

# FICHEROS Y BASES DE DATOS (E44)

## 3º INGENIERÍA EN INFORMÁTICA

### Tema 7.

#### Modelos de Bases de Datos

- 1.- Introducción.
- 2.- Modelo Jerárquico.
- 3.- Modelo de Red.
- 4.- Modelo Relacional.
- 5.- Otros Modelos.

(Capítulo 3 del Harrington)

# INTRODUCCIÓN

## Modelos de Datos y Esquemas

- Los Modelos de Datos (MD) son herramientas que permite describir la realidad.
- Los programadores los utilizan para construir Esquemas, que son representaciones de la realidad.
- La calidad de un esquema se relaciona con las propiedades del MD y la experiencia de los programadores.

## Representación de los Modelos de Datos

- Los MD tienen asociados una serie de conceptos, que describen un conjunto de datos y operaciones para manipular los datos.
- Dichos conceptos tienen asociados una construcción lingüística y una construcción gráfica.

## Clases de Modelos de Datos

- Existen dos tipos de MD:
  - Modelo Conceptual que representa la realidad en un alto nivel de abstracción.
  - Modelo Lógico, o Modelo de Base de Datos, que describen las relaciones lógicas entre los datos y la base de datos.
- El primero genera el Esquema Conceptual y el segundo el Esquema Lógico.

# MODELOS DE BASES DE DATOS (MBD)

## Características

- Los elementos de un modelo representan Entidades genéricas.
- Los valores concretos se denominan Instancias u Ocurrencias de una entidad.
- Cada SGBD se asocia a un MDB específico, aunque existen excepciones.

## Contenidos

- Por lo que respecta a los datos:
  - Datos o Entidades.
  - Propiedades de los Datos.
  - Relaciones de los Datos.
  - Restricciones de los Datos.

Se representa mediante el Lenguaje de Definición de Esquemas (LDE) del SGBD.

- Por lo que se refiere a las operaciones.
  - Operaciones de los Datos
  - Operaciones sobre Relaciones de los Datos.
  - Relaciones entre Operaciones.

Se representa mediante el Lenguaje de Manipulación de Datos (LMD) del SGBD.

# MODELOS DE BASES DE DATOS (MBD)

## Sistemas Prerrelacionales

- Modelo Jerárquico.  
Los datos se relacionan de modo jerárquico, y se representa mediante una estructura en árbol.
- Modelo de Red.  
Se entiende como una generalización del modelo jerárquico, en donde los nodos hijo pueden tener varios nodos padre.

## Sistemas Relacionales

- Modelo Relacional.  
Se utilizan conceptos matemáticos, como las relaciones, para representar los datos y las operaciones sobre estos.

## Sistemas Postrelacionales

- Modelo Orientado a Objetos.  
Los datos se representan mediante objetos, que contienen variables y métodos, y su manipulación se realiza mediante mensajes.
- Modelo Semántico.  
Tienen como objetivo describir de un modo más preciso la información contenida en la base de datos.
- Modelo Deductivo.  
Son capaces de deducir hechos a partir de las relaciones base y una serie de axiomas deductivos o reglas de inferencia.

# MODELO JERÁRQUICO

## Estructura Jerárquica

- Una Base de Datos Jerárquica se compone de un Conjunto Ordenado de Árboles.
- También se puede definir como el Conjunto Ordenado de instancias de un Tipo de Árbol.
- Un tipo de árbol se compone de un Tipo de Registro Raíz junto con un conjunto ordenado de cero, uno o más Subárboles Dependientes.
- Un subárbol es el árbol de un nivel inferior.
- El orden global de la base de datos se obtiene mediante el recorrido en Preorden del bosque.

## Definiciones

- Los tipos de registros representan las Entidades de la base de datos.
- Los arcos del árbol definen la relación existente entre las entidades.
- Los tipos de registros se descomponen en Campos.
- Aparece una dependencia entre los niveles del árbol.
- Dicha dependencia puede llevar asociada la Herencia de campos entre los niveles.
- Esa dependencia también tiene influencia en la integridad de los datos, ya que no puede existir un hijo sin su padre.

# MODELO JERÁRQUICO

## Diagramas de Ocurrencias (DO)

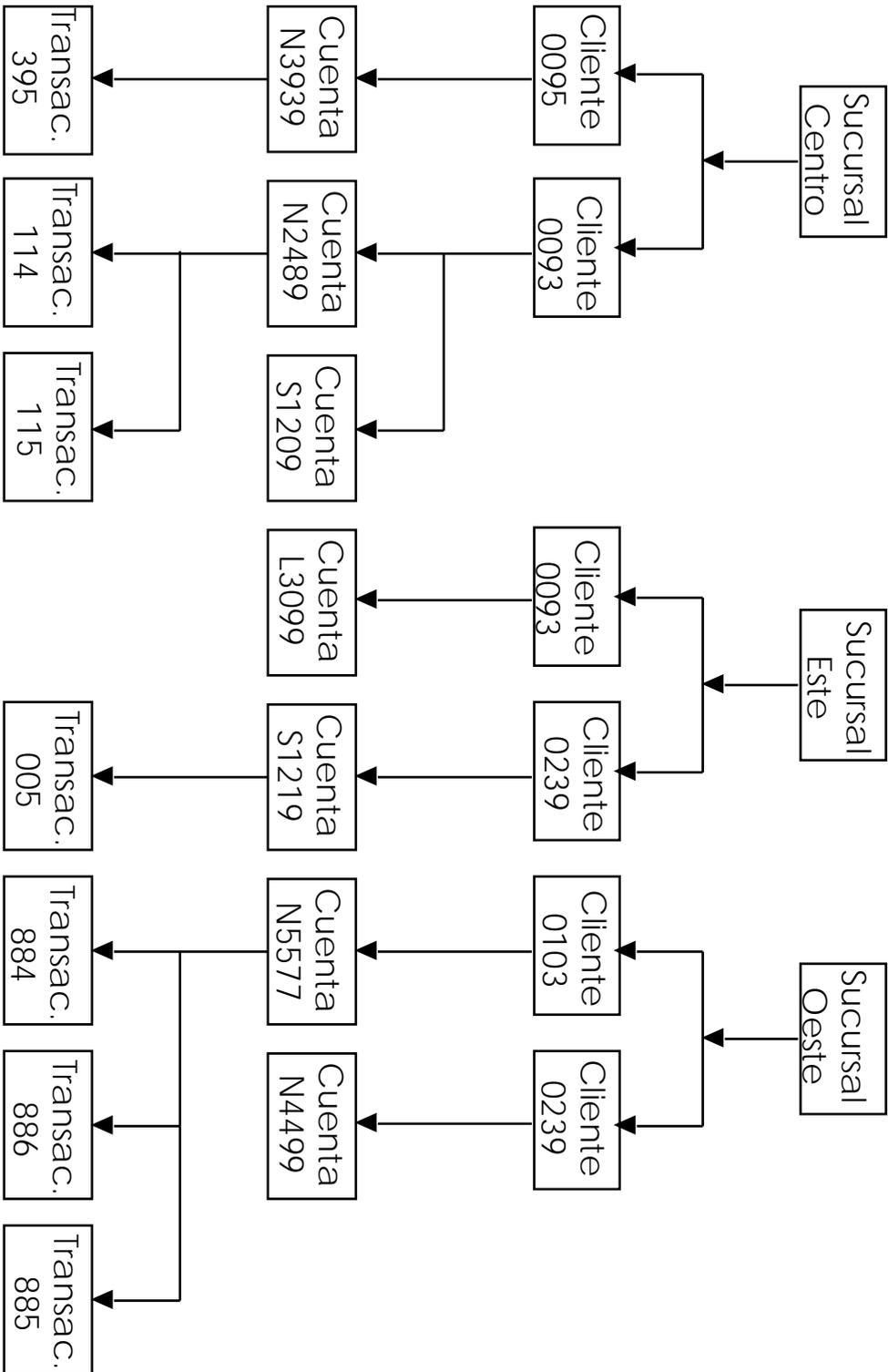
- Cada uno de los rectángulos representa la instancia de un tipo de registro.
- Los arcos definen la relación existente entre las instancias de dos niveles contiguos del árbol.
- No se suelen especificar los campos de los registros.
- Cuando la información almacenada en la base de datos es muy grande, se convierte en inmanejable.
- Si no aparece ninguna instancia de alguno de los niveles del árbol, la relación asociada no se puede representar.

## Diagrama de Bachman (DA)

- Describe gráficamente el esquema lógico de una base de datos jerárquica.
- Los rectángulos representan los tipos de registro.
- Los arcos representan la relación existente entre los tipos de registro.
- Los arcos se etiquetan porque puede existir más de una relación entre dos tipos de registro (no en el modelo jerárquico).
- La flecha simple-doble indica que existe un único padre para cada hijo, pero pueden existir cero o más hijos de un padre.

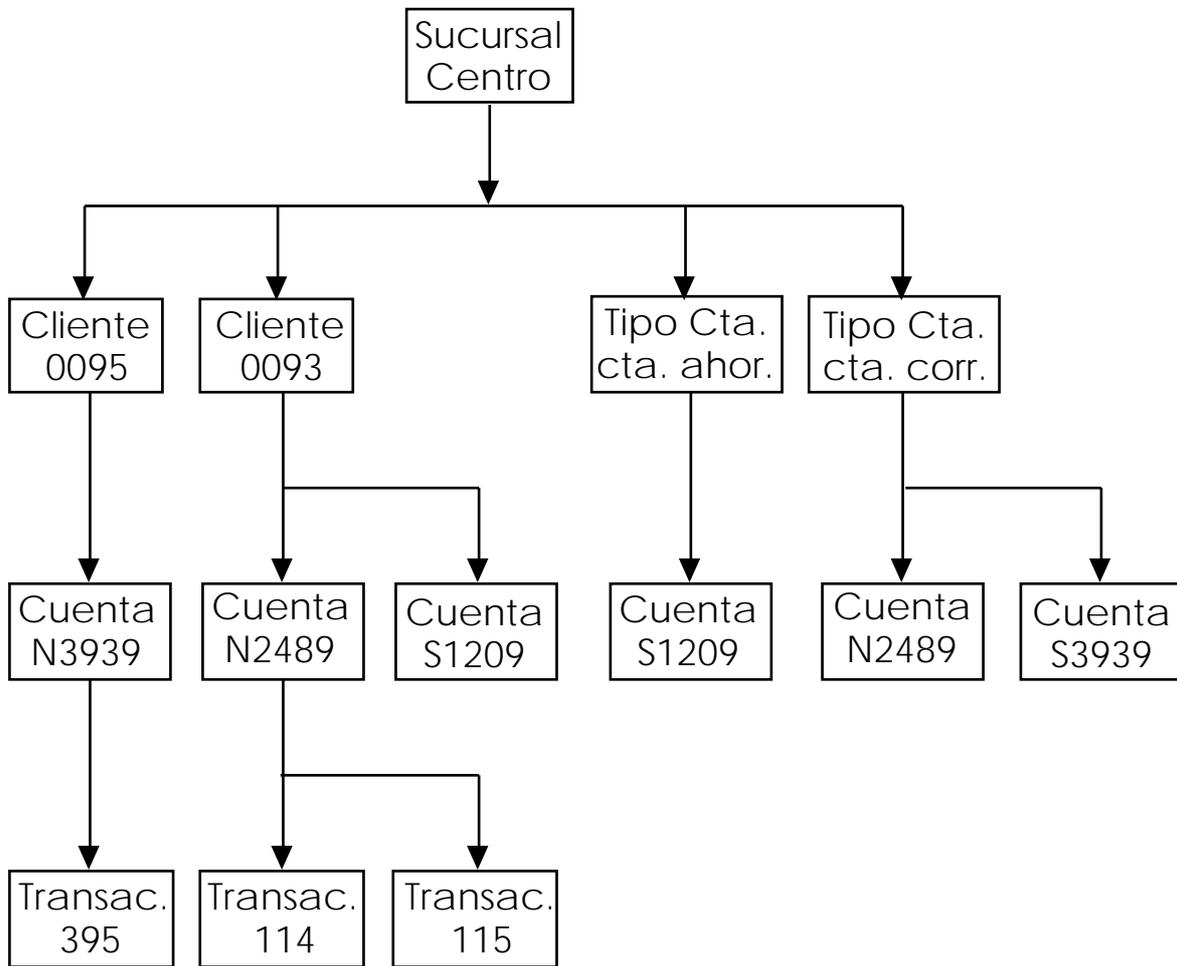
# MODELO JERÁRQUICO

## Ejemplo Básico con DO



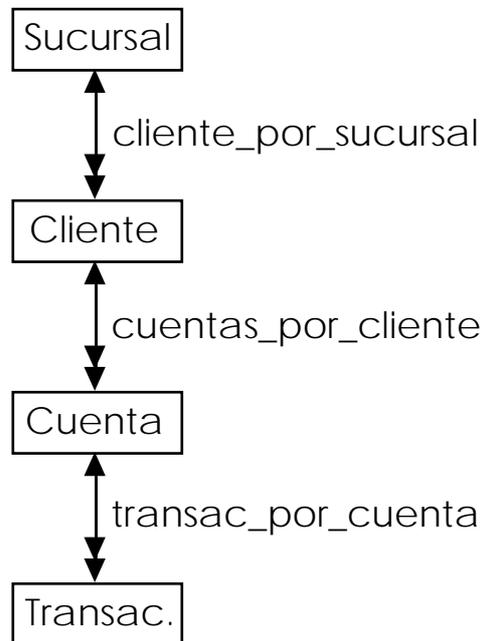
# MODELO JERÁRQUICO

## Ejemplo de Tipos de Cuentas con DO

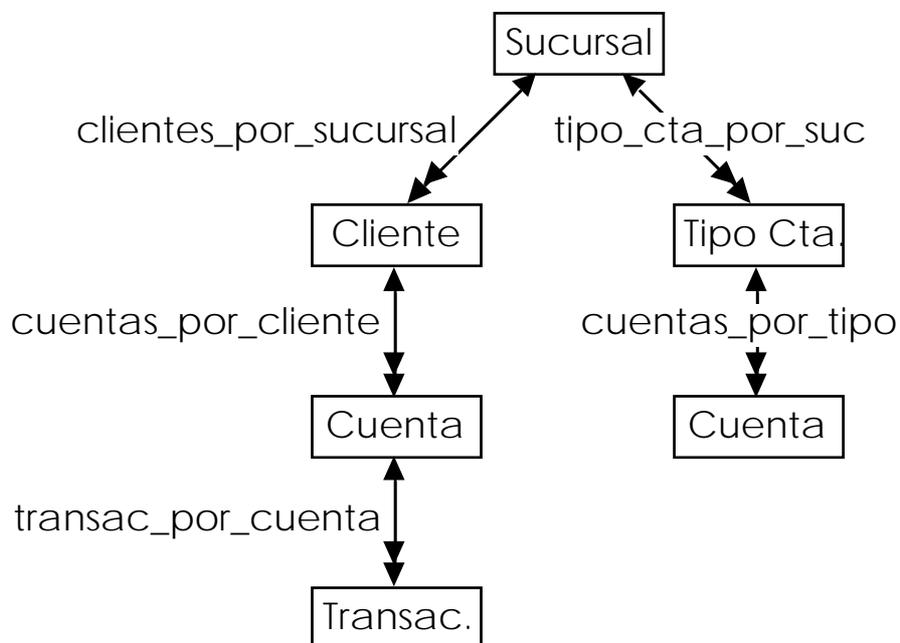


# MODELO JERÁRQUICO

## Ejemplo Básico con DA



## Ejemplo de Tipos de Cuentas con DA



# MODELO JERÁRQUICO

## Manipulación de los Datos

- Los lenguajes asociados contiene operadores que manejan datos almacenados en árboles ordenados.
- Deben contener los siguientes operadores:
  - Búsqueda de un árbol.
  - Movimiento en los árboles de un nivel.
  - Movimiento a un subárbol.
  - Movimiento dentro de la estructura jerárquica.
  - Inserción y borrado de registros.

## Problemas de las BD Jerárquicas

- Únicamente pueden representar relaciones de 1-a-Muchos.
- Las relaciones Muchos-a-Muchos requieren la redundancia de información.
- Si un registro tipo aparece como "hijo" en más de dos relaciones, se debe de replicar.
- Esto puede producir problemas de Integridad y Consistencia de los Datos.
- Los lenguajes de manipulación asociados son fuertemente navegacionales.
- Hace falta un Lenguaje de Programación Anfitrión en el que se inserten los operadores.

# MODELO JERÁRQUICO

## Esquema Físico (IMS de IBM, 1962)

DBD NAME = BANCO

SEGM NAME = SUCURSAL, PARENT = 0, BYTES = ...

FIELD NAME = Nombre, BYTES = ... , START = ...

FIELD NAME = Director, ...

FIELD NAME = Calle, ...

SEGM NAME = CLIENTE, PARENT = SUCURSAL, ...

FIELD NAME = Nombre, BYTES = ... , START = ...

FIELD NAME = DNI, ...

...

SEGM NAME = CUENTA, PARENT = CLIENTE, ...

...

SEGM NAME = TRANSACCION, PARENT = CUENTA, ...

...

## Lenguaje Operacional

"Datos de las cuentas del cliente de la sucursal  
centro con D.N.I. 01234567"

GET UNIQUE Sucursal WHERE Nombre = "Centro"  
Cliente WHERE DNI = "01234567"

GET NEXT WITHIN PARENT Cuenta

mientras no\_error

escribir (datos\_cuenta)

GET NEXT WITHIN PARENT Cuenta

fin mientras

# MODELO DE RED

## Estructura de Red

- El Modelo de Red se puede entender como una extensión del modelo jerárquico.
- También se presenta mediante un árbol, pero en este caso, cada hijo puede tener varios padres.
- De este modo se reducen, o eliminan, las redundancias.
- Pero desaparece la herencia de los campos.
- La integridad de datos, asociada a los arcos padre-hijo, se mantiene.
- Una Base de Datos de Red se compone de dos conjuntos:
  - El Conjunto de los Registros.  
Un conjunto de instancias múltiples de varios tipos de registros.
  - El Conjunto de las Relaciones.  
Un conjunto de instancias múltiples de varios tipos de relaciones.

## Representación de Red

- En el diagrama de ocurrencias, aparecen arcos que conectan los árboles.
- Los diagramas de Bachman se extienden, apareciendo flechas doble-doble que representan las relaciones muchos-a-muchos de las redes complejas.

# MODELO DE RED

## Tipos de Redes

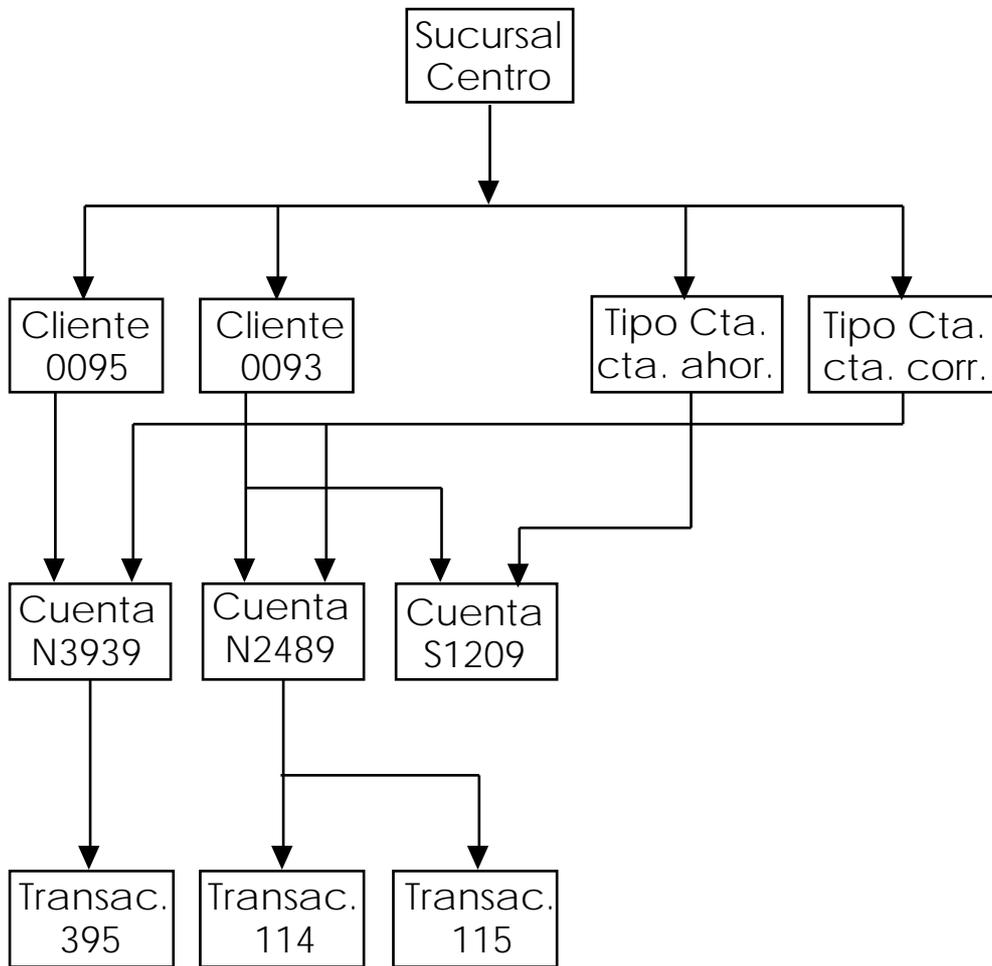
- Se dice que la Red es Simple si los padres de un hijo son instancias de registros tipo diferentes.
- Se dice que la Red es Compleja si los padres pueden ser instancias del mismo registros tipo.
- En las redes complejas puede desaparecer todo tipo de redundancia, pero perdiendo la herencia.
- En algunos casos, resulta interesante permitir cierto grado de redundancia, para evitar pérdida de información.
- Otra alternativa es convertir una red compleja en una red simple en donde no se pierda la información.

## Conversión Compleja-Simple

- Permite reducir el problema de la pérdida de información asociado a las redes complejas.
- La idea es convertir una relación muchos-a-muchos en dos relaciones uno-a-muchos, mediante la inserción de un nuevo tipo de registro.
- Este registro se denomina Registro Intersección si contiene algún tipo de información, que se denomina Datos de la Intersección.
- En otro caso, se denomina Registro de Enlace.

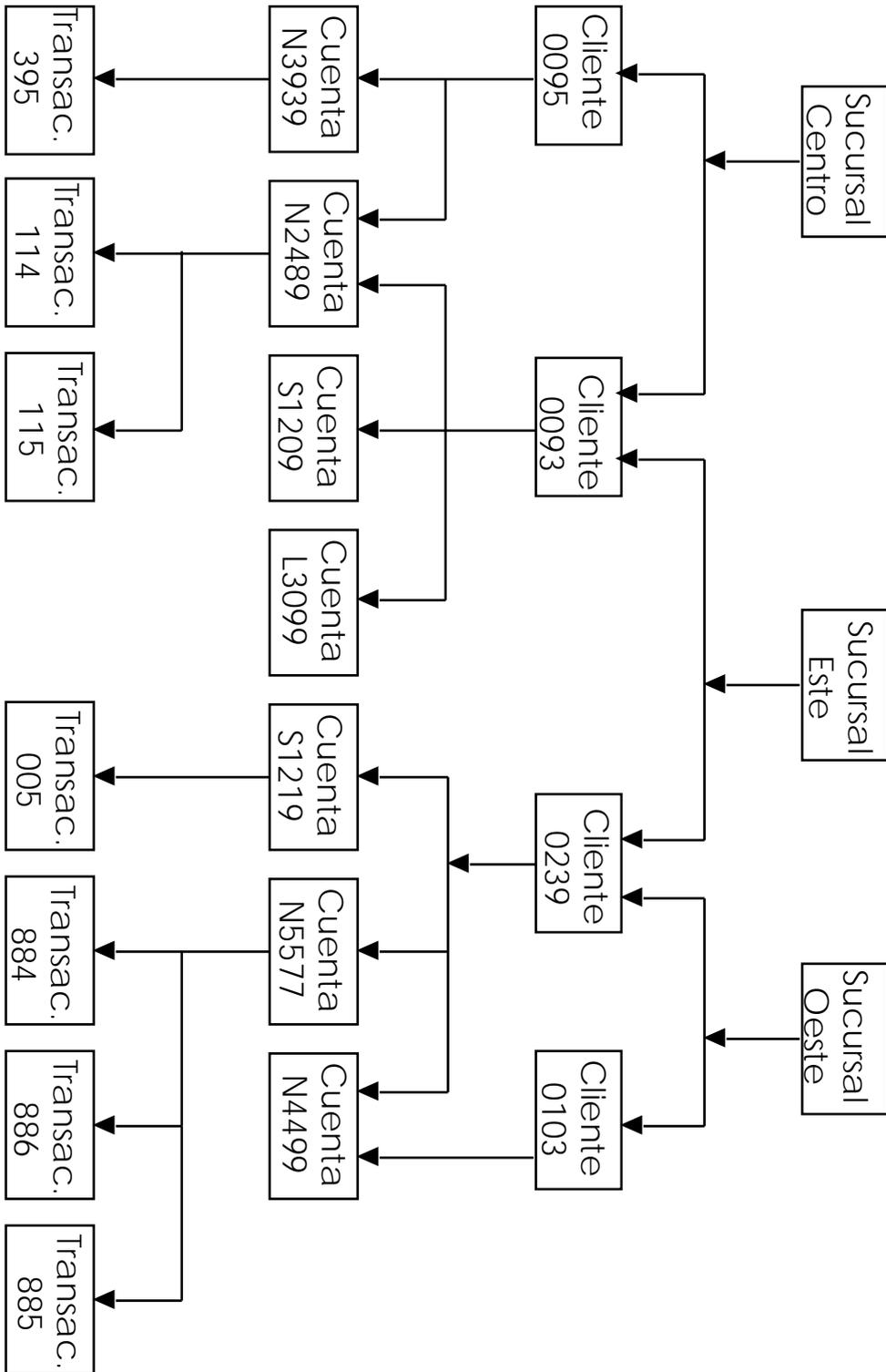
# MODELO DE RED

## Ejemplo de Red Simple con DO



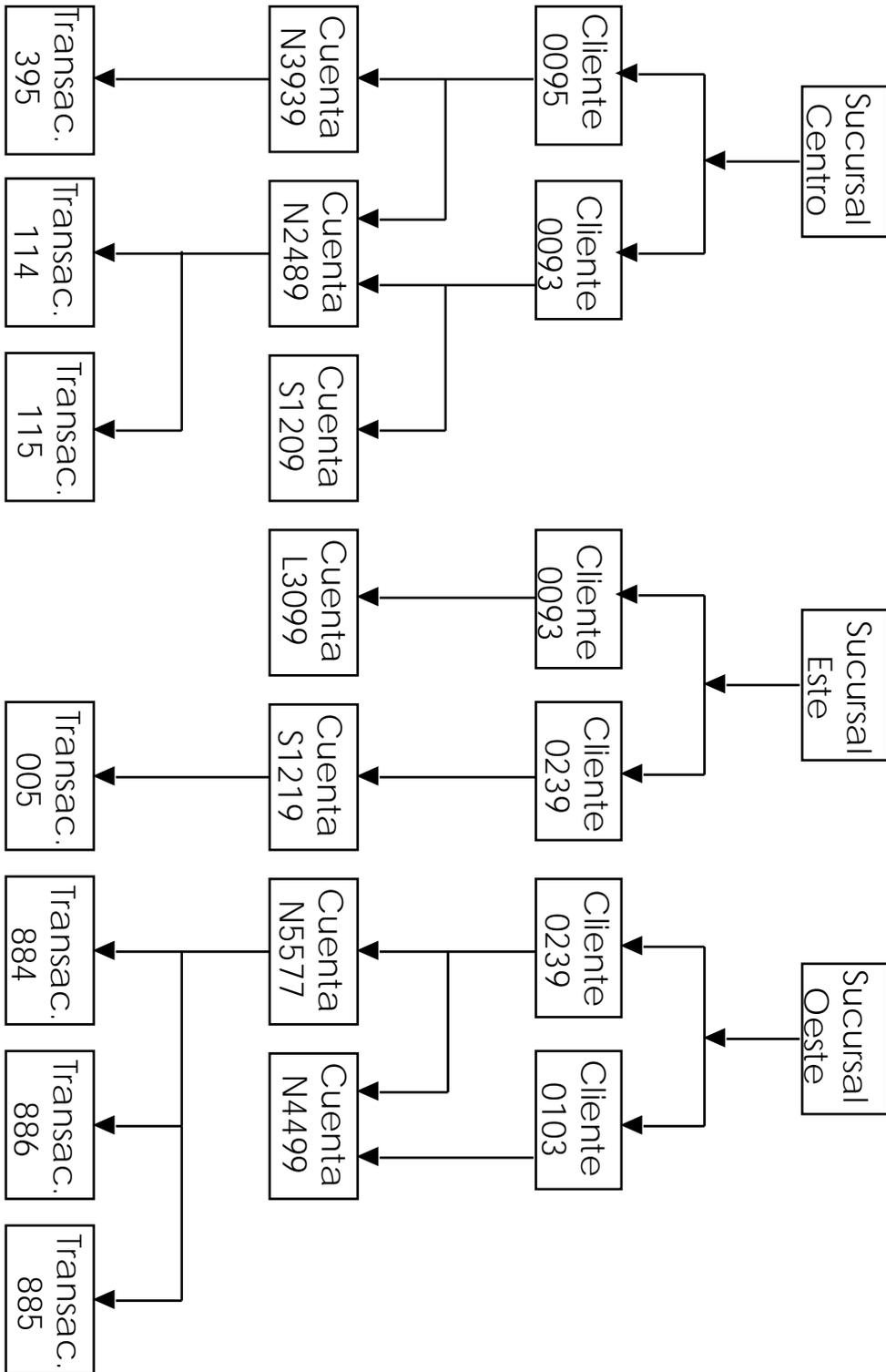
# MODELO DE RED

## Ejemplo de Red Compleja con DO



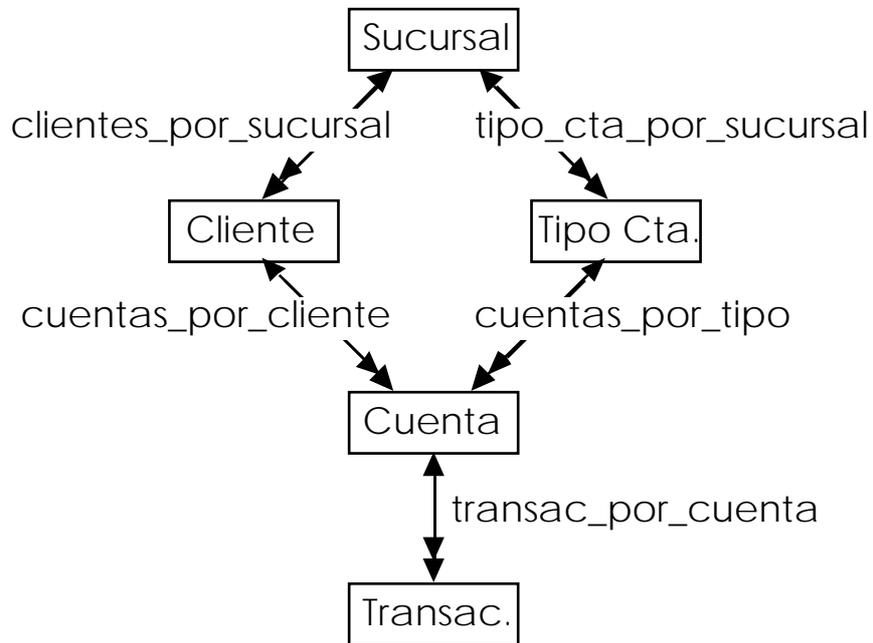
# MODELO DE RED

## Ejemplo de Red Compleja\* con DO

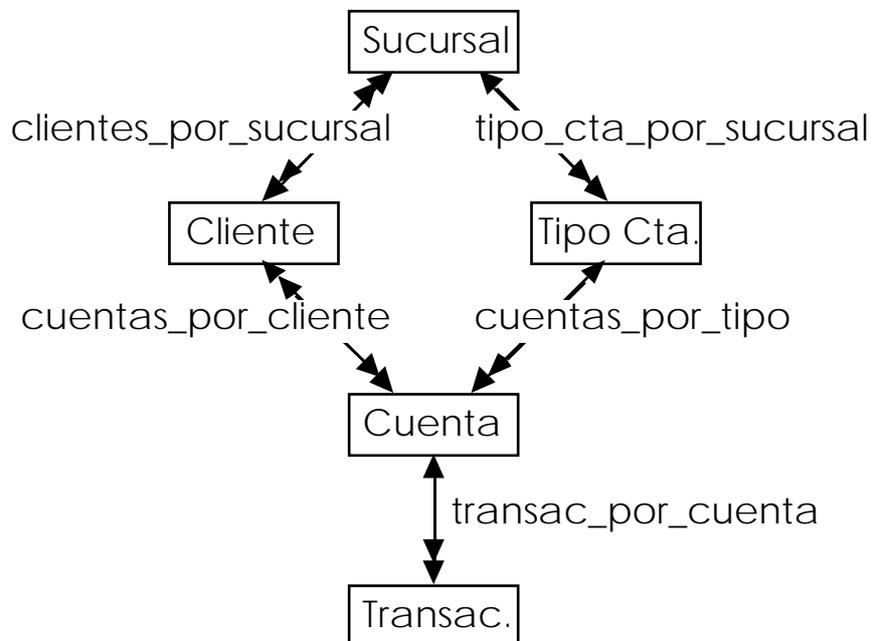


# MODELO DE RED

## Ejemplo de Red Simple con DA

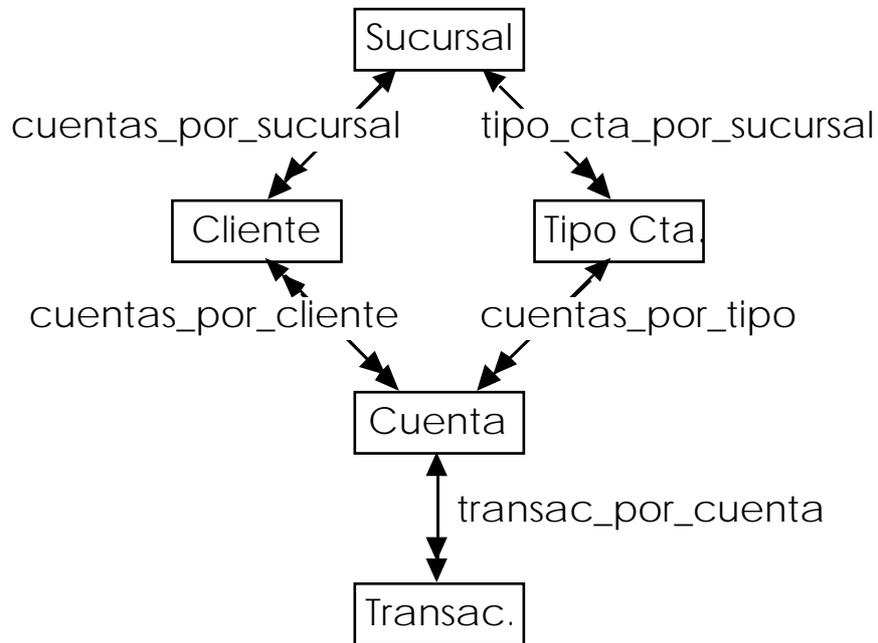


## Ejemplo de Red Compleja con DA

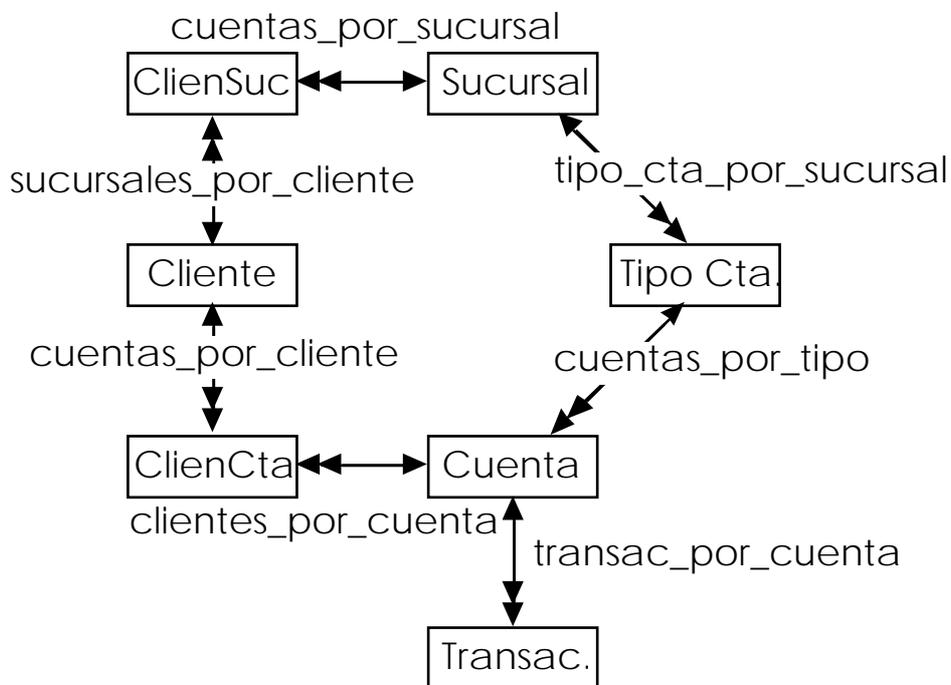


# MODELO DE RED

## Ejemplo de Red Compleja\* con DA



## Ejemplo Conversión Compleja-Simple con DA

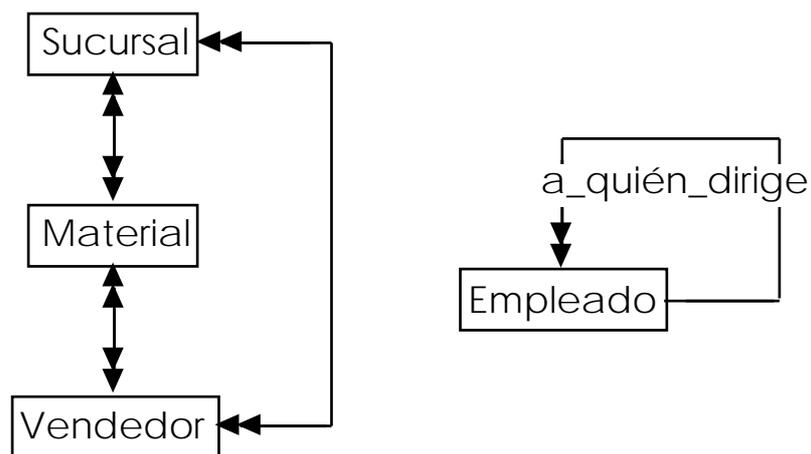


# MODELO DE RED

## Ciclos y Lazos

- Existen dos tipos de relaciones específicas, los Ciclos y los Lazos.
- En un ciclo, diferentes tipos de registro se relacionan de modo circular.
- Los lazos representan la relación de un tipo de registro consigo mismo.
- Pueden aparecer ciclos en redes complejas y en redes simples.
- En redes complejas, se puede aplicar a los ciclos la conversión compleja-simple.
- Los lazos sólo puede manejarse en redes complejas.

## Ejemplo de Ciclos y Lazos con DA



# MODELO DE RED

## Manipulación de los Datos

- Los lenguajes asociados contiene operadores que manejan datos almacenados en conjuntos de registros y relaciones.
- Deben contener los siguientes operadores:
  - Búsqueda de un registro específico.
  - Movimiento del padre al primer hijo de una relación.
  - Movimiento de un hijo al siguiente en una relación.
  - Movimiento del hijo al padre de una relación.
  - Creación, borrado y modificación de un registro.
  - Inserción y eliminación de un hijo en una relación.
  - Cambio de relación de un hijo.

## Problemas Asociados

- Los lenguajes de manipulación asociados son fuertemente navegacionales.
- Hace falta un lenguaje de programación Anfitrión en el que se inserten los operadores.
- Especialmente complicada en redes complejas.

# MODELO DE RED

## Ejemplo de Esquema Físico (CODASYL, 1971)

SCHEMA NAME = BANCO

RECORD NAME IS Sucursal

KEY código\_suc IS nombre

Nombre TYPE IS ...

Director TYPE IS ...

RECORD NAME IS Cliente

KEY código\_cli IS DNI

Nombre TYPE IS ...

DNI TYPE IS ...

...

SET NAME IS Clie\_Suc

OWNER IS Sucursal

MEMBER IS Cliente

## Ejemplo de Lenguaje Operacional

"Datos de los clientes de la sucursal Centro"

sucursal.nombre := "Centro"

FIND ANY sucursal USING nombre

FIND FIRST cliente WITHIN Clie\_Suc

mientras no\_error

    GET cliente; escribir (cliente)

    FIND NEXT cliente WITHIN Clie\_Suc

fin mientras

# MODELO RELACIONAL

## Estructura Relacional

- Definido por E.F. Codd en 1970, se fundamenta en conceptos matemáticos.
- La estructura de datos básicos es la Relación, denominada normalmente como Tabla.
- Una base de datos relacional se compone de una colección de relaciones.
- Cuando una tabla contiene datos se dice que es una instancia de la relación.
- Cada relación se asocia a una entidad, y se compone de una serie de atributos.
- Las filas de la tabla definen las instancias de la entidad, y son las Tuplas de la Relación.
- No se permiten tuplas duplicadas en una tabla, aunque los SGDB no suelen controlarlo.
- Las columnas de la tabla son las ocurrencias de los atributos de la entidad.
- Con cada atributo se asocia un Dominio que define el posible rango de valores.
- Dicho dominio define una Restricción de los atributos.
- La manipulación de los datos se realizan mediante Lenguajes de Especificación.
- De tal modo que el usuario indica que datos desea, sin especificar como obtenerlos.

# MODELO RELACIONAL

## Ejemplo de Relaciones

CLIENTE

Num_Clie	Nombre	Dirección	Teléfono	DNI
0093	Julio Gil	Oltra, 23-1	123456	00112233
0095	Ana García	Cuenca, 16-8	456789	11223344
0103	Pedro Ruíz	Barraca, 101-8	789123	22334455
0239	Luisa Medina	Reina, 3-3	456123	33445566

CTA\_CLIE

Num_Clie	Num_Cta
0093	N2489
0093	S1209
0093	L3099
0095	N3939
0103	N557
0239	N4499
0239	S1219
0095	N2489

TRANSACCION

Num_Cta	Num_Trans	Cantidad
S1209	156	+100000
L3099	023	+250850
N5577	309	-10000
N5577	310	-80000
1219	003	+189
.	.	.
.	.	.
.	.	.

# MODELO RELACIONAL

## Ejemplo de Esquema Relacional

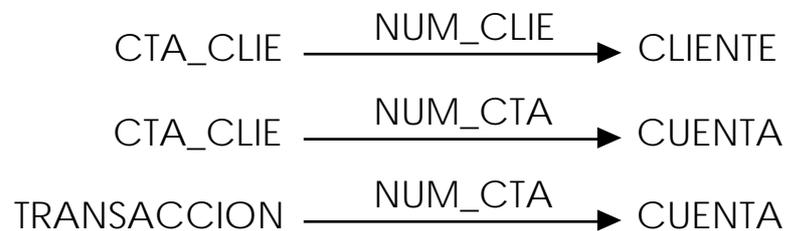
CLIENTE (NUM\_CLIE, NOMBRE, DIREC, TELEF, DNI)

CUENTA (NUM\_CTA, SALDO, TIPO)

CTA\_CLIE (NUM\_CLIE, NUM\_CTA)

TRANSACCION (NUM\_CTA, NUM\_TRANS, CANT)

...



## Lenguaje de Especificación

"Obtener el saldo y el tipo de cuentas del cliente con número 0095"

```
SELECT CUENT.NUM_CT, CUENTA.SALDO, CUENTA.TIPO
FROM CUENTA, CTA_CLIE
WHERE CTA_CLIE.NUM_CLIE = "0095"
AND CT_CLIE.NUM_CTA = CUENTA.NUM_CTA;
```

# OTROS MODELOS DE BASES DE DATOS

## Modelos Clásicos

- Los sistemas de Bases de Datos salieron como una alternativa a los sistemas de ficheros, con el objetivo de manejar grandes cantidades de información.
- Los modelos clásicos presentan una serie de características comunes.
- Uniformidad. Una gran cantidad de datos se estructuran de modo similar.
- Orientación en Registros. Los datos básicos se almacenan en registros de longitud fija.
- Datos Pequeños. Los registros son cortos, normalmente de pocos cientos de bytes.
- Campos Atómicos. Los campos de los registros son cortos, de longitud fija y no poseen ningún tipo de estructura interna.
- Transacciones Cortas. En las transacciones no existe interacción con el usuario y además su duración es de alguna fracción de segundo.
- Esquema Casi Estático. No suelen realizarse cambios en los esquemas, y si aparecen son de poca importancia.

# OTROS MODELOS DE BASES DE DATOS

## Nuevas Aplicaciones

- Cuando se definieron los modelos clásicos no se consideraron ciertas aplicaciones.
- Existen una serie de factores que han abierto el campo de aplicaciones:
  - Abaratamiento del hardware.
  - Aumento de la capacidad de proceso.
  - Mejora del entendimiento de la gestión de la base de datos posibilitan.
- Estas aplicaciones son:
  - Diseño Asistido por Computador (CAD). Almacena datos pertenecientes a un diseño de ingeniería.
  - Ingen. de Software Asistido por Computador (CASE). Contiene los datos requeridos para ayudar a los que desarrollan software.
  - Bases de Datos Multimedia. Contiene datos de tipo diverso, como dato de audio y video.
  - Sistema de Información de Oficina (OIS). Incluye herramientas para creación y recuperación de documentos, también mantiene calendarios con citas, etc.
  - Sistemas Expertos de Bases de Datos. Junto con los datos, almacena reglas explícitas que representan restricciones de integridad, disparadores y otros conocimientos sobre la empresa a modelar.

# OTROS MODELOS DE BASES DE DATOS

## Requisitos

- Todas estas aplicaciones requieren nuevos modelos de datos, lenguajes de consulta y modelos de transacciones.
- Los requisitos básicos son:
  - Objetos Complejos. Son objetos que se componen de otros objetos más simples, aunque se manejan como otro objeto. Se relacionan con las Bases de Datos Orientadas a Objetos y a las Bases de datos Relacionales Anidadas.
  - Datos de Comportamiento. Una misma operación puede actuar de modo diferente sobre distintos objetos, según como se define el Método asociado.
  - Metaconocimiento. Existen reglas concretas asociadas a la aplicación, que no pueden representarse en los modelos clásicos. Se asocian con las Bases de Datos Lógicas.
  - Transacciones de Larga Duración. Las aplicaciones CAD y CASE son interactivas, lo que produce problemas por el acceso concurrente de usuarios. Se deben introducir las Transacciones Anidadas y las Ejecuciones no Serializables Correctas.