MONASH University

MONASH
INFORMATION
TECHNOLOGY

Week 3 - Database Design I: Conceptual Modelling

FIT2094 - FIT3171 Databases

Clayton Campus S1 2019.

GROUP OF EIGHT AUSTRALIA

Overview

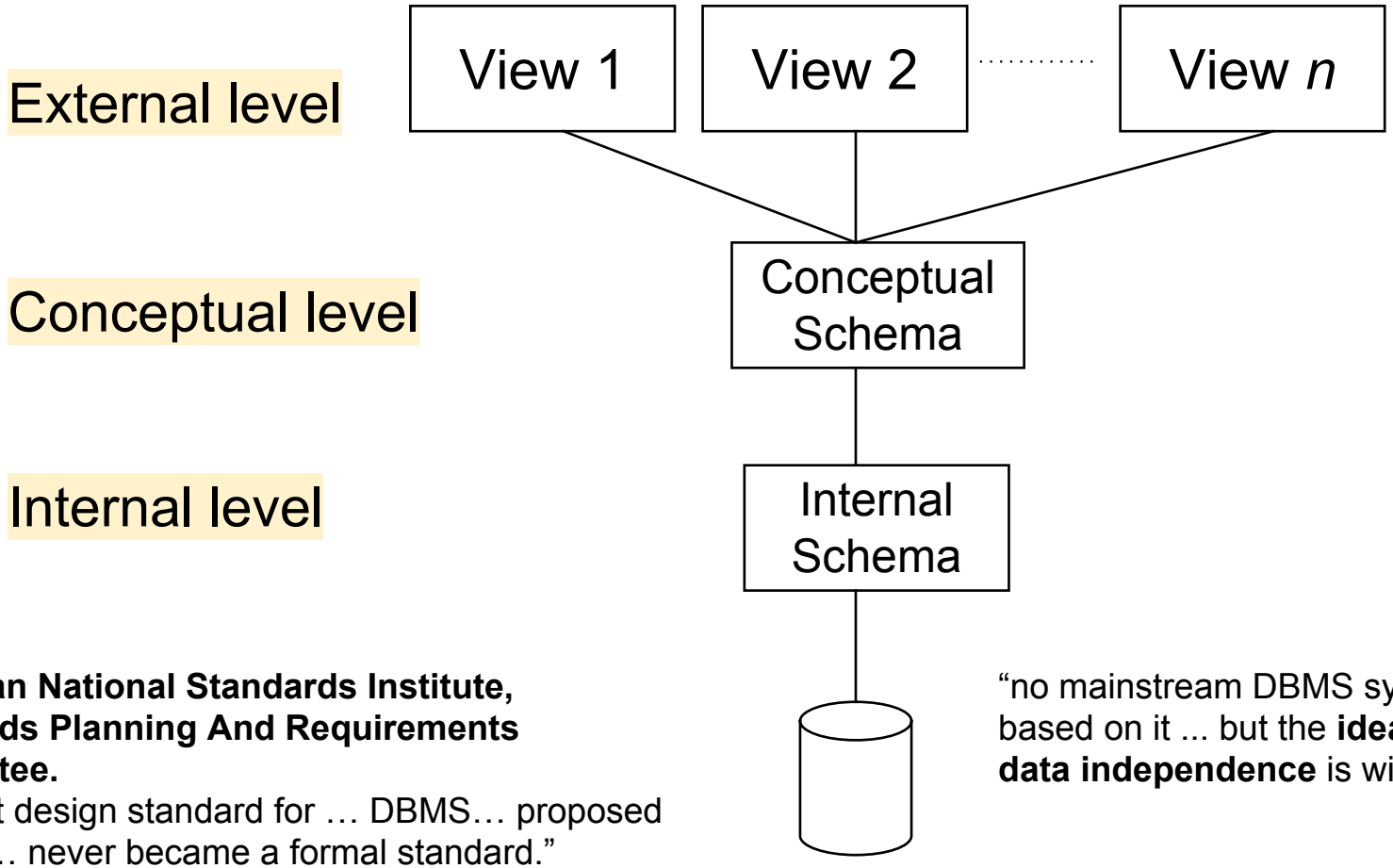
- **Hour 1 - Theory and Basics**
  - –ANSI/SPARC
  - –DBDLC: Conceptual Design
    - •Link → Conceptual level of ANSI/SPARC.
    - •Independent of all physical implementation considerations.
  - –DBDLC: Logical Design
    - •Target: particular database model (e.g. RDB).
    - •Independent of any implementation
    - •Link → Conceptual Level in ANSI/SPARC
  - –DBDLC: Physical Design
    - •Dependent on the particular DBMS environment in use.
    - •Link → Internal Level in ANSI/SPARC

… then COFFEE BREAK!

Overview

- **Hour 2 - Application**
  - Entity, Attributes and Relationships
  - More practical aspect
  - Refer to supplied case study
  - Lots of FLUX and Q&A!

ANSI/SPARC architecture

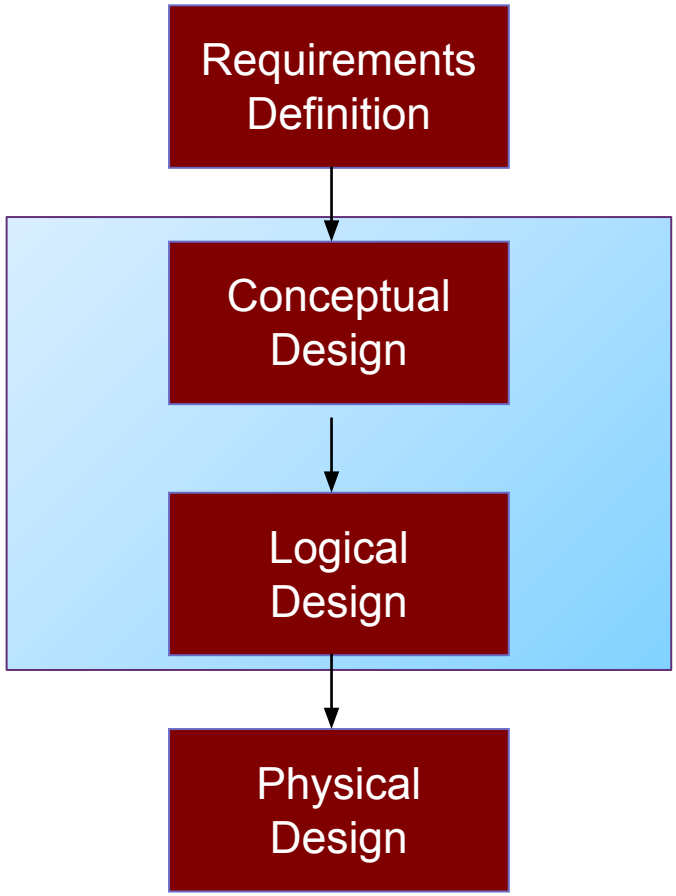


American National Standards Institute, Standards Planning And Requirements Committee.
"abstract design standard for … DBMS… proposed in 1975… never became a formal standard."

"no mainstream DBMS systems are fully based on it ... but the **idea of logical data independence** is widely adopted"

Src: Wikipedia.

The Database Design Life Cycle [DBDLC]



← ANSI/SPARC 'External'

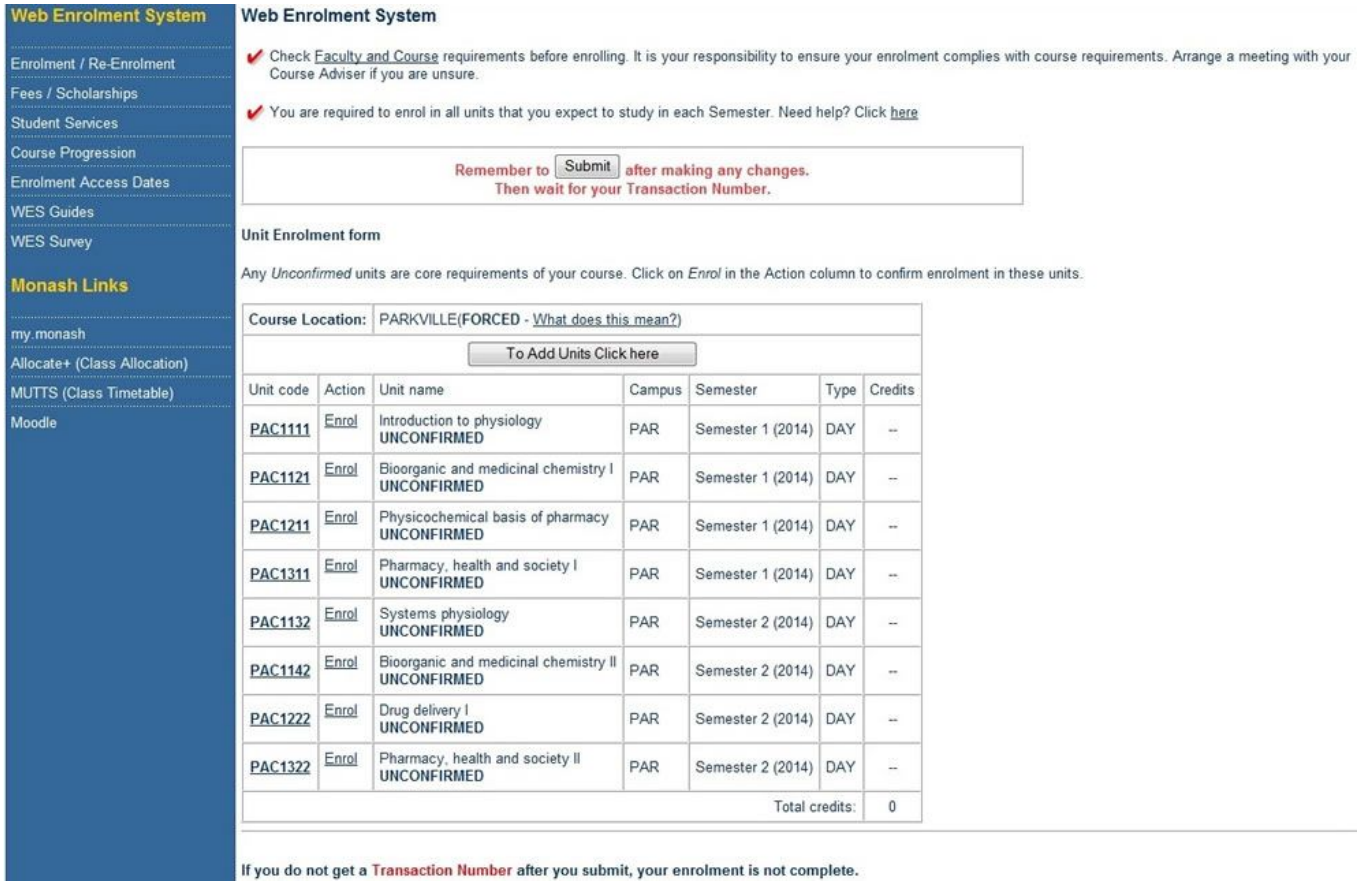
← ANSI/SPARC 'Conceptual'

← ANSI/SPARC 'Internal'

DBDLC: Requirements Definition

- Identify and analyse **user views**.
- A 'user view' may be a **report** to be produced or a particular type of **transaction** that should be supported.
- Corresponds to the **external** level of the ANSI/SPARC architecture.
- Output is a statement of **specifications** which describes the user views' particular **requirements and constraints.**

Case study: Monash - Student view



Web Enrolment System

- Enrolment / Re-Enrolment
- Fees / Scholarships
- Student Services
- Course Progression
- Enrolment Access Dates
- WES Guides
- WES Survey

Monash Links

- my.monash
- Allocate+ (Class Allocation)
- MUTTS (Class Timetable)
- Moodle

Web Enrolment System

✔ Check Faculty and Course requirements before enrolling. It is your responsibility to ensure your enrolment complies with course requirements. Arrange a meeting with your Course Adviser if you are unsure.

✔ You are required to enrol in all units that you expect to study in each Semester. Need help? Click here

Remember to Submit after making any changes.
Then wait for your Transaction Number.

Unit Enrolment form

Any *Unconfirmed* units are core requirements of your course. Click on *Enrol* in the Action column to confirm enrolment in these units.

Course Location: PARKVILLE(**FORCED** - What does this mean?)

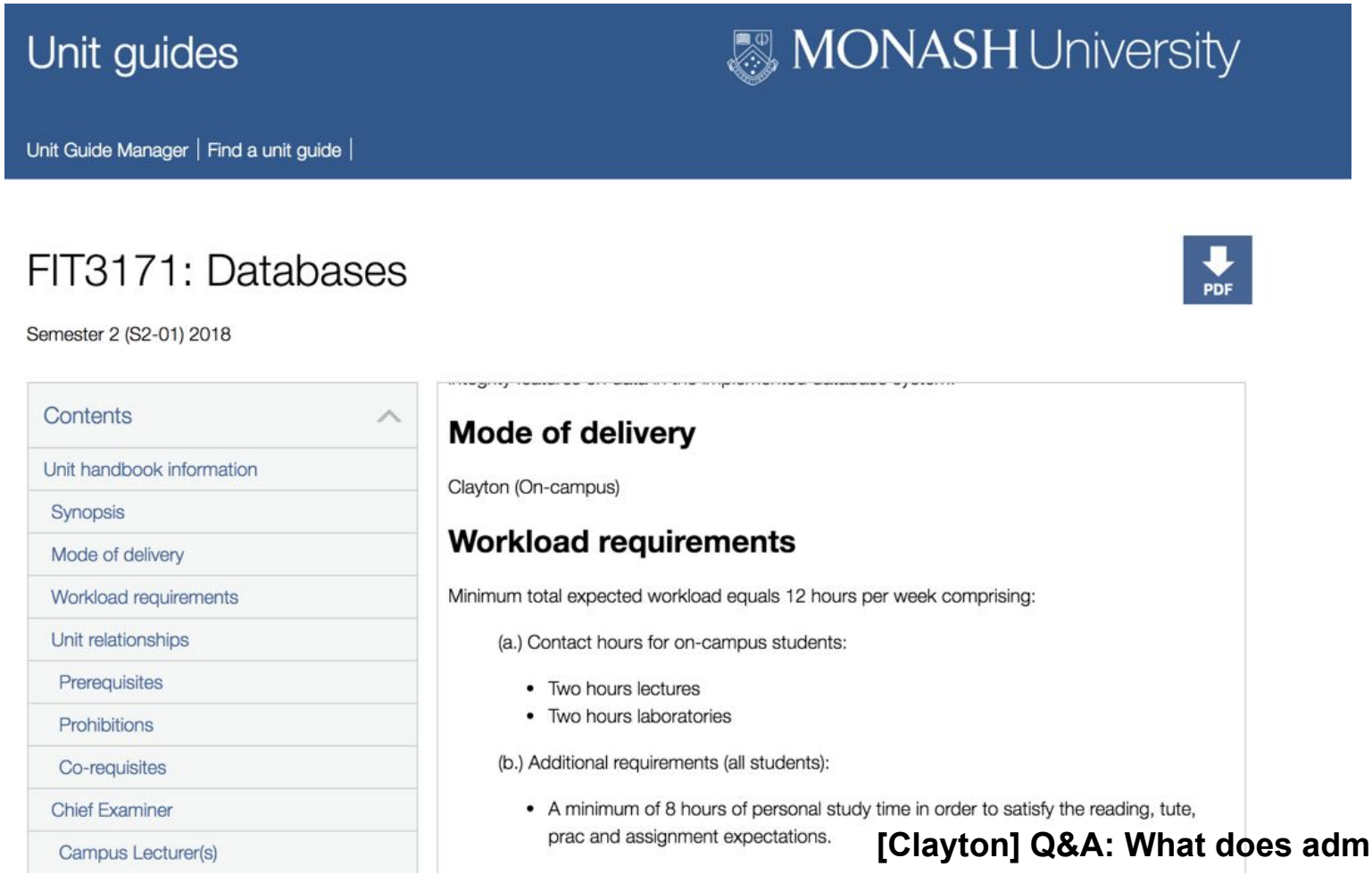
To Add Units Click here

Unit code	Action	Unit name	Campus	Semester	Type	Credits
PAC1111	Enrol	Introduction to physiology **UNCONFIRMED**	PAR	Semester 1 (2014)	DAY	--
PAC1121	Enrol	Bioorganic and medicinal chemistry I **UNCONFIRMED**	PAR	Semester 1 (2014)	DAY	--
PAC1211	Enrol	Physicochemical basis of pharmacy **UNCONFIRMED**	PAR	Semester 1 (2014)	DAY	--
PAC1311	Enrol	Pharmacy, health and society I **UNCONFIRMED**	PAR	Semester 1 (2014)	DAY	--
PAC1132	Enrol	Systems physiology **UNCONFIRMED**	PAR	Semester 2 (2014)	DAY	--
PAC1142	Enrol	Bioorganic and medicinal chemistry II **UNCONFIRMED**	PAR	Semester 2 (2014)	DAY	--
PAC1222	Enrol	Drug delivery I **UNCONFIRMED**	PAR	Semester 2 (2014)	DAY	--
PAC1322	Enrol	Pharmacy, health and society II **UNCONFIRMED**	PAR	Semester 2 (2014)	DAY	--
				Total credits:		0

If you do not get a Transaction Number after you submit, your enrolment is not complete.

[Clayton] Q&A: What do staff see?

Case study: Monash - Staff and Student View



[Clayton] Q&A: What does admin see?

Case study: Monash - Admin View

[Clayton] Q&A: What do students see?

[Clayton] Case study - Database design

Assume in 2050, you are building a new DB for Monash as per the prior case studies.

Were the previous examples - WES, Unit Guide, Allocate - constrained to only ONE table?

If yes - why?

If no - why?

Discuss, with reference to Lecture 1's discussion activity!

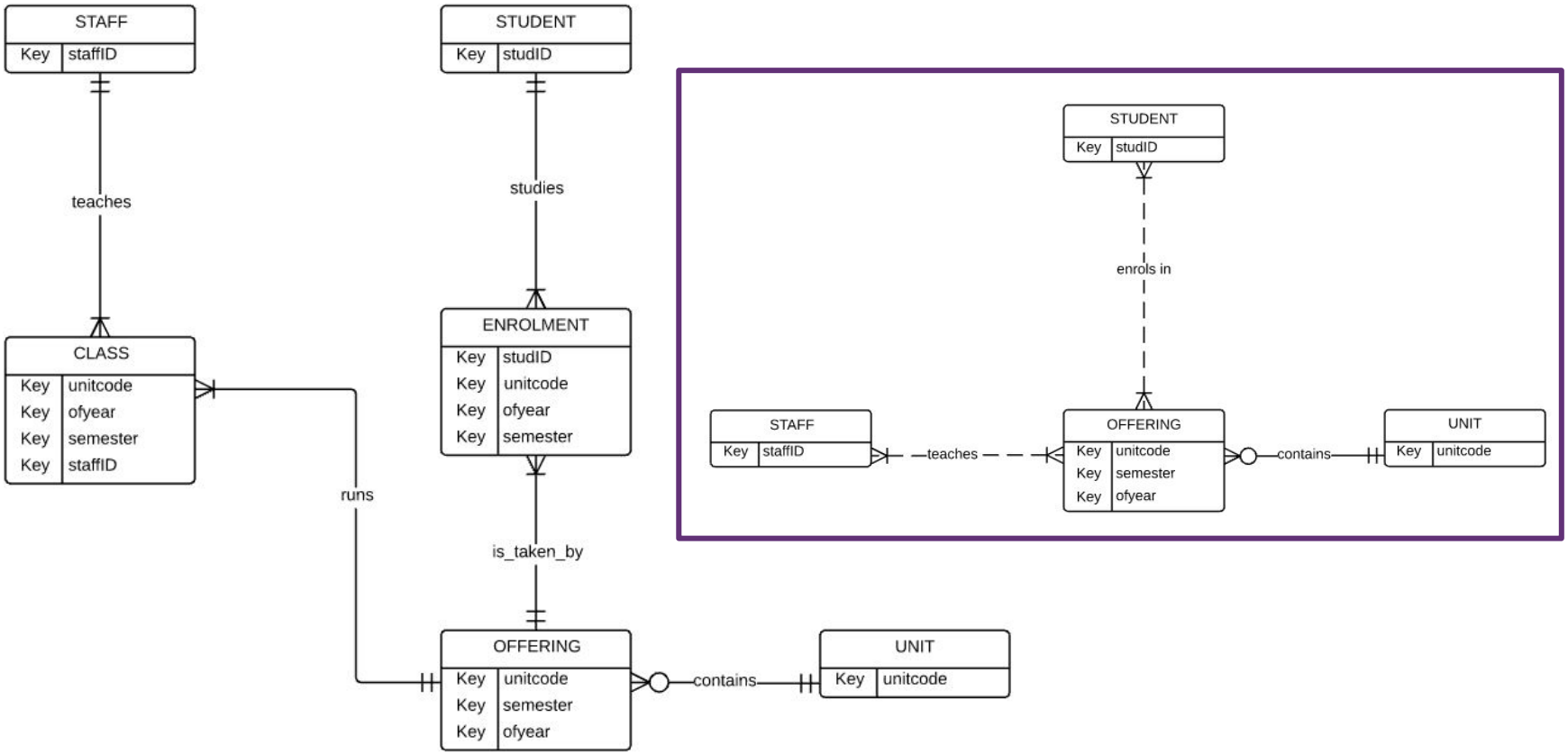
ER Modeling

- ER (Entity-Relationship) model developed by Peter **Chen** in 1976 to aid database design.
- May be used for conceptual (ERD)/logical design (ERD like).
- ER diagrams give a visual indication of the design.
- Basic components:
  - Entity
  - Attribute
  - Relationship

DBDLC: Conceptual Design

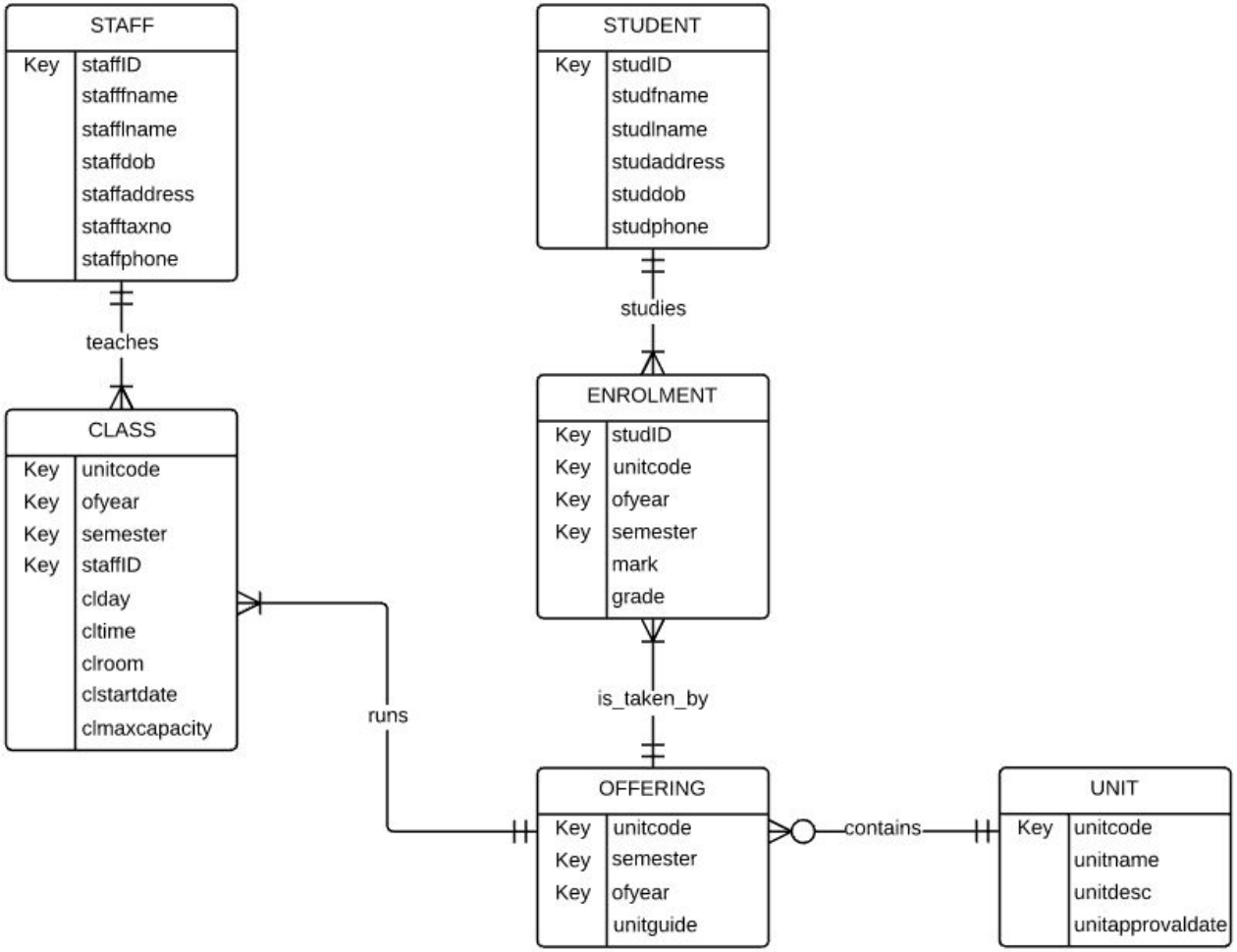
- Develop the enterprise data model.
- Corresponds to the conceptual level of the ANSI/SPARC architecture.
- **Independent of all physical implementation considerations.**
- Various design methodologies may be employed, including the ER (Entity-Relationship) approach.

Conceptual Level (ER Model): Keys only



[Clayton] Q&A: Why can't we just use the purple box?
What's the difference between that and the expanded edition to the left?

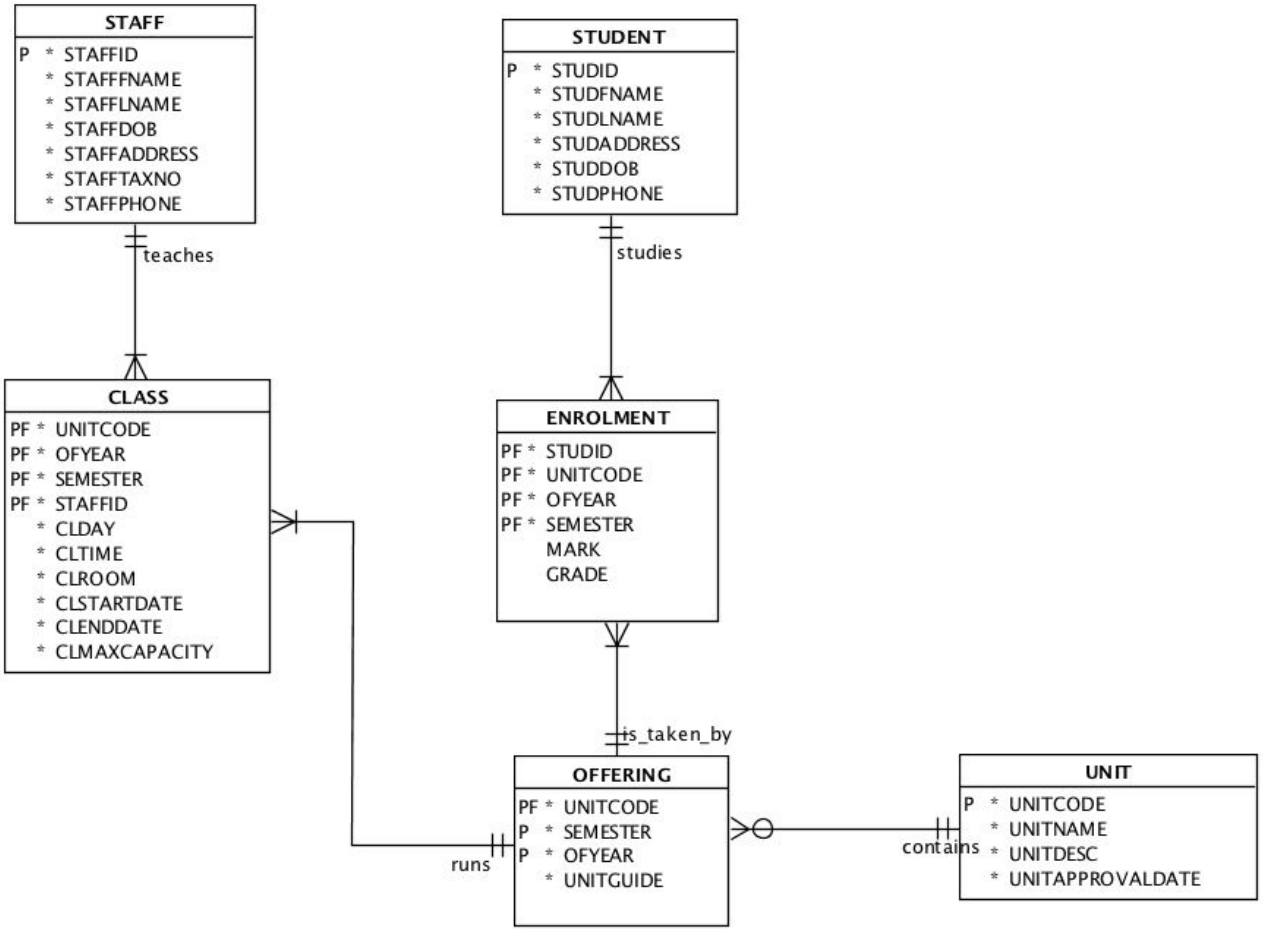
Conceptual Level (ER Model): All Attributes



DBDLC: Logical Design

- Develop a data model which **targets a particular database model** (e.g. **relational**, hierarchical, network, object-oriented).
  - –e.g. we're learning RDBs; e.g. Google uses NoSQL (Document-oriented DB)
- **Independent of any implementation details** which are **specific** to any particular DBMS package ('brand').
  - –e.g. our logical model can be used in Oracle, MySQL etc.
  - –e.g. Google's doc-DB model can be used in CouchDB, MongoDB, etc.
- Normalisation technique (see week 5) is used to test the correctness of the logical model.
- May also be considered to correspond to the conceptual level of the ANSI/SPARC architecture.

Logical Level (Logical Model)



DBDLC: Physical Design

- Develop a strategy for the **physical implementation of the logical data model.**
- Choose appropriate storage **structures**, **indexes**, **file organisations** and **access** methods which will most efficiently support the user requirements.
- Physical design phase is dependent on the **particular DBMS environment in use.**
- ANSI/SPARC internal level.
- Shown in SQL Developer Data Modeller as the Relational Model

Physical Level – Starting point in Oracle 11g

Oracle Database 11g | Relational_1 | Generate | Clear

```
create
  table enrolment
  (
    unitcode char (10) not null ,
    semester number (1) not null ,
    ofyear   date not null ,
    studid   number (10) not null ,
    mark     number (3) ,
    grade    char (2)
  ) ;


alter table enrolment add constraint enrol_mark_chk check (mark between 0 and
100) ;

alter table enrolment add constraint enrol_grade_chk check (grade in ('N','P',
'C','D','HD')) ;

alter table enrolment add constraint enrol_pk primary key ( semester, ofyear,
studid, unitcode ) ;

create
  table offering
  (
    unitcode  char (10) not null ,
    semester  number (1) not null ,
    ofyear    date not null ,
    chiefexam number (10) not null
  ) ;

alter table offering add constraint semester_chk check (semester between 1 and
3) ;


alter table offering add constraint offering_pk primary key ( unitcode,
semester, ofyear ) ;

create
  table prereq
  (
    unitcode      char (10) not null
```

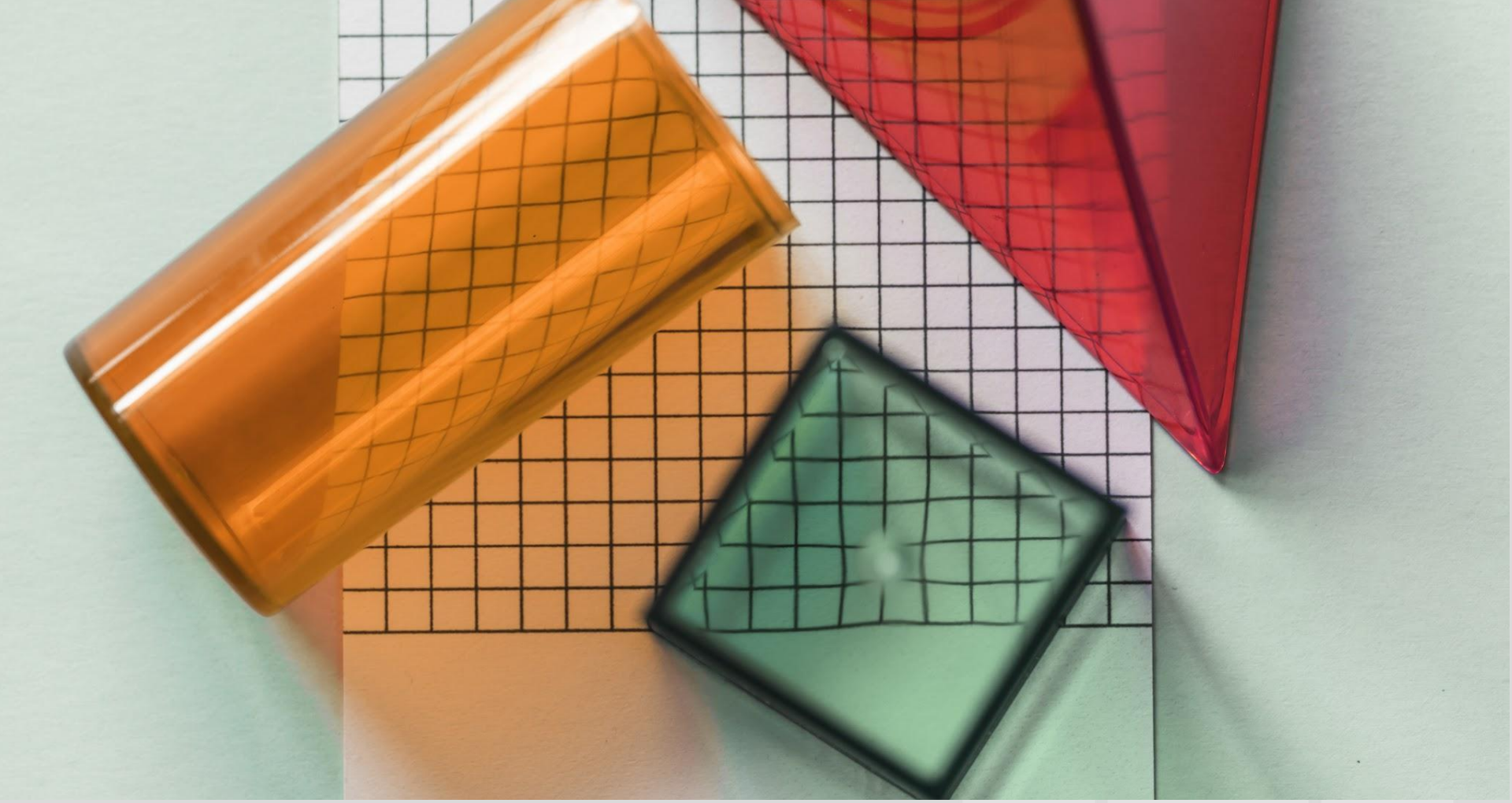
[Clayton] Example of different packages

Recall: **"Physical design phase is dependent on the particular DBMS environment in use."**

Let's study the creation of even a simple table, **enrolment.**

Oracle example as before - notice how MySQL rejects it, and requires correction...

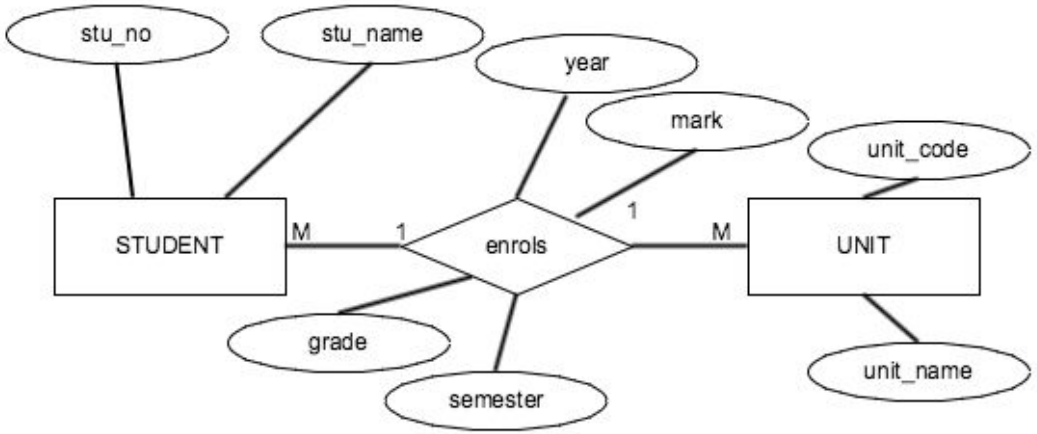
Img src: SQL Fiddle



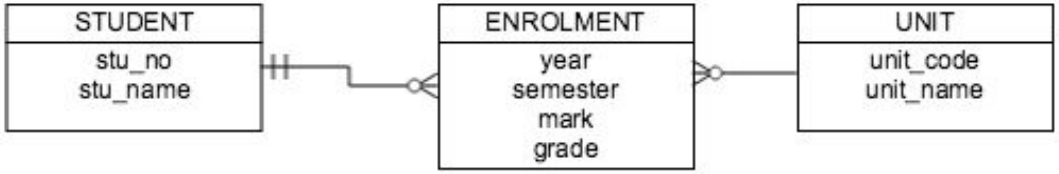
Entity Relationship Diagram (ERD)

Img src: RawPixel @rawpixel at Unsplash

ERD - Notation



Chen



Information Engineering/James Martin/Crows foot

ERD – Notation cont'd

Chen's Notation

- Semantically rich.
- Complex diagram.
- 'Pure' conceptual level.

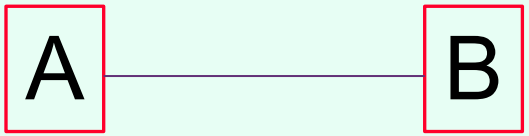
Information Engineering (Crow's Foot)

- Less semantics.
- Simpler diagram.
- **Mix** between conceptual and **logical** levels.

ENTITY, ATTRIBUTES, RELATIONSHIPS

Basic Theory!

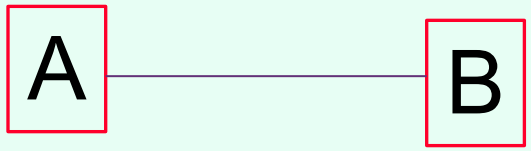
Simple questions now to test your understanding.
Actual **Case Study** work done after the break.



Please note this diagram is incomplete

Q1. How many entities are there in the above diagram?

A. 1
B. 2
C. 3
D. 4



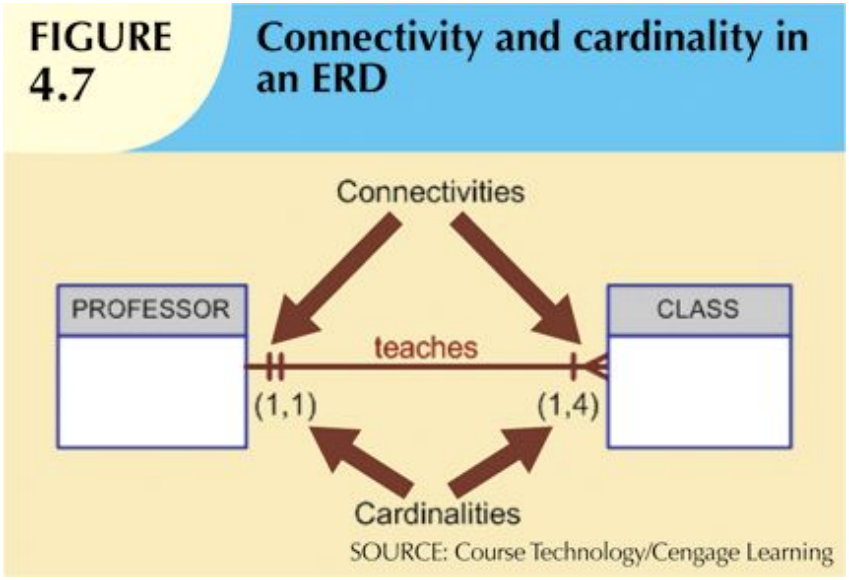
Please note this diagram is incomplete

Q2. How many relationship(s) are there on the above diagram?

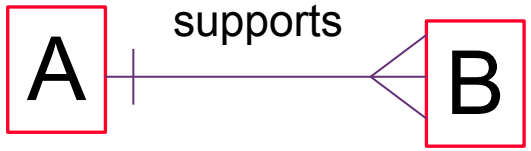
What is the degree of the relationship (the number of entities participating in the relationship) ?

A. 1, unary
B. 2, binary
C. 1, binary
D. 3, ternary

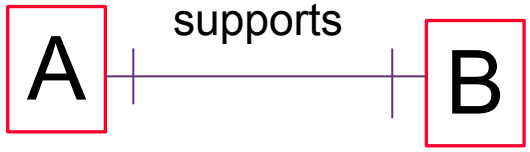
Crow's Foot: Cardinality vs Connectivity



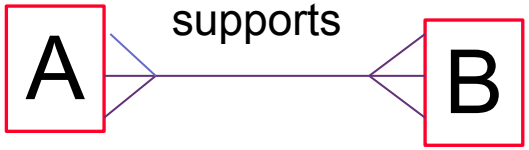
one to many

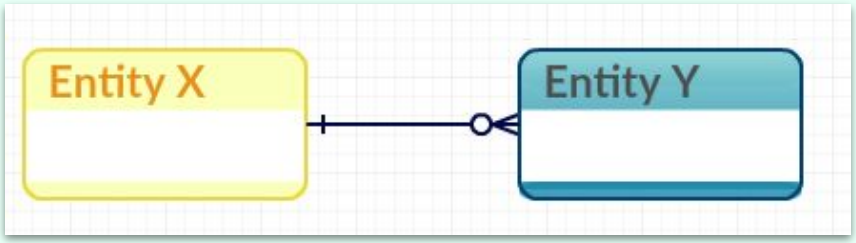


one to one



many to many





Please note this diagram is incomplete

Q3. Crow's Foot Notation.
Which of the following represents a CORRECT understanding of relationships (and cardinalities) between entities X and Y?

A. X = FacebookUser, Y = FacebookPost
B. X = Cell, Y = SpreadsheetFile
C. X = Student, Y = AcademicTranscript
D. X = TaxFileNumber, Y = Person
E. All of (A, B, C, D)
F. None of (A, B, C, D)

[Case Study] Read and discuss...

Q4. "An employee is assigned to be a member of a team. A team with more than 5 members will have a team leader. The members of the team elect the team leader."

List the entity(s) which you can identify in the above statement.

Each student only has to supply one entity.

(NB: any misuse of the FLUX system can be traced to your Authcate...)

[Case Study] Read and discuss...

Q5. "An employee is assigned to be a member of a team. A team with more than 5 members will have a team leader. The members of the team elect the team leader."

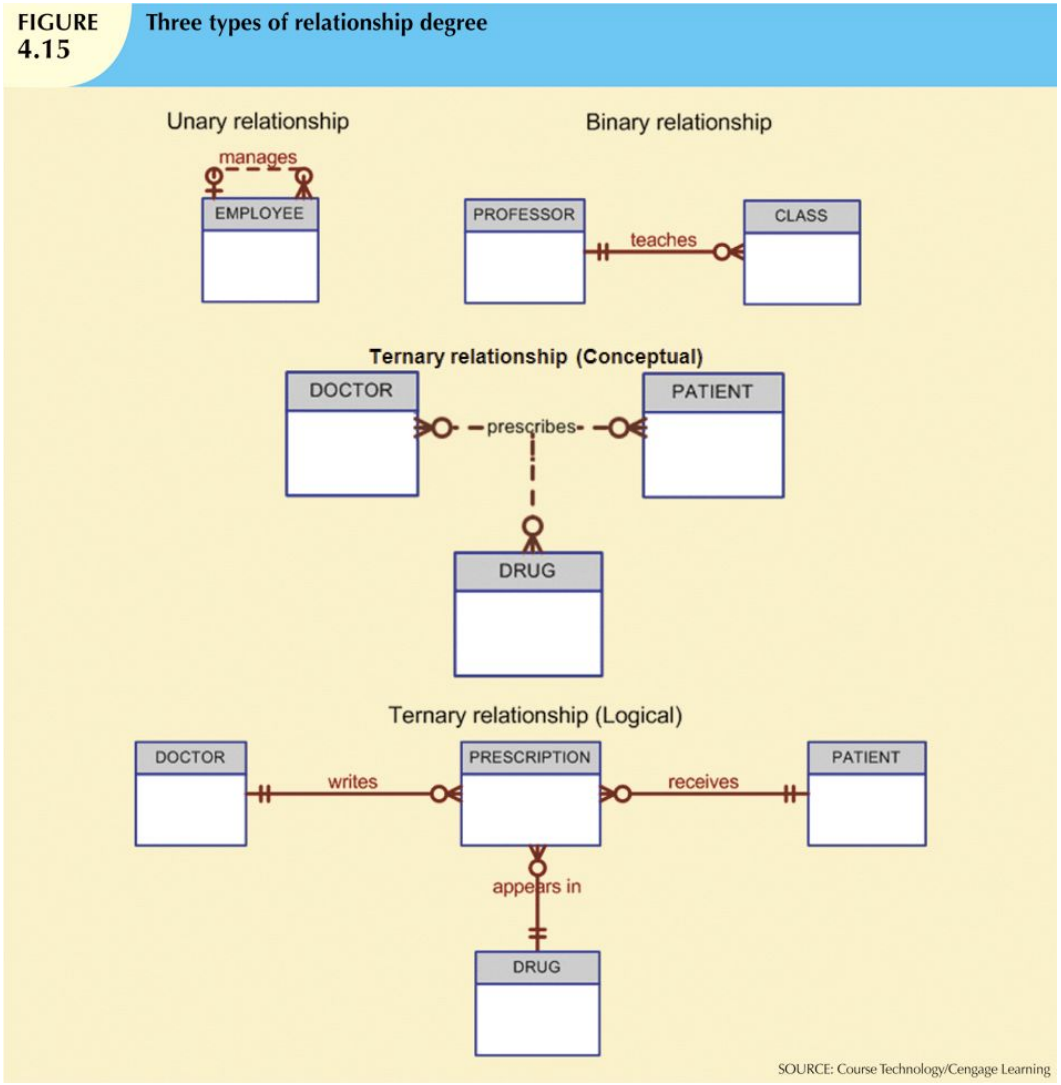
How many relationships connect TEAM and EMPLOYEE?

A. 0
B. 1
C. 2
D. 3
E. 4

Coffee break - see you in 10 minutes.

Img src: Jazmin Quaynor @jazminantoinette at Unsplash

Crow's Foot: Relationship Degrees

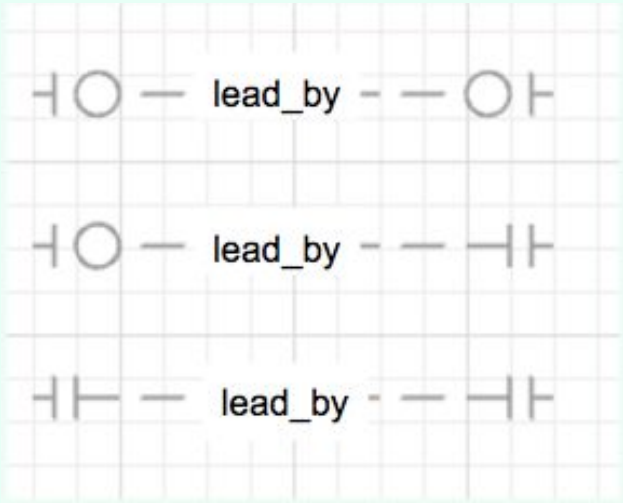


[Case Study] Read and discuss...

Relationship Participation:
Q6. "A Team with more than 5 members will have a team leader. The members of the team elect the team leader."

What is the relationship participation of the relationship
TEAM --- lead_by --- EMPLOYEE

A.



B.

C.

[Case Study] Read and discuss...

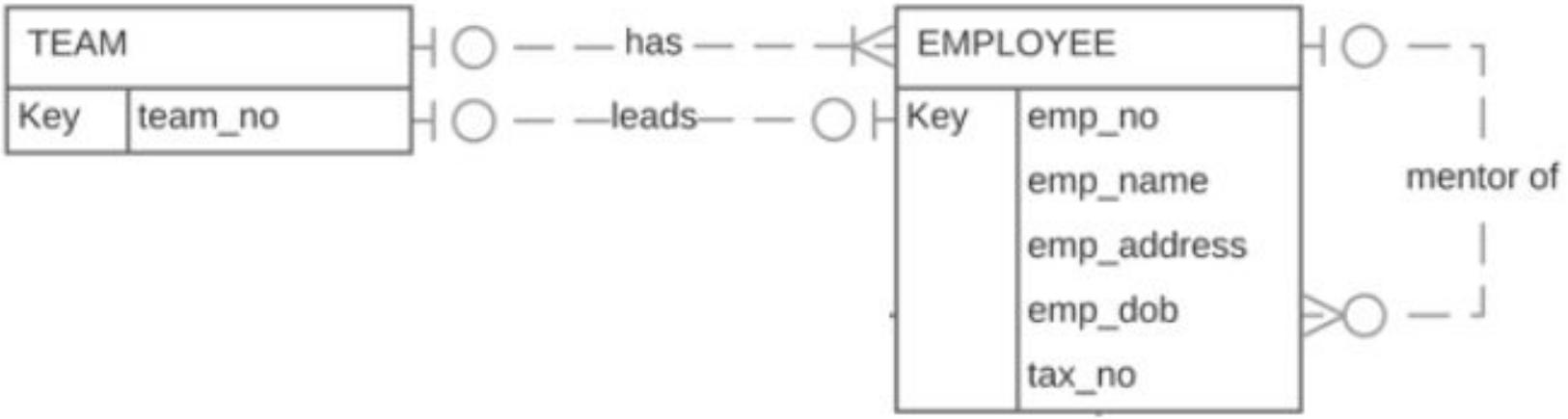
Q7. "The company also introduced a mentoring program, whereby a new employee will be paired with someone who has been in the company longer."

How many entity/ies do you need to model the mentoring program?

A. 0
B. 1
C. 2
D. 3

[Case Study]

Case study so far - **Team, Employee**



Discussion points: The connectivity and cardinality of Employee-Employee? Relationship degree?

[Case Study] Read and discuss...

Q8. "To attract high calibre talent, the company provides generous remuneration package as well as health insurance support for the employees and their family. To do this, the HR team needs to know the details of the family members. The family information will be recorded"

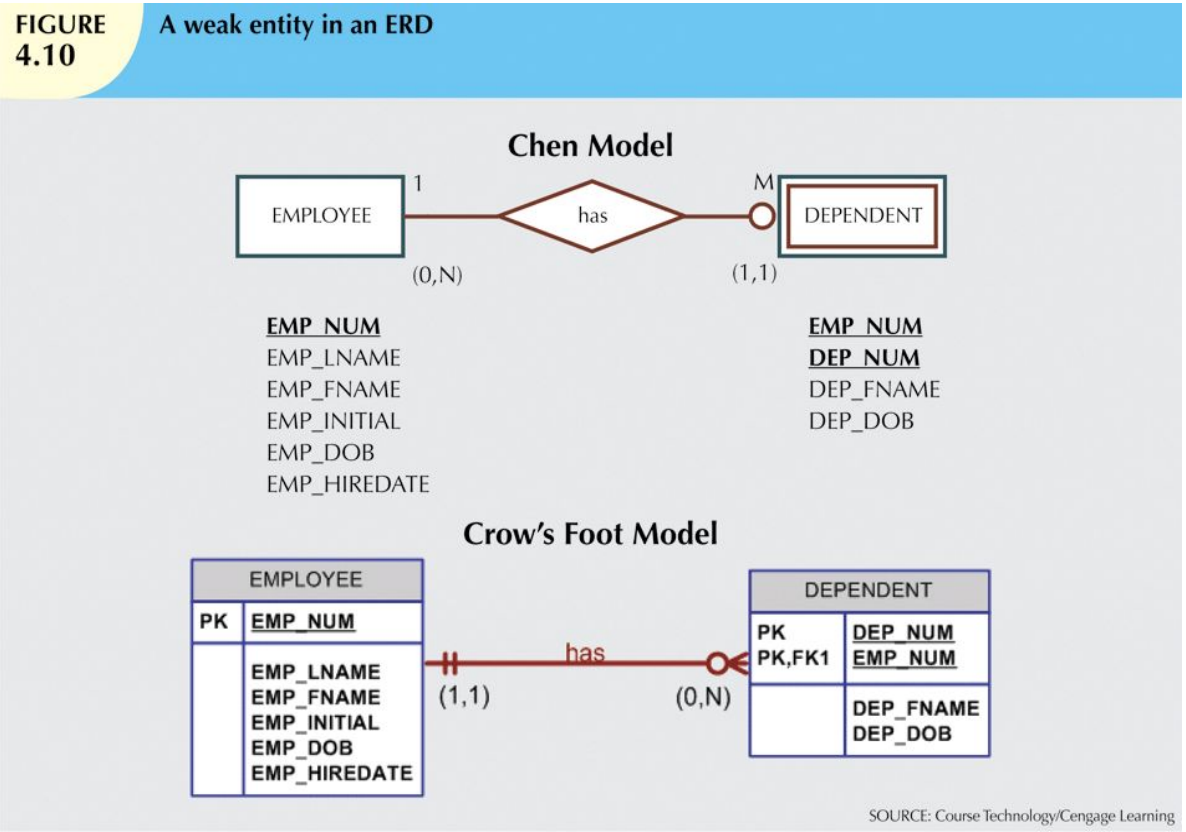
Choose a TRUE statement.

A. EMPLOYEE entity is a strong entity and FAMILY is a strong entity.
B. EMPLOYEE entity is a weak entity and FAMILY is a strong entity.
C. EMPLOYEE entity is a strong entity and FAMILY is a weak entity.
D. EMPLOYEE entity is a strong entity and FAMILY is strong entity.

Weak vs Strong Entity

- Strong entity
  - Has a key which may be defined without reference to other entities.
  - For example EMPLOYEE entity.
- Weak entity
  - Has a key which requires the existence of one or more other entities.
  - For example FAMILY entity - need to include the key of employee to create a suitable key for family
- Database designer often determines whether an entity can be described as weak based on business rules
  - customer pays monthly account
    - Key: cust_no, date_paid, or
    - Key: payment_no (surrogate? – not at conceptual level)

Weak vs Strong Entity



[Marc's Anecdote]

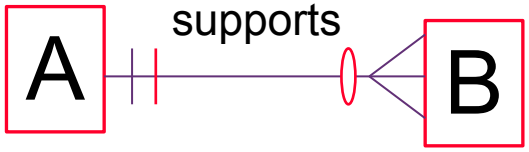
- Say you're a game designer.
- Weak vs Strong entity
  - Strong - "has a key which may be defined without reference to other entities" - e.g. Character
    - PK: **CharID = C001**, CharName = "Tracer"
  - Weak - "has a key… **requires the existence** of one or more other entities" - e.g. CharacterSkin
    - PK: **CharSkinID = T0004, CharID = C001**, Description="Overwatch Summer Games Tracer"
  - **NOTE: the Chen paper is quite particular on the rules of Strong vs Weak!**
    - e.g. long argument on https://stackoverflow.com/questions/4741967/example-of-a-strong-and-weak-entity-types
  - Lindsay's anecdote:

To be classified as a weak entity, ***two conditions must be met***:

1. The entity must be existence-dependent on its parent entity.
2. The entity must inherit at least part of its primary key from its parent entity.

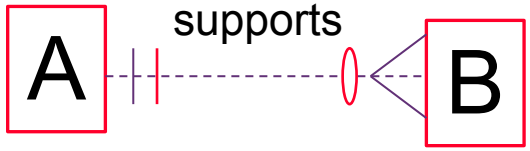
Identifying vs Non-Identifying Relationship

- **Identifying**
  - Identifier of A is part of identifier of B.



  - Shown with solid line
  - Enrolment's PK includes student id, which is an identifier of student.

- **Non-identifying**
  - Identifier of A is NOT part of identifier of B.



  - Shown with broken line
  - Department no (identifier of department) is not part of Employee's identifier.

Types of Attributes

- Simple
  - Cannot be subdivided
  - Age, sex, marital status
- **Composite**
  - Can be subdivided into additional attributes
  - Address into street, city, zip
- Single-valued
  - Can have only a single value
  - Person has one social security number
- **Multi-valued**
  - Can have many values
    - 0, 1, Many...
  - Person may have several college degrees
- Derived
  - Can be derived with algorithm
  - Age can be derived from date of birth

[Case Study] Read and discuss...

Q9. "The employee details that will be recorded are: Employee number, Full name, Address, Date of birth, Tax file number and Skill(s). Examples of skills are Java, Python, UNIX, Relational DB, Mongo DB, etc

Choose a TRUE statement.

A. ADDRESS is a multi-valued attribute and SKILL is a multi-valued attribute.

B. ADDRESS is a composite attribute and SKILL is a composite attribute.

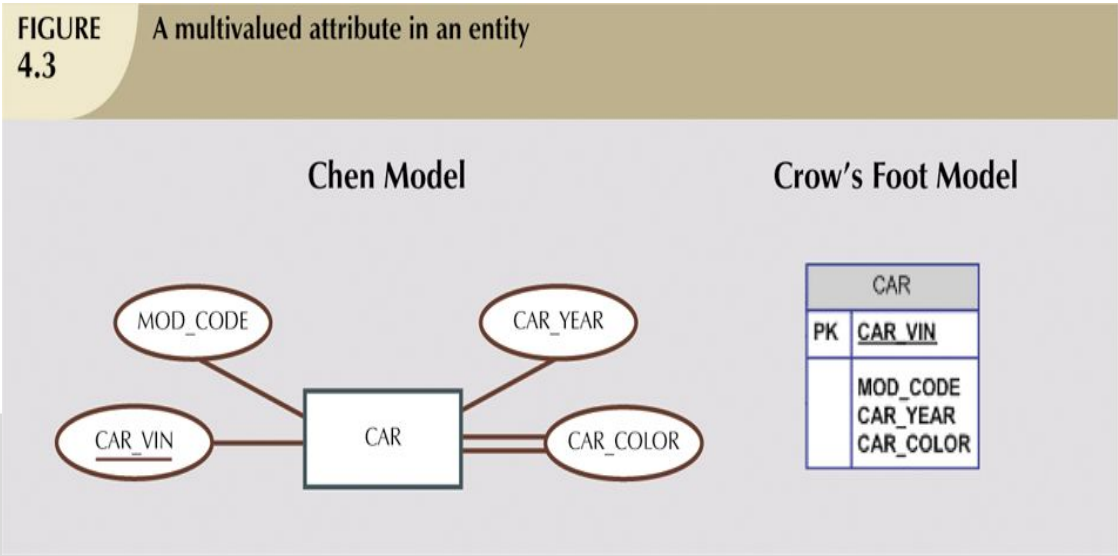
C. ADDRESS is a composite attribute and SKILL is a multi-valued attribute.

D. ADDRESS is a multi-valued attribute and SKILL is a composite attribute.

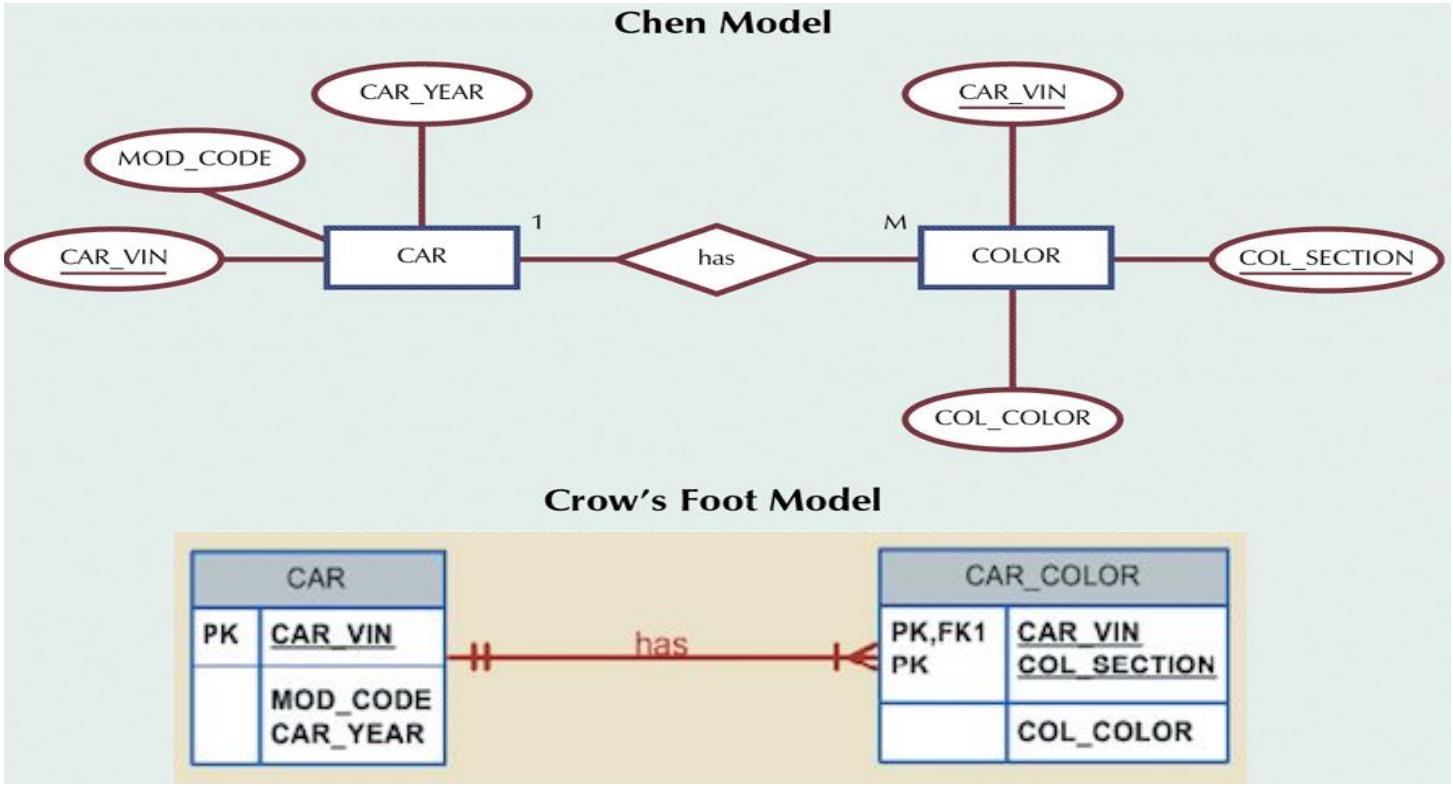
Multivalued Attribute

- An attribute that has a list of values.
- For example:
  - Car colour may consist of body colour, trim colour, bumper colour.
- Crow's foot notation does not support multivalued attributes. Values are listed as a separate attribute.

FIGURE 4.3 A multivalued attribute in an entity

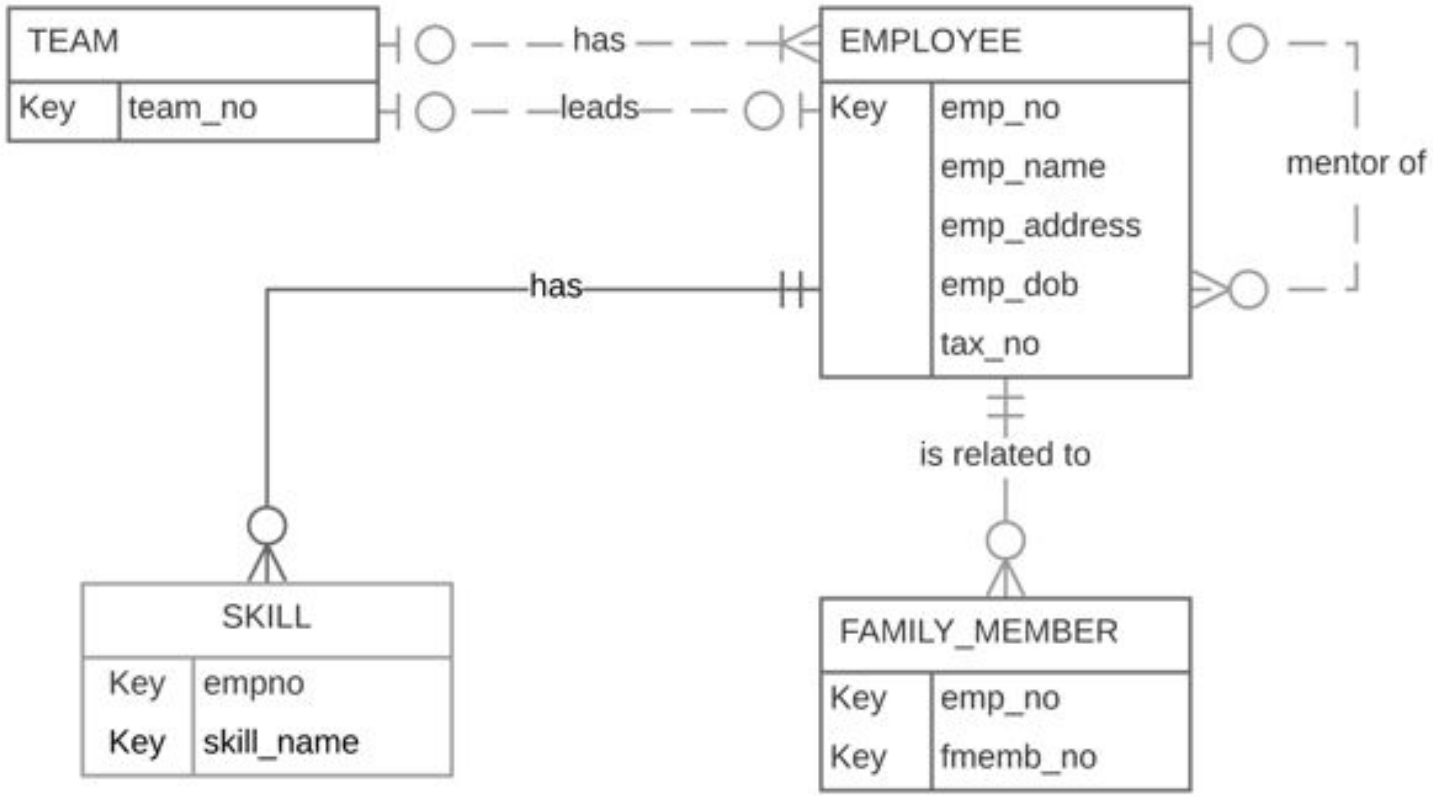


Resolving Multivalued Attributes



[Case Study]

Case study so far - Team, Employee, **Skill, Family_Member**



Discussion points: Family_Member as a weak entity (needs emp_no)?
Skill as a multi-valued attribute (not composite in Employee!).

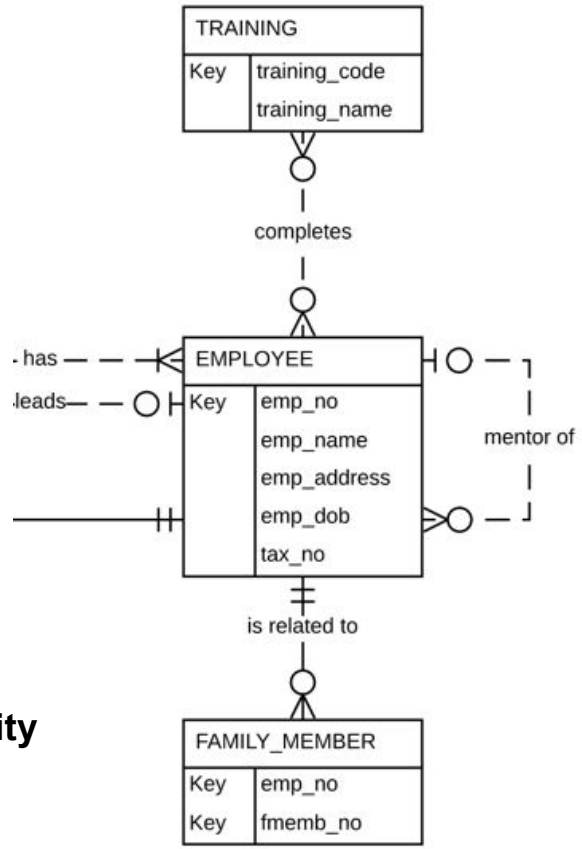
[Case Study]

Case study so far - ..., Employee, ..., Family_Member, **Training**

..."the company provides several in-house training programs. The HR team needs to keep track of the details about who has done what. An employee can do several training programs.

At the completion of a training, a certificate will be provided to the employee containing the training name and the completion date."...

Incomplete model



Discussion points: The connectivity and cardinality of Training-Employee? Relationship degree?

Incomplete model

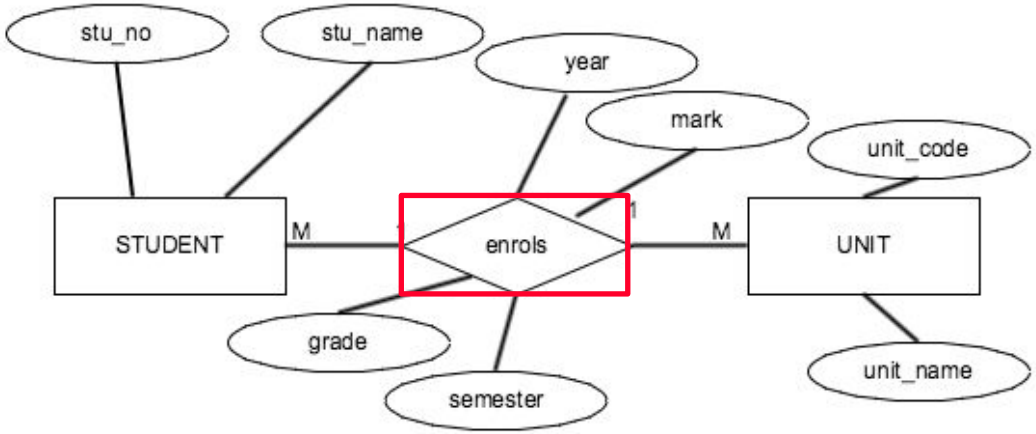
[Case Study] Read and discuss...

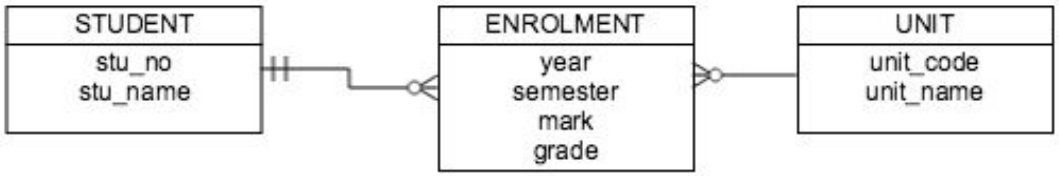
Q10. LAST QUESTION...

Which relationship requires an 'associative entity' (or colloquially, a 'junction' table)?

A. EMPLOYEE - TRAINING
B. EMPLOYEE - FAMILY_MEMBER
C. Both A and B...
D. None of A and B...

Associative (or Composite) Entity





Discussion points: A many-to-many rel with B -- cannot be represented in DBMS straightaway!
Needs an 'associative' entity (conceptual level, formally).
Other colloquialisms include: 'junction'/'bridge' table (physical level).

[Case Study]

Case study so far - Everything!