Database Systems
Lecture #3

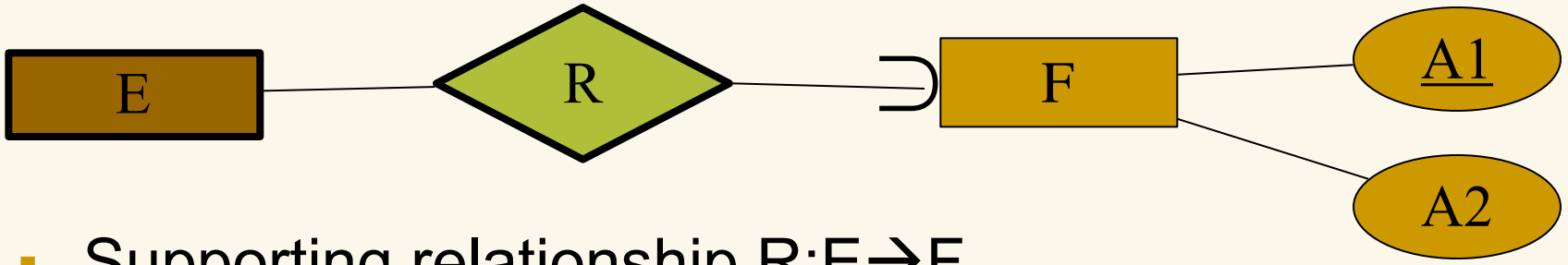
Agenda

- Last time:
  - A little on design
  - (nearly) finished E/R models
- This time:
  - Finish E/R
  - Constraints (some review)
  - Relational model
  - Converting E/R to relations
- Next time: Functional dependencies （Chapter19）

Quick topic: Weak entity sets

- Def: some or all key attributes belong to another ES
- Plays role in a connecting relationship
- The key consists of:
  - Possibly its own attributes and
  - All key attributes of entity sets from supporting relationships

Conditions for supporting relationships

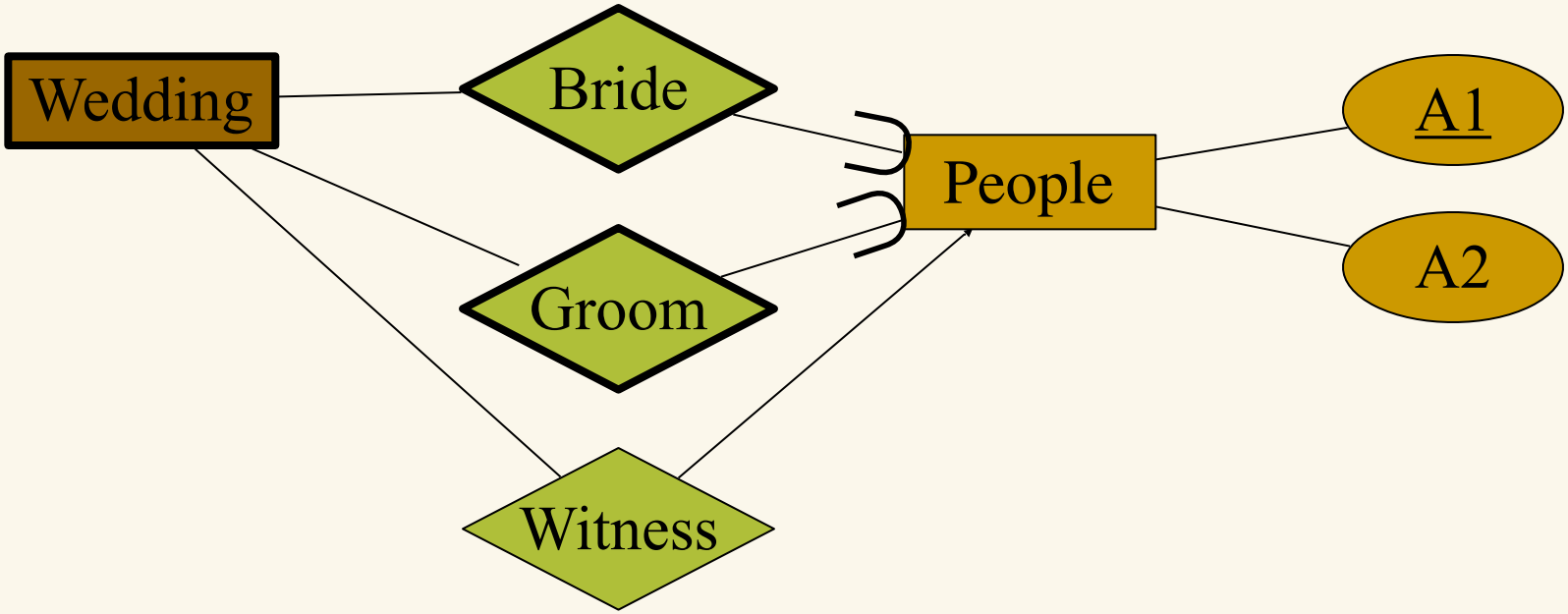


- Supporting relationship R:E→F
  - R is many-one (or 1-1) E→F
  - With referential integrity (rounded arrow)
  - R is binary
  - E receives key attributes of F

- F itself may be weak
  - Another entity set G, and so on recursively…

Conditions for weak entity sets

- For several supporting relships from E to F
  - Keys of each F role appear as foreign key of E



- Other, non-supporting many-one relationships are not affected

Weak entity set e.g.

- Example: Hierarchy – species & genus
- Idea: species name unique *per genus only*

- 物种(Species)名称一样，只能根据基因(Genus)不同区分

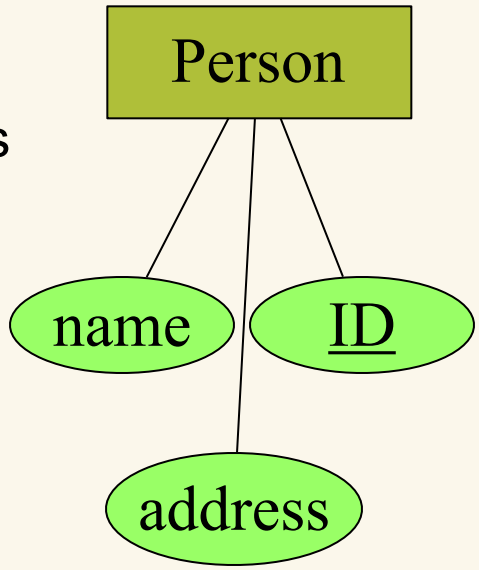
Next topic: Constraints (约束)

- Review: programmer-defined rules stating what should always be true about consistent databases
- Restrictions on data (egs?):
  - Keys
  - Single value constraints
  - Referential Integrity (参照完整性)
  - Domain constraints
  - General constraints

- Can't infer (推断) constraints from data
  - may hold "accidentally"
  - but they are a part of the *schema*

E/R keys

- Uniquely identify entity in ES
- Attribute or *set of* attributes
  - Two entities cannot agree on all key attributes
  - These attributes determine all others



- Every ES should have a key
  - possibly including all attributes

- <u>Primary key</u> attributes underlined
- More than one possible key:
  - *Candidate* keys, *primary* key

- Practical tip: create art key attribute
  - E.g. ID, course-id, employee-id, etc.
  - ID shorter than (name,address)

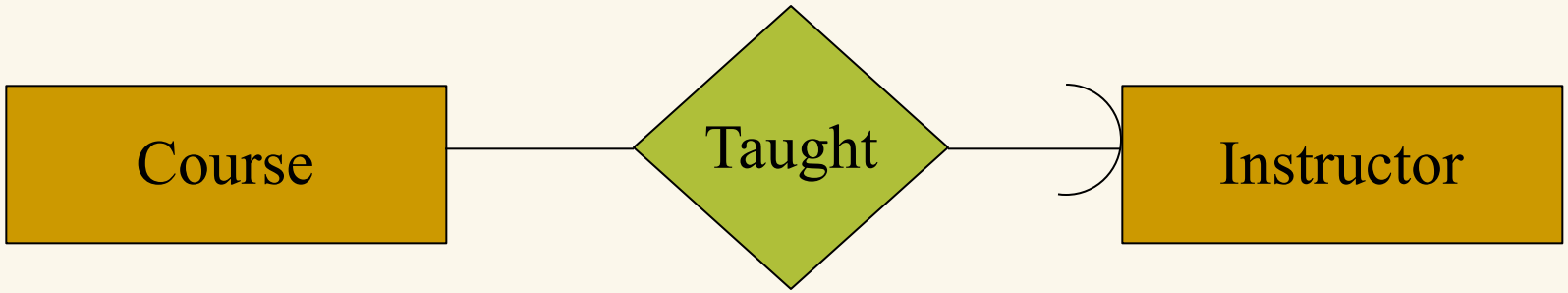
Single-valued constraints

- "at most one" value
  - Already saw sharp arrows for relationships

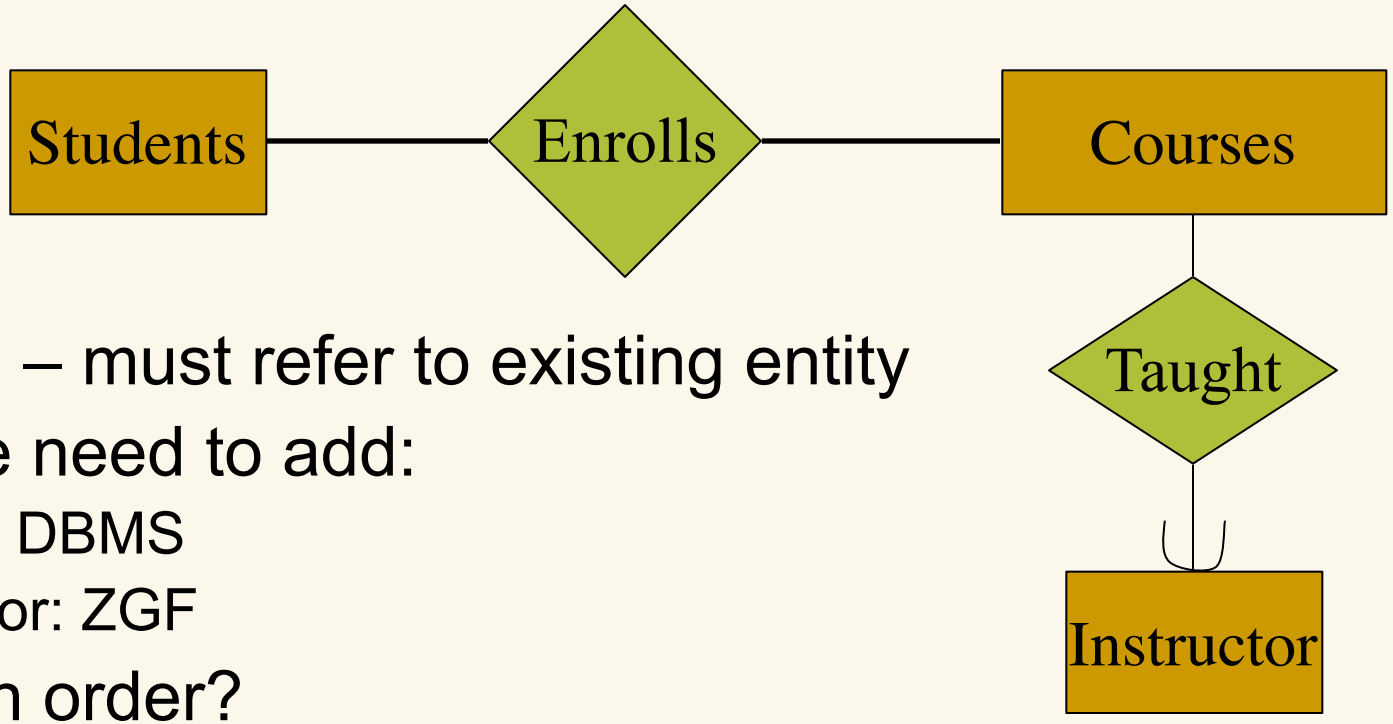
- Attributes have this automatically
  - could be null or one value
  - Can think of key atts as (non-null) single-valued

Referential integrity

- "Exactly one value"
  - NOT NULL & *foreign keys* in SQL
- Relationships
  - Non-null value refers to entity that exists
  - Refer to entity with foreign key
  - HTML analogy: no broken links
  - Programming analogy: no dangling pointers
  - Multiple ways of handling violations…



Referential integrity – E/R e.g.



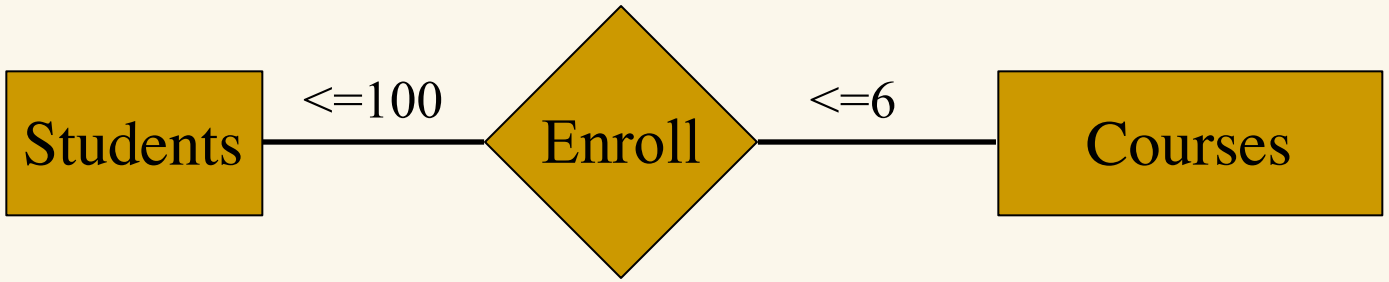
- Insertion – must refer to existing entity
- Suppose need to add:
  - course: DBMS
  - instructor: ZGF
- Q: Which order?

- Q: What if relship were exactly-exactly, say, M(Hs,Ws)?
  - i.e., referential integrity in both directions?
- A: Put both inserts in one xact – later

Other kinds of constraints

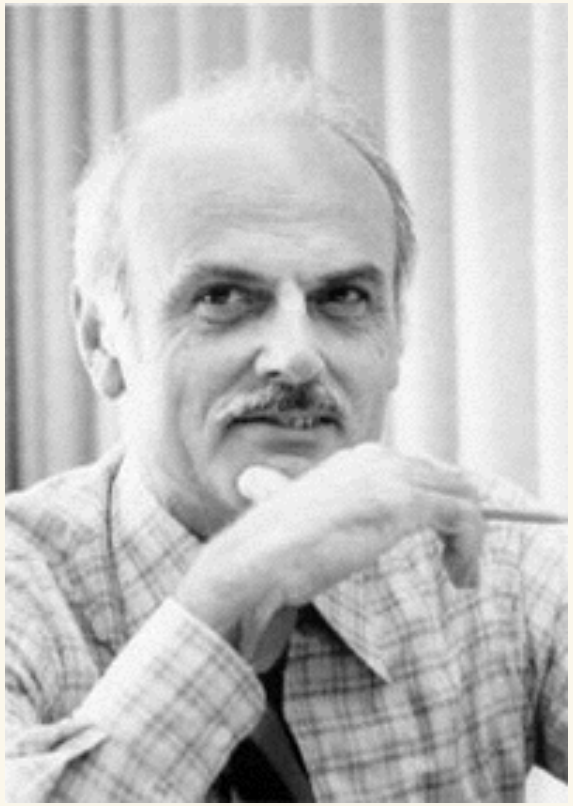
- Domain constraints
  - E.g. date: must be after 1980
  - Enumerated type: grades A through F, no E
  - No special E/R notation just write near line

- General constraints:
  - A class may have no more than 100 students; a student may not have more than 6 courses:

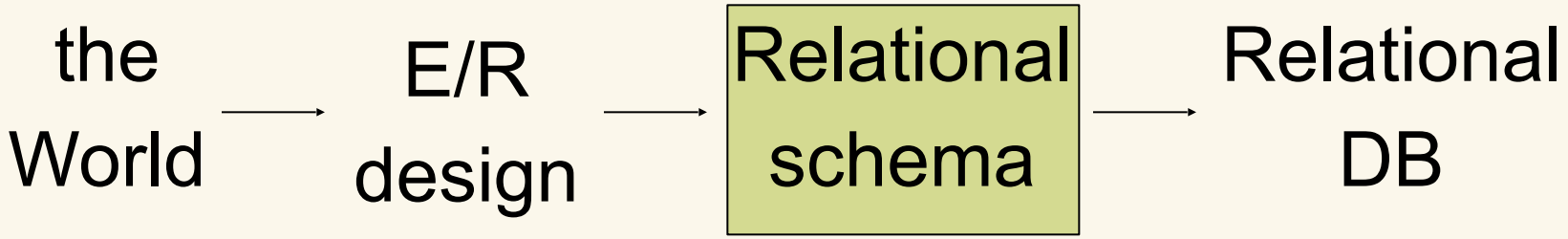


Next topic: the Relational Data Model

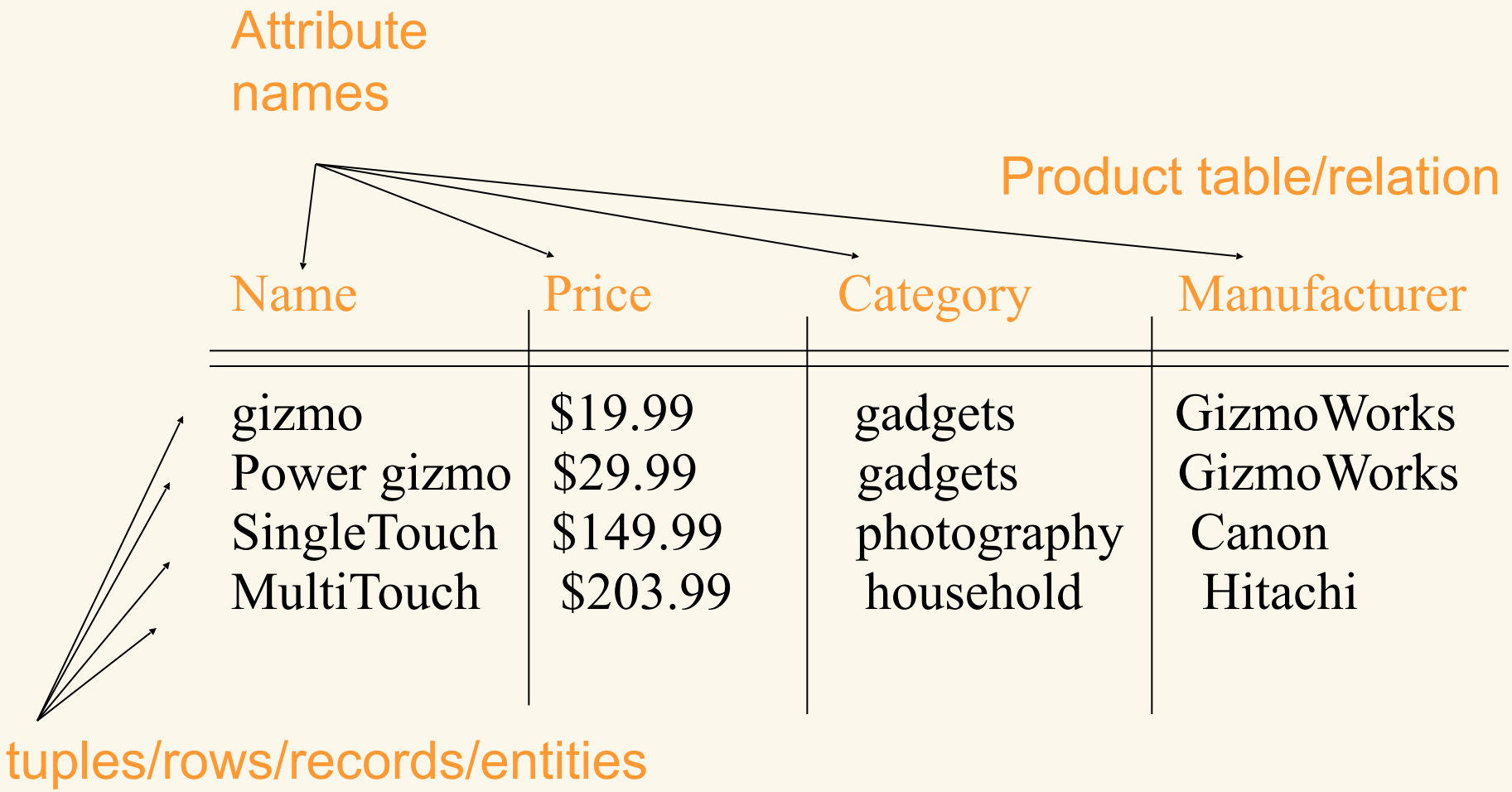
- Invented by Ted Codd
  - Researcher at IBM
  - We'll see his name again…
- Related work at Berkeley
- Introduced in a paper a paper published in June, 1970



DB development path



Relations as tables



Relational terminology

- Relation is composed of *tuples*
- Tuple = sequence of *attribute* values
  - Attribute has atomic types
- Relation *schema*:

  relation name + attribute names + attribute types
- Database schema: set of relation schemas

Relations as sets

- Recall: math relation is a subset of the cross-product of the attribute value sets
  - R subset-of S x T
  - Product subset-of Name x Price x Cat x Mft

- One member of Product relation:
  - (gizmo, $19.99, gadgets, GizmoWorks) in Product
  - Product(gizmo, $19.99, gadgets, GizmoWorks)

- Usual updates: add/delete/change a tuple in this set
- Updates to the *schema* are rare, painful (why?)

From E/R models to relations

- Recall justification:
  - design is easier in E/R
  - but implementation is easier/faster in R

- Analogy to program compilation:
  - design is easier in C/Java/whatever
  - implemen. is easier/faster in machine/byte code

- Strategy:
  1. apply semi-mechanical conversion rules
  2. improve by combining some relations
  3. improve by normalization
     - involves finding functional dependencies

E/R conversion rules

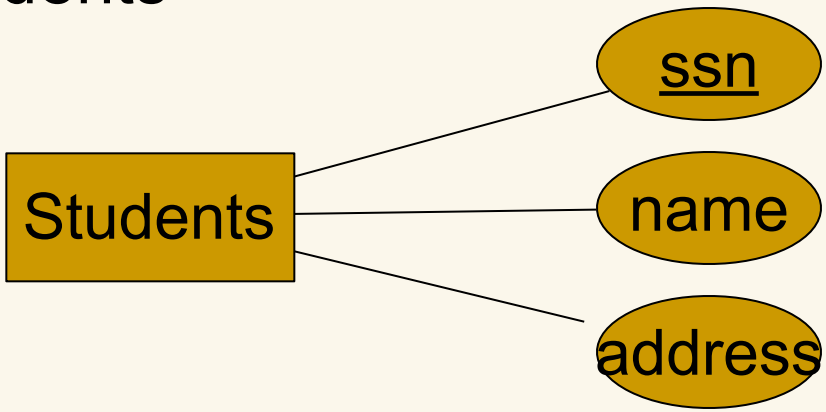
- Relationship → relation
  - attributes: keys of entity-sets/roles
  - key: depends on multiplicity
- Entity set → … relation
  - attributes: attributes of entity set
  - key: key of ES

- NB: mapping of types is not one-one
  - We'll see: mapping one tokens is also not one-one

- Special treatment:
  - Weak entity sets
  - Isa relations & subclasses

Entity Sets

- Entity set Students

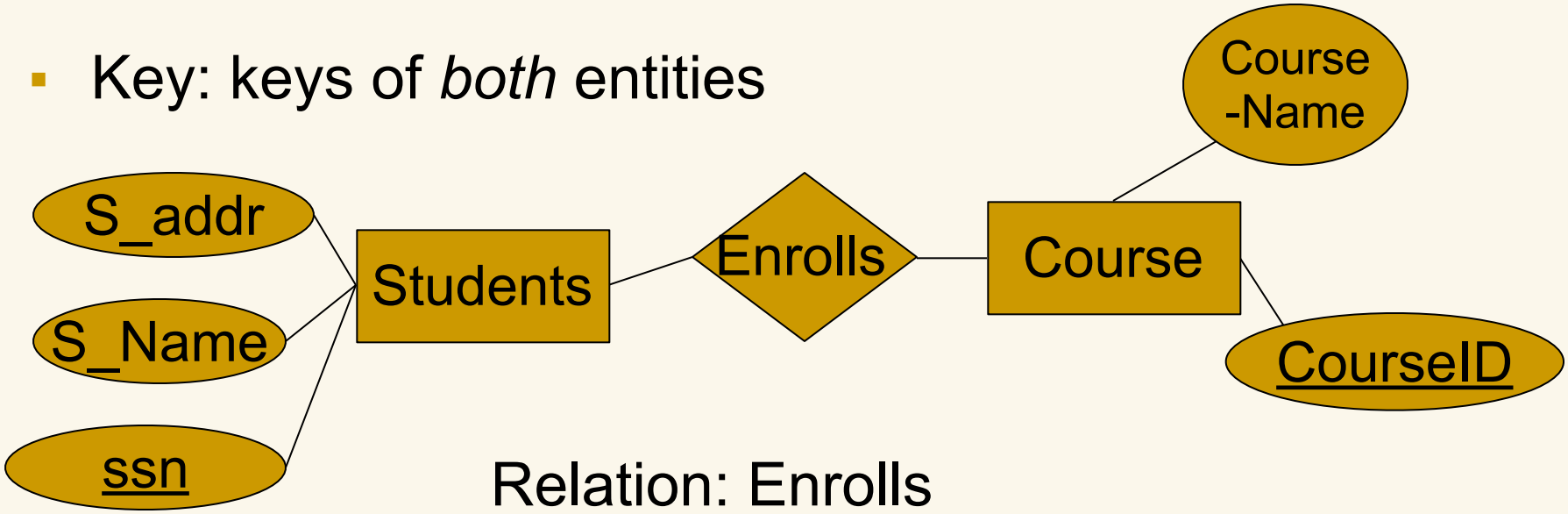


Rel: Students

SSN	Name	Address
111-222-3333	Howard	Park Avenue
444-555-6666	John	South Carolina

Binary many-to-many relationships

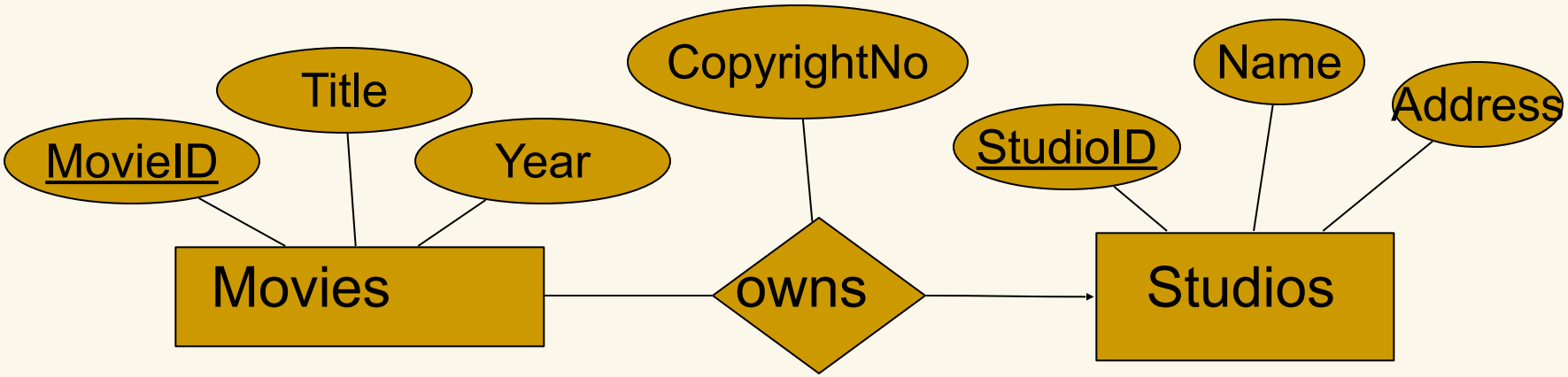
- Key: keys of *both* entities



Relation: Enrolls

ssn	CourseID
111-222-3333	SE-304
111-222-3333	C20.0056
444-555-6666	C30.0046

Many-to-one relationships



Movies

MovieID	Title	Year
M101	Mr. Ripley.	1999
M202	Sylia	2003

Studios

StudioID	Name	Address
S35	Miramax	NYC
S73	Disney	Orlando

- Key: keys of *many* entitiy

Owns

MovieID	StudioID	CopyrightNo
M101	S73	CN11111
M202	S35	CN22222

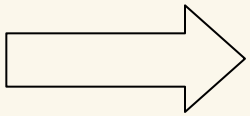
Many-to-one: a better design

Movies

MovieID	Title	Year
M101	Mr. Ripley.	1999
M202	Sylia	2003

Owns

MovieID	StudioID	CopyrightNo
M101	S73	CN11111
M202	S35	CN22222



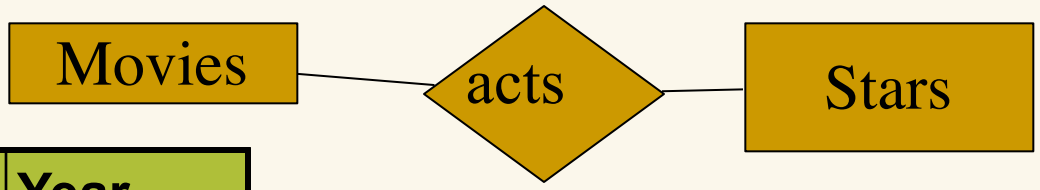
Movies'

MovieID	Title	Year	StudioID	CopyrightNo
M101	Talent Mr. Ripley	1999	S73	CN11111
M202	Sylia	2003	S35	CN22222

- Q: What if a movie's Owns row were missing?

Many-to-many relationships again

- NB: Won't work for many-many relationships



Movies

MovieID	Title	Year
M101	Mr. Ripley	1999
M202	Sylia	2003
M303	P.D. Love	2002

Stars

StarID	Name	Address
T400	Gwyneth P.	Bev.Hills
T401	P.S. Hoffman	Hollywood
T402	Jude Law	Palm Springs

Acts

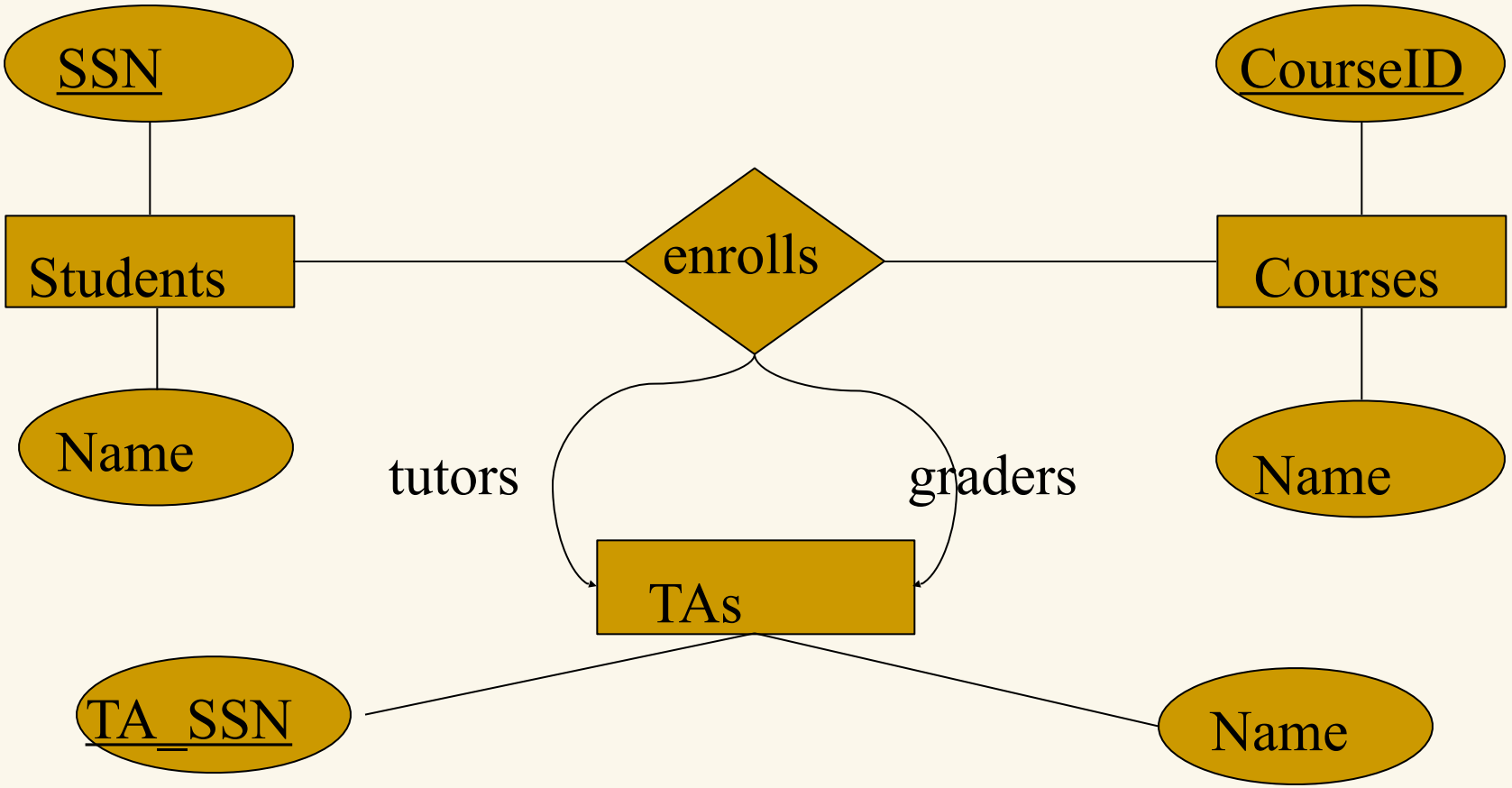
MovieID	StarID
M101	T400
M202	T400
M101	T401
M101	T402
M303	T401

Many-to-many relationships again

And here's why:

MovieID	Title	Year	StarID
M101	Talented Mr. Ripley	1999	T400
M101	Talented Mr. Ripley	1999	T401
M101	Talented Mr. Ripley	1999	T402
M202	Syvlia	2003	T400
M303	Punch Drunk Love	2003	T401

Multiway relationships & roles



- Different roles treated as different entity sets
- Key: keys of the *many* entities

Multiway relationships & roles

Students

SSN	Name
111-11-1111	George
222-22-2222	Dick

TAs

TA_SSN	Name
333-33-3333	Wesley
444-44-4444	Howard
555-55-5555	John

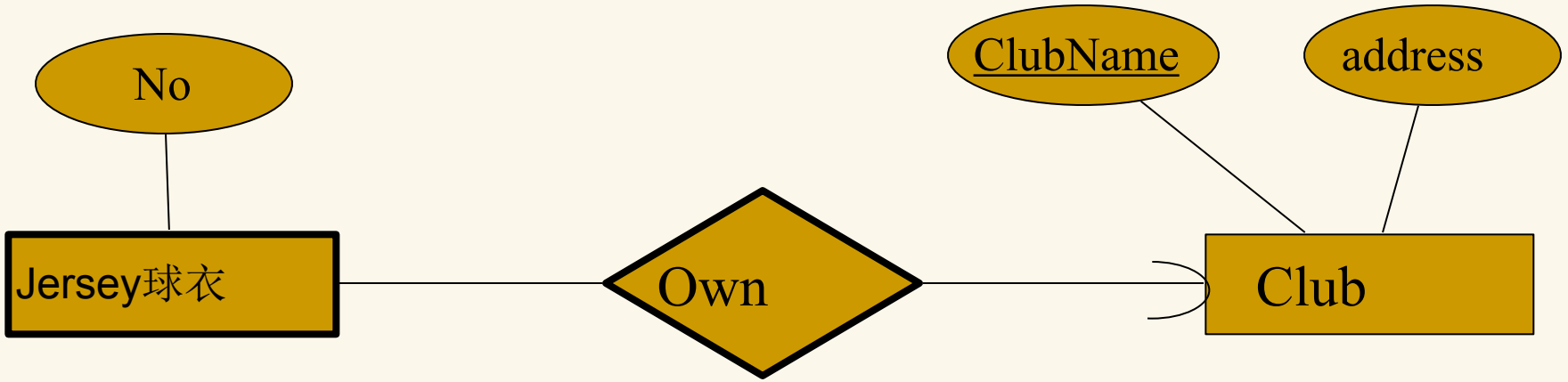
Courses

CourseID	Name
SE-304	Databases
C20.0056	Software

Enrolls(S_SSN, Course_ID, Tutor_SSN, Grader_SSN)

S_SSN	CourseID	Tutor_SSN	Grader_SSN
111-11-1111	SE-304	333-33-3333	444-44-4444
222-22-2222	SE-304	444-44-4444	555-55-5555

Converting weak ESs – differences



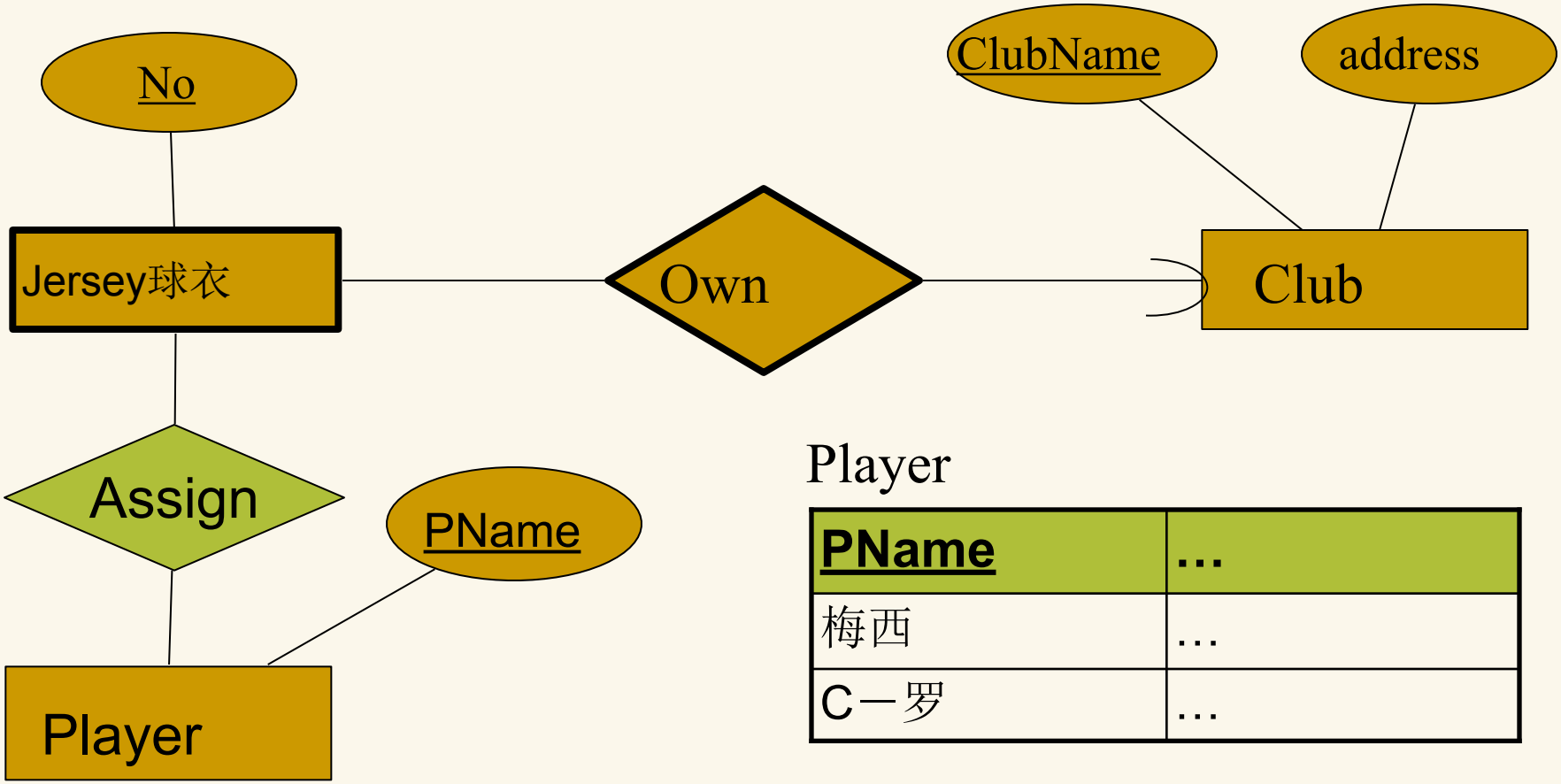
- Atts of Jersey Rel. are:
  - attributes of Jersey
  - key attributes of supporting ESs, Club

Jersey

ClubName	JerseyNo
巴萨	10
巴萨	9
皇马	10

- Supporting relships are omitted (*why?*)

Weak entity sets - relationships

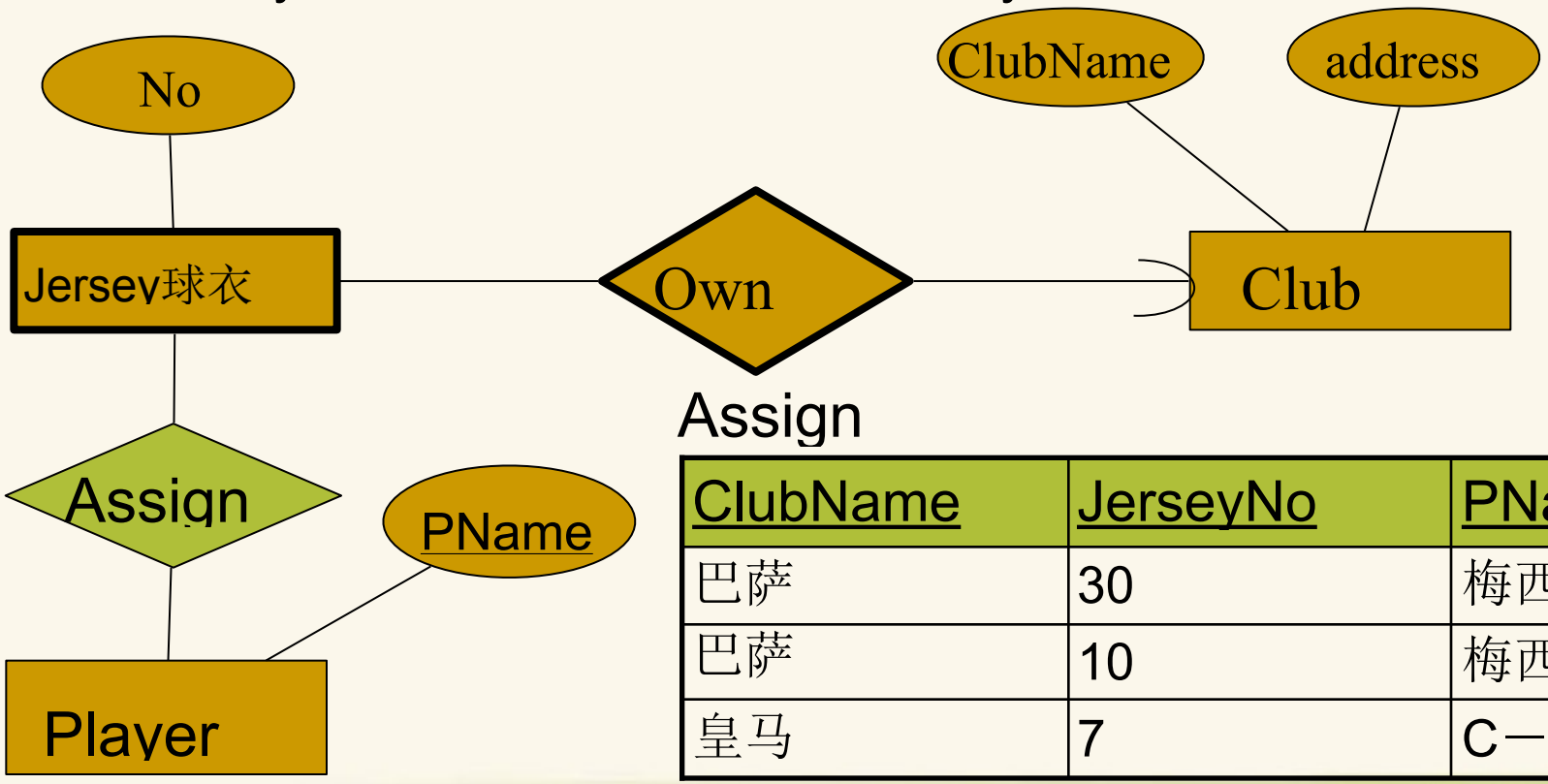


Player

PName	…
梅西	…
C一罗	…

Weak entity sets - relationships

- Non-supporting relationships for weak ESs *are* converted
  - keys include entire weak ES key



Assign

ClubName	JerseyNo	PName
巴萨	30	梅西
巴萨	10	梅西
皇马	7	C一罗

Next week

- For next week:
  - Review/skim Ch.3 section 5 (from today)
  - Read Ch.19 sections 1-3