

Unidad II: Diseño de Bases de Datos y el modelo E-R.

2.1 El Proceso de Diseño

El proceso de diseño para una base de datos consta básicamente de 7 pasos, los cuáles se describen en la siguiente imagen.



El proceso de diseño consta de los pasos siguientes:

- **Determinar la finalidad de la base de datos:** Esto le ayudará a estar preparado para los demás pasos.
- **Buscar y organizar la información necesaria:** Reúna todos los tipos de información que desee registrar en la base de datos, como los nombres de productos o los números de pedidos.

2.2 Modelo Entidad-Relación

Un modelo de datos es una colección de herramientas conceptuales para la descripción de datos, relaciones entre datos, semántica de los datos y

restricciones de consistencia. En esta parte se estudiarán dos modelos de datos —el modelo entidad-relación y el modelo relacional.

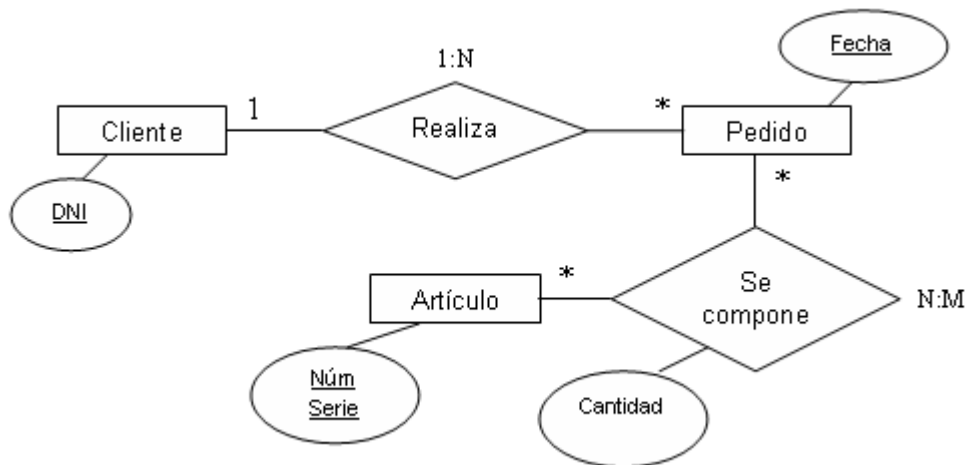
El modelo entidad-relación (E-R) es un modelo de datos de alto nivel. Está basado en una percepción de un mundo real que consiste en una colección de objetos básicos, denominados entidades, y de relaciones entre estos objetos.

El modelo relacional es un modelo de menor nivel. Usa una colección de tablas para representar tanto los datos como las relaciones entre los datos. Su simplicidad conceptual ha conducido a su adopción general; actualmente, una vasta mayoría de productos de bases de datos se basan en el modelo relacional. Los diseñadores formulan generalmente el diseño del esquema de la base de datos modelando primero los datos en alto nivel, usando el modelo E-R, y después traduciéndolo al modelo relacional.

Se estudiarán otros modelos de datos más tarde en este libro. El modelo de datos orientado a objetos, por ejemplo, extiende la representación de entidades añadiendo nociones de encapsulación, métodos (funciones) e identidad de objeto. El modelo de datos relacional orientado a objetos combina características del modelo de datos orientado a objetos y del modelo de datos relacional. Los Capítulos 8 y 9 tratan respectivamente estos dos modelos de datos.

El modelo de datos entidad-relación (E-R) está basado en una percepción del mundo real consistente en objetos básicos llamados entidades y de relaciones entre estos objetos. Se desarrolló para facilitar el diseño de bases de datos permitiendo la especificación de un esquema de la empresa que representa la estructura lógica completa de una base de datos. El modelo de datos E-R es uno de los diferentes

modelos de datos semánticos; el aspecto semántico del modelo yace en la representación del significado de los datos. El modelo E-R es extremadamente útil para hacer corresponder los significados e interacciones de las empresas del mundo real con un esquema conceptual. Debido a esta utilidad, muchas herramientas de diseño de bases de datos se basan en los conceptos del modelo E-R.



Ejemplo Modelo Entidad - Relacion

2.3 Restricciones

Las restricciones de dominio especifican que el valor de cada atributo A debe ser un valor atómico del dominio $dom(A)$ para ese atributo. Los tipos de datos asociados a los dominios por lo general incluyen los tipos de datos numéricos estándar de los números enteros (como entero- corto, entero, entero-largo) y reales (flotante y flotante de doble precisión). También disponemos de caracteres, cadenas de longitud fija y cadenas de longitud variable, así como tipos de datos de fecha, hora, marca de tiempo y dinero. Otros dominios posibles se pueden

describir mediante un intervalo de valores de un tipo de datos o como un tipo de datos enumerado en el que se listan explícitamente todos los valores posibles.

Restricción de Valores Nulos

Para determinados atributos, los valores nulos pueden ser inapropiados. Considérese una tupla en la relación cliente la que nombre-cliente es un valor vacío. Una tupla de este tipo da una calle y una ciudad para un cliente anónimo y, por tanto, no contiene información útil. En casos como éste, deseamos prohibir los valores nulos, restringiendo el dominio de ciudad-cliente para que excluya los valores nulos.

El SQL estándar permite que la declaración del dominio de un atributo incluya la especificación not null . Esto prohíbe la inserción de un valor nulo para este atributo. Cualquier modificación de la base de datos que causara que se insertase un valor nulo en un dominio not null genera un diagnóstico de error.

Hay muchas situaciones en las que la prohibición de valores nulos es deseable. Un caso particular en el que es esencial prohibir los valores nulos es en la clave primaria de un esquema de relación

Restricción de clave

Es una de las restricciones estándar que con frecuencia aparecen en las aplicaciones de bases de datos. Estas restricciones se manejan de formas ligeramente distintas en los diversos modelos de datos. En el modelo E-R, una clave es un atributo de un tipo de entidades que debe tener un valor único para cada entidad que pertenezca a dicho tipo en cualquier momento específico. Así el valor del atributo clave puede servir para identificar de manera única cada entidad.

Los atributos claves deben ser monovaluados, pero pueden ser simples o compuestos.

Un tipo de entidades normal puede tener una o más claves; un tipo de entidades débil no tiene clave, pero casi siempre tiene una clave parcial cuyos valores identifican de manera única las entidades débiles que están relacionadas a la misma entidad propietario a través de un vínculo identificador.

En general, un esquema de relación puede tener más de una clave. En tal caso, cada una de ellas se denominan clave candidata. Por ejemplo en una relación COCHE tiene dos claves candidatas: NumMatrícula y NumSerieMotor. Es común designar a una de las claves candidata como clave primaria de la relación. Ésta es la clave candidata cuyos valores sirven para identificar las tuplas en la relación.

Restricción de aserción

Una Técnica más formal para representar restricciones explícitas es con un lenguaje de especificación de restricciones , que suele basarse en alguna variación del cálculo relacional. Este enfoque declarativo establece una separación clara entre la base de restricciones (en la que las restricciones se almacenan en una forma codificada apropiada) y el subsistema de control de integridad del SGBD (que tiene acceso a la base de restricciones para aplicar estas últimas correctamente a las transacciones afectadas).

Cuando se usa esta técnica, las restricciones suelen llamarse aserciones . Se ha sugerido el uso de esta estrategia con SGBD relacionales. El subsistema de control de integridad compila las aserciones, que entonces se almacenan en el catalogo del SGBD, donde el subsistema de control de integridad puede consultarlas e imponerlas automáticamente. Esta estrategia es muy atractiva desde el punto de vista de los usuarios y programadores por su flexibilidad.

Restricción de Integridad:

Una fuente de restricciones de integridad son los conjuntos de entidades débiles. El esquema de relaciones para un conjunto de entidades débil debe incluir la clave esquema de relaciones de entidades de la cual depende. Así, pues, el esquema de relaciones para cada conjunto de entidades débil incluye una clave exterior que conduce a una restricción de integridad referencial.

2.4 Diagramas E-R

Denominado por sus siglas como: E-R; Este modelo representa a la realidad a través de un Esquema gráfico empleando los terminología de Entidades, que son objetos que existen y son los elementos principales que se identifican en el problema a resolver con el diagramado y se distinguen de otros por sus características particulares denominadas Atributos, el enlace que rige la unión de las entidades esta representada por la relación del modelo.

En un DER, cada entidad se representa mediante un rectángulo, cada relación mediante un rombo y cada dominio (conjunto donde toma valores el atributo) mediante un círculo. Mediante líneas se conectan las entidades con las relaciones, igual que las entidades con los dominios, representando a los atributos. Los Atributos Llaves se representan subrayando el correspondiente conjunto de valores.

En ocasiones, una entidad no puede ser identificada únicamente por el valor de sus propios atributos. En estos casos, se utilizan conjuntamente las relaciones con los atributos para lograr la requerida identificación unívoca. Estas entidades reciben el nombre de entidades débiles y se representan en el DER con un doble rectángulo. El MER restringe las relaciones a usar para identificar las entidades

débiles a relaciones binarias del tipo 1: N. Así, por ejemplo, una ocurrencia de "trabajador" puede tener N ocurrencias "persona-dependiente" asociadas, donde además, la existencia de las ocurrencias en la segunda entidad depende de la existencia de una ocurrencia que le corresponda en la primera entidad. Por ejemplo, en el modelo habrá personas dependientes de un trabajador sólo si ese trabajador existe. Para indicar esa dependencia en la existencia se usa una saeta en el DER. La llave de una entidad débil se forma combinando la llave de la entidad regular que la determina con algún otro atributo que defina unívocamente cada entidad débil asociada a una entidad regular dada. (Una entidad se denomina regular si no es débil).

En una relación, la llave es la combinación de las llaves de todas las entidades asociadas. Para cada relación se determina su tipo (simple o complejo) y en el DER se escribe el tipo de correspondencia. Por ejemplo, una empresa puede tener varios (n) trabajadores asociados y un trabajador pertenece a una sola empresa (1). En la relación Trabajador-Máquina-Pieza, un trabajador puede trabajar en n máquinas, produciendo p piezas, o una pieza puede ser producida por m trabajadores en n máquinas. Aquí, m, n y p no identifican un número específico, sino solamente el tipo de correspondencia que se establece en la relación.

2.5 Diseño con diagramas E-R.

Es la representación gráfica del Modelo Entidad-Relación y permite ilustrar la estructura de la base de datos del negocio modelado.

Escribe Johnson "los diagramas ER constituyen una notación para documentar un diseño tentativo de bases de datos. Los analistas los utilizan para facilitar el proceso de diseño".

Componentes y Diagrama E-R

Entidad Fuerte: Una Entidad fuerte (también conocida como entidad regular) es aquella que sí puede ser identificada unívocamente. En los casos en que se requiera, se puede dar que una entidad fuerte "preste" algunos de sus Atributos a una entidad débil para que, esta última, se pueda identificar.

Entidad débil: Es aquella que no puede existir sin participar en la relación, es decir, aquella que no puede ser unívocamente identificada solamente por sus atributos como Clave.

Conjunto de entidades Débiles. Es aquel conjunto de entidades que no tiene atributos que puedan identificar una entidad en forma única, o sea que no poseen atributos para conformar la llave primaria; por lo tanto dependen de una entidad fuerte.

Conjunto de entidades Fuerte. Conjunto de entidades que posee una clave primaria.

Relaciones: Una entidad se relaciona con otra entidad. Toda relación debe de llevar una cardinalidad. Una relación entre dos entidades siempre se va a dar por medio de un rombo. Cada entidad deberá tener sus elementos.

Atributos: Características o propiedades asociadas al conjunto de entidades o relaciones y que toman valor en una entidad en particular.

Los posibles valores puede tomar un atributo para un conjunto de entidades se denomina dominio.

2.6 Conjunto de entidades débiles.

ENTIDADES DÉBILES.

Un conjunto de entidades débiles es aquel que no tiene suficientes atributos para formar una clave primaria. Un conjunto que sí tiene una clave primaria se denomina conjunto de entidades fuertes.

Cada conjunto de entidades débiles debe estar asociada con un conjunto de entidades llamado conjunto de entidades identificadoras o propietarias.

Así, el conjunto de entidades débiles depende existencialmente del conjunto de entidades identificadoras. La relación que asocia el conjunto de entidades débiles con el conjunto de entidades identificadoras se denomina relación identificadora. La relación identificadora es varios a uno del conjunto de entidades débiles al conjunto de entidades identificadoras y la participación del conjunto de entidades débiles en la relación es total.

Aunque un conjunto de entidades débiles no tiene clave primaria, deben hacerse distinguir todas aquellas entidades del conjunto de entidades que dependen de una entidad fuerte particular. El discriminante de un conjunto de entidades débiles es un conjunto de atributos que permiten esta distinción.

La clave primaria de un conjunto de entidades débiles se forma con la clave primaria del conjunto de entidades identificadoras, más el discriminante del conjunto de entidades débiles.

El Discriminante: Es un conjunto de atributos que permite que esta distinción se haga. Por ejemplo; el discriminante del conjunto de entidades débiles pago es el atributo número-pago, ya que, para cada préstamo, un número de pago identifica de forma única cada pago para ese préstamo. El discriminante de un conjunto de entidades débiles se denomina la clave parcial del conjunto de entidades.

La clave primaria de un conjunto de entidades débiles se forma con la clave primaria del conjunto de entidades identificadoras, más el discriminante del conjunto de entidades débiles. Un conjunto de entidades débiles puede participar en relaciones distintas de relaciones identificadoras.

En algunos casos, el diseñador de la base de datos puede elegir expresar un conjunto de entidades débiles como un atributo compuesto multivalorado del conjunto de entidades propietarias. Un conjunto de entidades débiles se puede modelar más adecuadamente como un atributo si sólo participa en la relación identificadora y si tiene pocos atributos.

En pocas palabras podemos decir que:

CONJUNTO DE ENTIDADES DÉBILES: Es aquel conjunto de entidades que no tiene atributos que puedan identificar una entidad en forma única, o sea que no poseen atributos para conformar la llave primaria; por lo tanto dependen de una entidad fuerte.

Un conjunto de entidades débiles se indica en los diagramas E-R mediante un rectángulo dibujado con una línea doble y la correspondiente relación de identificación mediante un rombo dibujado con línea doble.

2.7 Modelo E-R extendido

Aunque los conceptos básicos de E-R pueden modelar la mayoría de las características de las bases de datos, algunos aspectos de una base de datos pueden ser más adecuadamente expresados mediante ciertas extensiones del modelo E-R básico. En este apartado se discuten las características E-R extendidas de especialización, generalización, conjuntos de entidades de nivel más alto y más bajo, herencia de atributos y agregación.

2.8 Otros aspectos del diseño de bases de datos

Dominio:

A veces es conveniente añadir información sobre el dominio de un atributo, los dominios se representan mediante hexágonos, con la descripción del dominio en su interior.

Diagrama:

Un diagrama E-R consiste en representar mediante estas figuras un modelo completo del problema, proceso o realidad a describir, de forma que se definan tanto las entidades que lo componen, como las interrelaciones que existen entre ellas. La idea es simple, aparentemente, pero a la hora de construir modelos sobre realidad es concreta cuando surgen los problemas. La realidad es siempre compleja. Las entidades tienen muchos atributos diferentes, de los cuales

debemos aprender a elegir sólo los que necesitamos. Lo mismo cabe decir de las interrelaciones.

Interrelación:

es la asociación o conexión entre conjuntos de entidades. Tengamos los dos conjuntos: de personas y de vehículos.

Grado:

número de conjuntos de entidades que intervienen en una interrelación. De este modo, en la anterior interrelación intervienen dos entidades, por lo que diremos que es de grado 2 o binaria. También existen interrelaciones de grado 3 pero las más frecuentes son las interrelaciones binarias. Podemos establecer una interrelación ternaria (de grado tres). Existen además tres tipos distintos de interrelaciones binarias, dependiendo del número de entidades del primer conjunto de entidades y del segundo. Así hablaremos de interrelaciones 1:1 (uno a uno), 1:N (uno a muchos) y N:M (muchos a muchos).

Clave:

es un conjunto de atributos que identifican de forma unívoca una entidad. Es muy importante poder identificar claramente cada entidad y cada interrelación. Esto es necesario para poder referirnos a cada elemento de un conjunto de entidades o interrelaciones, ya sea para consultarlo, modificarlo o borrarlo. No deben existir ambigüedades en ese sentido.

Claves candidatas:

Una característica que debemos buscar siempre en las claves es que contengan el número mínimo de atributos, siempre que mantengan su función. Diremos que una clave es mínima cuando si se elimina cualquiera de los atributos que la componen, deja de ser clave. Si en una entidad existe más de una de estas claves mínimas, cada una de ellas es una

clave candidata

Claves de interrelaciones:

Para identificar interrelaciones el proceso es similar, aunque más simple. Tengamos en cuenta que para definir una interrelación usaremos las claves primarias de las entidades interrelacionadas. De este modo, el identificador de una interrelación es el conjunto de las claves primarias de cada una de las entidades interrelacionadas.

Superclave:

Es un subconjunto de atributos que permite distinguir unívocamente cada una de las entidades de un conjunto de entidades. Si se añade un atributo al anterior subconjunto, el resultado seguirá siendo una super clave.

Clave primaria (Llave Primaria):

Es la clave candidata escogida por el diseñador. Atributo o conjunto de atributos que permiten identificar en forma única una tupla en la tabla (una entidad en un conjunto de entidades) y ningún subconjunto de ella posee esta propiedad.

Llave foránea:

Es un atributo que es llave primaria en otra entidad con la cual se relaciona. Las llaves foráneas son en últimas las que permiten relacionar las tablas en las bases de datos.

2.9 La Notación E-R con UML

El Lenguaje Unificado de Modelado preescribe un conjunto de notaciones y diagramas estándar para modelar sistemas orientados a objetos, y describe la semántica esencial de lo que estos diagramas y símbolos significan. Mientras que ha habido muchas notaciones y métodos usados para el diseño orientado a objetos, ahora los modeladores sólo tienen que aprender una única notación.

UML se puede usar para modelar distintos tipos de sistemas: sistemas de software, sistemas de hardware, y organizaciones del mundo real. UML ofrece nueve diagramas en los cuales modelar sistemas.

- Diagramas de Casos de Uso para modelar los procesos 'business'.
- Diagramas de Secuencia para modelar el paso de mensajes entre objetos.
- Diagramas de Colaboración para modelar interacciones entre objetos.
- Diagramas de Estado para modelar el comportamiento de los objetos en el sistema.

- Diagramas de Actividad para modelar el comportamiento de los Casos de Uso, objetos u operaciones.
- Diagramas de Clases para modelar la estructura estática de las clases en el sistema.
- Diagramas de Objetos para modelar la estructura estática de los objetos en el sistema.
- Diagramas de Componentes para modelar componentes.
- Diagramas de Implementación para modelar la distribución del sistema.