COMM1822

Term 2 2022

Introduction to Databases for Business Analytics

Week 1 Entity Relati
(ER) Modelling Part 1

Lecturer: Kam-Fung (Henry) Cheung
E-mail: kf.cheung@unsw.edu.au

UNSW SYDNEY

Copyright

- There are some file-sharing websites that specialise in buying and selling academic work to and from university students.

- If you upload your original wor other student downloads and presents it as their own eit **ight be found guilty of collusion — even years after graduatio**

- These file-sharing websites may also accept purchase of course materials, **such as copies of lecture slides and tutorial handouts. By law, the copyright on course materials, developed by UNSW staff in the course of their employment, belongs to UNSW. It constitutes copyright infringement, if not academic misconduct, to trade these materials.**

Acknowledgement of Country

UNSW Business School acknowledges the Bidjigal (Kensington campus) and Gadigal (City campus) the traditional custodians of the lands where each campus is located.

We acknowledge all Aboriginal and Torres St Islander Elders, past and present and thei communities who have shared and practiced their teachings over thousands of years including business practices.

We recognise Aboriginal and Torres Strait Islander people's ongoing leadership and contributions, including to business, education and industry.

UNSW Business School. (2022, May 7). *Acknowledgement of Country* [online video]. Retrieved from https://vimeo.com/369229957/d995d8087f

Agenda

- **Data Modelling**
  - ❑ Data model as a (relatively) simple abstraction of the complex real-world (for the purpose of creating a DB).
  - ❑ A good DBMS will perform atabase.
  - ❑ One modelling technique t **ationship Modelling**
- **Entity Relationship Mod**
  - ❑ Entity Types and Entity Instances
  - ❑ Attributes and Values
  - ❑ Keys
  - ❑ Relationships
  - ❑ Connectivity
  - ❑ Cardinality

Chapter 2
Data Models
2-1 to 2-6

r 4
elations
g
4-1 to 4-2

Data Modeling and Data Models

- **Model** - Abstraction of a real-world object or event
- **Data modeling**: Iterative and progressive process of creating a specific data model for a deter

To model and translate t ta m
be used to store data which to business can u

- **Data models**: Simple representations real-world data structures
  - Useful for supporting a specific problem domain

What would be the data model like for UNSW?

Importance of Data Models

- Are a communication tool
- Give an o e
- Organize data for various
- Are an abstraction for the creation of good database

Data Model Basic Building Blocks

- **Entity**: Unique and distinct object used to collect and store data
  - e.g., people, thing, event, …
  - ❑ **Attribute**: Characteristic of an entity
- **Relationship**: Des https://eduassistpro.github.io/ among entities
  - Cons g entities: professors, stude partments, research centres, etc.
  - ❑**One-to-many (1:M)**
  - ❑**Many-to-many (M:N or M:M)**
  - ❑**One-to-one (1:1)**
- **Constraint**: Set of rules to ensure data integrity

Violation examples:

- Enter an SID to STUDENT table, it finds two students (Entity integrity)
- STUDENT table says that student X's department code is Y. But in the DEPARTMENT Table, there is no code Y. (Referential integrity)

Business Rules – Design a Data Model

Brief, precise, and unambiguous description of a policy, procedure, or principle

Enable blocks

Describe main and distinguishing characteristics of the data

Translating Business Rules into Data Model Components

- Nouns translate into entities

Each student can take at most 3 courses each semester.

Each research student must have two supervisors at UNSW.

- Verbs translate into rel es

- Relationships are bidirectional

Ea ke at most 3 courses each semester.

Ea ave up to 300 students at UNSW.

- Questions to identify the relationship type: 1:1, 1:M or M:N
  - ❑How many instances of B are related to one instance of A?
  - ❑How many instances of A are related to one instance of B?

Naming Conventions

- Entity names - Required to
  - ❑ Be descriptive of the objects in the business environment
  - ❑ Use terminology that is familiar to the users

STUDENT, EMPLOYEE, DEPARTMENT, …

- Attribute name - Requi the data represented by the attribute

STUDENT: SID, D FIRST_NAME, LAST_NAME, …

- Proper naming
  - Facilitates communication between parties
  - Promotes self-documentation

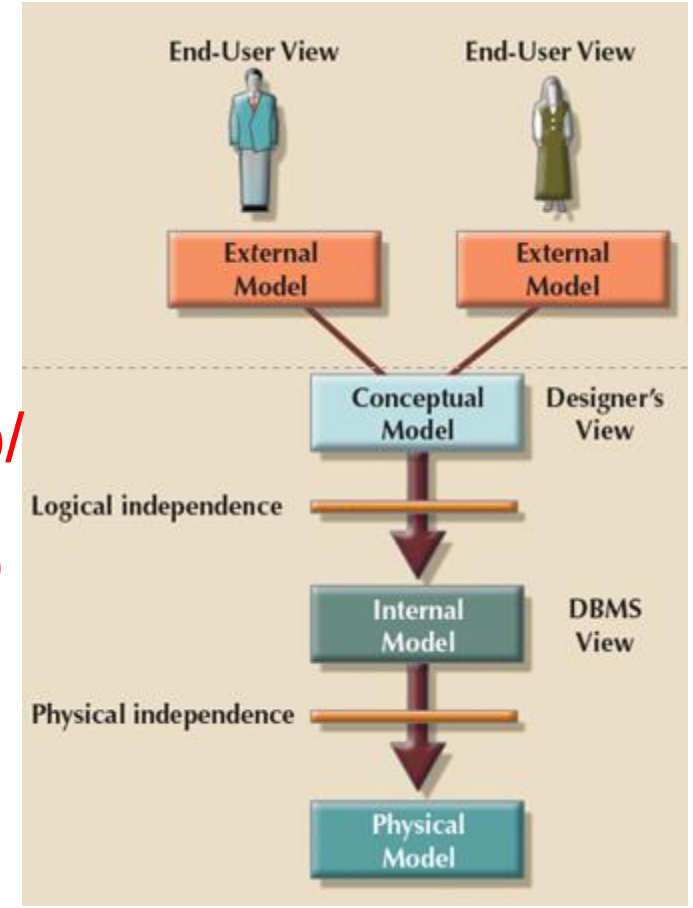
Evolution of Data Models

We will focus and discuss more on relational, entity relationship, and NoSQL in the last few weeks.

Source: Coronel, Morris, Rob 2017

Levels of Data Abstraction

Model	Degree of Abstraction	Focus	Independent of
External	High	End user views	Hardware and software
Conceptual		Glob mode	e and
Internal		Speci	e
Physical	Low	Storage and access metho	dware e



Source: Coronel, Morris, Rob 2017

Data abstraction is the reduction of a particular body of data to a simplified representation of the whole.

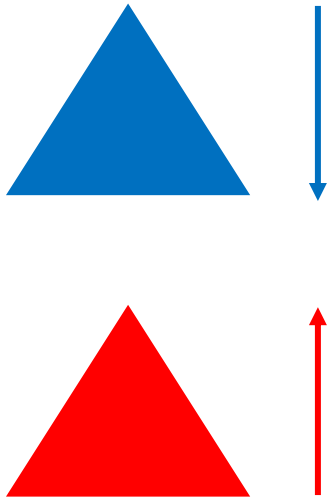
Abstraction, in general, is the process of taking away or removing characteristics from something in order to reduce it to a set of essential characteristic.

Conceptual Data Modelling Techniques

Two common techniques:

- **Entity-Relationship (ER** proach. Begins by looking for the data group

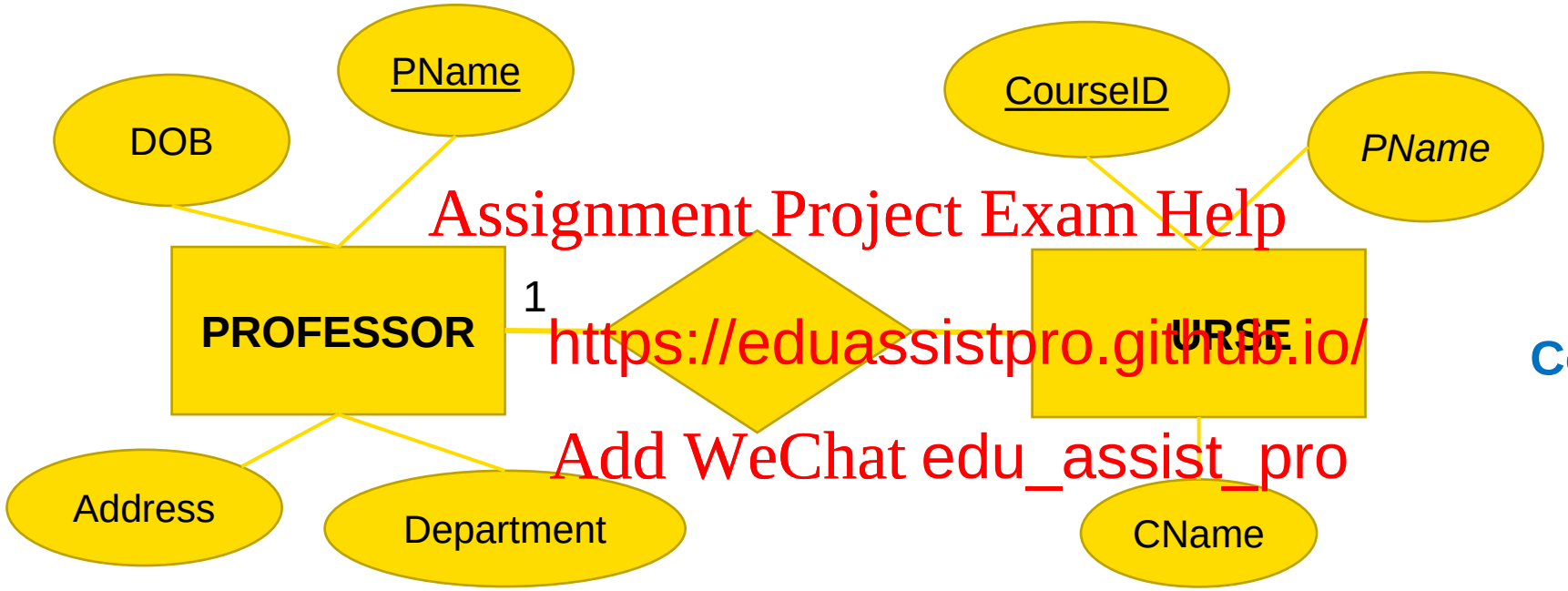
- **Normalization**: Bottom-up approach. Be ng at the smallest individual items of data recorded by the system.



Internal Model and Conceptual Model

- The internal model is the model that we used when database is **implemented**.
- The internal model maps the **conceptual model to the DBMS.**
- The internal model depend **ftware.**
- Hence, a change in DBMS software requires internal model be changed.
- **Logical independence**: you can change the internal model without affecting conceptual model!

Conceptual Model



Conceptual Model

PROFESSOR (**PName**, DOB, Address, Department)
COURSE (**CourseID**, CName, ***PName***)

Internal Model

Conceptual Modelling: ER Model

- An **Entity-Relationship (ER) model** is a detailed, logical representation of the data for an organisation or for a business area.
- The ER model is expressed in **entities** in the business enviro **relationships** or associations those entities, and the **attributes** of both the entities and their relationships.
- An ER model is normally expressed as an **ER diagram**, which is a graphical representation of an ER model. In this course we will follow **Chen's notation**.

ER Model Notations

Agenda

- **Data Modelling**
  - ❑ Data model as a (relatively) simple abstraction of the complex real-world (for the purpose of creating a DB).
  - ❑ A good DBMS will perform atabase.
  - ❑ One modelling technique t **ationship Modelling**
- **Entity Relationship Mod**
  - ❑ Entity Types and Entity Instances
  - ❑ Attributes and Values
  - ❑ Keys
  - ❑ Relationships
  - ❑ Connectivity
  - ❑ Cardinality

Entity Relationship Modelling (ERM)

- Basis of an entity relationship diagram (ERD)

- ERD depicts the:
  - Conceptual database
  - Database's main components
    - Entities (Tables)
    - Attributes (Columns of tables)
    - Relationships (Associations between tables)

- Entity - Refers to the entity set and not to a single entity occurrence

Entity (Type) and (Entity) Instance

- **en·ti·ty** /ˈentitē/ (Noun)
  - A thing with distinct and independent existence.
  - Existence; being: "entity and nonentity".
- Synonyms: being - existence - essence - thing

- **in·stance** /ˈinstəns/ (Noun)
  - An example or single occurrence of something: "an i tion".
  - A particular case: "in this instance".
- Synonyms: example - case - sample - event - occurrence - exemplar

- One type of things is a Person. Joe is an instance of Person.
- One type of things is a Drink. Espresso an instance of Drink.

Entity (Type) and (Entity) Instance

- **Entities**: “An **entity** is an object about which the system requires to hold data.”
- An **entity type** (entity class) is a collection of entities that share common properties or characteristics (similar to be grouped into one Entity Type).
- It is represented as a re **del** diagram with the name of the entity inside.
- An **entity instance** is a **single occurre** **tity type.**

STUDENT

Entity Type	Entity Instances		
STUDENT	z1234567	Martin, S.	BEng
	z1357926	Fong, L.	BSc

UNSW SYDNEY

ER Model with Only Appropriate Entities

A treasurer looks after researchers' research accounts. Each account pays more than one expenses. The treasure prints expense reports regularly, e.g., every month.

Attribute

Characteristics of entities

- **Required attribute**: Must have a value, cannot be left empty
- **Optional attribute**: Does not require a value, can be left empty
- **Domain**: Set of possible
- **Identifiers**: One or more attributes that . ntify each entity instance

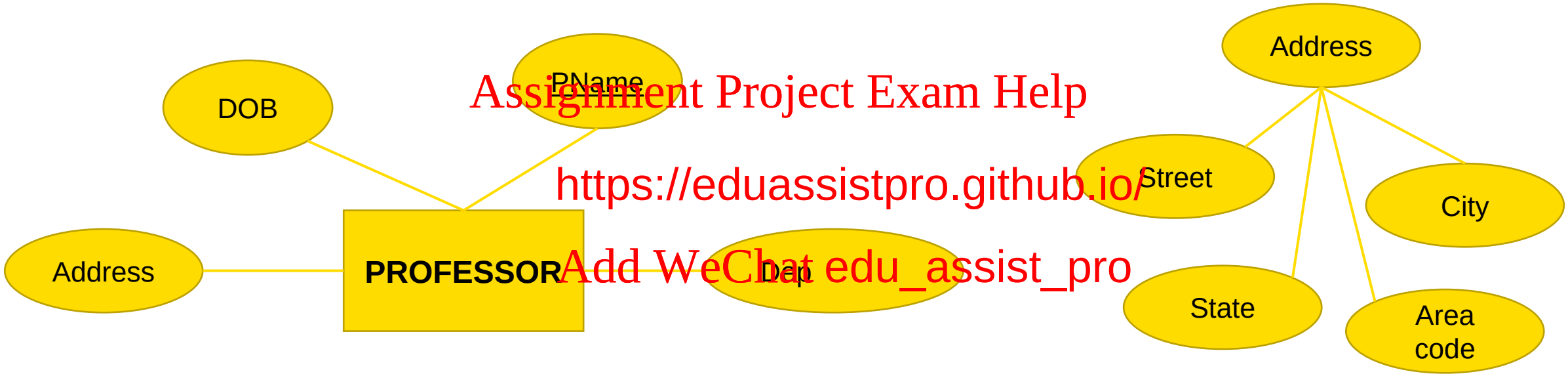
he relational model

Attributes

- **Simple attribute**: Attribute that cannot be subdivided
  - ❑ Example: zID
- **Composite attribute**: Attribute that can be subdivided to yield additional attributes
  - ❑ Example: Address (= str
- **Single-valued attribute**: Attribute th a single value
- **Multivalued attribute**: Attribute that have many values
- **Derived attribute:** Attribute that derived using an algorithm

- **A Key attribute** is unique so to identify the entity.

Example of Attributes



Entity: PROFESSOR
Attribute: PName, Department, Address, DOB

Composite attribute: Address

A Multivalued Attribute in an Entity

multiple values

Splitting the Multivalued Attributes into New Attributes

Is this a good idea?

Depends on your design!

Depiction of a Derived Attribute

Derived attribute is when the value is calculated from other attributes.
e.g., EMP_AGE can be calculated from EMP_DOB.

Do we need EMP_AGE?

Technically, we do not store the employee's age because we can calculate from date of birth of the employee. Otherwise, you have to update the age every day.

Advantages and Disadvantages of Storing Derived Attributes

	Derived Attribute: Stored	Derived Attribute: Not Stored
Advantage	• Saves CPU proc • Saves data acce • Data value is rea • Can be used to keep track of historical	es storage space putation always yields current value
Disadvantage	• Requires constant maintenance to ens derived value is current, especially if any values used in the calculation change	PU processing cycles • Increases data access time • Adds coding complexity to queries

Keys

- Consist of one or more attributes that determine other attributes
- Used to
  - ❑ Ensure that each row in a table is uniquely identifiable
  - ❑ Establish relations d to ensure the integrity of the data
- **Primary key (PK)**: Attribute or co of attributes that uniquely identifies any given row

For example, STUDENTS table, the PK is zID; EMPLOYEES table, PK is employee ID.
A PK may contain more than one attribute.

Find the Primary Keys

Table: STUDENT	Example
zID	z1234567
Email	z1234567@student.unsw.edu.au
LastName	Bold
FirstName	Alice
DOB	28/02/2010

Table: CLASS_ENR	Example
zID	z1234567
CourseID	COMM1822
	2022T2
	W16A
	M18A

Table: COURSE	Example
CourseID	COMM1822
CourseName	Intro to DB for BusAn

Table: DEGREE	Example
zID	z1234567
ProgramID	3347

For simplicity, not all attributes are included in the tables. Also, I used COMM1822 as the course ID.

Answer for the Primary Keys

Table: STUDENT	Example	Keys
zID	z1234567	PK
Email	z1234567@student.unsw.edu.au	
LastName	Bold	
FirstName	Alice	
DOB	28/02/2010	

Table: CLASS_ENR	Example	Keys
zID	z1234567	PK
CourseID	COMM1822	PK
	2022T2	PK
	W16A	
	M18A	

Table: COURSE	Example	Keys
CourseID	COMM1822	PK
CourseName	Intro to DB for BusAn	

Table: DEGREE	Example	Keys
zID	z1234567	PK
ProgramID	3347	PK

Types of Keys

- **Composite key**: Key that is composed of more than one attribute

  e.g., the CLASS_ENR table has (zID, CourseID, TermID) is a composite key

- **Key attribute**: Attribute that is a part of a key

- **Superkey:** Key that can uni table

  e.g., zID, {zID, Last_Name}, {zID, First_Name, Last_Name}, … in STUDENT table

- **Candidate key**: Minimal superkey

  e.g., zID in STUDENT table; mobile number can identify you if you forget your rewards card.

- **Entity integrity**: Condition in which each row in the table has its own unique identity
  - ❑ All of the values in the primary key must be unique
  - ❑ No key attribute in the primary key can contain a null

Types of Keys

- **Null**: Absence of any data value that could represent
  - ❑ An unknown attribute value
  - ❑ A known, but missing, attribute value
  - ❑ An inapplicable conditi
- **Referential integrity**: tity instance by another entity instance is valid
- **Foreign key (FK)**: Primary key of one has been placed into another table to create a common attribute
- **Secondary key:** Key used strictly for data retrieval purposes

  e.g., people do not remember their membership no. (PK), the secondary key can be their name, which may not be unique.

Example of Foreign Key

Table: CLASS_ENR	Example	Keys
zID	z1234567	PK, FK
CourseID	COMM1822	PK
TermID	2022T2	PK
Lab	W16A	
Lecture	M18A	

Table: COURSE	Example	Keys
CourseID	COMM1822	PK
CourseName	Intro to DB for BusAn	

Table: DEGREE	Example	Keys
zID	z1234567	PK, FK
ProgramID	3347	PK

Table: STUDENT	Example	Keys
	z1234567	PK
	z1234567@student.unsw.edu.au	
	Id	
FirstName	Alice	
DOB	28/02/2010	

Relationships

A relationship is a link between two entities which is significant for the system.

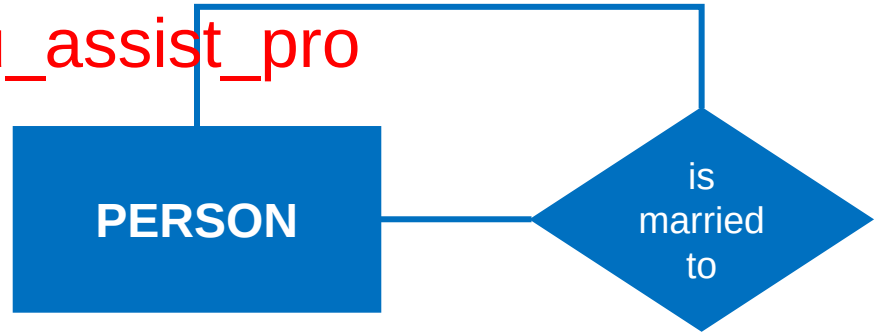
- The **degree of a relati** **f entity types** that participate in that relationship.

- The most common relationships are **u** **y, ternary, and quaternary**.
- The relationships between entities can be
  - ❑ One-to-One 1:1
  - ❑ One-to-Many 1:M
  - ❑ Many-to-Many M:N

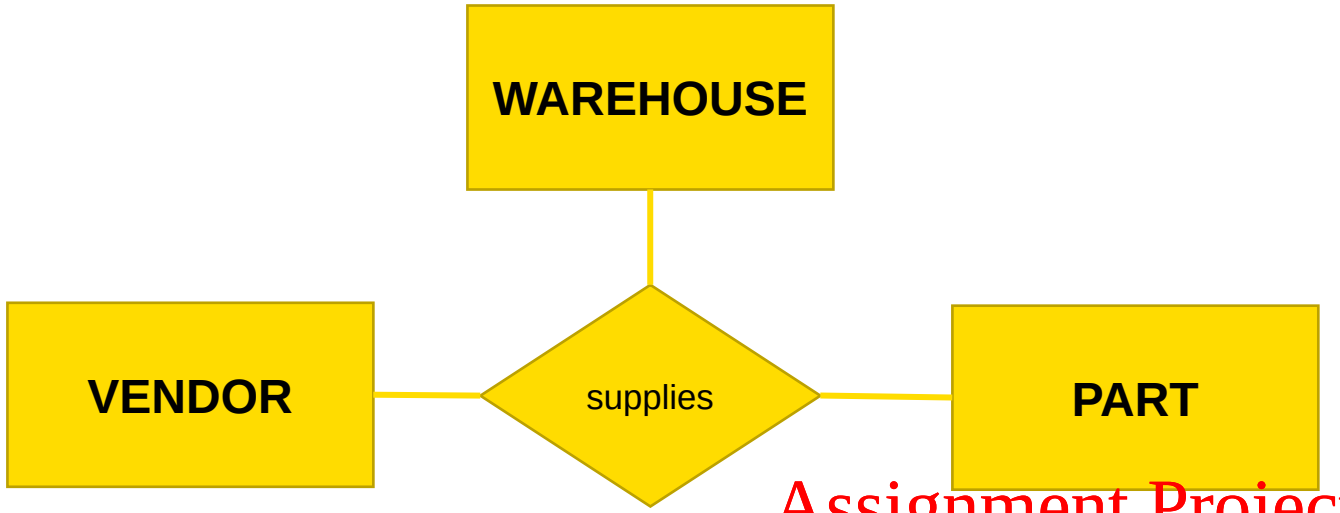
Entity Relationships

Entity 1 Relationshi

Binary Relationship (2 entities)

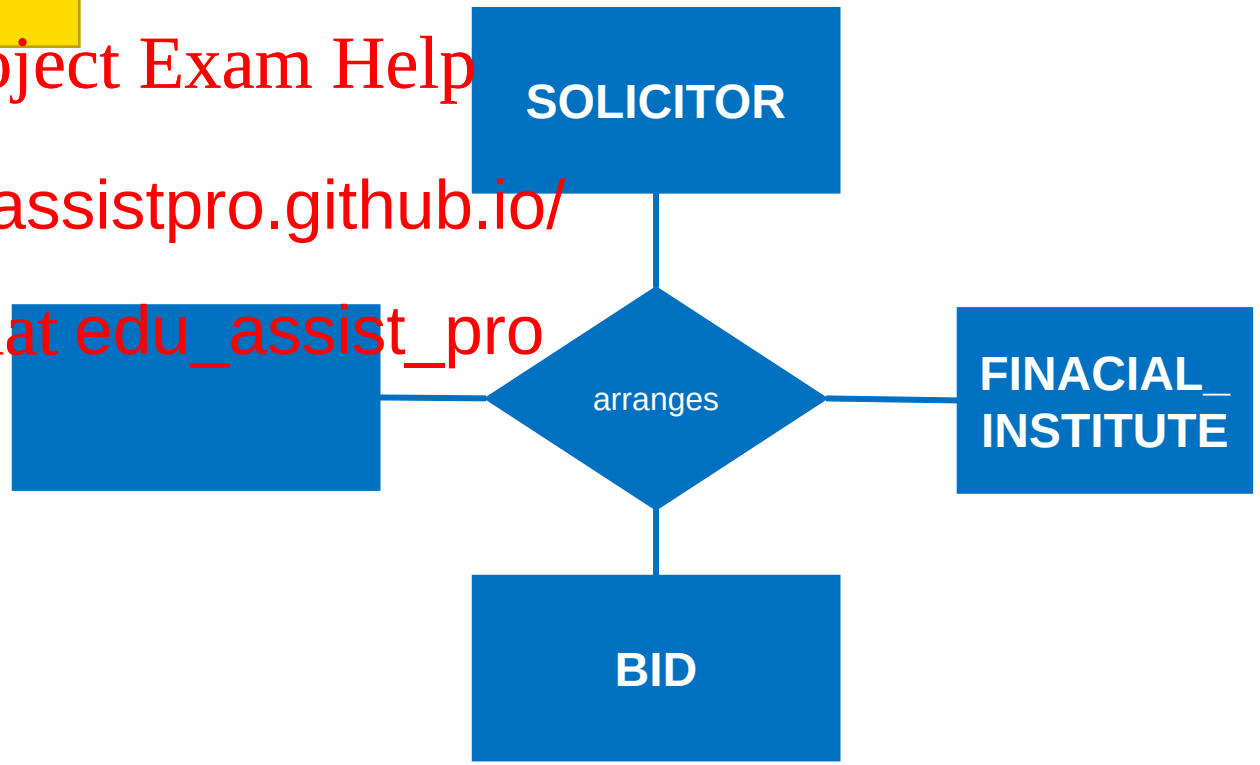


Unary Relationship (1 entity)



Ternary Relationship (3 e

Quaternary Relationship (4 entities)



Connectivity

- **Connectivity** is used to describe the **relationship classification**.
- The ER diagram indi using a numeric notation.

Basic Relationship (One-to-One 1:1)

State		Capital City
NSW	----------	Sydney
Victoria	----------	Melbourne

Basic Relationship (One-to-Many 1:M)

- A movie (e.g., Avengers) can be stocked a ys (e.g., 30 copies)
- All blurays contain a film.
- There is “one-to-many” relationship between film and bluray.

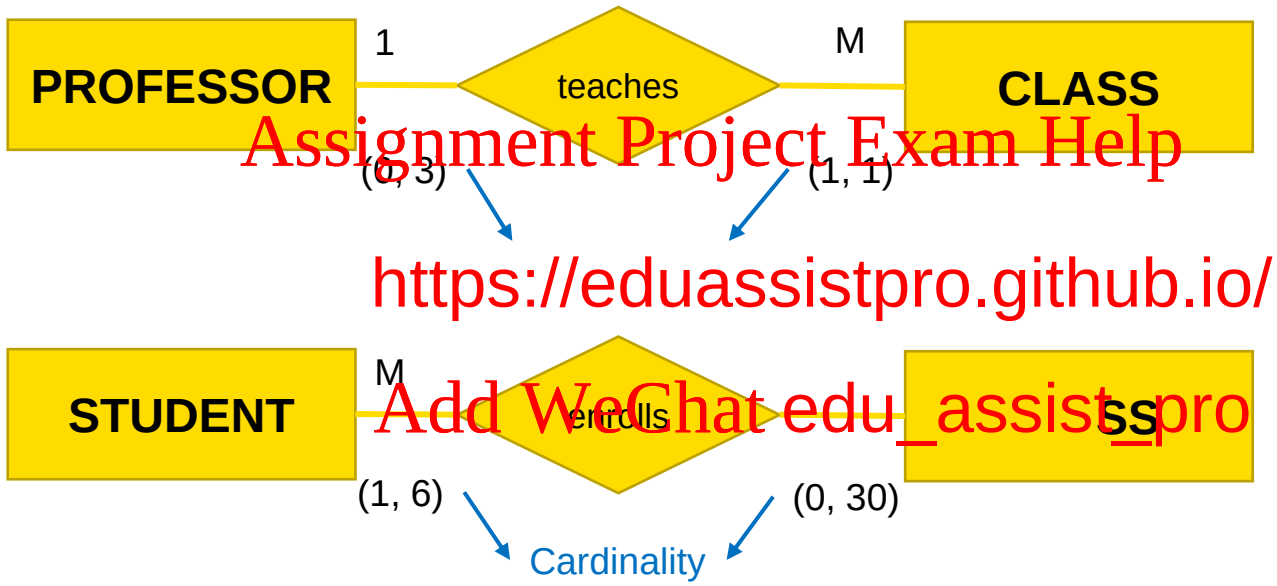
Basic Relationship (Many-to-Many M:N)

- A student enrolls many courses.
- Each course is enrolled by many students.

Cardinality

- **Cardinality** expresses the **specific number of entity occurrences associated with one occurrence of related entity.**
- A **cardinality constraint** specifies the number of instances of entity A that can be associated with each instance of entity B. Car rom business rules.
- **Business rules**: They are derived from organisati onment.
- **Minimum cardinality** is the minimum number of instances of one entity that may be associated with each instance of another entity.
- **Maximum cardinality** is the maximum number of instances of one entity that may be associated with each instance of another entity.

Examples of Cardinality



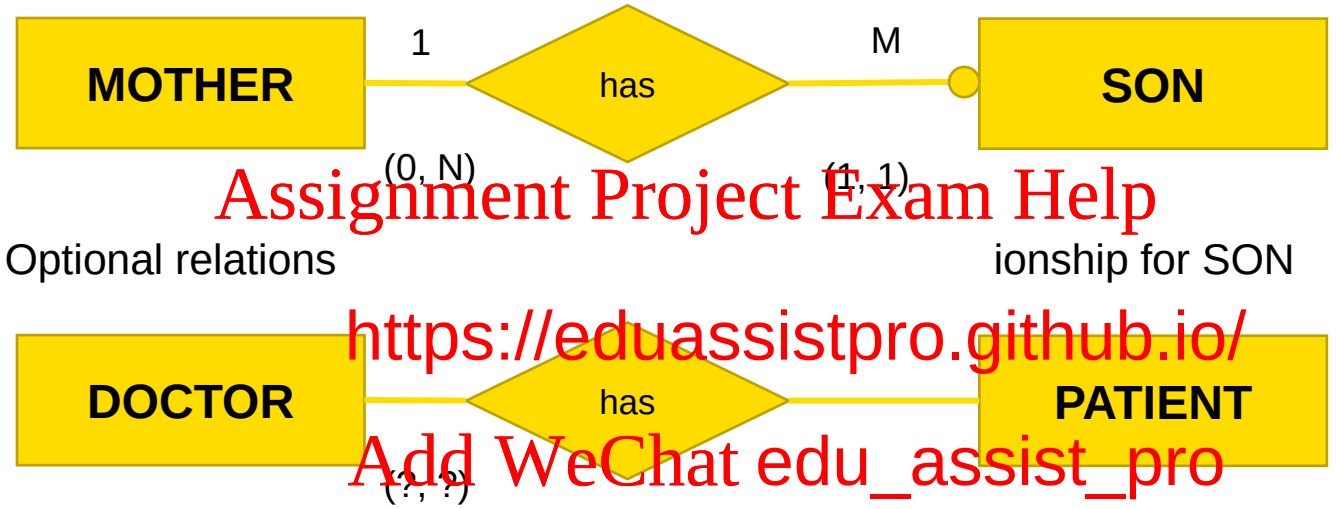
One-to-Many Relationship

Many-to-Many Relationship

How to read this?

- A professor teaches (0, 3) classes. A class is taught by (1, 1) professors.
- A student enrolls in (1, 6) classes. A class has enrolled in it (0, 30) students.

Relationship Participation



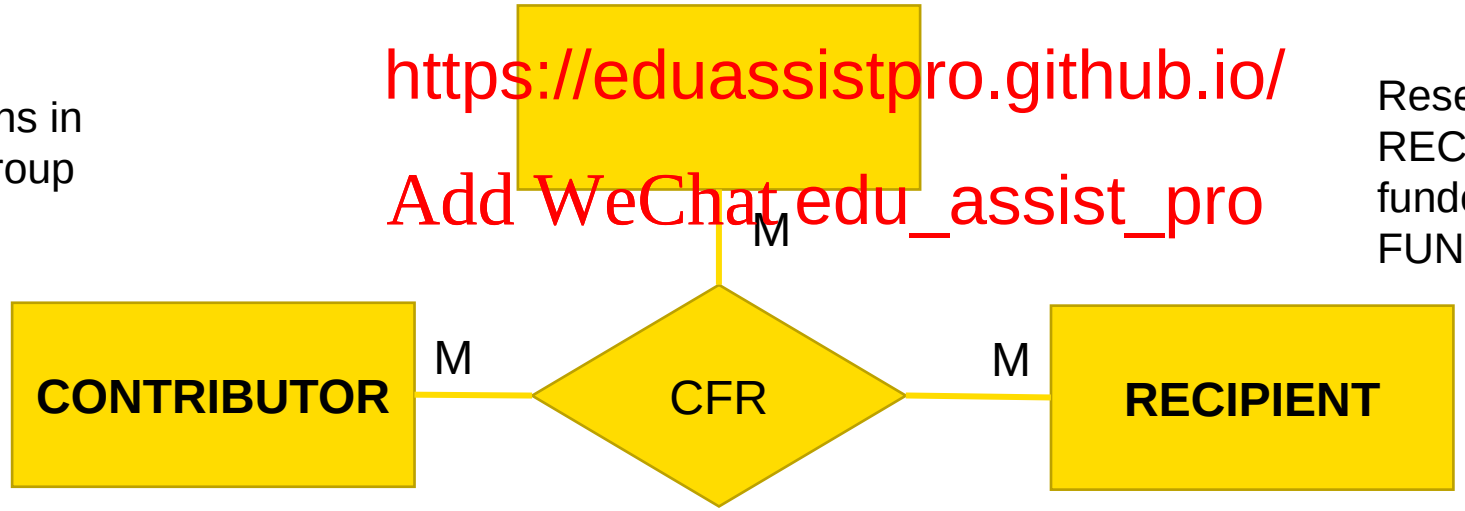
Mandatory relationship between DOCTOR & PATIENT

- A participating entity in a relationship can be either **optional** or **mandatory**.
- Determined by the **specific meaning of the terms used**.
  - ❑ Depends on context.
  - ❑ Need to state assumptions.

Ternary Relationships

Research fund may be placed into several categories.

People or institutions in CONTRIBUTOR group donate money to a special research FUND.



Researchers found in RECIPIENT are funded through the FUND contents.

Relationship, Connectivity, Cardinality

- **Relationship:** Association between entities that always operate in both directions
  tables
- **Participants**: Entities
  - The most common relati
    tionship
    rnary, and quaternary.
- **Connectivity**: Describes the relationship classification
  1:1, 1:M and M:N
- **Cardinality**: Expresses the minimum and maximum number of entity occurrences associated with one occurrence of related entity
  e.g., how many classes at most one professor can teach.

Recap: ER Modelling Part 1

- **Data Modelling**
  - ❑ Data model as a (relatively) simple abstraction of the complex real-world (for the purpose of creating a DB).
  - ❑ A good DBMS will perform atabase.
  - ❑ One modelling technique t **ationship Modelling**
- **Entity Relationship Mod**
  - ❑ Entity Types and Entity Instances
  - ❑ Attributes and Values
  - ❑ Keys
  - ❑ Relationships
  - ❑ Connectivity
  - ❑ Cardinality

Questions

Source: tryinteract.com