

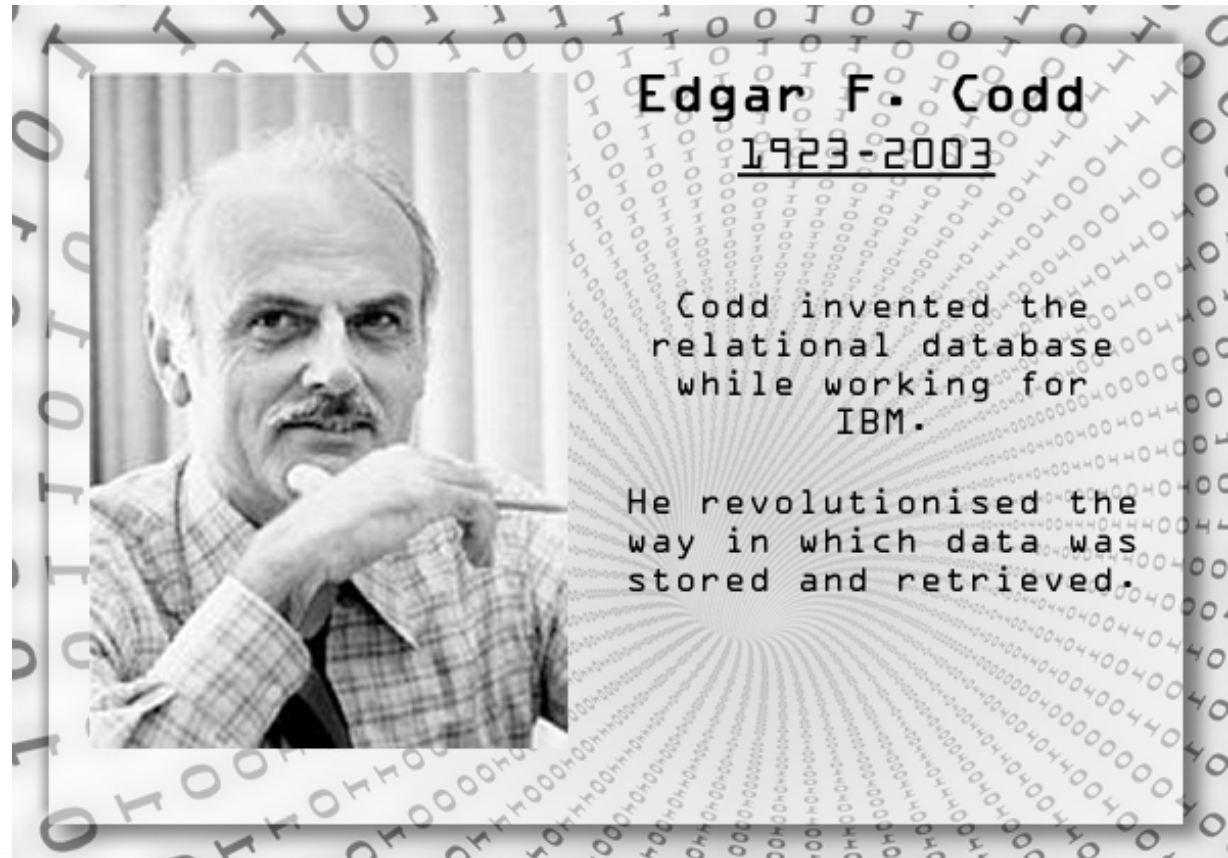
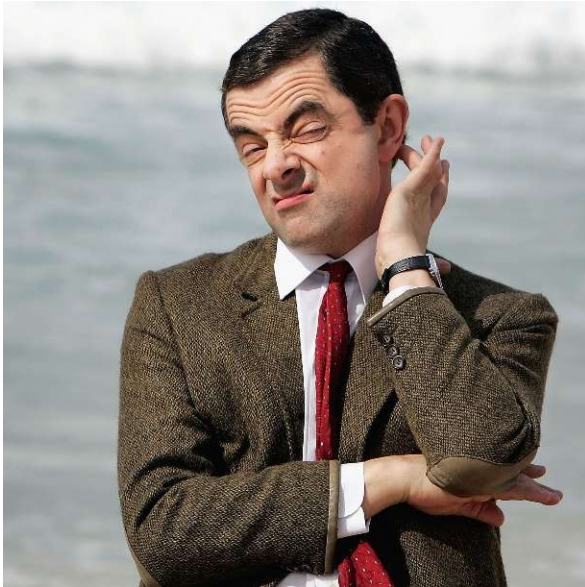


UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE CIENCIAS
FUNDAMENTOS DE BASES DE DATOS

El Modelo Relacional

Gerardo Avilés Rosas
gar@ciencias.unam.mx

Modelo relacio... ¿qué?



Edgar F. Codd
1923-2003

Codd invented the relational database while working for IBM.

He revolutionised the way in which data was stored and retrieved.

| Años | Sucesos |
|------|--|
| 1968 | Surge el Modelo Relacional (Codd) |
| 1970 | Desarrollo teóricos, álgebra relacional (Codd, 1972) |
| 1973 | Prototipos (Ingres, Sistema R, etc.) |
| 1979 | Oracle |
| 1981 | SQL |
| 1982 | Sybase, Informix |
| 1984 | SQL/ANS |
| 1986 | SQL ISO |
| 1990 | Modelo Relacional versión 2 (RM/V2) |
| 1992 | SQL2 estándar |
| 1994 | SQL3 aún no estandarizado BDOO |

Information Retrieval

A Relational Model of Data for Large Shared Data Banks

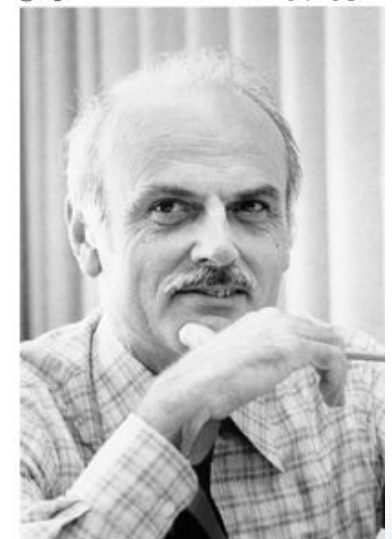
E. F. Codd

IBM Research Laboratory, San Jose, California

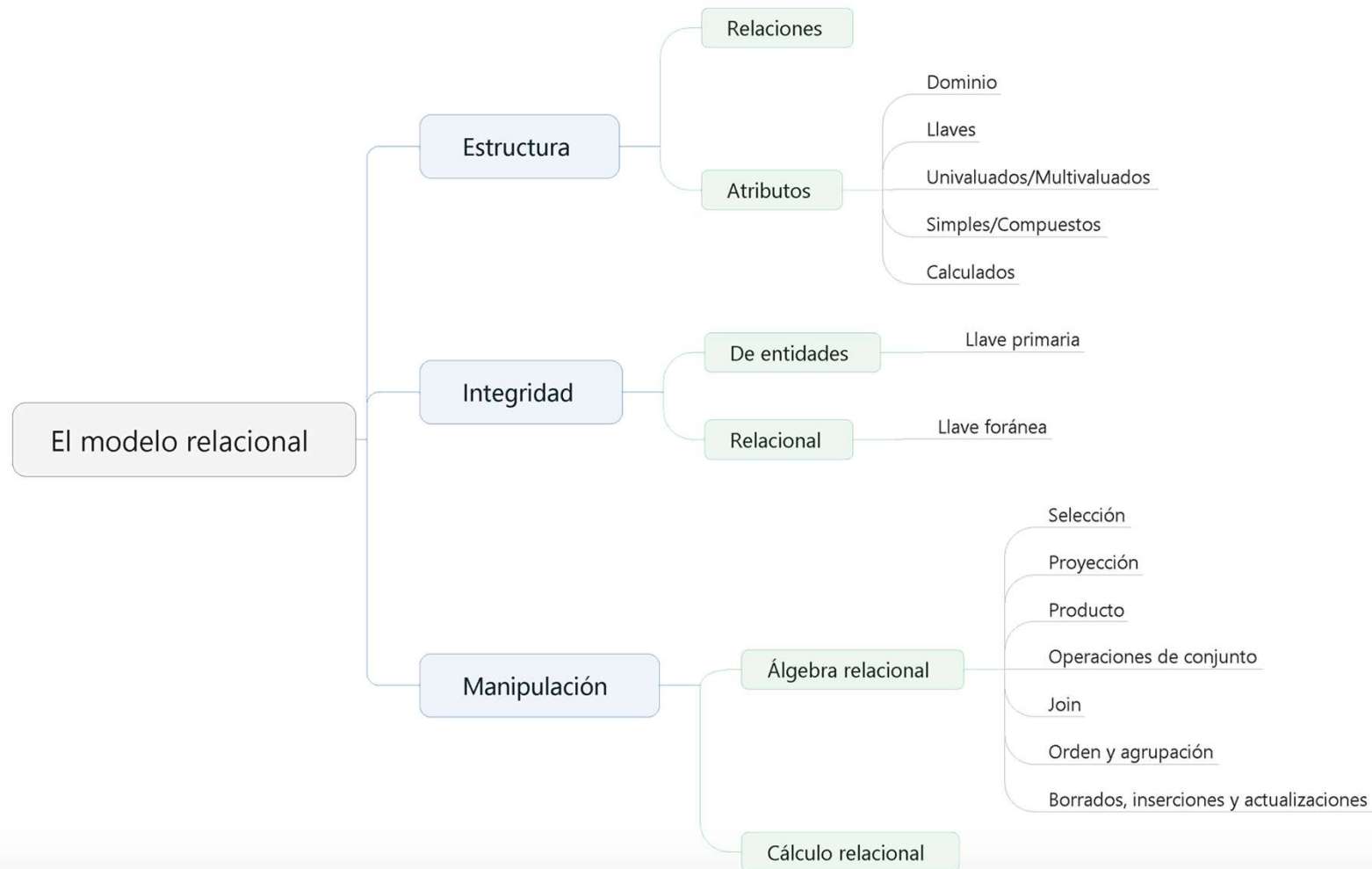
Future users of large data banks must be protected from having to know how the data is organized in the machine (the internal representation). A prompting service which supplies such information is not a satisfactory solution. Activities of users at terminals and most application programs should remain 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

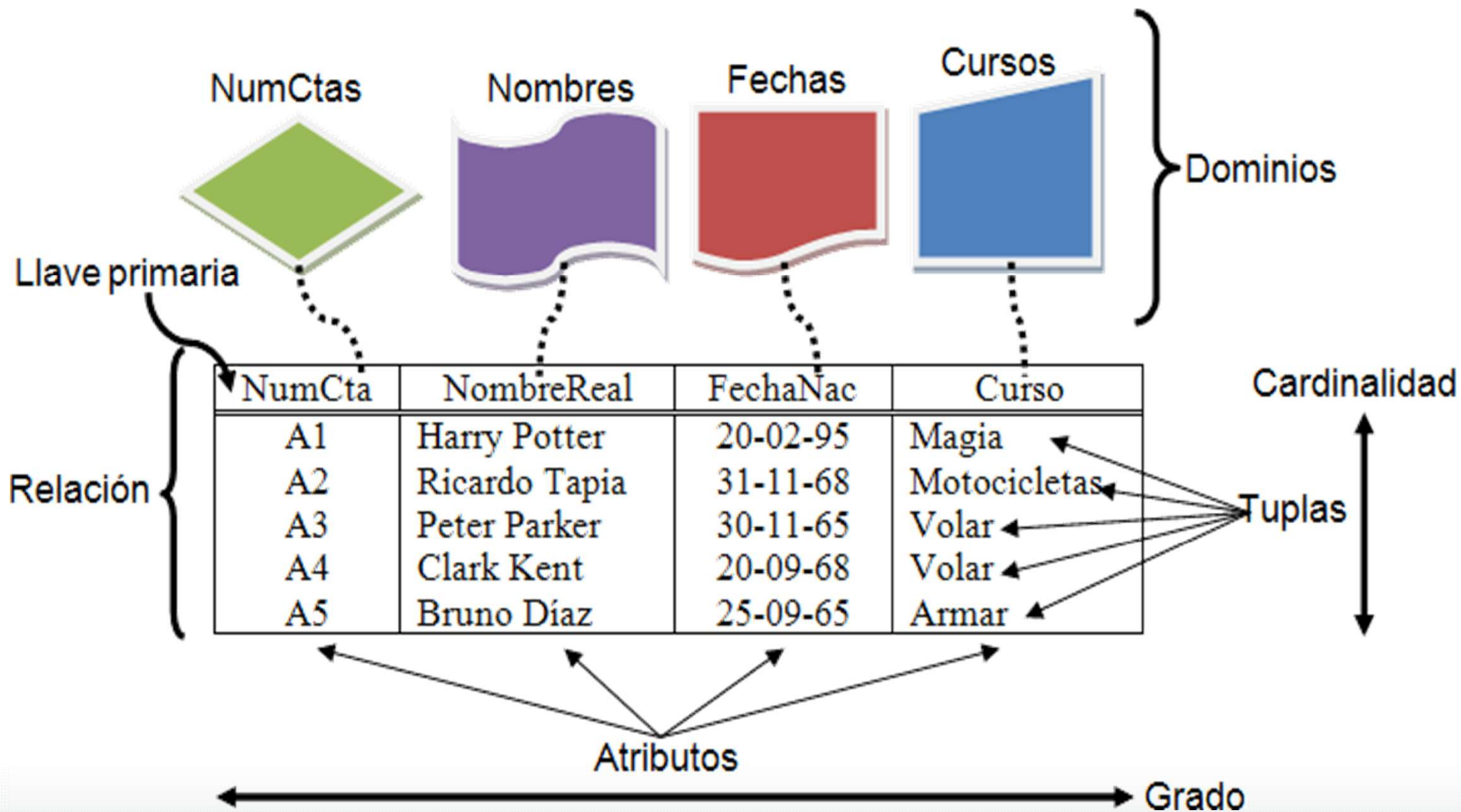
The relational view (or model Section 1 appears to be superior in graph or network model [3, 4] pre



Está basado en dos ramas de las matemáticas: la **teoría de conjuntos** y la **lógica de predicados** de primer orden. Su base matemática hace que el modelo sea **predecible, fiable y seguro** ¿qué más se puede pedir?



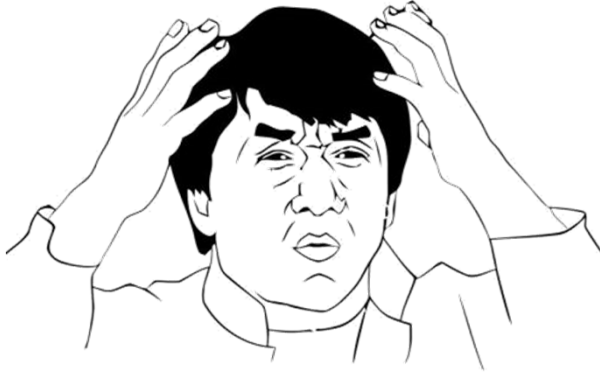
Una BDR (**Codd, 1968**) es una BD formada por una colección de relaciones:



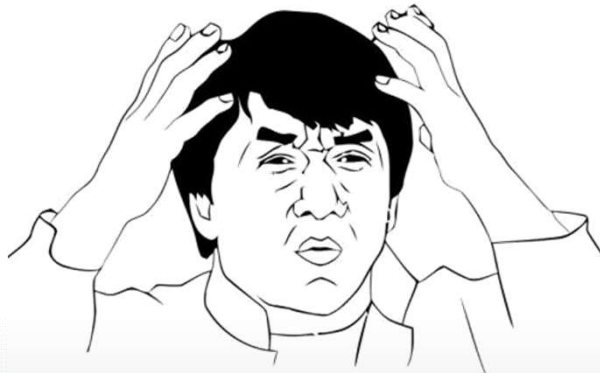
Objetivos del Modelo Relacional

- **Independencia física.** El modo en el que se almacenan los datos no influye en su manipulación lógica y por tanto, los usuarios que acceden a esos datos no tienen que modificar sus programas por cambios en el almacenamiento físico.
- **Independencia lógica.** El añadir, eliminar o modificar objetos de la base de datos no repercute en los programas y/o usuarios que están accediendo a subconjuntos parciales de los mismos (vistas).
- **Flexibilidad.** En el sentido de poder presentar a cada usuario los datos de la forma en que éste prefiera.
- **Uniformidad.** Las estructuras lógicas de los datos presentan un aspecto uniforme, lo que facilita la concepción y manipulación de la base de datos por parte de los usuarios.
- **Sencillez.** Las características anteriores, así como unos lenguajes de usuario muy sencillos hace que este modelo sea fácil de comprender y de utilizar por parte del usuario final.

Una relación R de los conjuntos $A_1, A_2, A_3, \dots, A_n$ es un subconjunto del **producto cartesiano** de los mismos: $R \subseteq A_1 \times A_2 \times A_3 \times A_n$



- **Producto cartesiano:** El producto cartesiano de los conjuntos $A_1, A_2, A_3, \dots, A_n$ representado por $A_1 \times A_2 \times A_3 \times A_n$ es el conjunto de todas las tuplas ordenadas $(a_1, a_2, a_3, \dots, a_n)$ tales que $a_1 \in A_1, a_2 \in A_2, a_3 \in A_3, \dots, a_n \in A_n$.



Vamos a suponer que tenemos información de **ropa (A_1)**, **colores (A_2)** y **tonos (A_3)**:

- ❑ $A_1 = \{\text{camisa, zapatos, calcetines, pantalón}\}$
- ❑ $A_2 = \{\text{rojo, azul, verde, amarillo}\}$
- ❑ $A_3 = \{\text{claro, intenso, oscuro, brillante}\}$

Es posible definir las siguientes relaciones:

- $R_1 = \{(\text{camisa, azul, claro}), (\text{zapatos, verde, brillante}), (\text{calcetines, rojo, intenso}), (\text{pantalón, azul, oscuro})\}$
- $R_3 = \{(\text{camisa, rojo, claro}), (\text{camisa, rojo, intenso}), (\text{zapatos, rojo, claro}), (\text{zapatos, rojo, oscuro})\}$
- $R_2 = \{(\text{pantalón, verde, claro}), (\text{camisa, verde, oscuro})\}$



- **Dominio:** Conjunto finito de valores **homogéneos** y **atómicos** caracterizados por un nombre.
- **Atributo:** Aquel que participa en la descripción de las entidades y que como tal constituye una pieza específica de información para un determinado dominio.
- **Llaves:** Conjunto no vacío de atributos que identifican unívoca y mínimamente cada *tupla*:
 - ❑ **Llave primaria:** Aquella llave que permite identificar *tuplas* de la relación de forma única.
 - ❑ **Llaves candidatas:** Son aquellas que no han sido escogidas como llaves primarias pero que también podrían identificar de manera única a una *tupla*.
 - ❑ **Llave foránea:** Conjunto no vacío de atributos cuyos valores han de coincidir con los valores de la llave primaria en una relación.
- **Restricciones:** Son estructuras no permitidas y hay de dos tipos: *inherentes* y *del usuario*.

Características de las relaciones

- ☐ No hay tuplas duplicadas
- ☐ Es irrelevante el orden de las tuplas
- ☐ Los atributos están desordenados

*Por ejemplo, la relación **estudiante(nombre,num_cta,edad,...)** o bien **(num_cta,...,nombre,...,edad,...)***

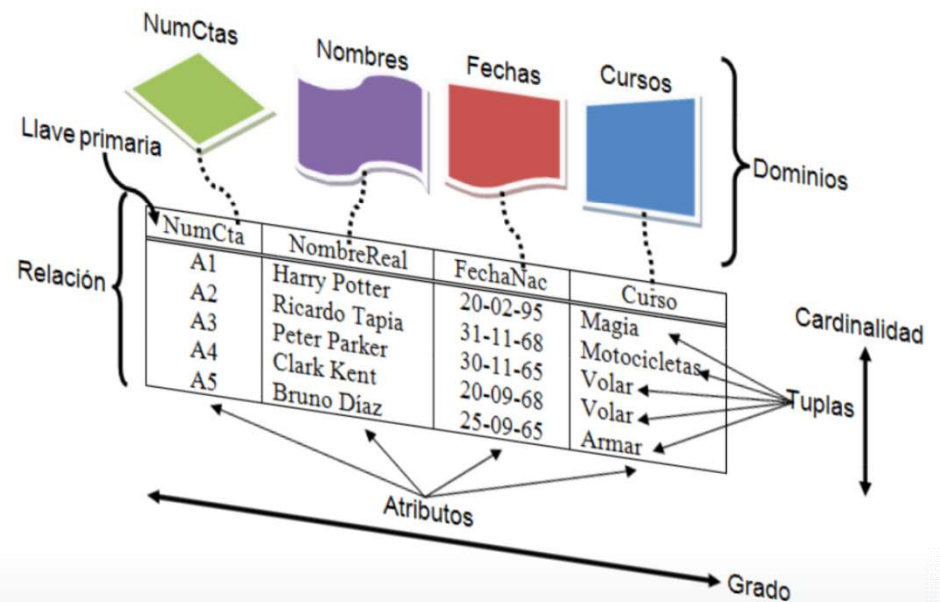
- ☐ Todos los atributos tienen **valores atómicos**

Esquema de BD = conjunto de esquemas de las relaciones de un diseño

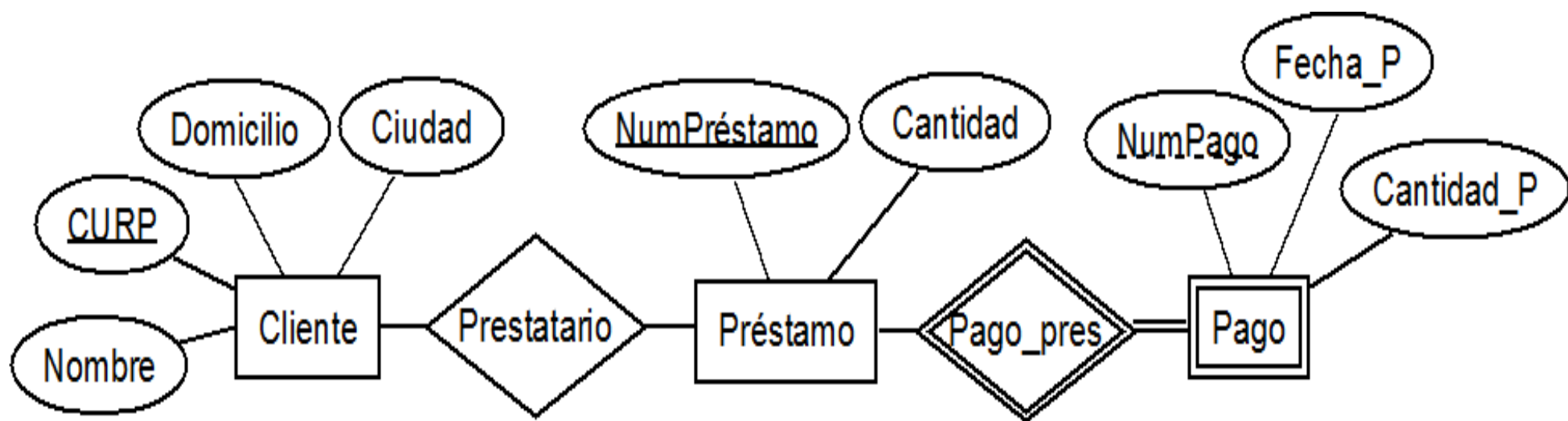
Conversión Modelo E/R a tablas

- Las llaves permiten que los conjuntos de entidades y los de relaciones se expresen de manera uniforme como **tablas** que representan el contenido de la BD.
- Una BD que conforma un diagrama E-R puede representarse como una colección de tablas.
- Para cada conjunto de entidades y cada conjunto de relaciones existe una tabla única con el nombre de tal conjunto.
- Cada tabla tiene una entidad de columnas fija, cada una con nombre único.
- Convertir un diagrama E-R a un formato de tablas es la base para derivar un diseño de BDR de un diagrama E-R

...Conversión Modelo E/R a tablas



Ejemplo de conversión



- Una **entidad fuerte** se convierte en una tabla con los mismos atributos.

Cliente:

| Nombre | <u>CURP</u> | Domicilio | Ciudad |
|-----------|-------------|-------------|------------|
| Santos | 19283746 | Mayor 78-2 | La Paz |
| Gómez | 19283756 | Carretas 65 | Querétaro |
| López | 67789901 | Mayor 67 | La Paz |
| Pérez | 96396396 | Carretas 89 | Querétaro |
| Jiménez | 19283746 | Azucena 124 | Cuernavaca |
| Ramírez | 32112312 | Montejo 567 | Mérida |
| Vázquez | 24466880 | Goya 67 | Cuetzalan |
| Fernández | 19283756 | Jazmín 111 | Cuernavaca |
| González | 33557799 | Arenal 234 | Sonora |

Préstamo:

| <u>NumPrestamo</u> | Importe |
|--------------------|---------|
| P-17 | 200,000 |
| P-23 | 400,000 |
| P-25 | 300,000 |
| P-14 | 300,000 |
| P-93 | 100,000 |
| P-11 | 150,000 |
| P-22 | 180,000 |
| P-15 | 250,000 |
| P-67 | 450,000 |

- En el caso de **entidades débiles**, sea **A** una entidad débil con atributos a_1, a_2, \dots, a_n y **B** la entidad fuerte de la que depende **A**, con llave primaria formada por los atributos b_1, b_2, \dots, b_m . Esta entidad débil se representa mediante una tabla C con columnas para cada atributo del conjunto
 - $\{a_1, a_2, \dots, a_n\} \cup \{b_1, b_2, \dots, b_m\}$

Pago:

| <u>NumPrestamo</u> | <u>NumPago</u> | <u>Fecha_P</u> | <u>Cantidad_P</u> |
|--------------------|----------------|----------------|-------------------|
| P-17 | 5 | 10-12-00 | 1,000 |
| P-23 | 11 | 17-12-00 | 1,500 |
| P-25 | 22 | 23-12-00 | 600 |
| P-14 | 69 | 28-12-00 | 10,000 |
| P-93 | 103 | 03-01-01 | 18,000 |
| P-11 | 6 | 08-01-01 | 1,000 |
| P-22 | 53 | 08-01-01 | 2,500 |
| P-15 | 104 | 25-01-01 | 4,000 |
| P-67 | 7 | 26-01-01 | 2,000 |

Sean:

- R un conjunto de relaciones,
- $\{a_1, a_2, \dots, a_n\}$ el conjunto de atributos formados por la unión de llaves primarias de cada uno de los conjuntos de entidades que participan en R,
- $\{b_1, b_2, \dots, b_m\}$ el conjunto de atributos de R.

El conjunto de relaciones se representa mediante una tabla llamada R con una columna por cada atributo de:

$$\{a_1, a_2, \dots, a_n\} \cup \{b_1, b_2, \dots, b_m\}$$

Prestatario:

| <u>CURP</u> | <u>NumPrestamo</u> |
|-------------|--------------------|
| 19283746 | P-17 |
| 19283756 | P-23 |
| 67789901 | P-25 |
| 96396396 | P-14 |
| 19283746 | P-93 |
| 32112312 | P-11 |
| 24466880 | P-22 |
| 19283756 | P-15 |
| 33557799 | P-67 |

Para **relaciones 1:1** se tienen tres posibilidades:

- **Relación parcial**



Se sigue la regla anterior aunque se pierde la semántica

- **Relación total de un lado**



Incluir en B los atributos de R y la llave de A

- **Relación total de ambos lados**



Incluir los atributos de A, B y de R en una sola relación adicional

- Para **relaciones 1:N** como la siguiente:



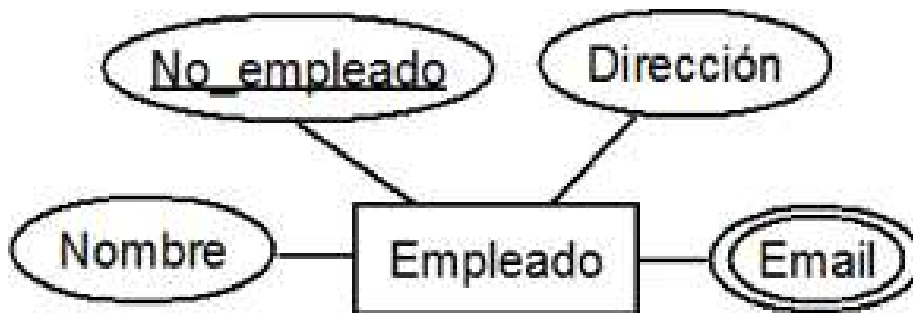
- En la relación **A** se incluye la llave de la relación **B** más los atributos de la relación **R**.
- Un conjunto de relaciones que asocia un **conjunto de entidades débiles** a un conjunto de entidades fuertes

| <u>NumPrestamo</u> | <u>NumPago</u> |
|--------------------|----------------|
| ... | ... |

- En general, la tabla para esta relación es redundante y por tanto no necesita representarse de esta manera.

Atributos multivaluados

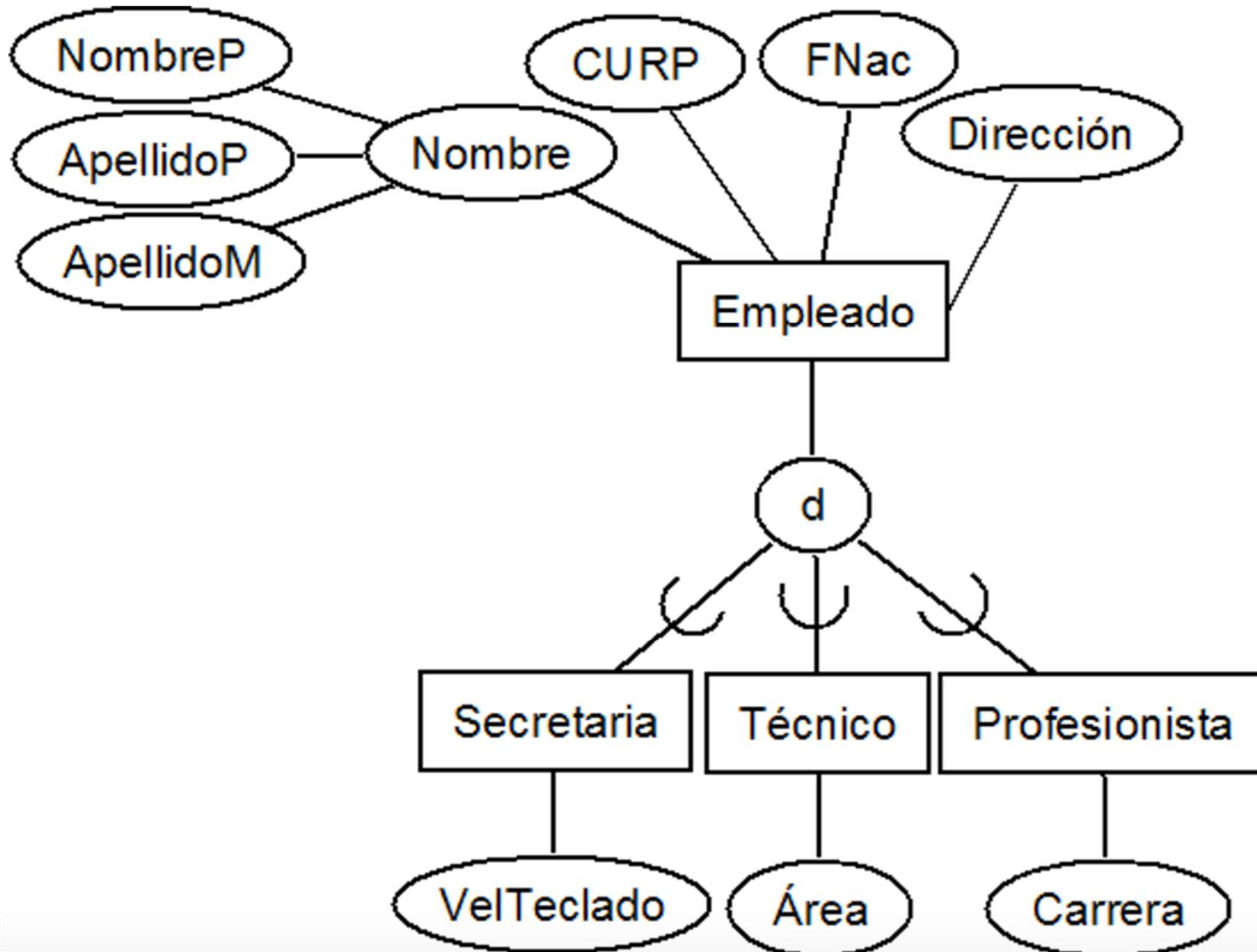
- Los **atributos multivaluados** se convierten en **tablas** no en columnas.
- Si **M** es un **atributo multivaluado**, se crea una tabla **T** con una columna que corresponde a la llave primaria del conjunto de entidades o conjunto de relaciones del que M es atributo y otra para el atributo.



| <u>No_empleado</u> | <u>email</u> |
|--------------------|--------------|
| ... | ... |

Generalización/Especialización

- En algunos casos se tiene una jerarquía como la siguiente:



...Generalización/Especialización

Caso general: Especialización/generalización parcial con disyunción.

- Se crea una tabla para la super entidad con todos sus atributos y se crea una tabla Ti para cada subentidad que contenga los atributos de esa subentidad más el atributo llave de la entidad superior.

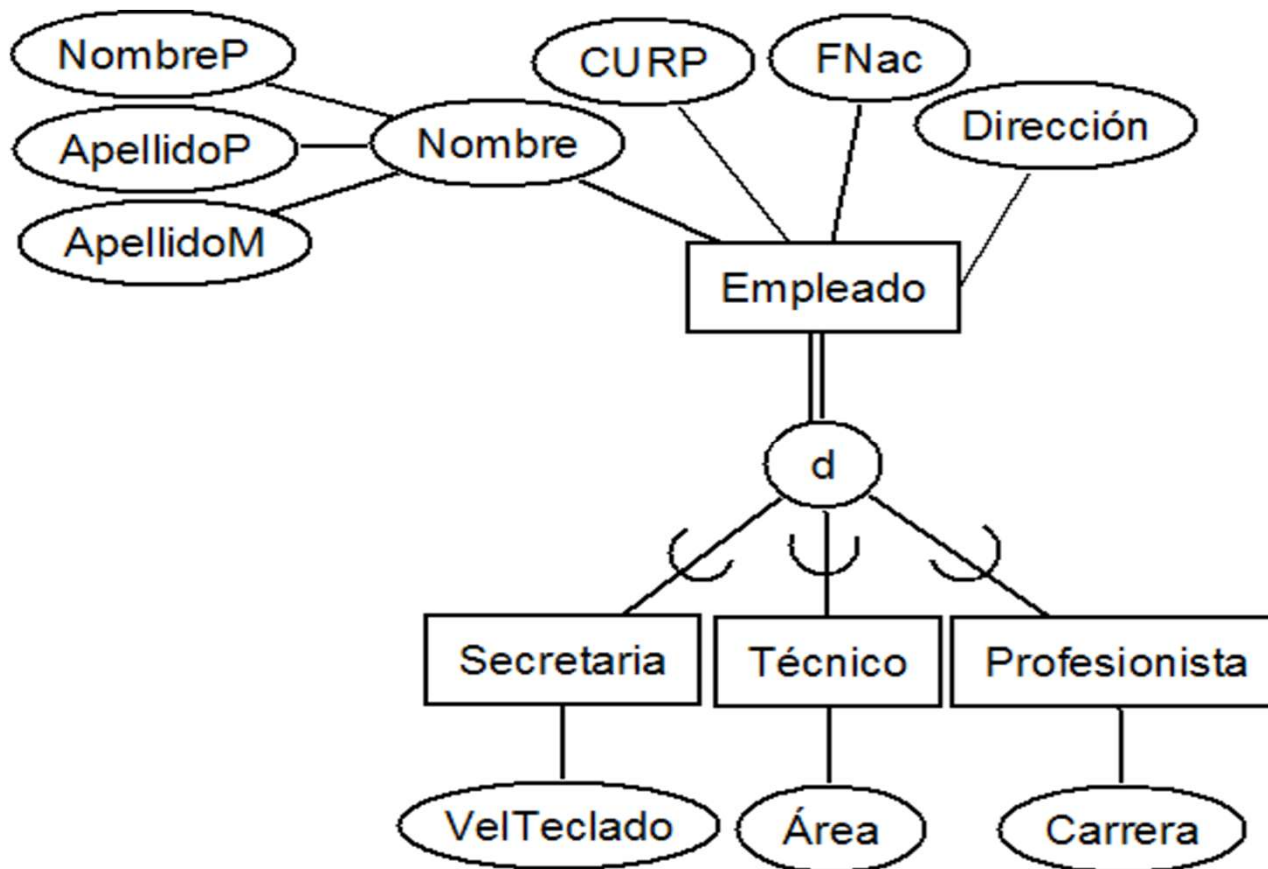
EMPLEADO: (CURP, NombreP, ApellidoP, ApellidoM, FNac, Dirección)

SECRETARIA: (CURP, VelTeclado)

TÉCNICO: (CURP, Área)

INGENIERO: (CURP, Carrera)

...Generalización/Especialización



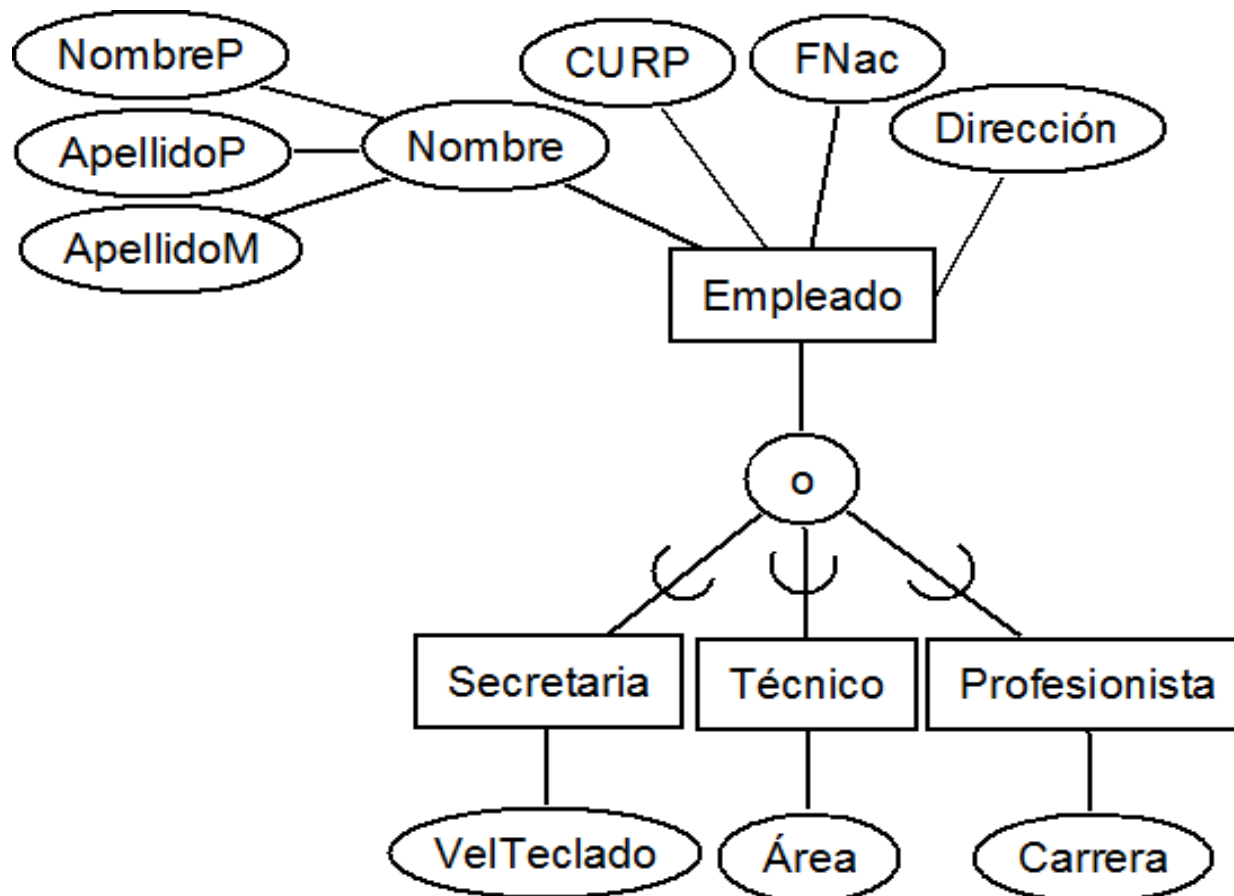
Se crea una tabla T_i para cada subentidad con sus atributos más los de la super entidad. La llave es la misma de la entidad super entidad.

SECRETARIA: (CURP, NombreP, ApellidoP, ApellidoM, FNac, Dirección, VelTeclado)

TÉCNICO: (CURP, NombreP, ApellidoP, ApellidoM, FNac, Dirección, Área)

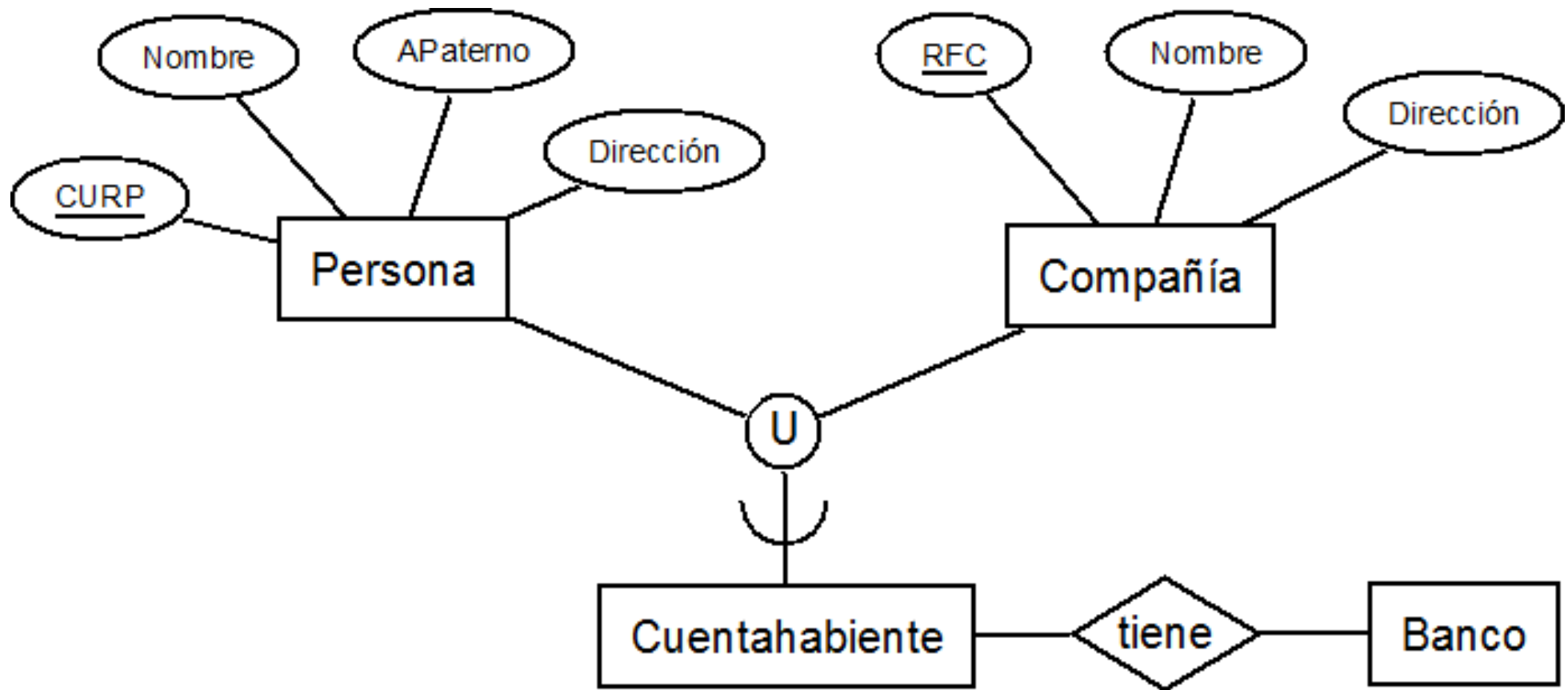
INGENIERO: (CURP, NombreP, ApellidoP, ApellidoM, FNac, Dirección, Carrera)

...Generalización/Especialización



Se crea una relación con un conjunto de atributos de tipo booleano para saber si se tiene ese tipo de la especialización.

EMPLEADO: (CURP, NombreP, ApellidoP, ApellidoM, FNac, Dirección, VelTeclado, Área, Carrera, EsSecre, EsTécnico, EsProfesionista)



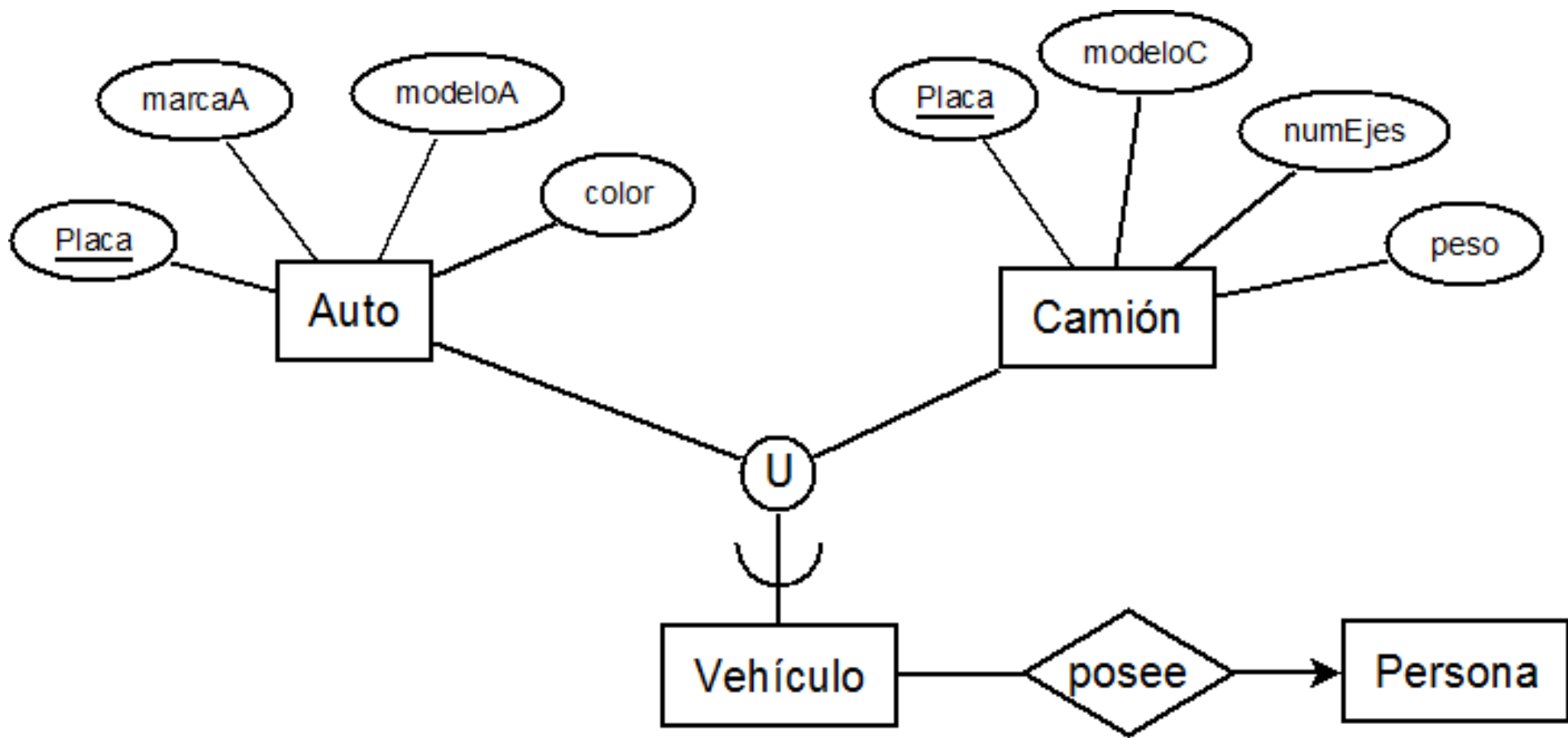
Si las superclases de la categoría tienen diferentes llaves primarias

- Se crea una relación R que corresponda a la categoría y se le asigna una llave sustituta arbitraria
- Se añade a la llave sustituta a modo de llave foránea en cada una de las relaciones R_i que corresponden a las superclases de la categoría.

Persona(**RFC**, Nombre, APaterno, Dirección, **IdCuentaHabiente**)

Compañía (**IdEmpleado**, Nombre, dirección, **IdCuentahabiente**)

Cuentahabiente (**idCuentahabiente**)



Si las superclases de la categoría tienen la misma llave primaria

- Se crea una relación R que corresponda a la categoría y se le asigna como atributo llave primaria la llave común a todas las superclases de la categoría.

Vehículo_Registrado(**NoPlaca**)

Auto (**NoPlaca**, marcaA, modeloA, color)

Camión (**NoPlaca**, modeloC, numEjes, peso)

Reglas de integridad

- ☐ Particulares de una BD específica
Edad entre 18 y 60
- ☐ Generales
 - Dominio
 - Llaves
 - Llaves externas



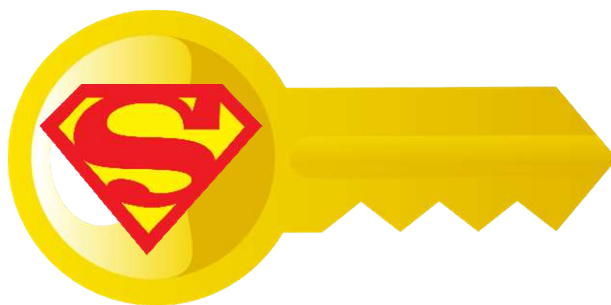
- ❑ **Dominios.** Todo atributo debe tomar un valor atómico de un dominio.
- ❑ **Valores nulos.** En ocasiones se debe asignar un valor para un componente de una tupla, pero no se puede decir cuál. En este caso se asigna el valor nulo **NULL**.

Existen diferentes interpretaciones para ellos:

- Valor desconocido
- Valor inaplicable
- Valor perdido.

En general, un atributo podría permitírsele o no tener valor nulo.

- Una **llave** es un conjunto **no vacío** de atributos que identifican de manera única a cada tupla.
- Una **superllave** es cualquier conjunto de atributos cuyos valores no se repiten en tuplas distintas de la relación



| <u>StudentId</u> | firstName | lastName | courseId |
|------------------|-----------|-----------|----------|
| L0002345 | Jim | Black | C002 |
| L0001254 | James | Harradine | A004 |
| L0002349 | Amanda | Holland | C002 |
| L0001198 | Simon | McCloud | S042 |
| L0023487 | Peter | Murray | P301 |
| L0018453 | Anne | Norris | S042 |

- Sea R una relación, entonces una **llave candidata** para R es un subconjunto del conjunto de atributos de R , digamos k tal que:
 - ☐ **Unicidad.** No existe par de tuplas distintas en R con el mismo valor para K .
 - ☐ **Irreductibilidad.** Ningún subconjunto propio de K tiene la propiedad de unicidad.
- Una llave candidata que involucra más de un atributo se llama **compuesta**, en otro caso se llama **simple**.



| <u>StudentId</u> | firstName | lastName | courseId |
|------------------|-----------|-----------|----------|
| L0002345 | Jim | Black | C002 |
| L0001254 | James | Harradine | A004 |
| L0002349 | Amanda | Holland | C002 |
| L0001198 | Simon | McCloud | S042 |
| L0023487 | Peter | Murray | P301 |
| L0018453 | Anne | Norris | S042 |

- Una **llave primaria** es una llave candidata elegida, casi siempre, arbitrariamente.



| <u>StudentId</u> | firstName | lastName | courseId |
|------------------|-----------|-----------|----------|
| L0002345 | Jim | Black | C002 |
| L0001254 | James | Harradine | A004 |
| L0002349 | Amanda | Holland | C002 |
| L0001198 | Simon | McCloud | S042 |
| L0023487 | Peter | Murray | P301 |
| L0018453 | Anne | Norris | S042 |

- Cuando una llave c en una relación R aparece como atributo en otra relación S , se dice que c es una **llave externa** en S .
- Sea R_2 una relación, se define como una **llave externa** en R_2 como un subconjunto FK, del conjunto de atributos de R_2 , tales que:
 - ❑ Existe una relación R_1 con una llave FK, y
 - ❑ Cada valor de FK en R_2 es idéntico al valor de FK en alguna tupla de R_1 .
 - ❑ Terminología:
 - ❑ El valor de una llave externa representa una referencia a la tupla que contiene el valor de la llave.
 - ❑ El problema de asegurar que la BD no incluye cualquier llave inválida se conoce como problema de **integridad referencial**.



Regla de integridad referencial

- La BD no debe contener valores de llave externa que no se correspondan con un valor de la llave candidata (Si B referencia a A, entonces A debe existir).
- Llave externa e integridad referencial, se definen una en términos de la otra.

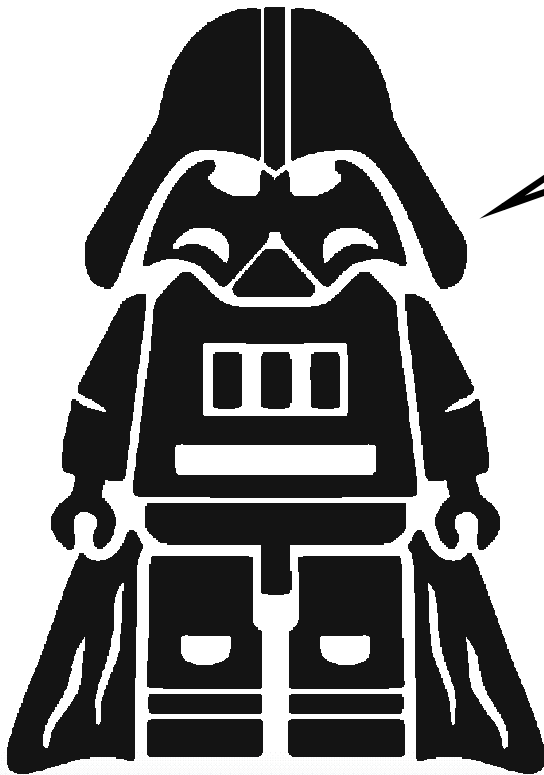
| <u>studentId</u> | firstName | lastName | courseId |
|------------------|-----------|-----------|----------|
| L0002345 | Jim | Black | C002 |
| L0001254 | James | Harradine | A004 |
| L0002349 | Amanda | Holland | C002 |
| L0001198 | Simon | McCloud | S042 |

Llave foránea

Relación

Llave primaria

| <u>courseId</u> | courseName |
|-----------------|--------------|
| A004 | Accounts |
| C002 | Computing |
| P301 | History |
| S042 | Short Course |



No estés muy orgulloso de haber comprendido estas notas.

La habilidad para manejar las bases de datos es insignificante comparado con el poder de la Fuerza.