

Bases de Datos I

Unidad II

Diseño de bases de datos
Modelo conceptual



Modelado de datos

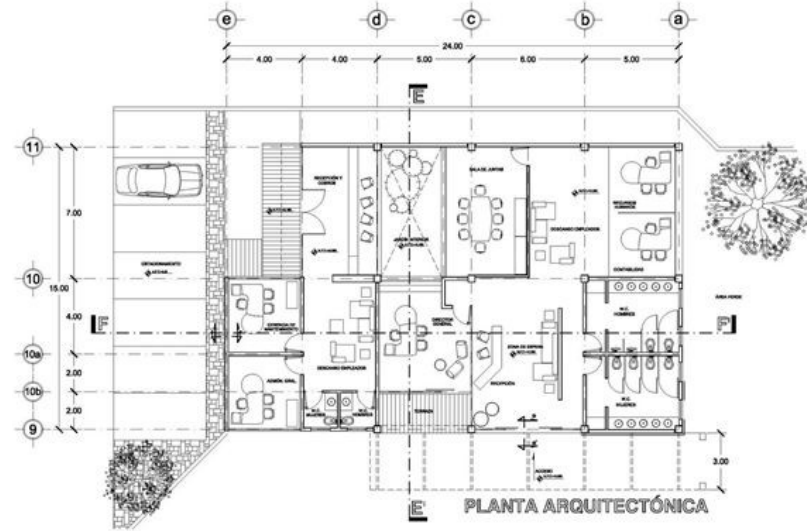
Se modela para:

- **Obtener la perspectiva de cada persona implicada en el problema.**
- **Observar la naturaleza y necesidad de cada dato.**
- **Observar cómo cada persona utiliza cada dato.**

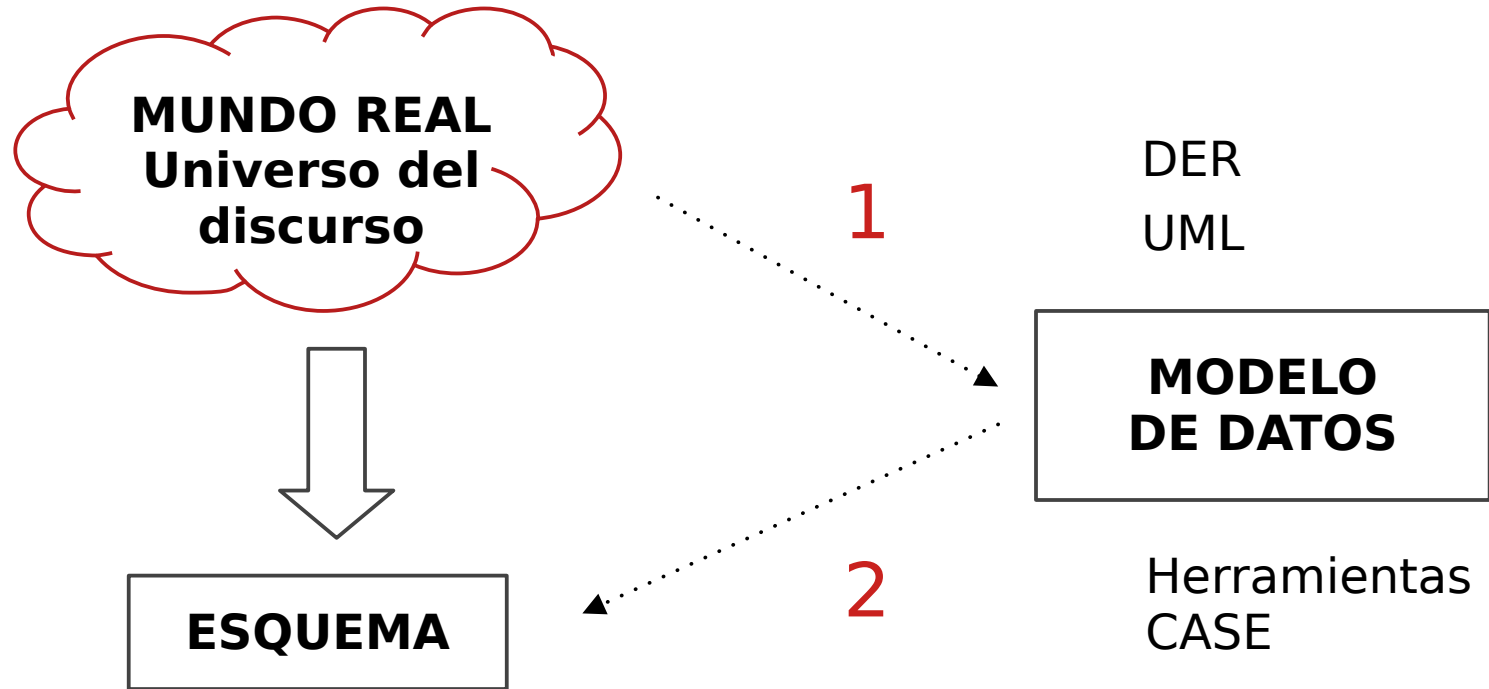
Modelos

Los modelos son medios para describir la realidad.

- **Las personas que diseñan usan los modelos para construir esquemas, que sintetizan la realidad del problema.**



Modelado de datos



Universo del discurso

- **El dominio de discurso, también llamado universo de discurso, o simplemente dominio, es el conjunto de cosas acerca de las cuales se habla en un determinado contexto.**



<https://www.flaticon.com/free-icons/universe>

Modelo y esquema

- **La descripción de un cierto mundo real o universo del discurso (p.e. una Universidad, un Comercio) por medio de un modelo de datos da como resultado un esquema.**
- **El esquema es la descripción de la estructura de la base de datos (tablas, atributos, restricciones, etc., “metadatos”) y es relativamente invariante en el tiempo.**

Diseño de bases de datos

- **La metodología de diseño propuesta en este curso consiste en construir un modelo de datos en etapas:**
- **Diseño conceptual:** se lleva a cabo luego de la etapa de especificación de requerimientos, es independiente del SGBD a utilizar y da como resultado el esquema conceptual.
- **Diseño lógico:** tiene dos fases, una de refinamiento del esquema conceptual y otra para obtener el esquema de la BD que depende del SGBD a utilizar. Usaremos el Modelo Relacional.
- **Diseño Físico:** es necesario tomar decisiones específicas, que tienen que ver con el SGBD específico a utilizar (índices, clusters, archivos).

Diseño de bases de datos

- **Vamos a seguir una aproximación similar a la que se llevaría en un proyecto arquitectónico, para la construcción de una casa.**
- **Vamos a hacerlo en tres etapas:**
 - 1.- Maquetamos (Diseño conceptual)**
 - 2.- Hacemos los planos (Diseño lógico)**
 - 3.- Construimos la casa (Diseño físico)**

Metáfora tomada del profesor Rodolfo Bertone.

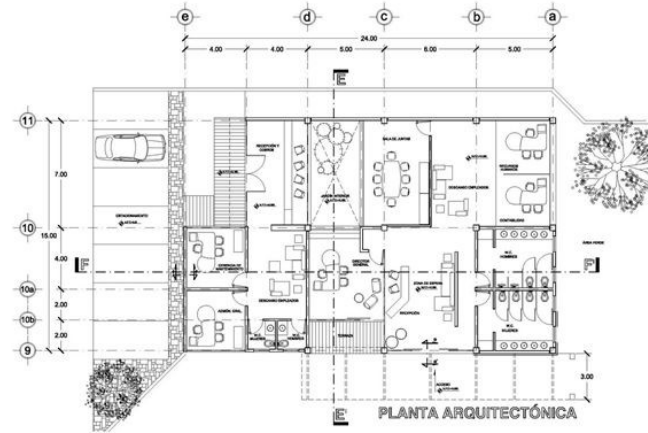
Diseño de bases de datos

MAQUETA



Diseño conceptual

PLANOS



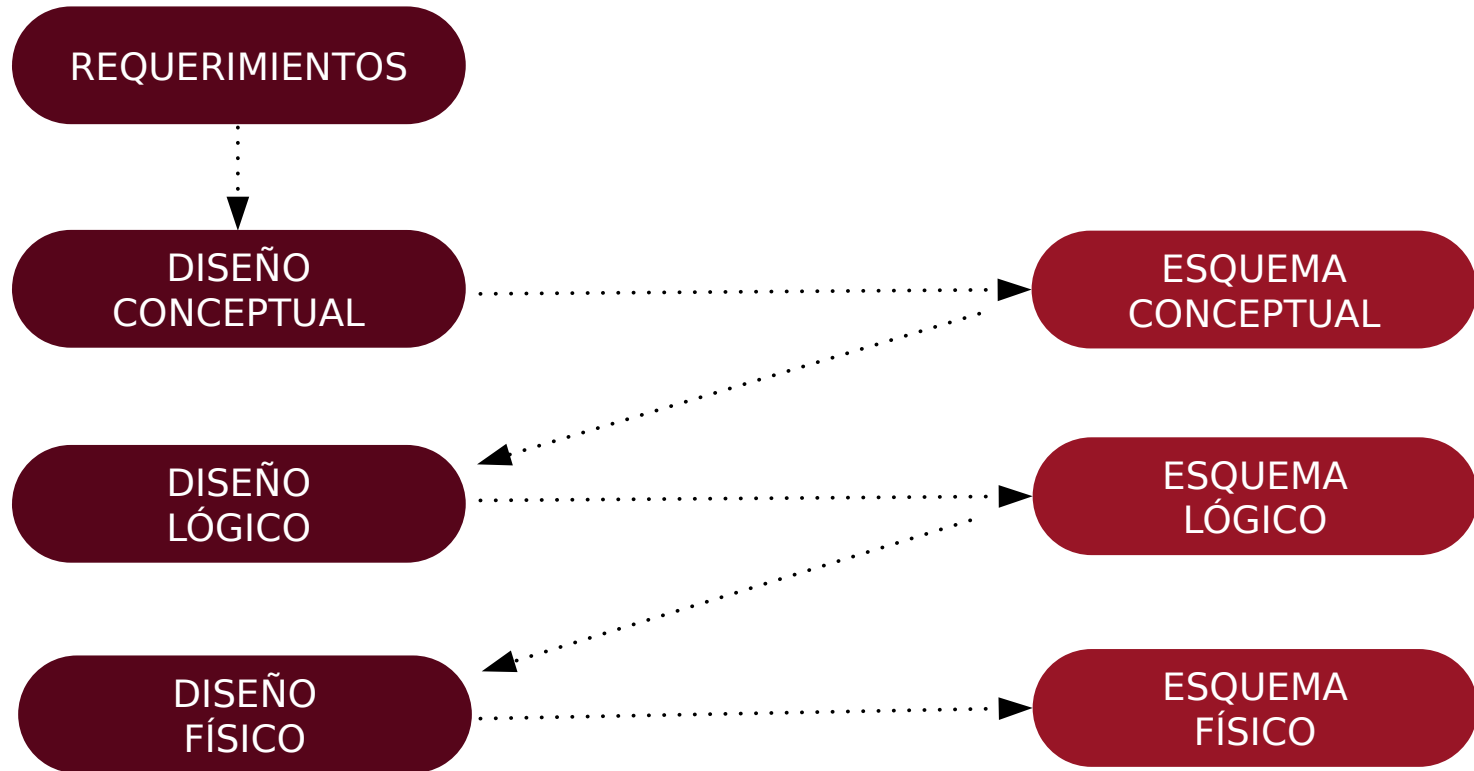
Diseño lógico

CASA



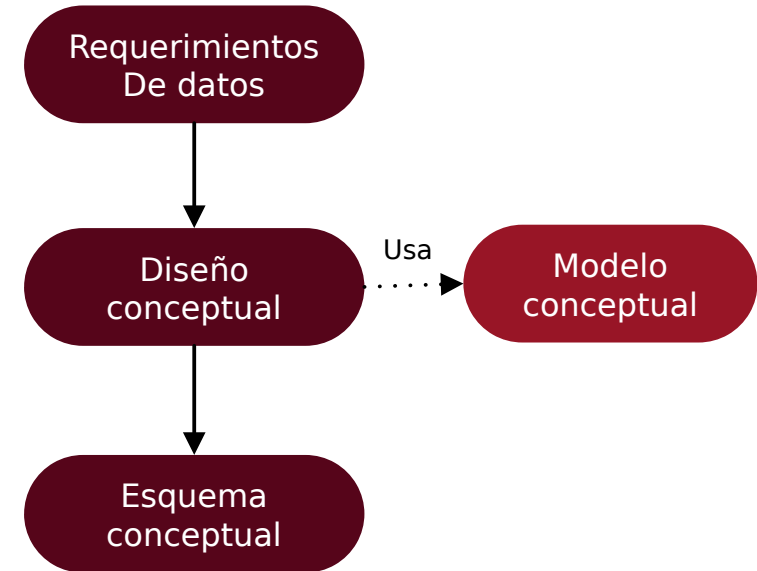
Diseño físico

Diseño de Sistemas de Información



Diseño conceptual

- El propósito del diseño conceptual es describir el contenido de información de la base de datos, más que las estructuras de almacenamiento que se necesitarán para manejar esta información.
- Parte de la especificación de requerimientos y su resultado es el esquema conceptual de la base de datos.



Modelo conceptual

- **Características que debe cumplir el modelo conceptual:**

EXPRESIVIDAD

Capturar y presentar de la mejor forma posible la semántica de los datos del problema a resolver

FORMALIDAD

Cada elemento representado en el modelo sea preciso, bien definido y con una sola interpretación posible

MINIMALIDAD

Cada elemento del modelo tiene una única forma de representación.

SIMPLICIDAD

El modelo debe ser fácil de entender por todas las personas involucradas (clientes, analistas, etc.)



Diseño conceptual

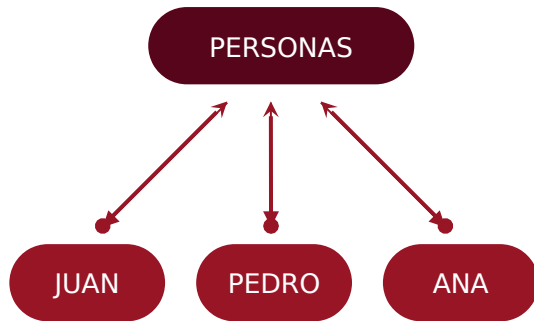
- **Un esquema conceptual, es una descripción de alto nivel de la estructura de la base de datos, independiente del software de SGBD que se use para manipularla.**
- **Un modelo conceptual, es un lenguaje que se usa para describir esquemas conceptuales.**
- **Para construir el esquema conceptual vamos a utilizar el Modelo Entidad Interrelación (ERM).**

Diseño y abstracción

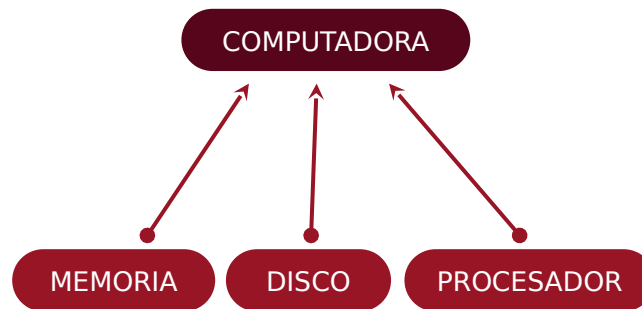
- **En el diseño, vamos a realizar un proceso de abstracción.**
- **La abstracción es un proceso que permite seleccionar algunas características de un conjunto de objetos (físicos o conceptuales) del mundo real, dejando de lado características que no son de interés o no son relevantes.**
- **La abstracción es un proceso que permite seleccionar algunas características del mundo real (universo del discurso), dejando de lado aquellos rasgos que no son de interés.**

Diseño y abstracción

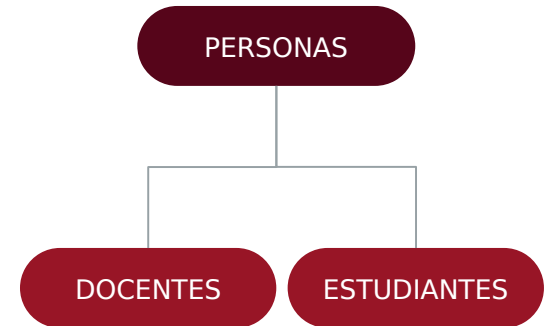
Para diseñar o modelar se utilizan distintas clases de abstracciones:



CLASIFICACIÓN



AGREGACIÓN



GENERALIZACIÓN

Clasificación

- **Se utiliza para definir una clase (declaración de objetos).**
- **La clase se origina a partir de las características comunes que tienen los objetos que la comparten.**
- **La clasificación permite identificar los atributos de los elementos.**
- **Un objeto puede clasificarse de varias formas.**

Agregación

- **Una abstracción de agregación define una nueva clase a partir de un conjunto de otras clases que representan sus partes componentes.**
- **Si se elimina la clase agregada, se eliminan además las otras clases que la forman.**

Generalización

- **Una abstracción de generalización define una relación de subconjunto entre los elementos de dos o más clases.**
- **Es muy útil por su cualidad fundamental de herencia: en una generalización, todas las abstracciones definidas para la clase genérica son heredadas por las clases subconjunto.**

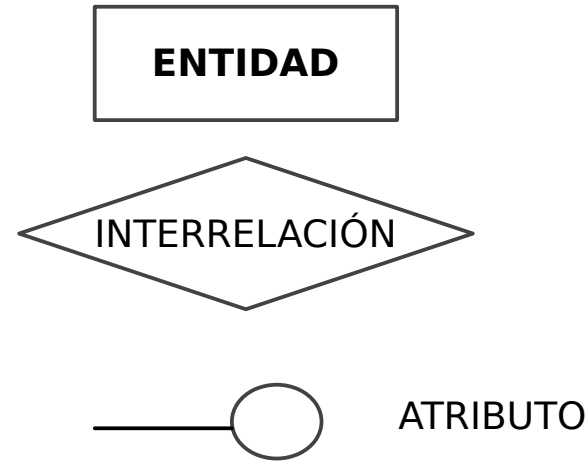
Modelo de entidad interrelación (DER)

- **El modelo entidad interrelación fue propuesto en 1976 por Peter Chen en su trabajo:**
“The Entity-Relationship Model - Toward a Unified View of Data”.
- **Este modelo puede ser usado como base para una vista unificada de los datos, adoptando un enfoque natural del mundo real que consiste en entidades e interrelaciones.**
- **Posteriormente otros autores lo han extendido con importantes aportes, lo que ha dado lugar a una familia de modelos de datos.**

Modelo de entidad interrelación (DER)

- Es un modelo gráfico que consta de los siguiente elementos:

- **Entidades**
- **Interrelaciones**
- **Atributos**

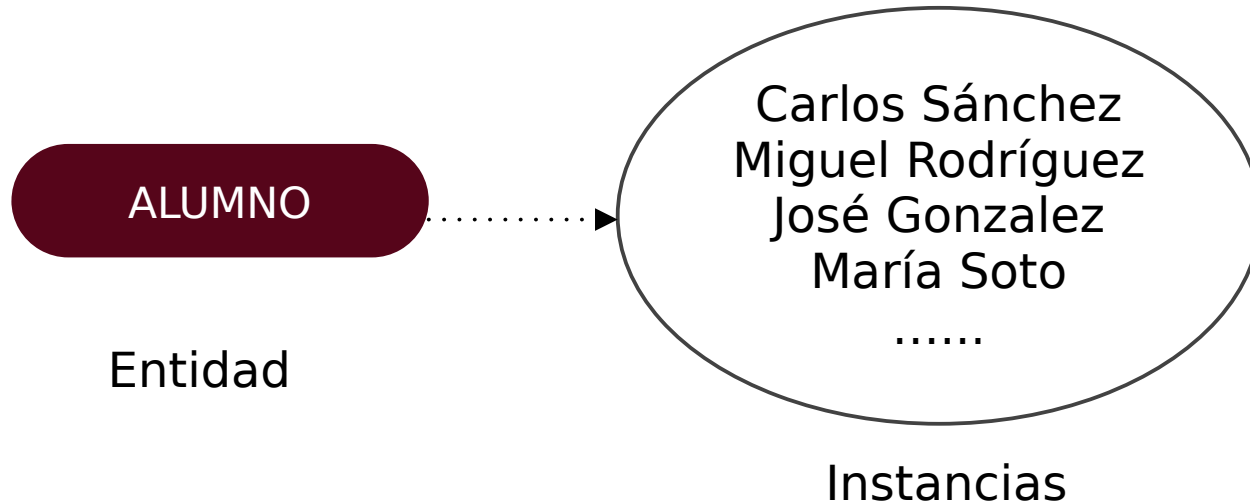


Entidades

- **Elemento del modelo que representa una cosa u objeto (sustantivo) del mundo real que existe, y pertenece al problema en cuestión.**
- **Puede tener existencia física (Ej: un automóvil, un empleado) o existencia conceptual (Ej: una categoría, una materia).**
- **Una entidad se caracteriza por un conjunto de atributos.**

Entidades

- Para definir o encontrar entidades, aplicamos un proceso de abstracción de clasificación.



Entidades

- **Aparecen normalmente reflejadas como sustantivos.**
- **A cada una de las posibles ocurrencias de la entidad se la denomina “ejemplar”.**
- **Tiene características o propiedades específicas (atributo) que la describe.**

Entidades

- **Ejemplo**, existe un estudiante con:

Nombre: Juan

Apellido: Pérez

DNI: 20.123.456 **Legajo:** 234569/5

Dirección: Rivadavia 456 **Teléfono:** 2901-444897

Mail: juan_perez@gmail.com

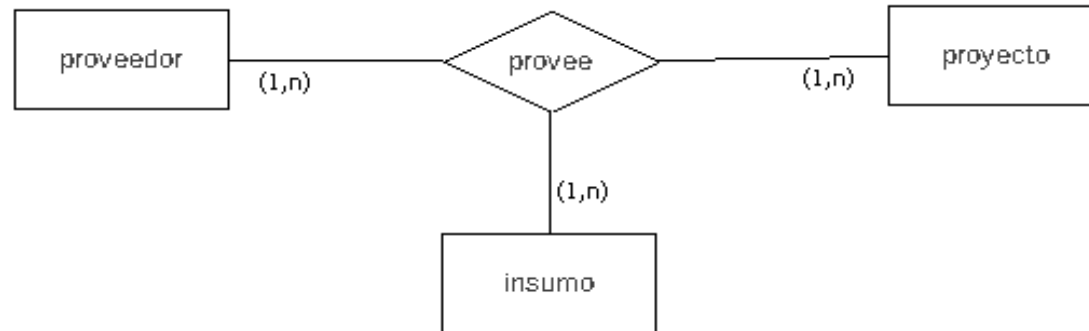
Interrelaciones

- **Una interrelación es una asociación, vinculación o correspondencia entre instancias de entidades.**
- **El nombre de una interrelación debe ser significativo y único para el esquema.**
- **Las interrelaciones se denotan con un rombo.**
- **Además del nombre, cada interrelación tiene:**
 - **Grado**
 - **Cardinalidad**
 - **Dependencia**
 - **Conectividad**

Grado de una interrelación

- **El grado, es el número de instancias de entidades que vincula la misma.**
- **Si clasificamos las interrelaciones según su grado, podemos dividir las en:**
 - **Interrelaciones binarias (grado = 2)**
 - **Interrelaciones n-arias (grado > 2)**

Grado de una interrelación



Cardinalidad

- Es el número mínimo y máximo de veces que la instancia de una entidad, participa de una interrelación.
- La cardinalidad mínima o participación (dependencia):
 - Total (participación obligatoria)
 - Parcial (participación no obligatoria u opcional)

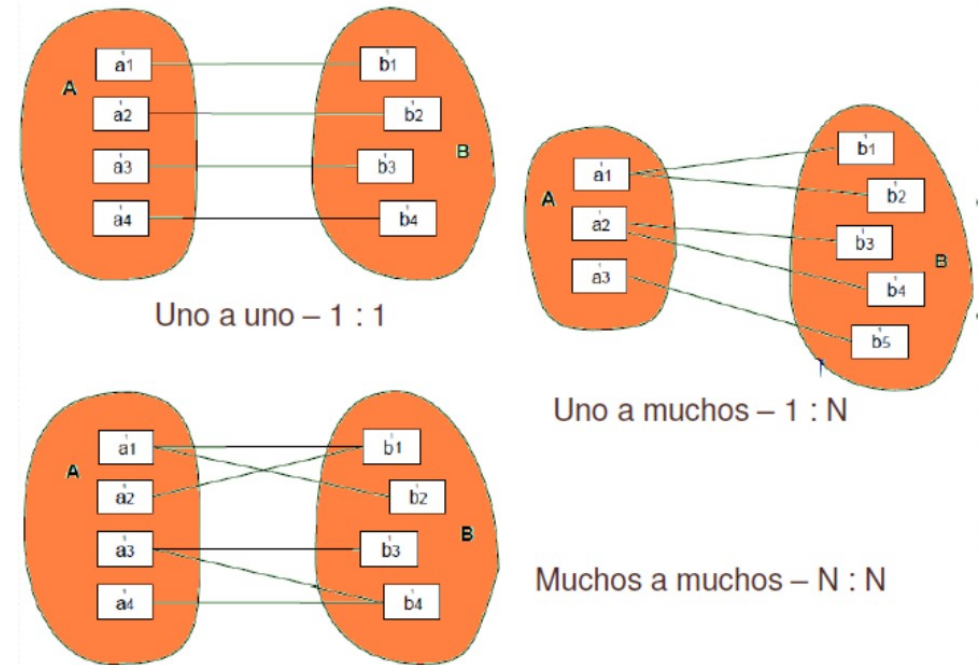
Como mínimo puede ser 0 (participación opcional) y como máximo 1 (participación obligatoria).

- La cardinalidad máxima (conectividad):
 - Uno a uno - 1:1
 - Uno a muchos - 1:N
 - Muchos a muchos - N:M
 -

Como mínimo 1 y como máximo N.

Cardinalidad

Cuando encontramos una abreviación como 1:1, 1:N o N:N; esta hace referencia sólo a las cardinalidades máximas.



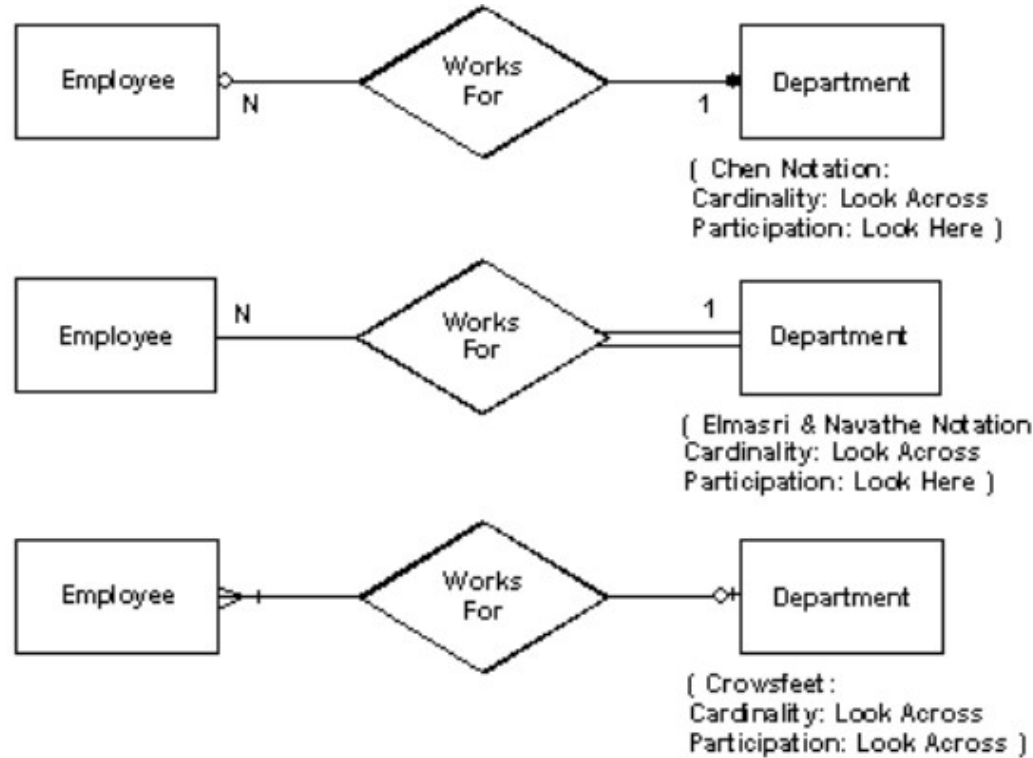
Cardinalidad

- **Existen distintas notaciones para especificar la mínima y máxima:**
 - Numérica
 - Gráfica
- **En cada notación varia la ubicación de la cardinalidad, lo que determina la forma de leer la misma:**
 - Mira al otro lado (Look across)
 - Mira aquí (Look here)

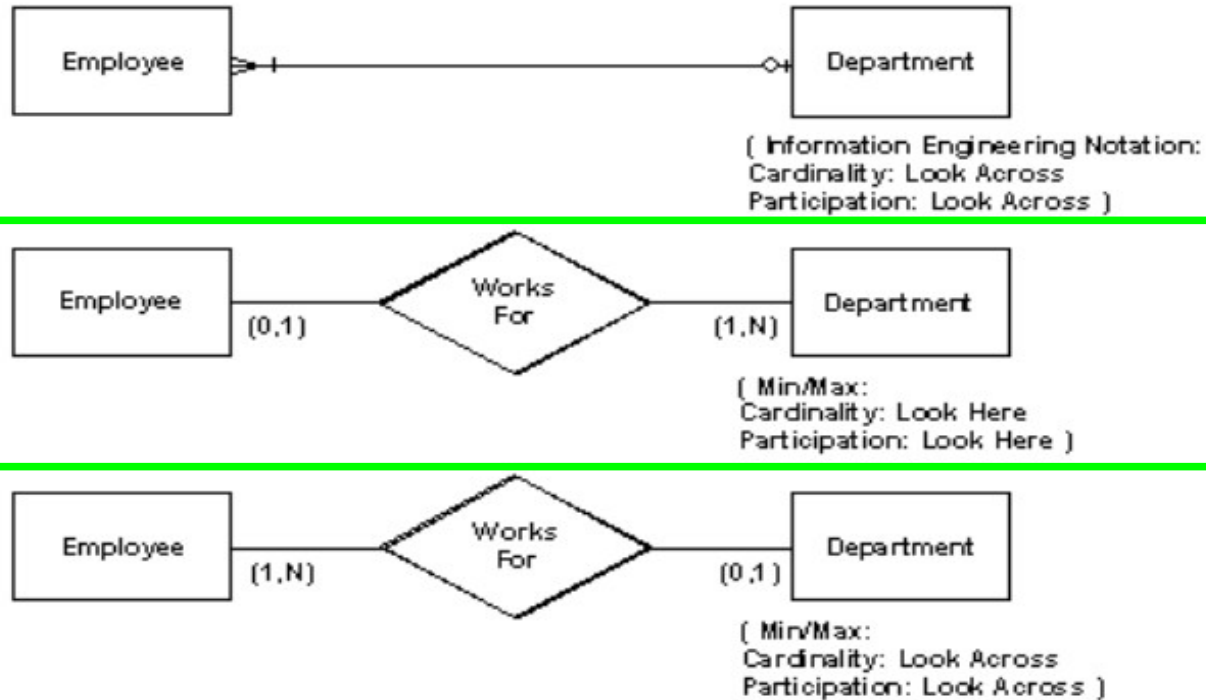
Notación

- **Vamos a expresar la siguiente interrelación:**
 - Un **empleado** *trabaja* para **departamento**.
- **Cardinalidad:**
 - Un empleado **puede o no** trabajar **como máximo** en un departamento.
 - En cada departamento trabaja **como mínimo** un empleado, aunque por lo general **trabajan varios**.

Notación

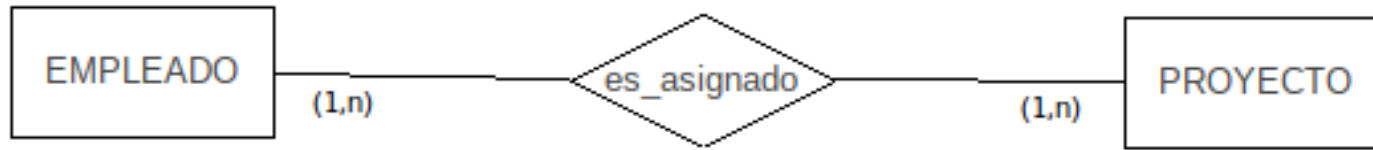


Notación



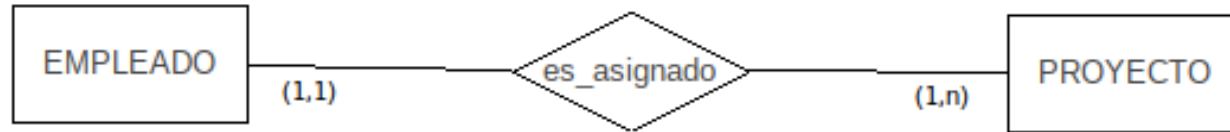
Interrelación N:M

- Los empleados son asignados a proyectos.



Interrelación 1:N

- Cada empleado trabaja en un único departamento de forma obligatoria.



Interrelación 1:N

- Los empleados pueden o no afiliarse a sindicatos.



Interrelación 1:1

- Los empleados deben tener licencia de conducir.



Interrelación 1:1

- Los empleados pueden tener licencia de conducir.



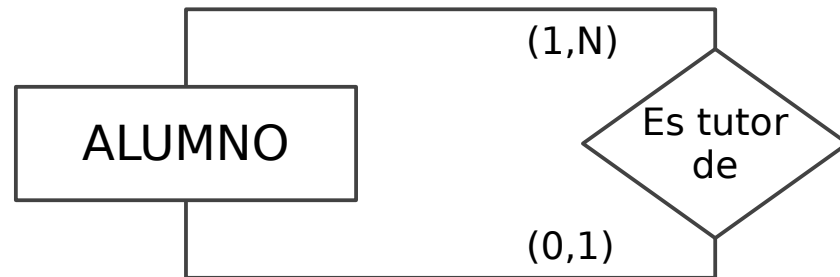
Interrelación 1:1

- Los empleados pueden tener asignada una computadora. Las computadoras pueden no estar asignadas a nadie.



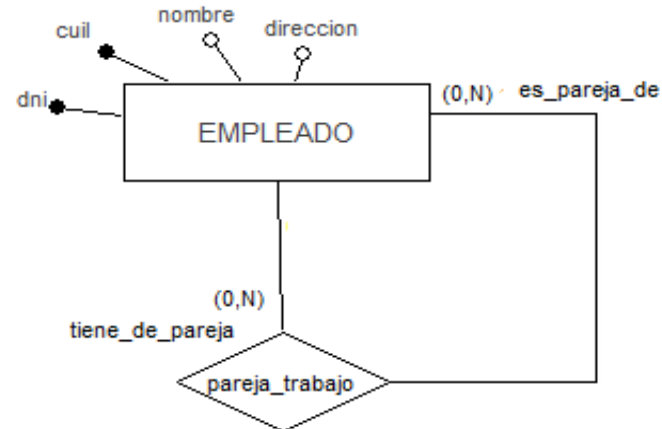
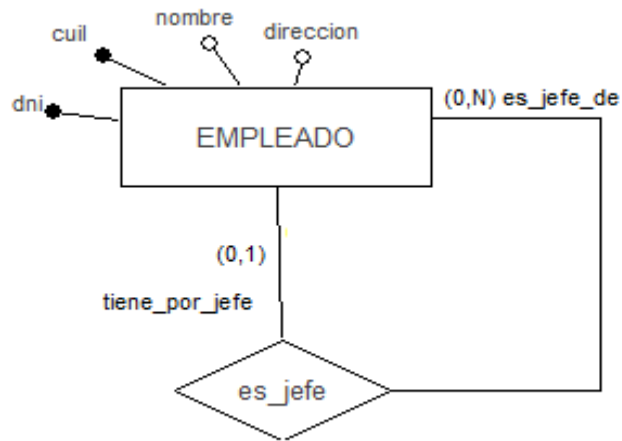
Interrelación recursivas

- Un caso especial de interrelación binaria, entre instancias de la misma entidad.
- Como se interpreta:
 - Cada alumno **puede tener o no, como máximo** un tutor.
 - Cada alumno debe ser tutor de **al menos un** alumno, aunque **puede ser de más de uno**.



Interrelación recursivas

- **Ejemplos:**
- Hay empleados que son jefes de otros empleados.
- Los empleados tienen una pareja de trabajo en equipo.



Atributos

- **Los atributos describen las características o propiedades relevantes de cada entidad o interrelación.**
- **Cada instancia de una entidad o interrelación tiene un valor particular para cada atributo. Conjunto de valores permitidos para un atributo (dominio).**
- **Todas las instancias de una entidad o interrelación se describen mediante el mismo conjunto de atributos.**
- **Los atributos tienen dos funciones: identificadores o descriptores**

Ejemplo

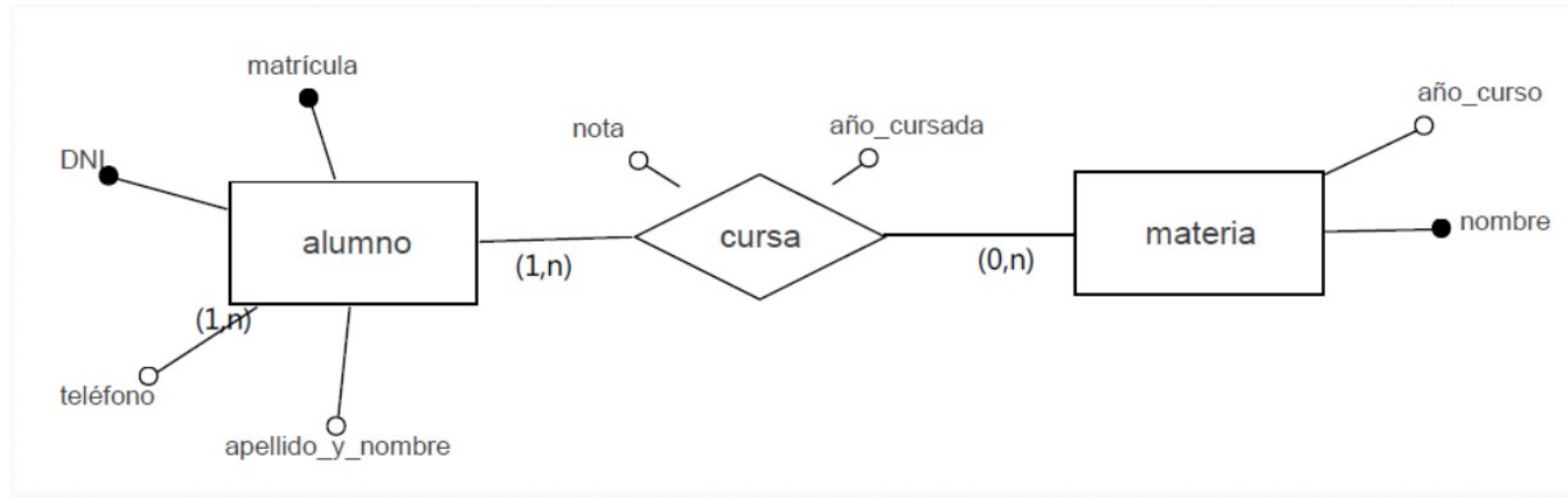
- **Ejemplo: Entidad ESTUDIANTES**
- **Atributos: dni, legajo, email, nombre, apellido, teléfono, dirección**
- **Atributos descriptores: nombre, apellido, teléfono, dirección**
- **Atributos identificadores: dni, legajo, mail**

Dominio

- **El dominio se puede definir:**
- **Por intención: indicando el tipo de dato**
precio: “Números reales mayores que cero”.
- **Por extensión: definiendo los valores del conjunto**
sexo = {M, F, X}

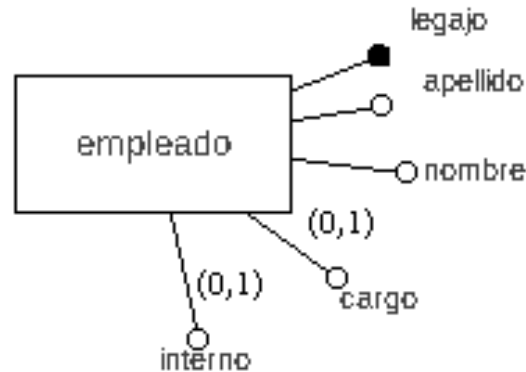
Atributos

- Para definir o encontrar atributos, aplicamos un proceso de abstracción de clasificación y agregación.



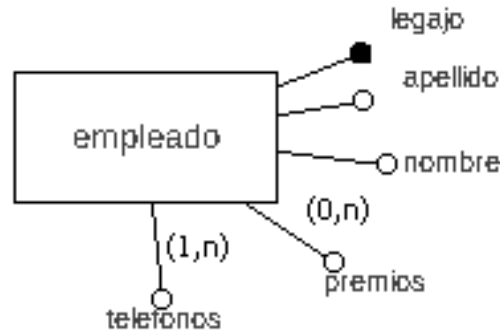
Atributos: cardinalidad

- Si la cardinalidad **mínima** es **1**, se dice que el atributo es **obligatorio**.
- Si la cardinalidad **mínima** es **0**, se dice que el atributo es **opcional**.



Atributos: cardinalidad

- Si la cardinalidad **máxima** es **1**, se dice que el atributo es **monovaluado** o **univaluado**.
- Si la cardinalidad **máxima** es **N**, se dice que el atributo es **multivaluado**.



Atributos: cardinalidad

Algunos atributos pueden no tener un valor específico para una entidad, estos atributos tendrán **valores nulos**.

¿Qué significa aceptar un valor nulo?

- **Valor no aplicable:** no tiene sentido dar un valor de atributo para el ejemplar de la entidad.
- **Valor perdido:** el valor existe pero no se tiene.
- **Valor no conocido:** no se sabe si existe o no hay valor para el ejemplar.

Modelo de entidad interrelación (DER)

- **Otros elementos del modelo:**

- **Atributos compuestos**



- **Atributos identificadores**



- **Atributos calculados**

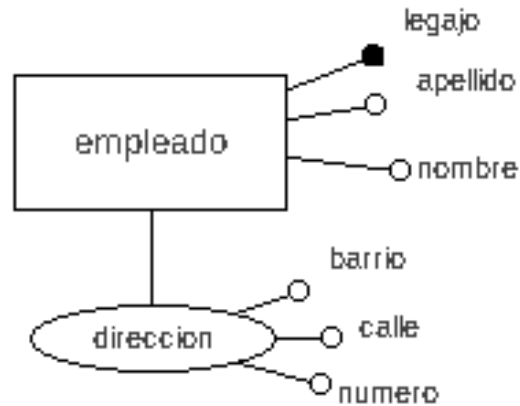


- **Generalizaciones**



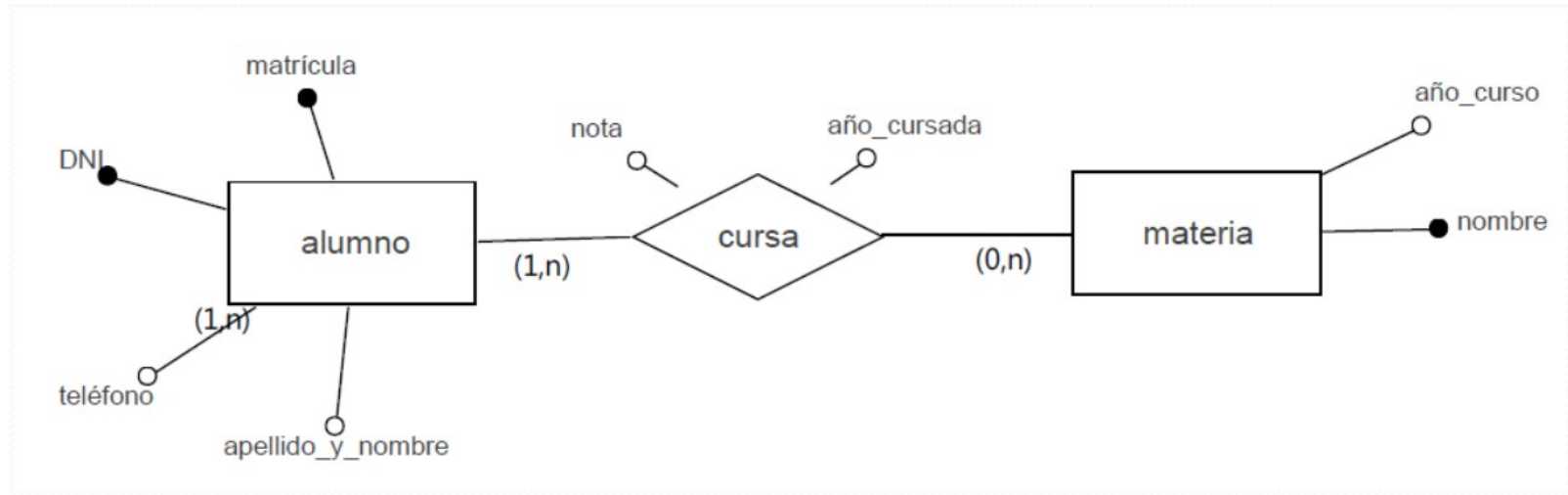
Atributos compuestos

- **Son aquellos atributos que se pueden dividir en componentes más pequeños. Están formados a su vez por atributos simples o compuestos.**



Atributos identificadores

- **Atributo (simple o compuesto) cuyo valor siempre es distinto para cada instancia particular de una entidad.**



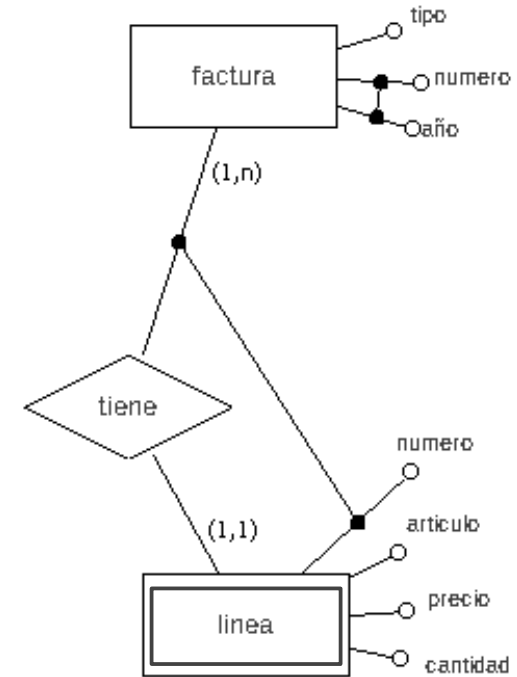
Atributos identificadores

- **Tiene la siguientes características:**
 - **Restricción de unicidad del valor del atributo.**
 - **Sirve para identificar de manera única a cada ejemplar de una entidad.**
 - **Tiene que tener un valor conocido para cada entidad, no puede ser nulo.**



Identificador externo

- **Existen entidades que necesitan de otras para poder identificarlas.**
- **Se identifican con identificadores externos.**
- **Se conocen como entidades débiles.**

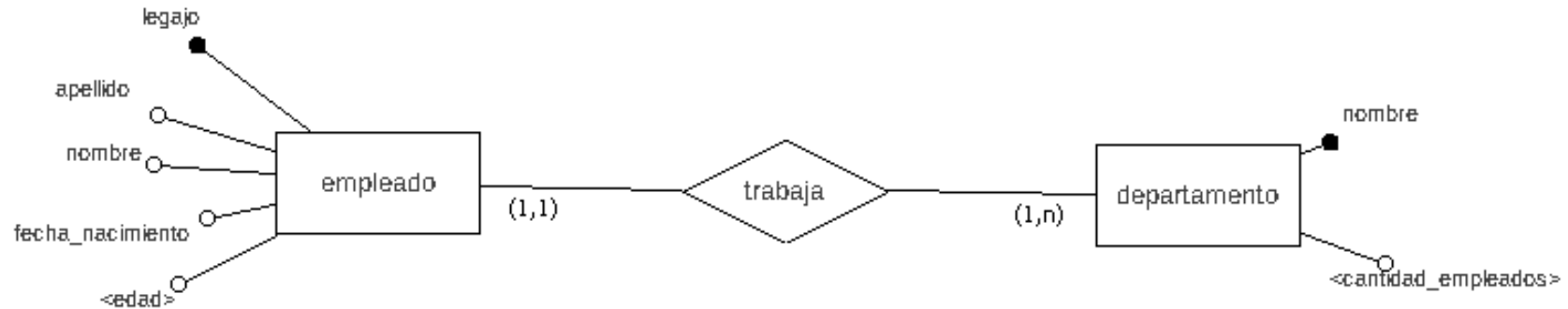


Entidades débiles

- **La dependencia de identificación se da entre dos entidades E1 y E2, donde E1 tiene dependencia de existencia de E2.**
- **Una instancia de E1 depende de la existencia de una instancia de E2.**
- **Si una instancia de E2 se borra, se borrarán todas las instancias dependientes de ella en E1.**

Atributos calculados o derivados

- **Son aquellos que pueden ser calculados a partir de otros atributos de la misma entidad o entidades interrelacionadas a la misma.**

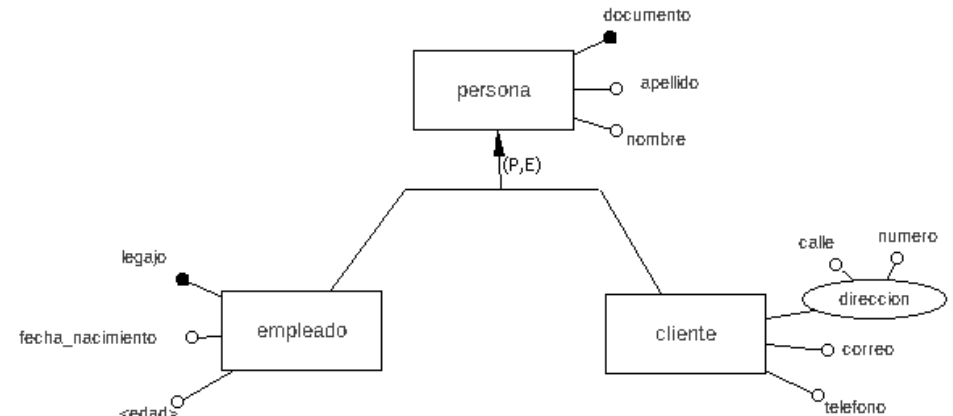


Generalización

- **La generalización permite extraer propiedades comunes de varias entidades y generar con ellas una superentidad (o supertipo) que las agrupe.**
- **Las características compartidas se expresan una única vez en la superentidad o supertipo.**
- **Las características particulares se definen en la subentidad o subtipo.**

Generalización

- Con la generalización, creamos jerarquías de entidades.
- Solo las entidades supertipo requieren de identificadores, las entidades subtipo no.



Propiedades de la generalización

- **Propiedad de cobertura:**

Indica el grado de relación entre supertipos y subtipos.

- **Existen dos tipos de coberturas:**

- Total o Parcial
- Exclusiva o Superpuesta

Cobertura

- **Cobertura total y parcial:**

En la **cobertura total**, cada instancia del padre, es instancia de alguno de los hijos.

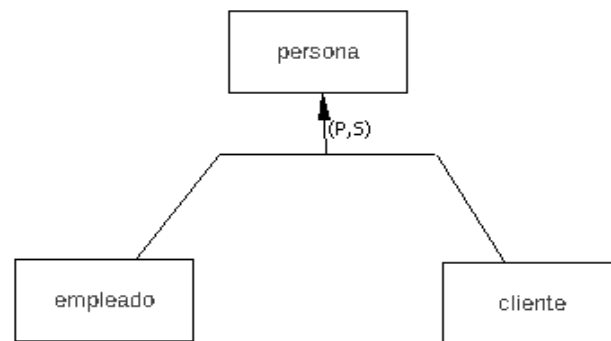
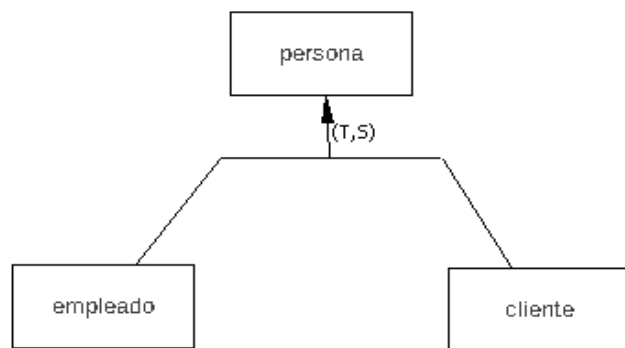
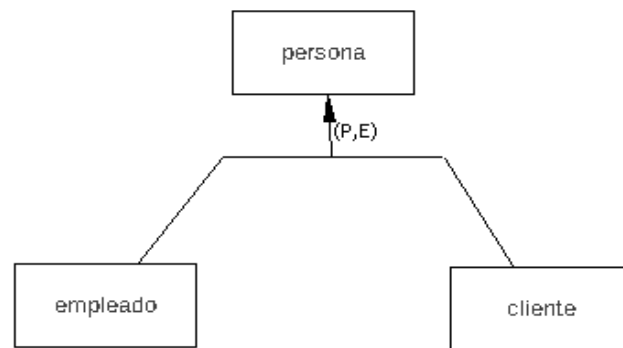
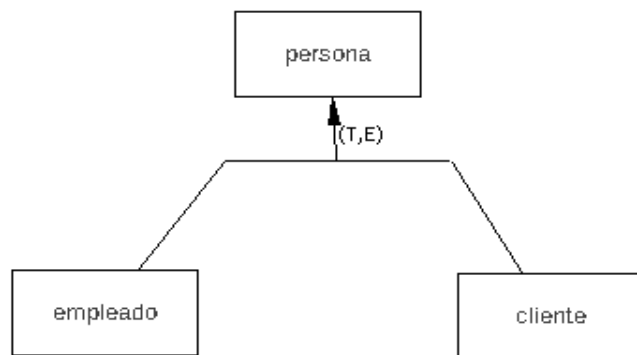
En la **cobertura parcial**, pueden existir instancias del padre que no son instancia en ninguno de los hijos.

- **Cobertura exclusiva y superpuesta:**

La **cobertura exclusiva**, implica que una instancia del padre puede ser instancia de un solo hijo.

La **cobertura superpuesta**, implica que una instancia del padre puede ser instancia de más de un hijo.

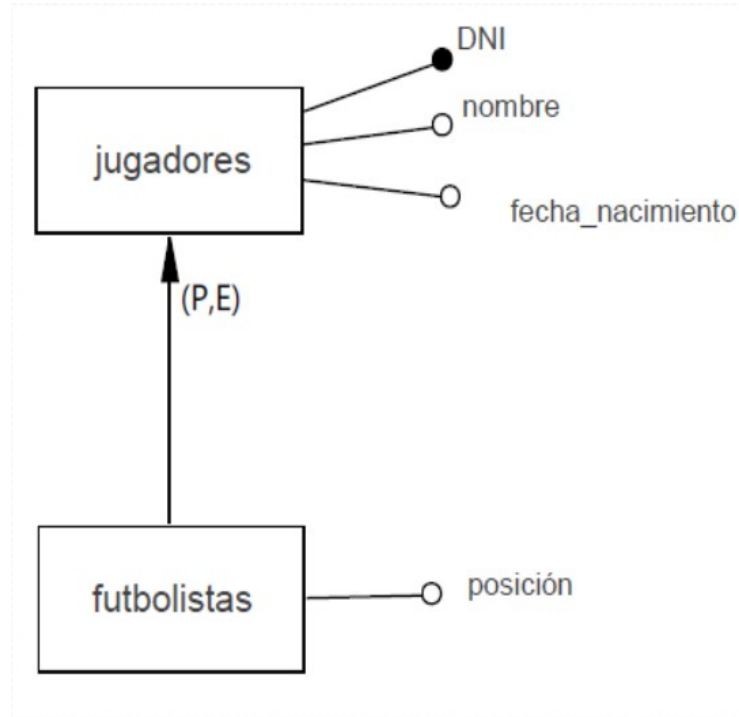
Cobertura



Subconjuntos

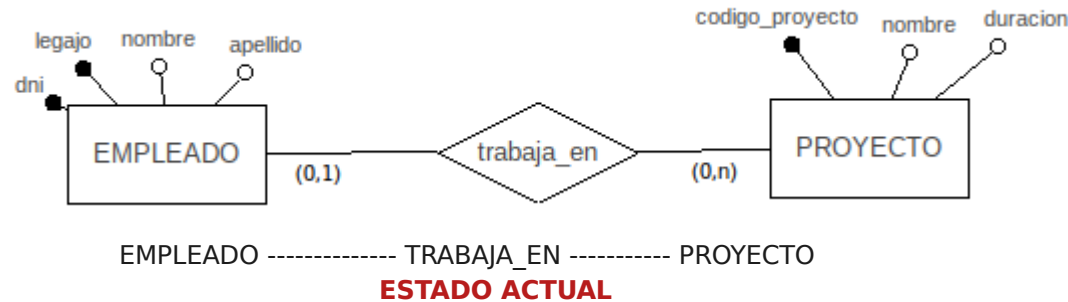
- **Los subconjuntos constituyen un caso especial de las jerarquías de generalización.**
- **Si la generalización contiene un único subtipo, se denomina subconjunto.**
- **La cobertura de los subconjuntos siempre es Parcial y Exclusiva.**

Ejemplo de subconjunto



Incluyendo la historia

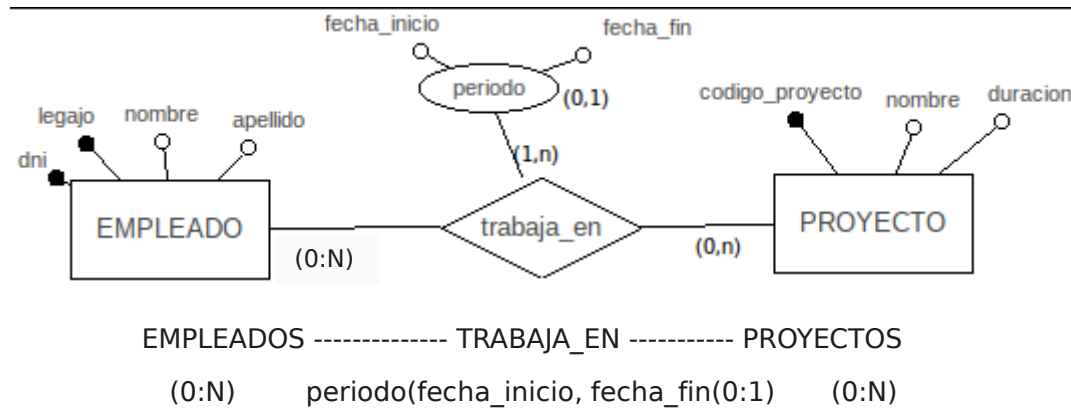
- Es un elemento del DER extendido
- Se utiliza para modelar fechas o períodos de tiempo.
 - Un empleado trabaja en un proyecto durante un período de tiempo. La situación actual no refleja el paso del tiempo (relación 1:N)



- Cuando el empleado deje de trabajar en el proyecto, deja de participar de la relación

Incluyendo la historia

- Se refleja el paso del tiempo (relación N:M)



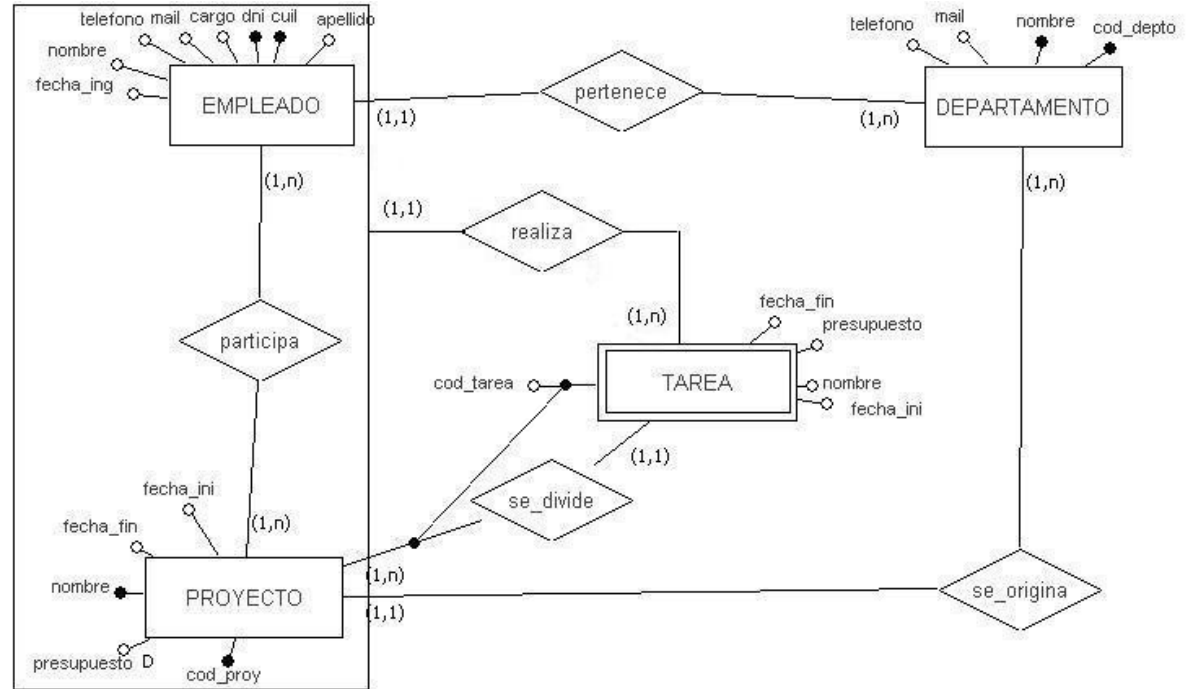
CON HISTORIA

Agregaciones

- **Cuando se genera una relación ternaria, tres entidades confluyen en ella. En ciertos casos esta representación pierde claridad.**
- **La agregación permite visualizar a un par de entidades y una relación como una gran entidad que a su vez se relaciona con otra.**

Ejemplo

- **El empleado participa en muchos proyectos, pero en cada proyecto, el empleado realiza una tarea particular.**
- **Pero una tarea en un proyecto, puede ser realizada por más de un empleado.**



Consideraciones finales

- **Evitar en lo posible interrelaciones de grado > 2 .**
- **Totalmente prohibido utilizar interrelaciones de grado > 3 .**
- **Tener cuidado con los atributos en las interrelaciones.**
- **Todas las entidades deben tener al menos un atributo identificador (Menos las subtipo).**

Ejemplo

- **Se debe modelar el esquema de una facultad. En este esquema los profesores, de los cuales se conocen los datos personales, el cuil y los títulos obtenidos, se relacionan con los departamentos (que se identifican por un nombre único) donde cumplen sus actividades. Un docente solo trabaja en un departamento, pero un departamento puede tener varios docentes. Además cada docente debe indicar en que ciudad ha nacido.**

Ejemplo

- **El esquema debe incluir la representación de los estudiantes, descritos por sus datos personales más la matrícula de alumno, que es única en toda la universidad.**
- **Los estudiantes deben indicar el lugar de nacimiento y de residencia.**

Ejemplo

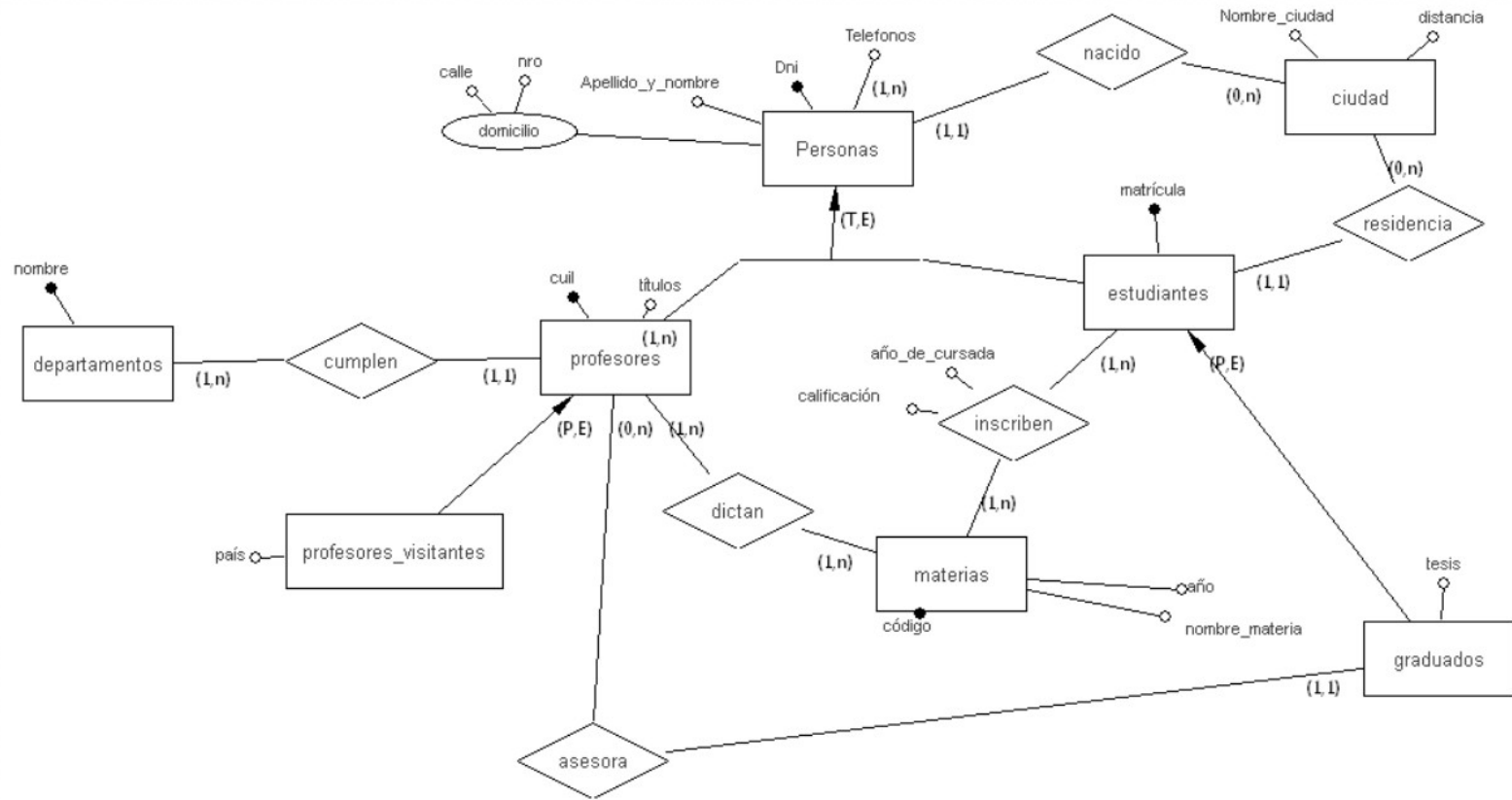
- **Además, se deben señalar las materias que se dictan en la facultad (de cada materia se sabe un código, el nombre y el año dentro del plan de estudio). Los alumnos se inscriben en materias, las cuales son dictadas por profesores bajo los siguientes supuestos:**
 - Un docente puede dictar más de una materia.
 - Una materia puede estar dictada por más de un docente.
 - Los alumnos pueden anotarse en más de una materia y se debe indicar la calificación obtenida.

Ejemplo

- **Continuación:**

- Si un alumno no aprueba la materia puede inscribirse nuevamente, por lo que debe registrarse el año de cursada.
- Los estudiantes pueden ser graduados. En este caso tendrán definidos un trabajo de tesis y un asesor, que es profesor de la casa.
- Pueden existir además profesores visitantes, para los cuales debe indicarse el país de procedencia.

Ejemplo



Otro ejemplo: campeonato de ajedrez

- **Se desea construir una base de datos para registrar campeonatos:**
 - En el campeonato participan jugadores y árbitros. De ambos se requiere conocer el número de asociado, nombre, dirección, teléfono de contacto y campeonatos en los que ha participado (como jugador o como árbitro).
 - De los jugadores se requiere además el nivel de juego en una escala de 1 a 10.
 - Ningún árbitro puede participar como jugador.

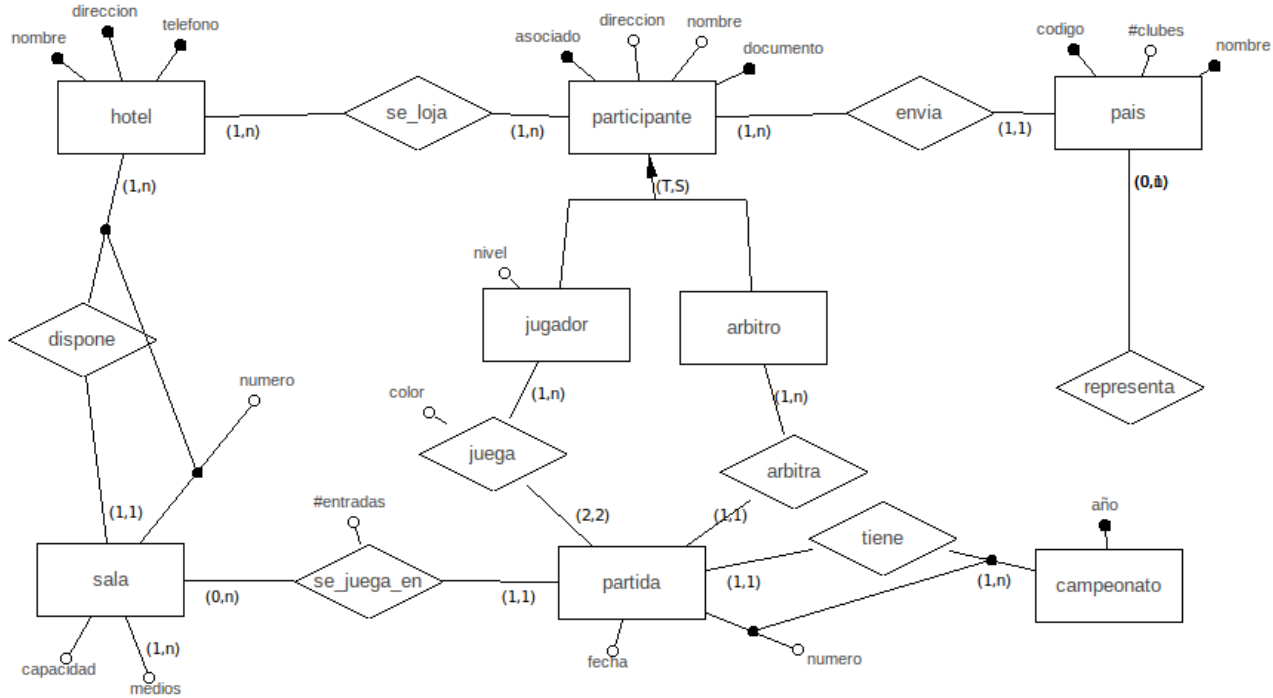
Otro ejemplo: campeonato de ajedrez

- Los países envían al campeonato un conjunto de jugadores y árbitros, aunque no todos los países envían participantes. Todo jugador y árbitro es enviado por un único país. Un país puede ser representado por otro país.
- Cada país se identifica por un código ISO3166-1 e interesa conocer también su nombre y el número de clubes existentes en el mismo.
- Cada partida se identifica por un número correlativo, la juegan dos jugadores y la arbitra un árbitro. Interesa registrar las partidas que juega cada jugador y el color con el que juega. Ha de tenerse en cuenta que un árbitro no puede arbitrar a jugadores enviados por el mismo país que le ha enviado a él.

Otro ejemplo: campeonato de ajedrez

- Todo participante participa de al menos una partida.
- Tanto los jugadores como los árbitros se alojan en un hotel. De cada hotel, se desea conocer el nombre, la dirección y el número de teléfono.
- El campeonato se desarrolla a lo largo de una serie de jornadas y cada partida tiene lugar en una de las jornadas aunque no tengan lugar partidas todas las jornadas.
- Cada partida se celebra en una de las salas de las que disponen los hoteles, se desea conocer el número de entradas vendidas para cada partida. De cada sala se desea conocer la capacidad.

Otro ejemplo: campeonato de ajedrez



Bibliografía

- **Carlo Batini y otros. Diseño conceptual de bases de datos, un enfoque de entidades-interrelaciones. 1992. Adison Wesley.**
- **Il-Yeol Song and others. A Comparative Analysis of Entity-Relationship Diagrams. 1995. Journal of Computer and Software Engineering Volumen 3 Number 4.**
- **Introducción a las Bases de Datos. Fundamentos y Diseños (Bertone, Thomas)**
- **Desarrollo de Bases de Datos: casos prácticos desde el análisis a la implementación (Cuadra,Castro,Iglesias,Martinez, Calle, Moreno)**
- **Diseño de Bases de Datos relacionales (De Miguel, Piattini, Marcos)**