CAPÍTULO 3

El modelo de datos relacionales y
Restricciones de bases de datos relacionales

Bosquejo del capítulo

- Conceptos del modelo relacional
- Características de una relación
- Restricciones del modelo relacional y relaciones
 Esquemas de bases de datos
- Actualizar operaciones y lidiar con restricciones
 Violaciones

Conceptos del modelo relacional

- El Modelo relacional de Datos se basa en el concepto de Relación
 - La fuerza del enfoque relacional para la gestión de datos proviene de la base formal proporcionada por la teoría de las relaciones.
- Revisamos los elementos esenciales del modelo relacional formal en este capítulo.
- En la práctica, existe un modelo estándar basado en SQL; se describe en los Capítulos 6 y 7 como un lenguaje
- Nota: Existen varias diferencias importantes entre el modelo formal y el modelo práctico, como veremos

Conceptos del modelo relacional

- Una relación es un concepto matemático basado en las ideas de conjuntos
- El modelo fue propuesto por primera vez por el Dr. EF Codd de IBM Research en 1970 en el siguiente artículo:
 - "Un modelo relacional para grandes bancos de datos compartidos"
 Comunicaciones de la ACM, junio de 1970
- El artículo anterior provocó una gran revolución en el campo de la gestión de bases de datos y le valió al Dr.
 Codd el codiciado premio ACM Turing

Definiciones informales

- Informalmente, una relación parece una tabla de valores.
- Una relación normalmente contiene un conjunto de filas.
- Los elementos de datos en cada fila representan ciertos hechos que corresponden a una entidad o relación del mundo real.
 - En el modelo formal, las filas se denominan tuplas.
- Cada columna tiene un encabezado de columna que proporciona una indicación del significado de los elementos de datos de esa columna.
 - En el modelo formal, el encabezado de la columna se denomina nombre de atributo (o simplemente atributo).

Ejemplo de una relación

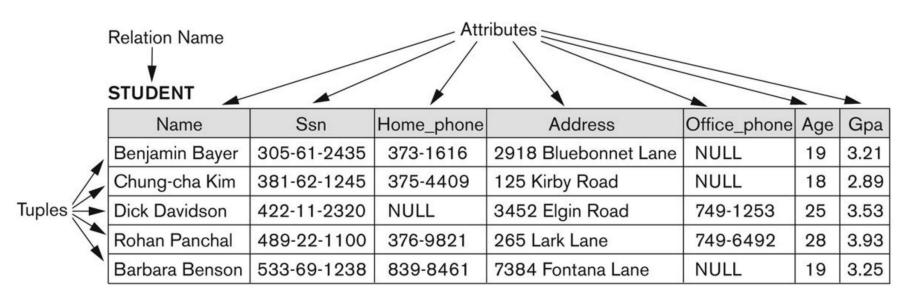


Figure 5.1

The attributes and tuples of a relation STUDENT.

Definiciones informales

- Clave de una relación:
 - Cada fila tiene un valor de un elemento de datos (o conjunto de elementos) que identifica de forma única esa fila en la tabla.
 - Llamada la clave
 - En la tabla ESTUDIANTE, SSN es la clave
 - A veces se asignan identificadores de fila o números secuenciales como claves para identificar las filas de una tabla.
 - Llamada clave artificial o clave sustituta

Definiciones formales: esquema

- El esquema (o descripción) de una relación:
 - Denotado por R(A1, A2,An)
 - R es el nombre de la relación
 - Los atributos de la relación son A1, A2, ..., An
- Ejemplo:

CLIENTE (ID del cliente, Nombre del cliente, Dirección, Número de teléfono)

- CLIENTE es el nombre de la relación
- Definido sobre los cuatro atributos: Cust-id, Cust-name, Dirección, número de teléfono
- Cada atributo tiene un dominio o un conjunto de valores válidos.
 - Por ejemplo, el dominio de Cust-id son números de 6 dígitos.

Definiciones formales: tupla

- Una tupla es un conjunto ordenado de valores (entre corchetes angulares '<... >')
- Cada valor se deriva de un dominio apropiado.
- Una fila en la relación CLIENTE es una tupla de 4 y constaría de cuatro valores, por ejemplo:
 - <632895, "John Smith", "101 Main St. Atlanta, GA 30332", "(404) 894-2000">
 - Esto se llama 4-tupla ya que tiene 4 valores
 - Una tupla (fila) en la relación CLIENTE.
- Una relación es un conjunto de tales tuplas (filas)

Definiciones formales: dominio

- Un dominio tiene una definición lógica:
 - Ejemplo: "USA_phone_numbers" son el conjunto de números de teléfono de 10 dígitos válidos en EE. UU.
- Un dominio también tiene un tipo de datos o un formato definido para él.
 - Los números_teléfonos_de_EE.UU. pueden tener el formato: (ddd)ddd-dddd donde cada d es un dígito decimal.
 - Las fechas tienen varios formatos, como año, mes y fecha. como aaaa-mm-dd, o como dd mm,aaaa, etc.
- El nombre del atributo designa el papel desempeñado en una relación:
 - Se utiliza para interpretar el significado de los elementos de datos correspondientes a ese atributo.
 - Ejemplo: El dominio Fecha se puede utilizar para definir dos atributos denominados "Fecha-factura" y "Fecha-pago" con significados diferentes.

Definiciones formales - Estado

- El estado de relación es un subconjunto del producto cartesiano de los dominios de sus atributos
 - cada dominio contiene el conjunto de todos los valores posibles que puede tomar el atributo.
- Ejemplo: el atributo Cust-name se define sobre el dominio de cadenas de caracteres de longitud máxima 25
 - dom(nombre-cliente) es varchar(25)
- El papel que desempeñan estas cadenas en la relación CLIENTE es el del nombre de un cliente.

Definiciones formales - Resumen

- Formalmente,
 - Dado R(A1, A2,, An)
 - r(R) dom (A1) X dom (A2) XX dom(An)
- R(A1, A2, ..., An) es el esquema de la relación
- R es el nombre de la relación
- A1, A2, ..., An son los atributos de la relación
- r(R): un estado específico (o "valor" o "población") de la relación R; este es un conjunto de tuplas (filas)
 - r(R) = {t1, t2, ..., tn} donde cada ti es una tupla
 - ti = <v1, v2, ..., vn> donde cada elemento vj de dom(Aj)

Definiciones formales: ejemplo

- Sea R(A1, A2) un esquema de relación:
 - Sea dom(A1) = {0,1}
 - Sea dom(A2) = {a,b,c}
- Entonces: dom(A1) X dom(A2) son todas las combinaciones posibles:

- El estado de relación r(R) dom(A1) X dom(A2)
- Por ejemplo: r(R) podría ser {<0,a> , <0,b> , <1,c> }
 - este es un posible estado (o "población" o "extensión") r de la relación R, definida sobre A1 y A2.
 - Tiene tres 2-tuplas: <0,a>, <0,b>, <1,c>

Resumen de definición

<u>Términos informales</u>	Términos formales
Mesa	Relación
Encabezado de la columna	Atributo
Todas las columnas posibles Valores	Dominio
Fila	tupla
Definición de tabla	Esquema de una relación
Tabla poblada	Estado de la relación

Ejemplo – Una relación ESTUDIANTE

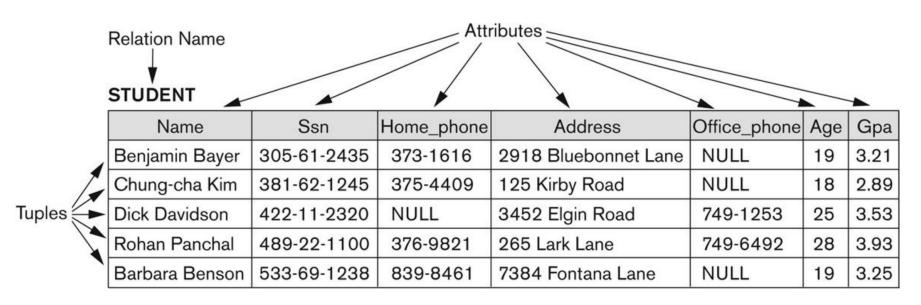


Figure 5.1

The attributes and tuples of a relation STUDENT.

Mismo estado que la figura anterior (pero con diferente orden de tuplas)

Figure 5.2

The relation STUDENT from Figure 5.1 with a different order of tuples.

STUDENT

Name	Ssn	Home_phone	Address	Office_phone	Age	Gpa
Dick Davidson	422-11-2320	NULL	3452 Elgin Road	749-1253	25	3.53
Barbara Benson	533-69-1238	839-8461	7384 Fontana Lane	NULL	19	3.25
Rohan Panchal	489-22-1100	376-9821	265 Lark Lane	749-6492	28	3.93
Chung-cha Kim	381-62-1245	375-4409	125 Kirby Road	NULL	18	2.89
Benjamin Bayer	305-61-2435	373-1616	2918 Bluebonnet Lane	NULL	19	3.21

Características de las relaciones

- Ordenación de tuplas en una relación r(R):
 - Las tuplas no se consideran ordenadas, aunque parezcan estar en forma tabular.
- Ordenación de atributos en un esquema de relación R (y de valores dentro de cada tupla):
 - Consideraremos los atributos en R(A1, A2, ..., An) y los valores en t=<v1, v2, ..., vn> para ordenar.
 - (Sin embargo, una definición alternativa más general de relación no requiere este pedido. Incluye tanto el nombre como el valor de cada uno de los atributos).
 - Ejemplo: t= { <nombre, "John" >, <SSN, 123456789> }

Características de las relaciones

- Valores en una tupla:
 - Todos los valores se consideran atómicos (indivisibles).
 - Cada valor de una tupla debe ser del dominio de el atributo para esa columna
 - Si la tupla t = <v1, v2,..., vn> es una tupla (fila) en el estado de relación r de R(A1, A2,..., An)
 - Entonces cada vi debe ser un valor de dom(Ai)
 - Se utiliza un valor nulo especial para representar valores.
 que son desconocidos o no están disponibles o no son aplicables en determinadas tuplas.

Características de las relaciones

■ Notación:

Nos referimos a los valores de los componentes de una tupla t

mediante:

t[Ai] o t.Ai ■ Este es el valor vi del atributo Ai para la tupla t

RESTRICCIONES

Las restricciones determinan qué valores están permitidos y cuáles no están en la base de datos.

Son de tres tipos principales:

- 1. Restricciones inherentes o implícitas: se basan en el propio modelo de datos. (Por ejemplo, el modelo relacional no permite una lista como valor para ningún atributo)
- 2. Restricciones basadas en esquemas o explícitas: Se expresan en el esquema utilizando las facilidades proporcionadas por el modelo. (Por ejemplo, restricción de relación de cardinalidad máxima en el modelo ER)
- 3. Restricciones semánticas o basadas en aplicaciones: están más allá del poder expresivo del modelo y deben ser especificadas y aplicadas por los programas de aplicación.

Restricciones de integridad relacional

- Las restricciones son condiciones que deben cumplirse en todos los estados de relación válidos.
- Hay tres tipos principales de (basados en esquemas explícitos) restricciones que se pueden expresar en el modelo relacional:
 - Limitaciones clave
 - Restricciones de integridad de la entidad
 - Restricciones de integridad referencial
- Otra restricción basada en esquemas es el dominio restricción
 - Cada valor en una tupla debe ser del dominio de su atributo (o podría ser nulo, si está permitido para ese atributo)

Limitaciones clave

- Superclave de R:
 - Es un conjunto de atributos SK de R con la siguiente condición:
 - No hay dos tuplas en cualquier estado de relación válido r(R) que tengan la mismo valor para SK
 - Es decir, para cualquier tupla distinta t1 y t2 en r(R), t1[SK] ≠ t2[SK]
 - Esta condición debe cumplirse en cualquier estado válido r(R)
- Clave de R:
 - Una superclave "mínima"
- Una clave es una superclave, pero no al revés

Restricciones clave (continuación)

- Ejemplo: Considere el esquema de relación CAR:
 - COCHE (estado, número de registro, número de serie, marca, modelo,
 - año) COCHE tiene dos claves:
 - Clave1 = {Estado, Reg#}
 - Clave2 = {SerialNo}
 - Ambas también son superclaves de CAR
 - {SerialNo, Make} es una superclave pero no una clave.

En general: ■

Cualquier clave es una superclave (pero no al revés)

Cualquier conjunto de atributos que incluya una clave es una superclave

Una superclave mínima también es una clave

Restricciones clave (continuación)

- Si una relación tiene varias claves candidatas, se elige una arbitrariamente para ser la clave principal.
 - Los atributos de la clave principal están subrayados.
- Ejemplo: Considere el esquema de relación CAR:
 - COCHE (estado, número de registro, número de serie, marca, modelo, año)
 - Elegimos SerialNo como clave principal
- El valor de la clave principal se utiliza para identificar de forma única cada tupla en una relación.
 - Proporciona la identidad de la tupla
- También se utiliza para hacer referencia a la tupla de otra tupla.
 - Regla general: elija como clave principal la más pequeña de las claves candidatas.

Tabla CAR con dos claves candidatas – Número de licencia elegido como clave principal

CAR

Figure 5.4 The CAR relation, with two candidate keys: License_number and Engine_serial_number.

<u>License_number</u>	Engine_serial_number	Make	Model	Year
Texas ABC-739	A69352	Ford	Mustang	02
Florida TVP-347	B43696	Oldsmobile	Cutlass	05
New York MPO-22	X83554	Oldsmobile	Delta	01
California 432-TFY	C43742	Mercedes	190-D	99
California RSK-629	Y82935	Toyota	Camry	04
Texas RSK-629	U028365	Jaguar	XJS	04

Integridad de la entidad

- Integridad de la entidad:
 - Los atributos de clave primaria PK de cada esquema de relación R en S no pueden tener valores nulos en ninguna tupla de r(R).
 - Esto se debe a que se utilizan valores de clave primaria para identificar las tuplas individuales.
 - t[PK] ≠ nulo para cualquier tupla t en r(R)
 - Si PK tiene varios atributos, no se permite nulo en ninguno de estos. atributos
 - Nota: Otros atributos de R pueden estar restringidos para no permitir valores nulos, aunque no sean miembros de la clave principal.

Integridad referencial

- Una restricción que involucra dos relaciones
 - Las restricciones anteriores implican una única relación.
- Se utiliza para especificar una relación entre tuplas en dos relaciones:
 - La relación referencial y el referenciado. relación.

Integridad referencial

- Las tuplas en la relación de referencia R1 tienen atributos FK (llamados atributos de clave externa) que hacen referencia a los atributos de clave primaria PK de la relación referenciada R2.
 - Se dice que una tupla t1 en R1 hace referencia a una tupla t2 en R2 si t1[FK] = t2[PK].
- Una restricción de integridad referencial se puede mostrar en un esquema de base de datos relacional como un arco dirigido desde R1.FK a R2.

Integridad referencial (o clave externa) Restricción

- Declaración de la restricción
 - El valor en la columna (o columnas) de clave externa FK de la relación de referencia R1 puede ser :
 - (1) un valor de un valor de clave primaria existente de una clave primaria PK correspondiente en la relación referenciada R2, o
 - **■** (2) un nulo.
- En el caso (2), la FK en R1 no debería ser parte de su propia clave primaria.

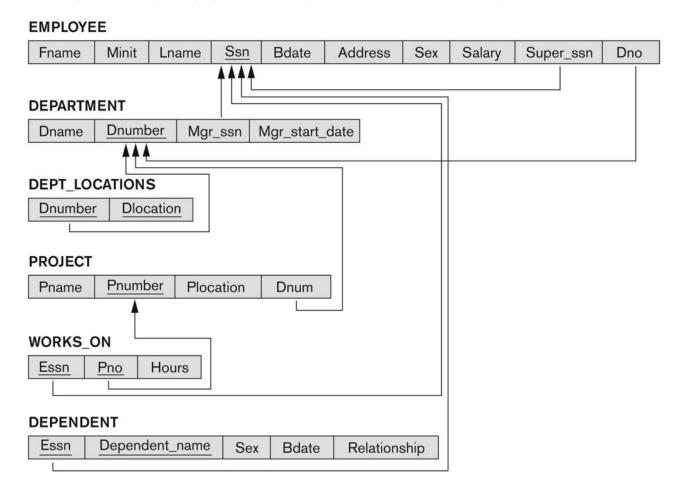
Mostrar un esquema de base de datos relacional y sus restricciones

- Cada esquema de relación se puede mostrar como una fila de nombres de atributos
- El nombre de la relación se escribe encima del atributo. nombres
- El atributo (o atributos) de la clave principal estará subrayado.
- Las restricciones de una clave externa (integridad referencial) se muestran como un arco dirigido (flecha) desde los atributos de la clave externa hasta la tabla referenciada.
 - También puede señalar la clave principal de la relación referenciada.

 para mayor claridad
- La siguiente diapositiva muestra el diagrama del esquema relacional EMPRESA con restricciones de integridad referencial.

Restricciones de integridad referencial para la base de datos de EMPRESA

Figure 5.7Referential integrity constraints displayed on the COMPANY relational database schema.



Otros tipos de restricciones

- Restricciones de integridad semántica:
 - basado en la semántica de la aplicación y no puede expresarse mediante el modelo per se
 - Ejemplo: "el máximo. No. El número de horas por empleado para todos los proyectos en los que trabaja es de 56 horas por semana".
- Es posible que sea necesario utilizar un lenguaje de especificación de restricciones para expresar estas
- SQL-99 permite CREAR DISPARADOR y CREAR ASERCIÓN para expresar algunas de estas restricciones semánticas.
- Claves, permisibilidad de valores nulos, claves candidatas (Único en SQL), claves externas, integridad referencial, etc. se expresan mediante la declaración CREATE TABLE en SQL.

Esquema de base de datos relacional

- Esquema de base de datos relacional:
 - Un conjunto S de esquemas de relación que pertenecen al misma base de datos.
 - S es el nombre de todo el esquema de la base de datos.
 - S = {R1, R2, ..., Rn} y un conjunto IC de restricciones de integridad.
 - R1, R2,..., Rn son los nombres de los esquemas de relación individuales dentro de la base de datos S
- La siguiente diapositiva muestra una base de datos de EMPRESA esquema con 6 esquemas de relación

Esquema de base de datos de EMPRESA

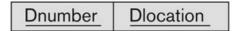
EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
							,		

DEPARTMENT

Dname	Dnumber	Mgr ssn	Mgr_start_date
		0 -	0

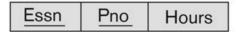
DEPT_LOCATIONS



PROJECT

Pname	Pnumber	Plocation	Dnum

WORKS_ON



DEPENDENT

Essn Dep	endent_name	Sex	Bdate	Relationship
----------	-------------	-----	-------	--------------

Figure 5.5

Schema diagram for the COMPANY relational database schema.

Estado de la base de datos relacional

- Un estado de base de datos relacional DB de S es un conjunto
- IC. El estado de una base de datos relacional a veces se denomina instantánea o instancia de una base de
- datos relacional. No utilizaremos el término instancia ya que también se aplica a tuplas individuales.
- Un estado de base de datos que no cumple con las restricciones es un estado no válido

Estado de la base de datos poblada

- Cada relación tendrá muchas tuplas en su relación actual estado
- El estado de la base de datos relacional es una unión de todos los estados de relación individuales
- Cada vez que se cambia la base de datos, surge un nuevo estado
- Operaciones básicas para cambiar la base de datos:
 - INSERTAR una nueva tupla en una relación
 - ELIMINAR una tupla existente de una relación
 - MODIFICAR un atributo de una tupla existente
- La siguiente diapositiva (Fig. 5.6) muestra un estado de ejemplo para el esquema de base de datos COMPANY que se muestra en la Fig. 5.5.

Estado de la base de datos completa para EMPRESA

Figure 5.6

One possible database state for the COMPANY relational database schema.

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	В	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	М	30000	333445555	5
Franklin	Т	Wong	333445555	1955-12-08	638 Voss, Houston, TX	М	40000	888665555	5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	М	38000	333445555	5
Joyce	Α	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	М	25000	987654321	4
James	Е	Borg	888665555	1937-11-10	450 Stone, Houston, TX	М	55000	NULL	1

DEPARTMENT

Dname	Dnumber	Mgr_ssn	Mgr_start_date
Research	5	333445555	1988-05-22
Administration	4	987654321	1995-01-01
Headquarters	1	888665555	1981-06-19

DEPT_LOCATIONS

Dnumber	Dlocation
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

WORKS_ON

Essn	Pno	Hours
123456789	1	32.5
123456789	2	7.5
666884444	3	40.0
453453453	1	20.0
453453453	2	20.0
333445555	2	10.0
333445555	3	10.0
333445555	10	10.0
333445555	20	10.0
999887777	30	30.0
999887777	10	10.0
987987987	10	35.0
987987987	30	5.0
987654321	30	20.0
987654321	20	15.0
888665555	20	NULL

PROJECT

Pname	Pnumber	Plocation	Dnum
ProductX	1	Bellaire	5
ProductY	2	Sugarland	5
ProductZ	3	Houston	5
Computerization	10	Stafford	4
Reorganization	20	Houston	1
Newbenefits	30	Stafford	4

DEPENDENT

Essn		Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	М	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	М	1942-02-28	Spouse
123456789	Michael	М	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse

Actualizar operaciones sobre relaciones

- INSERTAR una tupla.
- BORRAR una tupla.
- MODIFICAR una tupla.
- Las operaciones de actualización no deben violar las restricciones de integridad.
- Es posible que sea necesario agrupar varias operaciones de actualización.
- Las actualizaciones pueden propagarse para generar otras actualizaciones automáticamente. Esto puede ser necesario para mantener las restricciones de integridad.

Actualizar operaciones sobre relaciones

- En caso de violación de la integridad, se pueden tomar varias acciones. ser tomado:
 - Cancelar la operación que causa la infracción (opción RESTRINGIR o RECHAZAR)
 - Realice la operación pero informe al usuario del violación
 - Activar actualizaciones adicionales para que la infracción sea corregido (opción CASCADE, opción SET NULL)
 - Ejecutar una rutina de corrección de errores especificada por el usuario

Posibles violaciones para cada operación.

- INSERT puede violar cualquiera de las restricciones:
 - Restricción de dominio:
 - si uno de los valores de atributo proporcionados para la nueva tupla no es del dominio de atributo especificado
 - Restricción clave:
 - si el valor de un atributo clave en la nueva tupla ya existe en otra tupla en la relación
 - Integridad referencial:
 - si un valor de clave externa en la nueva tupla hace referencia a una clave principal valor que no existe en la relación referenciada
 - Integridad de la entidad:
 - si el valor de la clave principal es nulo en la nueva tupla

Posibles violaciones para cada operación.

- BORRAR puede violar sólo la integridad referencial:
 - Si el valor de clave principal de la tupla que se está eliminando tiene referencias de otras tuplas en la base de datos
 - Puede solucionarse mediante varias acciones: RESTRICT, CASCADE, SET NULL (consulte el Capítulo 6 para obtener más detalles)
 - Opción RESTRICT: rechazar la eliminación
 - Opción CASCADE: propaga el nuevo valor de clave principal al claves foráneas de las tuplas de referencia
 - Opción SET NULL: establece las claves externas de las tuplas de referencia a NULO
 - Se debe especificar una de las opciones anteriores durante el diseño de la base de datos para cada restricción de clave externa.

Posibles violaciones para cada operación.

- ACTUALIZAR puede violar la restricción de dominio y la restricción
 NOT NULL en un atributo que se está modificando
- También se puede infringir cualquiera de las otras restricciones, según el atributo que se actualiza:
 - Actualización de la clave principal (PK):
 - Actualización de una clave externa (FK):
 - Puede violar la integridad referencial
 - Actualización de un atributo ordinario (ni PK ni FK):
 - Sólo puede violar las restricciones del dominio

Resumen

- Conceptos del modelo relacional presentados
 - Definiciones
 - Características de las relaciones
- Se discutieron las restricciones del modelo relacional y los

Esquemas de bases de datos

- Restricciones de dominio
- Limitaciones clave
- Integridad de la entidad
- Integridad referencial
- Describió las operaciones de actualización relacional y el manejo de violaciones de restricciones.

Ejercicio en clase

(Tomado del Ejercicio 5.15)

Considere las siguientes relaciones para una base de datos que realiza un seguimiento de la inscripción de estudiantes en cursos y los libros adoptados para cada curso:

ESTUDIANTE(SSN, Nombre, Especialización, Fecha de Licenciatura)

CURSO(Curso#, Nombre, Departamento)

INSCRIBIRSE (SSN, Curso#, Trimestre, Grado)

BOOK_ADOPTION(Curso#, Trimestre, Libro_ISBN)

TEXTO(Libro_ISBN, Título_del_libro, Editorial, Autor)

Dibuje un diagrama de esquema relacional que especifique las claves externas para este esquema.