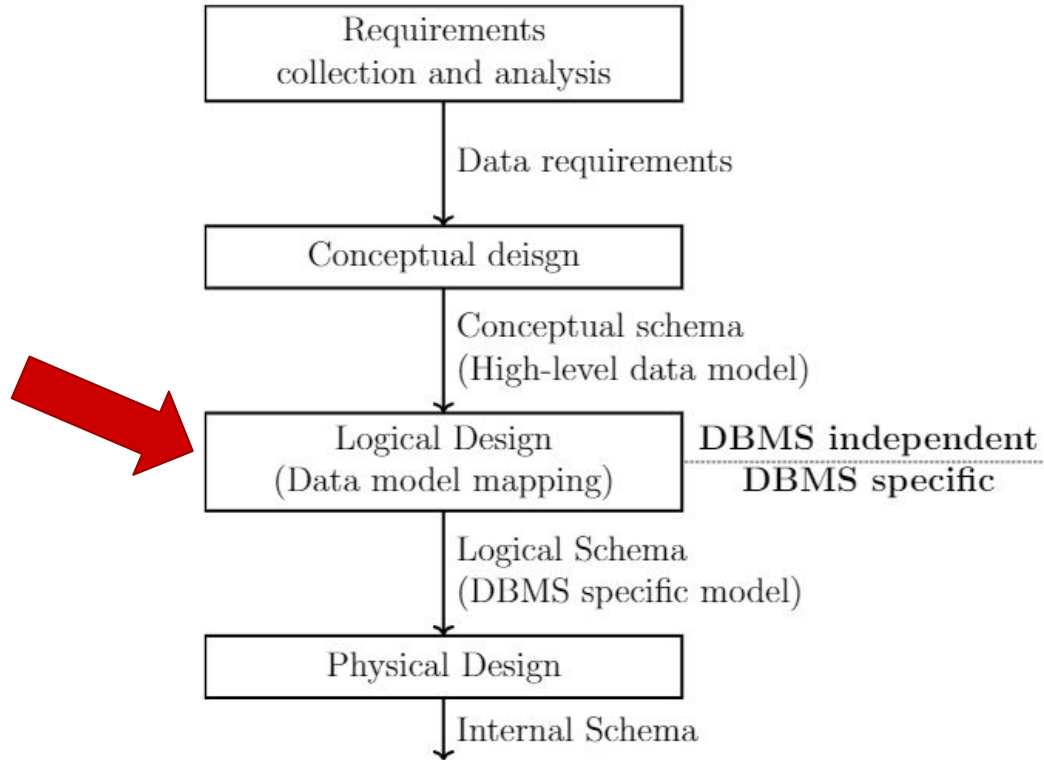


Diseño lógico: Modelo Relacional

Sivana Hamer - sivana.hamer@ucr.ac.cr
Escuela de Ciencias de la Computación
Licencia: CC BY-NC-SA 4.0

Vamos a seguir diseñando...



¿Por qué ocupamos un diseño lógico?

Vamos a diseñar con el modelo relacional



Information Retrieval

P. BAXENDALE, Editor

A Relational Model of Data for Large Shared Data Banks

E. F. CODD
IBM Research Laboratory, San Jose, California

Future users of large data banks must be protected from having to know how the data is organized in the machine (the internal representation). A prompting service which supplies such information is not a satisfactory solution. Activities of users at terminals and most application programs should remain unaffected when the internal representation of data is changed and even when some aspects of the external representation are changed. Changes in data representation will often be needed as a result of changes in query, update, and report traffic and natural growth in the types of stored information.

Existing noninferential, formatted data systems provide users with tree-structured files or slightly more general network models of the data. In Section 1, inadequacies of these models are discussed. A model based on n -ary relations, a normal form for data base relations, and the concept of a universal data sublanguage are introduced. In Section 2, certain operations on relations (other than logical inference) are discussed and applied to the problems of redundancy and consistency in the user's model.

KEY WORDS AND PHRASES: data bank, data base, data structure, data organization, hierarchies of data, networks of data, relations, derivability, redundancy, consistency, composition, join, retrieval language, predicate calculus, security, data integrity

CR CATEGORIES: 3.70, 3.73, 3.75, 4.20, 4.22, 4.29

The relational view (or model) of data described in Section 1 appears to be superior in several respects to the graph or network model [3, 4] presently in vogue for non-inferential systems. It provides a means of describing data with its natural structure only—that is, without superimposing any additional structure for machine representation purposes. Accordingly, it provides a basis for a high level data language which will yield maximal independence between programs on the one hand and machine representation and organization of data on the other.

A further advantage of the relational view is that it forms a sound basis for treating derivability, redundancy, and consistency of relations—these are discussed in Section 2. The network model, on the other hand, has spawned a number of confusions, not the least of which is mistaking the derivation of connections for the derivation of relations (see remarks in Section 2 on the “connection trap”).

Finally, the relational view permits a clearer evaluation of the scope and logical limitations of present formatted data systems, and also the relative merits (from a logical standpoint) of competing representations of data within a single system. Examples of this clearer perspective are cited in various parts of this paper. Implementations of systems to support the relational model are not discussed.

1.2. DATA DEPENDENCIES IN PRESENT SYSTEMS

The provision of data description tables in recently developed information systems represents a major advance toward the goal of data independence [5, 6, 7]. Such tables facilitate changing certain characteristics of the data representation stored in a data bank. However, the variety of data representation characteristics which can be changed *without logically impairing some application programs* is still quite limited. Further, the model of data with which users interact is still cluttered with representational prop-

Se basa en **relaciones**, un conjunto de atributos sin orden

SCHOOL



Es como una **tabla**

Name	Acronym	Phone_number	Number_students
Ciencias de la Computación e Informática	ECCI	2511-8000	888
Ciencias de la Comunicación Colectiva	ECCC	2511-3600	999
Lenguas Modernas	ELM	2511 8391	NULL
Administración de Negocios	EAN	2511-9180	3000
Antropología	EAT	2511-6458	500
Matemática	EMat	2511-6551	1500

***Se crean relaciones que tienen algún sentido lógico**

El nombre de la relación describe la relación y es único

SCHOOL

R



Name	Acronym	Phone_number	Number_students
Ciencias de la Computación e Informática	ECCI	2511-8000	888
Ciencias de la Comunicación Colectiva	ECCC	2511-3600	999
Lenguas Modernas	ELM	2511 8391	NULL
Administración de Negocios	EAN	2511-9180	3000
Antropología	EAT	2511-6458	500
Matemática	EMat	2511-6551	1500

El esquema de la relación define la relación

$R(A_1, A_2, \dots, A_n)$



SCHOOL

Name	Acronym	Phone_number	Number_students
Ciencias de la Computación e Informática	ECCI	2511-8000	888
Ciencias de la Comunicación Colectiva	ECCC	2511-3600	999
Lenguas Modernas	ELM	2511 8391	NULL
Administración de Negocios	EAN	2511-9180	3000
Antropología	EAT	2511-6458	500
Matemática	EMat	2511-6551	1500

$SCHOOL(Name, Acronym, Phone_number, Number_students) =$

$S(Name : string, Acronym : string, Phone_number : string, Number_students : integer)$

El **estado de la relación** describe las tuplas que se encuentran actualmente dentro de la relación

SCHOOL

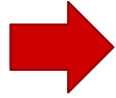
Name	Acronym	Phone_number	Number_students
Ciencias de la Computación e Informática	ECCI	2511-8000	888
Ciencias de la Comunicación Colectiva	ECCC	2511-3600	999
Lenguas Modernas	ELM	2511 8391	NULL
Administración de Negocios	EAN	2511-9180	3000
Antropología	EAT	2511-6458	500
Matemática	EMat	2511-6551	1500



$$r = r(R) = \{t_1, t_2, \dots, t_m\}$$

Un **atributo** describe un rol de un conjunto de valores en R

Son como las
columnas



A_1

A_2

A_3

A_4



SCHOOL

Name	Acronym	Phone_number	Number_students
Ciencias de la Computación e Informática	ECCI	2511-8000	888
Ciencias de la Comunicación Colectiva	ECCC	2511-3600	999
Lenguas Modernas	ELM	2511 8391	NULL
Administración de Negocios	EAN	2511-9180	3000
Antropología	EAT	2511-6458	500
Matemática	EMat	2511-6551	1500

El grado de la relación describe el número de atributos

Es cuatro

SCHOOL



Name	Acronym	Phone_number	Number_students
Ciencias de la Computación e Informática	ECCI	2511-8000	888
Ciencias de la Comunicación Colectiva	ECCC	2511-3600	999
Lenguas Modernas	ELM	2511 8391	NULL
Administración de Negocios	EAN	2511-9180	3000
Antropología	EAT	2511-6458	500
Matemática	EMat	2511-6551	1500

Los atributos tienen un **dominio** que son los valores atómicos posibles

SCHOOL

Name	Acronym	Phone_number	Number_students
Ciencias de la Computación e Informática	ECCI	2511-8000	888
Ciencias de la Comunicación Colectiva	ECCC	2511-3600	999
Lenguas Modernas	ELM	2511 8391	NULL
Administración de Negocios	EAN	2511-9180	3000
Antropología	EAT	2511-6458	500
Matemática	EMat	2511-6551	1500

$$D = \text{dom}(A_i)$$

para cada atributo



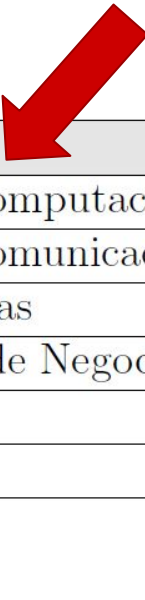
String

nnnn – nnnn

Una **tupla** representa una colección de datos relacionados

$$t = \{v_1, v_2, \dots, v_m\}$$

SCHOOL



Name	Acronym	Phone_number	Number_students
Ciencias de la Computación e Informática	ECCI	2511-8000	888
Ciencias de la Comunicación Colectiva	ECCC	2511-3600	999
Lenguas Modernas	ELM	2511 8391	NULL
Administración de Negocios	EAN	2511-9180	3000
Antropología	EAT	2511-6458	500
Matemática	EMat	2511-6551	1500

Son como las **filas**

Podemos tener valores NULL en las tuplas

SCHOOL

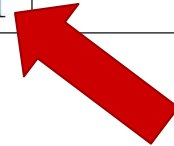


Name	Acronym	Phone_number	Number_students
Ciencias de la Computación e Informática	ECCI	2511-8000	888
Ciencias de la Comunicación Colectiva	ECCC	2511-3600	999
Lenguas Modernas	ELM	2511 8391	NULL
Administración de Negocios	EAN	2511-9180	3000
Antropología	EAT	2511-6458	500
Matemática	EMat	2511-6551	1500

Podemos referenciar un valor de una tupla

SCHOOL

Name	Acronym	Phone_number	Number_students
Ciencias de la Computación e Informática	ECCI	2511-8000	888
Ciencias de la Comunicación Colectiva	ECCC	2511-3600	999
Lenguas Modernas	ELM	2511 8391	NULL
Administración de Negocios	EAN	2511-9180	3000
Antropología	EAT	2511-6458	500
Matemática	EMat	2511-6551	1500



$$t[A_i] = t.A_i$$

Haciendo analogía con el vocabulario...

Modelo relacional

Analogía

Relación



Tabla

Esquema relacional



Definición tabla

Estado relacional



Filas actuales en la tabla

Atributo



Columna

Dominio



Todos los valores posibles

Tupla



Fila

Hablemos de llaves



Una **super llave** SK especifica un conjunto de atributos que ninguna tupla en R puede tener los mismos valores

Propiedad de ser **único**



Name	Acronym	Phone_number	Number_students
Ciencias de la Computación e Informática	ECCI	2511-8000	888
Ciencias de la Comunicación Colectiva	ECCC	2511-3600	999
Lenguas Modernas	ELM	2511 8391	NULL
Administración de Negocios	EAN	2511-9180	3000
Antropología	EAT	2511-6458	500
Matemática	EMat	2511-6551	1500

Una **llave** es única y a su vez tiene el conjunto de atributos mínimo

Propiedad de **único** y **mínimo**



Name	Acronym	Phone_number	Number_students
Ciencias de la Computación e Informática	ECCI	2511-8000	888
Ciencias de la Comunicación Colectiva	ECCC	2511-3600	999
Lenguas Modernas	ELM	2511 8391	NULL
Administración de Negocios	EAN	2511-9180	3000
Antropología	EAT	2511-6458	500
Matemática	EMat	2511-6551	1500

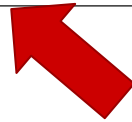
Todas las llaves posibles son **llaves candidatas**

Name	Acronym	Phone_number	Number_students
Ciencias de la Computación e Informática	ECCI	2511-8000	888
Ciencias de la Comunicación Colectiva	ECCC	2511-3600	999
Lenguas Modernas	ELM	2511 8391	NULL
Administración de Negocios	EAN	2511-9180	3000
Antropología	EAT	2511-6458	500
Matemática	EMat	2511-6551	1500

Una de las llaves candidatas se escoge de **llave primaria** como el identificador de la relación

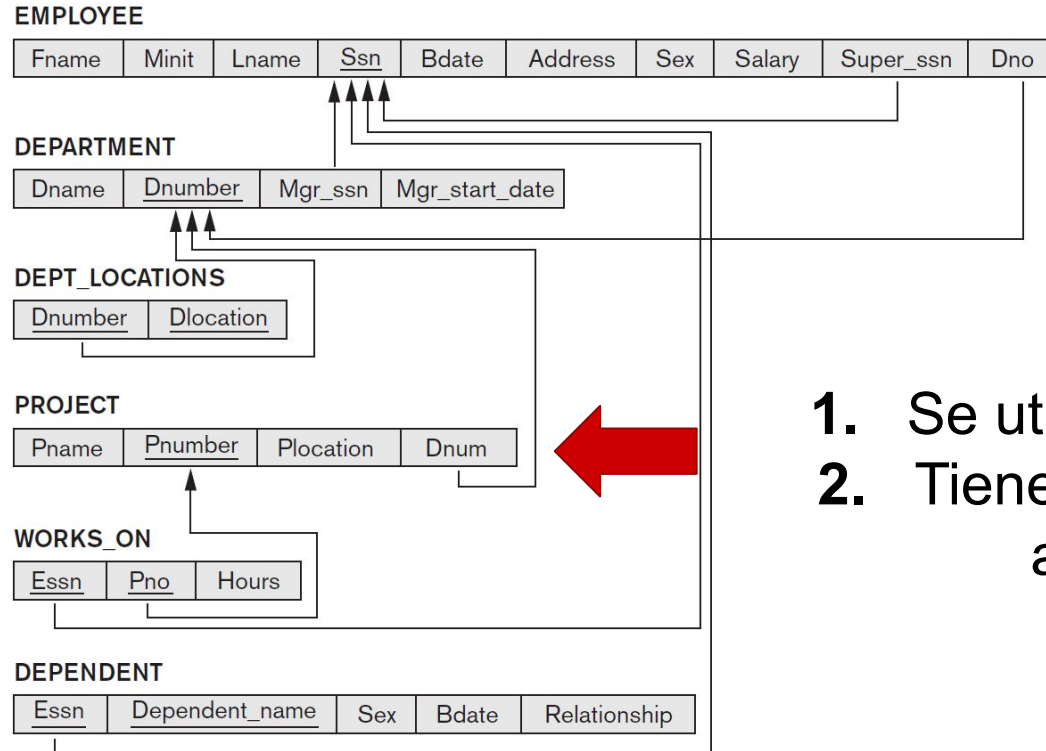


Name	<u>Acronym</u>	Phone_number	Number_students
Ciencias de la Computación e Informática	ECCI	2511-8000	888
Ciencias de la Comunicación Colectiva	ECCC	2511-3600	999
Lenguas Modernas	ELM	2511 8391	NULL
Administración de Negocios	EAN	2511-9180	3000
Antropología	EAT	2511-6458	500
Matemática	EMat	2511-6551	1500



Queda el resto como **llaves
únicas**

Una **llave foránea** referencia que existe una relación entre R_1 y una relación de referencia R_2



1. Se utiliza la llave primaria
2. Tienen el mismo dominio ambos atributos

Hablemos de restricciones...

Existen diversos niveles de **restricciones** que se pueden definir para los valores en las bases de datos

Inherentes del modelo

EMPLOYEE									
Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1965-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Alicia	J	Zelley	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Walton	987654321	1941-06-20	291 Berry, Bellaire, TX	F	42000	888665555	4
Ramesh	K	Narsany	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

DEPARTMENT			
Dname	Dnumber	Mgr_ssn	Mgr_start_date
Research	5	333445555	1968-05-22
Administration	4	987654321	1955-01-01
Headquarters	1	888665555	1981-06-19

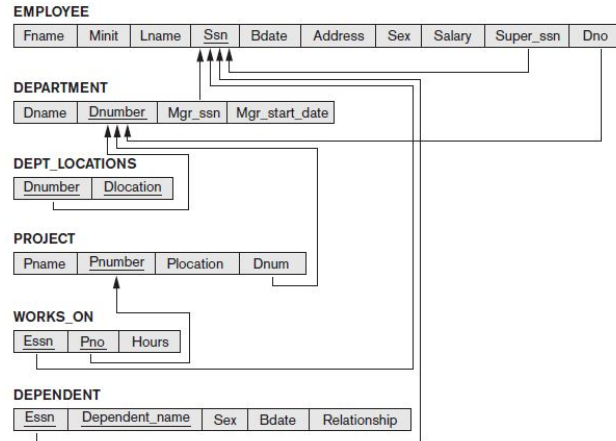
DEPT_LOCATIONS			
Dnumber	Dlocation		
1	Houston		
4	Stafford		
5	Bellaire		
5	Sugarland		
5	Houston		

WORKS_ON			
Essn	Pno	Hours	
123456789	1	22.5	
123456789	2	7.5	
666884444	3	40.0	
453453453	1	20.0	
453453453	2	20.0	
333445555	2	10.0	
333445555	3	10.0	
333445555	10	10.0	
333445555	20	10.0	

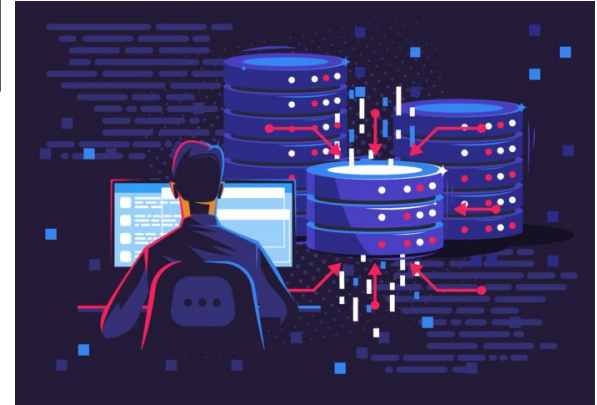
PROJECT				
Pname	Pnumber	Plocation	Dnum	
ProductX	1	Bellaire	5	
ProductY	2	Sugarland	5	
ProductZ	3	Houston	5	
Computerization	10	Stafford	4	
Reorganization	20	Houston	1	
Newbenefits	30	Stafford	4	

DEPENDENT				
Essn	Dependent_name	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
987987987	Joy	F	1958-05-03	Spouse
987987987	Abner	M	1942-02-28	Spouse
987654321	Michael	M	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse

Basados en el esquema



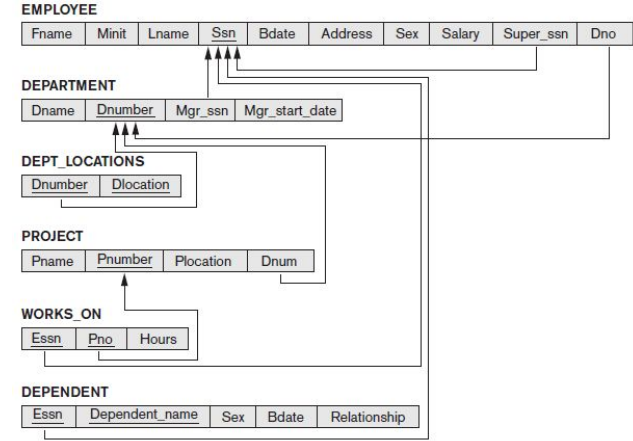
Basados en la aplicación



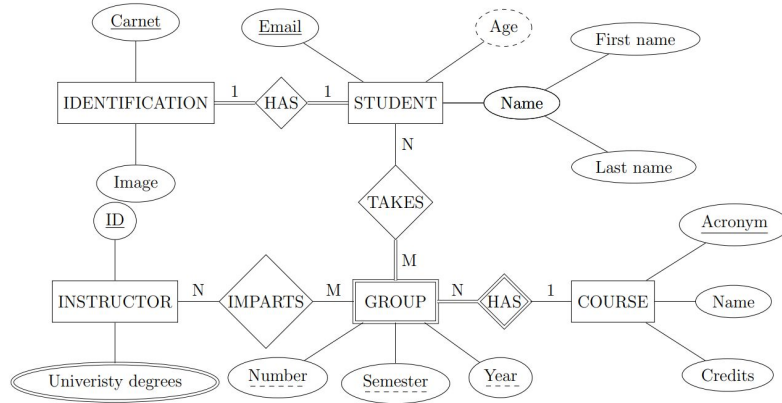
Mejor definir en el nivel más bajo

Existen las siguientes restricciones en el esquema

- Restricciones de dominio.
- Restricciones de llave.
- Restricciones de no NULL.
- Restricciones de integridad de las identidades.
- Restricciones de integridad en las referencias.



¿Cómo podemos pasar de un EER al modelo relacional?



STUDENT

<u>Email</u>	First_name	Last_name	Carnet	Image
--------------	------------	-----------	--------	-------

INSTRUCTOR

<u>ID</u>

COURSE

<u>Acronym</u>	Name	Credits
----------------	------	---------

GROUP

<u>Number</u>	<u>Semester</u>	<u>Year</u>	<u>Acronym</u>
---------------	-----------------	-------------	----------------

IMPARTS

<u>ID</u>	<u>Number</u>	<u>Semester</u>	<u>Year</u>	<u>Acronym</u>
-----------	---------------	-----------------	-------------	----------------

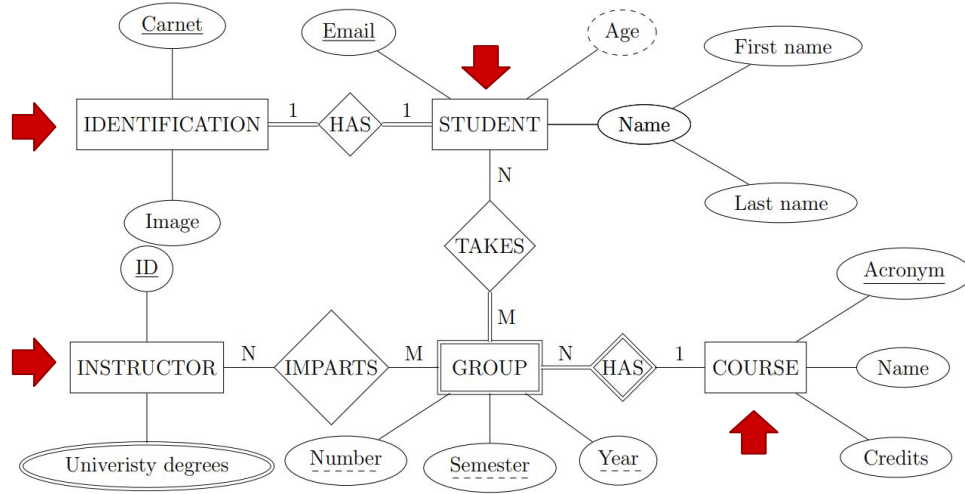
TAKES

<u>Email</u>	<u>Number</u>	<u>Semester</u>	<u>Year</u>	<u>Acronym</u>
--------------	---------------	-----------------	-------------	----------------

INSTRUCTOR.UNIVERSITY_DEGREES

<u>ID</u>	<u>University_degree</u>
-----------	--------------------------

Paso 1: Entidades fuertes



STUDENT

<u>Email</u>	First_name	Last_name
--------------	------------	-----------

IDENTIFICATION

<u>Carnet</u>	Image
---------------	-------

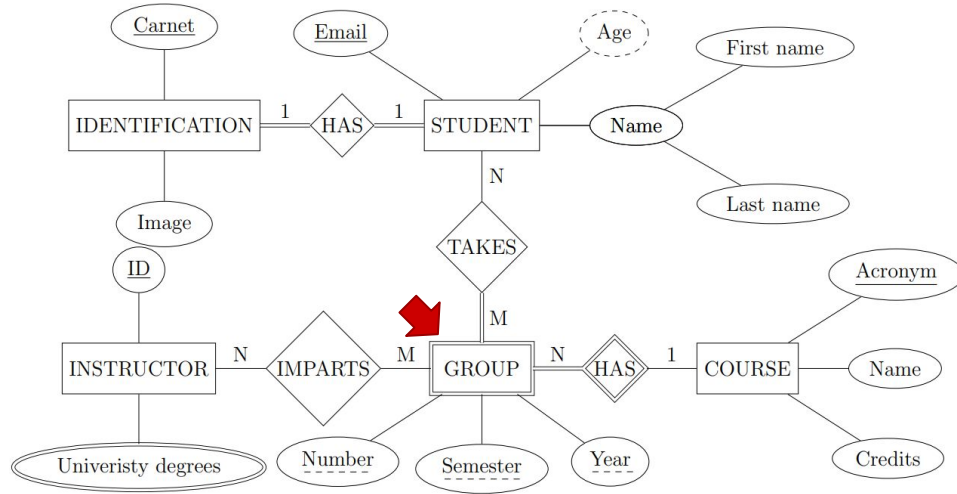
INSTRUCTOR

<u>ID</u>

COURSE

<u>Acronym</u>	Name	Credits
----------------	------	---------

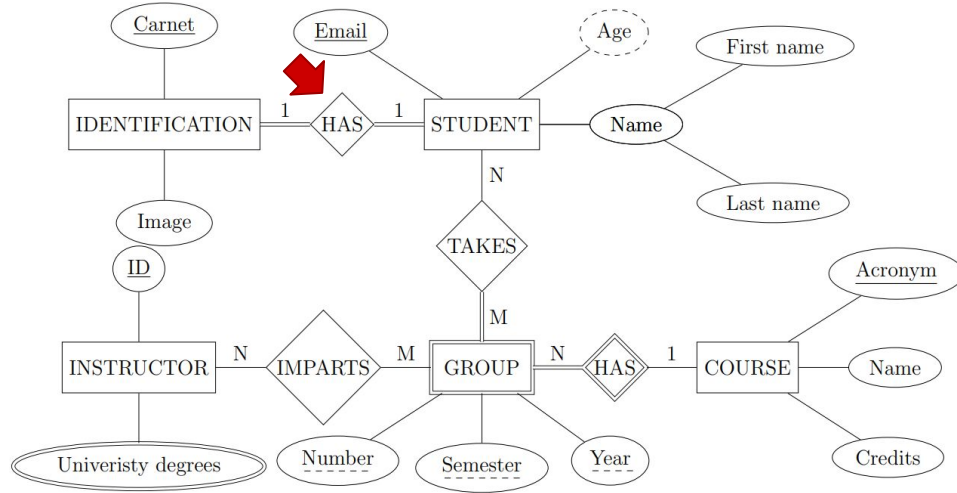
Paso 2: Entidades débiles



GROUP

<u>Number</u>	<u>Semester</u>	<u>Year</u>	<u>Acronym</u>
---------------	-----------------	-------------	----------------

Paso 3: Binarias 1:1



Foreign key approach (lado total)*

IDENTIFICATION

<u>Carnet</u>	Image	Student_email
---------------	-------	---------------

Merged relation (ambas total)

STUDENT

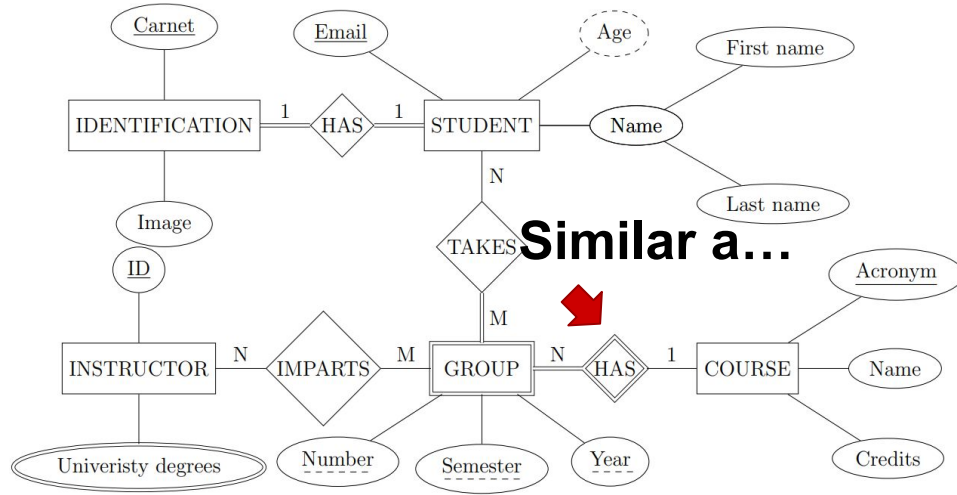
<u>Email</u>	First_name	Last_name	Carnet	Image
--------------	------------	-----------	--------	-------

Cross-reference (ambas total)

HAS

<u>Identification_Carnet</u>	Student_Email
------------------------------	---------------

Paso 4: Binarias 1:N

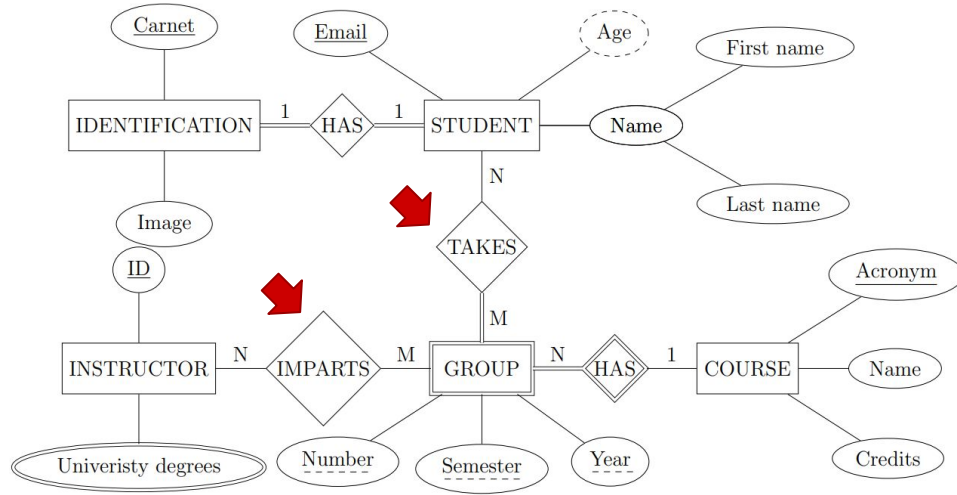


Similar a...

Dos opciones:

- Foreign key approach*
- Cross-reference

Paso 5: Binarias N:M



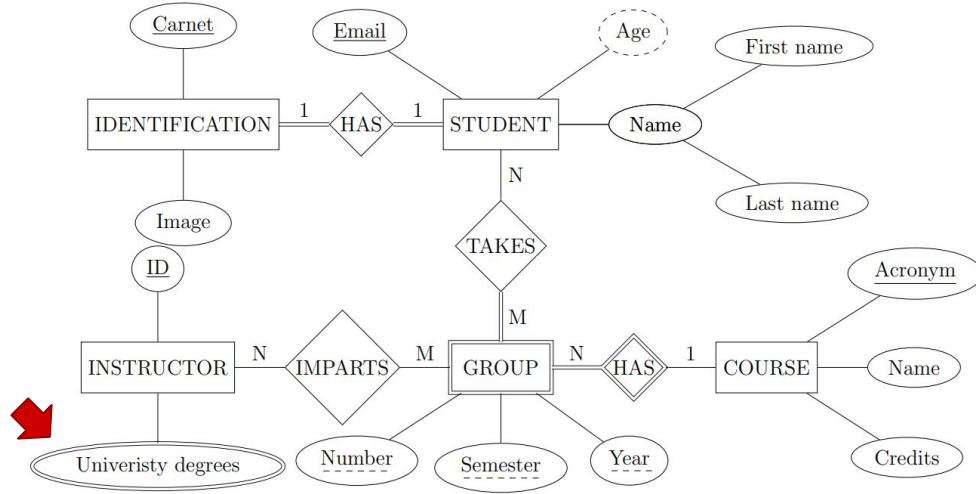
IMPARTS

<u>ID</u>	<u>Number</u>	<u>Semester</u>	<u>Year</u>	<u>Acronym</u>
-----------	---------------	-----------------	-------------	----------------

TAKES

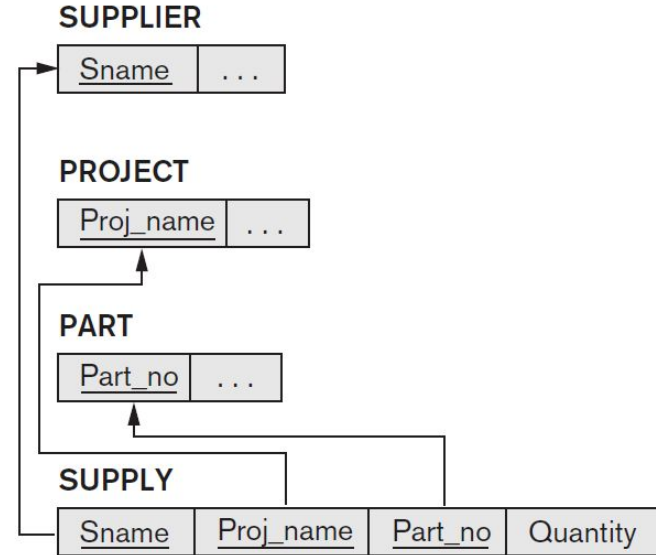
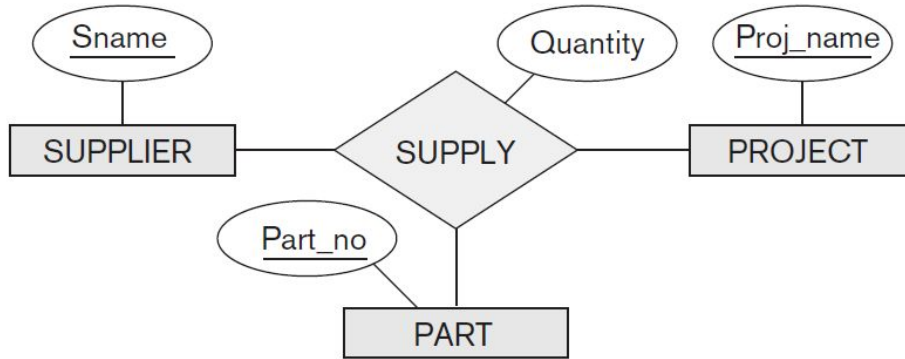
<u>Email</u>	<u>Number</u>	<u>Semester</u>	<u>Year</u>	<u>Acronym</u>
--------------	---------------	-----------------	-------------	----------------

Paso 6: Atributos multivalor



INSTRUCTOR_UNIVERSITY_DEGREES	
<u>ID</u>	<u>University_degree</u>

Paso 7: Relaciones n-arias ($n > 2$)



* Si alguna entidad tiene de cardinalidad 1, no es parte de la llave primaria

Se puede resumir el mapeo del modelo ER a relacional...

ER MODEL

Entity type

1:1 or 1:N relationship type

M:N relationship type

n -ary relationship type

Simple attribute

Composite attribute

Multivalued attribute

Value set

Key attribute

RELATIONAL MODEL

Entity relation

Foreign key (or *relationship* relation)

Relationship relation and *two* foreign keys

Relationship relation and n foreign keys

Attribute

Set of simple component attributes

Relation and foreign key

Domain

Primary (or secondary) key

Paso 8: Superclases y subclases

Superclase = Relación R

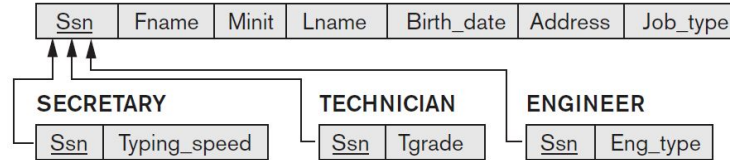
Subclases = Relación con PK y FK a superclase R

Subclases = Relación con todos atributos de superclase

Superclase = Relación con todos atributos de subclases y *un atributo de tipo*

Superclase = Relación con todos atributos de subclase y *varios atributos booleanos de tipo*

(a) EMPLOYEE



(b) CAR

<u>Vehicle_id</u>	License_plate_no	Price	Max_speed	No_of_passengers
-------------------	------------------	-------	-----------	------------------

TRUCK

<u>Vehicle_id</u>	License_plate_no	Price	No_of_axles	Tonnage
-------------------	------------------	-------	-------------	---------

(c) EMPLOYEE

<u>Ssn</u>	Fname	Minit	Lname	Birth_date	Address	Job_type	Typing_speed	Tgrade	Eng_type
------------	-------	-------	-------	------------	---------	----------	--------------	--------	----------

(d) PART

<u>Part_no</u>	Description	Mflag	Drawing_no	Manufacture_date	Batch_no	Pflag	Supplier_name	List_price
----------------	-------------	-------	------------	------------------	----------	-------	---------------	------------

La relación entre restricciones de herencia y esquema de mapeo de subclases es...

	Disjoint	Overlapping
8A	✓	✓
8B	✓	✗
8C	✓	✗
8A	✓	✓

How normal people see ~~relationships~~ relations



How database engineers see ~~relationships~~ relations



Referencias

- R. Elmasri and S. Navathe, Fundamentals of database systems, 7th ed. Pearson, 2016, chapters 5 and 9.
- https://www.youtube.com/watch?v=P8n_rwPzdBc&t=470s