CMU SCS

Carnegie Mellon Univ.
Dept. of Computer Science
15-415/615 - DB Applications

C. Faloutsos - A. Pavlo

Lecture#2: E-R diagrams

CMU SCS

Administrivia

- Course url:
  - http://15415.courses.cs.cmu.edu/
- Course policies
  - http://15415.courses.cs.cmu.edu/fall2016/policies.html
- Foils in pps:
  - http://15415.courses.cs.cmu.edu/fall2016/slides-pps/

CMU SCS

Course Topics

✔Introduction to DBMSs

- Data Models
- Query Language (SQL)
- Database Design
- Query Optimization & Indexing
- Transaction Management
- Advanced Topics

CMU SCS

Problem

- Develop an application for U.G. admin:
  - Student info
  - Who-takes-what class
  - Class rosters
  - Transcripts

- How do you proceed?
  - (Which role(s) are you playing?)

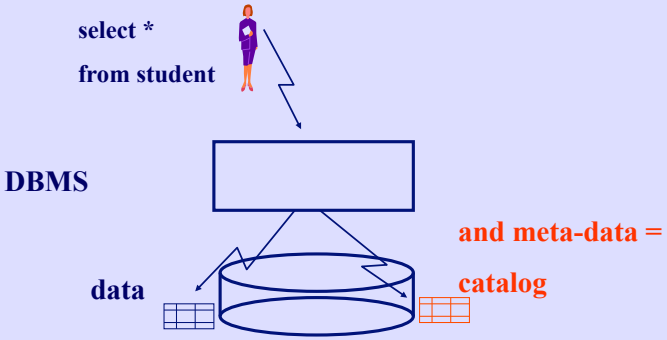
CMU SCS

Database users

- 'naive' users
- casual users
- application programmers
- [DBA (Data base administrator)]

CMU SCS

Casual users



CMU SCS

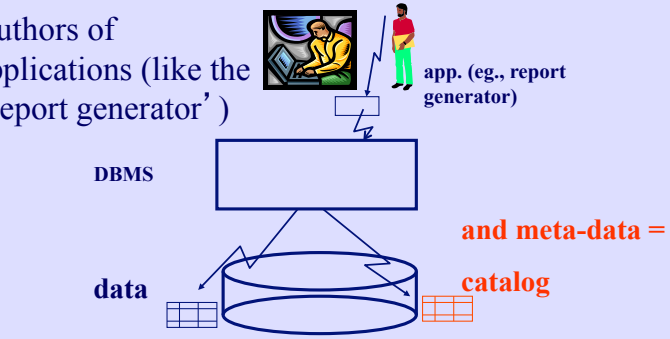
``Naive'' users

Pictorially:

CMU SCS

App. programmers

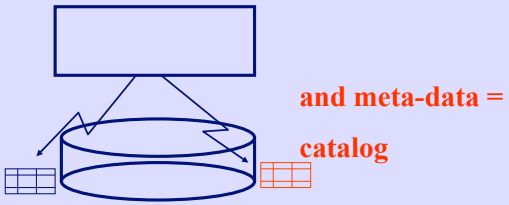
- Authors of applications (like the 'report generator')



CMU SCS

DB Administrator (DBA)

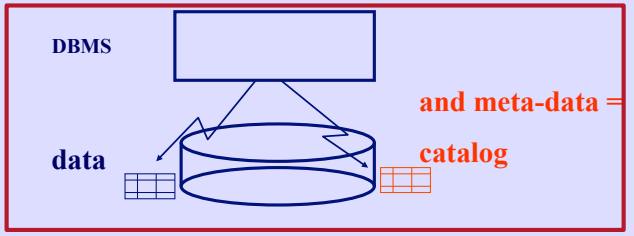
- Duties?



CMU SCS

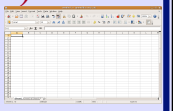
DB Administrator (DBA)

- Duties?



CMU SCS

DB Administrator (DBA)



- schema definition ('logical' level)
- physical schema (storage structure, access methods
- schema modifications
- granting authorizations
- integrity constraint specification

CMU SCS

Problem

- Develop an application for U.G. admin:
  - Student info
  - Who-takes-what class
  - Class rosters
  - Transcripts
- How do you proceed?
  - (Which role(s) are you playing?)

CMU SCS

Database users

- 'naive' users
- casual users
- application programmers
- [DBA (Data base administrator)]

CMU SCS

Database Design

- Requirements Analysis
- Conceptual Design
- Logical Design
- Schema Refinement
- Physical Design
- Security Design

CMU SCS

Database Design

✔Requirements Analysis	user's needs
• **Conceptual Design**	high level (ER)
• Logical Design	Tables
• Schema Refinement	Normalization
• Physical Design	Indices etc
• Security Design	Access controls

CMU SCS

Problem'

Maintain

- Develop an application for U.G. admin:
  - Student info
  - Who-takes-what class
  - Class rosters
  - Transcripts

- If you are the *new* DBA, what would you rather inherit:

CMU SCS

This or this ?

```
drop table if exists student;
create table student
   (ssn fixed,
   name char(20));
drop table if exists takes;
create table takes
    (ssn fixed,
    cid char(10),
   grade fixed);
```

CMU SCS

True story

- Health insurance company
- Wants to catch (some of the abundant) fraud
- Schema:
  - patients, visit doctors, get medicine,
  - Doctors perform operations, …
  - Nurses monitor patients, …
  - etc etc
- Q: How many tables do you think it spans?

CMU SCS

True story

- Schema:
  - patients, visit doctors, get medicine,
  - Doctors perform operations, …
  - Nurses monitor patients, …
  - etc etc
- Q: How many tables do you think it spans? 10? 20? 30?

CMU SCS

True story

- Schema:
  - patients, visit doctors, get medicine,
  - Doctors perform operations, …
  - Nurses monitor patients, …
  - etc etc
- Q: How many tables do you think it spans? 10? [illegible] 30?
- A: **120 PAGES** of schema

CMU SCS

Motivation & upcoming conclusion:

- E-R diagrams are excellent documentation tools


```
drop table if exists student;
create table student
  (ssn fixed,
  name char(20));
drop table if exists takes;
create table takes
  (ssn fixed,
  cid char(10),
  grade fixed);
```


CMU SCS

Overview

- concepts
  - Entities
  - Relationships
  - Attributes
  - Specialization/Generalization
  - Aggregation
  - ER modeling questions

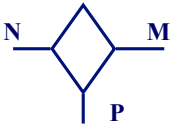
Basic

Advanced/ rare

CMU SCS

Tools

Entities ('entity sets')



Relationships ('rel. sets')

and mapping constraints

attributes

CMU SCS

Example

Students, taking courses, offered by instructors; a course may have multiple sections; one instructor per section

nouns -> entity sets

verbs -> relationship sets

CMU SCS

Example

Students, taking courses, offered by instructors; a course may have multiple sections; one instructor per section

nouns -> entity sets

verbs -> relationship sets

CMU SCS

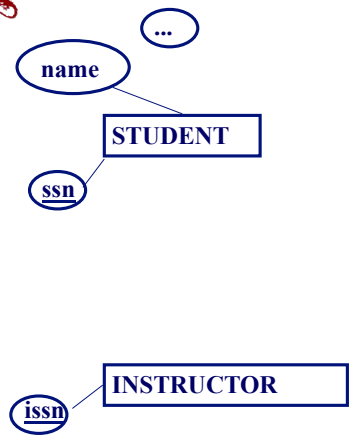
Example

Students, taking courses, offered by instructors; a course may have multiple sections; one instructor per section

nouns -> entity sets

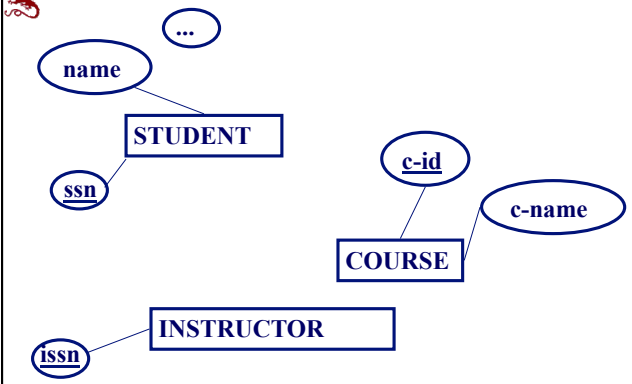
verbs -> relationship sets

CMU SCS

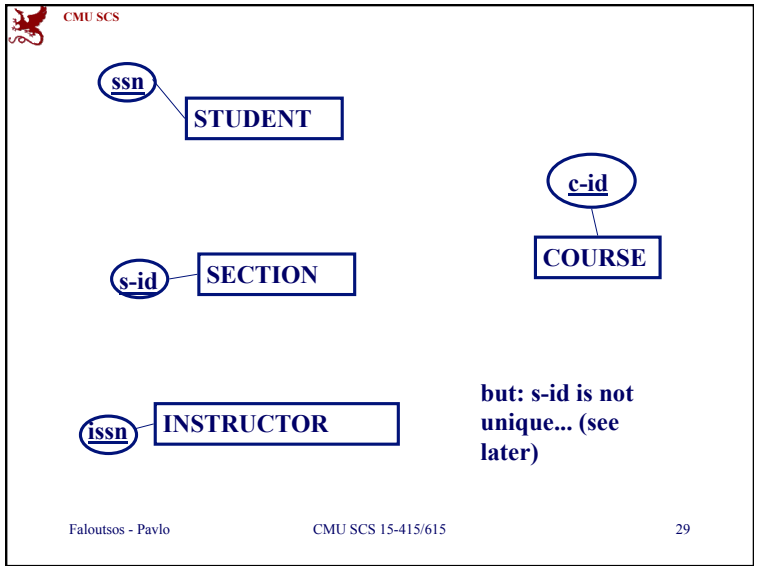


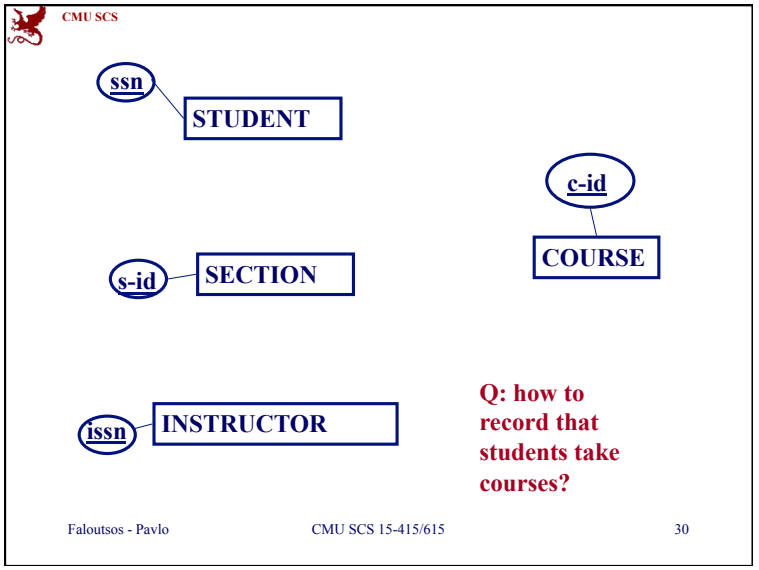
primary key = unique identifier -> underline

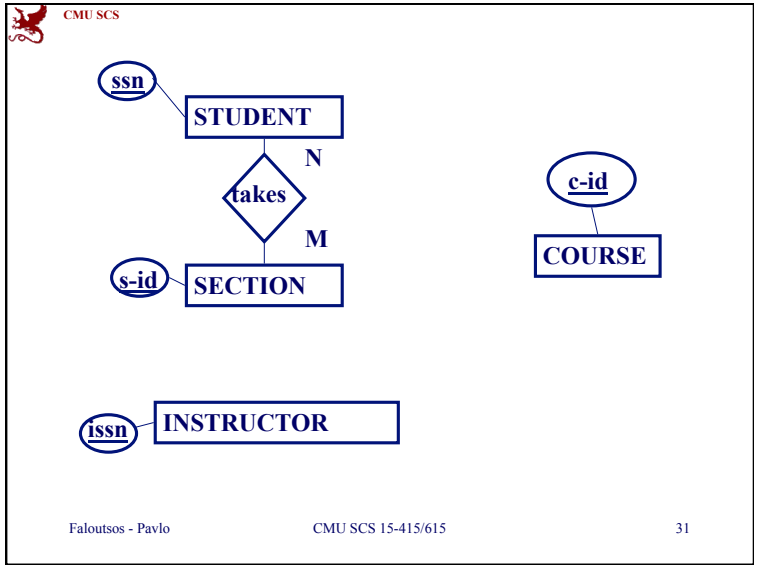
CMU SCS

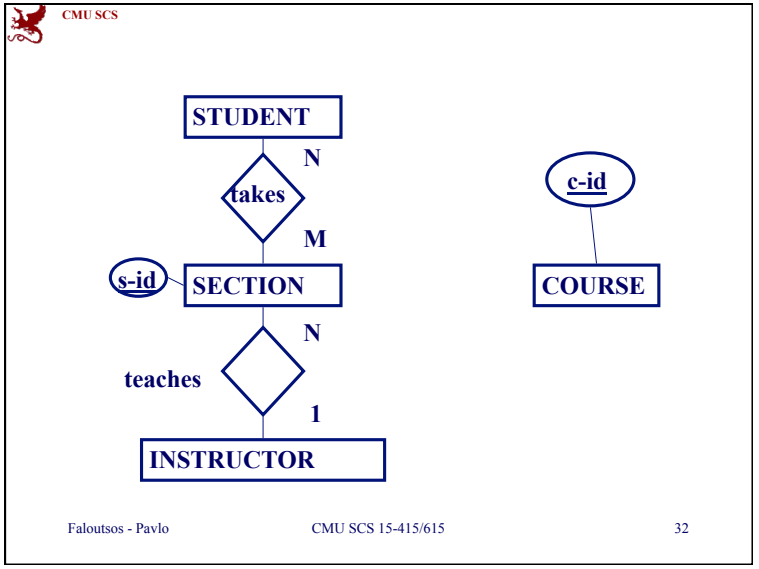


but: sections of course (with different instructors)?

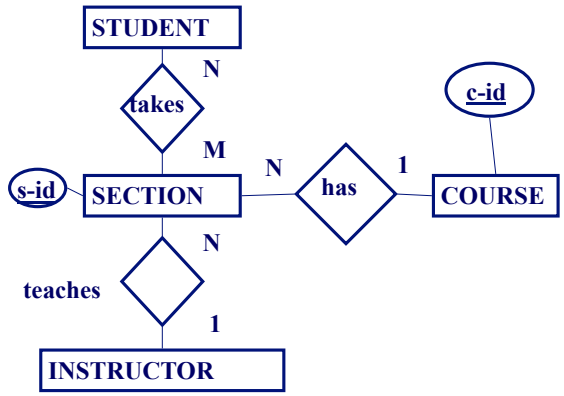

CMU SCS
ssn
STUDENT
c-id
COURSE
s-id
SECTION
issn
INSTRUCTOR
but: s-id is not unique... (see later)
Faloutsos - Pavlo CMU SCS 15-415/615 29


CMU SCS
ssn
STUDENT
c-id
COURSE
s-id
SECTION
issn
INSTRUCTOR
Q: how to record that students take courses?
Faloutsos - Pavlo CMU SCS 15-415/615 30


CMU SCS
ssn
STUDENT
N
takes
M
c-id
COURSE
s-id
SECTION
issn
INSTRUCTOR
Faloutsos - Pavlo CMU SCS 15-415/615 31


CMU SCS
STUDENT
N
takes
M
c-id
COURSE
s-id
SECTION
N
teaches
1
INSTRUCTOR
Faloutsos - Pavlo CMU SCS 15-415/615 32

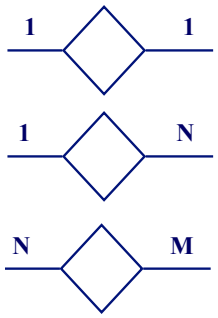
CMU SCS



CMU SCS

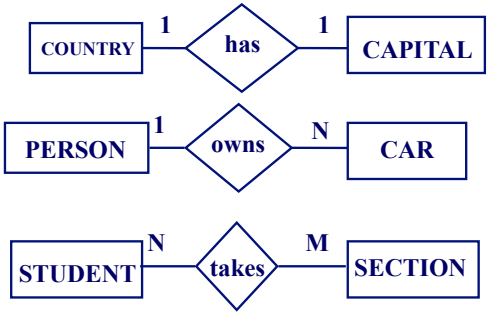
Cardinalities

- 1 to 1 (example?)
- 1 to N
- N to M



CMU SCS

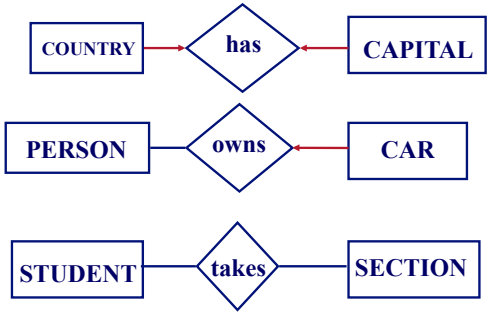
Cardinalities



CMU SCS

Cardinalities

Book's notation:



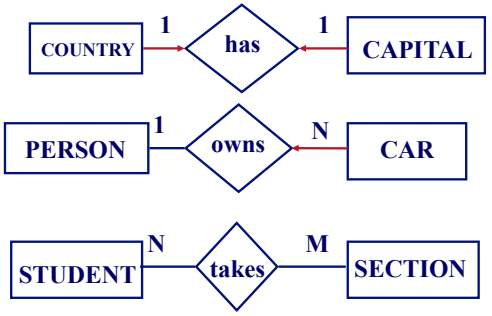
CMU SCS

Cardinalities

Book's notation

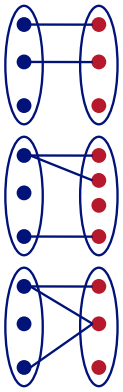
vs

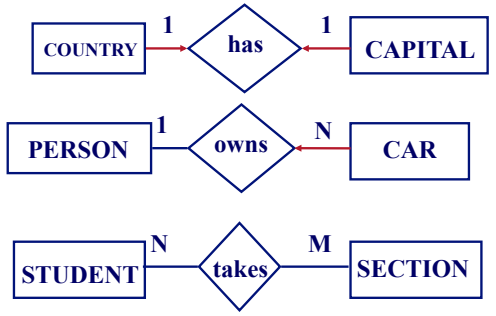
***1 to N* notation**



CMU SCS

Cardinalities





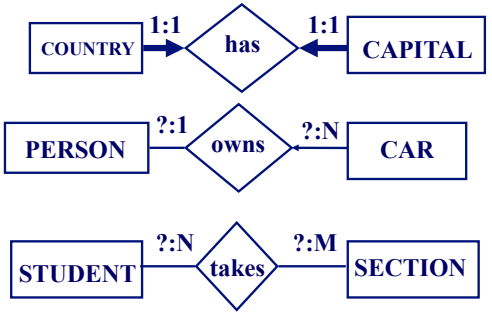
CMU SCS

'Total/partial' participation

total, total

??

??



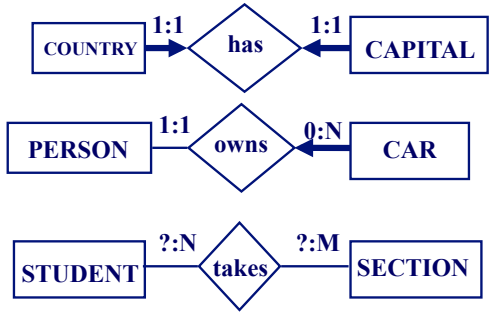
CMU SCS

'Total/partial' participation

total, total

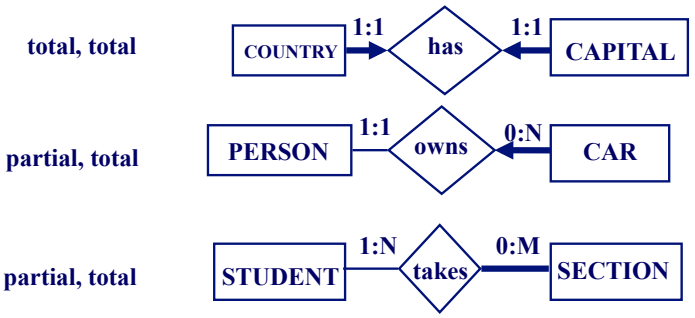
partial, total

??



CMU SCS

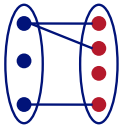
'Total/partial' participation



CMU SCS

'Total/partial' participation

Is it 'legal'?



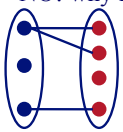
partial, total

CMU SCS

'Total/partial' participation

Is it 'legal'?

NO! why not?

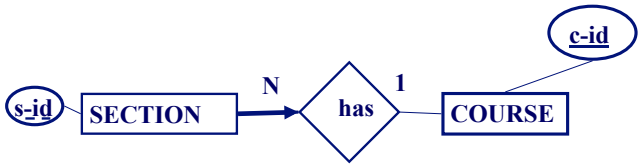


partial, total

CMU SCS

Subtle concept: Weak entities

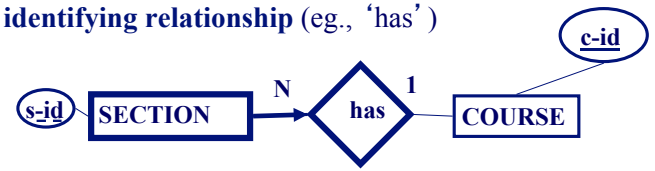
- 'section' has no unique-id of its own!(?)



CMU SCS

Weak entities

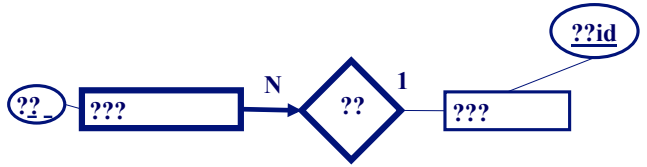
- 'weak' entities: if they need to borrow a unique id from a 'strong entity - **thick** box.
- 'c-id' + 's-id': unique id for SECTION
- **partial key** (eg., 's-id') - dashed-underline
- **identifying relationship** (eg., 'has')



CMU SCS

Weak entities

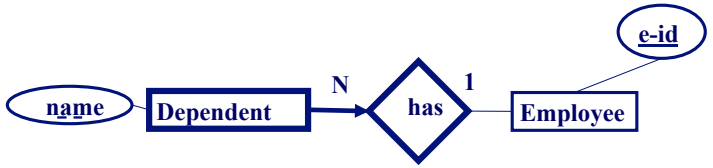
- Other example(s) of weak entities?



CMU SCS

Weak entities

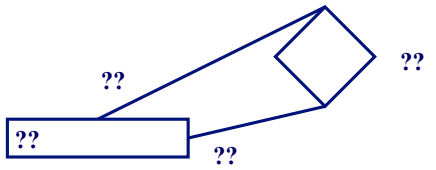
- Other example(s) of weak entities?



CMU SCS

More details

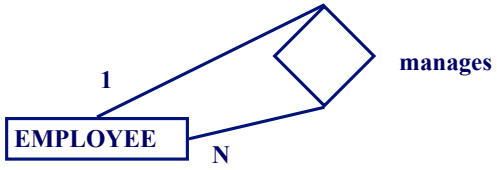
- self-relationships - example?



CMU SCS

More details

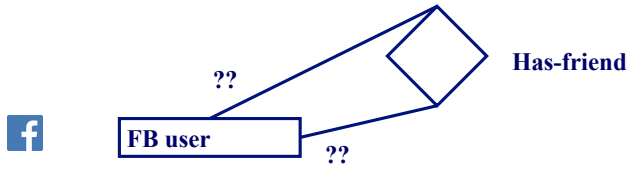
- self-relationships - example?



CMU SCS

More details

- self-relationships - example?



CMU SCS

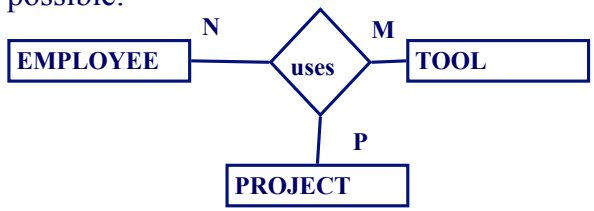
More details

- 3-way and k-way relationships?

CMU SCS

More details

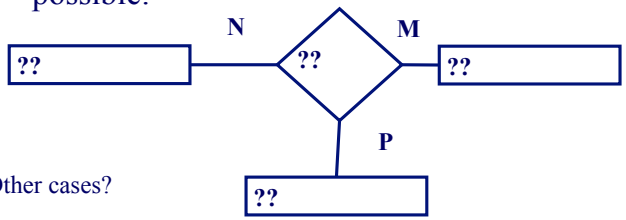
- 3-way and k-way relationships? Rare, but possible:



CMU SCS

More details

- 3-way and k-way relationships? Rare, but possible:

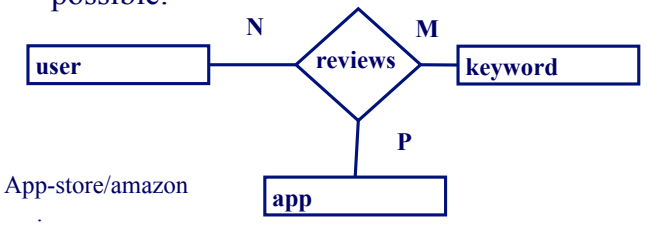


Other cases?

CMU SCS

More details

- 3-way and k-way relationships? Rare, but possible:



App-store/amazon reviews

CMU SCS

Overview

- concepts
  - Entities
  - Relationships
  - Attributes
  - Specialization/Generalization
  - Aggregation
  - ER modeling questions

CMU SCS

More details - attributes

- **key** (or **primary key**): unique identifier
- underlined, in the ER diagram
- [not in textbook - FYI:
  - **multivalued** or set-valued attributes (eg., 'dependents' for EMPLOYEE)
  - **derived** attributes (eg., 15% tip)

  ]

CMU SCS

Overview

- concepts
  - Entities
  - Relationships
  - Attributes
  - Specialization/Generalization
  - Aggregation
  - ER modeling questions

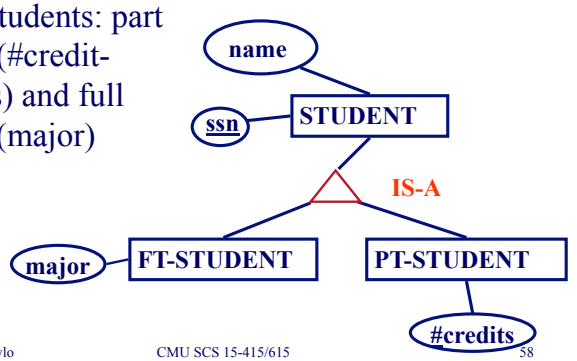
Basic

Advanced/ rare

CMU SCS

Specialization

- eg., students: part time (#credit-hours) and full time (major)



CMU SCS

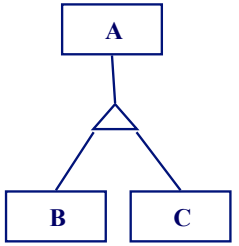
Observations

- Generalization: exact reverse of 'specialization'
- attribute inheritance
- could have **many** levels of an IS-A hierarchy

CMU SCS

More details

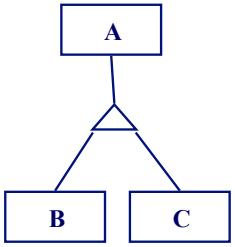
- Overlap constraints
- Covering constraints



CMU SCS

More details

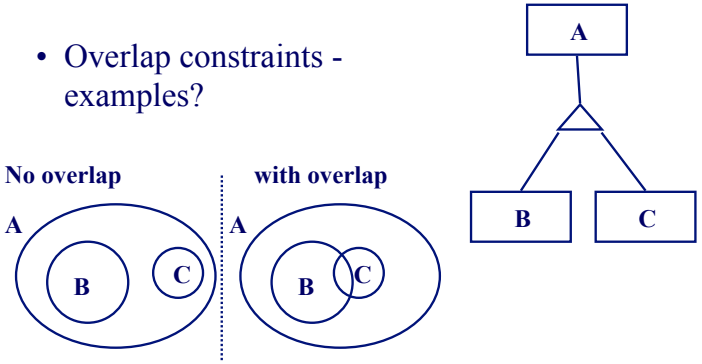
- Overlap constraints
  - can an entity belong to both 'B' and 'C'?
- Covering constraints
  - can an 'A' entity belong to neither 'B' nor 'C'?



CMU SCS

More details

- Overlap constraints - examples?



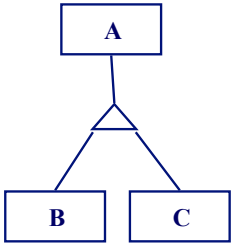
CMU SCS

More details

- Covering constraints - examples?

Total coverage

Partial coverage



CMU SCS

Overview

- concepts
  - Entities
  - Relationships
  - Attributes
  - Specialization/Generalization
  - Aggregation
  - ER modeling questions

CMU SCS

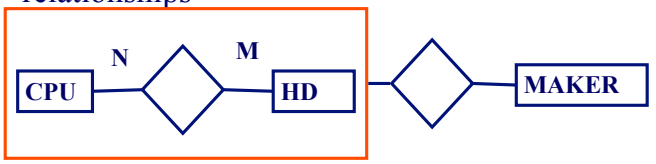
Aggregation

- computer model (w/ CPU and HD)
- and Maker (eg., Dell, HP)

CMU SCS

Aggregation

- treat a relationship as an entity
- used to express a relationship among relationships



CMU SCS

Overview

- concepts
  - Entities
  - Relationships
  - Attributes
  - Specialization/Generalization
  - Aggregation
  - ER modeling questions

CMU SCS

Conceptual design

- Entity vs attribute
- Entity vs relationship
- Binary or ternary relationships?
- Aggregation?

CMU SCS

Entity vs. attribute

- Entity EMPLOYEE (w/ emp#, name, job_code, ...)
- Q: How about 'spouse' - entity or attribute?
- Q: How about 'dependents'?

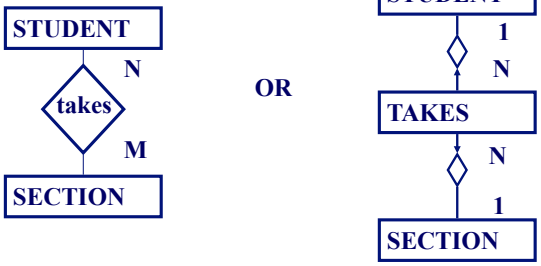
CMU SCS

Entity vs. attribute

- Entity EMPLOYEE (w/ emp#, name, job_code, ...)
- Q: How about 'spouse' - entity or attribute?
- A: probably, 'attribute' is enough
- Q: How about 'dependents'?
- A: Entity - we may have many dependents

CMU SCS

Entity vs. Relationship



CMU SCS

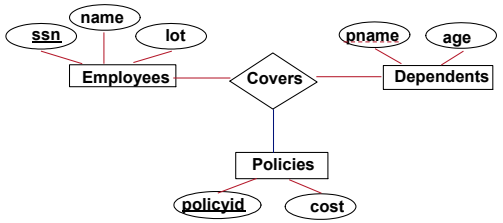
Binary vs Ternary Relationships

- usually, binary relationships are 'cleaner':

CMU SCS

Binary vs. Ternary Relationships

If each policy is owned by just 1 employee:

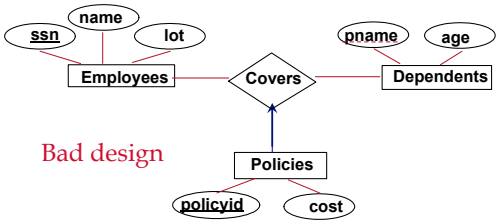


CMU SCS

Binary vs. Ternary Relationships

If each policy is owned by just 1 employee:

Bad design

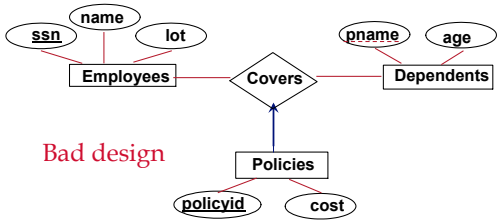


CMU SCS

Binary vs. Ternary Relationships

If each policy is owned by just 1 employee:

Bad design



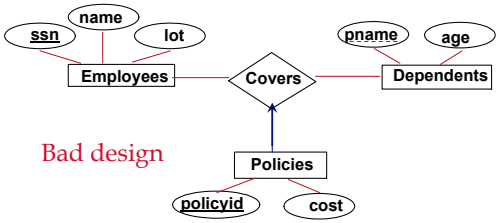
Key constraint on Policies would mean policy can only cover 1 dependent!

CMU SCS

Binary vs. Ternary Relationships

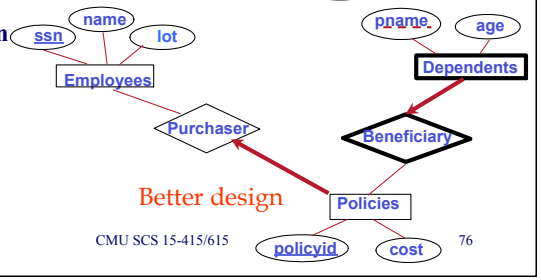
If each policy is owned by just 1 employee:

Bad design



Key constraint on Policies would mean policy can only cover 1 dependent!

What are the additional constraints in the 2nd diagram?



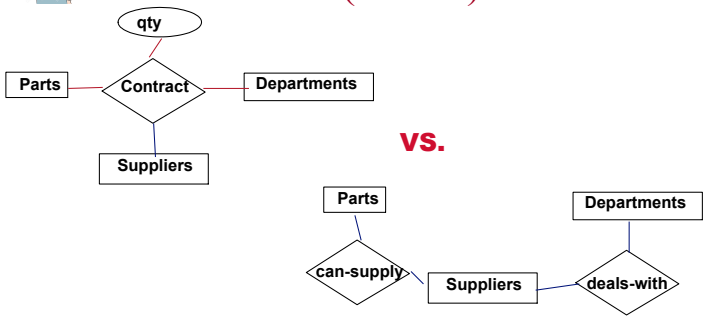
CMU SCS

Binary vs Ternary Rel.

- But sometimes ternary rel. can not be replaced by a set of binary rel's:

CMU SCS

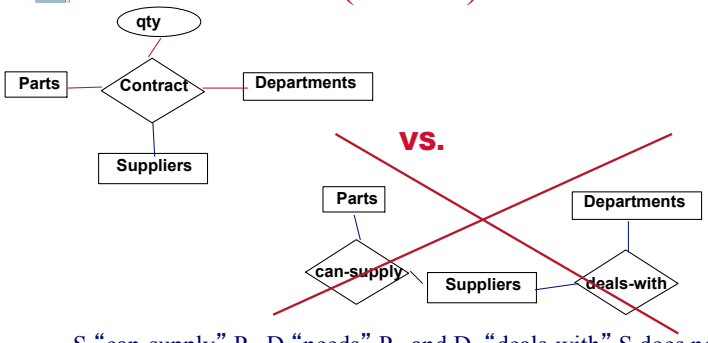
Binary vs. Ternary Relationships (Contd.)



why is it bad?

CMU SCS

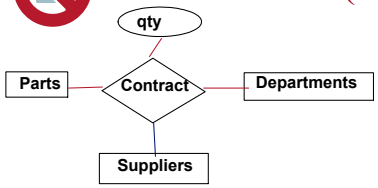
Binary vs. Ternary Relationships (Contd.)



- S "can-supply" P, D "needs" P, and D "deals-with" S does not imply that D has agreed to buy P from S.
- How do we record *qty*?

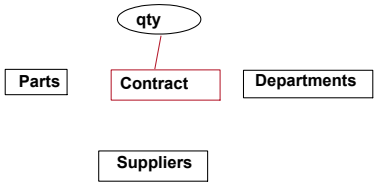
CMU SCS

Binary vs. Ternary Relationships (Contd.)



Not in textbook:

in practice, often:

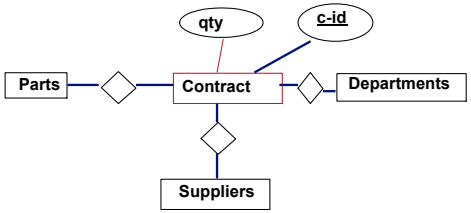


CMU SCS

Binary vs. Ternary Relationships (Contd.)

Not in textbook:

in practice, often:

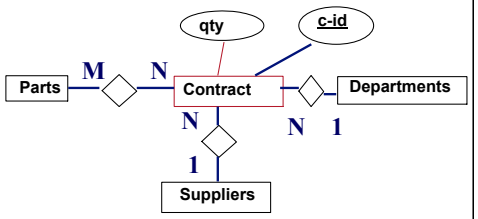


CMU SCS

Binary vs. Ternary Relationships (Contd.)

Not in textbook:

in practice, often:



CMU SCS

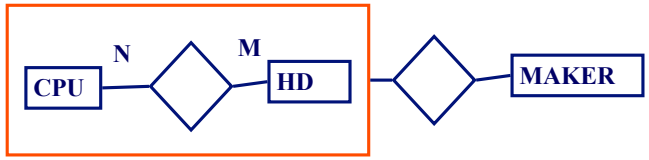
Ternary vs. aggregation

- use aggregation, if we want to attach a relationship to a relationship
- (see book for example)
- (in practice, again we create a unique-id and resort to binary relationships)

CMU SCS

Ternary vs. aggregation

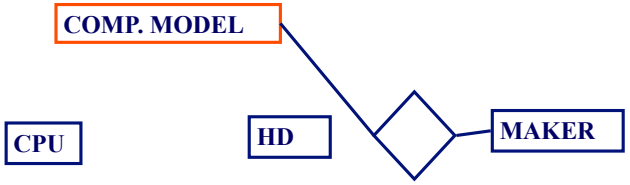
- How would you handle this case?



CMU SCS

Ternary vs. aggregation

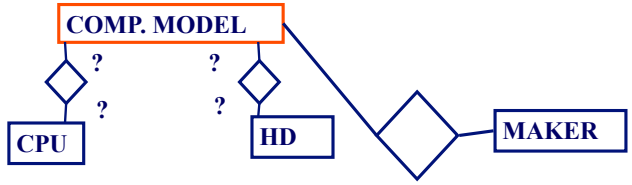
- How would you handle this case?



CMU SCS

Ternary vs. aggregation

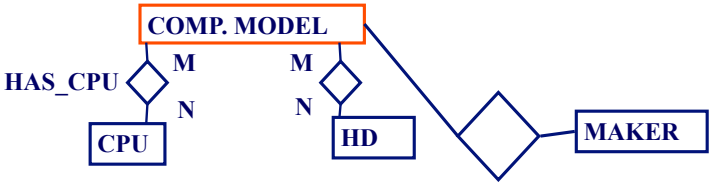
- How would you handle this case?



CMU SCS

Ternary vs. aggregation

- How would you handle this case?



CMU SCS

Summary

- E-R Diagrams: a powerful, user-friendly tool for data modeling:
  - Entities (strong, weak)
  - Attributes (primary keys, discriminators, derived, multivalued)
  - Relationships (1:1, 1:N, N:M; multi-way)
  - Generalization/Specialization; Aggregation

CMU SCS

Summary

- E-R Diagrams: a powerful, user-friendly tool for data modeling:
  - POPULAR
    - Entities (strong, weak)
    - Attributes (primary keys, discriminators, derived, multivalued)
    - Relationships (1:1, 1:N, N:M; multi-way)
    - Generalization/Specialization; Aggregation

CMU SCS

Summary - cont'd


(strong) entity set


weak entity set


relationship set


identifying rel. set for weak entity

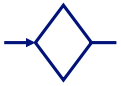

attribute

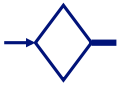
primary key

partial key

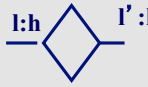
CMU SCS

Summary - cont'd


cardinalities


partial/total

cardinalities



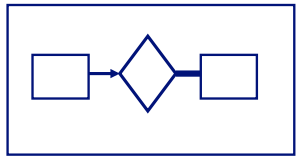
cardinalities with limits

(not in textbook - FYI)

CMU SCS

Summary - cont'd


IS-A


aggregation