

*W4111 – Introduction to Databases
Section 002, Fall 2023*

Lecture 7: ER, Relational, SQL (VI)



Contents

Contents

- Codd's Rules:
 - Overview
 - NULL examples
 - Metadata
- ER design patterns and advanced concepts.
- Applications: REST and web applications.

Codd's Rules Discussion

Codd's 12 Rules

Rule 1: Information Rule

The data stored in a database, may it be user data or metadata, must be a value of some table cell. Everything in a database must be stored in a table format.

Rule 2: Guaranteed Access Rule

URL Like

Every single data element (value) is guaranteed to be accessible logically with a combination of table-name, primary-key (row value), and attribute-name (column value).

No other means, such as pointers, can be used to access data.

Rule 3: Systematic Treatment of NULL Values

The NULL values in a database must be given a systematic and uniform treatment. This is a very important rule because a NULL can be interpreted as one the following – data is missing, data is not known, or data is not applicable.

Rule 4: Active Online Catalog

The structure description of the entire database must be stored in an online catalog, known as data dictionary, which can be accessed by authorized users. Users can use the same query language to access the catalog which they use to access the database itself.

Rule 5: Comprehensive Data Sub-Language Rule

A database can only be accessed using a language having linear syntax that supports data definition, data manipulation, and transaction management operations. This language can be used directly or by means of some application. If the database allows access to data without any help of this language, then it is considered as a violation.

Rule 6: View Updating Rule

All the views of a database, which can theoretically be updated, must also be updatable by the system.

Codd's 12 Rules

Rule 7: High-Level Insert, Update, and Delete Rule

A database must support high-level insertion, updation, and deletion. This must not be limited to a single row, that is, it must also support union, intersection and minus operations to yield sets of data records.

Rule 8: Physical Data Independence

The data stored in a database must be independent of the applications that access the database. Any change in the physical structure of a database must not have any impact on how the data is being accessed by external applications.

Rule 9: Logical Data Independence

The logical data in a database must be independent of its user's view (application). Any change in logical data must not affect the applications using it. **For example, if two tables are merged or one is split into two different tables, there should be no impact or change on the user application. This is one of the most difficult rule to apply.**

Rule 10: Integrity Independence

A database must be independent of the application that uses it. All its integrity constraints can be independently modified without the need of any change in the application. This rule makes a database independent of the front-end application and its interface.

Rule 11: Distribution Independence

The end-user must not be able to see that the data is distributed over various locations. Users should always get the impression that the data is located at one site only. This rule has been regarded as the foundation of distributed database systems.

Rule 12: Non-Subversion Rule

If a system has an interface that provides access to low-level records, then the interface must not be able to subvert the system and bypass security and integrity constraints.

ER Modeling

Design Patterns and Advanced Concepts

Model and Concepts



Weak Entity Sets

- Consider a *section* entity, which is uniquely identified by a *course_id*, *semester*, *year*, and *sec_id*.
- Clearly, section entities are related to course entities. Suppose we create a relationship set *sec_course* between entity sets *section* and *course*.
- Note that the information in *sec_course* is redundant, since *section* already has an attribute *course_id*, which identifies the course with which the section is related.
- One option to deal with this redundancy is to get rid of the relationship *sec_course*; however, by doing so the relationship between *section* and *course* becomes implicit in an attribute, which is not desirable.



Weak Entity Sets (Cont.)

- An alternative way to deal with this redundancy is to not store the attribute *course_id* in the *section* entity and to only store the remaining attributes *section_id*, *year*, and *semester*.
 - However, the entity set *section* then does not have enough attributes to identify a particular *section* entity uniquely
- To deal with this problem, we treat the relationship *sec_course* as a special relationship that provides extra information, in this case, the *course_id*, required to identify *section* entities uniquely.
- A **weak entity set** is one whose existence is dependent on another entity, called its **identifying entity**
- Instead of associating a primary key with a weak entity, we use the identifying entity, along with extra attributes called **discriminator** to uniquely identify a weak entity.



Weak Entity Sets (Cont.)

- An entity set that is not a weak entity set is termed a **strong entity set**.
- Every weak entity must be associated with an identifying entity; that is, the weak entity set is said to be **existence dependent** on the identifying entity set.
- The identifying entity set is said to **own** the weak entity set that it identifies.
- The relationship associating the weak entity set with the identifying entity set is called the **identifying relationship**.
- Note that the relational schema we eventually create from the entity set *section* does have the attribute *course_id*, for reasons that will become clear later, even though we have dropped the attribute *course_id* from the entity set *section*.

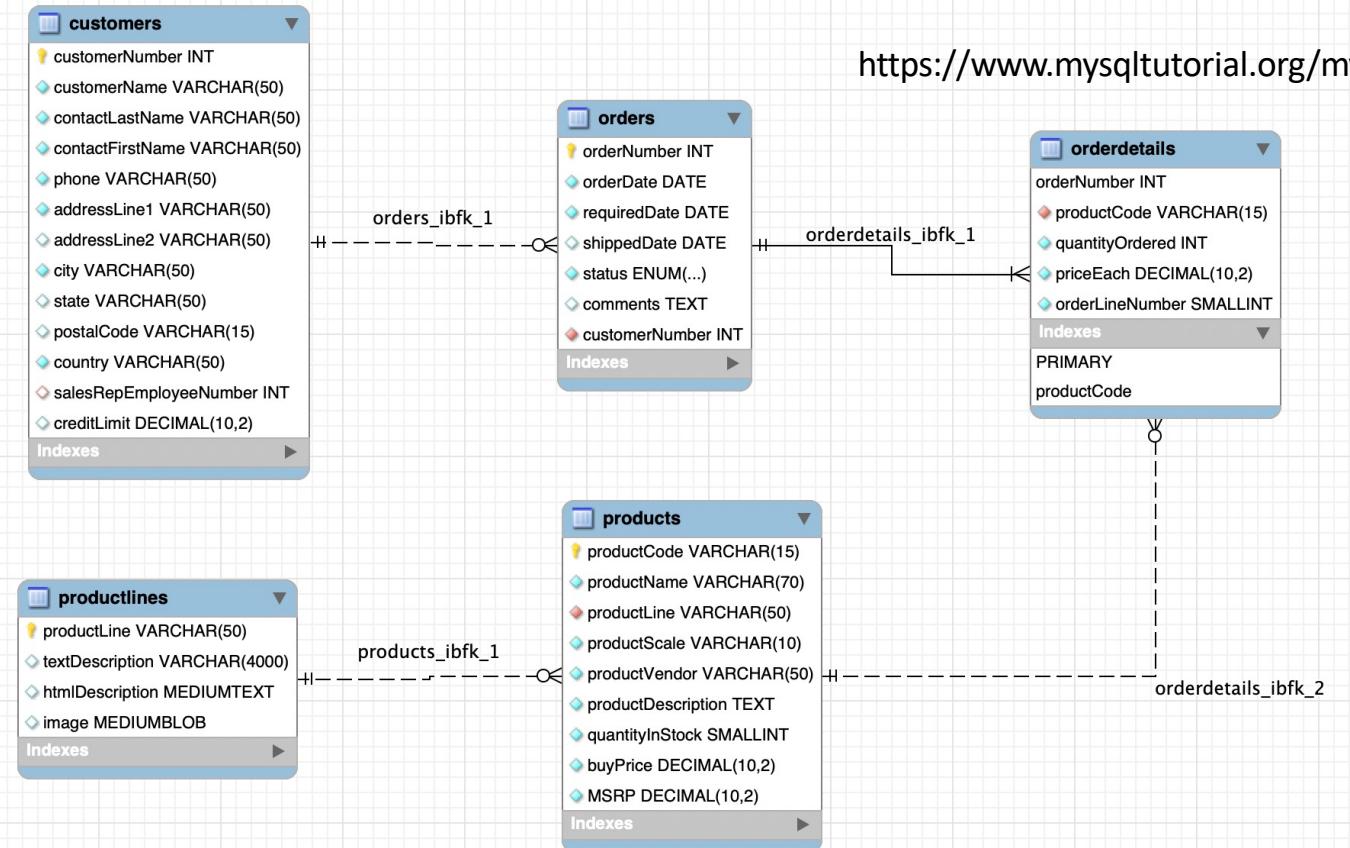


Expressing Weak Entity Sets

- In E-R diagrams, a weak entity set is depicted via a double rectangle.
- We underline the discriminator of a weak entity set with a dashed line.
- The relationship set connecting the weak entity set to the identifying strong entity set is depicted by a double diamond.
- Primary key for *section* – (*course_id*, *sec_id*, *semester*, *year*)



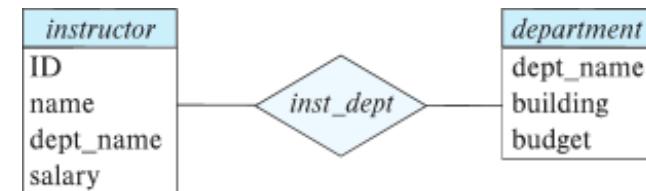
An Example – Classic Models



<https://www.mysqltutorial.org/mysql-sample-database.aspx/>

Redundant Attributes

- Suppose we have entity sets:
 - *instructor*, with attributes: *ID, name, dept_name, salary*
 - *department*, with attributes: *dept_name, building, budget*
- We model the fact that each instructor has an associated department using a relationship set *inst_dept*
- The attribute *dept_name* in *instructor* replicates information present in the relationship and is therefore redundant
 - and needs to be removed.
- BUT: when converting back to tables, in some cases the attribute gets reintroduced, as we will see later.





Design Alternatives

- In designing a database schema, we must ensure that we avoid two major pitfalls:
 - Redundancy: a bad design may result in repeat information.
 - **Redundant representation of information may lead to data inconsistency among the various copies of information**
 - Incompleteness: a bad design may make certain aspects of the enterprise difficult or impossible to model.
- Avoiding bad designs is not enough. There may be a large number of good designs from which we must choose.

**Emphasis
Added**



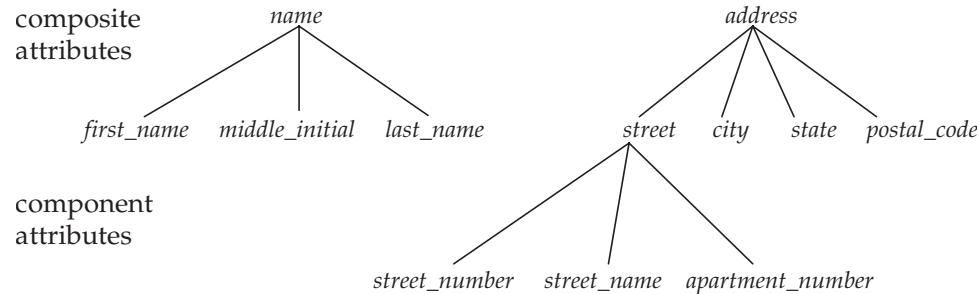
Complex Attributes

- Attribute types:
 - **Simple** and **composite** attributes.
 - **Single-valued** and **multivalued** attributes
 - Example: multivalued attribute: *phone_numbers*
 - **Derived** attributes
 - Can be computed from other attributes
 - Example: age, given date_of_birth
- **Domain** – the set of permitted values for each attribute



Composite Attributes

- Composite attributes allow us to divide attributes into subparts (other attributes).



Example from IMDB

- Consider name_basics

	nconst	primaryName	birthYear	deathYear	primaryProfession	knownForTitles
1	nm0000001	Fred Astaire	1899	1987	soundtrack,actor,miscellaneous	tt0050419,tt0031983,tt0072308,tt0053137
2	nm0000002	Lauren Bacall	1924	2014	actress,soundtrack	tt0071877,tt0117057,tt0037382,tt0038355
3	nm0000003	Brigitte Bardot	1934	<null>	actress,soundtrack,music_department	tt0049189,tt0056404,tt0057345,tt0054452
4	nm0000004	John Belushi	1949	1982	actor,soundtrack,writer	tt0077975,tt0072562,tt0080455,tt0078723
5	nm0000005	Ingmar Bergman	1918	2007	writer,director,actor	tt0050986,tt0060827,tt0069467,tt0050976
6	nm0000006	Ingrid Bergman	1915	1982	actress,soundtrack,producer	tt0077711,tt0038109,tt0034583,tt0036855
7	nm0000007	Humphrey Bogart	1899	1957	actor,soundtrack,producer	tt0043265,tt0034583,tt0042593,tt0037382
8	nm0000008	Marlon Brando	1924	2004	actor,soundtrack,director	tt0078788,tt0068646,tt0070849,tt0047296
9	nm0000009	Richard Burton	1925	1984	actor,soundtrack,producer	tt0061184,tt0087803,tt0057877,tt0059749
10	nm0000010	James Cagney	1899	1986	actor,soundtrack,director	tt0029870,tt0035575,tt0042041,tt0055256

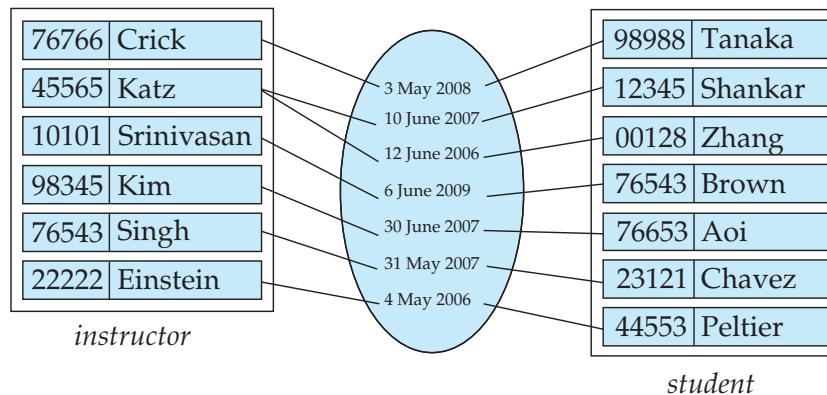
- There
 - Is one composite attribute, primaryName.
 - Are two multivalued attributes: primaryProfession, knownForTitles
 - knownForTitles is also tricky, which we will see.
 - Names are also a little tricky

Switch to notebook



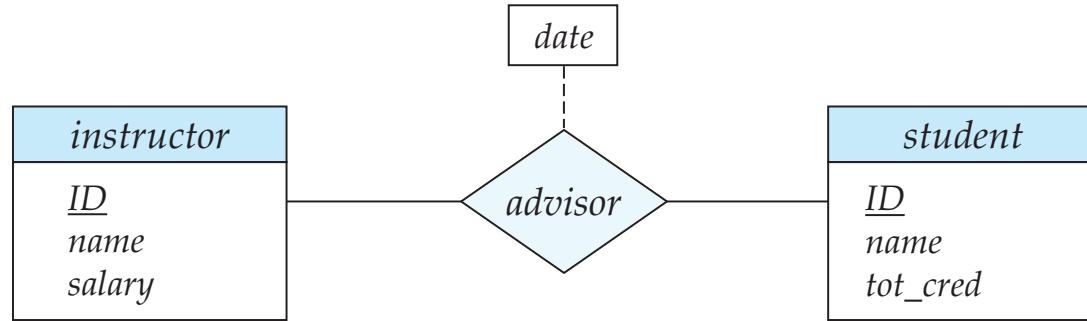
Relationship Sets (Cont.)

- An attribute can also be associated with a relationship set.
- For instance, the *advisor* relationship set between entity sets *instructor* and *student* may have the attribute *date* which tracks when the student started being associated with the advisor





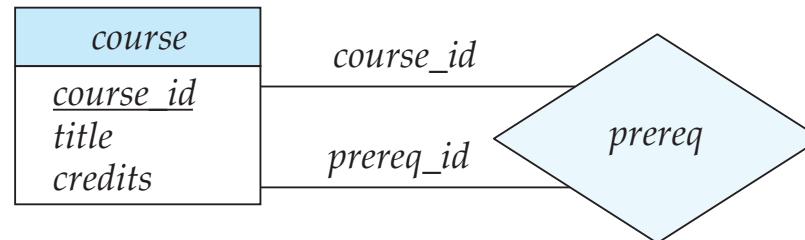
Relationship Sets with Attributes





Roles

- Entity sets of a relationship need not be distinct
 - Each occurrence of an entity set plays a “role” in the relationship
- The labels “*course_id*” and “*prereq_id*” are called **roles**.





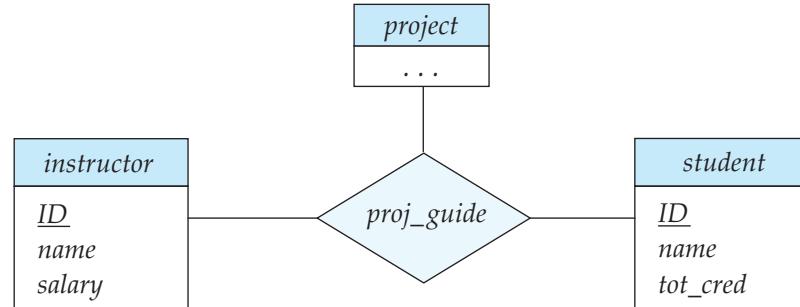
Degree of a Relationship Set

- Binary relationship
 - involve two entity sets (or degree two).
 - most relationship sets in a database system are binary.
- Relationships between more than two entity sets are rare. Most relationships are binary. (More on this later.)
 - Example: *students* work on research *projects* under the guidance of an *instructor*.
 - relationship *proj_guide* is a ternary relationship between *instructor*, *student*, and *project*



Non-binary Relationship Sets

- Most relationship sets are binary
- There are occasions when it is more convenient to represent relationships as non-binary.
- E-R Diagram with a Ternary Relationship





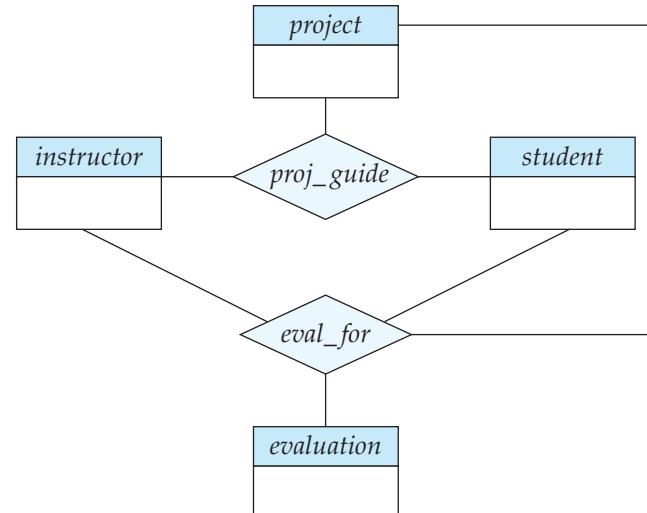
Mapping Cardinality Constraints

- Express the number of entities to which another entity can be associated via a relationship set.
- Most useful in describing binary relationship sets.
- For a binary relationship set the mapping cardinality must be one of the following types:
 - One to one
 - One to many
 - Many to one
 - Many to many



Aggregation

- Consider the ternary relationship *proj_guide*, which we saw earlier
- Suppose we want to record evaluations of a student by a guide on a project





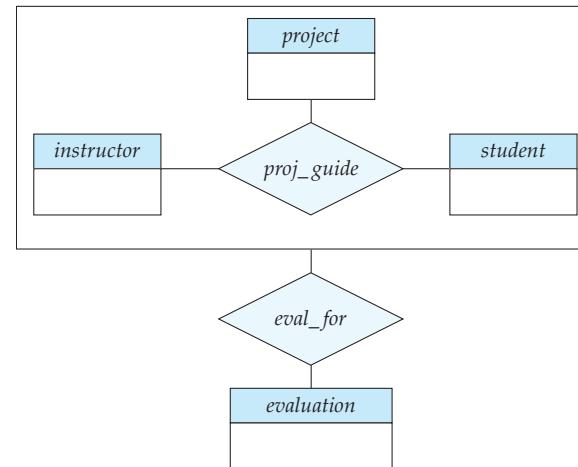
Aggregation (Cont.)

- Relationship sets *eval_for* and *proj_guide* represent overlapping information
 - Every *eval_for* relationship corresponds to a *proj_guide* relationship
 - However, some *proj_guide* relationships may not correspond to any *eval_for* relationships
 - So we can't discard the *proj_guide* relationship
- Eliminate this redundancy via *aggregation*
 - Treat relationship as an abstract entity
 - Allows relationships between relationships
 - Abstraction of relationship into new entity



Aggregation (Cont.)

- Eliminate this redundancy via *aggregation* without introducing redundancy, the following diagram represents:
 - A student is guided by a particular instructor on a particular project
 - A student, instructor, project combination may have an associated evaluation





Reduction to Relational Schemas

- To represent aggregation, create a schema containing
 - Primary key of the aggregated relationship,
 - The primary key of the associated entity set
 - Any descriptive attributes
- In our example:
 - The schema *eval_for* is:
$$\text{eval_for} (s_ID, project_id, i_ID, evaluation_id)$$
 - The schema *proj_guide* is redundant.



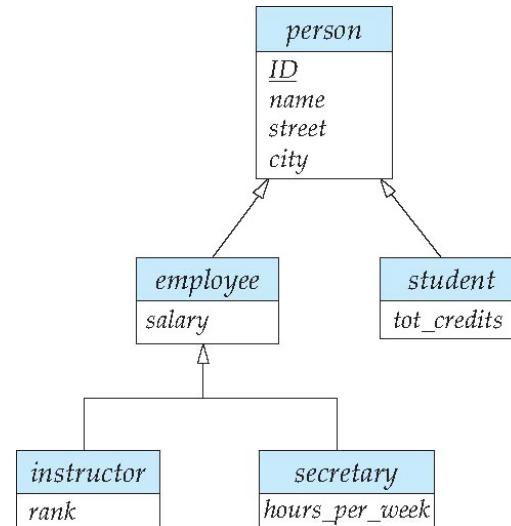
Specialization

- Top-down design process; we designate sub-groupings within an entity set that are distinctive from other entities in the set.
- These sub-groupings become lower-level entity sets that have attributes or participate in relationships that do not apply to the higher-level entity set.
- Depicted by a *triangle* component labeled ISA (e.g., *instructor* “is a” *person*).
- **Attribute inheritance** – a lower-level entity set inherits all the attributes and relationship participation of the higher-level entity set to which it is linked.



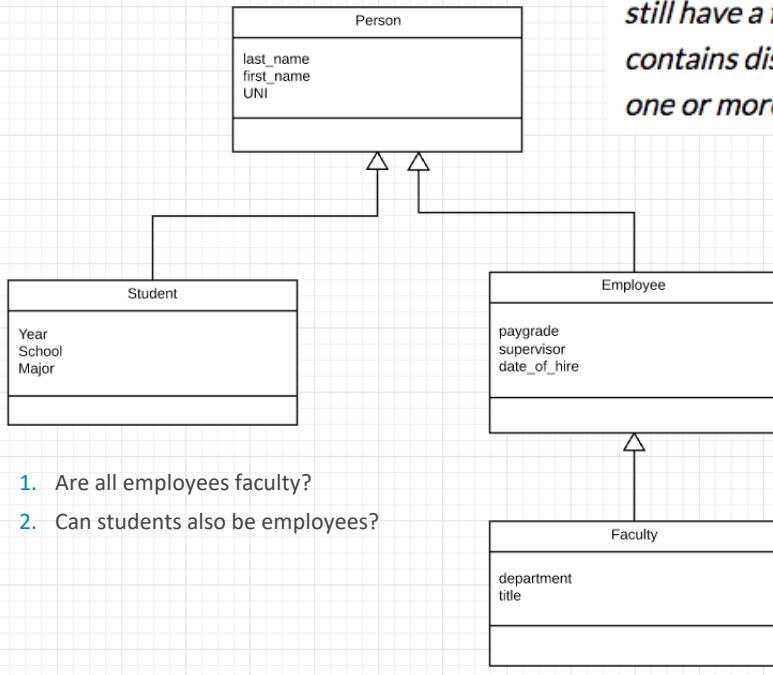
Specialization Example

- **Overlapping** – *employee* and *student*
- **Disjoint** – *instructor* and *secretary*
- Total and partial



Inheritance/Specialization

In the process of designing our entity relationship diagram for a database, we may find that attributes of two or more entities overlap, meaning that these entities seem very similar but still have a few differences. In this case, we may create a subtype of the parent entity that contains distinct attributes. A parent entity becomes a supertype that has a relationship with one or more subtypes.



1. Are all employees faculty?
2. Can students also be employees?

The subclass association line is labeled with specialization constraints. Constraints are described along two dimensions:

1 incomplete/complete

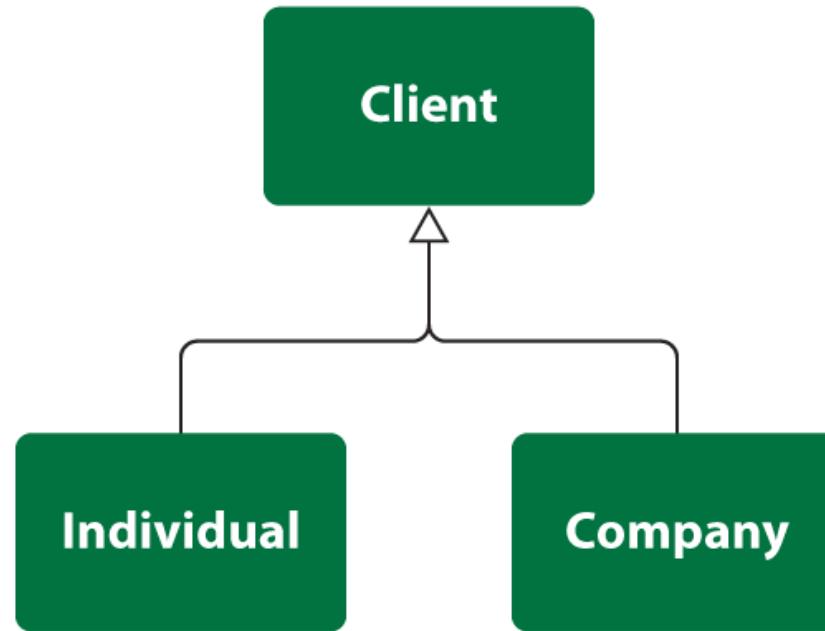
- In an **incomplete** specialization only some instances of the parent class are specialized (have unique attributes). Other instances of the parent class have only the common attributes.
- In a **complete** specialization, every instance of the parent class has one or more unique attributes that are not common to the parent class.

2 disjoint/overlapping

- In a **disjoint** specialization, an object could be a member of only one specialized subclass.
- In an **overlapping** specialization, an object could be a member of more than one specialized subclass.

Specialization

In class Client we distinguish two subtypes: Individual and Company. This specialization is disjoint (client can be an individual or a company) and complete (these are all possible subtypes for supertype).

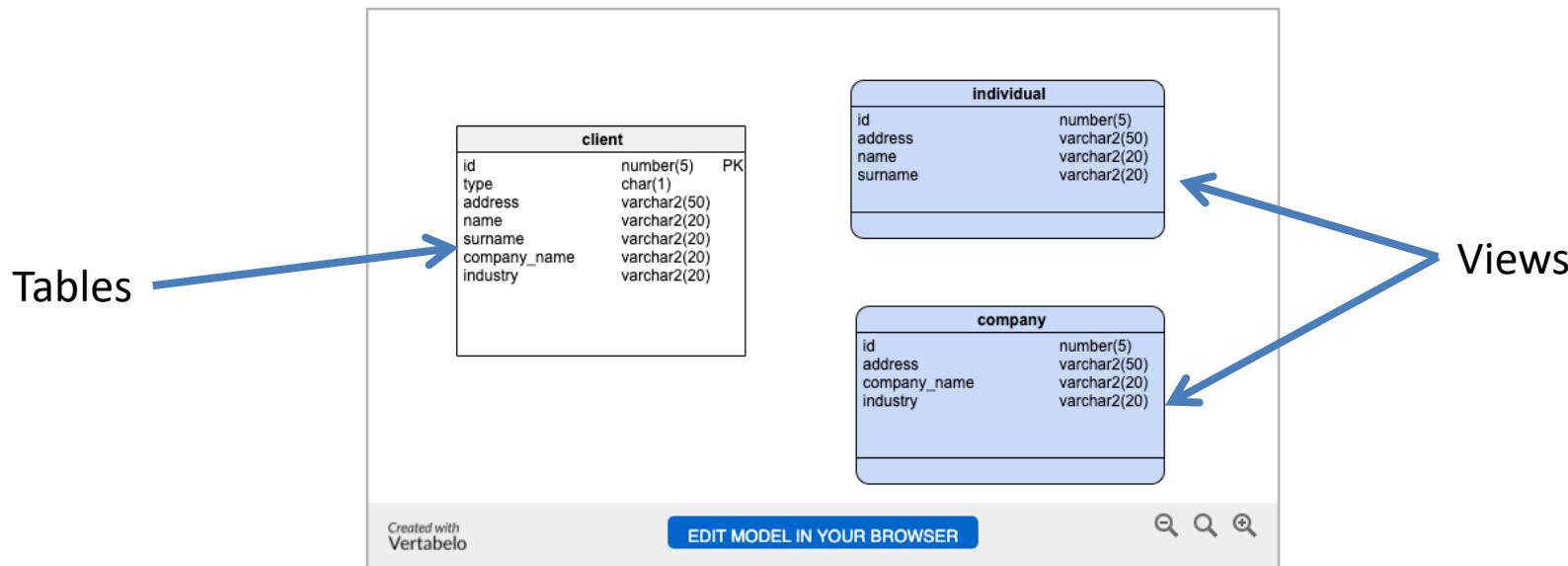


One Table

One table implementation

In a one table implementation, table `client` has attributes of both types.

The diagram below shows the table `client` and two views: `individual` and `company`:

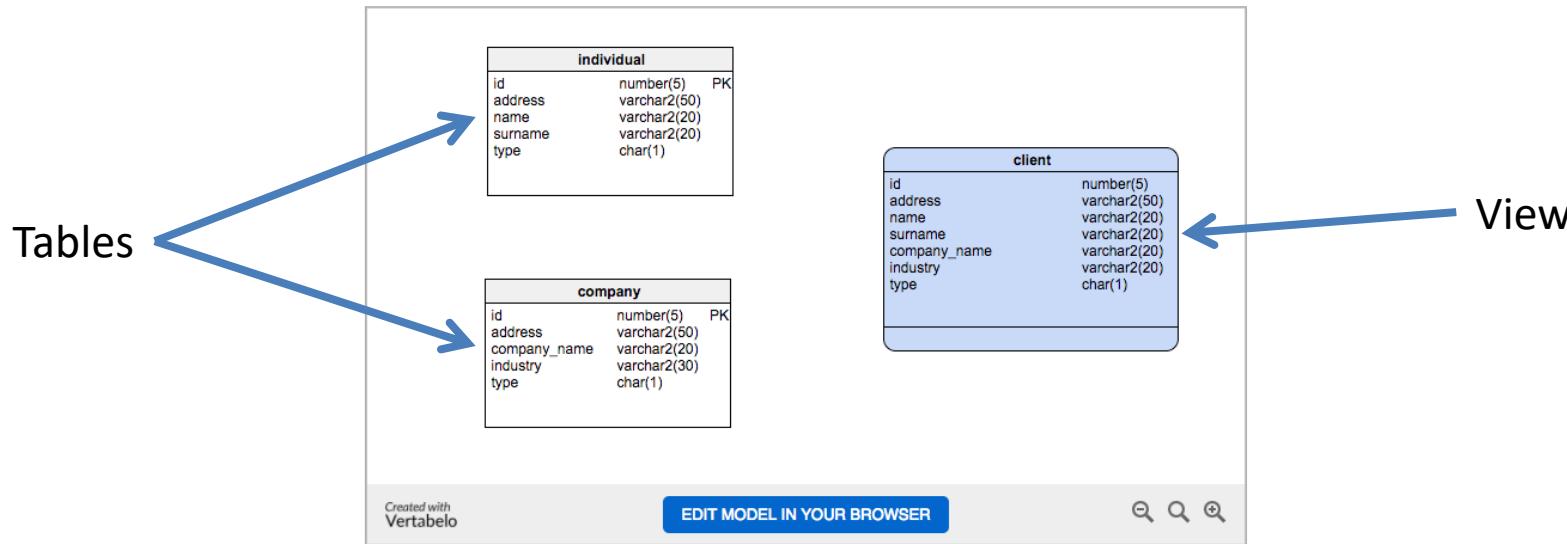


Two Table

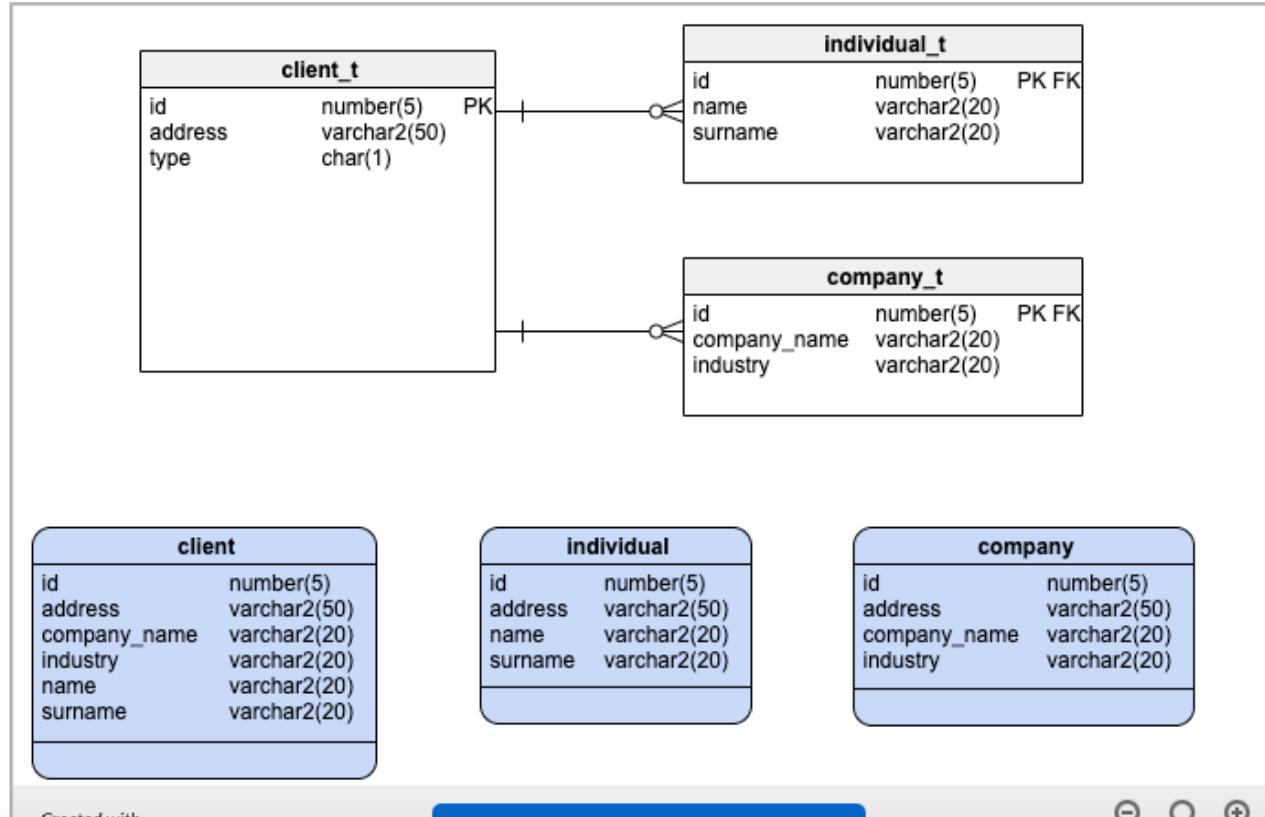
Two-table implementation

In a two-table implementation, we create a table for each of the subtypes. Each table gets a column for all attributes of the supertype and also a column for each attribute belonging to the subtype. Access to information in this situation is limited, that's why it is important to create a view that is the union of the tables. We can add an additional attribute called 'type' that describes the subtype.

The diagram below presents two tables, `individual` and `company`, and a view (the blue one) called `client`.

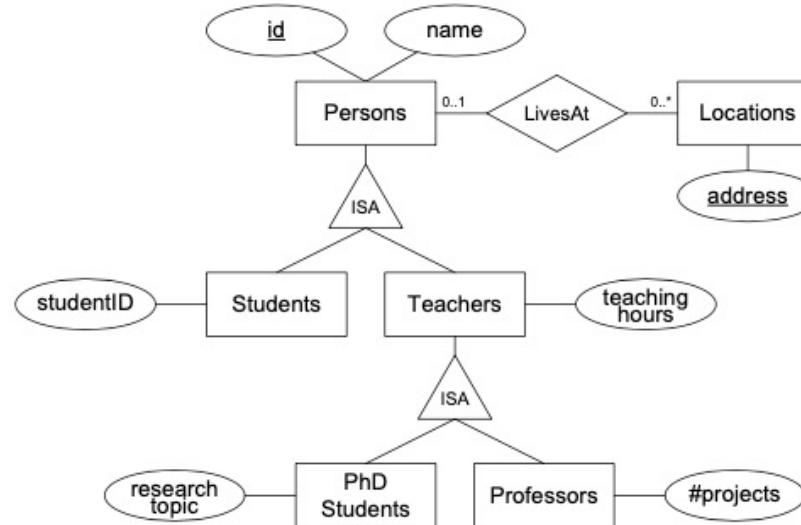


Three Table





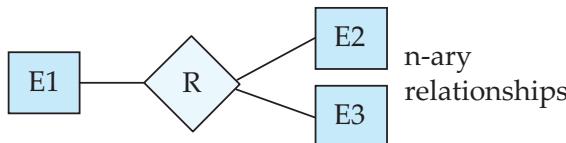
ISA Relationship





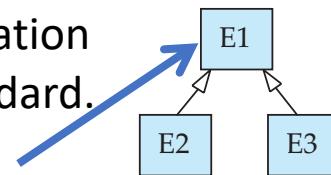
ER vs. UML Class Diagrams

ER Diagram Notation

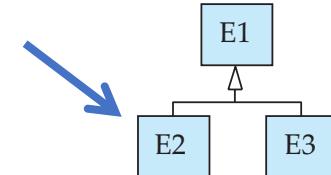


n-ary
relationships

I use this approach
in Crow's Foot Notation
but that is not standard.

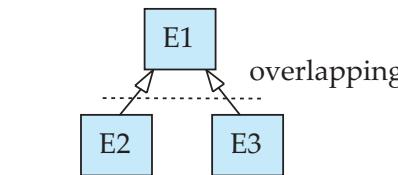
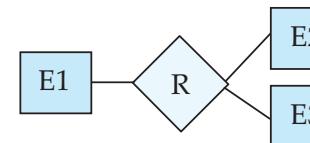


overlapping
generalization

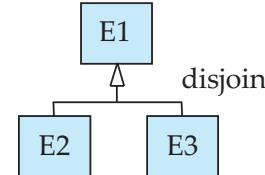


disjoint
generalization

Equivalent in UML



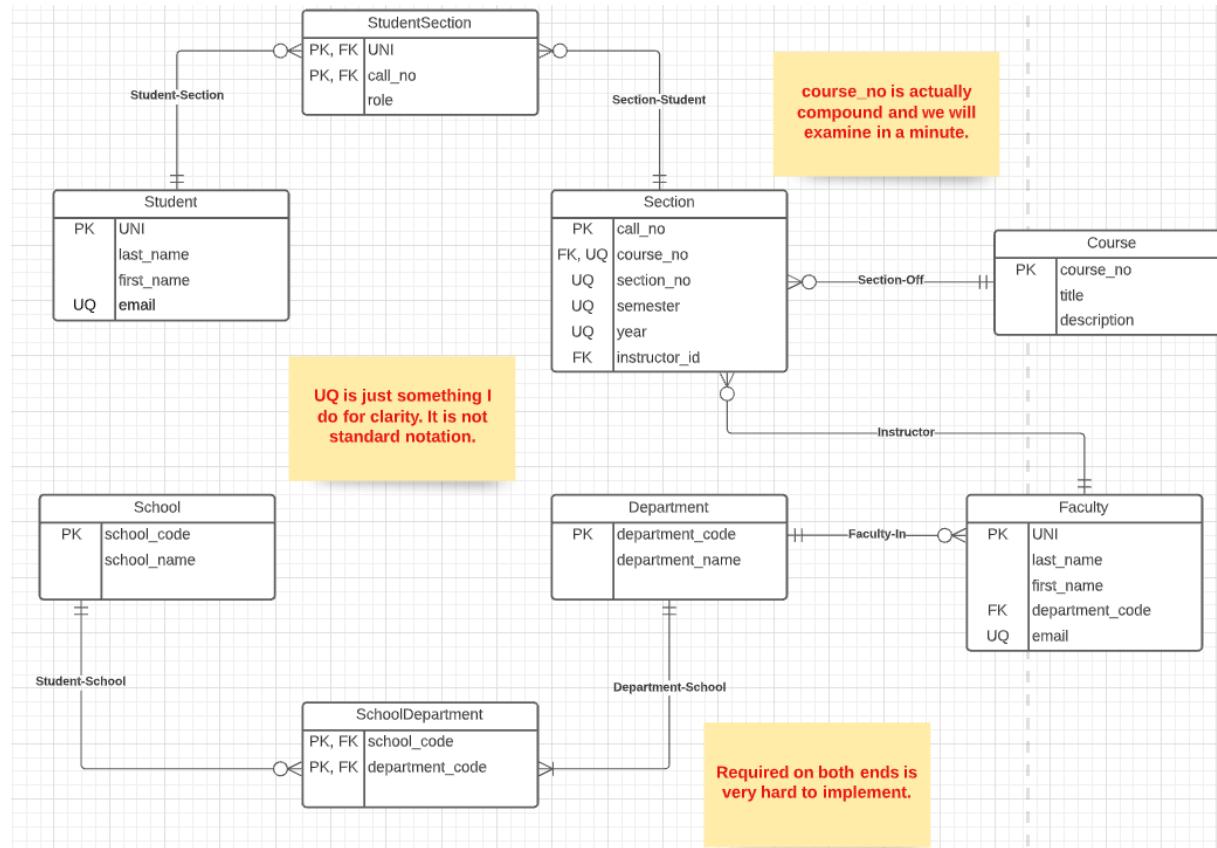
overlapping



disjoint

- * Generalization can use merged or separate arrows independent of disjoint/overlapping

Approximate Model



Course Number

Key to Columbia Course Listings

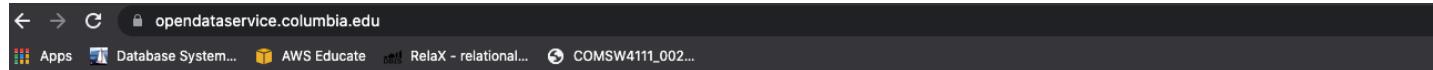
<https://www.cc-seas.columbia.edu/sites/dsa/files/handbooks/Columbia%20Key%20to%20Course%20Listing.pdf>

(Example: ECON W1105 001 Principles of Economics, 4 pts)

A	Architecture, Planning, and Preservation*
B	Business*
BC	Barnard College
C	Columbia College
D	Dentistry**
E	Engineering and Applied Science
F	General Studies
G	Graduate School of Arts and Sciences
H	Reid Hall, Paris**
I	Berlin Consortium Program**
J	Journalism*
K	Continuing Education**
L	Law**
M	Medicine**
N	Nursing**
O	Union Theological**
P	School of Public Health*
R	School of the Arts*
S	Summer Session
T	Social Work*
TA-TZ	Teachers College*
U	International and Public Affairs*
V	Interschool course with Barnard
W	Interfaculty course
X	Barnard College
Z	American Language Program(no credit)**

	Example	Description																																
Call # (5 digit number)	16238	This 5 digit code is assigned to individual courses and is specific to each semester.																																
Department Code (4 letter code)	ECON	This 4 letter code represents the Academic Department that manages the course.																																
Course Number (1 capital letter followed by 4 digit code)	W 1105 001 4	<p>The <i>capital letters</i> indicate the instructor teaching the course and their affiliation with a division, school or affiliate of the University.</p> <p>Unless otherwise noted, courses numbers beginning with the following letters are generally open to CC/SEAS undergraduate students:</p> <table border="1"> <tbody> <tr><td>BC</td><td>Barnard College</td></tr> <tr><td>C</td><td>Columbia College</td></tr> <tr><td>E</td><td>Engineering and Applied Science</td></tr> <tr><td>F</td><td>General Studies</td></tr> <tr><td>G</td><td>Graduate School of Arts and Sciences</td></tr> <tr><td>V</td><td>Interschool course with Barnard</td></tr> <tr><td>W</td><td>Interfaculty course</td></tr> <tr><td>X</td><td>Barnard College</td></tr> </tbody> </table> <p>The first <i>digit</i> indicates the level of the course. Generally, levels are indicated as:</p> <table border="1"> <tbody> <tr><td>0</td><td>Course that cannot be credited toward any degree</td></tr> <tr><td>1</td><td>Undergraduate course, introductory</td></tr> <tr><td>2</td><td>Undergraduate course, intermediate</td></tr> <tr><td>3</td><td>Undergraduate course, advanced</td></tr> <tr><td>4</td><td>Graduate course that is open to qualified undergraduates</td></tr> <tr><td>6</td><td>Graduate course</td></tr> <tr><td>8</td><td>Graduate course, advanced</td></tr> <tr><td>9</td><td>Graduate research course or seminar</td></tr> </tbody> </table>	BC	Barnard College	C	Columbia College	E	Engineering and Applied Science	F	General Studies	G	Graduate School of Arts and Sciences	V	Interschool course with Barnard	W	Interfaculty course	X	Barnard College	0	Course that cannot be credited toward any degree	1	Undergraduate course, introductory	2	Undergraduate course, intermediate	3	Undergraduate course, advanced	4	Graduate course that is open to qualified undergraduates	6	Graduate course	8	Graduate course, advanced	9	Graduate research course or seminar
BC	Barnard College																																	
C	Columbia College																																	
E	Engineering and Applied Science																																	
F	General Studies																																	
G	Graduate School of Arts and Sciences																																	
V	Interschool course with Barnard																																	
W	Interfaculty course																																	
X	Barnard College																																	
0	Course that cannot be credited toward any degree																																	
1	Undergraduate course, introductory																																	
2	Undergraduate course, intermediate																																	
3	Undergraduate course, advanced																																	
4	Graduate course that is open to qualified undergraduates																																	
6	Graduate course																																	
8	Graduate course, advanced																																	
9	Graduate research course or seminar																																	
Course Section	001	Based on course demand, some academic departments offer the same course during 2 or more different time slots. Each time slot is assigned a different section number.																																
Points/Credits	4	The term "points" and "credits" are often used interchangeably and is generally related to the number of classroom contact hours.																																

Columbia Open Data Service



COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

OPEN DATA SERVICE

[News and Updated](#) [About Data Feeds](#) [Request a New Feed](#)

Data Feed Service

Columbia University offers data feeds in programming-friendly formats for research and academic purposes. The Open Data Feed Service currently offers the following data feeds:

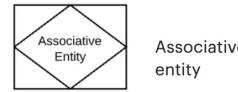
- [Course Information](#)
- [Athletic Schedule](#)
- [Academic Commons](#)
- [CLIO - Library Catalog Data](#)
- [Textbooks Feed](#)

Data feed information includes the feed's refresh schedule, data diagrams, table and data element documentation, data training, and data security. More details on each feed is available on each feed's page.



Associative Entity

- The ER model represents “associations/relationships” as
 - First class “things”
 - That are different from “entities.”
- The SQL model does not have “relationships” or associations as first class types. You have
 - Tables
 - Columns
 - Keys
 - Constraints
 -
- You can implement some “relationships” using foreign keys. Others require something more complex – an *associative entity*.

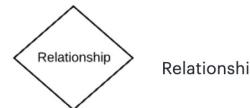


Associative entities relate the instances of several entity types. They also contain attributes specific to the relationship between those entity instances.

ERD relationship symbols

Within entity-relationship diagrams, relationships are used to document the interaction between two entities. Relationships are usually verbs such as assign, associate, or track and provide useful information that could not be discerned with just the entity types.

Relationship Symbol	Name	Description
---------------------	------	-------------



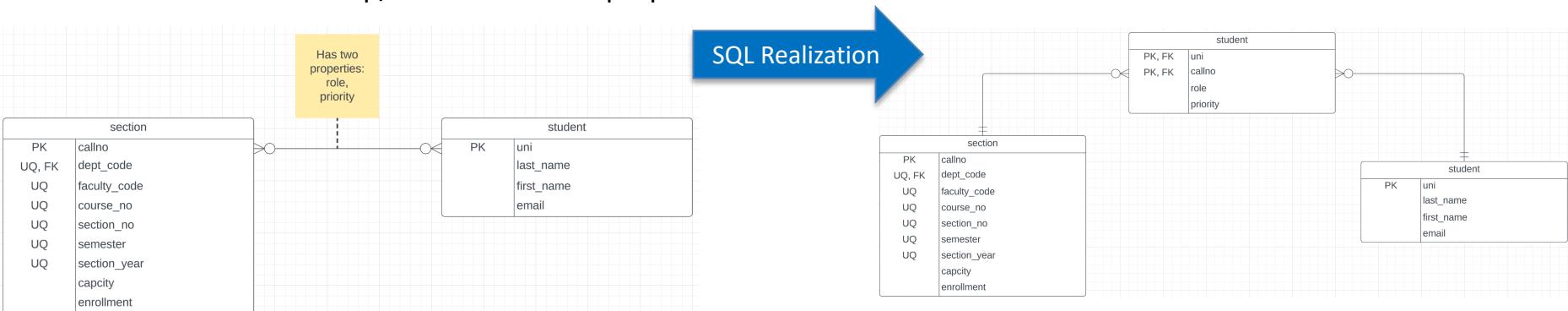
Relationships are associations between or among entities.



Weak Relationships are connections between a weak entity and its owner.

Associative Entity

- “An associative entity is a term used in relational and entity–relationship theory. A relational database requires the implementation of a base relation (or base table) to resolve many-to-many relationships. A base relation representing this kind of entity is called, informally, an associative table.” (https://en.wikipedia.org/wiki/Associative_entity)
- Consider *Students – Sections*:
 - This is many-to-many. There is not way to implement in SQL. You see this in the *Advise*s table in the sample database.
 - The “relationship/association” has properties that are not attributes of the connected entities.



Switch to notebook diagram.

Do Some Examples

Do Some Examples

- Composite and Multi-Valued Attributes
 - Composite:
 - *primaryName*.
 - Show simple example of splitting with a function.
 - Show HumanName and flow.
 - Multi-Valued:
 - *knownFor*.
 - Show pattern: Create a table. Associative entity.
- Aggregation: Class-project example.
- Inheritance examples.

Functions, Procedures, Triggers

Some Concepts



Functions and Procedures

- Functions and procedures allow “business logic” to be stored in the database and executed from SQL statements.
- These can be defined either by the procedural component of SQL or by an external programming language such as Java, C, or C++.
- The syntax we present here is defined by the SQL standard.
 - Most databases implement nonstandard versions of this syntax.

Note:

- The programming language, runtime and tools for functions, procedures and triggers are not easy to use.
- My view is that calling external functions is an anti-pattern (bad idea).
 - External code degrades the reliability, security and performance of the database.
 - Databases are often mission critical and the heart of environments.



Language Constructs for Procedures & Functions

- SQL supports constructs that gives it almost all the power of a general-purpose programming language.
 - Warning: most database systems implement their own variant of the standard syntax below.
- Compound statement: **begin ... end**,
 - May contain multiple SQL statements between **begin** and **end**.
 - Local variables can be declared within a compound statements
- While and repeat statements:
 - **while** boolean expression **do**
sequence of statements ;
end while
 - **repeat**
sequence of statements ;
until boolean expression
end repeat



(Core) Language Constructs (Cont.)

- **For** loop
 - Permits iteration over all results of a query
- Example: Find the budget of all departments

```
declare n integer default 0;
for r as
    select budget from department
        where dept_name = 'Music'
do
    set n = n + r.budget
end for
```

Note:

- There are various other looping constructs.



(Core) Language Constructs – if-then-else

- Conditional statements (**if-then-else**)

```
if boolean expression
    then statement or compound statement
    elseif boolean expression
        then statement or compound statement
    else statement or compound statement
end if
```

Note:

- We will not spend a lot of time writing functions, procedures, or triggers.
- The language and development environment are not easy to use.

Functions



Declaring SQL Functions

- Define a function that, given the name of a department, returns the count of the number of instructors in that department.

```
create function dept_count (dept_name varchar(20))
    returns integer
begin
    declare d_count integer;
    select count (*) into d_count
        from instructor
        where instructor.dept_name = dept_name
    return d_count;
end
```

- The function *dept_count* can be used to find the department names and budget of all departments with more than 12 instructors.

```
select dept_name, budget
from department
where dept_count (dept_name) > 12
```



Table Functions

- The SQL standard supports functions that can return tables as results; such functions are called **table functions**
- Example: Return all instructors in a given department

```
create function instructor_of (dept_name char(20))  
    returns table (  
        ID varchar(5),  
        name varchar(20),  
        dept_name varchar(20),  
        salary numeric(8,2))  
  
return table  
(select ID, name, dept_name, salary  
from instructor  
where instructor.dept_name = instructor_of.dept_name)
```

- Usage

```
select *  
from table (instructor_of ('Music'))
```

Procedures



SQL Procedures

- The *dept_count* function could instead be written as procedure:

```
create procedure dept_count_proc (in dept_name varchar(20),
                                   out d_count integer)
begin
    select count(*) into d_count
    from instructor
    where instructor.dept_name = dept_count_proc.dept_name
end
```

- The keywords **in** and **out** are parameters that are expected to have values assigned to them and parameters whose values are set in the procedure in order to return results.
- Procedures can be invoked either from an SQL procedure or from embedded SQL, using the **call** statement.

```
declare d_count integer;
call dept_count_proc ('Physics', d_count);
```



SQL Procedures (Cont.)

- Procedures and functions can be invoked also from dynamic SQL
- SQL allows more than one procedure of the same name so long as the number of arguments of the procedures with the same name is different.
- The name, along with the number of arguments, is used to identify the procedure.

Triggers



Triggers

- A **trigger** is a statement that is executed automatically by the system as a side effect of a modification to the database.
- To design a trigger mechanism, we must:
 - Specify the conditions under which the trigger is to be executed.
 - Specify the actions to be taken when the trigger executes.
- Triggers introduced to SQL standard in SQL:1999, but supported even earlier using non-standard syntax by most databases.
 - Syntax illustrated here may not work exactly on your database system; check the system manuals



Triggering Events and Actions in SQL

- Triggering event can be **insert**, **delete** or **update**
- Triggers on update can be restricted to specific attributes
 - For example, **after update of takes on grade**
- Values of attributes before and after an update can be referenced
 - **referencing old row as** : for deletes and updates
 - **referencing new row as** : for inserts and updates
- Triggers can be activated before an event, which can serve as extra constraints. For example, convert blank grades to null.

```
create trigger setnull_trigger before update of takes
referencing new row as nrow
for each row
    when (nrow.grade = ' ')
begin atomic
    set nrow.grade = null;
end;
```



Trigger to Maintain credits_earned value

- **create trigger** *credits_earned* **after update of** *takes* **on** (*grade*)
referencing new row as *nrow*
referencing old row as *orow*
for each row
when *nrow.grade* \neq 'F' **and** *nrow.grade* **is not null**
and (*orow.grade* = 'F' **or** *orow.grade* **is null**)
begin atomic
 update *student*
 set *tot_cred*= *tot_cred* +
 (**select** *credits*
 from *course*
 where *course.course_id*= *nrow.course_id*)
 where *student.id* = *nrow.id*;
end;



Statement Level Triggers

- Instead of executing a separate action for each affected row, a single action can be executed for all rows affected by a transaction
 - Use **for each statement** instead of **for each row**
 - Use **referencing old table** or **referencing new table** to refer to temporary tables (called ***transition tables***) containing the affected rows
 - Can be more efficient when dealing with SQL statements that update a large number of rows



When Not To Use Triggers

- Triggers were used earlier for tasks such as
 - Maintaining summary data (e.g., total salary of each department)
 - Replicating databases by recording changes to special relations (called **change** or **delta** relations) and having a separate process that applies the changes over to a replica
- There are better ways of doing these now:
 - Databases today provide built in materialized view facilities to maintain summary data
 - Databases provide built-in support for replication
- Encapsulation facilities can be used instead of triggers in many cases
 - Define methods to update fields
 - Carry out actions as part of the update methods instead of through a trigger



When Not To Use Triggers (Cont.)

- Risk of unintended execution of triggers, for example, when
 - Loading data from a backup copy
 - Replicating updates at a remote site
 - Trigger execution can be disabled before such actions.
- Other risks with triggers:
 - Error leading to failure of critical transactions that set off the trigger
 - Cascading execution

Summary

Comparison

comparing triggers, functions, and procedures

	triggers	functions	stored procedures
change data	yes	no	yes
return value	never	always	sometimes
how they are called	reaction	in a statement	exec

lynda.com

Comparison – Some Details

A *trigger* has capabilities like a procedure, except ...

- You do not call it. The DB engine calls it before or after an INSERT, UPDATE, DELETE.
- The inputs are the list of incoming new, modified rows.
- The outputs are the modified versions of the new or modified rows.

Sr.No.	User Defined Function	Stored Procedure
1	Function must return a value.	Stored Procedure may or not return values.
2	Will allow only Select statements, it will not allow us to use DML statements.	Can have select statements as well as DML statements such as insert, update, delete and so on
3	It will allow only input parameters, doesn't support output parameters.	It can have both input and output parameters.
4	It will not allow us to use try-catch blocks.	For exception handling we can use try catch blocks.
5	Transactions are not allowed within functions.	Can use transactions within Stored Procedures.
6	We can use only table variables, it will not allow using temporary tables.	Can use both table variables as well as temporary table in it.
7	Stored Procedures can't be called from a function.	Stored Procedures can call functions.
8	Functions can be called from a select statement.	Procedures can't be called from Select/Where/Having and so on statements. Execute/Exec statement can be used to call/execute Stored Procedure.
9	A UDF can be used in join clause as a result set.	Procedures can't be used in Join clause

Some Examples

Some Examples

- UNIs – Function, Trigger
- Inheritance
-
- Updating information.

Security

Security Concepts (Terms from Wikipedia)

- Definitions:
 - “A (digital) identity is information on an entity used by computer systems to represent an external agent. That agent may be a person, organization, application, or device.”
 - “Authentication is the act of proving an assertion, such as the identity of a computer system user. In contrast with identification, the act of indicating a person or thing's identity, authentication is the process of verifying that identity.”
 - “Authorization is the function of specifying access rights/privileges to resources, ... More formally, "to authorize" is to define an access policy. ... During operation, the system uses the access control rules to decide whether access requests from (authenticated) consumers shall be approved (granted) or disapproved.
 - “Within an organization, roles are created for various job functions. The permissions to perform certain operations are assigned to specific roles. Members or staff (or other system users) are assigned particular roles, and through those role assignments acquire the permissions needed to perform particular system functions.”
 - “In computing, privilege is defined as the delegation of authority to perform security-relevant functions on a computer system. A privilege allows a user to perform an action with security consequences. Examples of various privileges include the ability to create a new user, install software, or change kernel functions.”
- SQL and relational database management systems implementing security by:
 - Creating identities and authentication policies.
 - Creating roles and assigning identities to roles.
 - Granting and revoking privileges to/from roles and identities.



Authorization

- We may assign a user several forms of authorizations on parts of the database.
 - **Read** - allows reading, but not modification of data.
 - **Insert** - allows insertion of new data, but not modification of existing data.
 - **Update** - allows modification, but not deletion of data.
 - **Delete** - allows deletion of data.
- Each of these types of authorizations is called a **privilege**. We may authorize the user all, none, or a combination of these types of privileges on specified parts of a database, such as a relation or a view.



Authorization (Cont.)

- Forms of authorization to modify the database schema
 - **Index** - allows creation and deletion of indices.
 - **Resources** - allows creation of new relations.
 - **Alteration** - allows addition or deletion of attributes in a relation.
 - **Drop** - allows deletion of relations.



Authorization Specification in SQL

- The **grant** statement is used to confer authorization
grant <privilege list> on <relation or view > to <user list>
- <user list> is:
 - a user-id
 - **public**, which allows all valid users the privilege granted
 - A role (more on this later)
- Example:
 - **grant select on department to Amit, Satoshi**
- Granting a privilege on a view does not imply granting any privileges on the underlying relations.
- The grantor of the privilege must already hold the privilege on the specified item (or be the database administrator).



Privileges in SQL

- **select**: allows read access to relation, or the ability to query using the view
 - Example: grant users U_1 , U_2 , and U_3 **select** authorization on the *instructor* relation:

```
grant select on instructor to U1, U2, U3
```
- **insert**: the ability to insert tuples
- **update**: the ability to update using the SQL update statement
- **delete**: the ability to delete tuples.
- **all privileges**: used as a short form for all the allowable privileges



Revoking Authorization in SQL

- The **revoke** statement is used to revoke authorization.
revoke <privilege list> on <relation or view> from <user list>
- Example:
revoke select on student from U₁, U₂, U₃
- <privilege-list> may be **all** to revoke all privileges the revoker may hold.
- If <revoker-list> includes **public**, all users lose the privilege except those granted it explicitly.
- If the same privilege was granted twice to the same user by different grantors, the user may retain the privilege after the revocation.
- All privileges that depend on the privilege being revoked are also revoked.



Roles

- A **role** is a way to distinguish among various users as far as what these users can access/update in the database.
- To create a role we use:
 - **create a role <name>**
- Example:
 - **create role instructor**
- Once a role is created we can assign “users” to the role using:
 - **grant <role> to <users>**



Roles Example

- **create role** instructor;
- **grant** *instructor* **to** Amit;
- Privileges can be granted to roles:
 - **grant select on** *takes* **to** *instructor*;
- Roles can be granted to users, as well as to other roles
 - **create role** teaching_assistant
 - **grant** *teaching_assistant* **to** *instructor*;
 - *Instructor* inherits all privileges of *teaching_assistant*
- Chain of roles
 - **create role** dean;
 - **grant** *instructor* **to** *dean*;
 - **grant** *dean* **to** Satoshi;



Authorization on Views

- `create view geo_instructor as
(select *
from instructor
where dept_name = 'Geology');`
- `grant select on geo_instructor to geo_staff`
- Suppose that a `geo_staff` member issues
 - `select *
from geo_instructor;`
- What if
 - `geo_staff` does not have permissions on `instructor`?
 - Creator of view did not have some permissions on `instructor`?



Other Authorization Features

- **references** privilege to create foreign key
 - **grant reference** (*dept_name*) **on** *department* **to** Mariano;
 - Why is this required?
- transfer of privileges
 - **grant select on** *department* **to** Amit **with grant option**;
 - **revoke select on** *department* **from** Amit, Satoshi **cascade**;
 - **revoke select on** *department* **from** Amit, Satoshi **restrict**;
 - And more!

Note:

- Like in many other cases, SQL DBMS have product specific variations.

Switch to notebook.

Some Advanced SQL

Recursive Queries

I do not recommend this.



Recursive Queries



Recursion in SQL

- SQL:1999 permits recursive view definition
- Example: find which courses are a prerequisite, whether directly or indirectly, for a specific course

```
with recursive rec_prereq(course_id, prereq_id) as (
    select course_id, prereq_id
    from prereq
    union
    select rec_prereq.course_id, prereq.prereq_id,
    from rec_rereq, prereq
    where rec_prereq.prereq_id = prereq.course_id
)
select *
from rec_prereq;
```

This example view, *rec_prereq*, is called the *transitive closure* of the *prereq* relation



The Power of Recursion

- Recursive views make it possible to write queries, such as transitive closure queries, that cannot be written without recursion or iteration.
 - Intuition: Without recursion, a non-recursive non-iterative program can perform only a fixed number of joins of *prereq* with itself
 - This can give only a fixed number of levels of managers
 - Given a fixed non-recursive query, we can construct a database with a greater number of levels of prerequisites on which the query will not work
 - Alternative: write a procedure to iterate as many times as required
 - See procedure *findAllPrereqs* in book



Example of Fixed-Point Computation

<i>course_id</i>	<i>prereq_id</i>
BIO-301	BIO-101
BIO-399	BIO-101
CS-190	CS-101
CS-315	CS-190
CS-319	CS-101
CS-319	CS-315
CS-347	CS-319

<i>Iteration Number</i>	<i>Tuples in c1</i>
0	
1	(CS-319)
2	(CS-319), (CS-315), (CS-101)
3	(CS-319), (CS-315), (CS-101), (CS-190)
4	(CS-319), (CS-315), (CS-101), (CS-190)
5	done