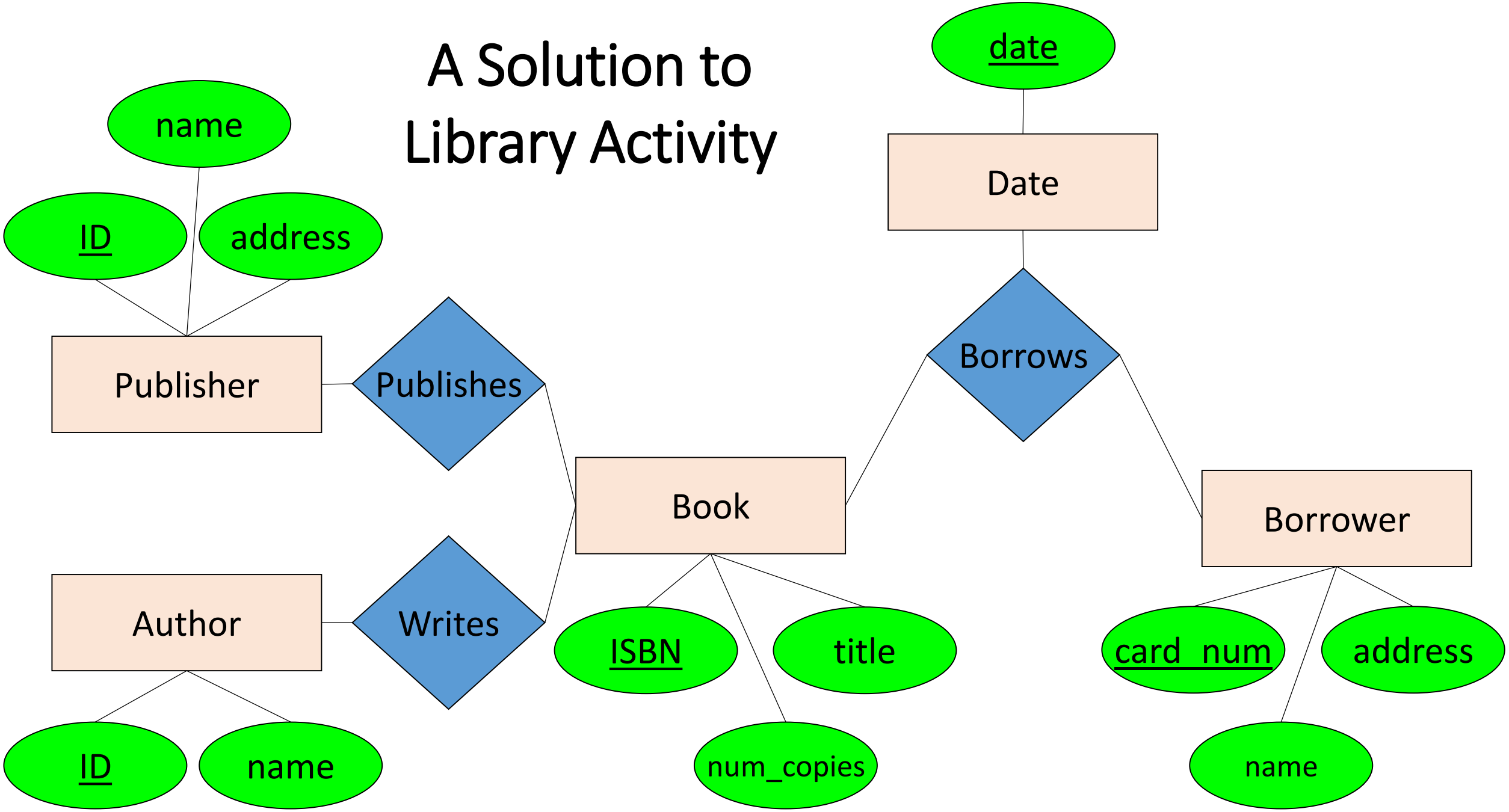


Books have ISBNs and titles. The library keeps track of how many copies it has of the book. Each book is written by authors and published by a publisher. We want to know every time it is checked out by a borrower.

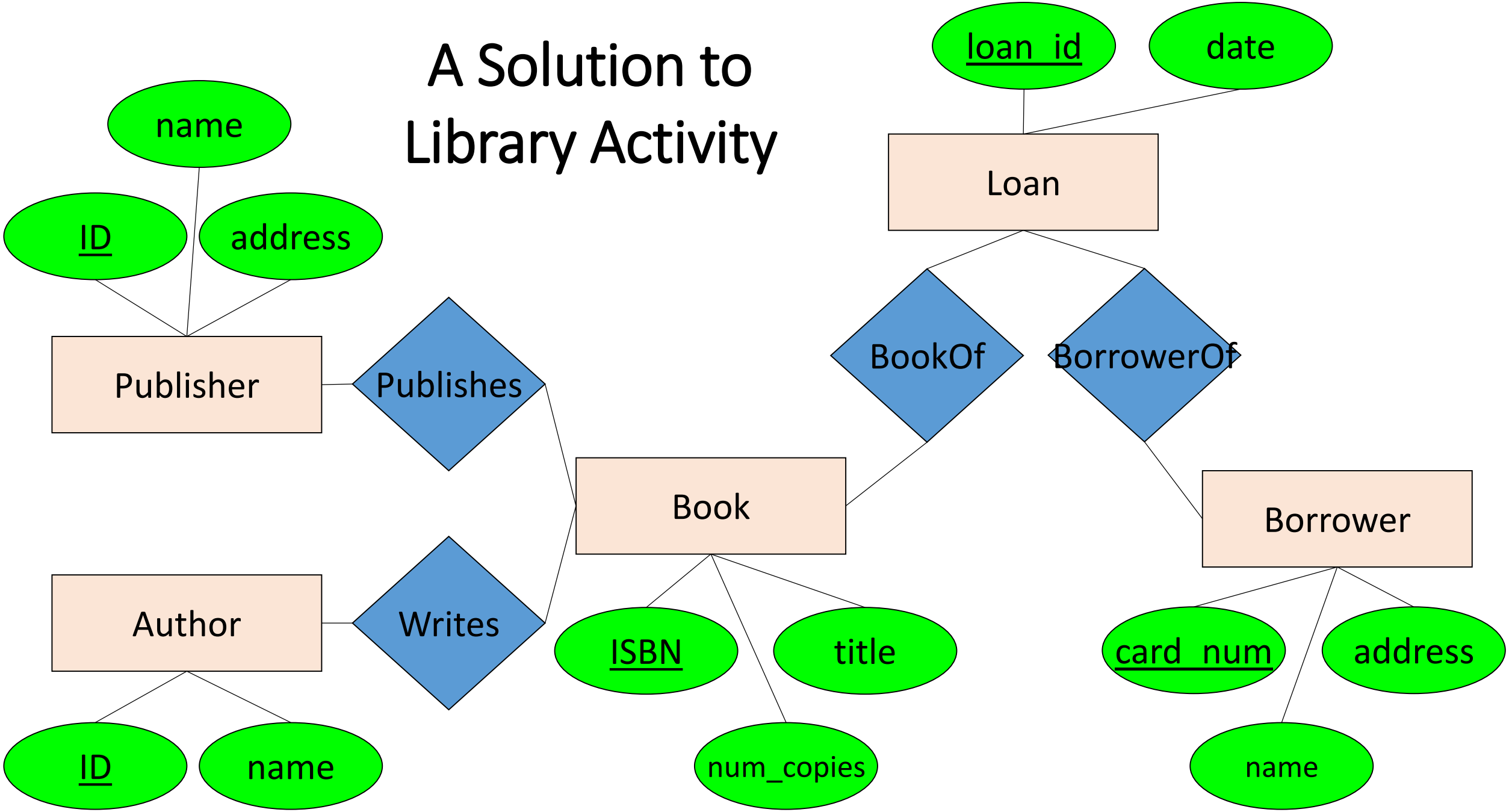
Borrowers have a library card number, name, and address. They can check out a book on a particular date.

Publishers have an ID, name and address. They publish books.

A Solution to Library Activity

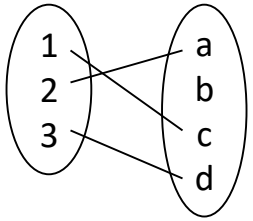


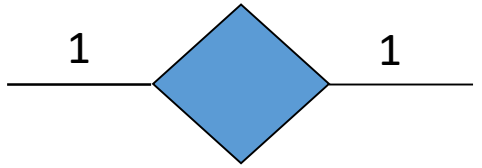
A Solution to Library Activity



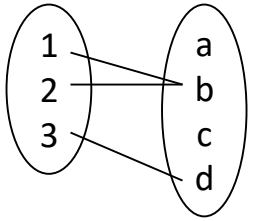
Maximum cardinality

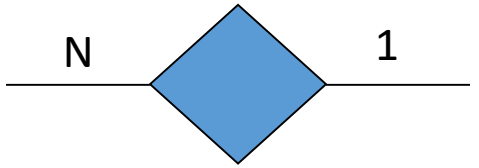
One-to-one:



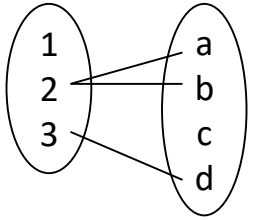


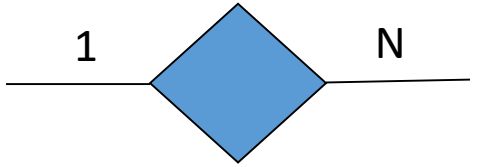
Many-to-one:



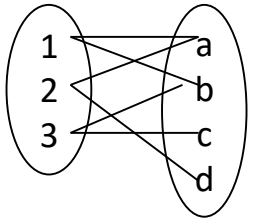


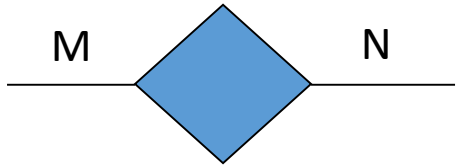
One-to-many:

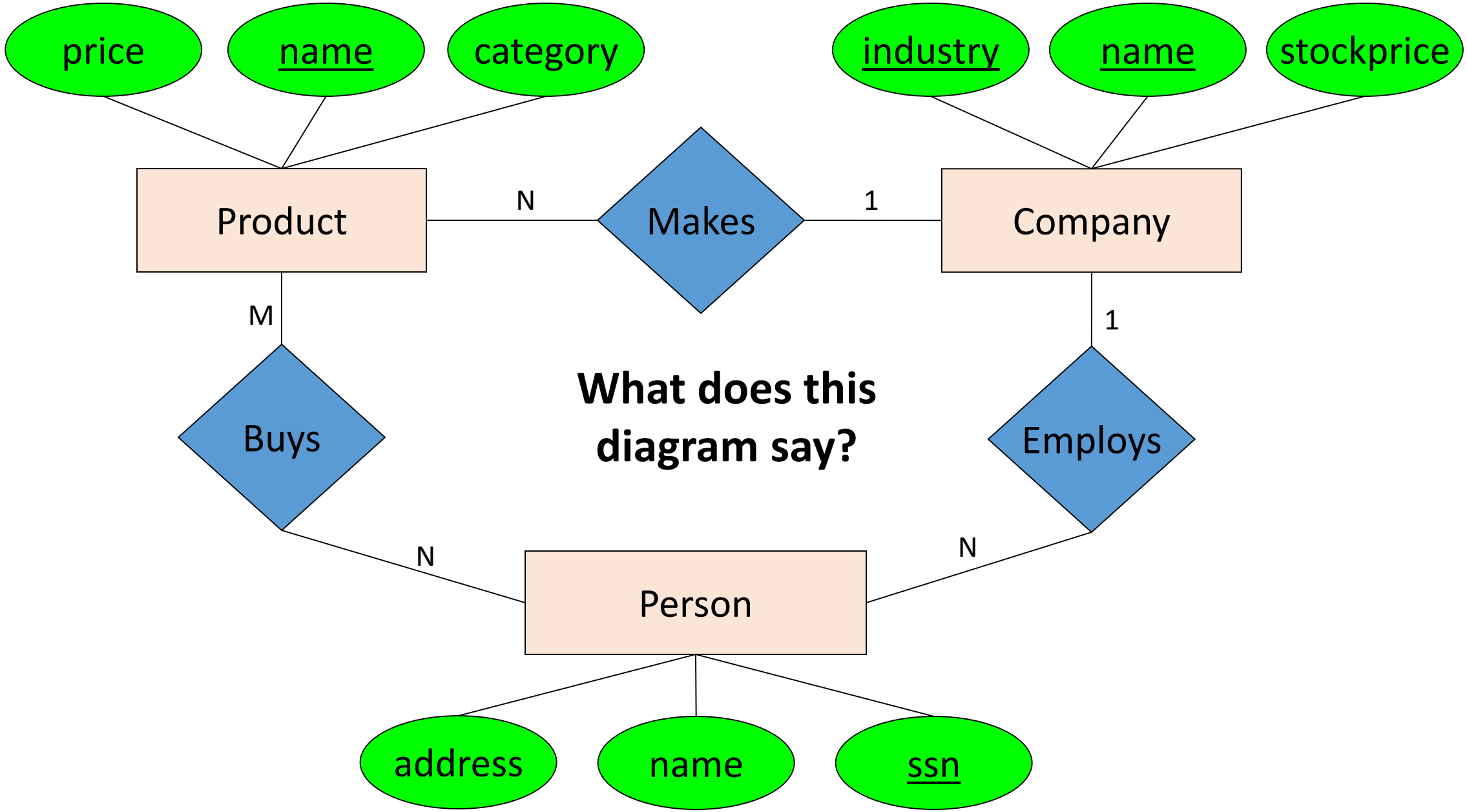




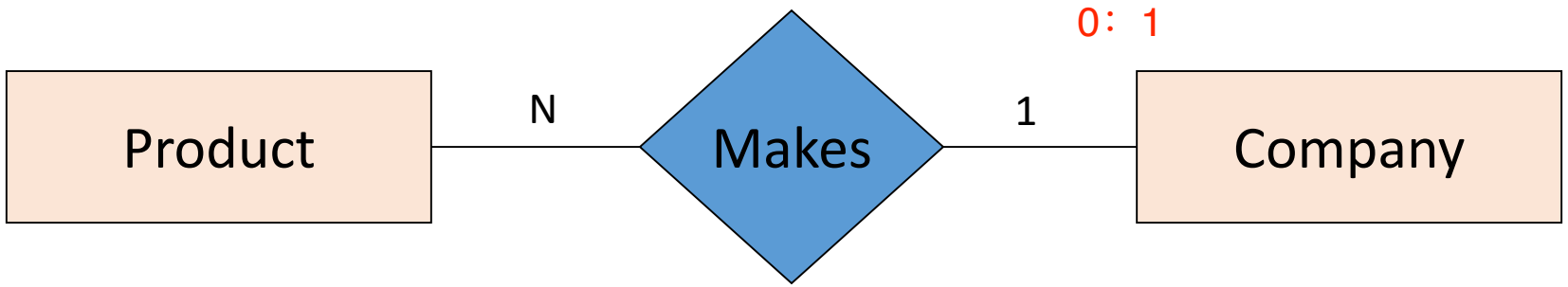
Many-to-many:





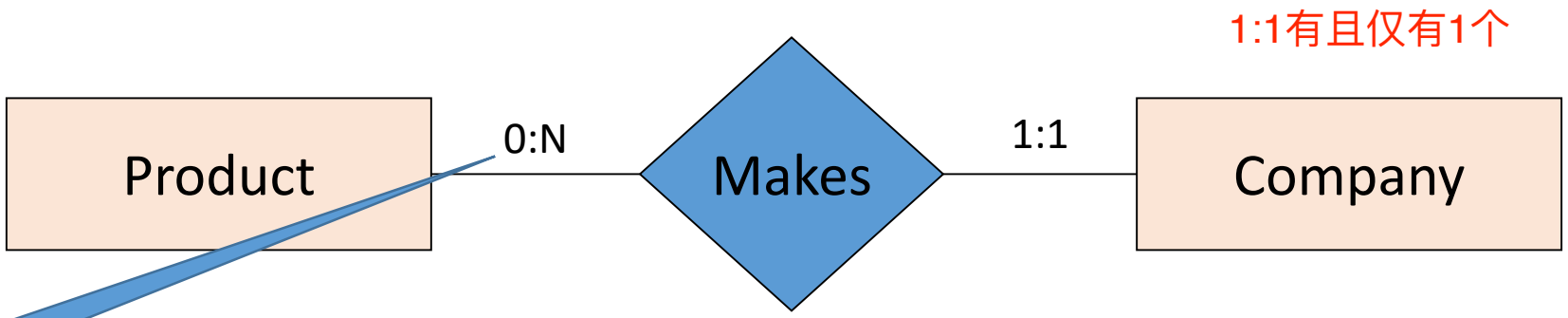

price
name
category
industry
name
stockprice
Product
N
Makes
1
Company
M
1
Buys
What does this diagram say?
Employs
N
N
Person
address
name
ssn

Minimum cardinality



Are there products made by no company?

Does every company make a product?

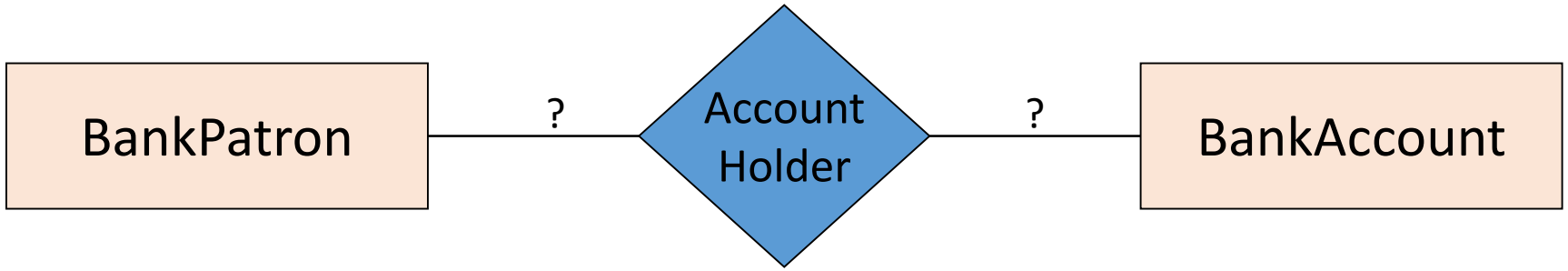


默认左边是0

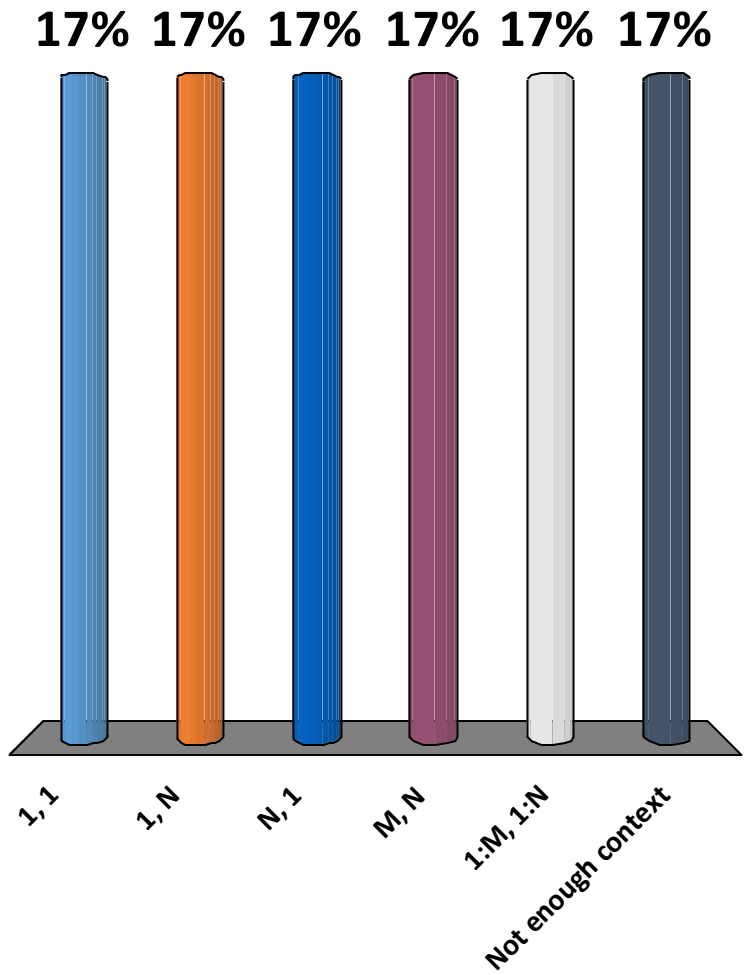
Or, simply N.

Each product maps to $1 \geq c \geq 1$ company.

Each company maps to $0 \geq p$ products.



Assume typical real-world context.



A. 1, 1

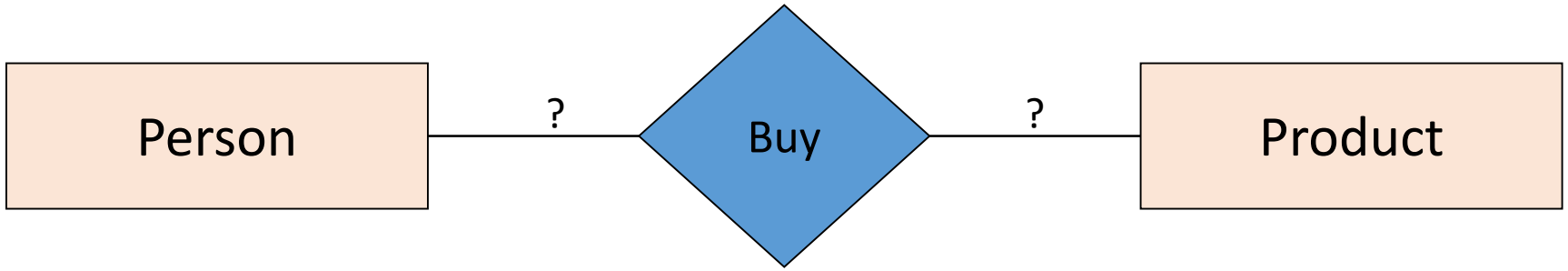
B. 1, N

C. N, 1

D. M, N

✓ E. 1:M, 1:N

F. Not enough context



Assume typical real-world context.

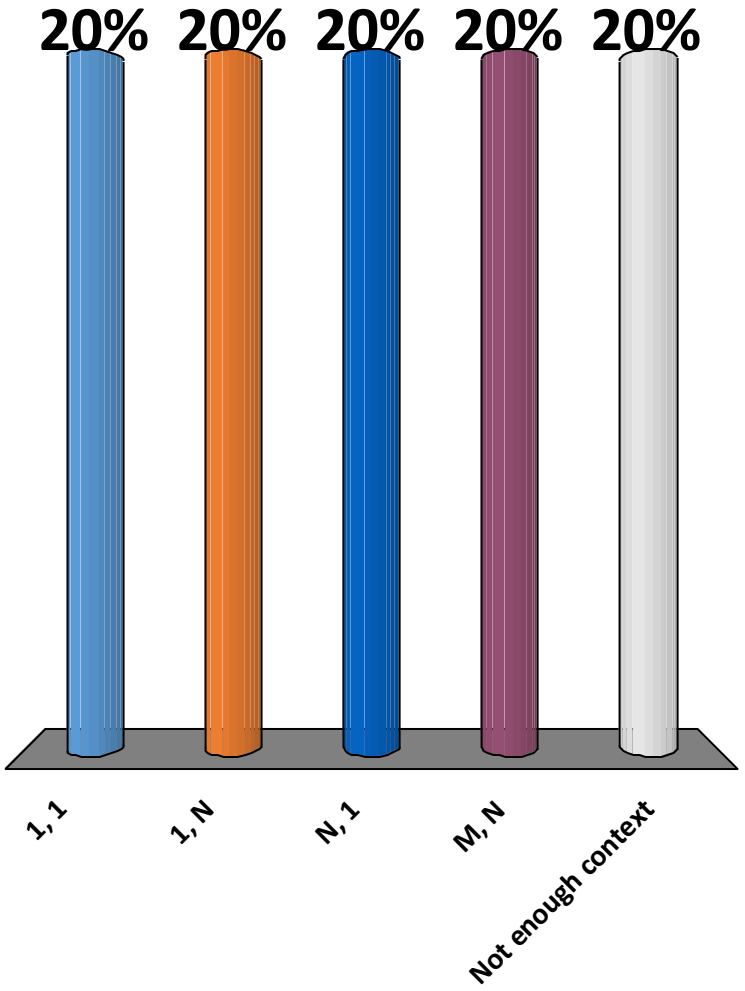
A. 1, 1

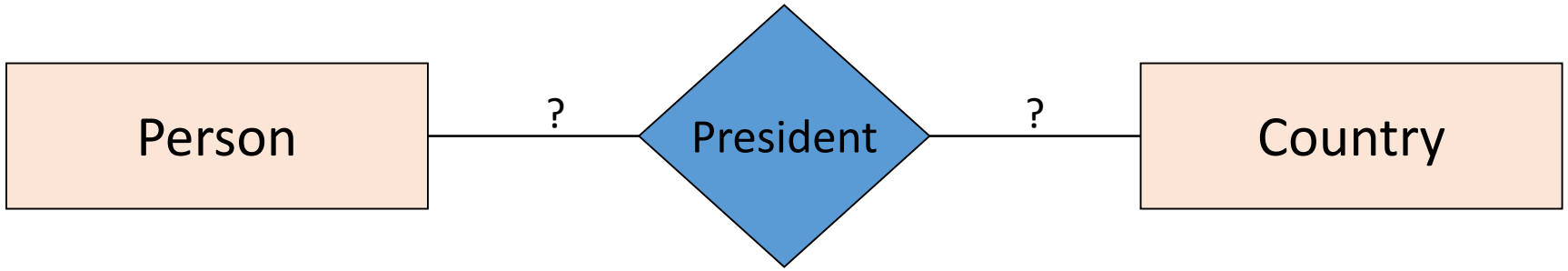
B. 1, N

C. N, 1

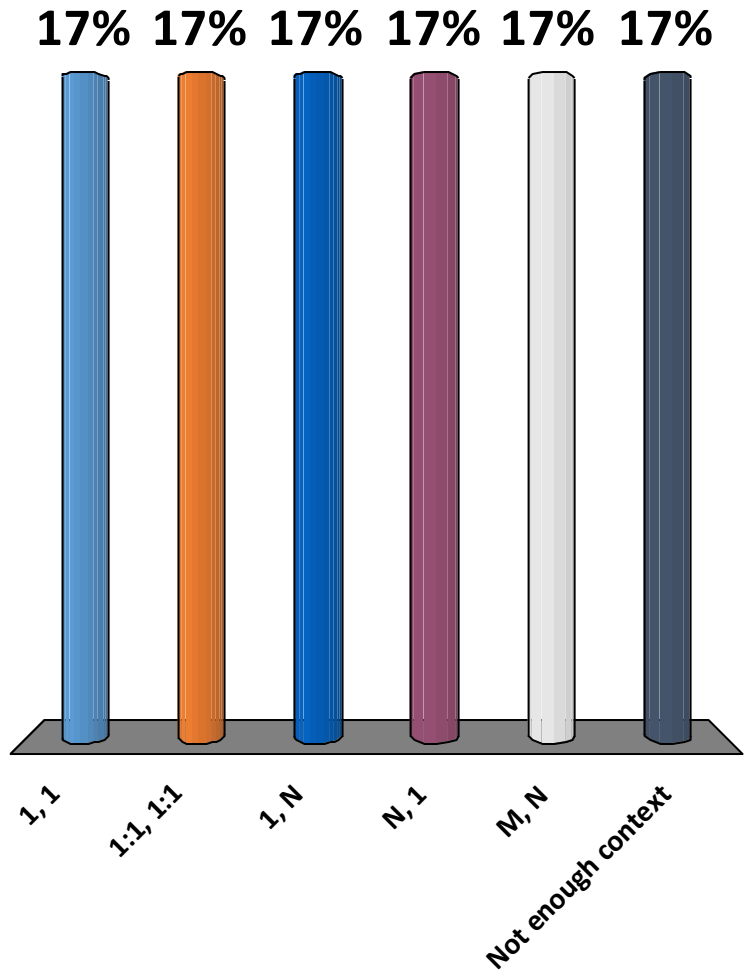
D. M, N

✓ E. Not enough context





Assume typical real-world context.



✓ A. 1, 1

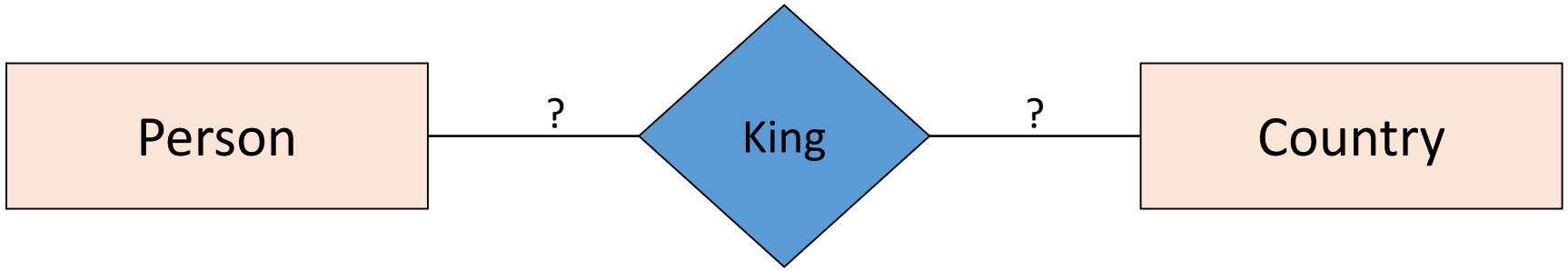
B. 1:1, 1:1

C. 1, N

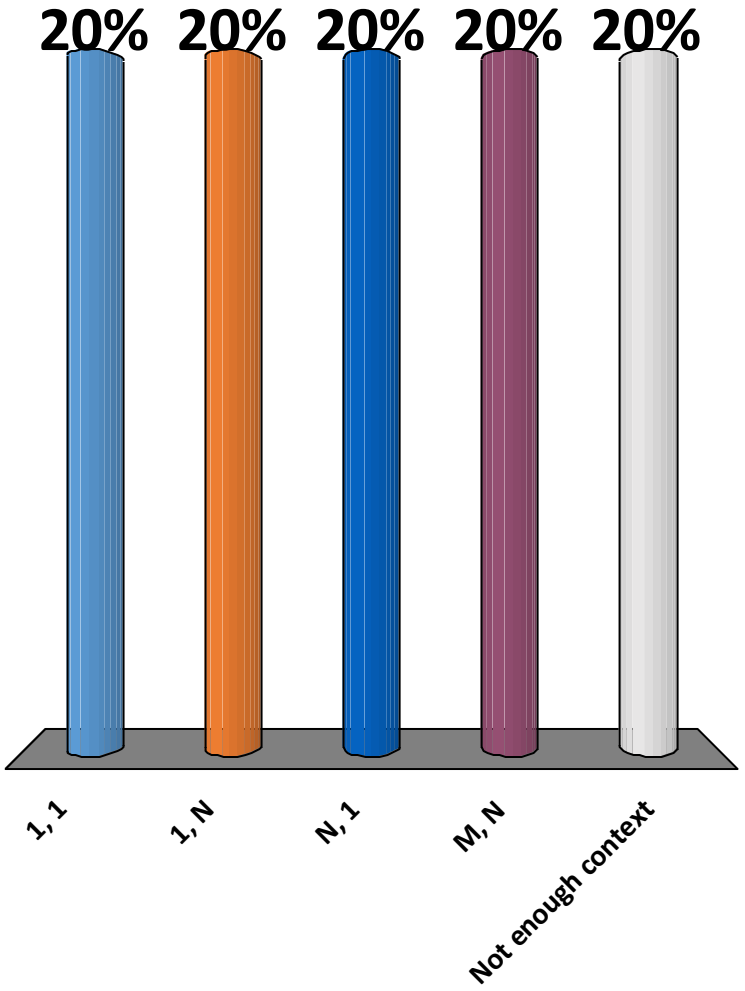
D. N, 1

E. M, N

F. Not enough context



Assume typical real-world context.



A. 1, 1

✓ B. 1, N

C. N, 1

D. M, N

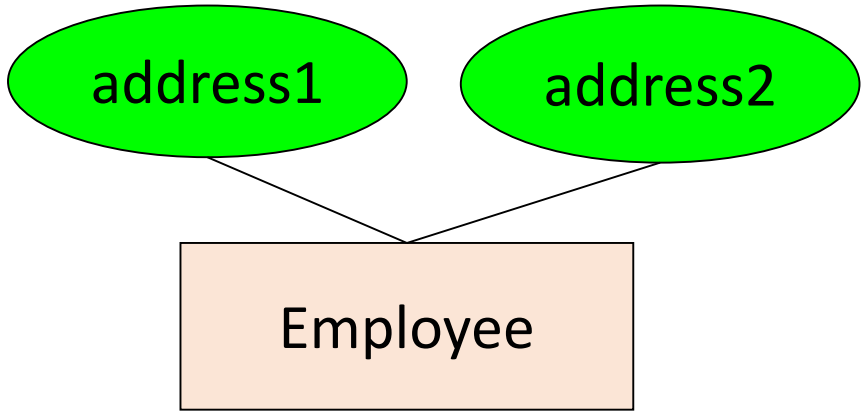
E. Not enough context

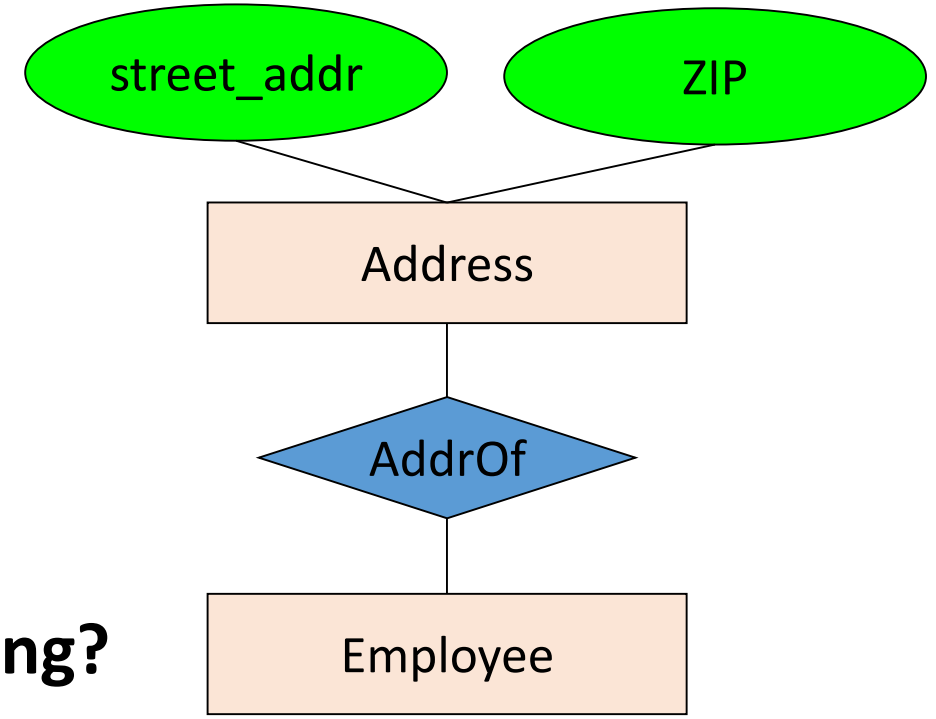
Discussion activity

Make ERD to represent people and their biological parents.

Pros/cons of different approaches?

Similar ex.: attribute vs. related entity set?





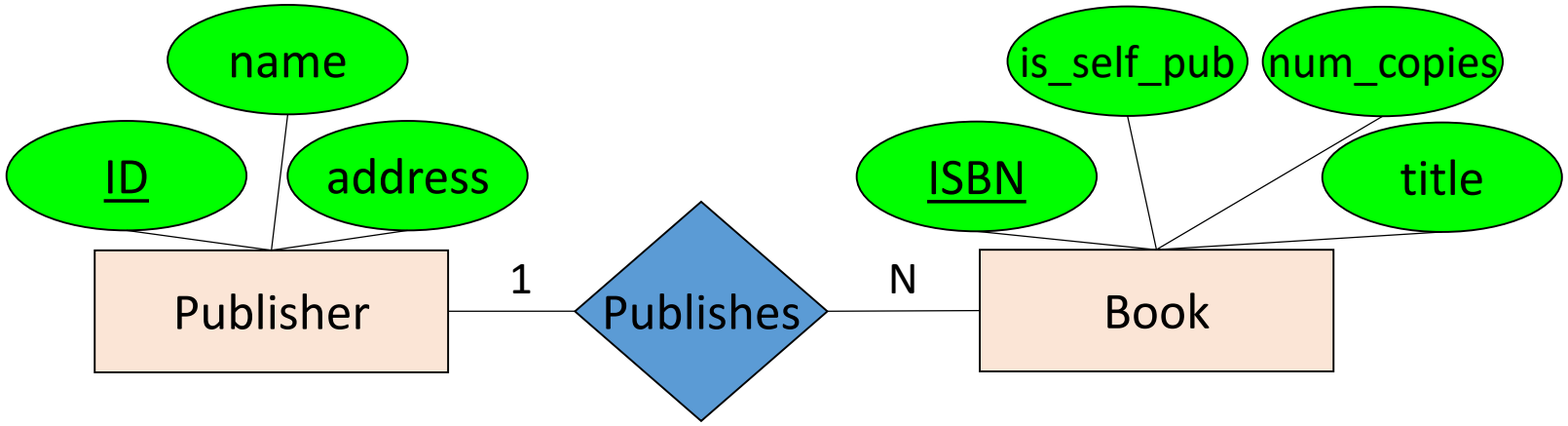
Difference in meaning?

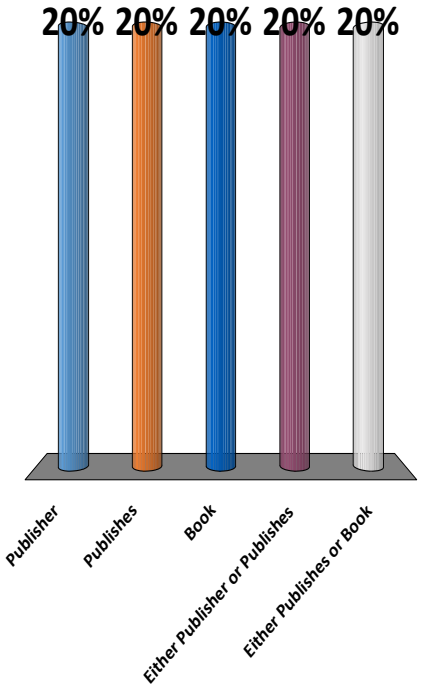
Advantages of each?

Cardinality?

Response Counter

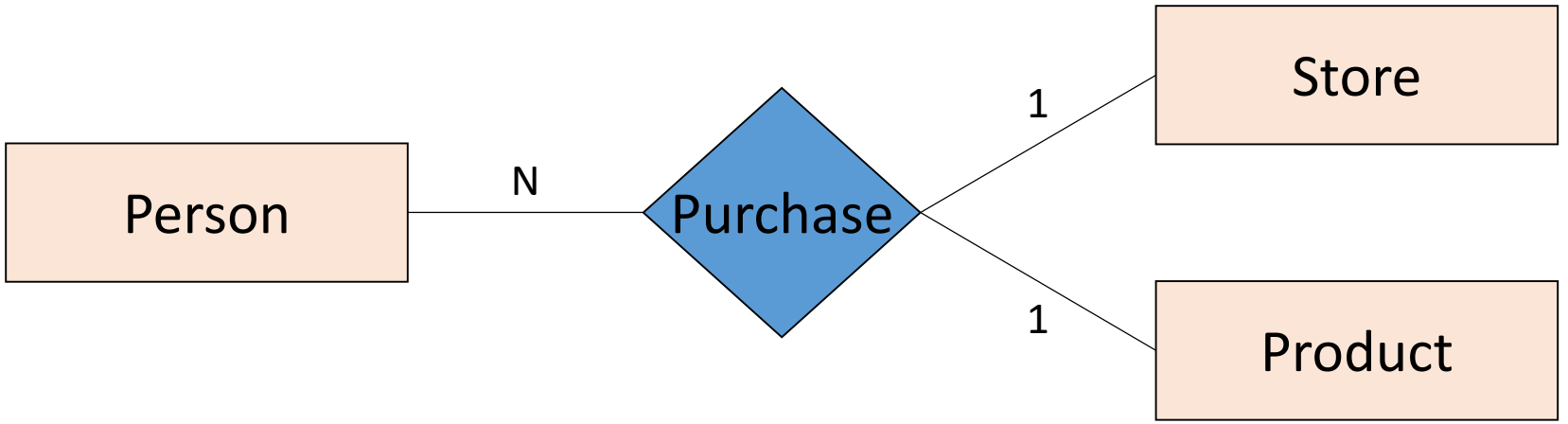
Where to add publication date attribute?





A. Publisher
B. Publishes
C. Book
D. Either Publisher or Publishes
✓ E. Either Publishes or Book

Cardinality in n-ary relationships



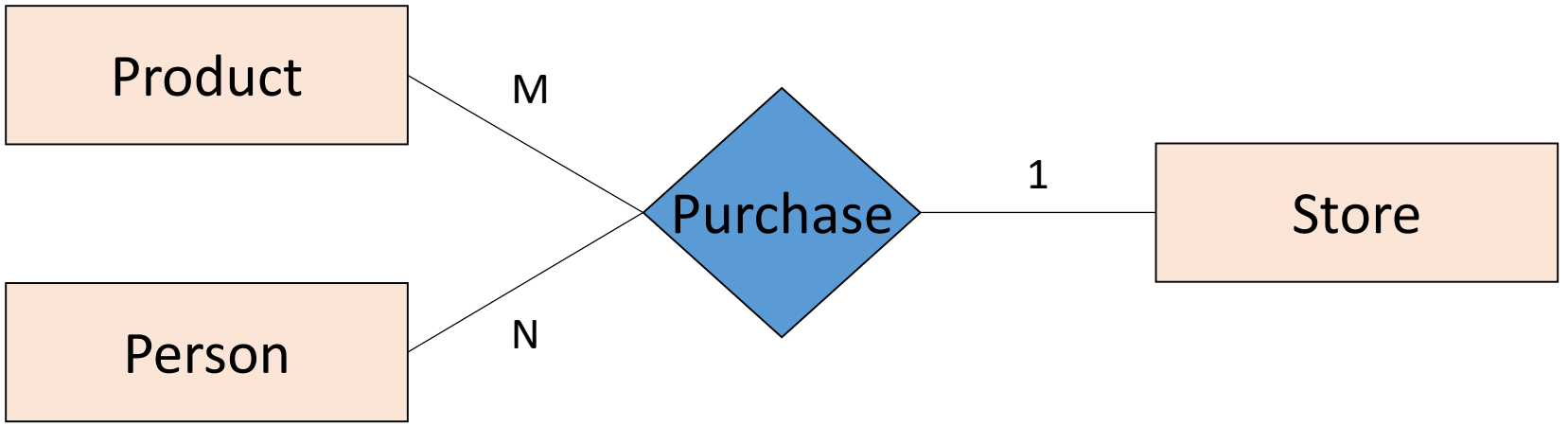
Purchase

person	store	product
Alice	Target	Jeans
Bob	Target	Shirt
Charles	Macys	Jeans
Dana	Amazon	Books

Given **Person**, then **Store** & **Product** are determined.

Each person can make one purchase – and thus of one product at one store.

Cardinality in n-ary relationships



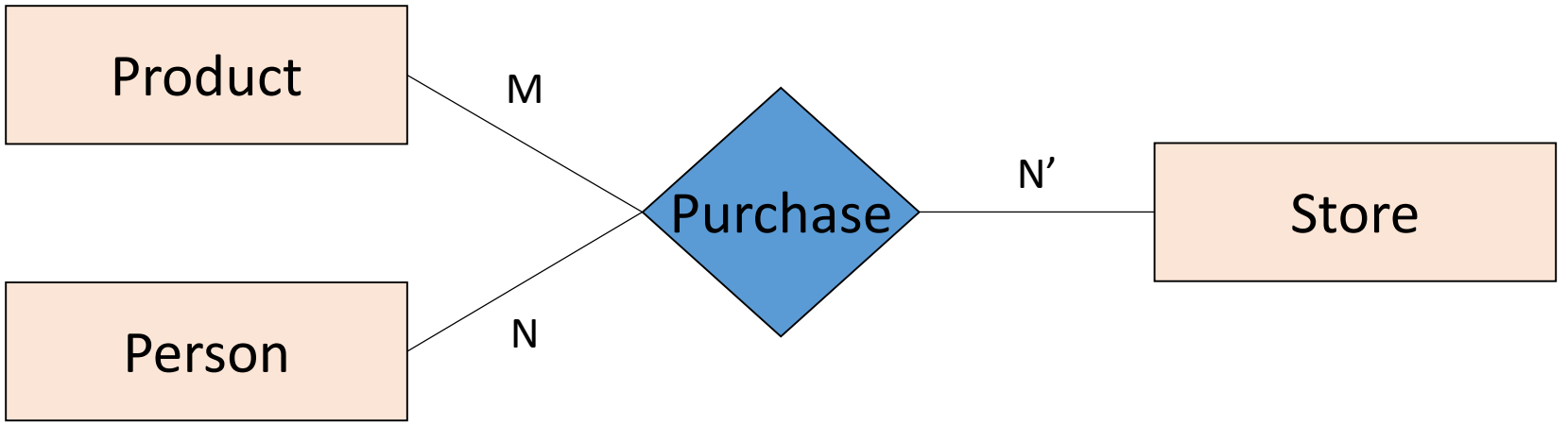
Purchase

person	store	product
Alice	Target	Jeans
Alice	Powells	Books
Bob	Target	Shirt
Charles	Macys	Jeans
Charles	Target	Shirt
Dana	Amazon	Books

Given **Product** & **Person**, then **Store** is determined.

Any person can buy any given product at most once – and thus at one store.

Cardinality in n-ary relationships

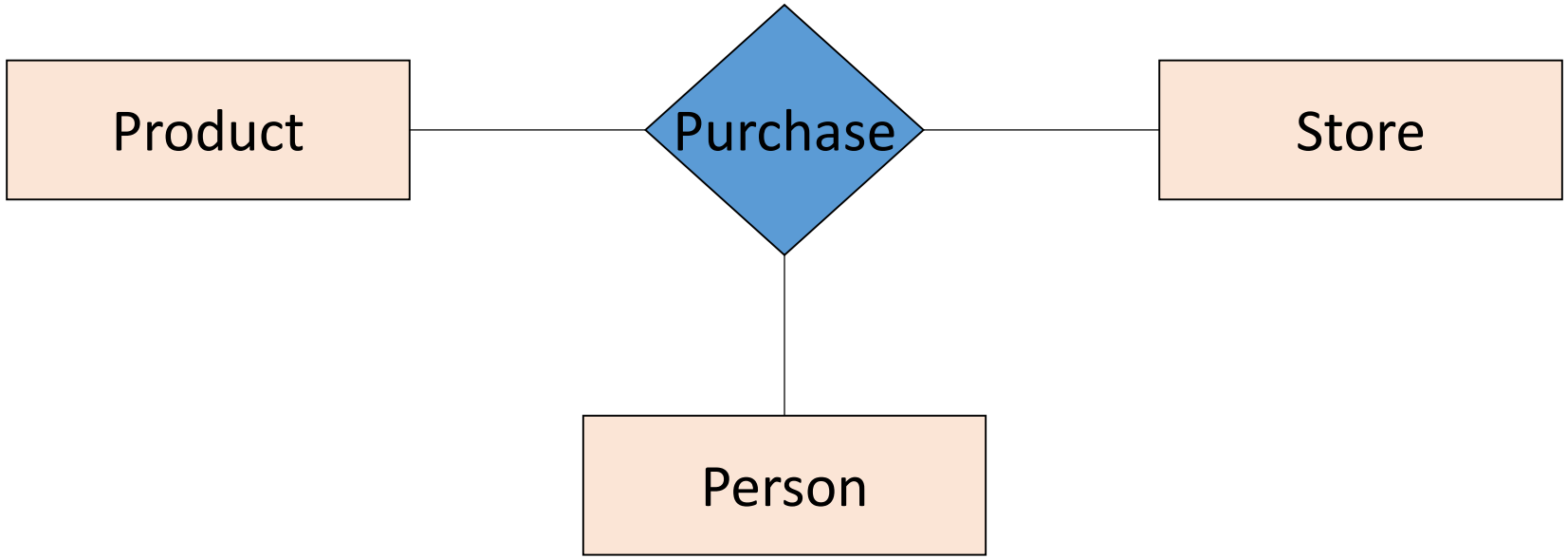


Purchase

person	store	product
Alice	Target	Jeans
Alice	Powells	Books
Alice	Target	Books
Bob	Target	Shirt
Charles	Macys	Jeans
Charles	Target	Shirt
Dana	Amazon	Books
Dana	Powells	Books

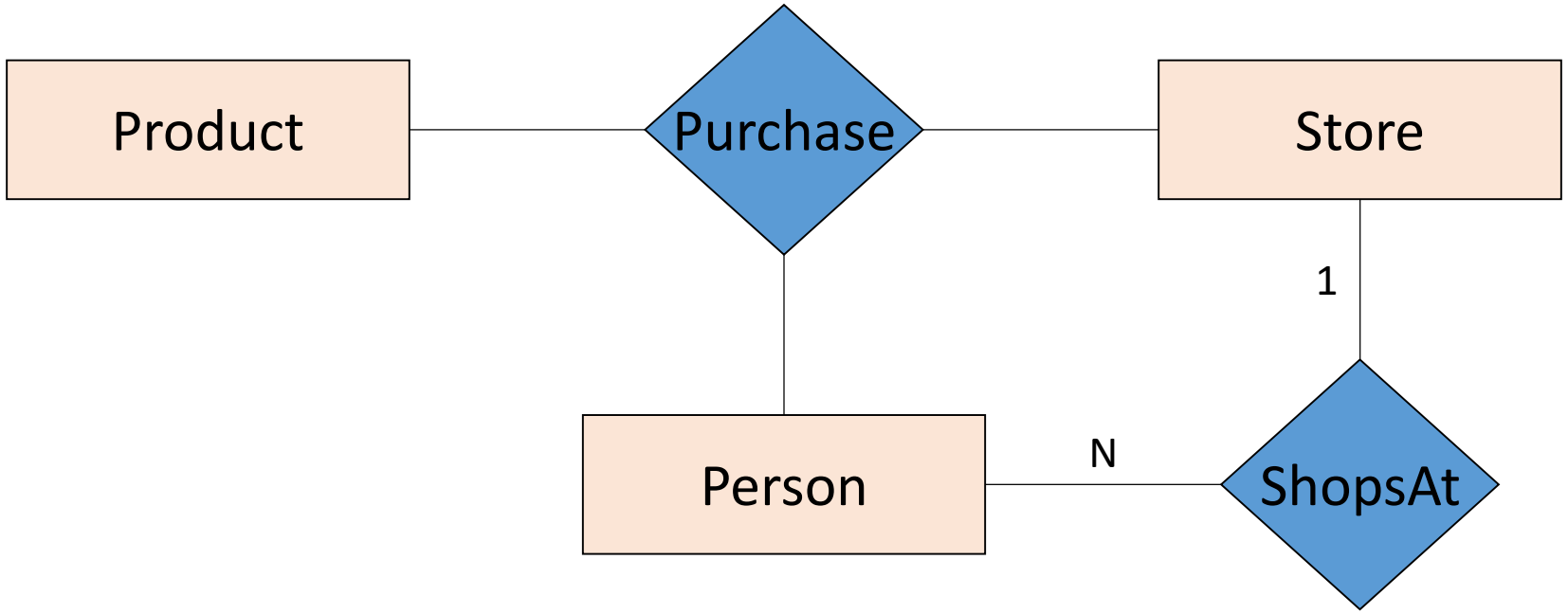
Any combination.

Cardinality in n-ary relationships



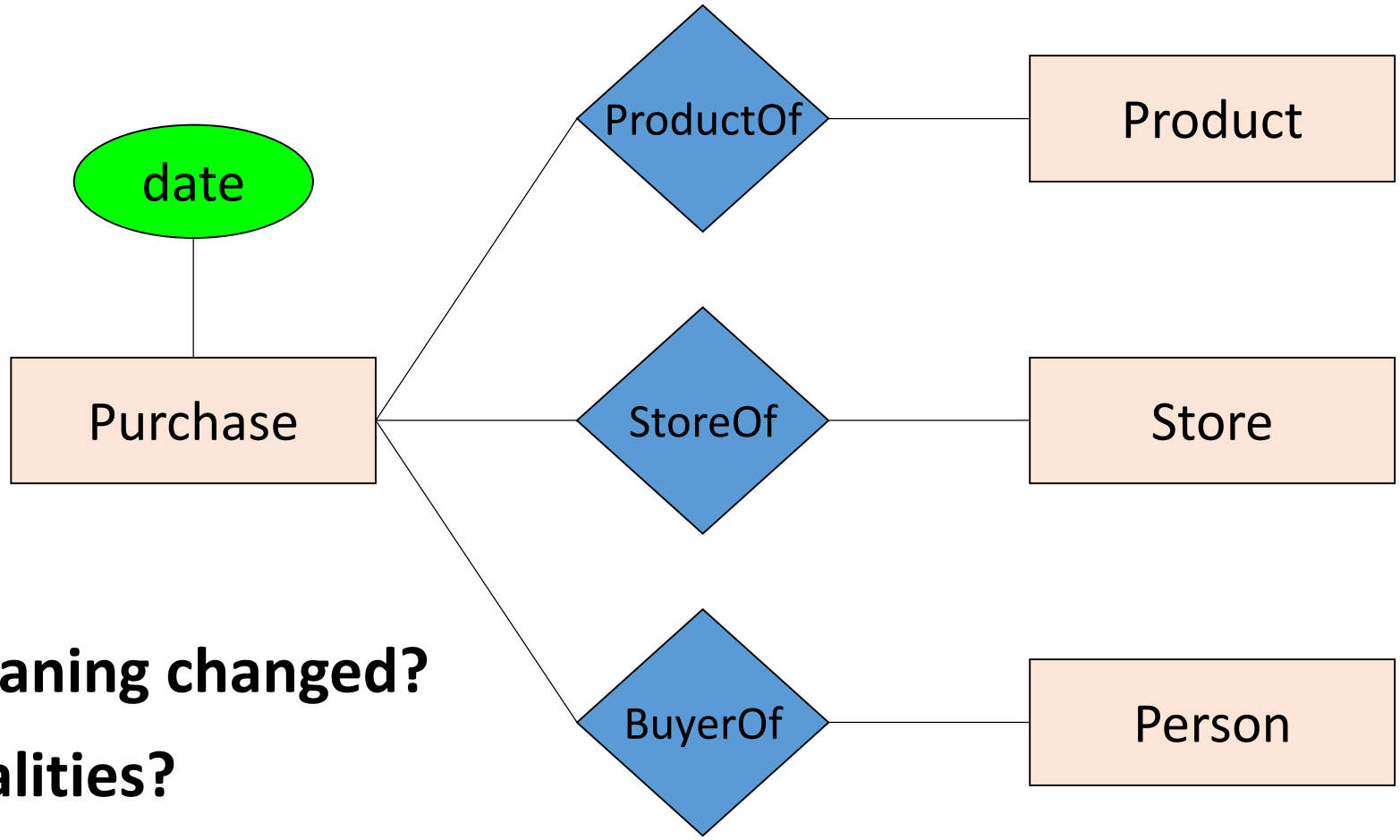
How to say: “Each person shops in at most one store.”?

Some constraints require extra relationships



How to say: “Each person shops in at most one store.”.

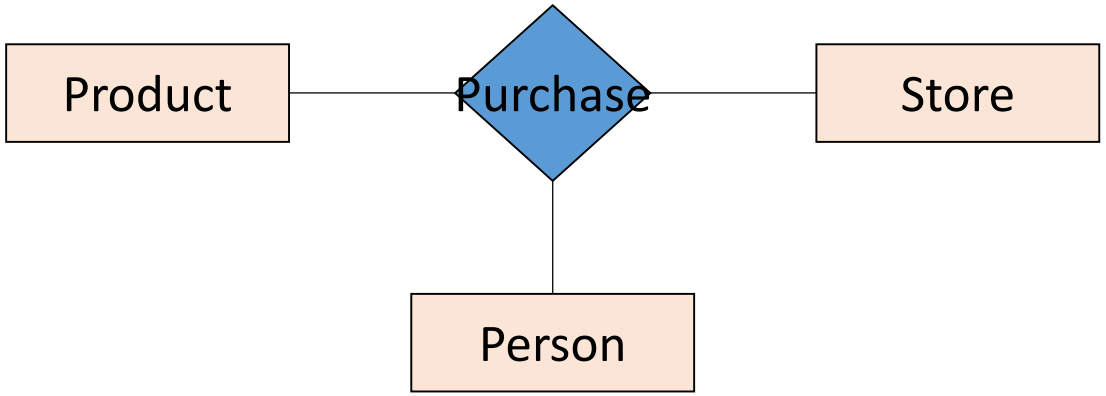
Can convert n-ary to binary

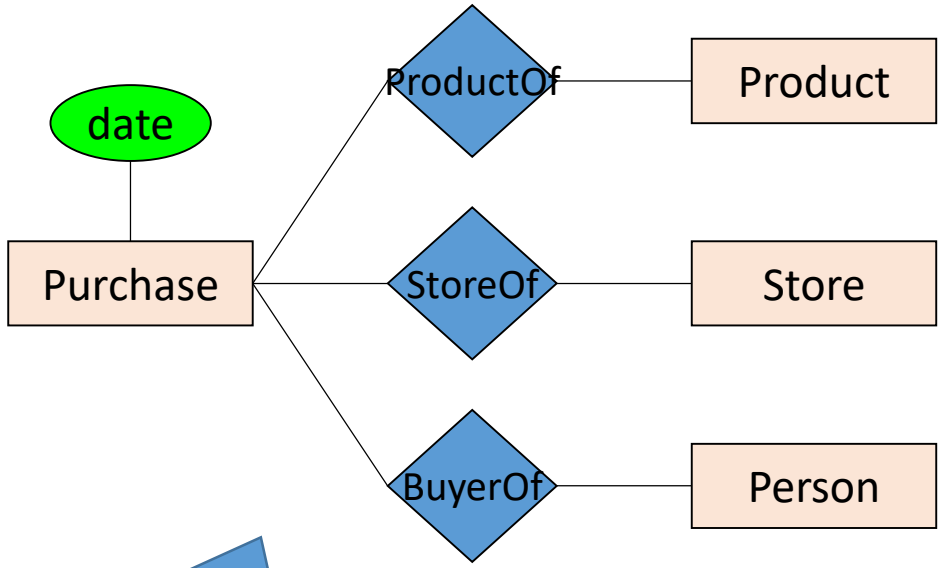


How has the meaning changed?

Cardinalities?

Decision: n-ary or binary?

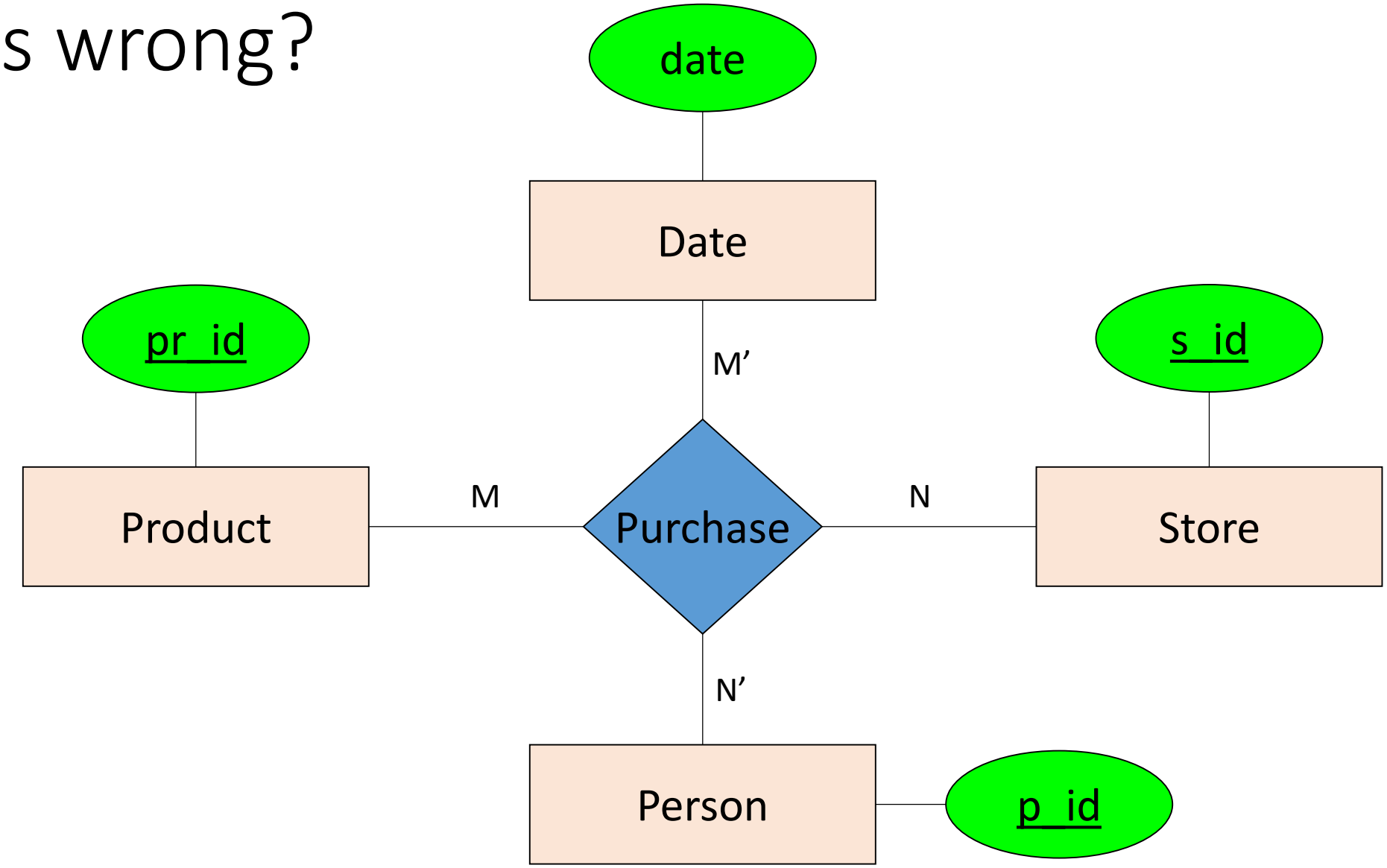




Generally best when relationship really is between multiple entities.

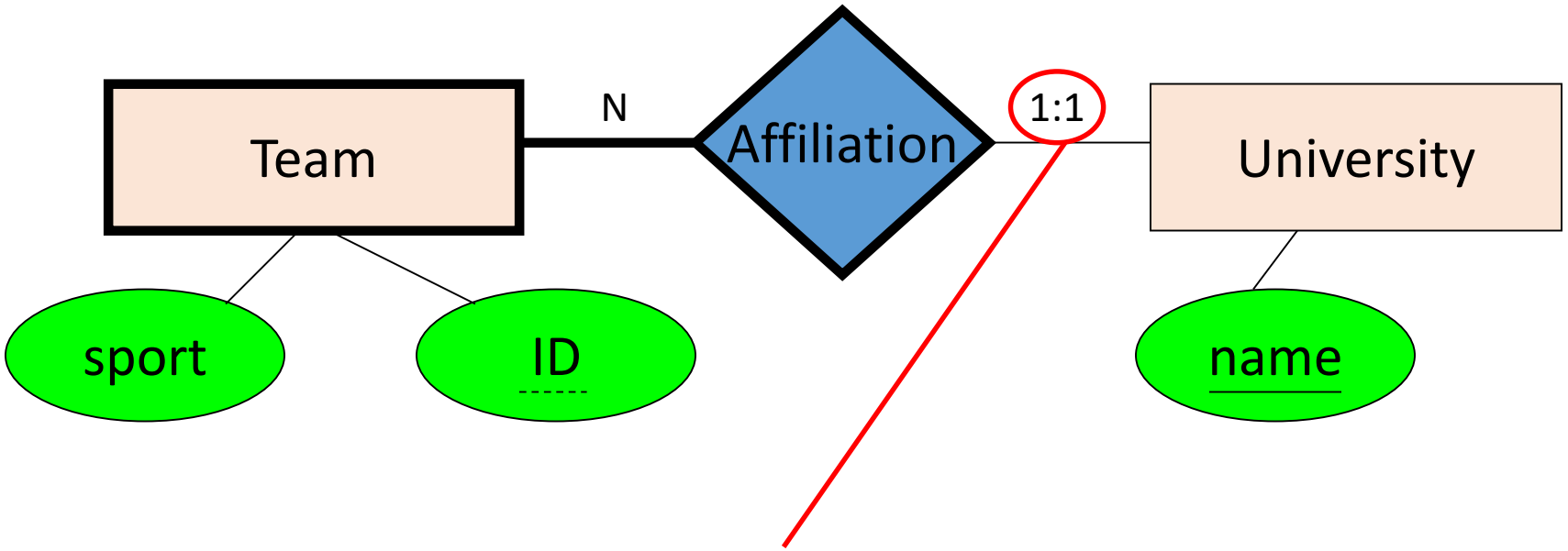
- Allows multiple purchases per Product-Store-Person combination.
- Allows more control on cardinality.
  - “A person who shops in only one store.”

What's wrong?



Weak entity set

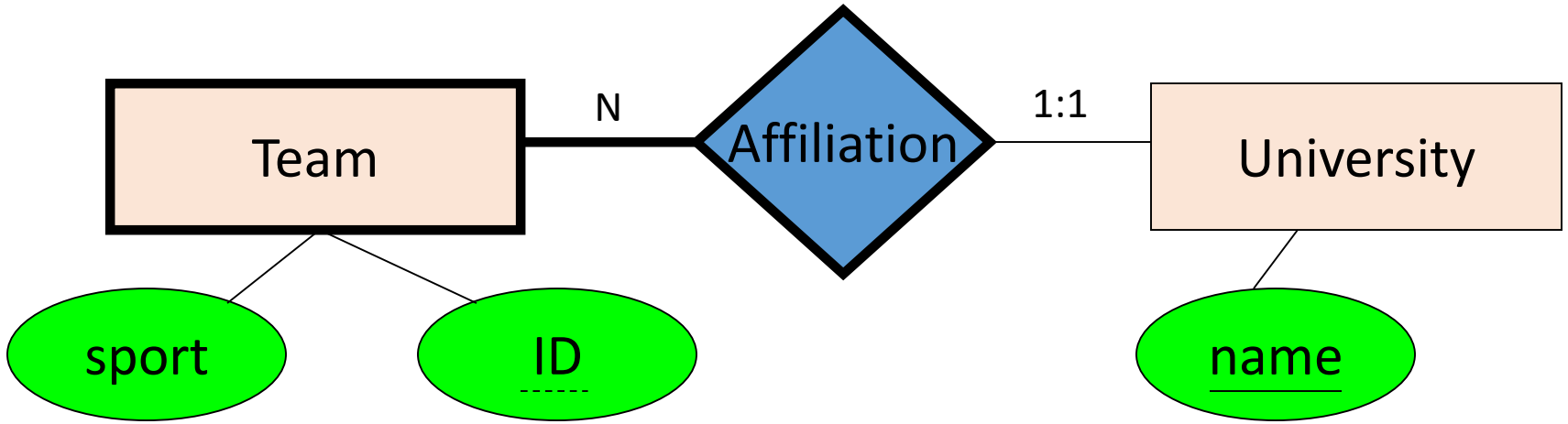
In a relational database, a weak entity is an entity that cannot be uniquely identified by its attributes alone; therefore, it must use a foreign key in conjunction with its attributes to create a primary key



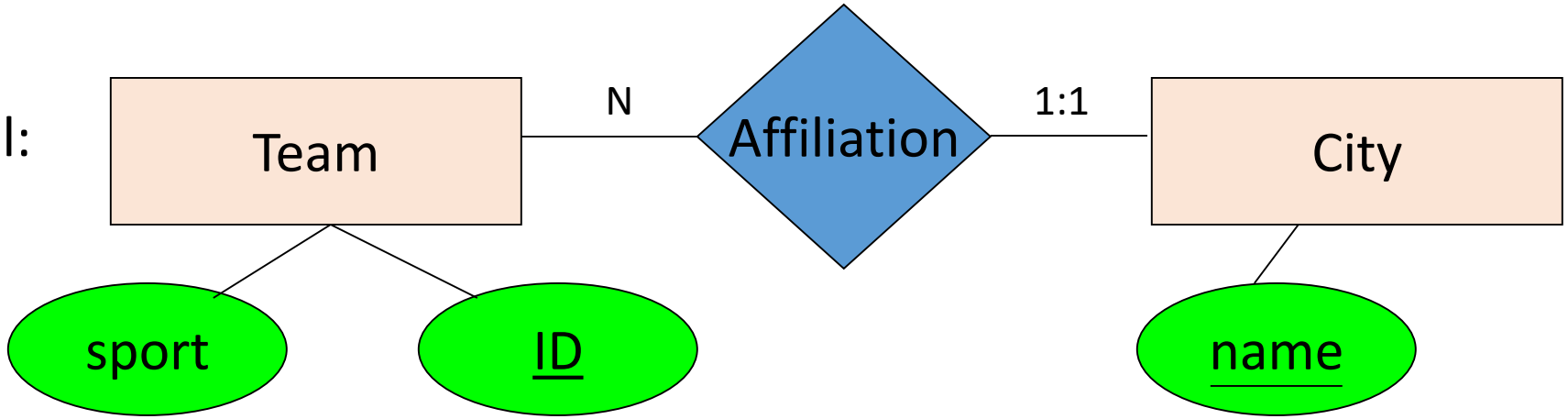
- Existence/meaning is dependent on another entity set(s).
- Part of its key comes from that other entity set(s)

Weak entity – a **subtle** semantic distinction

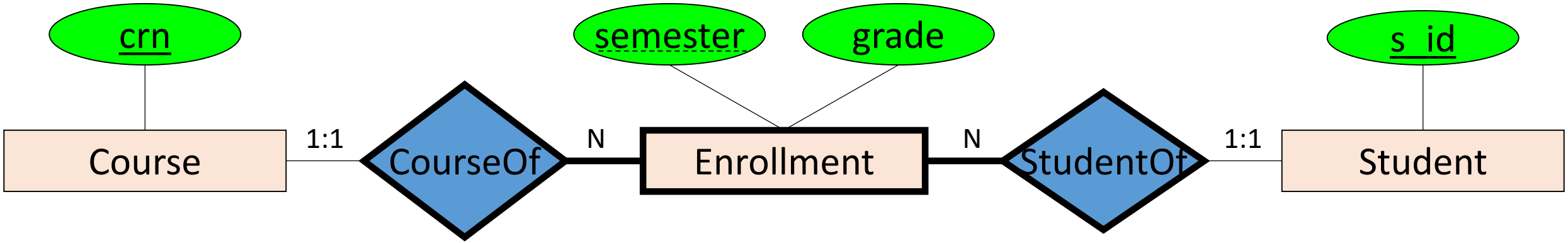
College:



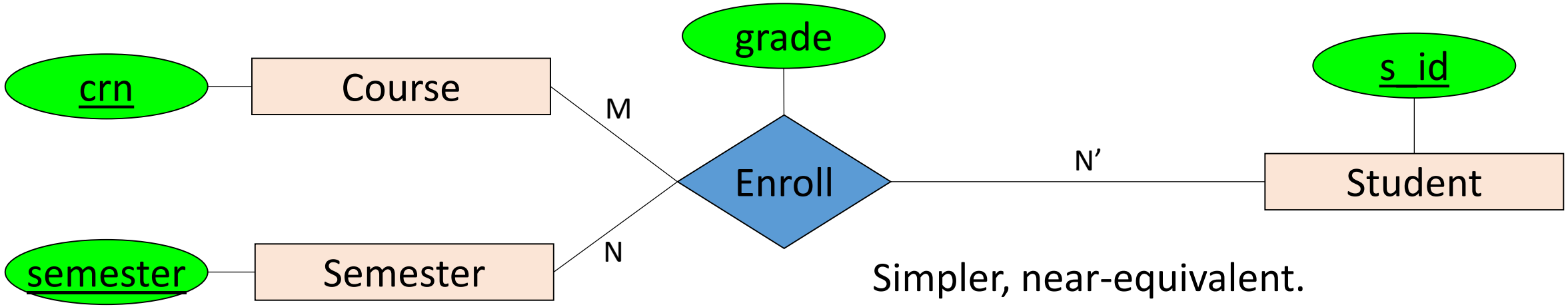
Professional:



Many-to-many junctions often weak

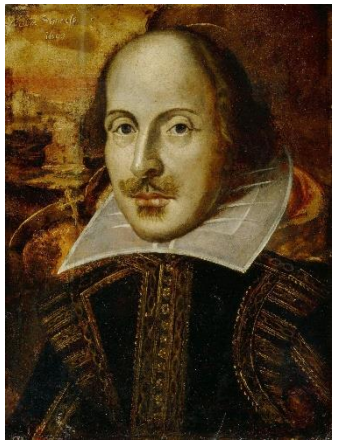


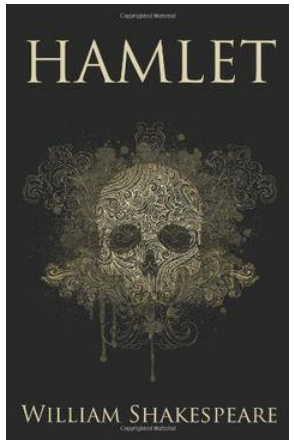
Enrollment isn't interesting/useful away from its connections to **Course** & **Student**.



Simpler, near-equivalent.

Activity – Add multiplicity to ER diagram





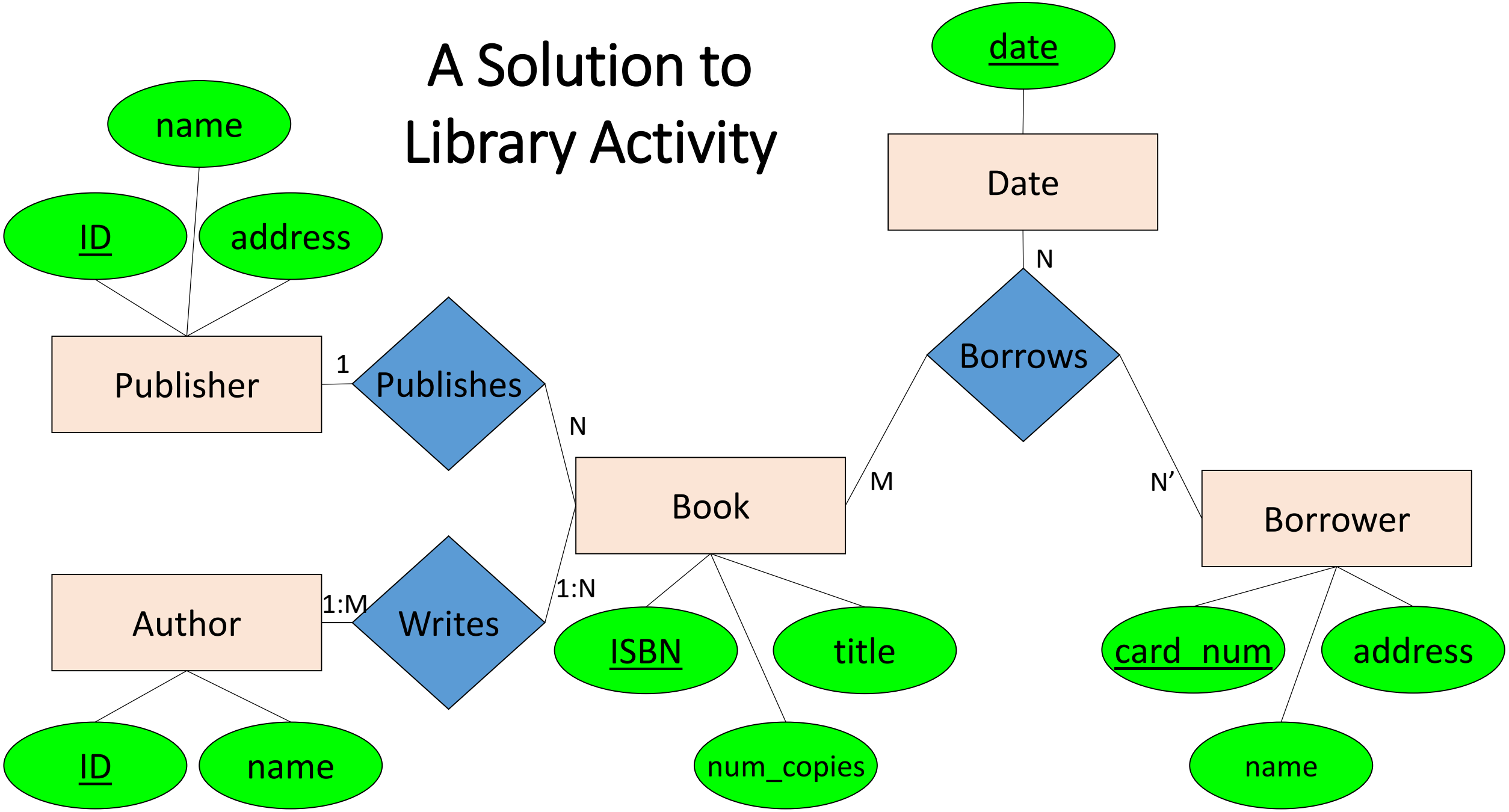
Authors have IDs and names. They write books.

Books have ISBNs and titles. The library keeps track of how many copies it has of the book. Each book is written by authors and published by a publisher. We want to know every time it is checked out by a borrower.

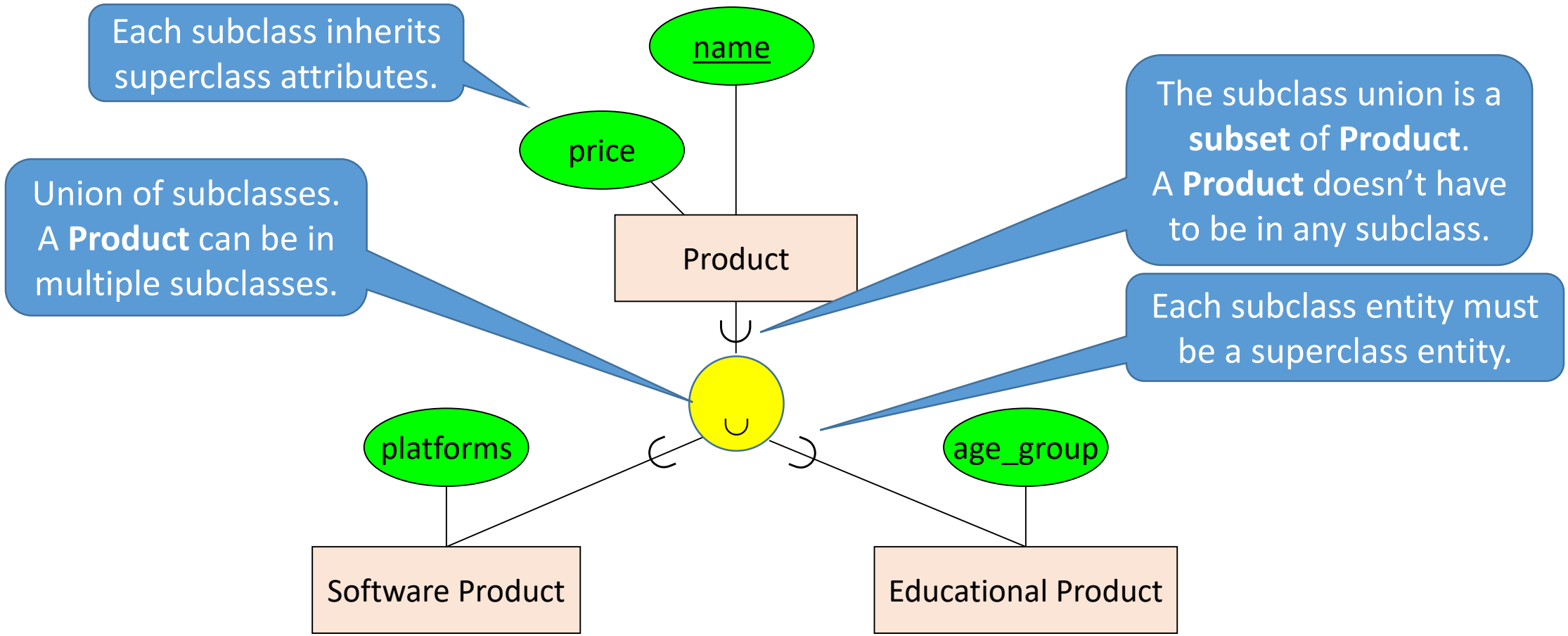
Borrowers have a library card number, name, and address. They can check out a book on a particular date.

Publishers have an ID, name and address. They publish books.

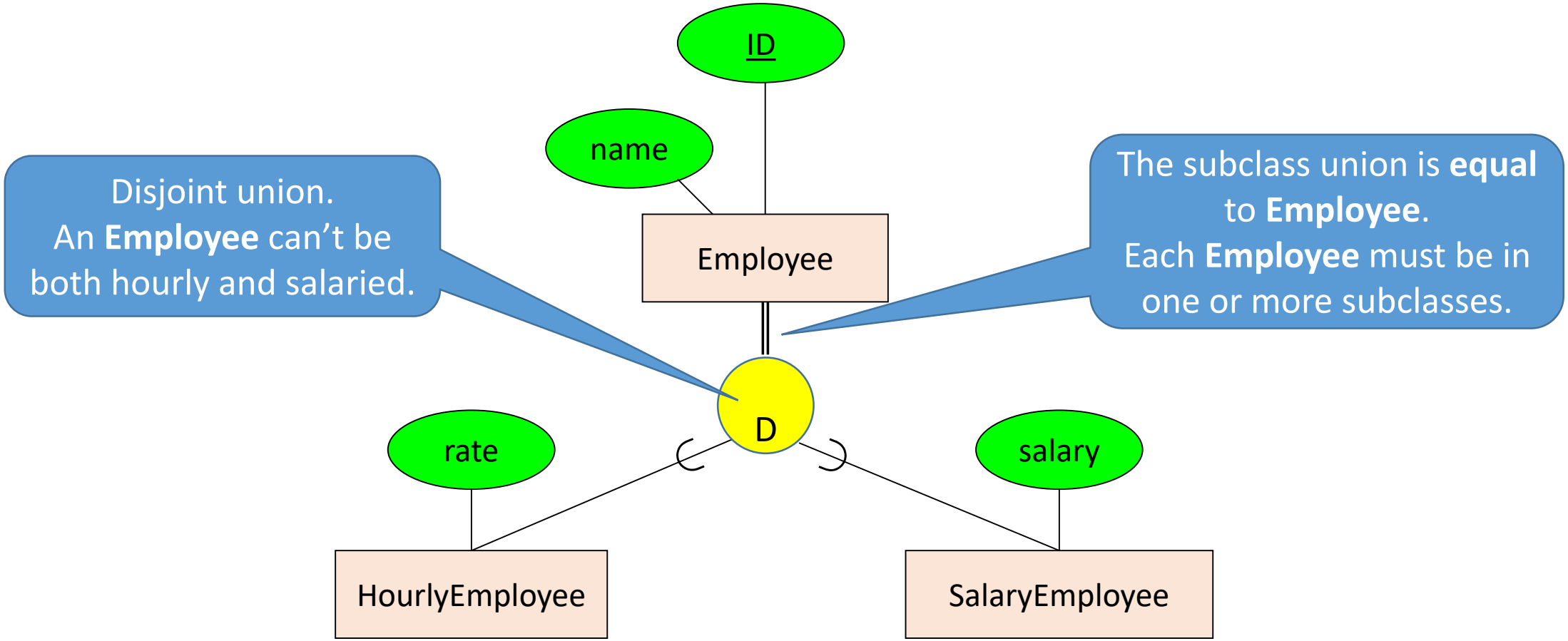
A Solution to Library Activity



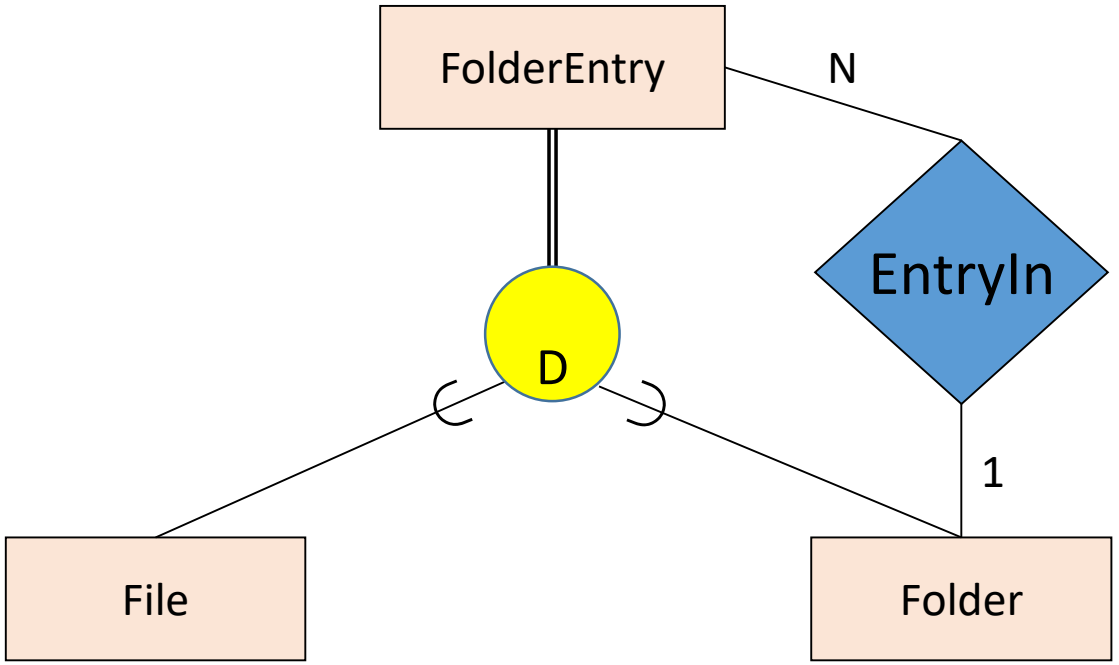
Subclasses: unions



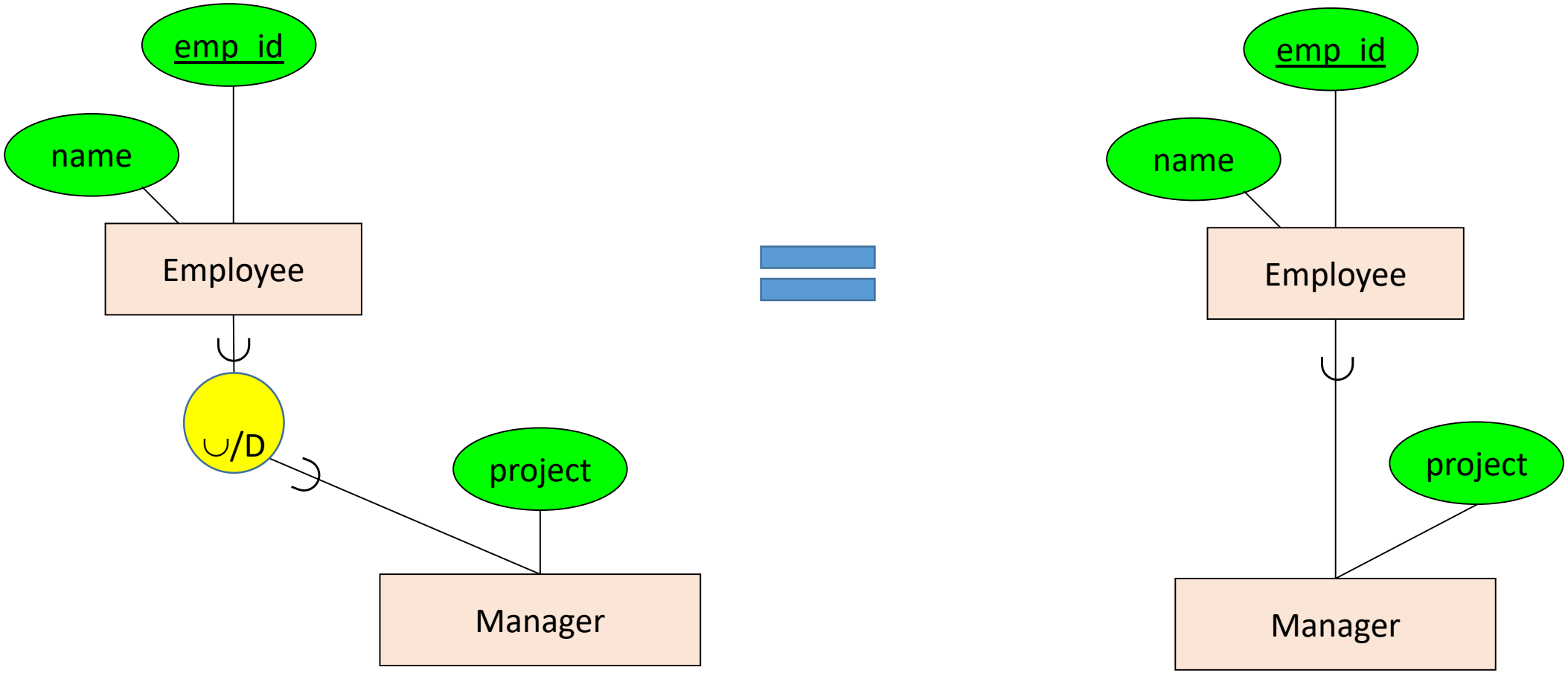
Subclasses: disjoint unions, equality



Composite pattern example

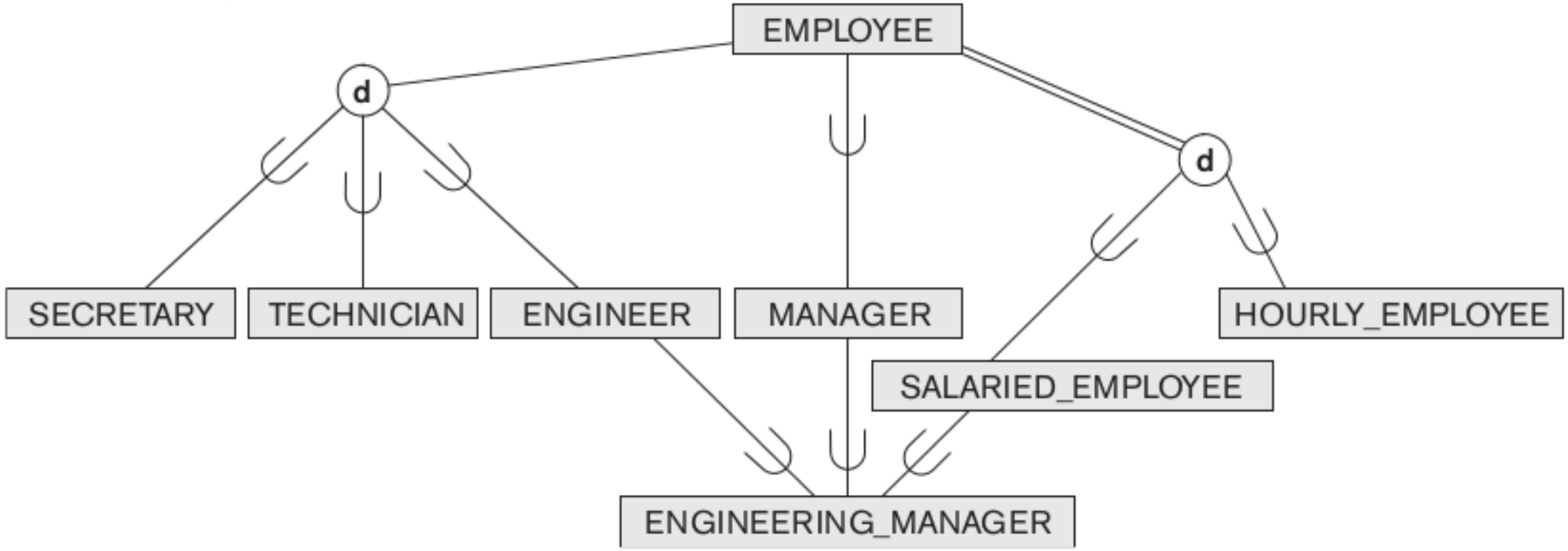


Shorthand for singleton subclass

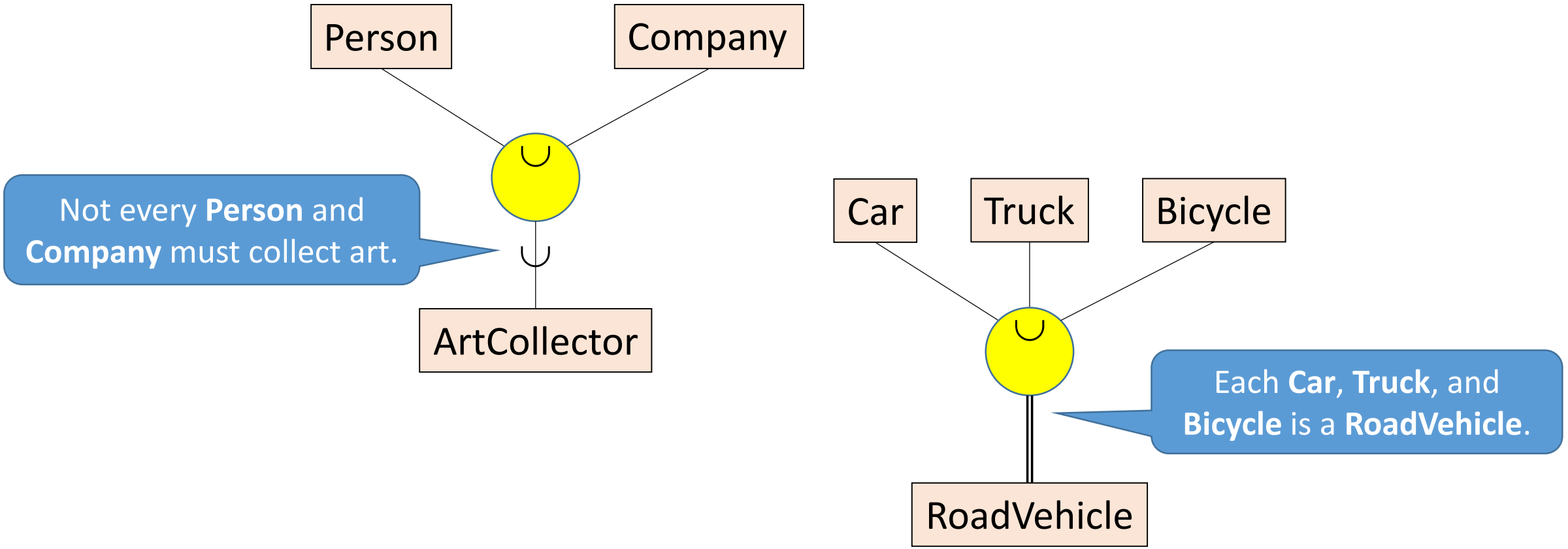


Subclasses: multiple inheritance hierarchy

Figure 8.6
A specialization lattice with shared subclass ENGINEERING_MANAGER.

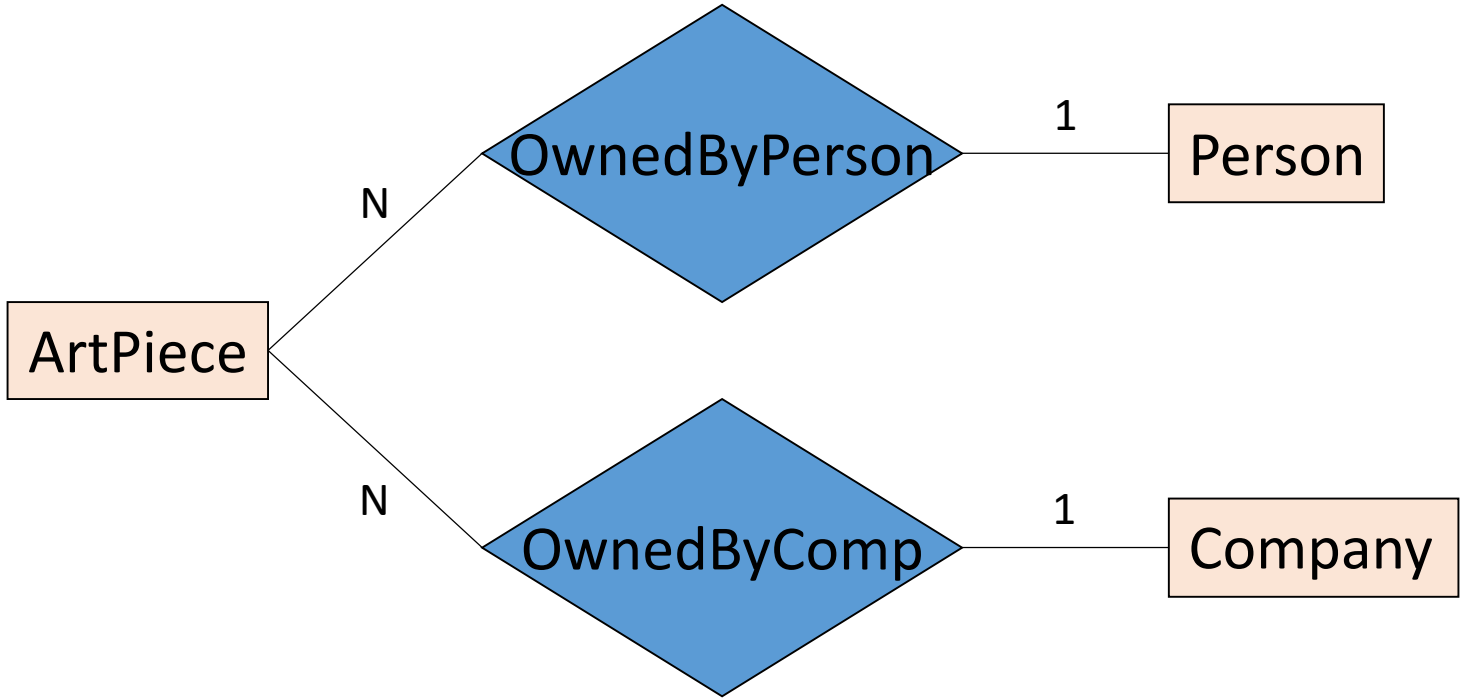


Subclasses: unions, equality – “categories”



The need for categories/unions

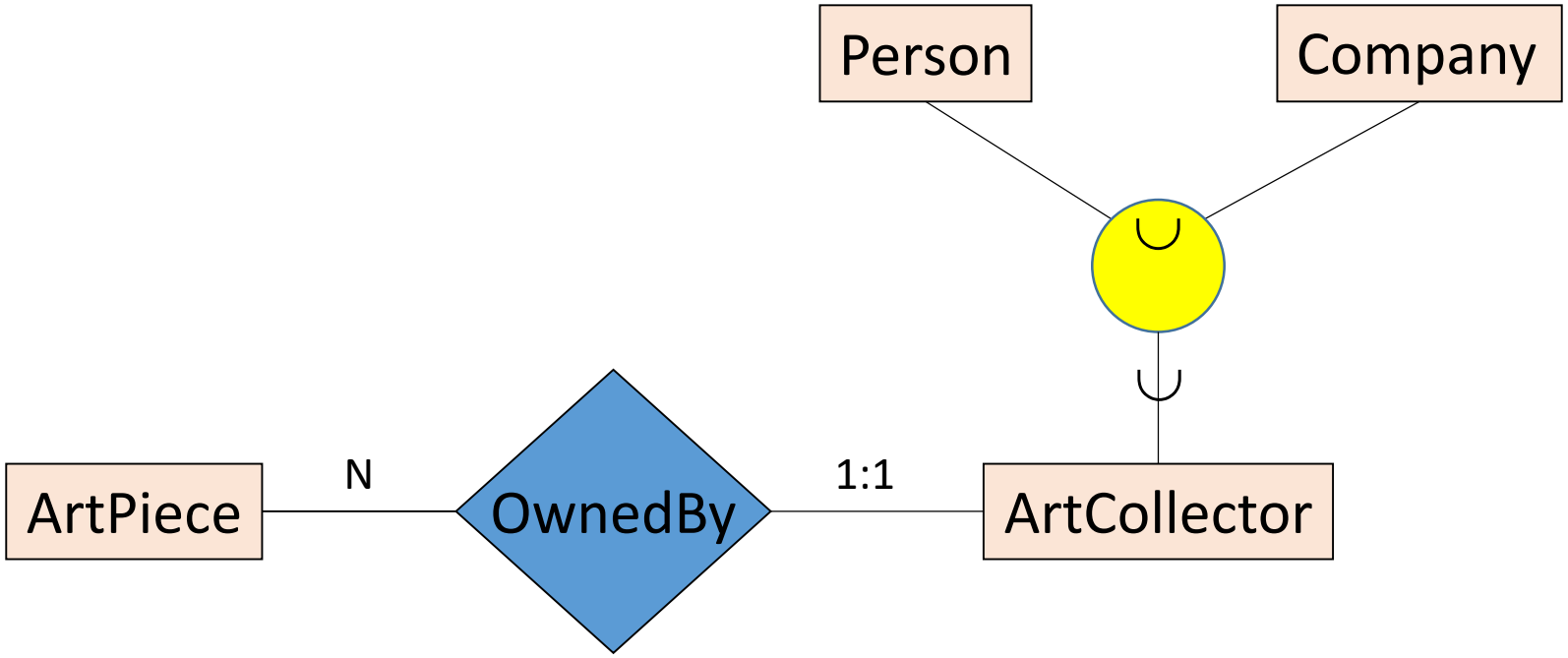
Goal: “Every art piece is owned by a person or company.”



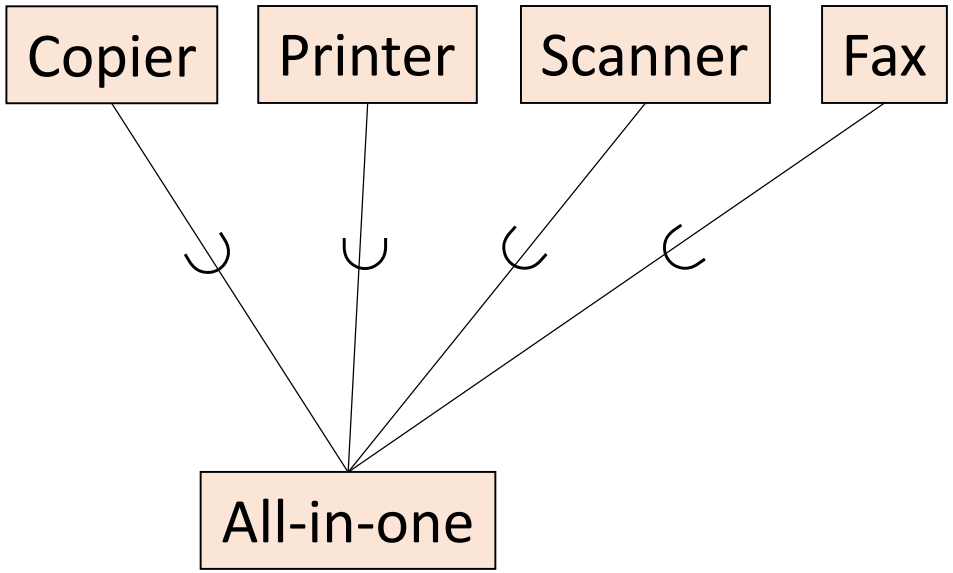
What is the problem?

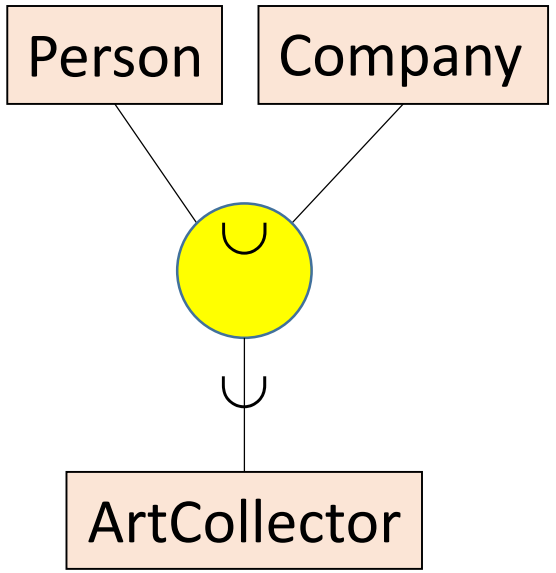
Solution with categories

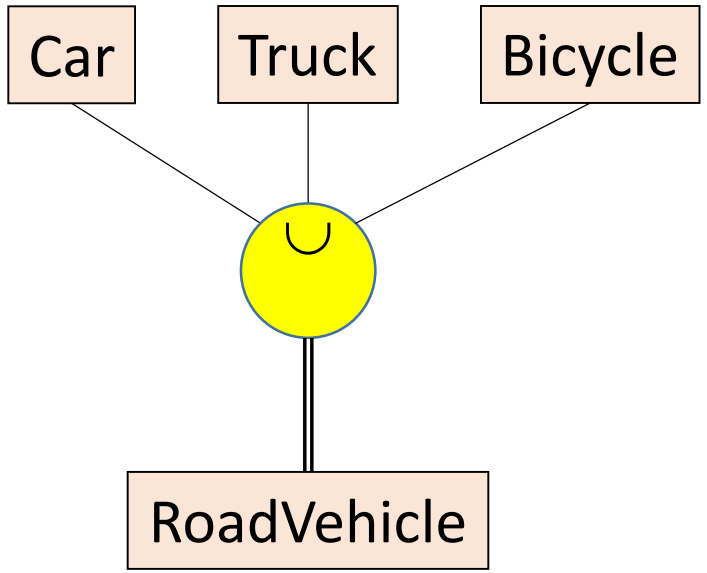
Goal: “Every art piece is owned by a person or company.”



Different options for multiple superclasses



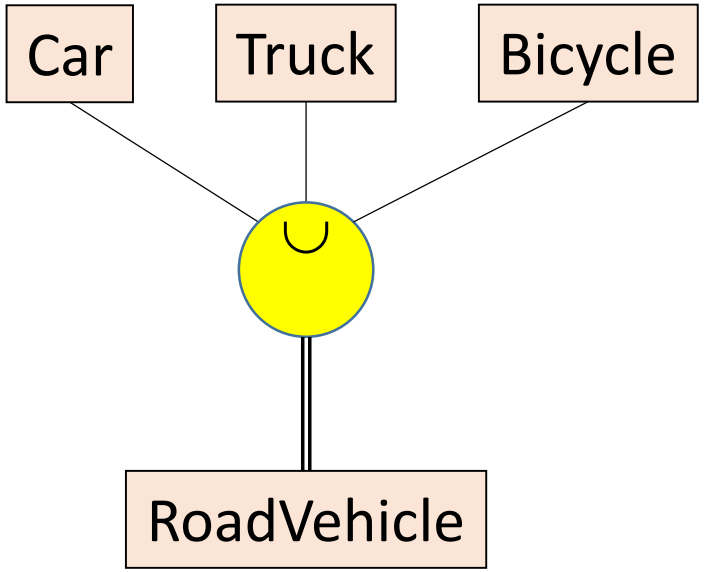


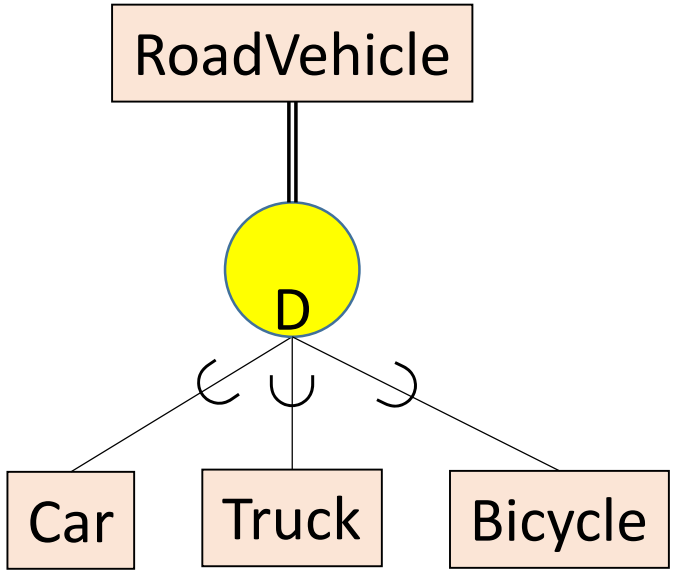


Each subclass entity belongs to & inherits from **all** superclasses.

Each subclass entity belongs to & inherits from **the appropriate one** superclass.

What's the difference?





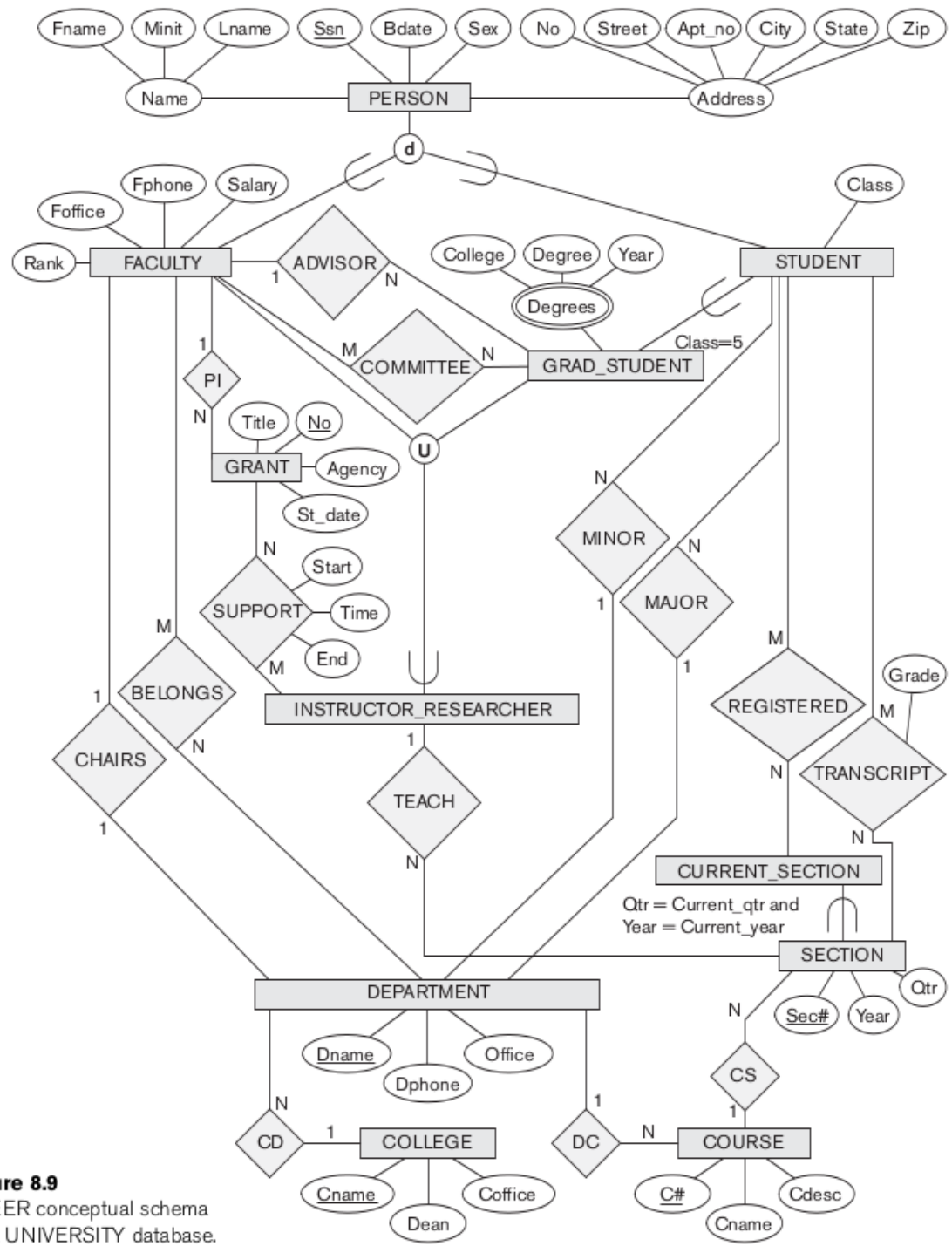


Figure 8.9
An EER conceptual schema for a UNIVERSITY database.

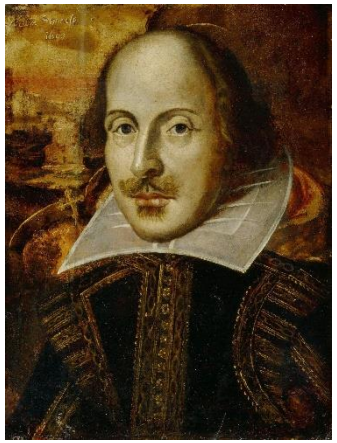
ER design summary

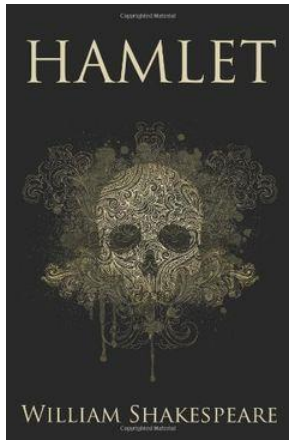
Semi-formal design technique based upon informal semantics

Goal: Schemas that accurately represent and formalize semantics

1. Design ER diagram
   a. Identify entity sets, their attributes, and sub/superclasses
   b. Identify relationships between entity sets, and their attributes
   c. Identify max. & min. cardinality of relationships
   d. Identify weak entity sets
2. Convert to schemas – next topic
3. Normalize schemas – coming soon

Activity – Add participation, sub/superclasses





Authors have IDs and names. They write books or chapters of book.

Books have ISBNs and titles. The library keeps track of how many copies it has of the book. Each book is written by authors and published by a publisher. We want to know every time it is checked out by a borrower.

Borrowers have a library card number, name, and address. They can check out a book on a particular date.

Publishers are typically a company with an ID, name and address. They publish books. Alternatively, authors can self-publish.

A Solution to Library Activity