

TEMARIO

- Tema 1.– Introducción a las Bases de Datos.
- Tema 2.– Modelos de datos.
- Tema 3.– Modelo Entidad / Relación. (E/R ; E/R Extendido).
- Tema 4.– Modelo relacional.
- Tema 5.– Teoría de Normalización.

BIBLIOGRAFIA

`Conceptos y diseño de bases de datos.'

Adoración de Miguel / Mario Riattini.

Editorial Ra–ma.

`Fundamentos de bases de daos'

Korth /Siberschats.

Editorial McGraw–Hill

`Bases de datos'

G.Gardarin.

Editoril Paraninfo.

`Principles of Database System'

Ullman

Editorial. Computer science Press.

`Fundamentals of database systems'

Elmasri / Navate

editorial. Benjamin.

TEMA 1.

Introducción a las Bases de Datos.

- Justificación de utilizar Bases de datos.
- Analisis de sistemas tradicionales (Problemas).
- Justificación.
- Definición de Bases de datos.

- Definiciones (Otros autores).
- Definición.
- Ventajas de las Bases de Datos.
- Oposición a los sistemas tradicionales.
- Desventajas.
- Sistema Gestor de Bases de Datos (SGBD).
- Definición.
- Justificación.
- Evaluación.

Justificación de Bases de datos

- Analisis de los sistemas tradicionales.
- Inconvenientes de los sistemas tradicionales :

EJ :

- Sistemas orientados al tratamiento en los que se fija el proceso y luego se gestionan los datos apocados a existir en ficheros.
- Se desarrollan aplicaciones independientes entendidas como programas y datos (repetición de datos).
- La primera situación problemática seria se plantea con el coste del almacenamiento.
- Podría existir un problema de actualización de datos, al existir datos duplicados en los ficheros.
- Peticiones sorpresivas, que se han de resolver en poco tiempo.

Desventajas de sistemas tradicionales. –

- Redundancia (copia innecesaria). Implica desperdicio de almacenamiento.
- Dificultad de mantenimiento (Actualización).
- Consistencia de datos (Actualización).
- excesiva dependencia del soporte y los datos. – Un cambio sutil en los datos acarreará el cambio total del programa.
- Peticiones inesperadas. – Tendencia a utilizar sistemas orientados a la toma de decisión.
- Aumento del tiempo de CPU

Estos problemas anteriores justifican la utilización de un sistema diferente.

El sistema tradicional se define como un esquema horizontal y en cada estrato se encuentra cada aplicación con todos los ficheros que necesita, aunque estos estarán duplicados.

Nuevo enfoque.

El error del enfoque antiguo consistía en un enfoque al programa. El enfoque nuevo esta orientado a los datos.

- Estos serán un conjunto estructurado independiente de aplicaciones.
- Objetivo de satisfacer necesidades de información de la aplicación.

Definición de las Bases de Datos.

Historia.–

- 1963.– Se acuña el termino de Bases de Datos en el Simposio de Santa Monica ('DATA BASE')
- 1967.– Codasyl, cambia su nombre por el de 'Data Base Taskgroup'.
- Definiciones.–
 - Martin (1975).
 - Conferencia de estadística (1977).
 - Flory (1982)
 - Howe (1983)
 - Deen (1985)

Términos.–

- La base de datos debería ser una colección o conjunto de datos.
- Tan importante es almacenar el contenido como tambien la relación entre ellos (Guardan documentación de los datos / Características especiales).
- No redundantes.
- Se distinguen dos tipos de redundancia :

LOGICA Una unica actualización de datos (unica copia).

FISICA Se permite de forma transparente pero por eficiencia.

- Puedan servir a cualquier aplicación.
- Procedimientos seguros para modificación :
- Consulta Genera errores pero no inconsistencia.
- Altas / Bajas (SGBD).
- Usuarios multiples.
- Los datos deben llegar a tiempo.
- Se deben almacenar restricciones (seguridad).
- Independencia.–

Datos / Procedimiento.

Datos / Soporte fisico.

- Soporte Acceso directo.
- **Definición de Bases de Datos.–** Un conjunto de información almacenada en memoria auxiliar que permite acceso directo y un conjunto de programas que manipulan esos datos

Base de Datos es un conjunto exhaustivo no redundante de datos estructurados organizados independientemente de su utilización y su implementación en máquina accesibles en tiempo real y compatibles con usuarios concurrentes con necesidad de información diferente y no predicable en tiempo.

Arquitectura del SGDB

- Ampliación de los niveles tradicionales.–

- Físico.
- Lógico.
- Objetivo de la base de datos de independizar.–

- Datos de programas.
- Datos (visión) / Soporte físico.

- Nivel externo Visión de los usuarios particulares.

Tantos niveles externos como usuarios distintos.

Tomando parte del nivel externo (Subconjuntos del nivel conceptual).

- Nivel conceptual.– Contiene no solamente datos sino también las relaciones y las restricciones

Intermedio entre nivel físico y el nivel externo.

- Nivel interno.– Como se gestionan los datos a nivel físico.
- Depende del sistema (SGBD).
- Contiene.–
- La estrategia de almacenamiento.
- Espacio de almacenamiento.
- Desbordamiento.
- Vías de acceso. (Claves / Índices).

Los padres de esta arquitectura de niveles es el grupo ANSI/X3/SPARC

- En las bases de datos la información esta integrada y compartida.
- La información la comparten :
- Muchos usuarios.
- Muchas aplicaciones (Diferentes intereses y objetivos).
- En cualquier momento.
- Muchas en cada momento.
- Todos comparten el mismo nivel conceptual y cada uno tiene su propio nivel externo. (Aunque pueden compartir sus objetivos e intereses).
- Se garantiza la compartición de datos.
- La información de la base de datos debe estar integrada.– La integración consiste en unificar datos y eliminar la redundancia añadiendo las propiedades de BD.
- **Definición de SGBD.–**

Clasificación de usuarios.–

- Usuarios de la información (Solo leer).
- Personas que gestionan la BD.– Se encargan de la actualización, gestión de permisos, volcado histórico, backup, ...

- Programadores.

Definición.- Conjunto coordinado de programas, procedimientos, lenguajes, etc. que suministran, tanto a los usuarios como a los analistas, programadores o administrador los medios necesarios para describir, recuperar y manipular los datos almacenados en la base, manteniendo su integridad, confidencialidad y seguridad.

* **Ventajas de las bases de datos.-**

- Independencia de datos y tratamiento.
 - Cambio en datos no implica cambio en programas y viceversa (Menor coste de mantenimiento).
- Coherencia de resultados.
 - Reduce redundancia :
 - Acciones logicamente unicas.
 - Se evita inconsistencia.
- Mejora en la disponibilidad de datos
 - No hay dueño de datos (No igual a ser publicos).
 - Ni aplicaciones ni usuarios.
 - Guardamos descripción (Idea de catalogos).
- Cumplimiento de ciertas normas.
 - Restricciones de seguridad.
 - Accesos (Usuarios a datos).
 - Operaciones (Operaciones sobre datos).
- Otras ventajas :
 - Más eficiente gestión de almacenamiento.
 - Efecto sinergico.
- Desventajas.-
 - Situación Sistema trad. Sist. De BD.
 - Fuerte coste inicial :
 - Programa
 - Personal
 - equipos.
 - Rentable a medio o largo plazo.
 - No hay standard
 - No solo se puede cambiar datos sino tambien el enfoque del sistema

TEMA 2.

Modelos de Datos.

- 2.1.– Evolución del enfoque de datos.
- 2.2.– Introducción.
- Definiciones.
- Definición formal del Modelo de datos.
- Estática.
- Dinámica.
- Estimaciones de modelos.
- Modelos físicos y lógicos.
- Modelos conceptuales y convencionales.
- Comparación de los modelos relacionales.

Evolución del enfoque de datos.

- Sistemas orientados al proceso.
- Sistemas informáticos orientados a los datos. (Bases de datos / SGBD)
- Realización del 'Modelo de datos'

Interpretar realidad

Comprender significado.

Introducción.

Definiciones (Según Flory / 1982).–

'**Modelar**' Consiste en crear un mundo teórico y abstracto tal que las conclusiones que se puedan sacar de él coinciden con las manifestaciones aparentes de la realidad.

'**Modelo**' Conjunto de conceptos que permiten construir una representación.

'**Modelo de Datos**' Dispositivo de abstracción que nos permite ver la información de los datos más que su valor concreto.

Modelo Vs. Esquema

Definición de Modelo (Universo del Discurso).– Instrumento que se aplica a una parcela del mundo real (Universo del discurso), para obtener una estructura de datos que llamamos ESQUEMA.

Universo del Discurso Vs. Mundo real.

- El universo del discurso es la visión del mundo real del diseñador de la Base de datos.
- Por sus objetivos.
- Primer paso en el diseño es definir el Universo del Discurso (Objetivos del diseñador).

Ej.– Hospital.

Definición de modelo de datos.– Conjunto de conceptos, reglas y convenciones que nos permiten describir los datos del universo del discurso, constituyendo una herramienta que facilita la interpretación de nuestro universo del discurso y su representación en forma de datos en nuestro sistema de información.

- Modelo de datos y lenguaje de datos.–
- Modelos es en lo que se basan los lenguajes de datos.
- Ejs :

SQL.– Modelo relacional + Sintaxis (SQL).

DL/I.– Modelo jerárquico + Sintaxis (DL/I).

Objetivos de los modelos de datos.

Formalización .– Permite definir formalmente las estructuras permitidas y sus restricciones a fin de representar los datos, y también porque establece las bases para un lenguaje de datos.

Diseño.– El modelo de datos es uno de los elementos básicos (Herramienta obligada) en el diseño de Bases de datos.

Propiedades del Universo del discurso :

- Estáticas Invariantes en el tiempo (Estructuras).
- Dinámicas Tienen que ver con el contenido de las estructuras.

Modelo de datos = $\langle S, O \rangle$

S Cto de reglas de generación que permiten representar la componente estática, es decir, las estructuras del UD. Se corresponden con el Lenguaje de Definición de Datos.

O conjunto de operaciones autorizadas sobre la estructura que permite representar la componente dinámica. Se corresponde con el lenguaje de la manipulación de datos.

- ESTATICA

Tipos de objetos permitidos.–

- Entidades Objetos.
- Relaciones Asociaciones entre objetos.
- Atributos.
- Dominios
- Representación.
- Restricciones Asociaciones prohibidas.

Proposiciones que tienen que cumplir los objetos.

Clases de restricciones.–

- Inherentes .– La da el modelo. Introduce rigideces a la hora de modelar.
- De usuario .– Facilitan al modelizar. Permiten expresar mejor la semántica del universo del discurso.

Definición Formal $S = \langle Se, Sr \rangle$

Se Reglas de definición de estructuras.

Sr Reestricciones de usuario, porque las restricciones inherentes van ligadas a las reglas de definición.

- La aplicación de la parte estática al UD da como resultado un ESQUEMA o estructura de datos que representa dicho UD en el correspondiente Modelo.

- Dinamica

Ocurrencia valores de esquema en un momento daod.

$O(BDi) = BD$;

Operación

Tipos de operación :

- Selección Localización de ocurrencias de una entidad determinada.
- Camino Especif. (Leng. Navegacionales).
- Codición Sistemas de especificación.
- Acción Tras localización se permite modificación (selección).

Insercción.

Borrado

Modificación.

Representación (en los lenguajes).

SELECCIÓN <condición>

<condición> := <expresion logica> | <camino>

Accion

SELECCIÓN <cond> ACCION <objetivo>

<objetivo> := entidad que se alige para realizar acción.

- La distinción entre selección y accion es de tipo formal.

MODELOS

Fisicos y Logicos

Niveles

Externo

Conceptual

Modelos logicos con niveles

Externo

Conceptual

Fisicos con nivel interno

- Modelos Logicos
- Conceptuales (CASE)
- Convencionales (Paquetes).
- Conceptuales. –
 - Mayor nivel de abstracción
 - Interface entre el usuario y el informatico.
 - Primeras fases del diseño.
- Modelos convencionales
 - Menos abstractos.
 - Interface entre usuario y el ordenador o sistema.

Ej. –

RED

JERARQUICO

RELACIONAL.

- Comparación entre modelos convencionales. –

Características :

Sencillez

- Relacional más sencillo (Trata de igual modo las entidades y relaciones) Inconveniente.

Independencia Fisica – Logica

- Mayor en relacional.
- En el jerarquico hay que especificar camino fisico.

Eficiencia

- Relacional es menos eficiente (Depende del equipo).

TEMA 3

Modelo Entidad – Relación.

- Trabajo Real con BD
- Modelo Entidad – Relación
 - Esquema Entidad – Relación.
- Elección de modelo convencional.
- Transformación del esquema Entidad – Relación al modelo convencional elegido.

Diagrama Programas (PDL)

- Programa Construcción de BD.
 - Observaciones.–

Modelo basado en 2 conceptos.

Entidad

Relación.

Idea intuitiva.

- Entidades.– Objetos sobre los cuales queremos guardar información y la característica es que tengan existencia por si mismos.
- Relaciones .– Asociaciones entre entidades.

Estática del MER

- Elementos.
- Entidades
- Atributos.
- Relaciones.
- Valores y dominios.
- Restricciones.

Entidades .– Objeto real o abstracto del que al menos queremos guardar información en la BD.

Representación Gráfica

<Nombre entidad>

Características.–

- Que tenga existencia propia.
- Que se pueda distinguir de los demás elementos del UD.
- Todas las ocurrencias tiene los mismos atributos.

$E = \{ \text{elementos} / \text{Prod} \}$

Clases :

- Fuertes Tienen existencias por si mismos.
- Debiles Existen por otra entidad.

Representación.

Relaciones .– Correspondencia entre entidades. (Matematicamente existe relación).

Representaciones

Etiqueta (Nombre de relación).

Arcos del rombo a las entidades relacionadas.

- Entre dos entidades pueden existir más de un tipo de relación.

Relaciones reflexivas.

Solo interesa características de personas sin hacer distinción.

Relaciones entre más de dos entidades. {Tener cuidado con relaciones engañosas}

- Pueden ser falsas relaciones. (Puede haber parte de la relación que no se cumpla semanticamente)
- El criterio para saber si son falsas es que emite SEMANTICA ENGAÑOSA .

Son falsas respecto de la realidad Los libros son publicados por los autores y editoriales.

- Relación semanticamente correcta.
- Existirian relaciones que darían problemas al descomponer porque se pierde la relación directa.
- La determinación de si una relación es falsa o no es cuestión del diseñador.
- En el ejemplo se quiere guardar información precisa de en que libro concreto, un autor concreto escribe sobre un tema concreto (Seguir un criterio claro indicado en el enunciado).

Elementos de relaciones.

- Nombre.
- Grado .– Numero de tipos de entidad que participan en una relación.

Ejs.– Grado 1.– Reflexivas.

Grado 2.

Grado $N > 2$ Falsas / Ciertas.

Tipos de correspondencia o de relación.

- Def.– Es el nº maximo de ocurrencias de cada tipo de entidad que pueden intervenir en una ocurrencia del tipo de relación
- Representación.– Se etiqueta con

1 : 1

1 : N

N : M

Ej :

Papel.– Relaciones son bidireccionales y cada dirección debe tener un nombre.

Def.– Función que cada tipo de entidad realiza en la relación.

Dominios, Valores y Atributos

. Características de las entidades y relaciones.

- Atributos Propiedades características que tiene un tipo de entidad o relación.
- Valores Contenido concreto de los atributos.
- Dominios.– Conjunto de las posibles valores de los atributos.
- Representación grafica :

Solo se representan graficamente los atributos :

- Un ovalo por atributo.
- Un arco de atributo a la entidad o relación a la que pertenece.

Ej :

- Decidir por la semantica que atributos pertenecen a la entidad y cuales a la relación.

Tipos de atributos. (Atributos de identificación).

- AIP (Atrib. De identificación principal)
- AIC (Atrib de Identificación candidato).
- AIA (Atrib de indentificación Alternativos).

AIP.– Atributos que identifican univocamente cada ocurrencia de la entidad o relación.

- Tiene que ser minimo, de modo que si quito una parte ya no es AIP.
- No son unicos.

AIC.– Para tipos de entidad todos los posibles de AIP.

AIA.– Los AIC no AIP.

- Observación AIP de relaciones.– Habitualmente son la concatenación de los AIP de las entidades que intervienen.

- Dependiendo de la semántica del problema (Criterio de enunciado) se toman unos u otros atributos como AIP para identificar unívocamente las ocurrencias de la relación.
- El único objetivo de guardar información de entidades es la de identificar unívocamente la relación entre determinadas ocurrencias.

4) **Restricciones.**

- Inherentes .– No hay.
- De usuario.
 - Valores / Dominios.– Contenido de los atributos. Condiciones sobre los valores de los atributos.

Ej .– Atrib Edad _socio [1, 120]

- Nº de ocurrencias (entidades / Relaciones).– Grado de relaciones y tipo.
- Restricciones.– No tiene una representación gráfica precisa.

Modelo E/R Extendido.

Observaciones.– Completa el modelo E/R Básico en algunos aspectos que trataba mal.

Extensiones.–

- Cardinalidades.
- Dependencias.
- Existencia.
- Identificación
- Relaciones exclusivas.
- Generación y herencia.
- Dimensión temporal
- CARDINALES.

Definición .– Nº máximo y mínimo de un tipo de entidad que pueden estar interrelacionados con una ocurrencia del otro u otros tipos que participan en el tipo de relación.

Representación

(Cardinal min, Cardinal max) .– en cada arco que va a una entidad.

Valores.

(0,1), (1,1), 0,N), (1,N)

Ej :

* Para rellenar un cardinal se ha de mirar desde el otro extremo de la relación.

- DEPENDENCIAS.
 - Al igual que para los tipos de entidad, los tipos de relación pueden ser regulares y débiles, según se

asocian a dos entidades fuertes (regulares) o una fuerte y una debil (Debil), respectivamente.

- Los tipo de relación regular o fuerte no tienen ningún tipo de dependencia.
- En los tipo de interrelación debil se distinguen dos tipos de dependencia.–
- **En existencia .–** Cuando las ocurrencias de un tipo de entidad débil no pueden existir si desaparece la ocurrencia de la entidad fuerte de la que dependen.
- **En identificación.–** Cuando, además de ser una dependencia en existencia, las ocurrencias de la entidad débil no pueden identificarse únicamente mediante los atributos propios de la misma, y hay que añadir el AI (AIP) del tipo de entidad regular de la que dependen.

- **RELACIONES EXCLUSIVAS**

- Dos o más tipos de relaciones son exclusivas cuando cada ocurrencia de un tipo de entidad sólo puede pertenecer a un tipo de relación.
- No se puede utilizar el modelo básico de representación.
- O se produce una relación o se produce otra pero nunca ambas a la vez.

- **GENERALIZACION Y HERENCIA.**

- La generalización responde a la necesidad de descomposición de tipos de entidad en varios subtipos. La interrelación que se establece entre un **supertipo** y sus **subtipos** corresponde al concepto **ES UN TIPO DE**.
- Esta relación tiene la característica de que toda ocurrencia del subtipo es una ocurrencia del supertipo, pero no al contrario, de modo que las **cardinalidades** serán siempre **(1,1)** en el caso del **Supertipo** y **(0,1)** en el caso de los **Subtipos**.
- Podemos distinguir entre dos tipos distintos de solapamiento atendiendo a los siguientes criterios.–

Solapamiento.– Cuando una ocurrencia del supertipo puede pertenecer a más de un subtipo (No Exclusividad).

Exclusividad = No solapamiento.

Representación gráfica similar a la de relación exclusiva.

Totalidad.– Si toda ocurrencia del supertipo tiene que pertenecer a algún subtipo.

Representación Añade un círculo al arco Supertipo Triángulo.

Ej 1.– Generalización total sin solapamiento:

Ej 2.– Generalización parcial sin solapamiento.

Ej 3.– Generalización total con solapamiento.

Ej.– 4 Generalización parcial sin solapamiento.

- **Herencia.–** Otra característica muy importante de este tipo de relaciones.

- Todo atributo del supertipo pasa a ser atributo de los subtipos
- Los atributos comunes a todos los subtipos se asignan al supertipo, mientras que los atributos específicos se asignan al subtipo correspondiente. Lo mismo ocurre con las interrelaciones.

EJ.– Las personas se dividen en decentes e indecentes. Interesa guardar Nombre, Apellido y DNI (único).

Supertipo.– Atributos comunes a todos los subtipos (Ej. Anterior). Relaciones con todos los subtipos se establecen con el supertipo.

Se añade una semántica.

Las personas tienen una serie de derechos humanos.

Las personas indecentes se relacionan con las comisarías (fichan).

Las personas decentes se relacionan con las empresas (contratos).

Subtipo .– Atributos y relaciones que pertenecen o afectan a subtipo.

Atributos discriminantes .– Aquellos atributos en función de valores que deciden la condición de `Ser un ` subtipo u otro.

Representación .– Atributo de la relación `Es un `.

Dimensión temporal.

- Estudio del tiempo es poco frecuente en el modelo E/R.

Primera aprox.

- Atributo tipo fecha.

** Cambia el cálculo de cardinalidades.

- No es semántica Establece limitación temporal a relación.

Otras aprox.–

- Bases de datos históricas.
- Incorporar tiempo como entidad.
- Estado.

Atributos derivados.– Aquellos que se calculan a partir de otros.

- En el modelo E/R se recomienda su eliminación.

Relaciones Redundantes.– Además de la existencia de atributos redundantes se ha de estudiar en los ciclos, si pudieran existir relaciones redundantes

- Se suele mirar si la relación que cierra el ciclo es redundante.
- Se determina que una relación es redundante si la semántica de la relación puede ser determinada a partir de otras.

- La relación redundante se puede mantener si nos obligan a almacenar información de la relación.

TEMA 4.

El Modelo Relacional. Estática.

HISTORIA Y OBJETIVOS DEL MODELO.

Codd, (1970), propone un modelo de datos basados en la Teoría de las relaciones, donde los datos se estructuran lógicamente en forma de relaciones (TABLAS), siendo un objetivo fundamental mantener la independencia de la estructura lógica respecto al modelo de almacenamiento y a otras características del tipo físico.

Estimación del modelo .-

- Modelo Teorico (1970)
- Prototipos :
- Sistema R (IBM)
- Ingres.
- Falta de eficiencia.
- Prototipo de investigación no comercializado.
- espera de desarrollo tecnológico.

Objetivos .-

- Independencia física .- Que el modo en que se almacenan los datos no influya en su manipulación lógica, y por tanto, no sea necesario modificar los programas por cambios en el almacenamiento físico. (Codd concede mucha importancia a este aspecto Independencia **de ordenación** , independencia **de indexación** e independencia **en criterios de acceso**).
- Independencia Lógica .- Que la modificación de objetos en la base de datos no repercuta en los programas y/o usuarios que esten accediendo al subconjunto parcial de la base de datos.
- Flexibilidad .- Poder presentar a cada usuario los datos de la forma que prefiera.
- Uniformidad.- Las estructuras logicas de datos presentan un estado uniforme.
- Sencillez.

Para conseguir esto Codd introduce el concepto de **Relación Tabla** como estructura básica del modelo. Todos los datos (Entidades / Interrelaciones) de un base de datos se representan en forma de Tablas cuyo contenido varia en el tiempo. Una tabla , es un conjunto de filas (Misma estructura , cabecera) en la terminologia relacional.

Surge la **Teoría de la Normalización** cuyas tres primeras formas normales fueron introducidas por Codd y constituyen un soporte para un diseño de bases de datos relacionales.

Tablas Regulares / Uniformes . Deben evitar anomalías.

Con respecto a la componente dinamica del modelo se proponen una serie de operadores que se aplican a las relaciones (Algunos clasicos de la Teria de conjuntos). Todos ellos forman el **Algebra relacional** (Proporciona la mayoría de los objetivos del modelo).

Estatica del Modelo Relacional.

El elemento basico del modelo relacional es la Tabla (Relación).

- Se puede distinguir un conjunto de columnas, (Atributos), que representan las propiedades de la misma y que estan caracterizadas por un nombre.
 - Un conjunto de filas llamadas Tuplas, que son las ocurrencias de la Tabla.
 - El numero de filas de la relación corresponde a la cardinalidad.
 - el número de filas es el Grado (Nº de atributos).
 - Dominios Donde los atributos toman sus valores (Casillas).
-
- Una tabla siempre tiene un nombre y en ella es posible distinguir :
-
- Intensión o Cabecera Esquema de la relación o intensidad. (Define la estructura de la Tabla).
 - Extensión conjunto de Tuplas que varían en el tiempo.

Dominio y Atributo.

Dominio Conjunto de valores homogeneos y atomicos, caracterizados por un nombre. Todo dominio debe tener un nombre por el que referimos a el y un tipo de datos. Los dominios pueden definirse por intensidad y por extensión.

Un atributo en el papel tiene un determinado dominio en la relación.

En el UD de un base de datos, esta compuesto por un conjunto finito de relaciones. Cada atributo toma sus valores de un unico dominio y varios atributos pueden tener el mismo dominio.

Relación

- Intensión o esquema de la relación .– Es la parte definitoria y estatica de la relación. (Cabecera)
- Extension, ocurrencia o instancia de la relación.– conjunto de tuplas, que en una instante determinado satisface el esquema correspondiente.

CLAVES .

- **Clave candidata** Conjunto no vacio de atributos que identifican univoca y minimamente cada tupla.
- **Clave primaria** La que el usuario escoge de las claves candidatas.
- **Claves alternativas** Claves candidatas que no han sido escogidas.
- **Clave ajena** Conjunto de atributos de la tabla cuyos valores han de coincidir con los de la clave primaria de otra tabla. (Clave ajena y primaria debe estar definida sobre los mismos dominios).

RESTRICCIONES

INHERENTES.

- No hay dos tuplas iguales.
- el orden de las tuplas no es significativo.
- El orden de los atributos no es significativo.
- Cada atributo solo puede tomar un valor del dominio, no admitiendose por tanto dos grupos repetitivos.
- Se debe cumplir la regla de Integridad de entidad ` Ningún atributo que forme parte de la clave primaria de una relación puede tomar un valor desconocido o inexistente'.

DE USUARIO.

Integridad referencial Si la relación R2 tiene un descriptor que referencia a la clave primaria de la relación R1

todo valor de dicha clave (Clave ajena R2) debe concordar con los valores de la clave primaria R1 o ser NULO. R1 y R2 son claves necesariamente distintas. Además, la clave ajena puede formar parte de la relación R2.

También se deben definir las acciones a tomar en caso de acciones de modificación y borrado :

- Operación RESTRINGIDA .– Solo se puede borrar una fila de la tabla que tiene clave primaria referenciada si no existen filas con esa clave en la tabla referenciada.
- Operación con transmisión en cascada.– El borrado o la modificación de una fila de la tabla que contiene la clave primaria lleva consigo la modificación de las tablas cuya clave ajena coincida con la clave primaria modificada.
- Operación con puesta a nulos.– El borrado o la modificación de una fila de la tabla que contiene la clave primaria lleva consigo la puesta a nulos de los valores de la clave ajena de las filas de la tabla que referencia la clave coincida con el valor de la clave primaria de la tabla referenciada.

Otras restricciones.–

- Definir predicados sobre :
 - Atributo de una relación.
 - Tuplas de una relación.
 - Atributos de varias relaciones.
 - Dominios.
 - Evaluación de restricciones.
 - Con cada operación
 - Diferido.

TRANSFORMACION DEL MODELO E/R AL MODELO RELACIONAL.

Tres principios :

- Toda entidad se transforma en una tabla.
- Toda interrelación M :N se transforma en una tabla.
- Toda interrelación del tipo 1 :N se traduce en el fenómeno de propagación de clave o se crea una nueva tabla.

Reglas de transformación del Modelo Básico.

- Transformación de Dominios .– Hay sistemas relacionales que no lo permiten.
- Trans. de entidades .– Toda entidad se transforma en una tabla.
- Transformación de atributos .– Cada atributo de una entidad se transforma en una columna de la tabla.
- AIP Clave primaria de la relación.
- Transformación interrelaciones .– Según tipo :
 - N :M Se transforma en una nueva tabla cuyos atributos son la concatenación de los AIP de las entidades que asocian.
 - Cada uno de estos atributos es clave ajena de la tabla donde estos atributos son clave primaria.
 - CLAVE AJENA (<Nombre columna>)

REFERENCIA TABLA (<Nombre tabla>)

TABLA ESCRIBE (Cod_Libro, Nombre)

CLAVE PRIMARIA (Cod_Libro, Nombre)

CLAVE AJENA (Cod_Libro)

REFERENCIA TABLA LIBRO

EN BORRADO : CASCADA

EN ACTUALIZACION : CASCADA

CLAVE AJENA (Nombre)

REFERENCIA TABLA AUTOR

EN BORRADO : NULOS

EN ACTUALIZACION : CASCADA

- 1 :N Dos soluciones :
- Propagación :
 - Propaga AIP de entidad con cardinalidad Maxima 1
 - Perdida de semantica
- Considerarla como una relación M :_N

¿Cuándo A y cuándo B ?

- Si la relación al final resultará una M :N mejor B
- Si hay atributos propios en la relación B
- Si el nº de ocurrencias interrelacionadas de la entidad que produce la clave es pequeña o bien la cardinalidad mínima es 0 (Nulos) B
- Perdida de semantica en cardinalidades :
- Cardinalidad maxima (Predicados)

1 :N) Maximo es conocido (En tablas se permite limitar el número de filas)

- cardinalidad minima (Tabla distinta para la relación) Predicado
- Car minima Propagando claves (Permitir NULOS en clave ajena)
 - Interrelaciones 1 :1
 - Caso particular de las relaciones 1 : y M :N
 - No existe una regla fija para su transformación , pudiendose crear una tabla nueva o propagar la clave. En este último caso, la propagación de la clave puede realizarse en ambas direcciones.
- Criterios para la selección de una solución :

- Cardinalidad minima.
- Semantica
- Eficiencia (Nulos o no Nulos / Accesos más frecuentes)
- Si la cardinalidad es (0,1) en ambas entidades la interrelación se transforma en una nueva tabla.
- Si la cardinalidad de una de ellas es (0,1) y la otra es (1,1) conviene propagarla la clave de la entidad con cardinalidad (1,1) a la tabla resultante de la entidad con cardinalidad (0,1)
- Si en ambas la cardinalidad es (1,1) se puede propagar clave en ambos sentidos o incluso propagar ambas claves.

Solución en el modelo Relacional.

Ej 1

EMPLEADO (Cod_emp,)

DEPARTAMENTO (Cod_Dep,, Cod_emp)

Ej 2.-

MATRIMONIO (Cod_H, Cod_ M)

HOMBRE (Cod_H ,)

MUJER (Cod_M,)

Caso 1 : 1 Cardinalidades en ambos (1,1) , (1,1)

- Propagar las dos entidades.
- Accesos más frecuentes (Primar estos accesos)

Transformación de los atributos de interrelaciones.

- Si la interrelación se transforma en una relación , todos sus atributos pasan a ser columnas de la relación .
- Si se transforma mediante propagación de clave, sus atributos migran junto con la clave de la relación correspondiente, aunque suele ser mejor crear una nueva tabla para realizar esta migración de la interrelación con sus atributos.

Transformación de restricciones de usuario.

- Se añaden como predicados al definir las tablas.
- Tipos :
 - Comprobación del rango de valores de los atributos.
 - Especificar todos los valores posibles de un atributo
 - Comprobar un predicado. (Ej : sobre valor de un atributo).
- NOTACION :
 - <atributo> RANGO ENTRE < Intervalo (Valores)>
 - <Atributo> EN < conjunto de valores>
 - COMPROBAR (<condición>, <Atributo>)

TRANSFORMACION DE REGLAS DEL MODELO EXTENDIDO.

Transformación de dependencias en identificación y en existencia.

(*) No reflejado por el modelo relacional

- Utilizar el mecanismo de propagación de clave, creando una clave ajena con NULOS no permitidos en la tabla de la entidad dependiente, obligando a la modificación y borrado en cascada.
- si la dependencia es en identificación la clave primaria de la tabla de la entidad débil, debe estar formada por la concatenación de las claves de las dos entidades que participan en la interrelación.

SOLUCION (Ej)

TABLA LIBRO (Cod_ Libro,)

CLAVE PRIMARIA (Cod_Libro)

TABLA EJEMPLAR (Cod_Libro ; Cod _ Ejem,)

CLAVE PRIMARIA (Cod_Libro, Cod_ejem)

CLAVE AJENA (Cod_Libro)

REFERENCIA A TABLA LIBRO

EN BORRADO : CASCADA

EN MODIFICACION : CASCADA

Relaciones exclusivas

- En cada caso según su semántica.

Edita 1 EXCLUYE Editar 2

- Propagando claves :

TABLA LIBRO (Cod_Libro, CodEd, Cod_Univ.....)

CLAVE PRIMARIA (Cod_libro)

CLAVE AJENA (CodEd)

REFERENCIA A TABLA EDITORIAL

EN BORRADO : CASCADA

EN ACTUALIZACION : CASCADA

CLAVE AJENA(Cod_Univ)

REFERENCIA TABLA UNIVERSIDAD

EN ACTUALIZACION Y BORRADO : CASCADA

COMPROBACION [(CodEd == NULO) y (Cod_univ != NULO)] o

[(CodEd != NULO) y(COD_univ == NULO)]

Transformación de TIPOS y SUBTIPOS.

- No esta recogido directamente en el modelo Relacional.
- Se pierde semantica.
- Opciones :
 - Una tabla que englobe los atributos del supertipo y de los subtipos.
 - Cuando los subtipos se diferencian en muy pocos atributos
 - Las interrelaciones que les asocian con otras entidades sean las mismas para todos los subtipos.
 - Se debe añadir a la tabla un atributo adicional. (DISCRIMINANTE).
 - Una sola tabla para el Supertipo y tantas tablas como subtipos haya con sus atributos correspondientes.
 - Cuando existen atributos diferentes en los subtipos y se quieren mantener los atributos comunes en una tabla.
 - Tablas distintas para cada subtipo que tengan además los atributos comunes
 - Cuando existan atributos distintos entre los subtipos
 - Cuando los accesos realizados a los datos de los subtipos siempre afectan a los atributos comunes.

Evaluación de eficiencia.

- Acceso a una fial que refleje todos los datos de una entidad es mucho más rapido.
- Mejor desde punto de vista semantico.
- Aumenta eficiencia en determinadas consultas. Se produce redundancia y perdida de semantica.

Dimensión Temporal.

OBS Puede ocurrir.

- Atributo temporal forma parte de las claves. (Fecha diferencia una ocurrencia de otras)

Atributos derivados.

- No hay representación directa.
- Opciones :
 - Considerarlo una columna más :
 - Procedimientos para recalcular valores cada vez que se modifique la tabla.
 - Disparadores .- Sin columnas en las tablas. Procedimientos que calculan los datos en el momento en que se piden.

TEMA 5

NORMALIZACION

Introducción.

Los modelos de datos son instrumentos (objetos y reglas) que nos ayudan a representar el UD. El proceso de diseño de una base de datos consiste en representar un determinado UD mediante los objetos que proporciona el modelo de datos (estructuras) aplicando para ello las reglas de dicho modelo (restricciones inherentes).

Cuando realizamos un diseño en el modelo relacional, existen diferentes alternativas, pudiendo obtener diferentes esquemas relacionales, no todos ellos serán equivalentes y unos representarán mejor la información que otros.

Propiedades que debe tener el esquema relacional para representar adecuadamente la realidad y que problemas se pueden derivar del diseño inadecuado.

La información de una base de datos puede representarse mediante un conjunto de objetos (estructuras), dominios (atributos) y relaciones y un conjunto de reglas de integridad (Restricciones).

No se cuentan con las restricciones del modelo E/S

En el modelo relacional el diseño puede realizarse de dos formas :

- Obtener el modelo relacional directamente de la observación del UD.
- Realizar el diseño ME/R y transformarlo al modelo relacional.

Las relaciones obtenidas pueden presentar problemas :

- Incapacidad para representar ciertos hechos.
- Redundancia en la información, e incoherencias en la misma.
- Ambigüedades.
- Aparición en la base de datos de estados no validos en el mundo real. (anomalías en modificación, inserción, borrado).

Ejemplo de la guía (pag 2).

ESCRIBE (autor, nacionalidad, cod_libro, titulo, editorial, año) ;

Presenta varios problemas :

- Gran cantidad de redundancia La nacionalidad del autor se repite en cada ocurrencia del mismo. Cuando un libro tiene más de un autor la editorial y el año se repiten tambien.
- Anomalías de modificación Puede ocurrir que se modifique el nombre de editorial en una fila sin modificarla en el resto que corresponden al mismo libro.
- Anomalías de inserción No sería posible la inserción de un autor del que no hubiera ningun libro (cod_libro Clave primaria) , tampoco podría haber obras anonimas. La inserción de un libro con más de un autor obligaría a la repetición de tuplas.
- Anomalías de borrado Si se quiere dar de baja un libro tambien se perdería información de los autores y viceversa.

Si se hubiera seguido la metodología de diseño de temas anteriores, no se habría presentado una relación de

este tipo. La realización de un modelo conceptual en ME/R y la transformación al modelo relacional evita estos problemas.(Aplicación de restricciones de usuario).

Ante cualquier duda de si el modelo es correcto, es preferible la aplicación de un método formal de análisis Tería de la Normalización.

Necesidad de Normalizar.

- Teóricamente están unidas el modelo relacional y la Normalización.
- Objetivo Eliminar anomalías.
- Redundancia Existen anomalías de Actualización (Se debe actualizar en todas las tuplas) y Borrado (Puesta a NULOS de todos los campos).

Aproximación por descomposición.

- Relación universal. (Todos los atributos).
- Proceso de depuración (Sucesiva).
- Operaciones sobre relaciones :
- Proyección Eliminar algunos atributos y eliminar tuplas repetidas.

Ej Proyección de propietario sobre los atributos nombre y apellido.

- Reunión natural
- Reunión de relaciones $R(A_1, \dots, A_n)$ y $S(B_1 \dots B_n)$

Unión de atributos.

Tuplas son concatenación de R y S con los mismos valores de atributos del mismo nombre.

Ej .-

- La proyección de la reunión sobre R da la relación R y la proyección de la reunión sobre S da la relación S.
- Descomposición Sustituir la relación $R(A_1 \dots A_n)$ en una serie de relaciones $R_1 \dots R_n$ de manera que la reunión de R_1 hasta R_n tiene el mismo esquema que R.
- Descomposición sin pérdida Sustituir la relación $R(A_1 \dots A_n)$ en una serie de relaciones $R_1 \dots R_n$ de manera que la reunión de R_1 hasta R_n tiene el mismo esquema que R, que cumple que para toda extensión de R se cumpla que extensión de R = Extensión de la reunión ($R_1 \dots R_n$).

Ej.-

- Si admitimos tipo 1 marca y 1 potencia.
- Comprobar que D1 si lo hace y D2 no permite recuperar el color.
- El diseño relacional desde este punto de vista, se puede ver como el proceso de descomposición sin pérdida.

Formas Normales.

La teoría de normalización consiste en obtener esquemas relacionales que cumplan unas determinadas condiciones y se centra en las determinadas Formas normales. Se dice que un esquema de relación está en una determinada forma normal si satisface un conjunto determinado de restricciones.

1FN Codd . Restricción inherente del modelo relacional . En una relación no puede haber grupos repetitivos.

2FN y 3FN Codd.

FNBC Boyce y Codd . Redefinición de 3FN.

4FN y 5Fn Fagin.

Relación entre formas normales.

Una relación esta en 5FN si tambien lo esta en todos las anteriores.

Una relación que esta en 1FN no tiene porque estar en 2FN.

- Objetivo Obtener la forma normal mayor posible.

Enfoque Dependencias .

Dependencias funcionales 2FN, 3Fn, FNBC

Dependencias multievaluadas 4FN

Dependencias de proyección / combinación 5FN

Normalización. Enfoque intuitivo.

2FN Una relación está en 2FN si además de estar en 1FN todos los atributos que no forman parte de ninguna clave candidata suministran información a cerce de la clave completa.

PRESTAMO (num_socio, nombre_socio, cod_libro, fecha_presentamo, editorial, pais)

Claves candidatas :

(num_socio, cod_libro)

(nombre_socio, cod_libro)

Para pasarlo a 2FN :

PRESTAMOS1 (num_socio, nombre_socio, cod_libro, fecha_presentamo)

LIBROS (cod_libro, editorial, pais)

PRESTAMOS1 Atributo fecha_prestamos, no forma parte de la clave pero suministra información de claves candidatas.

LIBROS Dos atributos no clave suministran información de la clave completa.

Toda relación cuya clave esta formada por un solo atributo está en 2FN.

3FN Si además de estar en 2FN, los atributos que no forman parte de ninguna clave candidata facilitan información solo acerca de las claves y no acerca de otros atributos.

PRESTAMOS1 3FN

LIBROS pais facilita información a cerca de editorial (No 3FN).

LIBROS1 (cod_libro, editorial)

EDITORIALES (editorial, pais)

- Todos los atributos que no forman parte de la clave candidata deben ser información referida a la clave, clave completa y nada más que la clave.

FNBC En PRESTAMOS1 nombre_socio y num_socio se repiten por cada libro y prestamo.

En FNBC si el conocimiento de las claves permite averiguar todas las relaciones existentes entre los datos de la relación. Las claves candidatas deben ser los únicos descriptores sobre los que se facilita información por cualquier otro atributo.

Nombre_socio se refiere a num_socio y viceversa. (Ninguno de ellos es clave aunque formen parte de ella).

Pasar a FNBC :

SOCIOS (num_socio, nombre_socio)

PRESTAMOS2 (num_socio, cod_libro, fecha_prestamo)

Esquema :

LIBROS1 (cod_libro, editorial)

EDITORIALES (editorial, pais)

SOCIOS (num_socio, nombre_socio)

PRESTAMOS2 (num_socio, cod_libro, fecha_prestamo)

Concepto de dependencia funcional.

Las dependencias son propiedades inherentes al contenido semántico de los datos que se han de cumplir para cualquier extensión del esquema de relación y forman parte de las restricciones de usuario del modelo relacional. Las dependencias, muestran interrelaciones existentes entre los atributos del mundo real cuya semántica tratamos de incorporar a nuestra base de datos. Son invariantes en el tiempo.

Puesto que las dependencias constituyen una parte importante de la semántica de nuestro UD, deben ser incluidas en los esquemas de relación. De modo que un esquema de relación será un par :

<A, DEP>

A Conjunto de atributos de la relación.

DEP Conjunto de dependencias existentes entre dichos atributos.

Las tres primeras formas normales tienen dependencias funcionales. En el proceso de normalización será fundamental identificar todas las dependencias funcionales del universo del discurso cuyo diseño estamos realizando y procuraremos conservar dichas dependencias a lo largo de todo el proceso.

26