



Diseño de Sistemas basado en TICS

Base de Datos

Bases de datos relacionales y RDBMS
Introducción al Modelo Entidad Relación
Introducción al lenguaje SQL

Aldo Di Biase Friedmann



Diseño de Sistemas basado en TICS

Bases de datos relacionales y RDBMS

Concepto de base de datos
Bases de datos relacionales
Productos comerciales

Aldo Di Biase Friedmann

Definición de Base de Datos

- Una base de datos es cualquier repositorio de información que permanece aunque el computador no esté encendido
 - Es información que se requiere mantener por un “período largo” de tiempo

Base de Datos v/s programas ejecutables

- Sin embargo, los programas ejecutables también satisfacen esta condición
 - Incluso el sistema operativo, controladores de impresora y los administradores de base de datos
 - De hecho es la definición de un “computador de von Neumann”
- Luego, una base de datos contiene los datos no los algoritmos o procesos para procesar esos datos
 - Aún la línea divisoria no es tan clara; por ejemplo, en el código puede estar en “duro” un parámetro o partes del algoritmo puede estar fuera del programa “juntos a los datos”
 - Ej: La nota mínima para aprobar es un 4,0

Tipos de base de datos

- Existen diferentes tipos de modelos para almacenar los datos
 - Jerárquico
 - De redes
 - Relacional
 - De estrellas
- En este curso nos interesa únicamente el modelo relacional

Ventajas del modelo relacional

- Está basado en la teoría (matemática) de conjuntos desde su inicio, recibiendo gran aporte teórico en su formulación y refinamiento
 - Establecido por el matemático Dr. Codd en 1972
- Es flexible ante cambios de diseño
 - Permite incorporar nuevos requerimientos de una manera “no traumática”
- Evita la duplicidad de datos
 - No hay errores de consistencia
- Es independiente de la plataforma de HW y SW sobre la que se vaya a construir

Productos comerciales

- Multiplataforma
 - Oracle
 - DB2
- Plataforma Windows (principalmente)
 - SQL Server
 - Access
- Plataforma Unix / Linux (principalmente)
 - Informix
 - My SQL

Características deseables de un RDBMS

- Fácil uso
- Confiable
 - No debe corromper los datos
 - % Tiempo en funcionamiento
 - Capacidad para realizar respaldos y recuperaciones de datos
- Uso de “pocos recursos”
- Buena performance
- Precio
- Complejidad de la base
 - Cantidad de usuarios
 - Tamaño de la base de datos
 - Escalable

Herramientas de diseño

- Además de las anteriores, existen una serie de herramientas de diseño
- Algunas de ellas de aplicación en cualquier RDBMS, otras dependientes del RDBMS específico
- Permiten documentar todo el proceso de diseño y construcción, facilitando la generación de los detalles que se agregan en cada etapa

Diseño de Sistemas basado en TICS

Introducción al Modelo Entidad Relación

Modelo lógico
Normalización
Modelo físico

Modelo lógico

Entidades
Atributos y Llaves
Relaciones y llaves foráneas
Diagramas

Aldo Di Biase Friedmann

Modelo lógico de datos

- Corresponde a la etapa de Diseño Lógico en el modelo en cascada
- Permite tener la visualización general de la solución, sin entrar a los detalles de su implementación en una herramienta específica
- En la práctica se realiza en forma conjunta con el diseño de la aplicación
 - Analizamos ambas áreas por separado sólo por simplicidad

Aldo Di Biase Friedmann
Página N° 12

Diseño de Sistemas basado en TICS
Universidad de Chile

Entidad

- Es cualquier cosa del “mundo real” sobre la que se quiere conservar información
- Corresponde a los sustantivos de la información recopilada de los usuarios
- Puede ser tangible (personas, lugares) o intangibles (cursos, facturas)
- Se tienen diversas instancias de cada entidad

Entidad – Ejemplo de U-cursos

- Cursos
- Secciones
- Profesores de cátedra
- Profesores auxiliares
- Profesores ayudantes
- Alumnos

Atributos

- Son los datos específicos que se desea almacenar de una entidad
- Pueden ser obligatorios o no (acepta valores “nulos”)
 - Toda persona en Chile tiene un N° de CI
 - No todas las mascotas están inscritas
- Puede o no aceptar duplicados
 - El CI es único
 - El nombre puede repetirse
- Pueden tener “subatributos”
 - Dirección: Calle, número, departamento, comuna, ciudad, región, código postal
 - Nombre: Nombres, apellido paterno, apellido materno

Diagramas

- El modelo Entidad – Relación puede dibujarse en un diagrama
- Facilita la comunicación y comprensión del modelo
- Existen varios estándares para este diagrama, dependiendo de la herramienta utilizada para generarlo

Diagramas de Entidades y Atributos

- La entidad es un cuadro (de cualquier tamaño) con bordes rectos o curvos
- El nombre de la entidad se coloca en mayúsculas
- Los atributos se colocan en minúsculas
 - Por claridad en modelos grandes podrían no colocarse
- Información adicional de los atributos puede o no ser agregada

Ejemplo de entidades y sus atributos

CURSO
Código
Nombre
Temario

PROFESORES
Rut
Nombre
Dirección
E-mail
Año de ingreso
AFP
Isapre
Título profesional
Grado académico
Categoría

Atributos y relaciones

- Se debe tener cuidado con no confundir atributos y relaciones
 - El profesor que dicta un determinado curso es una relación entre las entidades Profesor y Sección y no una propiedad de la sección

Entidades de código

- Aparecen entidades para definir la decodificación de información
- Ej: Regiones
 - RM = Región Metropolitana
 - I = Primera Región
 - II = Segunda Región
- Estas entidades permiten que la descripción utilizada sea siempre la misma aumentando la consistencia de la información almacenada
- Típicamente tienen los atributos
 - Código
 - Descripción
 - Opcionalmente una descripción corta o abreviatura
- Esto hace que algunos atributos sean en realidad relaciones

¿Debe considerarse la calle como una entidad de código?

- Si la calle es tratada como una entidad separada sólo podrán elegirse calles de esa entidad
 - Complica la administración de la lista de calles posibles a ¿Quién la administra?
 - Pero si es libre, el sistema no reconocerá como la misma calle a “Alameda” y “Alameda B. O’Higgins” y “Alameda Bernardo O’Higgins”
 - Dificulta la definición de rutas de recorrido
- Esta discusión puede generalizarse a otras entidades que en algunos casos puede ser necesario tratar en forma independiente y en otros casos no

¿Comuna o ciudad?

- Santiago es una ciudad formada por muchas comunas
- Pero, principalmente en sectores rurales, una comuna tiene varios poblados
- Luego, debe considerarse que la comuna contiene varias ciudades o al revés
- La respuesta dependerá del problema
 - En el caso de U-cursos, todas las direcciones son (en principio) de Santiago
 - Los clientes de un banco pueden vivir en cualquier parte y debe revisarse el tema con mayor detalle
- Esta situación no es frecuente, pero se debe tener presente que no siempre es posible definir una jerarquía clara entre las entidades

Información histórica

- ¿Es importante almacenar la dirección actual de cada estudiante, o se requiere conocer la información de cada una de las direcciones que ha tenido?
- La respuesta depende obviamente de los requerimientos de cada problema
- En caso se requiera, se tendrá una entidad con cada dirección y la fecha de inicio de vigencia de esa dirección
 - No se requiere la fecha de caducidad, ya que será la siguiente fecha de inicio (asumiendo que siempre vive en alguna dirección)

Ejemplos

- Curso
 - Código
 - Nombre
 - Temario
- Los requisitos no son un atributo, si no una relación con otros cursos (relaciones recursivas)
 - (Se asume que el único tipo de requisitos son otros cursos; lo analizaremos junto a las relaciones)

Ejemplos

- Profesores de cátedra
 - Rut
 - Nombre
 - Dirección
 - E-mail
 - Año de ingreso
 - AFP
 - Isapre
 - Título profesional
 - Grado académico
 - Categoría

Supertipo y subtipo

- Profesores auxiliares y ayudantes
 - Tienen los mismos atributos que los de cátedra
- Podrían considerarse como uno sólo o mantenerlos separados
 - Si comparten la mayoría de los atributos se dejan en una sola entidad con un atributo que indica el tipo de profesor
 - Si muchos atributos son diferentes se consideran entidades separadas
 - Existe también la alternativa intermedia; los atributos compartidos se dejan en la entidad principal y el resto en las entidades de los subtipos
- Al separarlas se debe tener cuidado que todos los casos están considerados
 - Puede requerirse el uso del subtipo “otros”
- En el ejemplo de U-cursos, pueden juntarse también con los alumnos en una entidad de Personas
 - Esta es la tendencia actual, tener una entidad principal para todas las personas con diferentes subtipos

Diagrama de supertipo / subtipo

SUPERTIPO

SUBTIPO 1

SUBTIPO 2

...

SUBTIPO n

Relaciones

- Corresponden a las asociaciones o interacciones entre las entidades
- En general corresponden a verbos en el levantamiento de la información de los usuarios
- Se identifican con una frase genérica y/o con dos frases, una por cada sentido de la relación
- Ejemplos
 - Los cursos son dictados por los profesores
 - Los profesores dictan cursos
 - Los alumnos toman los cursos
 - Cada curso tiene secciones

Tipos de relaciones

- Existen relaciones 1 a 1: Cada instancia de una entidad está relacionada con una (y sólo una) instancia de la otra
 - Ej: Cada empleado tiene un puesto de trabajo
- Relaciones 1 a M: Una instancia de una entidad puede estar asociada a muchas instancias de la otra; pero la inversa no es cierta
 - Ej: Un profesor dicta varios cursos; pero cada curso es dictado por un solo profesor
- Relaciones M a M: Cada instancia de una entidad puede estar relacionada con muchas instancia de la otra y viceversa
 - Un alumno toma varios cursos; un curso es tomado por varios alumnos

Aldo Di Biase Friedmann

Página N° 29

Diseño de Sistemas basado en TICS

Universidad de Chile

Relaciones obligatorias y opcionales

- Cada lado de la relación puede ser opcional u obligatorio
- En caso de ser obligatoria, cada instancia de esa entidad debe tener una relación (o al menos una, según el caso) con la otra entidad
 - Ej: Cada curso debe tener al menos una sección
- En caso de ser opcional es posible que hayan instancias de la entidad sin ninguna relación con la otra
 - Ej: ¡Una sección puede no tener alumnos!

Aldo Di Biase Friedmann

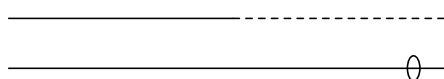
Página N° 30

Diseño de Sistemas basado en TICS

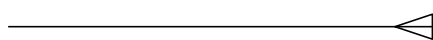
Universidad de Chile

Diagramas para relaciones

- Se usa una línea recta
 - Puede tener curvas o dobleces para facilitar la comprensión
- Una línea continua indica que es obligatoria; una línea punteada indica que es opcional
 - A veces se usa siempre continua con un 0 indicando que es opcional



- Si termina con una línea es 1 a ?; si termina con 3 líneas (“pata de gallo”) es M a ?



Relaciones obligatorias y opcionales

- Un tema interesante es que las relaciones se pueden forzar en la herramienta RDBMS, luego en algunos casos debe dejarse una relación como opcional para poder crear primero una entidad y luego la otra
 - Ej: Los cursos se crean antes que las secciones; luego pueden haber cursos sin secciones
 - Pero al crear una sección se debe relacionar inmediatamente a un curso

Información perteneciente a una relación (Asociaciones)

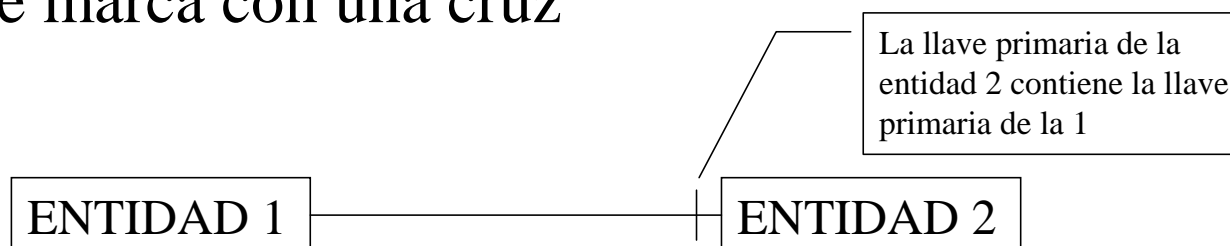
- Es posible querer almacenar información que depende de la relación, no de las entidades que la forman
- Por ejemplo, la nota de un alumno en una sección es un atributo de la relación, no del alumno ni de la sección
- En estos casos se crea una entidad que representa esta relación; en este caso, notas por alumno

Llave primaria

- Es la forma en que se diferencian dos instancias de una misma entidad
- Está constituido por:
 - Uno o más atributos
 - Una o más relaciones
 - Una combinación de atributos y relaciones
- No puede ser nulo ni estar repetido (el conjunto de atributos)
- No cambia en el tiempo
- En algunos casos existe más de una alternativa
 - Rut o N° Matrícula puede ser llave primaria para los alumnos
 - Se elige la más conveniente, generalmente la que tenga menos campos
- En algunos casos, por simplicidad, se utiliza un código inventado o simplemente un número serial

Diagramas para llaves primarias

- La llave primaria puede ser marcada en el diagrama de diferentes formas
 - Cursiva
 - Subrayado
 - Colocar un carácter especial
- Si la llave primaria está formada por la relación, se marca con una cruz

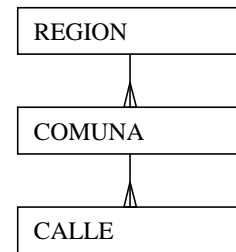


Llaves foráneas

- Las llaves foráneas corresponden a las llaves primarias de una entidad incorporadas como atributos a otra para considerar las relaciones

Jerarquías

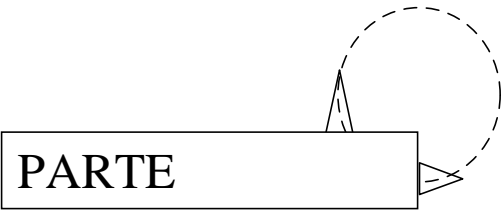
- Consideremos el ejemplo de las direcciones:
 - Toda calle pertenece a una comuna
 - Toda comuna pertenece a una región
- Llaves:
 - Llave primaria de la región: Un número entre 1 y 13
 - Llave primaria de la comuna: La llave de la región junto a un código de la propia comuna
 - Ej: Santiago Centro sería 13 (RM) + 001 (comuna)
 - Se debe definir si se pueden repetir los números de comuna o no
 - Llave primaria de la calle: La llave de la comuna (región + código) junto a la calle propiamente tal
 - Ej: Blanco sería 13 + 001 + Blanco
- Este tipo de relaciones se da con frecuencia en zonas geográficas (como el ejemplo anterior) y en la organización interna de una empresa (Universidad – Facultad – Departamento)



Relaciones con la misma entidad

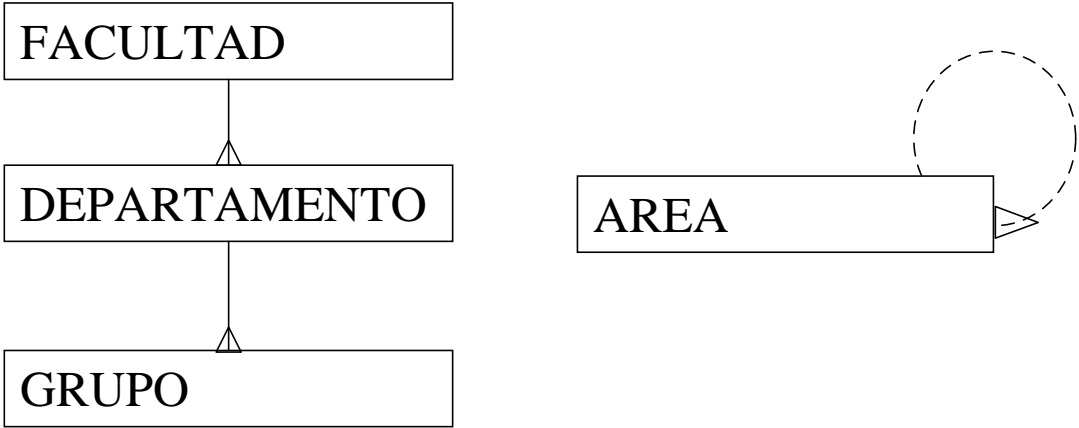
- En algunos casos se tienen relaciones recursivas; es decir, entre dos instancias de la misma entidad
- Por ejemplo, la lista de materiales para un producto:
 - El computador tiene monitor, CPU, teclado y mouse
 - La CPU tiene tarjeta madre, tarjeta de video, disco duro y gabinete
 - La tarjeta madre tiene varios chips, circuito impreso, bus, etc.
- Luego, se puede tener una relación recursiva (en este caso M a M, como es el caso típico) entre las piezas
- Este tipo de relaciones normalmente tiene algún atributo, por ejemplo la cantidad de cada parte requerida

Relaciones con la misma entidad



Relaciones jerárquicas y recursivas

- Una relación jerárquica (Facultad, Departamento, Grupo) puede ser modelada como una sola entidad con una relación recursiva
- Dependerá de cuán “parecidas” (cuán similares sean sus atributos) en cada nivel de la jerarquía



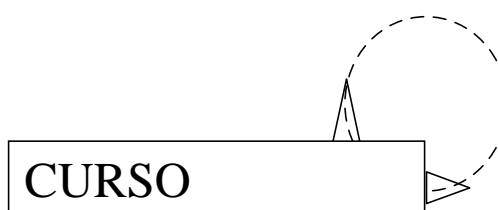
Relaciones con la misma entidad

Requisitos en los cursos

- Este es otro ejemplo de relaciones recursivas M a M
- Supuestos:
 - Los requisitos de los cursos son otro curso o una autorización
 - No existen requisitos de semestre o de autorización pura
 - Los requisitos pueden ser varios cursos en forma simultánea
 - Un grupo de cursos requisito (de otro) puede ser reemplazado por otro grupo

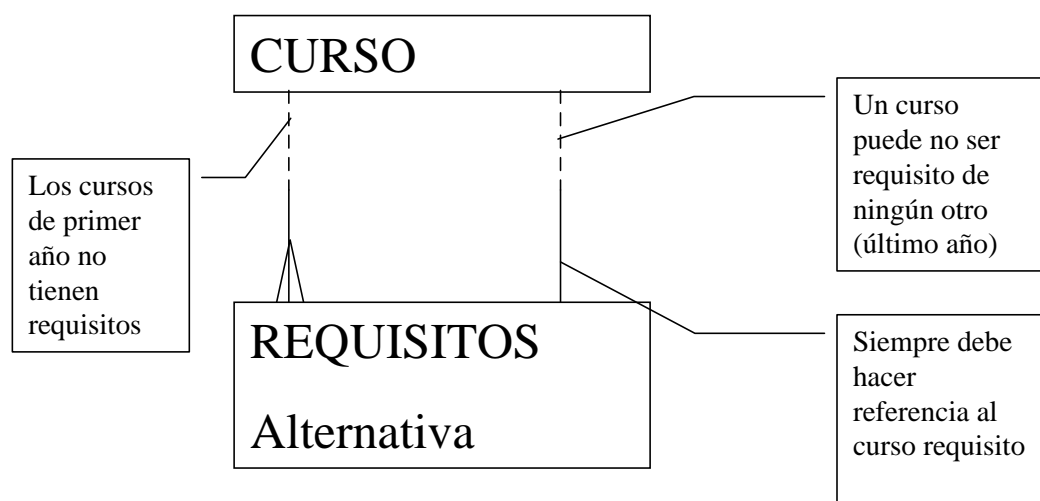
Requisitos entre cursos

- Modelo básico
 - Un curso puede tener muchos requisitos
 - Un curso puede ser requisito de varios cursos



Requisitos entre cursos

- Modelo avanzado I
 - Conjuntos de cursos requisitos alternativos
- Se incorpora una entidad REQUISITOS

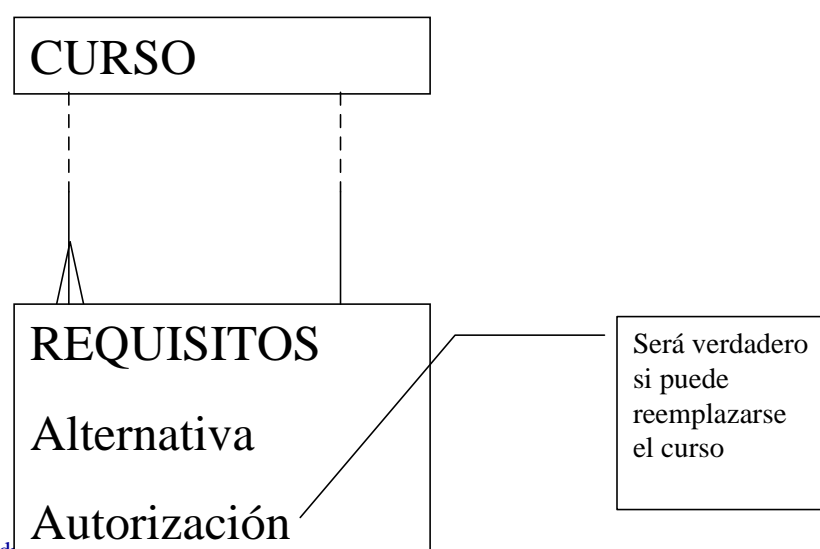


Aldo Di Biase Friedmann
Página N° 43

Diseño de Sistemas basado en TICS
Universidad de Chile

Requisitos entre cursos

- Modelo avanzado II
 - Requisito de un curso reemplazable por Autorización
- Se incorpora el atributo Autorización a la entidad REQUISITOS



Aldo Di Biase Friedmann
Página N° 44

Diseño de Sistemas basado en TICS
Universidad de Chile

Requisitos entre cursos

- Modelo avanzado III
 - Requisito de un conjunto de cursos reemplazables por Autorización
 - Es decir todas instancias de la “Alternativa” se puede reemplazar por una autorización
- ¡TAREA!

Diseño de Sistemas basado en TICS

Normalización

Concepto de normalización
Las tres formas normales principales
Desnormalización

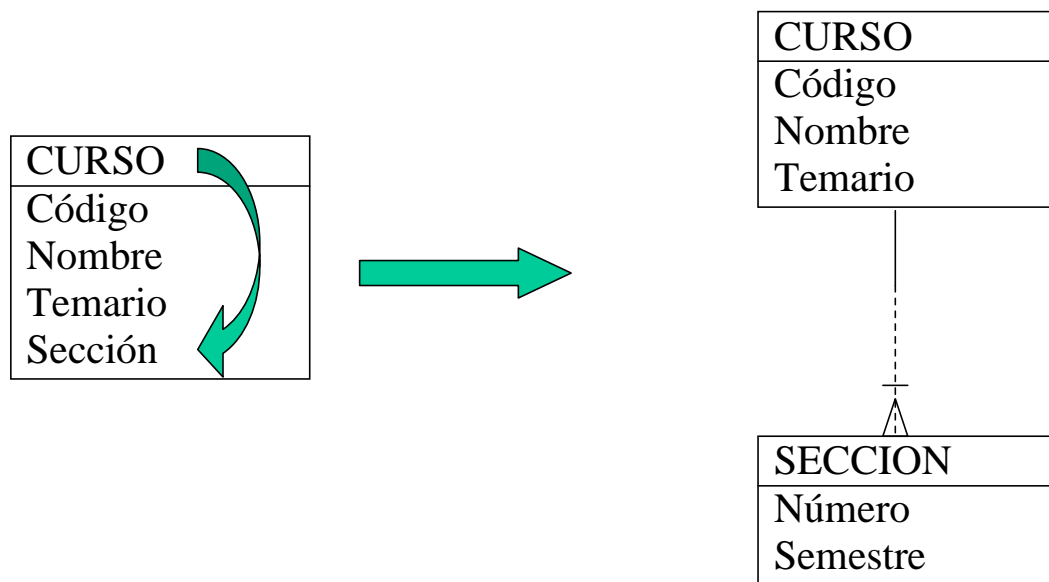
Normalización

- La normalización tiene raíces en la teoría de conjuntos y funciones
- Sin embargo, en la práctica se puede entender y aplicar sin todas las herramientas matemáticas asociadas
- Eliminan los errores de modelamiento, evitando las condiciones que producen corrupción en los datos
- Existen varias formas normales, que consideran problemas complejos de modelamiento de datos, pero en algunos casos poco frecuentes
- Por esto son importantes sólo las tres primeras

Primera forma normal

- La relación entre la llave primaria y cualquier atributo es 1 a 1 en esa dirección
- Es decir, para cada instancia se debe tener un solo valor para todos sus atributos (o ser nulos)
- Si un atributo puede tener más de un valor, se debe reemplazar por una entidad independiente y una relación M a 1

Primera forma normal



Segunda forma normal

- Un atributo debe ser dependiente de toda la llave primaria (y no sólo de una parte de ella)
 - Sólo tiene sentido revisarlo si la llave primaria es compuesta
- En este caso el atributo que depende de una parte de la llave corresponde a otra entidad
 - Puede ser necesario crear esa entidad



UNIVERSIDAD DE CHILE

Segunda forma normal

CURSO
Código
Nombre

SECCION
Número
Semestre
Temario



CURSO
Código
Nombre
Temario

SECCION
Número
Semestre

Aldo Di Biase Friedmann

Página N° 51

Diseño de Sistemas basado en TICS

Universidad de Chile



UNIVERSIDAD DE CHILE

Tercera forma normal

- La relación entre atributos (no llave primaria) no debe ser de 1 a 1 en ninguna dirección
- Es decir, los atributos deben depender de la llave primaria, no de otros atributos
- Se resuelve creando otra entidad con los atributos interdependientes y una relación con la entidad inicial

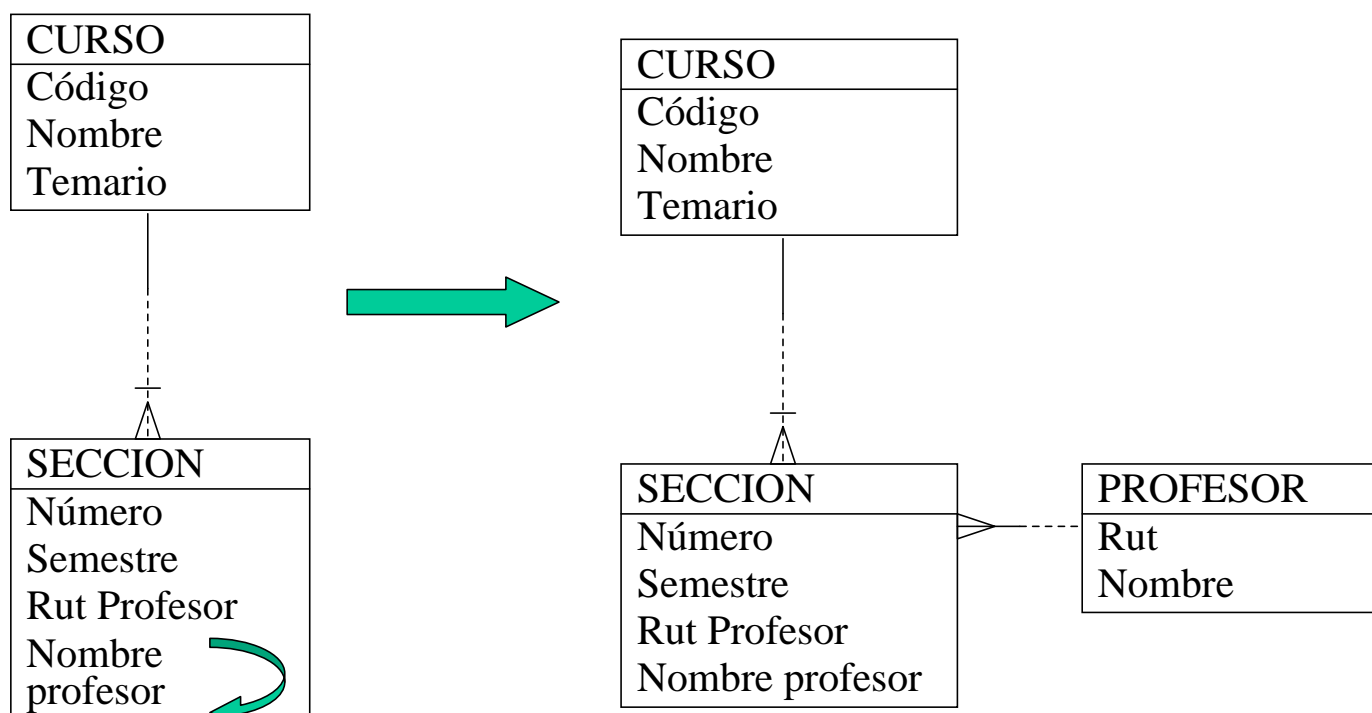
Aldo Di Biase Friedmann

Página N° 52

Diseño de Sistemas basado en TICS

Universidad de Chile

Tercera forma normal



Desnormalización

- Es la acción consiente y documentada de diseñar una base de datos no normalizada
- Esto debe ser una decisión del diseño físico, no lógico
- Se aplica para mejorar la performance
- Un ejemplo es colocar el código y la descripción en la entidad; de esta forma al recuperar la información sólo se debe acceder a una entidad
 - Esto es posible porque los códigos no cambian frecuentemente

Desnormalización usualmente utilizada

- Colocar el código y la descripción en la entidad; de esta forma al recuperar la información sólo se debe acceder a una entidad
 - Esto es posible porque los códigos no cambian frecuentemente
- Colocar el saldo de todos los movimientos, a pesar que se puede calcular a partir de todos los movimientos históricos
 - Se justifica por la alta frecuencia que se consulta el saldo (por ejemplo el saldo en la cuenta corriente bancaria)
- Totales y porcentajes que se desea almacenar de la misma forma en que fueron transmitidos (o enviados) a otras instituciones
 - Por ejemplo el total e IVA de una factura, que se pueden calcular a partir del valor de cada ítem
- Resúmenes y/o estadísticas, para mejorar la performance
 - Puede requerir un proceso periódico de cálculo o hacerlo en línea

Diseño de Sistemas basado en TICS

Base de Datos

Bases de datos relacionales y RDBMS
Introducción al Modelo Entidad Relación
Introducción al lenguaje SQL