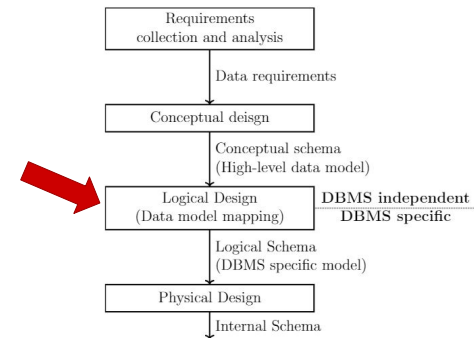


Diseño lógico: Modelo Relacional

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

Vamos a seguir diseñando...



UNIVERSIDAD DE COSTA RICA CI 0127 - sivana.hamer@ucr.ac.cr

2

¿Por qué ocupamos un diseño lógico?

Vamos a diseñar con el modelo relacional

Information Retrieval

A Relational Model of Data for Large Shared Data Banks

E. F. COO
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 of 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 nonrelational, 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 every relation, 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 representation, knowledge of data, networks of data, relations, derivability, redundancy, consistency, membership, join, universal language, predicate calculus, security, data integrity.

CR CATEGORIES: 3.70, 3.75, 3.79, 4.30, 4.32, 4.39

F. BARENDAISE, Editor



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 testing 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 DERIVATION IN PRAXIS 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 disturbing some application programs is still quite limited. Further, the model of data with which users interact is still cluttered with representational prop-

UNIVERSIDAD DE COSTA RICA CI 0127 - sivana.hamer@ucr.ac.cr

3

UNIVERSIDAD DE COSTA RICA CI 0127 - sivana.hamer@ucr.ac.cr

4

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

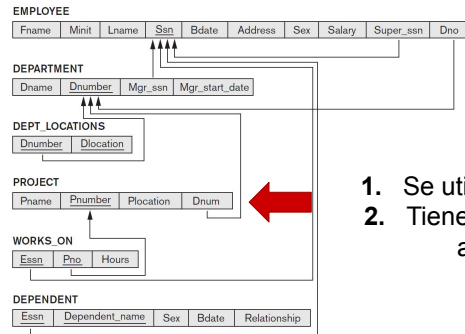


Name	Acronym	Phone_number	Number_students
Ciencias de la Computación e Informática	ECCI	2511-8000	888
Ciencias de la Comunicación Colectiva	ECCC	2511-3600	999
Lenguas Modernas	ELM	2511 8391	NULL
Administración de Negocios	EAN	2511-9180	3000
Antropología	EAT	2511-6458	500
Matemática	EMat	2511-6551	1500

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

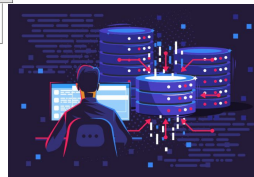
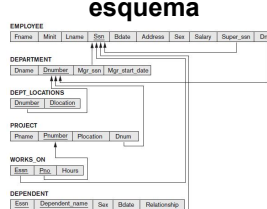
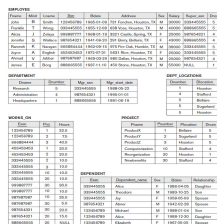
Hablemos de restricciones...

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

Inherentes del modelo

Basados en el esquema

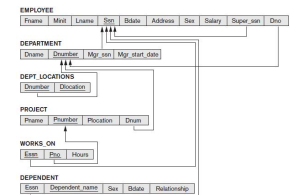
Basados en la aplicación



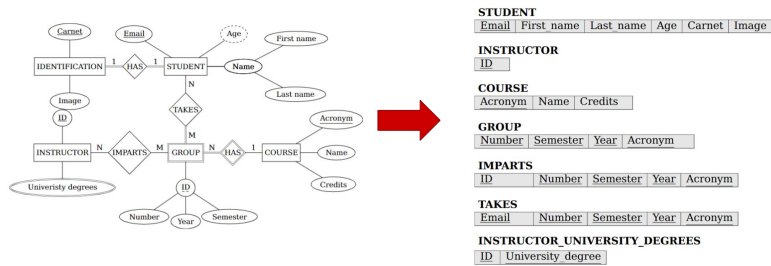
Mejor definir en el nivel más bajo

Existen las siguientes restricciones en el esquema

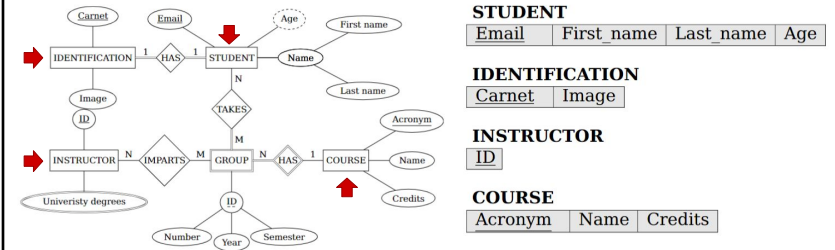
- Restricciones de dominio.
- Restricciones de llave.
- Restricciones de no NULL.
- Restricciones de integridad de las identidades.
- Restricciones de integridad en las referencias.



¿Cómo podemos pasar de un EER al modelo relacional?

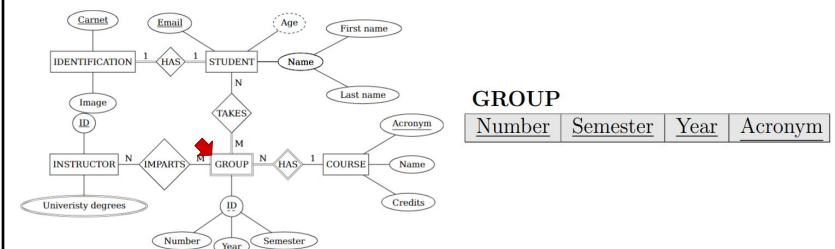


Paso 1: Entidades fuertes

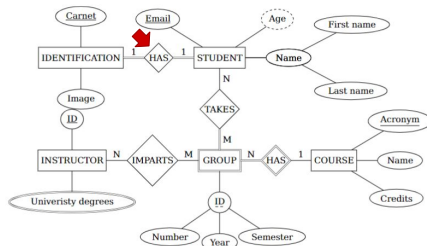


Con los **atributos derivados**, vamos a crear un atributo dentro de la relación PERO ocupamos implementar esa restricción por medio de SQL

Paso 2: Entidades débiles



Paso 3: Binarias 1:1



Foreign key approach (lado total)*

IDENTIFICATION

Carnet	Image	Student_email
--------	-------	---------------

Merged relation (ambas total)

STUDENT

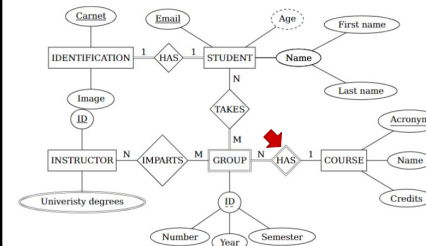
Email	First_name	Last_name	Age	Carnet	Image
-------	------------	-----------	-----	--------	-------

Cross-reference (ambas total)

HAS

Identification_Carnet	Student_Email
-----------------------	---------------

Paso 4: Binarias 1:N



Dos opciones:

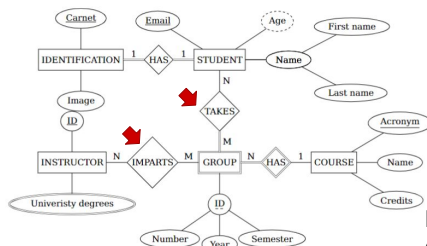
- Foreign key approach*
- Cross-reference

GROUP

Number	Semester	Year	Acronym
--------	----------	------	---------

Nota: Cuando hacemos el mapeo de entidades fuertes ya hacemos este mapeo

Paso 5: Binarias N:M



IMPARTS

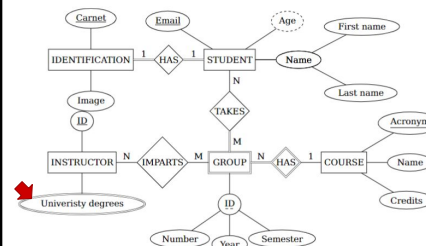
ID	Number	Semester	Year	Acronym
----	--------	----------	------	---------

TAKES

Email	Number	Semester	Year	Acronym
-------	--------	----------	------	---------

Nota: Igual que cross reference 1 a 1 donde ambas son llaves

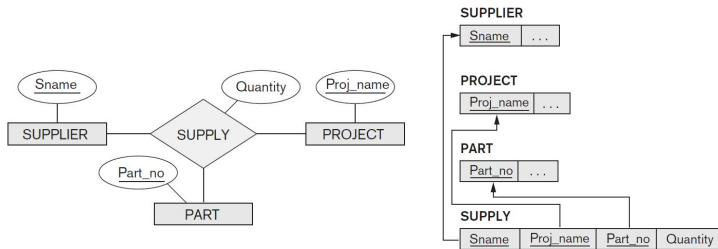
Paso 6: Atributos multivalor



INSTRUCTOR.UNIVERSITY_DEGREES

ID	University_degree
----	-------------------

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



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

STUDENT

Email First name Last name Age Carnet Image

INSTRUCTOR

ID

COURSE

Acronym Name Credits

GROUP

Number Semester Year Acronym

IMPARTS

ID Number Semester Year Acronym

TAKES

Email Number Semester Year Acronym

INSTRUCTOR_UNIVERSITY_DEGREES

ID University degree

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

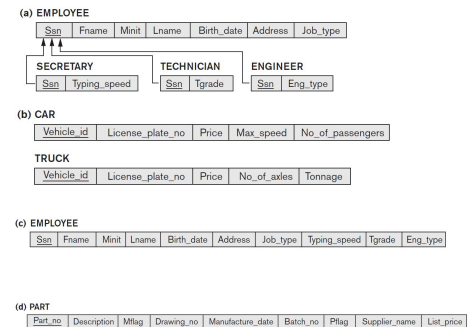
Superclase = Relación R

Subclases = Relación con PK y FK a superclase R

Subclases = Relación con todos atributos de superclase

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

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



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

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

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