

## Indice

<b>MODELIZACION CONCEPTUAL DE DATOS .....</b>	<b>2</b>
MODELOS DE DATOS: CONCEPTOS GENERALES .....	2
REPRESENTACIÓN SIMBÓLICA DE LA INFORMACIÓN.....	4
MODELOS DE DATOS .....	5
POTENCIA SEMÁNTICA .....	5
FILOSOFÍA DE LA MODELIZACIÓN CONCEPTUAL .....	6
<b>MODELO ENTIDAD-RELACIONAMIENTO .....</b>	<b>8</b>
ENTIDADES Y TIPO DE ENTIDADES .....	8
RELACIONAMIENTOS Y TIPOS DE RELACIONAMIENTOS.....	10
DIAGRAMAS E-R .....	12
RESTRICCIONES DE CARDINALIDAD DE CORRESPONDENCIA .....	13
RESTRICCIONES DE DEPENDENCIA EXISTENCIAL .....	16
CLAVES PRIMARIAS - RESTRICCIONES DE IDENTIFICACION.....	17
GENERALIZACION Y ESPECIALIZACION .....	19
AGREGACION.....	21
CASO DE ESTUDIO .....	23
Definición de tipos de entidades.....	24
Definición de tipos de relacionamientos.....	24
Identificación de atributos para tipos de entidades .....	28
Identificación de atributos para tipos de relacionamientos .....	29
Identificación de claves primarias para tipos de entidades .....	29
Identificación de claves primarias para tipos de relacionamientos.....	29
Diagrama ENTIDAD-RELACIONAMIENTO .....	30
EJERCITACION PROPUESTA MODELO ENTIDAD-RELACIONAMIENTO .....	31
<b>MODELO DE DATOS RELACIONAL.....</b>	<b>34</b>
INTRODUCCION.....	34
CONCEPTOS BASICOS DEL MODELO .....	34
DOMINIO .....	34
RELACION.....	34
ATRIBUTO.....	35
ESQUEMA DE RELACION .....	35
SUPERCLAVE, CLAVE CANDIDATA, CLAVE PRIMARIA, CLAVE FORANEA.....	36
SUPERCLAVE .....	36
CLAVE CANDIDATA: UNICIDAD - MINIMALIDAD .....	36
CLAVE PRIMARIA .....	36
CLAVE FORANEA .....	36
MODELO RELACIONAL: RESTRICCIONES .....	36
REGLA DE INTEGRIDAD DE ENTIDAD .....	37
REGLA DE INTEGRIDAD REFERENCIAL.....	37
ALGEBRA RELACIONAL.....	37
SELECCIÓN o RESTRICCION .....	37
PROYECCION .....	38
UNION .....	39
DIFERENCIA .....	40
PRODUCTO CARTESIANO .....	40
OPERACIONES ADICIONALES.....	41
JOIN .....	41
INTERSECCION .....	42
DIVISION .....	42
COMPOSICION DE RELACIONES.....	44
EJERCITACION PROPUESTA ALGEBRA RELACIONAL .....	45
PRACTICA ADICIONAL DE ALGEBRA RELACIONAL:.....	46
<b>APENDICE .....</b>	<b>50</b>
TRANSFORMACION DEL MODELO E-R A OTROS MODELOS IMPLEMENTABLES .....	50
TRANSFORMACION DE UN MODELO E-R A TABLAS .....	50
TIPOS DE ENTIDADES REGULARES .....	50
TIPOS DE ENTIDADES DÉBILES .....	51
TIPOS DE RELACIONAMIENTO .....	51
GENERALIZACION Y ESPECIALIZACION .....	52

## MODELIZACION CONCEPTUAL DE DATOS

### **MODELOS DE DATOS: Conceptos Generales**

Para que los datos sean útiles proveyendo información, deben ser organizados de manera tal que puedan ser procesados efectivamente.

En la modelización de datos tratamos de organizar los datos de manera que representen lo más cercanamente posible la situación del mundo real, siendo dicha representación manejable por computadoras; en general, estos dos requerimientos resultan conflictivos.

Dado que no es posible tener un conocimiento completo del mundo real, debemos centrar nuestra atención en el significado de los datos apropiado para su utilización, ocultando o desconociendo otros significados que resulten irrelevantes o inadecuados para el propósito requerido.

Se trata, entonces, de comprender las características de los datos que son importantes para capturar la esencia de su significado. Estas características permiten establecer cómo van a ser organizados y procesados, y pueden ser descriptas a través de enunciaciones generales. Un conjunto formal y consistente de tales enunciaciones define un modelo de datos.

En términos generales, podemos decir que un **MODELO DE DATOS** es *una herramienta intelectual que provee una interpretación del mundo real*, lo suficientemente poderosa para ayudar a entender cómo están relacionados los datos, y lo suficientemente *abstracta*, para ser mínimamente perturbable ante los cambios concernientes al aspecto evolutivo del mundo real.

Es un *dispositivo de abstracción*, que nos permite ver el contenido de información de los datos, en lugar de sus valores individuales.

Los MODELOS son usados extensivamente en muchas disciplinas para mejorar la comprensión y abstraer los detalles, entre otros podemos mencionar MODELOS FÍSICOS, MATEMÁTICOS, ECONÓMICOS, etc. Nuestro tema de interés se limita exclusivamente a los MODELOS DE DATOS, los cuales pueden ser codificados y manipulados fácilmente por computadoras.

Si intentamos determinar, al menos aproximadamente, qué es un dato, podemos suponer como una definición de trabajo que una pieza atómica de dato es la siguiente tupla:

< nombre de objeto, propiedad del objeto, valor de la propiedad del objeto, instante >

dado que un fenómeno o idea, usualmente refiere a *un objeto* y a algún *aspecto del objeto*, el cual captura un determinado *valor* en un cierto *momento*.

Sin embargo, el tiempo real implica cierta sincronización entre los fenómenos que no es del todo representable, en cambio nos interesa a menudo el tiempo relativo del fenómeno, lo cual puede ser adecuadamente capturado, ordenando los fenómenos antes que registrando el tiempo real. Podríamos entonces dejar de lado la noción de tiempo y reemplazarla por otras propiedades explícitas que permitan establecer el ordenamiento entre objetos.

Entonces nuestra definición de *dato elemental* se convertiría en la terna:

**< nombre de objeto, propiedad del objeto, valor de la propiedad del objeto >**

Hay muchas maneras en que estos aspectos de los datos pueden ser representados y relacionados, dando lugar a distintos MODELOS DE DATOS.

Una manera simple y potente, consiste en categorizar los datos, de modo tal que datos puestos en la misma categoría manifiesten similitudes; en consecuencia podemos establecer que esas similitudes constituyen propiedades de la categoría y, por lo tanto, son heredables por los datos que pertenecen a la categoría. Hacemos una clara distinción entre *datos* y *categorías de datos*.

Una desventaja de esta clase de MODELOS DE DATOS, que damos en llamar *tipificados estrictamente* (*strictly typed*), es que obligan a “encajar el mundo real” de modo que cada dato caiga dentro de una categoría; esta falta de flexibilidad, algunas veces, presenta dificultades para representar diferencias semánticas sutiles, o, en el peor de los casos, resultan datos del mundo real que no caben en categoría alguna y, por lo tanto, no pueden ser manejados en el MODELO DE DATOS.

Por otra parte, podemos justificar esta imposición, teniendo en cuenta la cantidad de objetos a tratar y las limitaciones de la inteligencia humana. Dado que en la modelización de datos se debe tratar con muchos objetos, es necesario cierto marco restrictivo que los comprenda en un conjunto limitado de ideas que resulten más fáciles para el entendimiento.

Concluyendo, la gran ventaja, entonces, es que las propiedades de los datos pueden ser abstraídas e investigadas en términos de sus categorías. Es decir, una teoría puede ser formulada en base a categorías que comprendan las propiedades de los datos en sí; luego, probar algo acerca de las categorías implica haberlo probado acerca de los datos.

## **REPRESENTACIÓN SIMBÓLICA DE LA INFORMACIÓN**

*Conocimiento* es el entendimiento de la naturaleza, cualidades y relaciones de las cosas. Cuando percibimos u observamos un fenómeno, obtenemos una porción incremental de conocimiento. Denominamos *información* a estos incrementos de conocimiento.

La información valiosa debe ser registrada en una forma simbólica inteligible para que pueda ser comunicada a otras personas. Para estos fines resulta de importancia fundamental contar con una adecuada representación simbólica de la información.

La pieza más elemental de información es el *dato* (ej., una cifra de sueldo). Un dato no es de utilidad por sí mismo, si no está vinculado con otros datos (ej., la identificación de un empleado), por lo que el almacenamiento y la comunicación de información requieren la estructuración de datos. La estructura y el contexto le dan *significado* a los datos posibilitando su entendimiento.

El lenguaje natural es nuestra forma primaria de representación y comunicación de información. Si bien es el medio más generalizado para ese propósito, no siempre constituye el mejor medio. En muchas situaciones resulta más útil establecer formas especializadas de representación de la información. Las fórmulas matemáticas, los mapas carreteros y las partituras musicales, constituyen ejemplos de *modelos* especializados de representación simbólica de la información.

Los datos de una base de datos se almacenan y procesan en un computador por medio de Sistemas de Gestión de Bases de Datos (SGBD o DBMS -Data Base Management Systems-). Se han desarrollado diferentes Modelos de Datos (Relacional, Jerárquico, Red) para la representación de la información en los sistemas computarizados. En este contexto, un Modelo de Datos describe la forma de representación de la información utilizada por un SGBD en particular. Así, decimos que un SGBD Relacional (RDBMS) responde a un Modelo de Datos Relacional.

Estos Modelos de Datos, que describen cómo se almacenan los datos en el computador, resultan inadecuados para representar el *significado* de los datos almacenados. Por este motivo han surgido metodologías de diseño de bases de datos sustentadas en la utilización de Modelos de Datos de mayor potencia semántica orientados a la descripción de la realidad que estará representada en las bases de datos. Estos Modelos, que se denominan *modelos conceptuales* o *semánticos*, tienden a describir la realidad captando en mayor medida el significado de los datos y de sus vinculaciones en el mundo real.

## MODELOS DE DATOS

Podemos decir que un MODELO DE DATOS es una herramienta formal para representar la realidad.

Los Modelos de Datos en general, especifican:

- la **estructura** de los datos.
- **restricciones** sobre las estructuras.
- **operaciones** sobre las estructuras.

Un Modelo de Datos establece reglas para la estructuración de los datos (ej, tablas, árboles, redes, etc.), permite establecer restricciones sobre las estructuras para que los datos mantengan su coherencia informativa o *integridad semántica* y define la naturaleza de las operaciones permitidas sobre los datos, dentro del contexto provisto por las estructuras.

Las restricciones pueden abarcar una gran variedad de aspectos, contribuyendo a la descripción más ajustada de la realidad, a través de la exclusión de lo imposible. Por ejemplo, pueden especificarse restricciones provenientes del mundo real tales como:

- los empleados no pueden ganar más que sus jefes.
- para inscribirse en un curso, un estudiante debe satisfacer ciertos requisitos.
- la cantidad de pacientes internados en un hospital no puede exceder la cantidad de camas disponibles.

Por otra parte, además de esta clase de *restricciones* que llamamos *explícitas*, un Modelo de Datos introduce ciertas *restricciones inherentes* como consecuencia de sus propias características y limitaciones de modelización.

## POTENCIA SEMÁNTICA

La *potencia semántica* de un Modelo de Datos está determinada por su capacidad de representar, además de la estructura de los datos, *el significado de los mismos y de sus interrelaciones*. Por ejemplo, en la frase “Su altura es de 5 metros”, el dato es el número “5” y el significado es “altura en metros”.

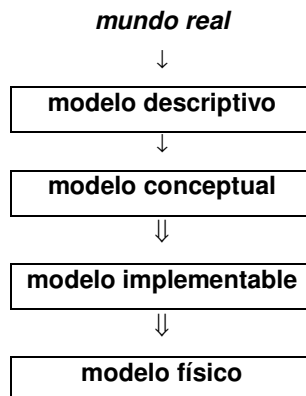
En general, la utilización del computador para el procesamiento de información ha resultado en una separación entre los datos y su significado. Mientras que las computadoras sólo procesan los datos, queda para el usuario o los programas, la tarea de interpretación. A medida que la utilización de computadoras evoluciona hacia aplicaciones más sofisticadas, resulta cada vez más necesario captar y utilizar activamente el significado de los datos (y de sus vinculaciones).

Consecuentemente, es necesario advertir que al traducir el *modelo conceptual* al modelo de datos del SGBD que se utilizará, inevitablemente habrá pérdida de información semántica.

## FILOSOFÍA DE LA MODELIZACIÓN CONCEPTUAL

La siguiente figura pretende reflejar el proceso mental mediante el cual pasamos de una realidad informacional a una representación asimilable por un computador. Es claro que deben existir varios niveles de abstracción entre el usuario final que trata con *objetos* como: vuelos, pilotos, aviones o pasajeros y el computador, que almacena e interpreta bits.

El mundo real	compañía de fabricación; líneas de transportes; entidades bancarias; servicios de telecomunicaciones; ciencias; arte; deportes ...
Modelos conceptuales	Modelos infológicos Modelos semánticos Modelo entidad-relacionamiento Modelo binario
Modelo de bases de datos	Modelo relacional Modelo CODASYL (en red) Modelo jerárquico Modelos híbridos
Estructuras de datos	árboles equilibrados; archivos secuenciales indexados; pilas; colas; listas invertidas
Almacenamiento físico	bits; bytes; código ASCII; EBCDIC



Algunas misceláneas a considerar en el proceso de modelización conceptual:

- Una base de datos es una colección de hechos registrados que reflejan el estado de ciertos aspectos de interés del mundo real. En todo momento el contenido (*extensión*) de la base de datos representa una instantánea del estado del sistema real. Cada cambio en un valor almacenado en la base de datos refleja un cambio ocurrido o planificado en el sistema real.
- Los Modelos de Datos permiten diseñar la estructura (*intensión o esquema*) de una base de datos, de manera que refleje la estructura del sistema que la misma modeliza.
- Un modelo de datos expresa la estructura y las reglas de gestión de un sistema. Por ejemplo, la estructura organizativa y las reglas de funcionamiento de un hospital.
- Un modelo de datos resulta definido por el sistema que representa.
- Como consecuencia de las consideraciones anteriores, se puede decir que la estructura y las reglas de funcionamiento (de gestión) del sistema real determinan objetivamente la estructura de la base de datos, con independencia del diseñador. El diseñador se limita a decidir el alcance (amplitud) y la granularidad (grado de detalle) del modelo y a modelizar conceptualmente el mundo real. La función del diseñador no es crear una estructura de información, sino descubrir y explicitar aquella que está implícita en la realidad. Del modelo conceptual resultante, se deriva objetivamente la estructura de la base de datos.
- Un ambiente de aplicación (un banco, un hospital, una universidad, una industria, etc.) es un sistema realimentado por información que evoluciona con el tiempo. Es posible identificar en él múltiples flujos (dinero, materiales, personas, equipos, órdenes, etc.) siendo el de información el que interconecta y posibilita el funcionamiento del sistema completo.
- Dentro de un ambiente de aplicación cada grupo, actividad, proceso, persona, función y sistema tiene su propia percepción dinámica de la realidad, su propio vocabulario y su propia valoración de lo que es importante. La empresa típica es una composición de estas visiones.
- El resultado del diseño conceptual es un modelo de datos conceptual construido desde el punto de vista de los usuarios y de la aplicación, con independencia de cómo los datos van a ser almacenados y procesados por el computador. En una etapa posterior de *diseño lógico* se traduce el modelo conceptual al modelo de datos que corresponde al SGBD a utilizar, adaptándolo a dicho SGBD y a los aspectos de procesamiento (procesos, frecuencias, tiempos de respuesta, prioridades, etc.). Finalmente, en la etapa de *diseño físico* se realiza la adaptación al computador contemplando las características y limitaciones del mismo.

## MODELO ENTIDAD-RELACIONAMIENTO

### ENTIDADES Y TIPO DE ENTIDADES

Una *Entidad* es algo que existe (concreto o abstracto) y es distinguible de algo similar. Por ejemplo, Juan Cancela con cédula de identidad 5.789.231 es una entidad pues identifica a una persona en particular del universo. La cuenta número 115 de la sucursal Boedo del Banco de la Nación es una entidad pues identifica unívocamente a una cuenta en particular. Una entidad puede ser algo concreto (una persona, un libro) o abstracto (un concepto).

Un *Conjunto de Entidades* (o *Tipo de Entidad*<sup>1</sup>) es un conjunto de entidades del mismo tipo. El conjunto de todas las personas que tienen cuentas en un banco, por ejemplo, puede definirse como el conjunto de entidades CLIENTE o el tipo de entidad CLIENTE. El conjunto de entidades CUENTA o el tipo de entidad CUENTA puede representar el conjunto de todas las cuentas de un banco en particular.

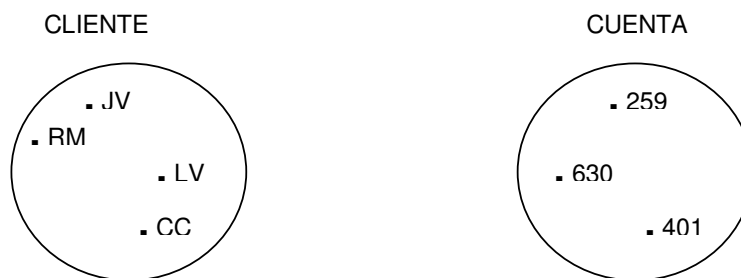


Fig. 1

Existe un predicado asociado a cada conjunto de entidades para probar si una entidad pertenece o no al conjunto. Matemáticamente un conjunto de entidades puede representarse como  $E = \{ e / p(e) \}$ , donde **E** es el conjunto de entidades **e** que satisfacen el predicado **p**.

Los conjuntos de entidades no necesitan ser disjuntos. Por ejemplo, es posible definir el conjunto de entidades de todos los empleados de un banco (EMPLEADO) y el conjunto de entidades de todos los clientes del banco (CLIENTE). Una entidad PERSONA puede ser una entidad EMPLEADO, una entidad CLIENTE, ambas, o ninguna.

<sup>1</sup> En la literatura se utilizan diferentes nombres para este concepto: conjunto de entidades (entity set), tipo de entidad (entity type), entidad-tipo, entidad genérica. Nosotros usaremos conjunto de entidades o tipo de entidad indistintamente.



Un Tipo de Entidad queda definido por:

- a) un NOMBRE
  - b) un SIGNIFICADO
  - c) ATRIBUTOS
  - d) un IDENTIFICADOR
- a) El NOMBRE debe ser:
- conciso: nombre en singular (Ej.: CLIENTE) o frase simple en singular (CUENTA CORRIENTE)
  - adecuado: de uso común en la organización y que no excluya ciertas ocurrencias
  - único: dentro del modelo evitar sinónimos y homónimos
- b) El SIGNIFICADO o definición debe ser expresado a través de un texto conciso, preciso y claro, indicando, si existieran, dependencias con otros tipos de entidades
- c) Un conjunto de ATRIBUTOS que caracterizan a todas las entidades de ese tipo.
- NOMBRE, CEDULA DE IDENTIDAD, CALLE y CIUDAD son posibles atributos del conjunto de entidades CLIENTE.
  - NUMERO DE CUENTA y SALDO son posibles atributos del conjunto de entidades CUENTA.

Para cada atributo existe un conjunto de valores permitidos, denominado el dominio de ese atributo:

- El dominio del atributo NOMBRE puede ser el conjunto de todos los nombres permitidos de personas.
- El dominio del atributo NUMERO DE CUENTA puede ser el conjunto de todos los números de cuenta permitidos en ese banco.

Formalmente un atributo es una función que hace corresponder un conjunto de entidades a un dominio o producto cartesiano de dominios. Así, toda entidad está descrita por un conjunto de pares (atributo, valor); un par por cada atributo del conjunto de entidades. Por ejemplo, una entidad EMPLEADO en particular, está descrita por el conjunto {(nombre, Juan), (apellido, Cancela), (fecha de ingreso, (2, enero, 1990)), (cédula de identidad, 5.789.231)}. Esto significa que esta entidad describe a un empleado llamado Juan Cancela con cédula de identidad número 5.789.231, que ingresó a la empresa el 2/1/1990 (ver figura 2).

Deben evitarse (o marcarse) los atributos derivables, esto es, atributos cuyos valores pueden calcularse a partir de otros atributos pertenecientes al mismo o a diferentes tipos de entidades. Por ejemplo un tipo de entidad FACTURA con atributos PRECIO UNITARIO y CANTIDAD, podría tener como atributo derivable IMPORTE TOTAL.

- d) El IDENTIFICADOR es un atributo o conjunto de atributos que permiten distinguir a cada entidad dentro del conjunto de entidades del mismo tipo. (ver más adelante CLAVES PRIMARIAS)

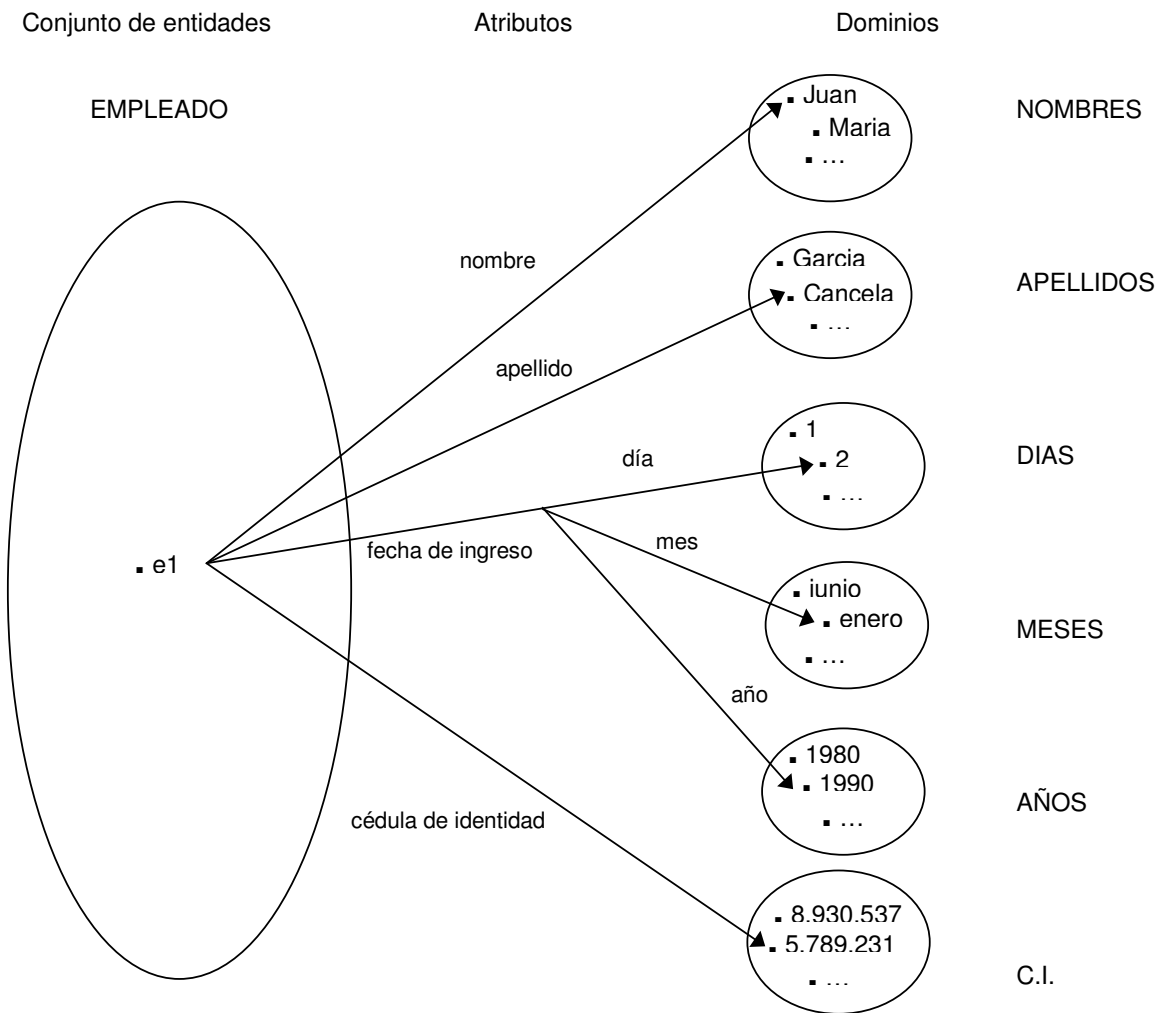


Fig. 2

## RELACIONAMIENTOS Y TIPOS DE RELACIONAMIENTOS

Un *Relacionamiento* es una asociación entre dos o más entidades. Por ejemplo, el relacionamiento que asocia el cliente "Pérez" con la cuenta 259.

Un *Conjunto de Relacionamientos* (o *Tipo de Relacionamiento*<sup>2</sup>) es un conjunto de relacionamientos del mismo tipo. Formalmente es una relación matemática sobre  $n \geq 2$  conjuntos (no necesariamente disjuntos) de entidades. Si  $E1, E2, \dots, En$  son conjuntos de entidades, entonces el conjunto de relacionamientos  $R$  es un subconjunto del producto cartesiano

$$\{(e1, e2, \dots, en) / e1 \in E1, e2 \in E2, \dots, en \in En\},$$

donde la tupla  $(e1, e2, \dots, en)$  es un relacionamiento.

<sup>2</sup> Es frecuente emplear el término relación para denominar el concepto que nosotros hemos llamado relacionamiento. La palabra en inglés es *relationship*. Hemos preferido "relacionamiento" para evitar confusión con las "relaciones" del Modelo Relacional. Usaremos conjunto de relacionamientos o tipo de relacionamiento indistintamente.

Consideremos los dos tipos de entidades **CLIENTE** y **CUENTA**. Definimos el tipo de relacionamiento **TENENCIA** que denomina la asociación entre los clientes y sus cuentas bancarias.

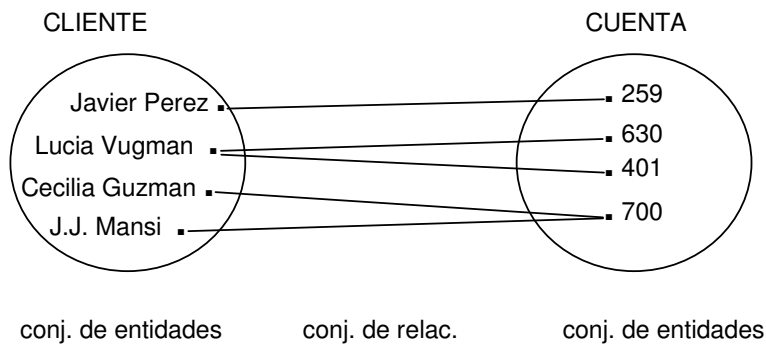
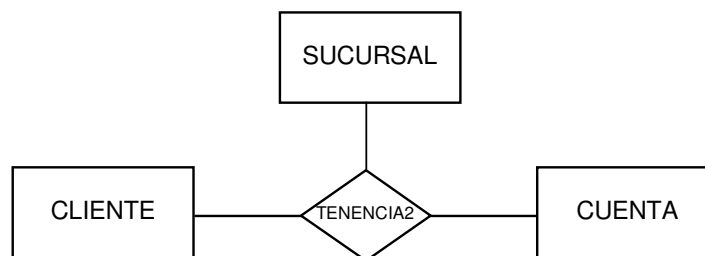


Fig. 3



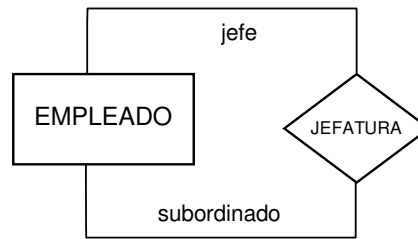
**TENENCIA** es un ejemplo de tipo de relacionamiento binario, esto es, un tipo de relacionamiento que involucra dos tipos de entidades.

Grado de un tipo de relacionamiento es el número de tipos de entidades que relaciona. **TENENCIA** es un tipo de relacionamiento de grado **dos**. Existen tipos de relacionamientos que involucran más de dos tipos de entidades. Como ejemplo, considérese el relacionamiento ternario (Pérez, 259, Boedo) que especifica que el cliente Pérez tiene la cuenta 259 en la sucursal Boedo. Este relacionamiento es una instancia del tipo de relacionamiento **TENENCIA2** que involucra los tipos de entidades **CLIENTE**, **CUENTA** y **SUCURSAL**.



La función que una entidad desempeña en un relacionamiento se denomina su *rol*. Normalmente los roles están implícitos y no son especificados. Al hablar del tipo de relacionamiento **TENENCIA** sobreentendemos el rol que juega cada tipo de entidad: un **CLIENTE** "tiene" una **CUENTA**, una **CUENTA** "es (man)tenida por" un **CLIENTE**. Sin embargo, puede resultar útil especificar los roles para esclarecer el significado del tipo de relacionamiento. Tal es el caso cuando los tipos de entidades que participan en un tipo de relacionamiento no son distintos. Por ejemplo, las relaciones

de subordinación/jefatura que existen entre los empleados de una empresa pueden representarse con el tipo de relacionamiento JEFATURA.



Un EMPLEADO "es jefe de"  $n$  EMPLEADOs, un EMPLEADO "es subordinado de" un EMPLEADO (su jefe). Cada relacionamiento del tipo JEFATURA puede modelizarse por un par ordenado de entidades EMPLEADO. Por ejemplo, el primer empleado del par toma el rol de subordinado, y el segundo toma el rol de jefe. De esta manera, todos los relacionamientos JEFATURA están caracterizados por pares (subordinado, jefe). Los pares (jefe, subordinado) están excluidos. JEFATURA es un tipo de relacionamiento grado **uno**.

También un tipo de relacionamiento puede tener atributos descriptivos. Por ejemplo, FECHA ULTIMO MOVIMIENTO puede ser un atributo del tipo de relacionamiento TENENCIA, indicando la última fecha en la que un cliente ha accedido su cuenta. El relacionamiento (Pérez, 259) estará descrito por {(fecha, 23 de mayo de 1998)} lo que significa que la ultima vez que Pérez accedió a la cuenta 259 fue el 23 de mayo de 1998.

## DIAGRAMAS E-R

Un Modelo E-R consiste en una colección de tipos de entidades y tipos de relacionamientos más los atributos, dominios y restricciones necesarios para describir las propiedades que interesan de esos conjuntos. Una forma gráfica de representar un modelo de este tipo es mediante un diagrama E-R que utiliza la siguiente simbología:

- **rectángulos**      que representan tipos de entidades
- **óvalos**            que representan atributos
- **rombos**            que representan tipos de relacionamientos
- **arcos**                que asocian atributos a los tipos de entidades y tipos de entidades a tipos de relacionamientos

Cada componente está rotulado con su correspondiente nombre. Los roles se indican rotulando los arcos que conectan los rectángulos con los rombos.

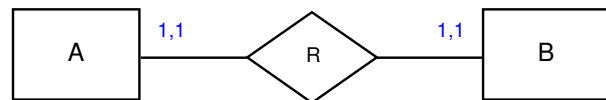
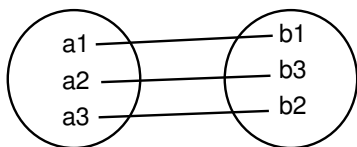
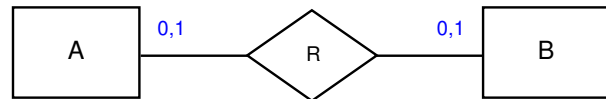
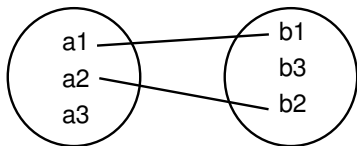
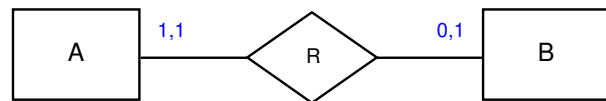
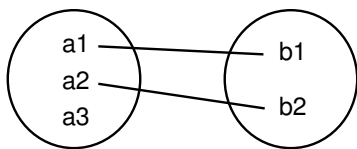
## RESTRICCIONES DE CARDINALIDAD DE CORRESPONDENCIA

Las cardinalidades de correspondencia son restricciones importantes que expresan la cantidad (mínima y máxima) de entidades a las cuales puede asociarse una entidad a través de un tipo de relacionamiento. En esta sección nos ocuparemos de las cardinalidades de correspondencia entre tipos de relacionamientos binarios (es decir, relacionamientos entre dos tipos de entidades).

Para un tipo de relacionamiento binario **R** entre dos tipos de entidades **A** y **B**, la cardinalidad puede ser una de las siguientes:

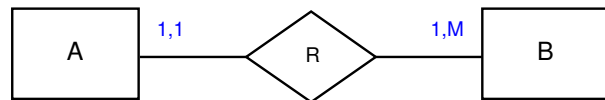
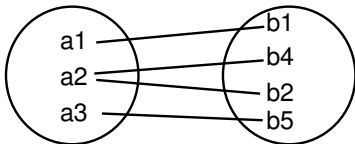
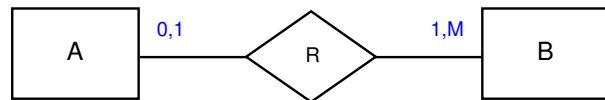
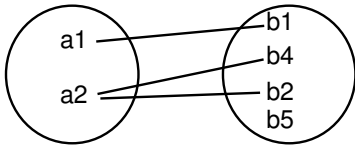
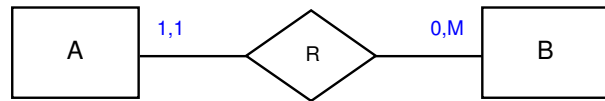
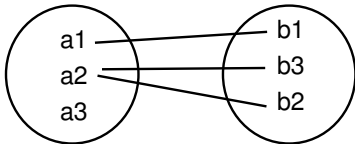
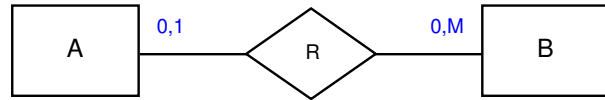
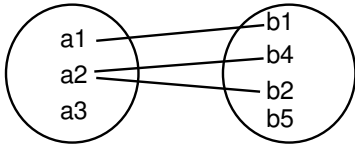
- **UNO A UNO (1:1)**

Una entidad en **A** está asociada como máximo a una entidad en **B**, y una entidad en **B** está asociada como máximo a una entidad en **A**.



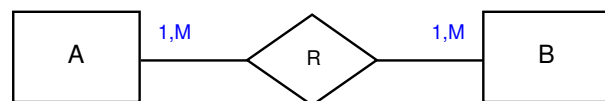
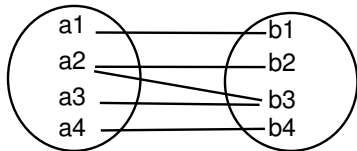
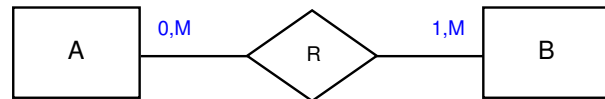
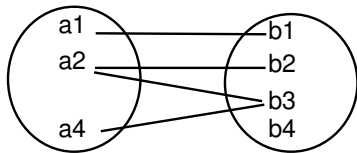
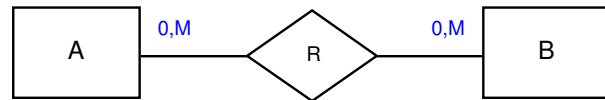
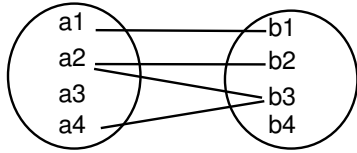
▪ **UNO A MUCHOS (1:M)**

Una entidad en **A** está asociada como máximo con cualquier cantidad de entidades en **B**, y una entidad en **B** está asociada como máximo a una entidad en **A**.



▪ **MUCHOS A MUCHOS (M:M)**

Una entidad en **A** está asociada como máximo con cualquier cantidad de entidades en **B**, y una entidad en **B** está asociada como máximo con cualquier cantidad de entidades en **A**.



La cardinalidad de correspondencia apropiada para un conjunto de relacionamientos particular depende, obviamente, del mundo real que está siendo modelizado. Considérese el tipo de relacionamiento TENENCIA. Si en un banco en particular una cuenta puede pertenecer a un cliente solamente, y un cliente puede tener varias cuentas, entonces el tipo de relacionamiento es **uno-a-muchos** de CLIENTE a CUENTA.



Si una cuenta puede pertenecer a varios clientes en forma conjunta, el tipo de relacionamiento es **muchos-a-muchos**.

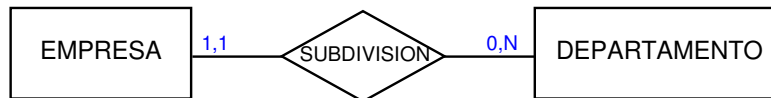


Veamos otros ejemplos:

- a) Todo empleado tiene asignado un escritorio para trabajar, en cada escritorio trabaja un solo empleado, no todos los escritorios tienen empleados asignados.



- b) Las empresas están divididas en departamentos, todo departamento pertenece a una empresa exclusivamente, no todas las empresas están formadas por departamentos.



- c) Todo empleado tiene habilidades, varios empleados pueden poseer la misma habilidad, pueden existir habilidades que no sean poseídas por ningún empleado.



### RESTRICCIONES DE DEPENDENCIA EXISTENCIAL

Las dependencias de existencia conforman otra clase importante de restricciones. Específicamente, si la existencia de la entidad **S** depende de la existencia de la entidad **D**, entonces se dice que **S** tiene dependencia existencial o depende existencialmente de **D**. Operacionalmente, esto significa que si **D** es borrado, también **S** es borrado. La entidad **D** es la entidad *dominante* y **S** es la entidad *subordinada*.

Considérese los tipos de entidades CUENTA y TRANSACCION, y el tipo de relacionamiento MOVIMIENTO entre éstos, que especifica que para una cuenta particular pueden existir varias transacciones. Este tipo de relacionamiento es **uno-a-muchos** de CUENTA a TRANSACCION. Cada entidad TRANSACCION debe estar asociada a una entidad CUENTA. Si una entidad cuenta es borrada, entonces todas las entidades TRANSACCION asociadas también deben ser borradas. Sin embargo, las entidades TRANSACCION pueden ser borradas de la base de datos sin afectar ninguna entidad CUENTA. El tipo de entidad CUENTA, por lo tanto, es *dominante* y el tipo de entidad TRANSACCION es *subordinado* del tipo de relacionamiento MOVIMIENTO.





## **CLAVES PRIMARIAS - RESTRICCIONES DE IDENTIFICACION**

Conceptualmente las entidades y relacionamientos individuales son distinguibles entre sí. Una **superclave** es un conjunto de uno a más atributos que, tomados colectivamente, permiten identificar unívocamente una entidad (o un relacionamiento) en el conjunto de entidades (o relacionamientos) del mismo tipo. Por ejemplo, el atributo NOMBRE DE CLIENTE del tipo de entidad CLIENTE no es una superclave, pues varios clientes pueden tener el mismo nombre. El atributo CEDULA DE IDENTIDAD del tipo de entidad CLIENTE es suficiente para distinguir una entidad CLIENTE de otra. Por lo tanto CEDULA DE IDENTIDAD es una superclave. Asimismo la combinación NOMBRE DE CLIENTE y CEDULA DE IDENTIDAD es una superclave para el tipo de entidad CLIENTE. Si **K** es una superclave, cualquier superconjunto de **K** también es una superclave.

Interesan generalmente las menores superclaves posibles. Esto es, estamos interesados en aquellas superclaves para las cuales ningún subconjunto propio es una superclave. Estas superclaves minimales se denominan **claves candidatas** o identificadores candidatos.

Vemos que aunque los atributos NOMBRE DE CLIENTE y CEDULA DE IDENTIDAD juntos son una superclave (pues permitirían distinguir entidades CLIENTE), su combinación no constituye una clave candidata porque el atributo CEDULA DE IDENTIDAD por sí solo es una superclave.

Es posible que varios conjuntos distintos de atributos puedan servir como claves candidatas. El atributo D.N.I., por ejemplo, podría ser suficiente para distinguir entre miembros del conjunto de entidades CLIENTE. Así, tanto {CEDULA DE IDENTIDAD} como {D.N.I.} son claves candidatas.

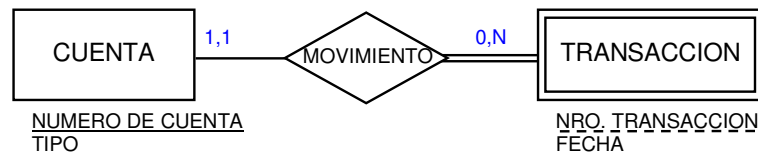
Usaremos el término **clave primaria** o identificador principal para denominar aquella clave candidata que es elegida, por el diseñador del modelo de datos, como medio principal para identificar entidades dentro de un conjunto de entidades.

Puede ocurrir que en un tipo de entidad ningún atributo (o conjunto de atributos) satisfaga las condiciones de identificador principal, esto es, permitir la identificación unívoca de una entidad dentro del conjunto. Por ejemplo, si para cada CUENTA las TRANSACCIONES se numeran secuencialmente desde 1, no será posible identificar por NRO.TRANSACCION una transacción dentro del conjunto de transacciones, diferentes transacciones sobre diferentes cuentas pueden tener el mismo número.

Se dice que una entidad tiene una *dependencia de identidad o identificación* si no es identificable por sus propios atributos. Este tipo de entidad que no tiene una clave primaria propia se denomina **entidad débil**, y normalmente está subordinada a otro tipo de entidad, que denominamos entidad fuerte, de la cual depende su identificación. Una entidad que tiene clave primaria propia es una **entidad fuerte**.

El concepto de entidades fuertes y débiles está relacionado con el concepto de *dependencia de existencia*. Una entidad débil es, por definición, una entidad *subordinada* a la entidad *dominante* de la cual depende su identidad. Toda *dependencia de identidad* es una *dependencia de existencia*, pero una *dependencia de existencia* no implica necesariamente una *dependencia de identidad*.

Un tipo de entidad débil se indica con un rectángulo doble. El tipo de relacionamiento que lo conecta con el tipo de entidad fuerte del cual depende se indica con arcos dobles.



Aunque un tipo de entidad débil no tiene clave primaria propia, es necesario distinguir, entre sí, las entidades débiles que dependen de una misma entidad fuerte. El **discriminador** de un tipo de entidad débil es el conjunto de atributos que permiten realizar esta distinción. Por ejemplo, el discriminador del tipo de entidad débil TRANSACCION es el atributo NRO. TRANSACCION, dado que para cada cuenta estos números de transacción identifican unívocamente una única transacción.

La clave primaria de un tipo de entidad débil está formada por la clave primaria de la entidad fuerte de la cual depende existencialmente, más su discriminador. En el caso del tipo de entidad TRANSACCION, su clave primaria es (NUMERO DE CUENTA, NRO. TRANSACCION), donde NUMERO DE CUENTA identifica la entidad dominante de una TRANSACCION, y NRO. TRANSACCION distingue las entidades TRANSACCION dentro de la misma cuenta.

Los tipos de relacionamiento también tienen claves primarias. En general, aunque no siempre, sus claves primarias están formadas por los atributos de las claves primarias de los tipos de entidad que asocian. Por ejemplo, CEDULA DE IDENTIDAD es la clave primaria de CLIENTE, y NUMERO DE CUENTA es la clave primaria de CUENTA. La clave primaria del tipo de relacionamiento TENENCIA es (CEDULA DE IDENTIDAD, NUMERO DE CUENTA). La clave primaria del tipo de relacionamiento MOVIMIENTO es (NUMERO DE CUENTA, NRO. TRANSACCION). Obsérvese que en este último caso la clave primaria del tipo de relacionamiento coincide, trivialmente, con la del tipo de entidad débil; en general esto sucede con cada tipo de relacionamiento binario *funcional* (cardinalidades 1:M o 1:1), en que la condición de *minimalidad* determina que su clave primaria coincida con la de uno de los tipos de entidad que asocia, según el sentido de la *funcionalidad* de la relación (se verán ejemplos en el Caso de Estudio).

Puede darse el caso de un tipo de relacionamiento que sea identificado por un atributo propio.

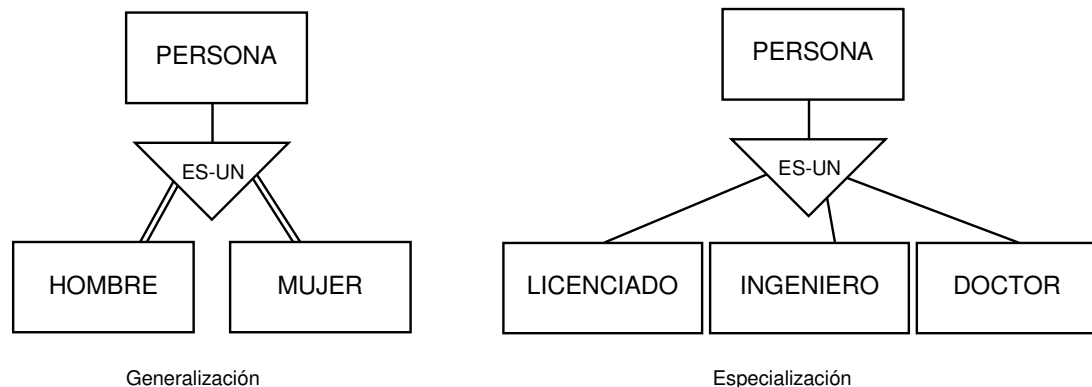
## GENERALIZACION Y ESPECIALIZACION

La generalización y la especialización son dos tipos de relacionamientos con características particulares.

**Generalización** es el resultado de unir dos o más conjuntos disjuntos de entidades de menor nivel (subtipos) para producir un conjunto de entidades de mayor nivel (supertipo). En una generalización toda entidad del supertipo debe ser una entidad de 1 y sólo 1 subtipo.

**Especialización** es el resultado de tomar uno o varios subconjuntos de un conjunto de entidades de mayor nivel (supertipo) para formar conjuntos de entidades de menor nivel (subtipos). En una especialización una entidad del supertipo puede pertenecer a 0, 1, o N subtipos.

En un diagrama E-R tanto la generalización como la especialización se representan a través de un triángulo ES-UN (ISA). La generalización se distingue de la especialización mediante arcos dobles entre el triángulo y los subtipos de entidades.

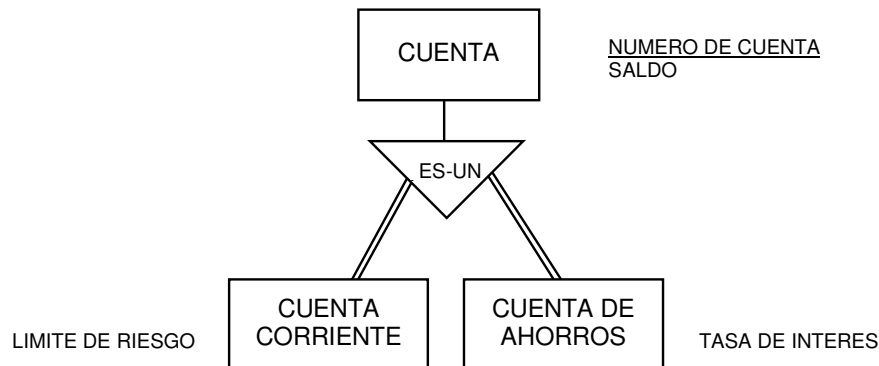


Las cardinalidades de estos tipos de relacionamientos se omiten en el diagrama porque por definición son siempre las mismas: 1 a 1 del supertipo a cada uno de los subtipos.

Cada uno de los subtipos de entidades puede tener atributos propios. Además de estos atributos propios, tanto en la generalización como en la especialización, los atributos de las entidades de mayor nivel son **heredados** por las entidades de menor nivel.

Veamos otros ejemplos:

- a) Podemos clasificar las cuentas de un banco en cuentas corrientes y cajas de ahorro. Toda cuenta tiene un número de cuenta y un saldo. Las cuentas de ahorro están adicionalmente descritas por una tasa de interés y las cuentas corrientes por un límite de riesgo.

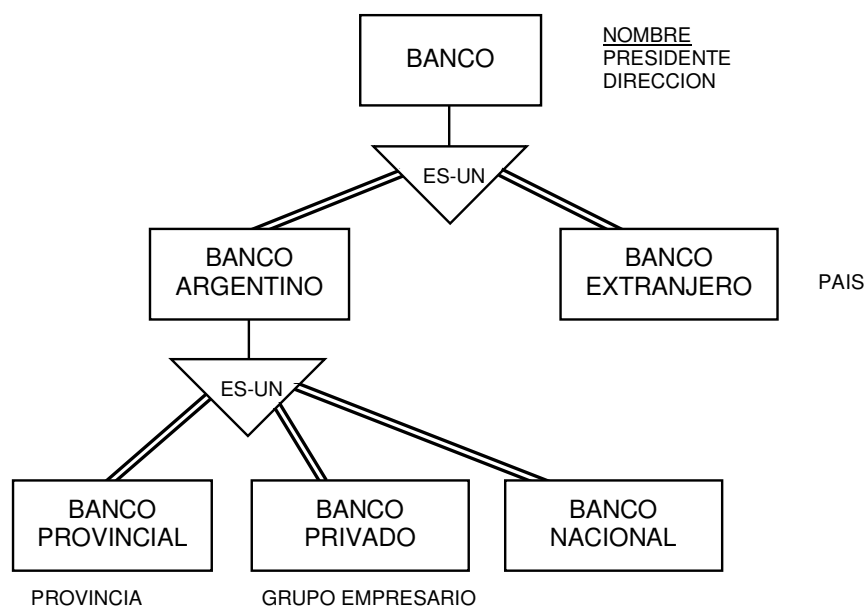


Por qué una generalización y no una especialización?

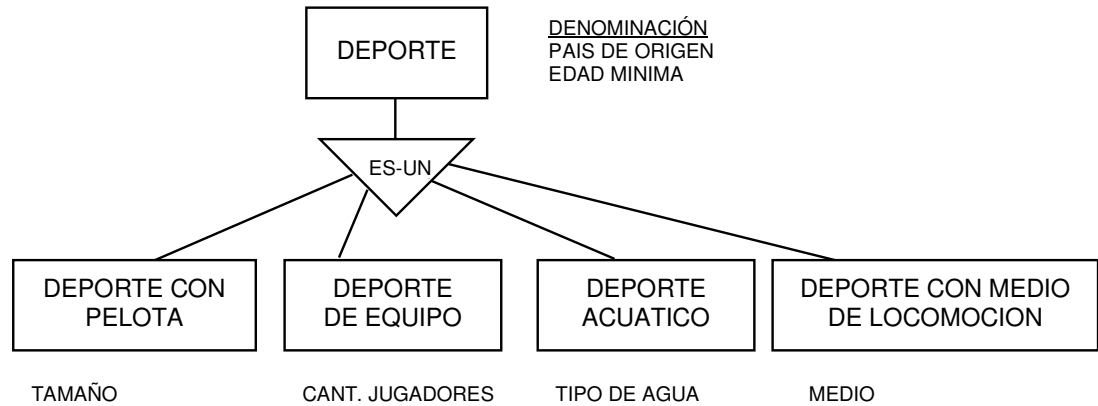
Existen similitudes entre el tipo de entidad CUENTA CORRIENTE y el tipo de entidad CUENTA DE AHORRO en el sentido que tienen varios atributos en común. Una generalización requiere que toda entidad CUENTA sea una entidad CUENTA DE AHORRO o (excluyente) una entidad CUENTA CORRIENTE.

Cada entidad del subtipo está descrita por un conjunto de atributos que incluyen todos los atributos del tipo de entidad CUENTA más los atributos propios del subtipo. Es decir, una CUENTA DE AHORRO está descrita por los atributos NUMERO DE CUENTA, SALDO y TASA DE INTERES. Una CUENTA CORRIENTE está descrita por los atributos NUMERO DE CUENTA, SALDO y LIMITE DE RIESGO.

- b) Considerando la plaza bancaria argentina, encontramos bancos nacionales, provinciales y privados, todos éstos argentinos; y bancos extranjeros. Podríamos generar el siguiente modelo:



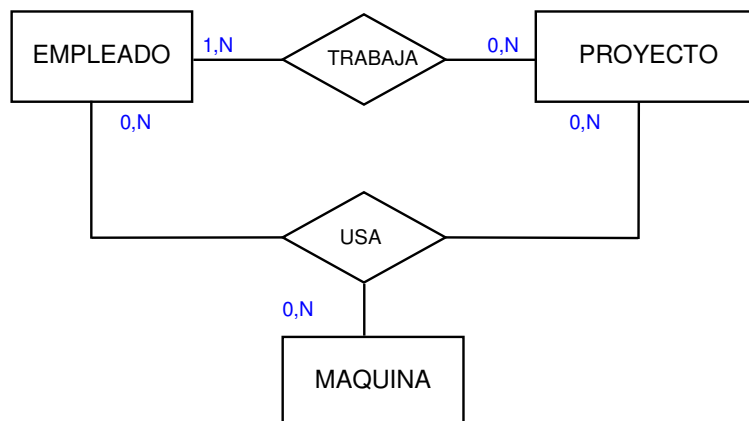
- c) Hablando de deportes podemos querer guardar distinto tipo de información de acuerdo a los distintos tipos de deportes. Cada uno de los subtipos (DEPORTE CON PELOTA, DEPORTE DE EQUIPO, etc.) es un subconjunto del conjunto de todos los deportes (DEPORTE), pero no son excluyentes el uno del otro, incluso pueden existir deportes que no pertenecen a ninguno de estos subconjuntos.



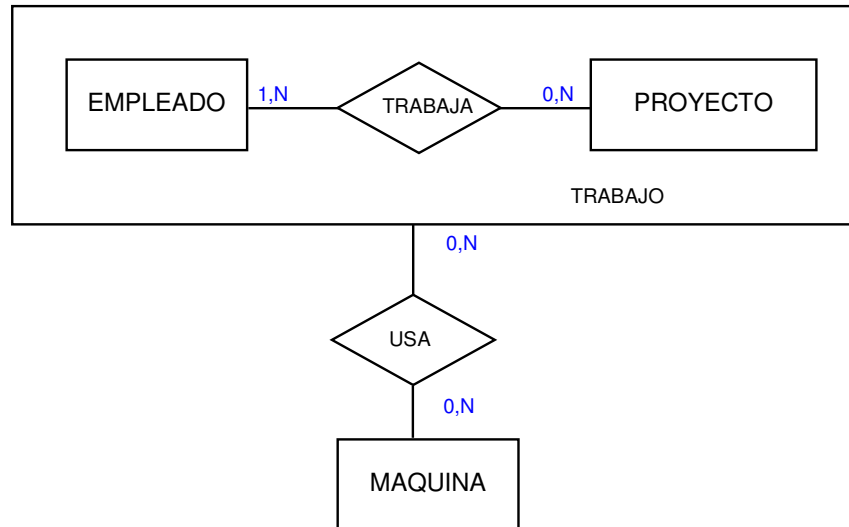
## AGREGACION

Una limitación del modelo E-R es la imposibilidad de expresar relacionamientos entre relacionamientos. La solución es usar la [agregación](#). La agregación es una abstracción que permite manejar relacionamientos como entidades de mayor nivel.

Considérese el siguiente caso: algunos empleados trabajan en proyectos y durante su trabajo usan diferentes máquinas. Usando las construcciones conocidas hasta ahora del modelo E-R, obtendríamos el siguiente diagrama:



Sin embargo, este diagrama no muestra claramente el hecho de que una máquina es usada por "un EMPLEADO asignado a un PROYECTO". Para ello, podemos realizar la agregación del tipo de relacionamiento TRABAJA y los tipos de entidades EMPLEADO y PROYECTO en un tipo de entidad de mayor nivel llamado TRABAJO. Este tipo de entidad puede ser tratado como cualquier otro tipo de entidad.



## CASO DE ESTUDIO

Sea una empresa de ferrocarriles destinada a prestar servicios de transporte ferroviario en distintos países. Los países están compuestos por provincias.

Los empleados de la empresa que no sean directivos son asignados a talleres (mecánicos, técnicos, etc. ) o trenes (conductores, inspectores, etc. ). Se asumirá que existen empleados dirigidos por otros.

Existen distintos tipos de trenes los cuales son mantenidos en talleres; éstos han de mantener al menos un tipo de tren.

A cada tren se le asigna una ruta, la cual pasa por un conjunto de estaciones, y que está compuesta por tramos para los cuales existen plazas (asientos disponibles) <sup>3</sup>.

Se admite que existan ciudades de igual nombre en distintas provincias, como así también que existan nombres de provincias iguales en países distintos.

Las estaciones están localizadas en ciudades al igual que los talleres, con la observación que estos últimos deben encontrarse en las inmediaciones de alguna estación.

Distinguiremos seis etapas en el diseño de un modelo E-R:

1. Identificar y definir los **tipos de entidades** que nos interesan
2. Identificar y definir los **tipos de relacionamientos** que nos interesan
3. Definir los **atributos** que nos interesan (y sus dominios) para los **tipos de entidades**
4. Definir los **atributos** que nos interesan (y sus dominios) para los **tipos de relacionamientos**
5. Determinar las **claves primarias** de los tipos de entidades
6. Determinar las **claves primarias** de los tipos de relacionamientos

---

<sup>3</sup> Al efecto de simplificar la solución se establecen las siguientes hipótesis:

- Se consideran los recorridos en un solo sentido (dirección)
- Se considera que un tren no puede realizar mas de un viaje por día.

**Definición de tipos de entidades**

PAIS	División política del territorio por donde transitan los TRENes
PROVINCIA	Subdivisión administrativa de los PAISes
CIUDAD	Todo lugar perteneciente a alguna PROVINCIA donde al menos exista una ESTACION ferroviaria
ESTACION	Lugar dentro de una CIUDAD donde pueden operar los TRENes
TALLER	Lugar donde los TRENes son atendidos tanto para mantenimiento preventivo como correctivo
TRAMO	Recorrido ferroviario entre ESTACIONes, sin paradas intermedias
ruta	Conjunto de TRAMOs consecutivos para ir desde la ESTACION origen hasta la ESTACION terminal de esa RUTA
TREN	Medio de transporte en el cual se ejecuta un recorrido
TIPO-TREN	Categorías en que son clasificados los TRENes
FECHA	Día, mes y año en que esta programado un viaje
EMPLEADO	Personal que trabaja en la empresa ferroviaria

**Definición de tipos de relacionamientos**

"Los países esta compuestos por provincias."



"Se admite que existan ciudades de igual nombre en distintas provincias, como así también que existan nombres de provincias iguales en países distintos."



"Existen distintos tipos de trenes...."

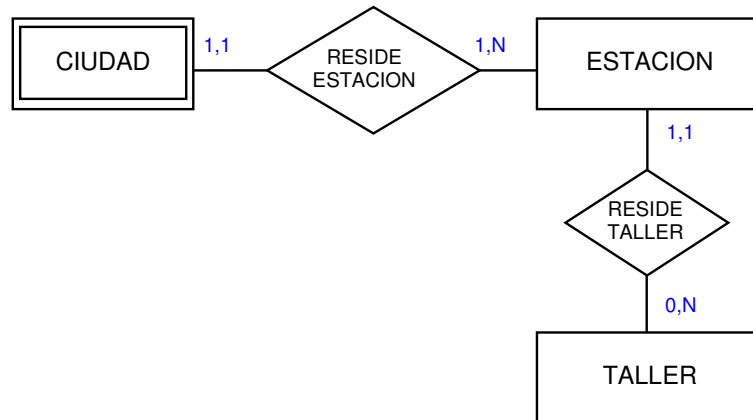




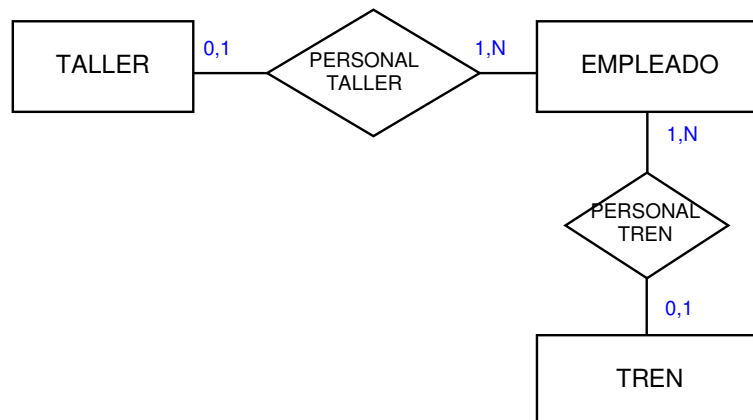
"...tipos de trenes los cuales son mantenidos en talleres; éstos han de mantener al menos a un tipo de tren."



"Las estaciones están localizadas en ciudades al igual que los talleres, con la observación que estos últimos deben encontrarse en las inmediaciones de alguna estación."

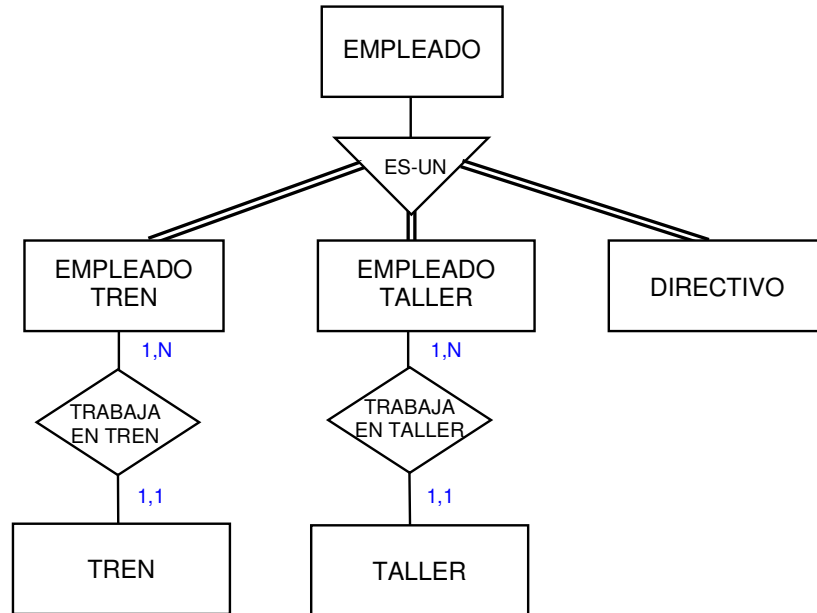


"Los empleados de la empresa que no sean directivos son asignados a talleres (mecánicos, técnicos, etc.) o trenes (conductores, inspectores, etc.)"

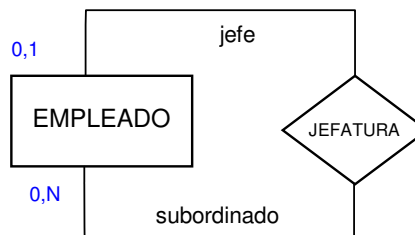


Con esta representación no estamos recogiendo la restricción de exclusión contenida en la afirmación de que todo empleado o es directivo o es de tren o es de taller. Si observamos con detenimiento el tipo de entidad EMPLEADO podemos apreciar que en su relación de trabajo se particiona en tres tipos de entidades (o subtipos de entidades): los que trabajan en talleres, los que trabajan en trenes y los directivos. Esta partición nos está señalando que son de interés los tipos de entidad DIRECTIVO, EMPLEADO TALLER y EMPLEADO TREN, que no habíamos detectado en el paso 1 (Definición de tipos de entidades), al cual deberíamos agregar las siguientes definiciones:

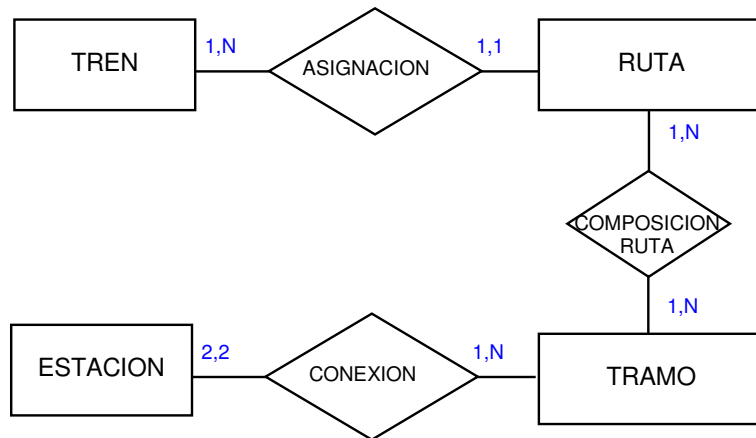
EMPLEADO TALLER      EMPLEADO asignado para trabajar en un TALLER  
 EMPLEADO TREN        EMPLEADO asignado para trabajar en un TREN  
 DIRECTIVO              EMPLEADO que cumple funciones directivas



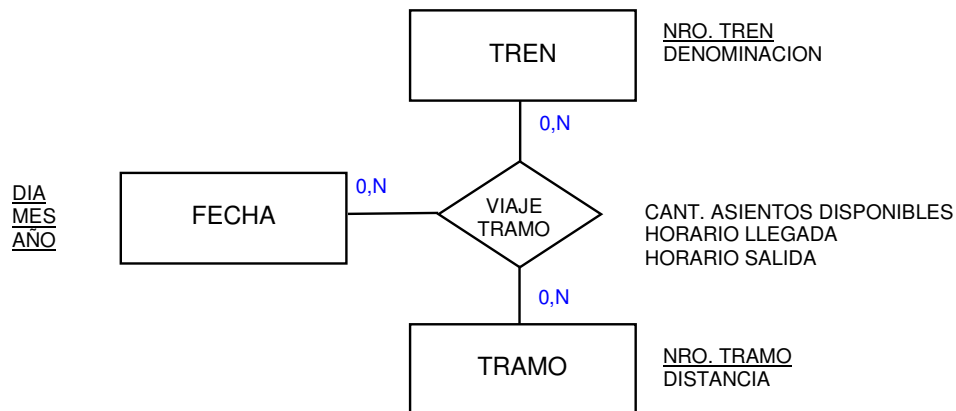
"Se asumirá que existen empleados dirigidos por otros."



"A cada tren se le asigna una ruta, la cual pasa por un conjunto de estaciones, y que está compuesta por tramos...."



"... tramos.. para los cuales existen plazas (asientos disponibles). "



Para determinar las cardinalidades de este tipo de relacionamiento ternario (intervienen tres tipos de entidades) procedemos de la siguiente forma:

- Cada par TRAMO-FECHA se asocia con 0 a N TRENes (es decir, en una determinada FECHA y por un determinado TRAMO, pasan 0 a N TRENes).
- Cada par FECHA-TREN se asocia con 0 a N TRAMOs (es decir, en una determinada FECHA un TREN particular, pasa por 0 a N TRAMOs).
- Cada par TRAMO-TREN se asocia con 0 a N FECHAs (es decir, un determinado TREN pasa por un determinado TRAMO, en 0 a N FECHAs).

En cuanto a la identificación de tipos relacionamientos realizada, la definición de cada uno de los tipos de relacionamientos del modelo propuesto es la siguiente:

COMPOSICION PAIS	Composición de países en provincias (dentro de un país los nombres de las provincias son únicos).
RESIDE CIUDAD	Composición de provincias en ciudades (dentro de una provincia los nombres de las ciudades son únicos).
CLASIFICACION	Clasificación de trenes según su tipo.
MANTENIMIENTO	Mantenimiento de los trenes por talleres, según el tipo de tren.
RESIDE ESTACION	Existencia de una estación en una ciudad.
RESIDE TALLER	Existencia de un taller en una estación.
TRABAJA EN TREN	Asignación de empleados a los trenes.
TRABAJA EN TALLER	Asignación de empleados a los talleres.
JEFATURA	Relación jerárquica ente empleados y sus jefes.
ASIGNACION	Asignación de trenes a sus rutas.
COMPOSICION RUTA	Composición de rutas por tramos.
CONEXION	Conexión de tramos con dos estaciones.
VIAJE TRAMO	Recorrido de un tren en un tramo en una fecha determinada.

#### ***Identificación de atributos para tipos de entidades***

PAIS	(nombre_país)
PROVINCIA	(nombre_provincia)
CIUDAD	(nombre_ciudad)
ESTACION	(nombre_estación, altura)
TALLER	(nro_taller, dirección)
TRAMO	(nro_tramo, distancia)
RUTA	(nro_ruta)
TREN	(nro_tren, denominación)
TIPO-TREN	(nombre_tipo, capacidad)
FECHA	(día, mes, año)
EMPLEADO	(nro_empleado, nombre_empleado, dirección_empleado)
EMPLEADO TALLER	(especialidad)
EMPLEADO TREN	(jornada, tipo_trabajo)
DIRECTIVO	(profesión, cargo)

**Identificación de atributos para tipos de relacionamientos**

VIAJE TRAMO (cant\_asientos\_disponibles, horario\_llegada, horario\_salida)

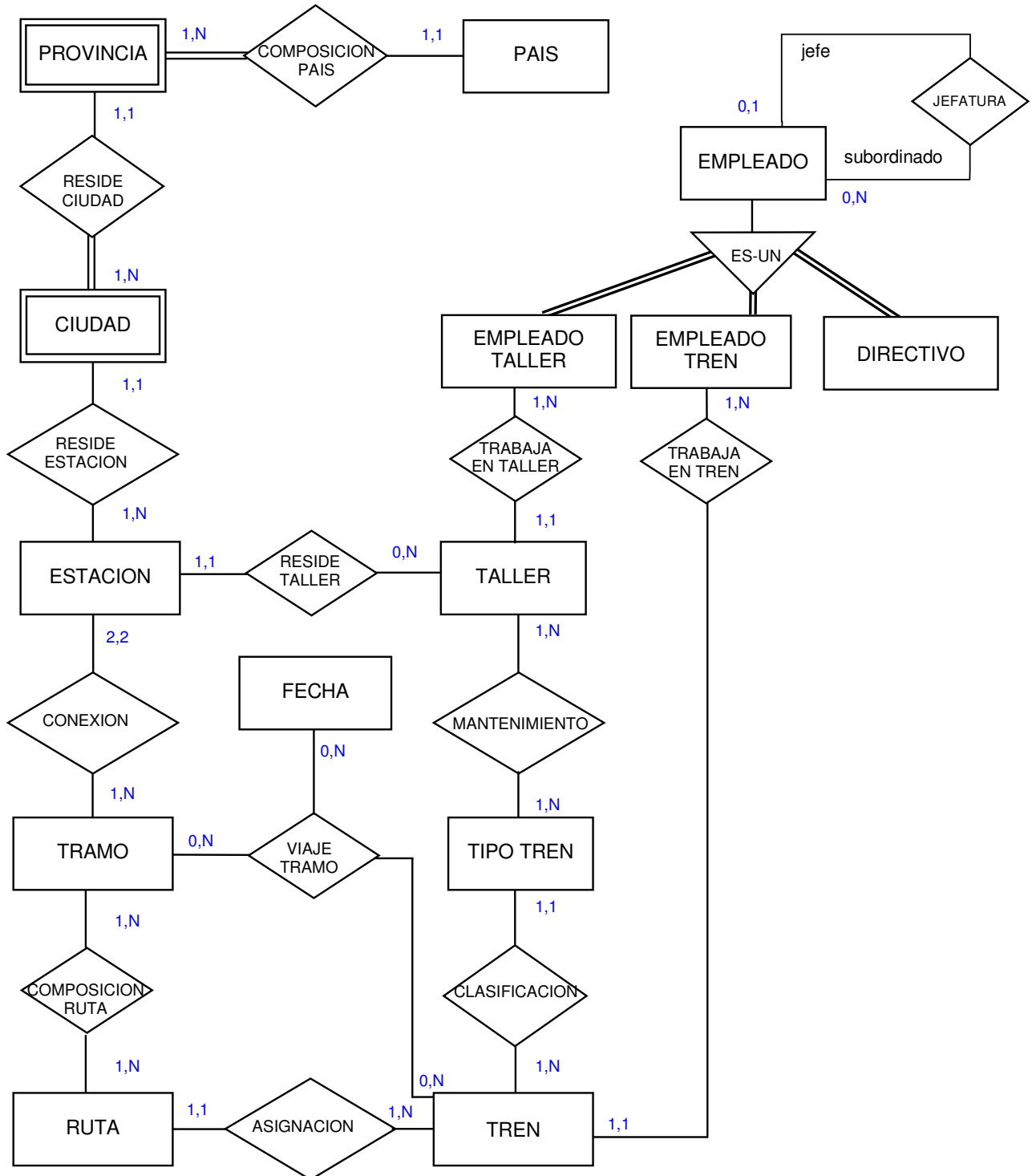
**Identificación de claves primarias para tipos de entidades**

PAIS (nombre\_país)  
 PROVINCIA (nombre\_país, nombre\_provincia)  
 CIUDAD (nombre\_país, nombre\_provincia, nombre\_ciudad)  
 ESTACION (nombre\_estación)  
 TALLER (nro\_taller)  
 TRAMO (nro\_tramo)  
 RUTA (nro\_ruta)  
 TREN (nro\_tren)  
 TIPO-TREN (nombre\_tipo)  
 FECHA (día, mes, año)  
 EMPLEADO (nro\_empleado)  
 EMPLEADO TALLER (nro\_empleado)  
 EMPLEADO TREN (nro\_empleado)  
 DIRECTIVO (nro\_empleado)

**Identificación de claves primarias para tipos de relacionamientos**

COMPOSICION PAIS (nombre\_país, nombre\_provincia)  
 RESIDE CIUDAD (nombre\_país, nombre\_provincia, nombre\_ciudad)  
 CLASIFICACION (nro\_tren)  
 MANTENIMIENTO (nombre\_tipo, nro\_taller)  
 RESIDE ESTACION (nombre\_estación)  
 RESIDE TALLER (nro\_taller)  
 TRABAJA EN TREN (nro\_empleado)  
 TRABAJA EN TALLER (nro\_empleado)  
 JEFATURA (nro\_empleado\_jefe, nro\_empleado\_subordinado)  
 ASIGNACION (nro\_tren)  
 COMPOSICION RUTA (nro\_ruta, nro\_tramo)  
 CONEXION (nro\_tramo, nombre\_estación)  
 VIAJE TRAMO (día, mes, año, nro\_tren, nro\_tramo)

**Diagrama ENTIDAD-RELACIONAMIENTO**



## **EJERCITACION PROPUESTA MODELO ENTIDAD-RELACIONAMIENTO**

Realizar modelos E-R en base a cada una de las descripciones que se enumeran a continuación.

Cada modelo debe incluir:

- **Diagrama E-R**, con indicación de **restricciones de cardinalidad de correspondencia, de existencia y de identificación**.
- Diccionario de Datos, con definición de **tipos de entidades y tipos de relacionamientos**, especificación de **atributos y claves primarias**.
- Especificación de supuestos que justifiquen el modelo.

1. Considérese una tienda comercial minorista compuesta por varias sucursales:
  - i. De cada empleado interesan sus datos personales como número de empleado, nombre, apellido, domicilio y la sucursal en la que trabaja
  - ii. De cada sucursal interesa su nombre, sus empleados, el gerente y los artículos que vende.
  - iii. Los datos de cada artículo son: descripción, fabricantes, precio, número de modelo (asignado por el fabricante) y número interno (asignado por la tienda comercial).
  - iv. Los datos de cada fabricante son: razón social, domicilio, artículos que suministra a la tienda y sus precios.
2. Una compañía aseguradora de autos mantiene información sobre las pólizas de sus clientes.
  - i. Las pólizas son unipersonales e intransferibles y corresponden a clientes que son personas físicas.
  - ii. A cada cliente se le asigna un número que lo identifica; además interesan sus datos personales tales como apellido, nombre, fecha de nacimiento, número de licencia de conductor y domicilio.
  - iii. Un cliente puede asegurar uno o más autos, a través de sendas pólizas.
  - iv. Una póliza queda caracterizada por un número que la identifica, la prima (costo del seguro), el premio (prima + gastos administrativos e impuestos), la suma asegurada, el período de vigencia, el número de matrícula del productor de la póliza y la cobertura de riesgos. Las clases de siniestros posibles están tipificadas con su correspondiente código de siniestro y descripción, de manera que la cobertura de riesgos especifica las clases de siniestros que cubre una póliza.
  - v. Se mantiene información descriptiva de los autos asegurados, como así también la fecha, el lugar y las clases de siniestro, de cada uno de los accidentes que éstos sufrieron durante la vigencia de la póliza.
  - vi. Además del dominio (número de patente), son datos relevantes de los automóviles el número de motor, el número de chasis, la marca, el modelo, el tipo y el color.
3. La AFA mantiene datos históricos acerca de sus jugadores y clubes incluyendo la siguiente información:
  - nombre del club
  - número de inscripción en AFA del jugador
  - apellido y nombre del jugador
  - club actual y clubes anteriores del jugador
  - años en que el jugador integró el equipo de un club
  - salarios promedio anuales del jugador en un equipo
  - posición en la que juega el jugador
  - cantidad de goles convertidos en cada año por el jugador
  - ciudad y provincia de origen del jugador
  - ciudad y provincia de ubicación del club
  - antigüedad del club
 Se debe tener en cuenta que:
  - \* Cada año se juega un único campeonato.
  - \* Los datos deben conservarse durante varios campeonatos.
  - \* Los jugadores sólo pueden cambiar de club una vez al año, durante el receso del campeonato (libro de pases abierto).
  - \* Un jugador juega en una única posición en cada campeonato.
  - \* Un jugador puede estar libre (sin pertenecer a club alguno) uno o más años.

## 4. Considérese una biblioteca.

- i. Una obra queda caracterizada por un título y un código de obra (ambos únicos para cada obra) y se considera escrita por uno o más autores, cada uno de los cuales se identifica por un nombre de autor, que no admite repeticiones; es de interés la nacionalidad (País de origen) de estos autores.
- ii. Por otra parte, cada obra hace referencia a una lista de temas asociados, los cuales se encuentran catalogados según el nombre del rubro que conceptualiza el tema y el número de ítem dentro de ese rubro; cada uno de estos temas tiene también una sucinta descripción aclaratoria del ítem que, sin embargo, no es exclusiva de cada tema.
- iii. Una obra puede haber sido editada en distintas oportunidades, por la misma o distintas editoriales cada vez, determinando así un número correlativo de edición; es posible que haya más de una edición de la misma obra en el mismo año.
- iv. Cada edición, a su vez, puede estar publicada en un único idioma, eventualmente distinto al de otras ediciones de la misma obra. La mayoría de los libros están en español, aunque también hay libros en inglés, francés y portugués; sin embargo, por norma de la biblioteca, no hay libros en otros idiomas que no sean los precitados.
- v. Cada ejemplar de la misma edición de una obra, mantiene un número correlativo de ejemplar, que ayuda a llevar un mejor control de las existencias; por otra parte, para lograr la rápida ubicación de los libros, se asigna a cada ejemplar un lugar específico en la biblioteca, caracterizado por un código de sector y un número de estante; es claro que algunos de los estantes pueden estar vacíos, pero de estar asignados, en cada estante no cabe más de un libro.
- vi. Cuando el único ejemplar de una obra se encuentra en mal estado, se retira a la obra de la disponibilidad para consulta; no obstante, se conservan los datos de sus autores y los de los temas que la obra abarca.

Hipótesis de trabajo: considerar que no existe codificación ISBN.

## 5. El departamento de Personal de una empresa guarda información de sus empleados para liquidación de sueldos y jornales.

- i. Tanto los empleados jornalizados como los asalariados están identificados por un número de legajo.
- ii. Cuando un asalariado asciende a Director, interesa la fecha de su designación.
- iii. Se cuenta con los datos personales de los empleados (apellido y nombre, DNI, CUIL, domicilio, fecha de nacimiento y sexo) así como los de sus familiares a cargo.
- iv. Los empleados jornalizados pueden solicitar préstamos a la empresa, pero este beneficio no está disponible para los asalariados.
- v. Tanto los asalariados como los jornalizados pueden convenir para trabajar por hora, fuera del horario de trabajo normal. Personal necesita tener información referente a los días y horas, del año en curso, en que los empleados están disponibles para trabajar extra.

## 6. La oficina de inscripción de alumnos de la facultad necesita contar con información, a saber, para verificar si un alumno está habilitado para inscribirse en una asignatura:

- nro. de padrón, apellido y nombre del alumno
- carrera y plan de estudios en que está inscripto
- régimen de correlatividades vigentes para esa carrera y plan
- código de las asignaturas que el alumno tiene aprobadas
- código de las asignaturas en las que está inscripto

Para responder a las consultas de los alumnos se precisa conocer:

- código y nombre de las asignaturas de cada carrera y plan
- régimen de correlatividades para cada carrera y plan
- profesores, horarios y vacantes de las asignaturas a dictarse en ese cuatrimestre

Considérese que:

- \* Un profesor puede estar a cargo de uno o más cursos de una o más asignaturas
- \* Una materia puede dictarse en distintos turnos, dando lugar en cada caso a un curso distinto. Cada turno, abarca dos días no consecutivos de la semana y una de las cuatro bandas horarias (7 a 11, 11 a 15, 15 a 19 , 19 a 23)



7. Analizando la organización de los trabajos prácticos de un cuatrimestre en la materia Base de Datos, se observaron las siguientes pautas:
- los auxiliares docentes constituyen grupos de trabajo integrados, cada uno de ellos, por un docente rentado y uno, dos o tres docentes ad honorem.
  - independientemente de su condición, los auxiliares docentes pueden, eventualmente, especializarse en uno o más temas de la materia; estos temas se encuentran catalogados como ítems de programa. Cada ítem de programa está caracterizado por una denominación que lo identifica y una descripción muy genérica del contenido que, por su forma abreviada, puede repetirse para distintos ítems de programa; para facilitar la identificación, existe también un número de ítem respectivo. Puede haber ítems de programa en los cuales no se especialice docente alguno.
  - en lo referente a los alumnos, caracterizados por su número de padrón, apellido, nombre, número de carrera y teléfono, se constituyen grupos de hasta cuatro integrantes. Cada grupo es coordinado por un docente rentado asignado desde el comienzo del cuatrimestre.
  - cada auxiliar docente queda descripto por su DNI, su apellido y nombre y su número de teléfono. En el caso de los docentes rentados, además se conoce su número de legajo, fecha de nombramiento y CUIL. También es un dato de interés, la fecha de incorporación de cada docente ad honorem.
  - cada alumno debe rendir tres evaluaciones parciales, numeradas correlativamente, en fechas preestablecidas en el cronograma del cuatrimestre. En caso que la calificación obtenida en cualesquiera de ellas sea insuficiente o, simplemente que el alumno no se haya presentado a rendir, le cabe la posibilidad de rendir hasta en dos oportunidades más el mismo número de parcial. La segunda oportunidad, como así también la tercera, las puede utilizar en cualesquiera de las fechas de recuperación que se establecen para el período de receso entre cuatrimestres.
  - se considera que la calificación obtenida por un alumno en una evaluación parcial, es asignada por un único docente rentado, independientemente de las consultas que puedan realizarse entre colegas antes de definir las calificaciones. Es claro que en los casos de alumnos ausentes, no existe la correspondiente evaluación parcial.
8. Aerolíneas Sudacas está desarrollando un sistema de mantenimiento de aviones, para ayudar a cubrir las normas de seguridad vigentes. Además de producir información acerca de los aviones individuales, tal como:
- número de serie del avión
  - tipo de avión y cantidad de asientos para pasajeros (capacidad)
  - procedimientos de mantenimiento a efectuar a cada avión
  - duración y frecuencia de cada procedimiento de mantenimiento
  - fecha de los procedimientos de mantenimiento efectuados sobre un avión
  - siguiente fecha en la que un avión tiene programado un procedimiento de mantenimiento
- también se espera que el sistema proporcione información sobre los mecánicos de cada taller y su nivel de capacitación.
- La empresa tiene un programa de capacitación para los mecánicos. Para cada procedimiento se realizan cursos idénticos en forma periódica. La duración de los cursos depende del procedimiento. La especialización de cada mecánico está dada por los correspondientes cursos de capacitación que ha tomado y aprobado. Además de los datos personales de cada mecánico, interesa conocer la fecha en que tomó sus cursos de capacitación con la respectiva calificación obtenida.
- Observaciones:
- \* Cada taller tiene un gerente y varios mecánicos
  - \* En un taller determinado pueden efectuarse varios procedimientos
  - \* Un determinado procedimiento sólo se puede efectuar en un taller (para todos los aviones)
  - \* Se desea mantener registros históricos de todos los procedimientos de mantenimiento realizados durante la vida útil de cada avión
  - \* Un mecánico puede estar calificado para efectuar varios procedimientos, pero sólo puede efectuar aquellos que sean factibles en el taller en el que se encuentra trabajando.

## MODELO DE DATOS RELACIONAL

### INTRODUCCION

El Modelo de Datos Relacional fue propuesto por E. F. Codd en 1970. La mayoría de los sistemas de bases de datos comercializados hasta 1980 utilizaban el Modelo Red o el Jerárquico, pero estos Modelos de Datos fueron siendo desplazados gradualmente por el Relacional debido a su sencillez y flexibilidad.

El Modelo Relacional, basado en conceptos muy simples, tiene asociada una teoría que no puede ser separada del mismo: la teoría de la Normalización de las Relaciones.

Esta teoría tiene por objeto suprimir los datos redundantes, eliminar los comportamientos anormales durante las actualizaciones y facilitar la comprensión de las relaciones semánticas entre los datos.

Todo Modelo de Datos se caracteriza por tres conjuntos:

- 1) ESTRUCTURAS
- 2) RESTRICCIONES (implícitas y explícitas)
- 3) OPERACIONES

En el Modelo RELACIONAL esta caracterización se manifiesta de la siguiente forma:

- |                             |                                    |
|-----------------------------|------------------------------------|
| 1) ESTRUCTURAS              | ⇒ dominios atómicos                |
|                             | ⇒ relaciones                       |
| 2) RESTRICCIONES IMPLICITAS | ⇒ regla de integridad de entidad   |
|                             | ⇒ regla de integridad referencial  |
| 3) OPERACIONES              | ⇒ álgebra relacional o equivalente |

### CONCEPTOS BASICOS DEL MODELO

#### DOMINIO

Un dominio es un conjunto homogéneo de valores.

##### Ejemplos:

El dominio de los números enteros, el dominio de los números reales, el dominio de las vocales, etc.

$$V = \{a, e, i, o, u\}$$

#### RELACION

El concepto “relación” deriva directamente de la teoría de conjuntos. Es un subconjunto del producto cartesiano de una lista de dominios. Está caracterizada generalmente por un nombre.

Recordemos que el producto cartesiano de un conjunto de dominios  $D_1, D_2, D_3, \dots, D_n$  es el conjunto de las tuplas  $(v_1, v_2, \dots, v_n)$  tales que  $v_i$  pertenece a  $D_i$ .

##### Ejemplo:

El producto cartesiano de los dominios:

$$D1 = \{\text{azul, blanco, rojo}\}$$

$$D2 = \{0, 1\}$$

está formado por las seis tuplas

$$D1 \times D2 = \{(\text{azul}, 0), (\text{azul}, 1), (\text{blanco}, 0), (\text{blanco}, 1), (\text{rojo}, 0), (\text{rojo}, 1)\}$$

A partir de  $D1 \times D2$  podemos componer una relación  $r$

$$r = \{(azul,0),(azul,1),(blanco,0)\}$$

De una forma más sencilla, podemos considerar una relación como una tabla de dos dimensiones en que las columnas corresponden a los dominios y las filas contienen las tuplas:

D1	D2
azul	0
azul	1
blanco	0

**Dominio D1**

**Tupla** (blanco,0)

### ATRIBUTO

Los matemáticos definen una relación como un subconjunto del producto cartesiano de una lista de dominios, usando el entero **1** para denotar el atributo cuyo dominio aparece en el primer lugar en la lista de dominios: el entero **2** para el atributo cuyo dominio aparece en segundo lugar, etc.

El Modelo Relacional, en cambio, otorga un nombre a cada columna sin conferir importancia alguna al orden en el que se encuentran las mismas.

Sea una relación  $r$  conteniendo  $n$  tuplas. Sea  $t$  una tupla de la relación. Se utilizará la notación  $t[A]$  para denotar el valor en  $t$  del atributo (o conjunto de atributos)  $A$ .

$t[A] = \text{valor en la tupla } t \text{ del atributo } A$
--

Adicionalmente se utilizará la notación  $r.A$  para referirse al atributo  $A$  perteneciente a  $r$ .

$r.A = \text{atributo } A/A \in r$
------------------------------------

### ESQUEMA DE RELACION

En un primer enfoque, un esquema de relación es el nombre de la relación seguido de la lista de los atributos con sus dominios.

Debe diferenciarse entre el *esquema de relación* y las *instancias* de la relación: el primero describe la intención, mientras que una instancia de la relación describe su extensión en un momento dado.

Se adoptará la convención de usar minúsculas para las relaciones y mayúsculas para los esquemas de relación.

#### Ejemplo:

Dada una relación *automóvil* que incluye los atributos *número de matrícula*, *marca*, *modelo*, y *color*, su esquema es:

AUTOMOVIL = (NM: entero,  
MARCA: char(10),  
MODELO: char(4),  
COLOR: color)

El hecho de que *automóvil* es una relación con el esquema AUTOMOVIL se expresa al escribir:

automóvil (AUTOMOVIL)

**SUPERCLAVE, CLAVE CANDIDATA, CLAVE PRIMARIA, CLAVE FORANEA**

Los conceptos *superclave*, *clave candidata* y *clave primaria* que se explicaron para el Modelo Entidad-Relacionamiento son aplicables al Modelo Relacional.

**SUPERCLAVE**

Es todo conjunto de uno o más atributos pertenecientes al esquema de relación **R** que permita identificar en forma unívoca a cada tupla **t** de la relación **r(R)**.

**CLAVE CANDIDATA: UNICIDAD - MINIMALIDAD**

Superclave para la cual ningún subconjunto propio es una superclave. En términos formales:

Sea **R** un *esquema de relación* con *atributos*  $A_1, A_2, \dots, A_n$ . El conjunto de atributos  $K = (A_1, A_2, \dots, A_k)$  es una *clave candidata* de **R** si, y sólo si, satisface estas dos propiedades independientes del tiempo:

**UNICIDAD**

En todo momento, no existen dos tuplas distintas de **r(R)** que tengan el mismo valor para  $A_1$ , el mismo valor para  $A_j$ , ..., el mismo valor para  $A_k$ .

**MINIMALIDAD**

Ningún atributo  $A_1, A_2, \dots, A_r$  puede ser eliminado del conjunto **K** sin que pierda la propiedad de unicidad.

**CLAVE PRIMARIA**

Se utiliza el término *clave primaria* para referir a la *clave candidata* que elija el diseñador de la base de datos como la forma principal de identificar a las tuplas dentro de una relación

**CLAVE FORANEA**

En una relación **r(R)** una *clave foránea* es un atributo, o combinación de atributos, que es *clave primaria* en otra relación **s(S)**.

Nótese que la *clave primaria* y la *clave foránea* deben estar definidas sobre el mismo dominio.

Los esquemas de relación **R** y **S** no son necesariamente distintos.

La relación **r(R)** se denomina relación referenciante y la relación **s(S)** relación referenciada.

Las *claves foráneas* representan las referencias de una relación a otra (es por esta razón que en algunas propuestas de Diagrama de Estructuras de Información basadas en el Diagrama Entidad-Relacionamiento surge lo que se denomina *atributo referencial*, para denotar, precisamente, una clave foránea del Modelo Relacional).

**MODELO RELACIONAL: RESTRICCIONES**

El Modelo Relacional tiene dos restricciones inherentes o implícitas, llamadas *reglas de integridad*:

- 1) REGLA DE INTEGRIDAD DE ENTIDAD
- 2) REGLA DE INTEGRIDAD REFERENCIAL

que refieren a las claves primarias y a las claves foráneas, respectivamente.

**REGLA DE INTEGRIDAD DE ENTIDAD**

Ningún atributo de la clave primaria de un esquema de relación puede tomar valor nulo.

Valor nulo significa que la información no se conoce o que la propiedad que describe el atributo no es aplicable.

**REGLA DE INTEGRIDAD REFERENCIAL**

Si la relación  $r(R)$  incluye una clave foránea  $F$  que es la clave primaria  $P$  de una relación  $s(S)$ , entonces todo valor de  $F$  en  $r(R)$  debe ser totalmente nulo o ser igual al valor de  $P$  en alguna tupla de  $s(S)$ .

$R$  y  $S$  no son necesariamente distintos.

**ALGEBRA RELACIONAL**

El álgebra relacional es un lenguaje de consulta procedimental. Existen cinco operaciones básicas que son:

Operaciones unarias:

1. Selección (o Restricción)
2. Proyección

Operaciones binarias:

3. Unión
4. Diferencia
5. Producto cartesiano

Todas ellas producen como resultado una nueva relación.

Además de las cinco operaciones fundamentales, introduciremos otras operaciones, a saber

Operaciones binarias:

6. Intersección
7. Join
8. División

Estas se definirán en términos de las operaciones fundamentales.

**SELECCIÓN o RESTRICCIÓN**

La selección de la relación  $r(R)$  por un predicado (o criterio de selección)  $P$ , es una relación  $r'(R)$ , con el mismo esquema  $R$ , en la que las tuplas son aquellas pertenecientes a  $r$  que satisfacen el predicado  $P$ .

Se utiliza la siguiente notación para la operación Selección aplicada sobre una relación  $r$  con un predicado  $P$ :

$$\sigma_P(r)$$

Para clarificar esta definición, introduzcamos el concepto de *predicado*.

Un *predicado atómico* es una expresión lógica simple formada por un nombre de atributo ligado a una constante, o a otro nombre de atributo, mediante un operador de comparación  $\theta$ , tal que  $\theta$  pertenece a  $\{ < , > , <= , >= , = , <> \}$ . Un predicado atómico es, pues, del tipo  $A_i \theta C$ , donde  $A_i$  es un nombre de atributo y  $C$  es una constante u otro nombre de atributo.

Los *predicados atómicos* se combinan mediante *conectores lógicos* conjuntivos y/o disyuntivos para formar *predicados compuestos*.

Conectores:

- conjunción (operador AND):  $\wedge$
- disyunción inclusiva (operador OR):  $\vee$

Operadores:

- Negación:  $\neg$

En general, un predicado contendrá expresiones del tipo:

$$P = (A_1 \theta C_1) \wedge ((A_2 \theta C_2) \vee \neg(A_3 \theta C_3))$$

Ejemplo de Selección:

Dada la relación **automóvil(AUTOMOVIL)** :

AUTOMOVIL	NM	MARCA	MODELO	COLOR
	SQY687	Chevrolet	Corsa 4p.	Verde
	AQA121	Audi	A3	Gris perla
	BEX369	Chrysler	Grand Cherokee	Verde
	RCP333	Renault	Megane	Azul océano
	UTU025	Renault	Clío	Azul océano

fig.2

La operación  $\sigma_{\text{COLOR} = \text{"Verde"}}(\text{automóvil})$  tiene como relación resultante a:

AUTOMOVIL	NM	MARCA	MODELO	COLOR
	SQY687	Chevrolet	Corsa 4p.	Verde
	BEX369	Chrysler	Grand Cherokee	Verde

fig.3

### PROYECCION

La proyección de una relación  $r$ , de esquema  $R=(A_1, A_2, \dots, A_n)$ , sobre los atributos  $A_{i1}, A_{i2}, \dots, A_{ip}$  (donde se cumple  $I_j < I_k$  y  $p < n$ ) es una relación  $r'$ , de esquema  $R'=(A_{i1}, A_{i2}, \dots, A_{ip})$ , cuyas tuplas se obtienen por eliminación en  $r$  de los atributos de  $R$  que no pertenecen a  $R'$  y por supresión de las tuplas duplicadas.

Se utiliza la siguiente notación para la operación Proyección de una relación  $r$  sobre los atributos:

$$\pi_{A_{i1}, A_{i2}, \dots, A_{ip}}(r)$$

Ejemplo de Proyección:

Dada la relación **automóvil(AUTOMOVIL)** de la fig.2, la operación  $\pi_{MARCA, COLOR}(r)$  tiene como relación resultante a:

MARCA	COLOR
Chevrolet	Verde
Audi	Gris perla
Chrysler	Verde
Renault	Azul océano

fig 4.

Nótese que se han eliminado las columnas correspondientes a los atributos NM y MODELO. Al eliminar estas dos columnas quedaría repetida la tupla (Renault, Azul océano), por lo tanto, para cumplir con la definición de relación, se elimina la tupla repetida.

**UNION**

La unión de dos relaciones **r1** y **r2** que tienen el mismo esquema, es una relación **r** con igual esquema, conteniendo el conjunto de tuplas que pertenecen a **r1** o a **r2** o a ambas relaciones.

El requisito que deben cumplir los esquemas de relación de ambos operandos, se sintetiza diciendo que **r1** y **r2** son *UNIONCOMPATIBLES*.

Se ha adoptado la siguiente notación para la operación unión entre dos relaciones **r1** y **r2**:

$r1 \cup r2$
--------------

Ejemplo de Unión:

Dadas las relaciones **auto1(AUTOMOVIL)** y **auto2(AUTOMOVIL)**:

**auto1**

AUTOMOVIL	NM	MARCA	MODELO	COLOR
	CSW698	Chevrolet	Corsa 4p.	Verde
	BMW666	Audi	A3	Negro
	AXA123	Renault	Megane	Blanco

**auto2**

AUTOMOVIL	NM	MARCA	MODELO	COLOR
	PEZ888	Honda	Accord	Gris Perla
	REG444	Mitsubishi	Colt	Verde
	BMW666	Audi	A3	Negro

fig.5

La relación resultante de la operación **auto1**  $\cup$  **auto2** es:

AUTOMOVIL	NM	MARCA	MODELO	COLOR
	CSW698	Chevrolet	Corsa 4p.	Verde
	AXA123	Renault	Megane	Blanco
	PEZ888	Honda	Accord	Gris Perla
	REG444	Mitsubishi	Colt	Verde
	BMW666	Audi	A3	Negro

fig.6

**DIFERENCIA**

La diferencia de dos relaciones **r1** y **r2**, ambas con el mismo esquema y en el orden **r1 - r2**, es una relación **r** con igual esquema, conteniendo las tuplas que pertenecen a **r1** y no pertenecen a **r2**.

Nótese que **r1** y **r2** deben ser *UNIONCOMPATIBLES*.

Se utiliza la siguiente notación:

$$r1 - r2$$

Ejemplo de Diferencia:

Dadas las relaciones **auto1(AUTOMOVIL)** y **auto2(AUTOMOVIL)** presentes en la figura 5, la relación resultante de la operación **auto1 - auto2** es:

AUTOMOVIL	NM	MARCA	MODELO	COLOR
	CSW698	Chevrolet	Corsa 4p.	Verde
	AXA123	Renault	Megane	Blanco

fig.7

**PRODUCTO CARTESIANO**

El producto cartesiano de dos relaciones **r1(R1)** y **r2(R2)** es una relación **r** que tiene por esquema la concatenación de **R1** con **R2** y cuyas tuplas son todas las concatenaciones posibles de una tupla de **r1** con una tupla de **r2**.

**Producto Cartesiano**

- **Esquema R** = concatenación de **R1** y **R2**
- **Tuplas** = todas las tuplas **t** para las cuales existe **t1** en **r1**, y **t2** en **r2** tales que **t[R1] = t1[R1]** y **t[R2] = t2[R2]**

Se utiliza la siguiente notación para representar el producto cartesiano de dos relaciones **r1** y **r2**:

$$r1 \times r2$$

Ejemplo de Producto cartesiano:

Dadas las relaciones **empleado(EMP)** y **departamento(DEPTO)**:

EMP	NROEMP	APELLIDO	NRODPTO	SUELDO
	4528	Gimenez	20	17500
	4356	Lopez	10	35890
	5678	García	10	8900

DEPTO	NRODPTO	NOMBRE
	10	Sistemas
	20	Finanzas
	30	Producción

fig.8



La relación resultante de la operación producto cartesiano **empleado X departamento** es:

NROEMP	APELLIDO	NRODTO	SUELDO	NRODTO	NOMBRE
4528	Gimenez	20	17500	10	Sistemas
4356	Lopez	10	35890	10	Sistemas
5678	García	10	8900	10	Sistemas
4528	Gimenez	20	17500	20	Finanzas
4356	Lopez	10	35890	20	Finanzas
5678	García	10	8900	20	Finanzas
4528	Gimenez	20	17500	30	Producción
4356	Lopez	10	35890	30	Producción
5678	García	10	8900	30	Producción

fig.9

Esquema **empleado X departameto** = (empleado.NROEMP,  
empleado.APELLIDO,  
empleado.NRODTO,  
empleado.SUELDO,  
departamento.NRODTO,  
departamento.NOMBRE)

### OPERACIONES ADICIONALES

Estas operaciones se deducen de las anteriores y son, por lo tanto, redundantes. No obstante, son esenciales para la creación práctica de los sistemas relacionales.

Por ejemplo, el producto cartesiano es una operación tan costosa en tiempo de ejecución y espacio físico requerido que es preferible realizar la operación de join.

Las operaciones adicionales que consideramos son las siguientes:

- \* JOIN y sus casos particulares
  - a) equijoin
  - b) autojoin
  - c) natural join
- \* INTERSECCION
- \* DIVISION

### JOIN

El join de dos relaciones **r1** y **r2** según un predicado **P** es una relación **r(R)** de esquema igual al del producto cartesiano **r1 X r2**, y cuyas tuplas son el conjunto de las tuplas de **r1 X r2** que satisfacen el predicado **P**.

En otros términos, podemos considerar al join de **r1** y **r2** según **P**, como la selección por un predicado **P** sobre **r1 X r2**.

Se empleará la siguiente notación para representar el join de dos relaciones **r1** y **r2** según **P**:

$$\boxed{r1 \bowtie_P r2}$$

Según lo expuesto:

$$r1 \bowtie_P r2 = \sigma_P(r1 \times r2)$$

Distinguiremos varios casos particulares del join:

- a) El equi-join de **r1** y **r2** sobre los atributos **r1.Ai** y **r2.Aj** es el join según el predicado **r1.Ai = r2.Aj**
- b) El auto-join de **r** sobre el atributo **Ai** es el join de **r** consigo misma según el predicado **1.Ai = 2.Ai**
- c) El join natural de **r1(R1)** y **r2(R2)** es una relación de esquema **R1 ∪ R2** con las tuplas del equi-join de **r1** y **r2** sobre todos los atributos que tienen el mismo nombre en **r1** y en **r2**.

Para indicar el join natural de **r1** y **r2** se utilizará la siguiente notación:

$$\boxed{r1 * r2}$$

Formalmente:

$$r1 * r2 = \pi_{R1 \cup R2}(r1 \bowtie_P r2)$$

Siendo **P** = (r1.A<sub>1</sub> = r2.A<sub>1</sub> ∧ r1.A<sub>2</sub> = r2.A<sub>2</sub> ∧ ... ∧ r1.A<sub>n</sub> = r2.A<sub>n</sub>)

y **R1 ∩ R2** = (A<sub>1</sub>, A<sub>2</sub>, ..., A<sub>n</sub>)

Ejemplo de Join natural:

Dadas las relaciones **empleado(EMP)** y **departamento(DEPTO)** presentes en la figura 8, el join natural **empleado \* departamento** tiene como relación resultante a:

NROEMP	APELLIDO	SUELDO	NRODPTO	NOMBRE
4356	Lopez	35890	10	Sistemas
5678	García	8900	10	Sistemas
4528	Gimenez	17500	20	Finanzas

fig.10

### INTERSECCION

La intersección de dos relaciones **r1** y **r2** con el mismo esquema, es una relación **r** con igual esquema, que contiene las tuplas que pertenecen a ambas relaciones, es decir, a **r1** y **r2** a la vez.

Nótese que **r1** y **r2** deben ser *UNIONCOMPATIBLES*.

Para representar la intersección de dos relaciones **r1** y **r2** se utiliza la siguiente notación:

$$\boxed{r1 \cap r2}$$

La intersección entre dos relaciones se puede obtener a partir de la operación diferencia como:

$$r1 \cap r2 = r2 - (r2 - r1)$$

### DIVISION

El cociente resultante de dividir la relación **r1**, de esquema **R1=(A<sub>1</sub>,A<sub>2</sub>,...,A<sub>n</sub>)**, por la relación **r2**, de esquema **R2=(A<sub>p+1</sub>,...,A<sub>n</sub>)**, es la relación **r**, de esquema **R=(A<sub>1</sub>,A<sub>2</sub>,A<sub>3</sub>,...,A<sub>p</sub>)**, formada por tuplas **t** tales que concatenada **t** con cada una de las tuplas de **r2** constituye una tupla existente en **r1**.

Nótese que, en términos de conjuntos, el esquema **R2** es un subconjunto de **R1**, y el esquema resultante **R** es la diferencia de conjuntos **R1 - R2**.

Si  $a_i$  es un valor cualquiera del atributo  $A_i$  podemos definir el cociente  $r$ , resultante de dividir  $r1$  por  $r2$  como:

$$r = \{ (a_1, a_2, \dots, a_p) / \forall (a_{p+1}, \dots, a_n) \in r2 : \exists (a_1, a_2, \dots, a_p, a_{p+1}, \dots, a_n) \in r1 \}$$

Para representar la división de dos relaciones  $r1$  y  $r2$  se utiliza la siguiente notación:

$$r1 \% r2$$

La división se puede obtener a partir de la diferencia, el producto cartesiano y la proyección de la forma siguiente:

$$r1 \% r2 = u1 - u2$$

siendo

$$u1 = \pi_{A1, A2, \dots, Ap}(r1)$$

$$u2 = \pi_{A1, A2, \dots, Ap}((u1 \times r2) - r1)$$

Ejemplo de División de relaciones:

Sean las relaciones  $r1(R1)$ ,  $r2(R2)$  y  $r3(R3)$ :

R1	S#	P#	Q#
	S1	P1	Q1
	S1	P2	Q2
	S1	P3	Q3
	S2	P1	Q1
	S2	P2	Q1
	S2	P3	Q3
	S3	P3	Q2
	S3	P4	Q1
	S4	P1	Q2
	S4	P2	Q2

R2	P#	Q#
	P1	Q1
	P2	Q2
	P3	Q3

R3	P#
	P1
	P2

fig.11

La figura 12 muestra los respectivos resultados de las operaciones  $r1 \% r2$  y  $r1 \% r3$ :

$r1 \% r2$	
S#	
S1	

$r1 \% r3$	
S#	Q#
S2	Q1
S4	Q2

fig.12

**COMPOSICION DE RELACIONES**

Las operaciones presentadas, encadenadas sucesivamente y aplicadas a las relaciones, permiten elaborar consultas complejas sobre una base de datos relacional.

A continuación presentaremos un conjunto de ejemplos que permiten visualizar la construcción de expresiones necesarias para consultar una base de datos relacional.

Sean las relaciones:

**pedido (PEDIDO)**  
**producto (PRODUCTO)**  
**proveedor (PROVEEDOR)**

PROVEEDOR	PROVE#	APELLIDO
	128	Gomez
	547	García
	328	Valenti

PRODUCTO	P#	NOMBRE	COLOR
	12	Rueda	Rojo
	33	Tuerca	Azul
	22	Perno	Rojo
	34	Tuerca	Verde

PEDIDO	P#	PROVE#	CANTIDAD
	34	547	100
	33	547	150
	12	128	360
	34	328	230
	12	547	70
	22	328	200
	22	547	240

fig.13

- a) Obtener los nombres de aquellos proveedores que suministran el producto número 34.  
 Planteamos la solución paso a paso:

$$\begin{aligned} r1 &= \pi_{P\#, APELLIDO} (\text{proveedor} * \text{pedido}) \\ r2 &= \sigma_{P\# = 34}(r1) \\ r &= \pi_{APELLIDO} (r2) \end{aligned}$$

Usando una expresión compuesta:

$$r = \pi_{APELLIDO} (\sigma_{P\# = 34} (\pi_{P\#, APELLIDO} (\text{proveedor} * \text{pedido})))$$

- b) Obtener los números de proveedor de aquellos proveedores que suministran al menos un producto de color rojo.  
 Solución:

$$r = \pi_{PROVE\#} (\sigma_{COLOR = "Rojo"} (\text{producto} * \text{pedido}))$$

- c) Obtener los nombres de los proveedores que suministran todos los productos.  
 Solución:

$$\begin{aligned} r1 &= \pi_{PROVE\#, P\#} (\text{pedido}) \% \pi_{P\#} (\text{producto}) \\ r2 &= r1 * \text{proveedor} \\ r &= \pi_{APELLIDO} (r2) \end{aligned}$$

**EJERCITACION PROPUESTA ALGEBRA RELACIONAL**

Considerando las relaciones:

**pedido (P#, PROVE#, CANTIDAD)**  
**producto (P#, NOMBRE, COLOR)**  
**proveedor (PROVE#, APELLIDO)**

escribir expresiones en Algebra Relacional que satisfagan los siguientes requerimientos de información:

1. Obtener los números de aquellos proveedores que suministran el producto 34.
2. Obtener el número y el apellido de aquellos proveedores que suministran el producto 34 o el producto 12.
3. Obtener los números de los productos que son tuercas o que son suministrados por el proveedor número 328.
4. Obtener los números de los productos que son tuercas y que son suministrados por el proveedor número 328.
5. Obtener los apellidos de aquellos proveedores que suministran al menos un producto que no sea el producto 33.
6. Obtener los apellidos de aquellos proveedores que no suministran el producto 33.
7. Obtener los apellidos de aquellos proveedores que suministran al menos un producto y que no suministran el producto 33.
8. Obtener el número y el apellido de aquellos proveedores que suministran más de 100 tuercas en algún pedido.
9. Obtener los números de aquellos proveedores que suministran todos los productos.
10. Obtener los apellidos de aquellos proveedores que suministran todos los productos menos el 12.
11. Obtener los apellidos de aquellos proveedores que suministran, al menos, todos los productos suministrados por el proveedor número 128.
12. Obtener los números de aquellos proveedores que suministran todos los productos en igual cantidad.

**PRACTICA ADICIONAL DE ALGEBRA RELACIONAL:**

**CASO 1:** Considerando el siguiente Esquema Relacional que describe las operaciones de préstamos y depósitos de una entidad bancaria:

sucursal ( SUC#, ACTIVO, CIUDAD )  
 cliente ( CLI#, NOMBRE, DIRECCION, CIUDAD )  
 cuenta ( SUC#, CTA#, CLI#, SALDO )  
 préstamo ( SUC#, PTMO#, CLI#, IMPORTE )

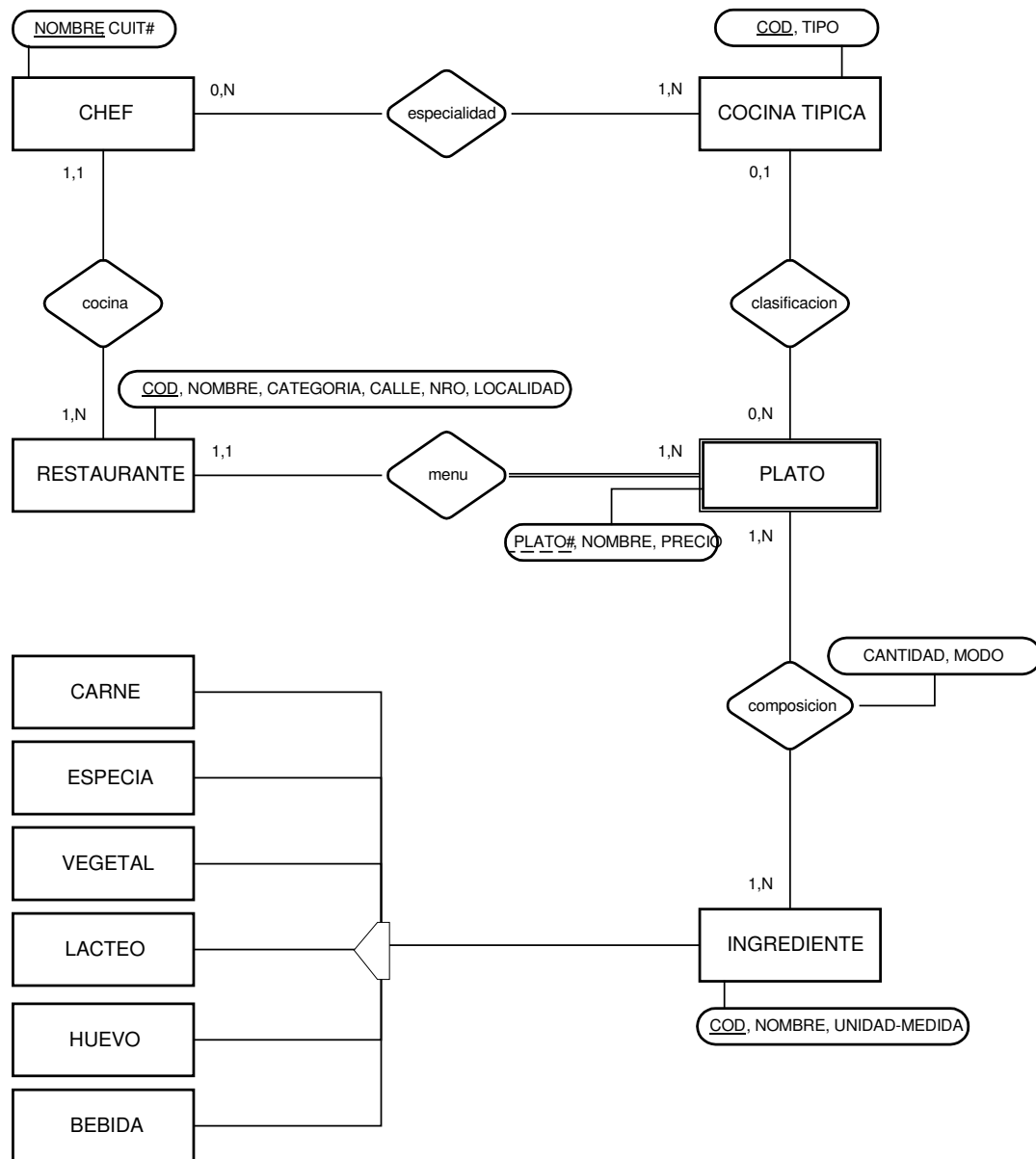
de acuerdo a los siguientes supuestos:

- el número de cliente es único para todas las sucursales,
- existe sólo un número de cliente por cuenta,
- existe sólo un número de cliente por préstamo,
- las cuentas y los préstamos se otorgan por sucursal,

escribir expresiones en Algebra Relacional que satisfagan los siguientes requerimientos de información:

1. El número de sucursal, número de préstamo, número de cliente e importe para los préstamos mayores a \$50.000.
2. El número, el nombre y la ciudad de residencia de todos los clientes que tienen una o más cuentas en la sucursal 17.
3. El nombre de todos los clientes que tienen préstamo, cuenta o ambas cosas en la sucursal 32.
4. El nombre de todos los clientes que tienen tanto préstamos como cuentas en la sucursal 32.
5. El nombre, dirección y ciudad de residencia de todos los clientes que tienen cuentas en todas las sucursales de Córdoba.
6. El número de sucursal y la ciudad de establecimiento de aquellas sucursales para las cuales no existe préstamo alguno.
7. El número de sucursal y la ciudad de establecimiento de los pares de sucursales establecidas en la misma ciudad.
8. El nombre y dirección de cada cliente que opere con sucursales establecidas en cualquier ciudad que no sea la ciudad de residencia del cliente.
9. El nombre y dirección del(los) cliente(s) titular(es) de la(s) cuenta(s) cuyo saldo es el mayor.

**CASO 2:** A partir de una modelización de datos referentes al circuito de Restaurantes “El Colo”, se obtuvo el siguiente Diagrama E-R:



Por transformación a un modelo implementable, se obtuvo el siguiente ESQUEMA RELACIONAL, del cual se ejemplifica parcialmente una posible instancia:

**chef**

NOMBRE_CHEF	CUIT#_TIPO	CUIT#_NRO	CUIT#_V
ANA LIA PERRY	27	33333333	3
CHABUCA PISCO	27	66666666	6
FRANCIS BIENMANN	20	22222222	2
HERR GABEL	20	44444444	4
PATO PLUMAS	20	11111111	1
PIERRE NITTA	20	55555555	5
...	...	...	...

**especialidad**

NOMBRE_CHEF	C_COCINA
ANA LIA PERRY	ARG
ANA LIA PERRY	MED
CHABUCA PISCO	PER
FRANCIS BIENMANN	ARG
HERR GABEL	ALM
HERR GABEL	HNG
PATO PLUMAS	ARG
PIERRE NITTA	FRN
PIERRE NITTA	MED
PIERRE NITTA	VSC
...	...

**cocina-típica**

C_COCINA	TIPO_COCINA
ALM	ALEMANA
ARG	ARGENTINA
FRN	FRANCESA
HNG	HUNGARA
MED	MEDITERRANEA
MEX	MEXICANA
PER	PERUANA
VSC	VASCA
...	...

**restaurante**

C_REST	NOMBRE_REST	CATEGOR	CALLE	NRO	LOCALIDAD	NOMBRE_CHEF
CHOLI	CHOLITO	3 TENED	ALEM	123	BS. AS.	CHABUCA PISCO
GRUNE	GRUNE WELLE	3 TENED	MITRE	757	BERNAL	HERR GABEL
LATAN	LA TANA	3 TENED	CHILE	150	BS. AS.	ANA LIA PERRY
LESPU	LA ESPUELA	3 TENED	P. COLON	850	BS. AS.	FRANCIS BIENMANN
LVEUV	LA VEUVE	3 TENED	MITRE	757	BS. AS.	PIERRE NITTA
MANDU	MAN DUKE	2 TENED	CONDE	50	BS. AS.	HERR GABEL
PITUC	PITUCO	3 TENED	BYRON	150	BS. AS.	PATO PLUMAS
...	...	...	...	...	...	...

**plato**

C_REST	PLATO#	NOMBRE_PLATO	PRECIO	C_COCINA
CHOLI	15	ANTICUCHOS	11	PER
CHOLI	16	TAMALES	10	PER
LESPU	16	LOCRO	8	ARG
LESPU	18	TAMALES	9	ARG
LESPU	24	HUEVOS FRITOS	5	FRN
MANDU	16	GULASCH	9	HNG
LVEUV	20	OMELETTE DE PUERROS	10	FRN
...	...	...	...	...

**composición**

C_REST	PLATO#	C_ING	CANT	MODO
CHOLI	15	A032	20	MOLIDO
CHOLI	15	C050	30	PICADO
CHOLI	15	C192	500	TROZADO
LVEUV	20	H100	5	BATIDO
LVEUV	20	P027	50	PICADO
LVEUV	20	V132	50	NATURAL
LVEUV	20	C050	50	ENTERO
...	...	...	...	...

**ingrediente**

C_ING	NOMBRE_ING	UN_MED	CLASE
A032	AJI ROCOTO	GRS.	ESPECIA
P027	PUERRO	GRS.	VEGETAL
C230	CEBOLLA	GRS.	VEGETAL
V132	VINO CHABLIS	CMS. 3	BEBIDA
M420	MAIZ	GRS.	VEGETAL
H100	HVO. CODORNIZ	UNIDAD	HUEVO
C050	CILANTRO	GRS.	ESPECIA
C192	CORAZON VACA	GRS.	CARNE
...	...	...	...



considerando que se cumplen las Reglas de Integridad del MODELO RELACIONAL, escribir expresiones en Algebra Relacional que satisfagan los siguientes requerimientos de información:

1. El NOMBRE de los ingredientes que componen los platos que clasifican en cocina típica “VASCA”.
2. El NOMBRE de los ingredientes que componen los platos que cocinan los chefs especializados en cocina típica “VASCA”.
3. El (o los) TIPO de cocina típica en que se especializan los chefs que cocinan platos incluyendo el ingrediente “CILANTRO”.
4. El TIPO de cocina típica en la cual no clasifica plato alguno, ni se especializa Chef alguno.
5. El NOMBRE de los restaurantes para los cuales se cumple que su menú incluye platos que clasifican en alguna de las cocinas típicas en las cuales se especializa el chef que cocina en el restaurante.
6. El NOMBRE de los restaurantes para los cuales se cumple que su menú incluye platos que clasifican en alguna cocina típica en la cual no se especializa el chef que cocina en el restaurante.
7. La CLASE de ingredientes que participa en la composición de todos los platos del circuito de restaurantes.
8. El NOMBRE de los ingredientes que son utilizados, para la composición de platos, por todos los chefs.
9. El NOMBRE de los ingredientes que participan, en la composición de platos, en todas las cocinas típicas.
10. Para cada uno de los restaurantes, el (o los) plato/s de su menú que en su composición incluye/n todos los ingredientes que utilizan los platos de ese restaurante.

## APENDICE

### TRANSFORMACION DEL MODELO E-R A OTROS MODELOS IMPLEMENTABLES

El problema que se nos presenta es el siguiente: dado un Modelo Conceptual con la forma de un MER ¿cómo diseñar estructuras para representar los datos correspondientes?.

En el Modelo Conceptual, los relacionamientos son pares (o ternas, etc.) abstractos, que relacionan puntos abstractos de conjuntos. En los Modelos Operacionales es preciso transformar esas abstracciones en datos concretos, representables en un computador.

En general, cada conjunto de entidades  $E$  estará representado en el Modelo Operacional por un conjunto de datos  $D$ . Cada entidad de  $E$  estará representada por un elemento de  $D$ .

En lo que refiere a los conjuntos de relacionamientos, existen dos maneras, según los distintos Modelos Operacionales, de representar los relacionamientos conceptuales:

- a) la representación *Implícita*: sin acrecentar los datos
- b) la representación *Explícita*: introduciendo datos específicos para representar el relacionamiento

### TRANSFORMACION DE UN MODELO E-R A TABLAS

#### TIPOS DE ENTIDADES REGULARES

Sea  $E$  un conjunto de entidades fuertes con atributos descriptivos  $A_1, A_2, \dots, A_n$ . Podemos representar este conjunto de entidades mediante una *tabla*  $T$  con  $n$  *columnas*, cada una de las cuales corresponde a un atributo de  $E$ . Cada *fila* (tupla) en esta tabla corresponde a una entidad del conjunto de entidades  $E$ .

Ejemplo:

PAÍS (nombre\_país, cant\_hab\_país)  
TALLER (nro\_taller, dirección)

#### Países

NOMBRE_PAÍS	CANT_HAB_PAÍS

#### talleres

NRO_TALLER	DIRECCIÓN

**TIPOS DE ENTIDADES DÉBILES**

Sea  $D$  un conjunto de entidades débiles con atributos descriptivos  $D_1, D_2, \dots, D_n$ . Sea  $E$  el conjunto de entidades fuertes del que depende  $D$ . Sean  $A_1, A_2, \dots, A_k$  los atributos que constituyen la clave primaria de  $E$ . Podemos representar el conjunto de entidades  $D$  mediante una *tabla T* con  $n + k$  *columnas*, cada una de las cuales corresponde a un atributo del conjunto  $\{(A_1, A_2, \dots, A_k) \cup (D_1, D_2, \dots, D_n)\}$ .

Ejemplo:

PAÍS (nombre\_país, cant\_hab\_país)

PROVINCIA (nombre\_provincia, cant\_hab\_provincia)

**Provincias**

NOMBRE_PAÍS	NOMBRE_PROVINCIA	CANT_HAB_PROVINCIA

**TIPOS DE RELACIONAMIENTO**

Sea  $R$  un conjunto de relacionamientos  $n$ -arios entre los conjuntos de entidades  $E_1, E_2, \dots, E_n$ .

Sea  $\{K E_i\}$  el conjunto de atributos que conforman la clave primaria del conjunto de entidades  $E_i$ .

Sean  $A_1, A_2, \dots, A_p$  los atributos descriptivos propios del relacionamiento  $R$ . Podemos representar el conjunto de relacionamientos  $R$  mediante una *tabla T* con una *columna* por cada atributo del siguiente conjunto :

$$\left\{ \bigcup_{i=1}^n \{K E_i\} \cup (A_1, A_2, \dots, A_p) \right\}$$

Ejemplo:

TREN (nro\_tren, denominación)

FECHA (día, mes, año)

TRAMO (nro\_tramo, distancia)

VIAJE\_TRAMO (cant\_asientos\_disponibles, horario\_llegada, horario\_salida)

**Viaje\_tramo**

NRO_TREN	NRO_TRAMO	DIA	MES	AÑO	CANT_ASIENT_DISP	H_LLEGA	H_SALE

**GENERALIZACION Y ESPECIALIZACION**

Existen dos variantes para la Generalización, sólo la primera de ellas es válida también para la Especialización :

1. Crear una *tabla* para el tipo de entidad de mayor nivel (más general), tal como se especificó para los tipos de entidades regulares. Crear una *tabla* para cada tipo de entidad de menor nivel (menos general), que incluya una *columna* para cada uno de los atributos descriptivos, más una *columna* por cada atributo de la clave primaria del tipo de entidad más general.

Ejemplo:

EMPLEADO (nro\_empleado, nombre\_empleado, dirección\_empleado)

EMPLEADO\_TALLER (especialidad)

EMPLEADO\_TREN (jornada, tipo\_trabajo)

DIRECTIVO (profesión, cargo)

**Empleados**

NRO_EMPLEADO	NOMBRE_EMPLEADO	DIRECCION_EMPLEADO

**Empleados\_taller**

NRO_EMPLEADO	ESPECIALIDAD

**Empleados\_tren**

NRO_EMPLEADO	JORNADA	TIPO_TRABAJO

**Directivos**

NRO_EMPLEADO	PROFESION	CARGO

2. Crear una tabla para cada tipo de entidad de menor nivel (menos general), que incluya una columna para cada uno de sus atributos descriptivos, más una columna por cada atributo del tipo de entidad más general. En este caso no se crea tabla para el tipo de entidad de mayor nivel (más general).

Ejemplo:

EMPLEADO (nro\_empleado, nombre\_empleado, dirección\_empleado)

EMPLEADO\_TALLER (especialidad)

EMPLEADO\_TREN (jornada, tipo\_trabajo)

DIRECTIVO (profesión, cargo)

### **Empleados\_taller**

NRO_EMPLEADO	NOMBRE_EMPLEADO	DIR_EMPLEADO	ESPECIALIDAD

### **Empleados\_tren**

NRO_EMPLEADO	NOMBRE_EMPLEADO	DIR_EMPLEADO	JORNADA	TIPO_TRABAJO

### **Directivos**

NRO_EMPLEADO	NOMBRE_EMPLEADO	DIR_EMPLEADO	PROFESION	CARGO