

# La entidad-relación

## Modelo-Hacia

a

## Vista unificada de datos

PETER PIN-SHAN CHEN  
Instituto de Tecnología de Massachusetts

Se propone un modelo de datos, llamado modelo entidad-relación. Este modelo incorpora algunos de la información semántica importante sobre el mundo real. Una técnica de diagramación especial es introducido como una herramienta para el diseño de bases de datos. Un ejemplo de diseño y descripción de base de datos usando se da el modelo y la técnica diagramática. Algunas implicaciones para la integridad de los datos, Se discuten la recuperación de información y la manipulación de datos. El modelo entidad-relación se puede utilizar como base para la unificación de diferentes vistas de datos: 1, el modelo de red, el modelo relacional y el modelo de conjunto de entidades. Ambigüedades semánticas en estos se analizan los modelos. Posibles formas de derivar sus puntos de vista de los datos a partir de la relación entidad se presentan el modelo.

Palabras y frases clave: diseño de bases de datos, vista lógica de datos, semántica de datos, modelos de datos, modelo entidad-relación, modelo relacional, grupo de tareas de base de datos, modelo de red, conjunto de entidades modelo, definición y manipulación de datos, integridad y consistencia de datos Categorías de CR: 3.50, 3.70, 4.33, 4.34

### 1. INTRODUCCIÓN

La visión lógica de los datos ha sido un tema importante en los últimos años. Tres principales Se han propuesto modelos de datos: el modelo de red [2, 3, 71, el modelo relacional [S] y el modelo de conjunto de entidades [25]. Estos modelos tienen sus propias fortalezas y debilidades. El modelo de red proporciona una vista más natural de los datos al separar entidades y relaciones (hasta cierto punto), pero su capacidad para obtener datos la independencia ha sido cuestionada [S]. El modelo relacional se basa en relacional teoría y puede lograr un alto grado de independencia de datos, pero puede perder alguna información semántica importante sobre el mundo real [12, 15, 231. El conjunto de entidades El modelo, que se basa en la teoría de conjuntos, también logra un alto grado de independencia de datos. Dependencia, pero su visualización de valores como "3" o "rojo" puede no ser natural para algunas personas [25].

Este artículo presenta el modelo entidad-relación, que tiene la mayoría de las ventajas de los tres modelos anteriores. El modelo entidad-relación adopta el visión más natural de que el mundo real se compone de entidades y relaciones. Eso Copyright © 1976, Association for Computing Machinery, Inc. Permiso general para volver a publicar, pero no con fines de lucro; todo o parte de este material se otorga siempre que el aviso de derechos de autor de ACM sea dado y que se hace referencia a la publicación, a su fecha de emisión y al hecho de que los privilegios de reimpresión fueron otorgados con permiso de la Asociación de Maquinaria de Computación.

Una versión de este documento se presentó en la Conferencia Internacional sobre Bases de Datos Muy Grandes, Framingham, Mass., 22-24 de septiembre de 1975. Dirección del autor: Center for Information System Research, Alfred P. Sloan School of Management, Instituto de Tecnología de Massachusetts, Cambridge, MA 02139. incorpora parte de la información semántica importante sobre el mundo real (se pueden encontrar otros trabajos sobre semántica de bases de datos en [1, 12, 15, 21, 23 y 29)). El modelo puede lograr un alto grado de independencia de datos y se basa en un conjunto teoría y teoría de las relaciones, El modelo entidad-relación se puede utilizar como base para una vista unificada de datos. La mayoría de Ivork en el pasado ha enfatizado la diferencia entre el modelo de red y el modelo relacional [22]. Recientemente, se han hecho varios intentos para reducir las diferencias de los tres modelos de datos [4, 19, 26, 30, 311. Este documento utiliza el modelo entidad-relación como marco a partir del cual los tres datos existentes se pueden derivar modelos. El lector puede ver el modelo entidad-relación como un generalización o extensión de modelos existentes. Este documento está organizado en tres partes (Secciones 2-4). La sección 2 presenta la entidad-relación modelo utilizando un marco de vistas de datos de varios niveles. La sección 3 describe la información semántica en el modelo y sus implicaciones para descripción y manipulación de datos. Una técnica de diagramación especial, la diagrama entidad-relación, se presenta como una herramienta para el diseño de bases de datos. Sección 4 analiza el modelo de red, el modelo relacional y el modelo de conjunto de entidades, y describe cómo pueden derivarse del modelo entidad-relación.

### 2. LA ENTIDAD-RELACIÓN

#### MODELO

##### 2.1 Vistas multinivel de Doto

En el estudio de un modelo de datos, debemos identificar los niveles de vistas lógicas de los datos. de lo que se trata el modelo. Ampliación del marco desarrollado en [IS, 251, podemos identificar cuatro niveles de vistas de datos (Figura 1):

- (1) Información sobre entidades y relaciones que existen en nuestras mentes.
- (2) Estructura de información, organización de urea de información en la que entidades y las relaciones están representadas por datos.
- (3) Estructura de datos independiente de la ruta de acceso: las estructuras de datos que no son involucrado con esquemas de búsqueda, esquemas de indexación, etc.
- (4) Estructura estándar de datos dependientes de la ruta de acceso.

En las siguientes secciones, desarrollaremos el modelo entidad-relación paso a paso. paso para el primero, dos niveles. Como veremos más adelante en el artículo, el modelo de red, tal como se aplica actualmente, se refiere

principalmente al nivel 4; el modelo relacional se ocupa principalmente de los niveles 3 y 2; el modelo de conjunto de entidades se refiere principalmente con niveles 1 y 2.

## 2.2 Información relativa a entidades y relaciones (Nivel 1)

En este nivel consideramos entidades y relaciones. Una entidad es una "cosa" que se puede identificar claramente. Una persona, empresa o evento específico es un ejemplo de una entidad. Una relación es una asociación entre entidades. Por ejemplo, "padre-hijo" es una relación entre dos "entidades" CLperson.

Itt. Es posible que algunas personas vean algo (por ejemplo, el matrimonio) como una entidad mientras que otras la gente puede verlo como una relación. Creemos que esta es una decisión que debe tomar el administrador de la empresa [27]. Debe definir qué son entidades y qué son relaciones para que la distinción sea adecuada a su entorno.

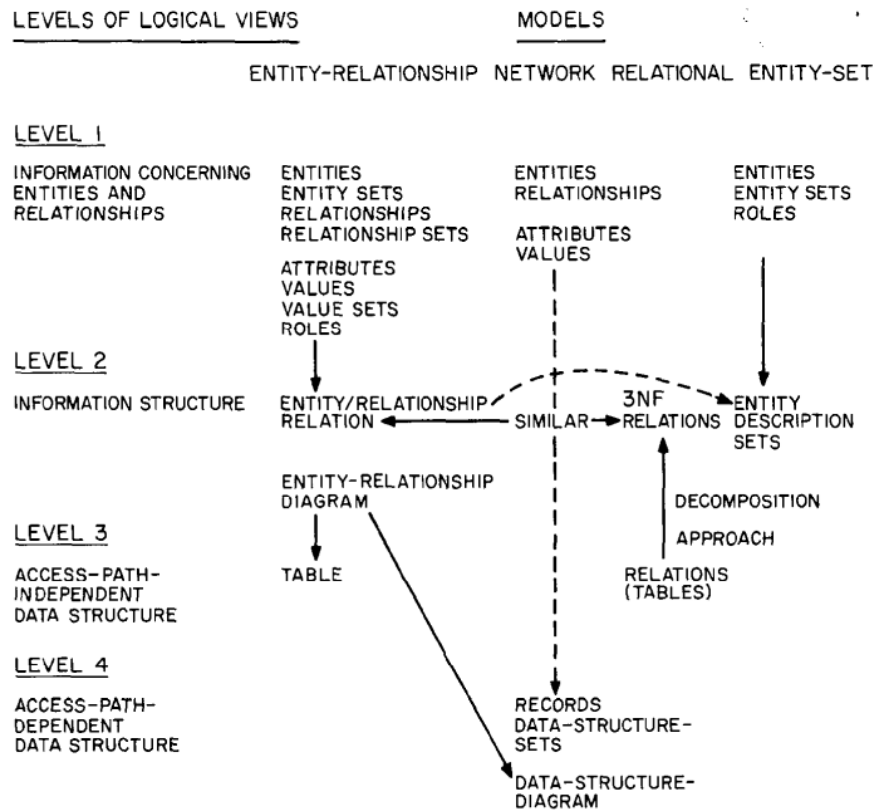


Fig. 1. Análisis de modelos de datos utilizando múltiples niveles de vistas lógicas

La base de datos de una empresa contiene información relevante sobre entidades y relaciones en las que está interesada la empresa. Una descripción completa de una entidad o relación no puede registrarse en la base de datos de una empresa. Es imposible (y, tal vez, innecesario) registrar toda la información sobre entidades y relaciones. De ahora en adelante, lo haremos considerar solo las entidades y relaciones (y la información que les concierne) que van a entrar en el diseño de una base de datos.

### 2.2.1 Entidad y conjunto de entidades. Denotemos una entidad que existe en nuestras mentes.

Las entidades se clasifican en diferentes conjuntos de entidades como EMPLEADO, PROYECTO, y DEPARTAMENTO. Hay un predicado asociado con cada conjunto de entidades para probar si una entidad le pertenece, por ejemplo, si sabemos que una entidad está en la entidad establece EMPLEADO, entonces sabemos que tiene las propiedades comunes a la otra entidades en el conjunto de entidades EMPLEADO. Entre estas propiedades se encuentra la anterior- predicado de prueba mencionado. Sea  $R_i$  los conjuntos de entidades. Tenga en cuenta que los conjuntos de entidades pueden no ser mutuamente inconexos. Por ejemplo, una entidad que pertenece al conjunto de entidades MALE-PERSON también pertenece al conjunto de entidades PERSON. En este caso, PERSONA MASCULINA es un subconjunto de PERSON.

### 2.2.2 Relación, rol y conjunto de relaciones. Considere las asociaciones entre entidades. Un conjunto de relaciones, $R_i$ , es una relación matemática [5] entre $n$ entidades, cada uno tomado de un conjunto de entidades:

$(Gel, e_2, \dots, E_n) \mid 1 \leq i \leq E_1, e_2 \in E_2, \dots, E_n \in E_n$ ,

y cada tupla de entidades,  $[e_1, e_2, \dots, e_n]$ , es una relación. Tenga en cuenta que el  $E_i$  en el La definición anterior puede no ser distinta. Por ejemplo, un "matrimonio" es una relación entre dos entidades en el conjunto de entidades PERSON.

El papel de una entidad en una relación es la función que realiza en la relación. LLHusband "y" esposa "son roles. El ordenamiento de entidades en el La definición de relación (tenga en cuenta que se utilizaron

corchetes) se puede eliminar si Los roles de las entidades en la relación se establecen explícitamente de la siguiente manera:  $(r_1 / e_1, r_2 / e_2, \dots, r_n / e_n)$ , donde  $r_i$  es el papel de  $e_i$  en la relación.

2.2.3 Atributo, valor y conjunto de valores. La información sobre una entidad o una La relación se obtiene por observación o medición, y se expresa mediante un conjunto de pares atributo-valor. "3", "rojo", "Peter" y "Johnson" son valores. Valores se clasifican en diferentes conjuntos de valores, como PIES, COLOR, PRIMER NOMBRE,

Y apellido. Hay un predicado asociado con cada conjunto de valores para probar si un valor le pertenece. Un valor en un conjunto de valores puede ser equivalente a otro valor en un conjunto de valores diferente. Por ejemplo, "12" en el valor establecido en PULGADAS es equivalente a "1" en el valor establecido PIES.

Un atributo se puede definir formalmente como una función que se asigna a una entidad conjunto o un conjunto de relaciones en un conjunto de valores o un producto cartesiano de conjuntos de valores:

f:  $E_i$  o  $R_i + V_i$  o  $V_i \times V_i$ ,  $X \in V_i$ ,  $X \rightarrow V_i$ .

La Figura 2 ilustra algunos atributos definidos en el conjunto de entidades PERSON. El atributo EDAD se asigna al conjunto de valores NO AÑOS. Un atributo puede mapear en un cartesiano producto de conjuntos de valores. Por ejemplo, el atributo NAME se asigna a conjuntos de valores Nombre y apellido. Tenga en cuenta que más de un atributo puede mapear del mismo conjunto de entidades en el mismo conjunto de valores (o el mismo grupo de conjuntos de valores). Por ejemplo, NAME y ALTERNATIVE-NAME mapa del conjunto de entidades EMPLEADO en valor establece FIRST-NAME y LAST-NAME. Por lo tanto, el atributo bute y value set son conceptos diferentes, aunque pueden tener el mismo nombre en algunos casos (por ejemplo, EMPLOYEE-NO mapea de EMPLOYEE a value establecer EMPLEADO-NO). Esta distinción no es clara en el modelo de red y en muchos sistemas de gestión de datos existentes. También tenga en cuenta que un atributo se define como Una función. Por lo tanto, asigna una entidad dada a un solo valor (o una sola tupla de valores en el caso de un producto cartesiano de conjuntos de valores).

Tenga en cuenta que las relaciones también tienen atributos. Considere el conjunto de relaciones PROYECTO-TRABAJADOR (Figura 3). El atributo PORCENTAJE DE TIEMPO, que es la porción de tiempo que un empleado en particular está comprometido con un proyecto, es un atributo definido en el conjunto de relaciones PROYECTO-TRABAJADOR. Eso no es un atributo del EMPLEADO ni un atributo del PROYECTO, ya que su el significado depende tanto del empleado como del proyecto involucrado. El concepto de El atributo de relación es importante para comprender la semántica de los datos y en la determinación de las dependencias funcionales entre los datos.

2.2.4 Estructura de información conceptual. Ahora nos preocupa cómo organizar la información asociada a entidades y relaciones. El método propuesto en este documento es separar la información sobre las entidades de la información

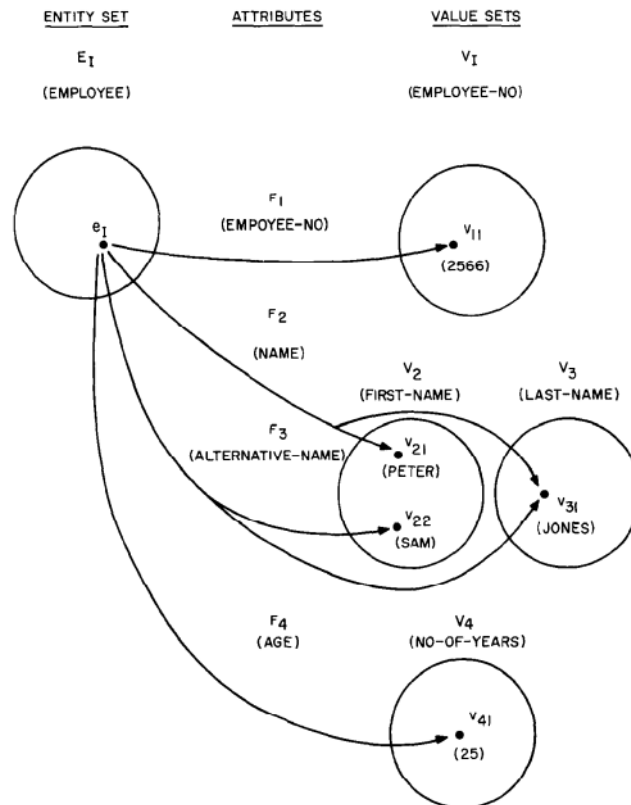


Fig. 2. Atributos definidos en el conjunto de entidades PERSONA

información sobre las relaciones. Veremos que esta separación es útil para identificar dependencias funcionales entre datos. La Figura 4 ilustra en forma de tabla la información sobre las entidades en un conjunto de entidades. Cada fila de valores está relacionada con la misma entidad y cada columna está relacionada con una conjunto de valores que, a su vez, está relacionado con un atributo. El orden de filas y columnas es insignificante. La Figura 5 ilustra información sobre las relaciones en un conjunto de relaciones. Nota que cada fila de valores está relacionada con una relación indicada por un grupo de entidades, cada una con un rol específico y perteneciente a un conjunto de entidades específico. Tenga en cuenta que las Figuras 4 y 2 (y también las Figuras 5 y 3) son formas diferentes de misma información. La forma de tabla se utiliza para relacionarse fácilmente con el modelo relacional.

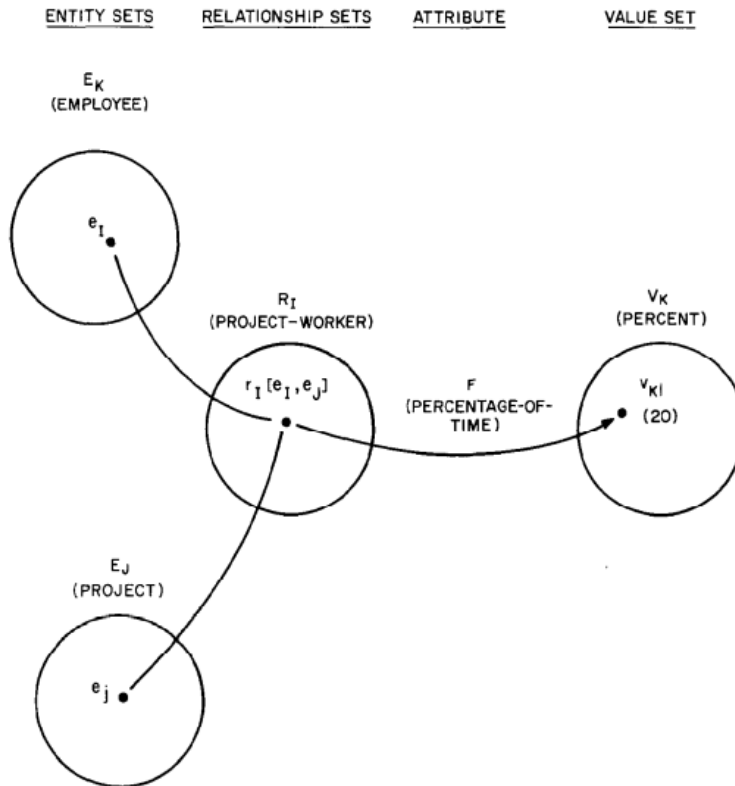


Fig. 3. Atributos definidos en el conjunto de relaciones PROYECTO-TRABAJADOR

### 2.3 Estructura de la información (nivel 2)

Las entidades, relaciones y valores en el nivel 1 (ver Figuras 2-5) son conceptuales. objetos en nuestras mentes (es decir, estábamos en el ámbito conceptual [IS, 27-J. En el nivel 2, consideramos representaciones de objetos conceptuales. Asumimos que existen representaciones directas de valores. A continuación, describiremos cómo representar entidades y relaciones.

2.3.1 Clave principal. En la Figura 2 los valores del atributo EMPLEADO-NO pueden ser utilizado para identificar entidades en el conjunto de entidades EMPLEADO si cada empleado tiene un número de empleado diferente. Es posible que se necesite más de un atributo para identificar las entidades en un conjunto de entidades. También es posible que varios grupos de los atributos pueden usarse para identificar entidades. Básicamente, una clave de entidad es un grupo de atributos tales que el mapeo del conjunto de entidades al grupo correspondiente de conjuntos de valores es uno a uno. Si no podemos encontrar tal mapeo uno a uno en disponibles datos, o si se desea simplicidad en la identificación de entidades, podemos definir un atributo y un valor establecido para que tal mapeo sea posible. En el caso donde

La entidad-relación

Modelo

(23)

ATTRIBUTE ENTITY SET AND VALUE SET	$F_1$	$F_2$	$F_3$	$F_4$
	(EMPLOYEE-NO)	(NAME)	(ALTERNATIVE-NAME)	(AGE)
$E_1$	$V_1$	$V_2$	$V_3$	$V_4$
(EMPLOYEE)	(EMPLOYEE-NO)	(FIRST-NAME)	(LAST-NAME)	(NO-OF-YEARS)
$e_1$	$v_{11}$ (2566)	$v_{21}$ (PETER)	$v_{31}$ (JONES)	$v_{41}$ (25)
$e_2$	$v_{12}$ (3378)	$v_{23}$ (MARY)	$v_{32}$ (CHEN)	$v_{42}$ (23)
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Fig. 4. Información sobre entidades en un conjunto de entidades (forma de tabla)

ROLE	WORKER	PROJECT	
			$F$ (PERCENTAGE-OF-TIME)
ENTITY SET	$E_I$ (EMPLOYEE)	$E_J$ (PROJECT)	$V_K$ (PERCENTAGE)
	$e_{11}$	$e_{j1}$	$v_{k1}$ (20)
	$\vdots$	$\vdots$	$\vdots$

Fig. 5. Información sobre las relaciones en un conjunto de relaciones (forma de tabla)

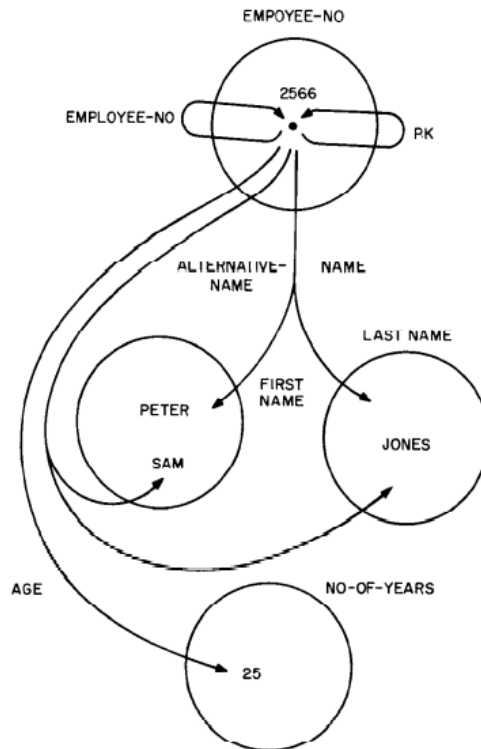


Fig. 6. Representación de entidades por valores (números de empleados)

Existen varias claves, normalmente elegimos una clave semánticamente significativa como entidad clave primaria (PK). La figura 6 se obtiene fusionando el conjunto de entidades EMPLEADO con el conjunto de valores NUMERO DE EMPLEADO en la Figura 2. Deberíamos notar algunas implicaciones semánticas de Figura 6. Cada valor en el conjunto de valores EMPLEADO-NO representa una entidad (empleado). Mapa de atributos del conjunto de valores EMPLEADO-NO a otro valor conjuntos. También tenga en cuenta que el atributo EMPLEADO-NO mapas del conjunto de valores NUMERO DE EMPLEADO a sí mismo.

### 2.3.2 Entidad / Relación

Relaciones. Información sobre entidades en una entidad El conjunto ahora se puede organizar en la forma que se muestra en la Figura 7. Tenga en cuenta que la Figura 7 es similar a la Figura 4 excepto que las entidades están representadas por los valores de su teclas. Toda la tabla en la Figura 7 es una relación de entidad, y cada fila es una entidad tupla. Dado que una relación es identificada por las entidades involucradas, la clave primaria de una La relación puede ser representada por las claves primarias de las entidades involucradas. En

ATTRIBUTE	PRIMARY KEY		ALTERNATIVE-NAME		AGE
	EMPLOYEE-NO	NAME	FIRST-NAME	LAST-NAME	
VALUE SET (DOMAIN)	EMPLOYEE-NO	FIRST-NAME	LAST-NAME	FIRST-NAME	LAST-NAME
ENTITY (TUPLE)	2566	PETER	JONES	SAM	JONES
	3378	MARY	CHEN	BARB	CHEN
	⋮	⋮	⋮	⋮	⋮

Fig. 7. Relación entidad regular EMPLEADO

Figura 8, las entidades involucradas están representadas por sus claves primarias EMPLEADO- NO y PROYECTO- NO. Los nombres de los roles proporcionan el significado semántico de los valores en las columnas correspondientes. Tenga en cuenta que EMPLEADO-NO es el primario clave para las entidades involucradas en la relación y no es un atributo de la relación. PORCENTAJE DE TIEMPO es un atributo de la relación. La tabla de la Figura 8 es una relación de relación, y cada fila de valores es una relación tupla.

En ciertos casos, las entidades en un conjunto de entidades no pueden ser identificadas de forma única por los valores de sus propios atributos; por lo tanto, debemos usar una (s) relación (es) para identificar ellos. Por ejemplo, considere a los dependientes de los empleados: se identifican los dependientes por sus nombres y por los valores de la clave principal de los empleados que apoyan ellos (es decir, por sus relaciones con los empleados). Tenga en cuenta que en la Figura 9,

ENTITY RELATION NAME	PRIMARY KEY			RELATIONSHIP ATTRIBUTE
	EMPLOYEE	PROJECT		
	WORKER	PROJECT		
	EMPLOYEE-NO	PROJECT-NO	PERCENTAGE-OF-TIME	
	EMPLOYEE-NO	PROJECT-NO	PERCENTAGE	
VALUE SET (DOMAIN)				
RELATIONSHIP TUPLE	2566	31	20	
	2173	25	100	
	⋮	⋮	⋮	
	⋮	⋮	⋮	

Fig. 8. Relación regular Relación PROYECTO-TRABAJADOR

ENTITY RELATION NAME	PRIMARY KEY			RELATIONSHIP ATTRIBUTE
	EMPLOYEE			
	SUPPORTER			
	EMPLOYEE-NO	NAME	AGE	
	EMPLOYEE-NO	FIRST-NAME	NO-OF-YEARS	
VALUE SET (DOMAIN)				
ENTITY TUPLE	2566	VICTOR	3	
	2173	GEORGE	6	
	⋮	⋮	⋮	
	⋮	⋮	⋮	

Fig. 9. Una entidad débil en relación, DEPENDIENTE de iones

NUMERO DE EMPLEADO no es un atributo de una entidad en el conjunto DEPENDIENTE sino es la clave principal de los empleados que apoyan a los dependientes. Cada fila de valores en la Figura 9 hay una tupla de entidad con EMPLOYEE-NO y NAME como it.s principal llave. Toda la tabla es una relación de entidad. En teoría, se puede utilizar cualquier tipo de relación para identificar entidades. Para simplicidad, nos limitaremos al uso de un solo tipo de relación:

las relaciones binarias con mapeo 1: n en el que la existencia de las n entidades en un lado de t, la relación depende de la existencia de una entidad en el otro lado de la relación. Por ejemplo, un empleado puede tener n (= 0, 1, 2,...) dependientes, y la existencia de los dependientes depende de la existencia del empleado correspondiente. Este método de identificación de entidades por relaciones con otras entidades puede ser aplicados de forma recursiva hasta que las entidades que pueden ser identificadas por su propia se alcanzan los valores tributos. Por ejemplo, la clave principal de un departamento en una



empresa puede constar del número de departamento y la clave principal del división, que a su vez consta del número de división y el nombre de la empresa. Por lo tanto, tenemos dos formas de relaciones entre entidades. Si las relaciones se utilizan para identificando las entidades, la llamaremos relación de entidad débil (Figura 9). Si relación- los barcos no se utilizan para identificar las entidades, lo llamaremos una relación de entidad regular (Figura 7). De manera similar, también tenemos dos formas de relaciones de relación. Me caigo entidades en la relación se identifican por sus propios valores de atributo, vamos a llámelo una relación de relación regular (Figura 8). Si algunas entidades en la relación se identifican por otras relaciones, la llamaremos relación de relación débil. Por ejemplo, cualquier relación entre entidades DEPENDIENTES y otras entidades resultará en relaciones de relación débiles, ya que se identifica una entidad DEPENDIENTE por su nombre y su relación con una entidad EMPLEADO. La distinción entre regular (entidad / relación) relaciones y débil (entidad / relación) Las relaciones serán útiles para mantener la integridad de los datos.

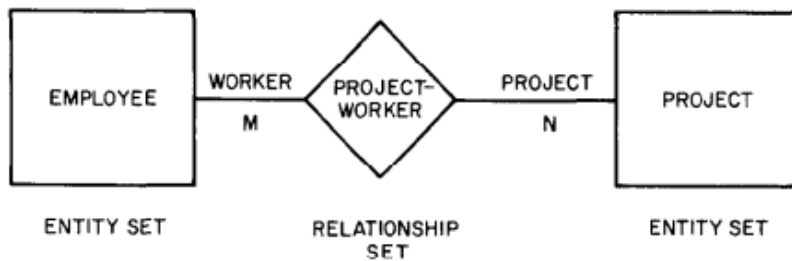


Fig. 10. Un diagrama entidad-relación simple

### 3. ENTIDAD-RELACIÓN ESQUEMA E INCLUSIÓN DE SEMÁNTICA EN DESCRIPCIÓN Y MANIPULACIÓN DE DATOS

#### 3.1 Análisis del sistema utilizando la relación entidad Diagrama

En esta sección presentamos una técnica diagramática para exhibir entidades y relaciones: el diagrama entidad-relación. La Figura 10 ilustra el conjunto de relaciones PROYECTO-TRABAJADOR y la entidad establece EMPLEADO y PROYECTO utilizando esta técnica de diagrama. Cada entidad El conjunto está representado por una caja rectangular, y cada conjunto de relaciones está representado por una caja en forma de diamante. El hecho de que la relación establezca PROYECTO-TRABAJADOR se define en los conjuntos de entidades EMPLEADO y PROYECTO está representado por las líneas que conectan las cajas rectangulares. Los roles de las entidades en la relación están expresados.

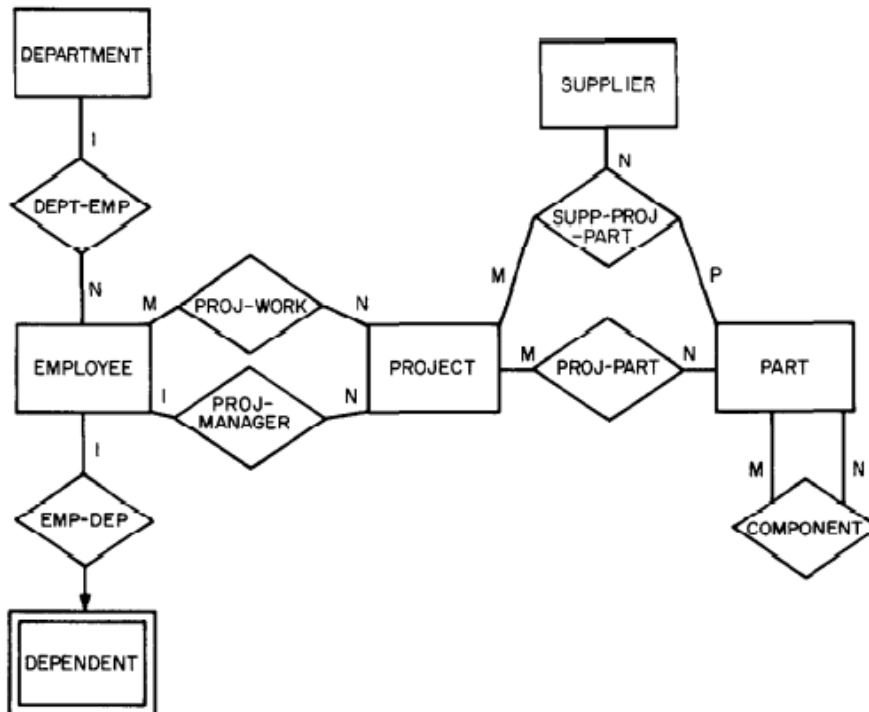


Fig. 11. Una entidad-relación

La figura 11 ilustra un diagrama más completo de algunos conjuntos de entidades y relaciones conjuntos que pueden ser de interés para una empresa de fabricación. DEPARTAMENTO, EMPLEADO, DEPENDIENTE, PROYECTO, PROVEEDOR y PARTE son entidades conjuntos. EMPLEADO DEPARTAMENTO, DEPENDIENTE DEL EMPLEADO, PROYECTO- TRABAJADOR, DIRECTOR DE PROYECTO, PROVEEDOR-PROYECTO-PARTE, PROYECTO-PART y COMPONENT son conjuntos de relaciones. El componente relación describe qué subpartes (y cantidades) se necesitan para hacer super- partes. No es necesario explicar el significado de los otros conjuntos de relaciones. Varias características importantes sobre las relaciones en general se pueden encontrar en Figura 11:

(1) Un conjunto de relaciones puede definirse en más de dos conjuntos de entidades. Por ejemplo, el PROYECTO-PROYECTO-PARTE El conjunto de relaciones se define en tres conjuntos de entidades, s: PROVEEDOR, PROYECTO y PARTE.

(2) Un conjunto de relaciones puede definirse en un solo conjunto de entidades. Por ejemplo, el El COMPONENTE del conjunto de relaciones se define en un conjunto de entidades, PARTE.

«(3) Puede haber más de un conjunto de relaciones definido en conjuntos de entidades dados. Por ejemplo, la relación establece PROYECTO-TRABAJADOR y PROYECTO- GERENTE se definen en los conjuntos de entidades PROYECTO y EMPLEADO.

(4) El diagrama puede distinguir entre asignaciones 1:12, m: 12 y 1: 1. Los conjunto de relaciones DEPARTMENT-EMPLOYEE es un mapeo 1: n, es decir, uno departamento puede tener 12 ( $n = 0, 1, 2, \dots$ ) empleados y cada empleado trabaja para solo un departamento. El conjunto de relaciones PROYECTO-TRABAJADOR

es un m: n mapeo, es decir, cada proyecto puede tener cero, uno o más empleados asignados a él y cada empleado pueden ser asignados a cero, uno o más proyectos. también es posible expresar asignaciones 1: 1 como el conjunto de relaciones MATRIMONIO. En para- información sobre el número de entidades en cada conjunto de entidades que se permite en una relación el armado del barco se indica especificando "I", "M", "% "en el diagrama. El relacional modelo y el conjunto de entidades model2 no incluyen este tipo de información; la red El modelo no puede expresar fácilmente un mapeo 1: 1.

(5) El diagrama puede expresar la dependencia de existencia de un tipo de entidad en otro. Por ejemplo, la flecha en el conjunto de relaciones EMPLOYEE-DEPEND-ENT indica que la existencia de una entidad, y en el conjunto de entidades DEPENDENT DEPENDENT depende de la entidad correspondiente en el conjunto de entidades EMPLEADO. Es decir, si un El empleado abandona la empresa, es posible que sus dependientes ya no sean de interés. Tenga en cuenta que el conjunto de entidades DEPENDENT se muestra como un cuadro rectangular especial. Esto indica que en el nivel 2 la información sobre las entidades de este conjunto está organizada como una relación de entidad débil (utilizando la clave primaria de EMPLEADO como parte de su Clave primaria).

Hay cuatro pasos para diseñar una base de datos utilizando el modelo entidad-relación:

(1) identificar los conjuntos de entidades y los conjuntos de relaciones de interés; (2) identificar información semántica en los conjuntos de relaciones, como si tiene un certificado en la set es un mapeo 1: n; (3) definir los conjuntos de valores y atributos; (4) organizar datos en las relaciones entidad / relación y decidir las claves primarias.

Usemos la empresa de fabricación discutida en la Sección 3.1 como ejemplo, Los resultados de los dos primeros pasos del diseño de la base de datos se expresan en una entidad:

diagrama de relaciones como se muestra en la Figura 11. El tercer paso es definir conjuntos de valores y atributos (ver Figuras 2 y 3). El cuarto paso es decidir el principal claves para las entidades y las relaciones y para organizar los datos como entidad / relación- relaciones marítimas. Tenga en cuenta que cada conjunto de entidad / relación en la Figura 11 tiene una correlación relación entidad / relación sponsori. Usaremos los nombres de los conjuntos de entidades. (en el nivel 1) como los nombres de las relaciones entidad / relaciones correspondientes (en nivel 2) siempre que no se produzca confusión.

Al final de la sección, ilustramos, calificamos un esquema (definición de datos) para un pequeño parte de la base de datos en el ejemplo de empresa de fabricación anterior (la sintaxis de la definición de datos no es importante). Tenga en cuenta que los conjuntos de valores se definen con especificaciones de representaciones y valores permitidos. Por ejemplo, los valores en EMPLEADO-NO se representan como números enteros de 4 dígitos y van de 0 a 2000. Luego declaramos tres relaciones de entidad: EMPLEADO, PROYECTO y DESEMBOLSO. PENDIENTE. Los atributos y conjuntos de valores definidos en los conjuntos de entidades, así como se indican las claves primarias. DEPENDIENTE es una relación de entidad débil ya que usa EMPLOYEE.PK como parte de su clave principal. También declaramos dos relaciones relaciones: PROYECTO-TRABAJADOR y EMPLEADO-DEPENDIENTE. Los roles y se especifican las entidades involucradas en las relaciones. Usamos EMPLOYEE.PK para indicar el nombre de la entidad relación (EMPLEADO) y cualquier atributo- Los pares de conjuntos de valores se utilizan como claves primarias en esa relación de entidad. El máximo se indica el número de entidades de un conjunto de entidades en una relación. Por ejemplo, PROYECTO WORKER es un mapeo m: n. Podemos especificar los valores de my n. Podemos también especifique el número mínimo de entidades además del número máximo. EMPLEADO-DEPENDIENTE es una relación de relación débil ya que uno de las relaciones de entidad relacionadas, DEPENDIENTE, es una relación

de entidad débil. Tenga en cuenta que el También se indica la existencia de dependencia de los dependientes del sustentador.

<u>DECLARE</u>	<u>VALUE-SETS</u>	<u>REPRESENTATION</u>	<u>ALLOWABLE-VALUES</u>
	EMPLOYEE-NO	INTEGER (4)	(0,2000)
	FIRST-NAME	CHARACTER (8)	ALL
	LAST-NAME	CHARACTER (10)	ALL
	NO-OF-YEARS	INTEGER (3)	(0,100)
	PROJECT-NO	INTEGER (3)	(1,500)
	PERCENTAGE	FIXED (5.2)	(0,100.00)

DECLARE      REGULAR ENTITY RELATION EMPLOYEE  
ATTRIBUTE/VALUE SET:  
                 EMPLOYEE-NO/EMPLOYEE-NO  
                 NAME/(FIRST-NAME, LAST-NAME)  
                 ALTERNATIVE-NAME/(FIRST-NAME,LAST-NAME)  
                 AGE/NO-OF-YEARS  
PRIMARY KEY:  
                 EMPLOYEE-NO

DECLARE      REGULAR ENTITY RELATION PROJECT  
ATTRIBUTE/VALUE-SET:  
                 PROJECT-NO/PROJECT-NO  
PRIMARY KEY:  
                 PROJECT-NO

DECLARE      REGULAR RELATIONSHIP RELATION PROJECT WORKER  
ROLE/ENTITY-RELATION.PK/MAX-NO-OF-ENTITIES  
                 WORKER/EMPLOYEE.PK/m  
                 PROJECT/PROJECT.PK/n      (m:n mapping)  
ATTRIBUTE/VALUE-SET:  
                 PERCENTAGE-OF-TIME/PERCENTAGE

DECLARE      WEAK RELATIONSHIP RELATION EMPLOYEE-DEPENDENT  
ROLE/ENTITY-RELATION.PK/MAX-NO-OF-ENTITIES  
                 SUPPORTER/EMPLOYEE.PK/1  
                 DEPENDENT/DEPENDENT.PK/n  
EXISTENCE OF DEPENDENT DEPENDS ON  
EXISTENCE OF SUPPORTER

DECLARE      WEAK ENTITY RELATION DEPENDENT  
ATTRIBUTE/VALUE-SET:  
                 NAME/FIRST-NAME  
                 AGE/NO-OF-YEARS  
PRIMARY KEY:  
                 NAME  
                 EMPLOYEE.PK THROUGH EMPLOYEE-DEPENDENT

### 3.3 Implicaciones sobre la integridad de los datos

Se ha realizado algún trabajo sobre la integridad de los datos para otros modelos [S, 14, 16, 28-j]. Con conceptos explícitos de entidad y relación, el modelo entidad-relación será útil para comprender y especificar las limitaciones para mantener la integridad de los datos. Por ejemplo, existen tres tipos principales de restricciones sobre los valores:

(1) Restricciones sobre los valores permitidos para un conjunto de valores. Este punto fue discutido en definiendo el esquema en la Sección 3.2.

(2) Restricciones sobre los valores permitidos para un determinado atributo. En algunos casos, no todos los valores permitidos en un conjunto de valores están permitidos para algunos atributos. Por ejemplo, es posible que tengamos una restricción de edades de los empleados entre 20 y 65 años. Es decir,  $EDAD(e) \in (20,65)$ , donde  $e \in EMPLEADO$ . Tenga en cuenta que usamos las notaciones de nivel 1 para aclarar la semántica. Dado que cada entidad / el conjunto de relaciones tiene una entidad / relación correspondiente relación, la expresión anterior si no se puede traducir fácilmente a notaciones de nivel 2.

(3) Restricciones sobre los valores de  $e$  &  $sting$  en la base de datos. Hay dos tipos de restricciones:

(i) Restricciones entre conjuntos de valores existentes. Por ejemplo,  $\{NOMBRE(e) \in E \text{ MASculino-PERSONA} \mid NOMBRE(e) \in E \text{ PERSONA}\}$ .

(ii) Restricciones entre valores particulares. Por ejemplo,

$TAX(e) \leq SALARY(e), e \in EMPLOYEE$

or  
 $BUDGET(e_i) = \sum_{e_j \in DEPARTMENT} BUDGET(e_j), \text{ where } e_i \in COMPANY$   
and  $[e_i, e_j] \in COMPANY-DEPARTMENT$ .

### 3.4 Semántica y operaciones de conjuntos de solicitudes de recuperación de información

La semántica de las solicitudes de recuperación de información se vuelve muy clara si las solicitudes se basan en el modelo de datos entidad-relación. Para mayor claridad, primero discutimos la situación en el nivel 1. Conceptualmente, los elementos de información se organizan como en las Figuras 4 y 5 (en las Figuras 2 y 3). Muchas solicitudes de recuperación de información pueden ser considerado como una combinación de los siguientes tipos básicos de operaciones:

(1) Selección de un subconjunto de valores de un conjunto de valores.

(2) Selección de un subconjunto de entidades de un conjunto de entidades (es decir, selección de ciertas filas en la Figura 4). Las entidades se seleccionan indicando los valores de ciertos atributos. (es decir, subconjuntos de conjuntos de valores) y / o sus relaciones con otras entidades.

(3) Selección de un subconjunto de relaciones de un conjunto de relaciones (es decir, selección de ciertas filas en la Figura 5).

(4) Las relaciones se seleccionan indicando los valores de ciertos atributos y / o identificando ciertas entidades en la relación. (4) Selección de un subconjunto de atributos (es decir, selección de columnas en las Figuras 4) y 5). Una solicitud de recuperación de información como "¿Cuáles son las edades de los empleados cuyas los pesos son superiores a 170 y que se asignan al proyecto con PROYECTO-

NO 254? se puede expresar como:

$\{AGE(e) \mid e \in EMPLOYEE, WEIGHT(e) > 170,$   
 $[e, e_j] \in PROJECT-WORKER, e_j \in PROJECT,$   
 $PROJECT-NO(e_j) = 254\};$

or

$\{AGE(EMPLOYEE) \mid WEIGHT(EMPLOYEE) > 170,$   
 $[EMPLOYEE, PROJECT] \in PROJECT-WORKER,$   
 $PROJECT-NO(EMPLOYEE) = 254\}.$

To retrieve information as organized in Figure 6 at level 2, "entities" and "relationships" in (2) and (3) should be replaced by "entity PK" and "relationship PK." The above information retrieval request can be expressed as:

$\{AGE(EMPLOYEE.PK) \mid WEIGHT(EMPLOYEE.PK) > 170$   
 $(WORKER/EMPLOYEE.PK, PROJECT/PROJECT.PK) \in [PROJECT-WORKER.PK],$   
 $PROJECT-NO (PROJECT.PK) = 254\}.$

Para recuperar información organizada en relaciones entidad / relación (Figuras 7, 8 y 9)) podemos expresarlo en un lenguaje similar a SEQUEL [S]:

Table I. Insertion

level 1	level 2
<i>operation:</i> insert an entity to an entity set	<i>operation:</i> create an entity tuple with a certain entity-PK <i>check:</i> whether PK already exists or is acceptable
<i>operation:</i> insert a relationship in a relationship set  <i>check:</i> whether the entities exist	<i>operation:</i> create a relationship tuple with given entity PKs <i>check:</i> whether the entity PKs exist
<i>operation:</i> insert properties of an entity or a relationship  <i>check:</i> whether the value is acceptable	<i>operation:</i> insert values in an entity tuple or a relationship tuple <i>check:</i> whether the values are acceptable

AND           EMPLOYEE.PK =  
SELECT     WORKER/EMPLOYEE.PK  
FROM       PROJECT-WORKER  
WHERE      PROJECT-NO = 254.

Es posible recuperar información sobre entidades en dos conjuntos de entidades diferentes sin especificar una relación entre ellos. Por ejemplo, una información solicitud de recuperación como "Enumere los nombres de empleados y barcos que tienen el mismo

Table II. Updating

level 1	level 2
<i>operation:</i> • change the value of an entity attribute	<i>operation:</i> • update a value <i>consequence:</i> • if it is not part of an entity PK, no consequence • if it is part of an entity PK, •• change the entity PKs in all related relationship relations •• change PKs of other entities which use this value as part of their PKs (for example, DEPENDENTS' PKs use EMPLOYEE'S PK)
<i>operation:</i> • change the value of a relationship attribute	<i>operation:</i> • update a value (note that a relationship attribute will not be a relationship PK)

Table III. Deletion

level 1	level 2
<i>operation:</i> <ul style="list-style-type: none"> <li>delete an entity</li> </ul> <i>consequences:</i> <ul style="list-style-type: none"> <li>delete any entity whose existence depends on this entity</li> <li>delete relationships involving this entity</li> <li>delete all related properties</li> </ul>	<i>operation:</i> <ul style="list-style-type: none"> <li>delete an entity tuple</li> </ul> <i>consequences (applied recursively):</i> <ul style="list-style-type: none"> <li>delete any entity tuple whose existence depends on this entity tuple</li> <li>delete relationship tuples associated with this entity</li> </ul>
<i>operation:</i> <ul style="list-style-type: none"> <li>delete a relationship</li> </ul> <i>consequences:</i> <ul style="list-style-type: none"> <li>delete all related properties</li> </ul>	<i>operation:</i> <ul style="list-style-type: none"> <li>delete a relationship tuple</li> </ul>

age'' can be expressed in the level 1 notation as:

$$\{(\text{NAME}(e_i), \text{NAME}(e_j)) \mid e_i \in \text{EMPLOYEE}, e_j \in \text{SHIP}, \text{AGE}(e_i) = \text{AGE}(e_j)\}.$$

No discutimos más la sintaxis del lenguaje aquí. Lo que deseamos enfatizar es que las solicitudes de información pueden expresarse utilizando nociones de conjuntos y operaciones de conjuntos [17 3, y la semántica de la solicitud es muy clara al adoptar este punto de vista.

### 3.5 Semántica y reglas para la inserción, eliminación y actualización

Es siempre un problema difícil de mantener inserción siguiente consistencia de los datos, eliminación y actualización de datos en la base de datos. Una de las principales razones es que la semántica y las consecuencias de las operaciones de inserción, eliminación y actualización generalmente no están claramente definidos; por lo tanto, es difícil encontrar un conjunto de reglas que puedan hacer cumplir la coherencia de los datos. Veremos que este problema de coherencia de datos se convierte en más simple usando el modelo entidad-relación. En las Tablas I-III, discutimos la semántica y las reglas<sup>3</sup> para la inserción, eliminación y actualización tanto en el nivel 1 como en el nivel 2. El nivel 1 se utiliza para aclarar la semántica.

## 4. ANÁLISIS DE OTROS MODELOS DE DATOS Y SU DERIVACIÓN DEL RELACIÓN ENTRE ENTIDADES MODELO

### 4.1 El modelo relacional

#### 4.1.1 La visión relacional de los datos y la ambigüedad en semántica. Ahí-modelo lacional, relación,

R, es una relación matemática definida en los conjuntos  $X_1, X_2, \dots, X_n$ :

$$R = I(a, 22, \dots, 2, 121EX1, 22EX2, \dots, 2, EK).$$

Los conjuntos  $X_1, X_2, \dots, X_n$  se denominan dominios y  $(x_1, x_2, \dots, x_n)$  se denomina tupla. La Figura 12 ilustra una relación denominada EMPLEADO. Los dominios en la relación 3 Nuestro principal objetivo es ilustrar la semántica de las operaciones de manipulación de datos. Por lo tanto, Es posible que estas reglas no estén completas. Tenga en cuenta que la consecuencia de las operaciones indicadas en las tablas puede ser realizado por el sistema en lugar de por los usuarios.

ROLE	LEGAL		ALTERNATIVE		
	EMPLOYEE-NO	FIRST-NAME	LAST-NAME	FIRST-NAME	LAST-NAME
DOMAIN	2566	PETER	JONES	SAM	JONES
	3378	MARY	CHEN	BARB	CHEN
TUPLE	25				
	23				

Fig. 12. Relación EMPLEADO

son El \$ PLOYEE-NO, PRIMER NOMBRE, APELLIDO, PRIMER NOMBRE, ÚLTIMO- NOMBRE, NO DEL AÑO. El orden de filas y columnas en la relación tiene sin importancia. Para evitar la ambigüedad de columnas con el mismo dominio en una relación, Los nombres de dominio están calificados por roles (para distinguir el rol del dominio en el relación). Por ejemplo, en relación EMPLEADO, dominios FIRST-NAME y APELLIDO puede ser calificado por roles LEGAL o ALTERNATIVO. Un atributo name en el modelo relacional es un nombre de dominio concatenado con un nombre de rol [lo]. Comparando la Figura 12 con la Figura 7, podemos ver que los "dominios" son básicamente equivalentes prestado a conjuntos de valores. Aunque "rol" o "atributo" en el modelo relacional parece sirven para el mismo propósito que "atributo" en el modelo entidad-relación, el se-Las mánticas de estos términos son diferentes. El "rol" o "atributo" en el relacional El modelo se utiliza principalmente para distinguir dominios con el mismo nombre en el mismo relación, mientras que "atributo" en el modelo entidad-relación es una función que mapas de un conjunto de entidades (o relaciones) en conjuntos de valores. El uso de operadores relacionales en el modelo relacional puede causar ambigüedades semánticas. güties. Por ejemplo, la unión de la relación EMPLEADO con la relación

EMPLEADO-PROYECTO

(Figura 13) en el dominio EMPLEADO-NO produce el

PROJECT-NO	EMPLOYEE-NO
7	2566
3	2566
7	3378

Fig. 13. Relación EMPLEADO-PROYECTO

		LEGAL	LEGAL	ALTERNATIVE	ALTERNATIVE	
PROJECT-NO	EMPLOYEE-NO	FIRST-NAME	LAST-NAME	FIRST-NAME	LAST-NAME	NO-OF-YEARS
7	2566	PETER	JONES	SAM	JONES	25
3	2566	PETER	JONES	SAM	JONES	25
7	3378	MARY	CHEN	BARB	CHEN	23

Fig. 14. Relación EMPLEADO-PROYECTO 'como "unión" de relaciones EMPLEADO y EMPLEADO-PROYECTO

relación EMPLEADO-PROYECTO ' (Figura 14). Pero, ¿cuál es el significado de un unión entre la relación EMPLEADO con la relación ENVÍO en el dominio NO, -DE-AÑOS (Figura 15)? El problema es que el mismo nombre de dominio puede tienen semánticas diferentes en relaciones diferentes (tenga en cuenta que un rol está destinado a dominios tingüísticos en una relación dada, no en todas las relaciones). Si el dominio NO-OF- AÑO de la relación EMPLEADO no se permite comparar con el dominio NO-DE-AÑO de la relación SHIP, se deben declarar diferentes nombres de dominio.

Pero si tal comparación es aceptable, ¿puede el sistema de base de datos advertir al usuario? En el modelo entidad-relación, la semántica de los datos es mucho más aparente. Por ejemplo, una columna en el ejemplo anterior contiene los valores de AGE de EMPLEADO y la otra columna contiene los valores de EDAD de NAVE. Si esta información semántica está expuesta al usuario, puede operar con más cautela (consulte los ejemplos de solicitudes de recuperación de información que se indican en la Sección 3.4). Ya que el sistema de base de datos contiene la información semántica, debería poder advertir el usuario de los problemas potenciales para una operación propuesta "tipo unión".

4.1.2 Semántica de las dependencias funcionales entre datos. En lo relacional modelo, el "atributo" B de una relación es funcionalmente dependiente del "atributo" A del misma relación si cada valor de A no tiene más de un valor de B asociado con en la relación. La semántica de las dependencias funcionales entre los datos se vuelve clara

SHIP-NO	NAME	NO-OF-YEARS
037	MISSOURI	25
056	VIRGINIA	10

Fig.15. Relación NAVE

en el modelo entidad-relación. Básicamente, hay dos tipos principales de funciones dependencias:

(1) Dependencias funcionales relacionadas con la descripción de entidades o relaciones. Dado que un atributo se define como una función, asigna una entidad en un conjunto de entidades a un valor único en un conjunto de valores (consulte la Figura 2). En el nivel 2, los valores de la clave primaria se utilizan para representar entidades. Por lo tanto, los conjuntos de valores no clave (dominios) son dependiente de los conjuntos de valores de clave primaria (por ejemplo, en las Figuras 6 y 7, NO-OF-AÑOS es funcionalmente dependiente de EMPLEADO-NO). Dado que una relación puede tener varias claves, los conjuntos de valores no clave dependerán funcionalmente de cualquier clave conjunto de valores. Los conjuntos de valores clave serán mutuamente funcionalmente dependientes de cada otro. De manera similar, en una relación de relación, los conjuntos de valores no clave serán funcionalmente depende de los conjuntos de valores de clave principal (por ejemplo, en la Figura 8, PORCENTAJE depende funcionalmente de EMPLOYEE-NO y PROJECT-NO).

(2) Dependencias funcionales relacionadas con entidades en una relación. Tenga en cuenta que En la Figura 11 identificamos los tipos de asignaciones (1: n, m: n, etc.) para la relación conjuntos. Por ejemplo, PROJECT-XfANAGER es un mapeo 1: n. Asumamos que PROYECTO-NO es la clave primaria en la relación de entidad PROYECTO. Ahí- relación de relación PROYECTO-GESTOR, el valor establecido EMPLEADO-NO será ser funcionalmente dependiente del valor establecido PROYECTO-NO (es decir, cada proyecto tiene solo un gerente). La distinción entre el nivel 1 (Figura 2) y el nivel 2 (Figuras 6 y 7) y la separación de la relación de entidad (Figura 7) de la relación de relación (Figura 8) aclara la semántica de las dependencias funcionales entre los datos.

4.1.3 Relaciones 3NF versus relaciones entidad-relación. De la definición de "relación", cualquier agrupación de dominios puede considerarse una relación. Para evitar propiedades indeseables en el mantenimiento de relaciones, se propone un proceso de normalización transformar relaciones arbitrarias en la primera forma normal, luego en la segunda forma normal, y finalmente en la tercera forma normal (3NF) [9, 111. Vamos a mostrar que las relaciones entidad y relación en el modelo entidad-relación son similares a las relaciones 3NF pero con una semántica más clara y sin utilizar la operación de transformación. Usemos una versión simplificada de un ejemplo de normalización descrito en [9].

Las siguientes tres relaciones están en primera forma normal (es decir, no hay dominio cuyos elementos son ellos mismos relaciones):

**EMPLOYEE (EMPLOYEE-NO)**

**PART (PART-NO, PART-DESCRIPTION, QUANTITY-ON-HAND)**

**PART-PROJECT (PART-NO, PROJECT-NO, PROJECT-DESCRIPTION,  
PROJECT-MANAGER-NO, QUANTITY-COMMITTED).**

Tenga en cuenta que el dominio PROJECT-MANAGER-NO realmente contiene el EMPLEADO-NO del director del proyecto. En las relaciones anteriores, claves primarias están subrayados.

Se aplican ciertas reglas para transformar las relaciones anteriores en una tercera normalidad. formulario:

**EMPLOYEE' (EMPLOYEE-NO)**

**PART' (PART-NO, PART-DESCRIPTION, QUANTITY-ON-HAND)**



PROJECT' (PROJECT-NO, PROJECT-DESCRIPTION, PROJECT-MANAGER-NO)  
 PART-PROJECT' (PART-NO, PROJECT-NO, QUANTITY-COMMITTED).

Using the entity-relationship diagram in Figure 11, the following entity and relationship relations can be easily derived:

entity	PART'' (PART-NO, PART-DESCRIPTION, QUANTITY-ON-HAND)
relations	PROJECT'' (PROJECT-NO, PROJECT-DESCRIPTION)
	EMPLOYEE '' (EMPLOYEE-NO)
relationship	PART-PROJECT'' (PART/PART-NO, PROJECT/PROJECT-NO,
relations	QUANTITY-COMMITTED)
	PROJECT-MANAGER'' (PROJECT/PROJECT-NO,
	MANAGER/EMPLOYEE-NO).

Se indican los nombres de roles de las entidades en las relaciones (como ADMINISTRADOR). Los nombres de relación de entidad asociados con los PK de entidades en relaciones y se han omitido los nombres de los conjuntos de valores.

Tenga en cuenta que en el ejemplo anterior, entidad / relación las relaciones son similares a las Relaciones 3NF. En el enfoque 3NF, PROJECT-MANAGER-NO está incluido en la relación PROYECTO 'desde PROYECTO-GESTOR-NO se supone que es funcionalmente dependiente de PROYECTO-NO. En la entidad-relación

modelo, PROYECTO-GERENTE-NO (es decir, EMPLEADO-NO de un director de proyecto) es incluido en una relación de relación PROYECTO-GESTOR desde EMPLOYEE-NO se considera una entidad PK en este caso.

También tenga en cuenta que en el enfoque 3NF, los cambios en las dependencias funcionales de los datos pueden causar que algunas relaciones no estén en 3NF. Por ejemplo, si hacemos una nueva as-Suponiendo que un proyecto puede tener más de un gerente, la relación PROYECTO 'ya no es una relación 3NF y debe dividirse en dos relaciones como PROYECTO "y PROYECTO-GERENTE! '. Usando el modelo entidad-relación, no es necesario tal cambio. Por tanto, podemos decir que al utilizar la entidad-modelo de relación podemos organizar los datos en una forma similar a las relaciones 3NF pero con significado semántico claro.

Es interesante notar que el enfoque de descomposición (o transformación) descrito anteriormente para la normalización de las relaciones puede verse como un enfoque en el diseño de bases de datos.4 Comienza con relaciones arbitrarias (nivel 3 en la Figura 1) y luego utiliza alguna información semántica (dependencias funcionales de los datos) para transformarlos en relaciones 3NF (nivel 2 en la Figura 1). La entidad-relación modelo adopta un enfoque de arriba hacia abajo, utilizando la información semántica para organizar datos en relaciones entidad / relación.

#### 4.2 El modelo de red

4.2.1 Semántica del diagrama de estructura de datos. Una de las mejores formas de explicar el modelo de red se basa en el diagrama de estructura de datos [3]. Figura 16 (a) ilus- Trata un diagrama de estructura de datos. Cada cuadro rectangular representa un tipo de registro.

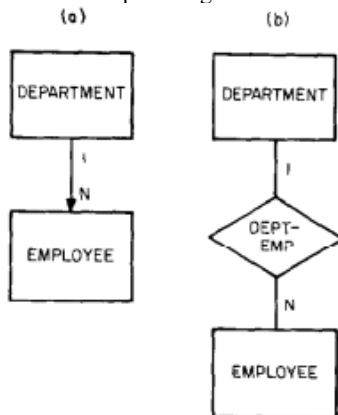


Fig. 16. Relationship DEPARTMENT-EMPLOYEE  
 (a) data structure diagram  
 (b) entity-relationship diagram

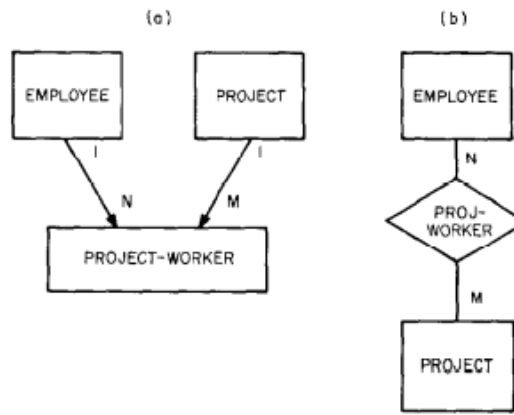


Fig. 17. Relationship PROJECT-WORKER  
 (a) data structure diagram  
 (b) entity-relationship diagram

La flecha representa un conjunto de estructura de datos en el que el DEPARTAMENTO registro es el registro de ozcner, y un registro de propietario puede poseer n (n = 0, 1, 2,...) miembro- registros. La figura 16 (b) ilustra el diagrama entidad-relación correspondiente. Se podría concluir que la flecha en el diagrama de estructura de datos representa una relación entre entidades en dos conjuntos de entidades. Esto no siempre es cierto. Cifras 17 (a) y 17 (b) son el diagrama de estructura de datos y el diagrama de entidad-relación expresando la relación PROYECTO-TRABAJADOR entre dos tipos de entidad EMPLEADO y PROYECTO. Podemos ver en la Figura 17 (a) que la relación TRABAJADOR DE PROYECTOS se

convierte en otro tipo de registro y que las flechas no ya representan relaciones entre entidades. ¿Cuáles son los significados reales de las flechas en diagramas de estructura de datos? La respuesta es que una flecha representa un 1: n relación entre dos tipos de registro (no entidad) y también implica la existencia de una ruta de acceso desde el registro de propietario a los registros de miembros. La estructura de datos El diagrama es una representación de la organización de los registros (nivel 4 en la Figura 1) y no es una representación exacta de entidades y relaciones.

4.2.2 Derivación del diagrama de estructura de datos. ¿En qué condiciones un La flecha en un diagrama de estructura de datos corresponde a una relación de entidades? Un acercamiento comparación de los diagramas de estructura de datos con el correspondiente entidad-relación- Los diagramas de barcos revelan las siguientes reglas:

1. Para relaciones binarias 1: n, se usa una flecha para representar la relación (ver Figura 16 (a)).
2. Para m: n relaciones binarias, se crea un tipo de "registro de relación" para representar enviado la relación y las flechas se dibujan desde el tipo "registro de entidad" al Tipo de "registro de relación" (ver Figura 17 (a)).
3. Para las relaciones k-ary (k > 3), se aplica la misma regla que (2) (es decir, crear un Tipo "registro de relación"). Dado que DBTG [7] no permite definir un conjunto de estructura de datos en un solo tipo de registro (es decir, la Figura 18 no está permitida aunque se ha implementado en[13]), se necesita un "registro de relación" para implementar tales relaciones (ver

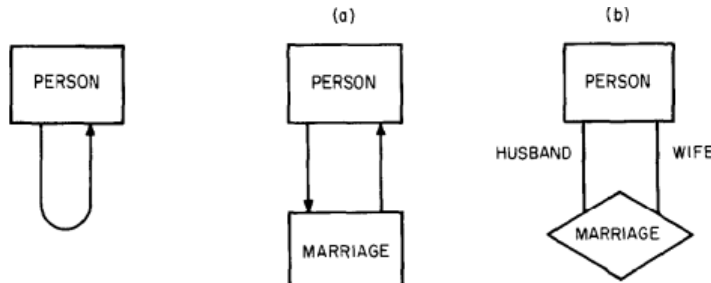


Fig. 18. Data-structure-set defined on the same record type

Fig. 19. Relationship MARRIAGE (a) data structure diagram (b) entity-relationship diagram

Figura 19 (a) [20]. El diagrama entidad-relación correspondiente se muestra en Figura 19 (b). Ahora está claro que las flechas en un diagrama de estructura de datos no siempre representan relaciones de entidades. Incluso en el caso de que una flecha represente una relación 1: n, ion- barco, la flecha sólo representa una relación unidireccional [20] (aunque es posible encontrar el registro de propietario de un registro de miembro). En la entidad-relación modelo, ambas direcciones de la relación están representadas (los roles de ambos en- tidades se especifican). Además de la ambigüedad semántica en sus flechas, la red El modelo es incómodo para manejar cambios en la semántica. Por ejemplo, si la relación entre DEPARTAMENTO y EMPLOYEE cambia de un mapeo 1: n a un mapeo m: n (es decir, un empleado puede pertenecer a varios departamentos), debemos crear un registro de relación DEPARTAMENTO-EMPLEADO en el modelo de red.

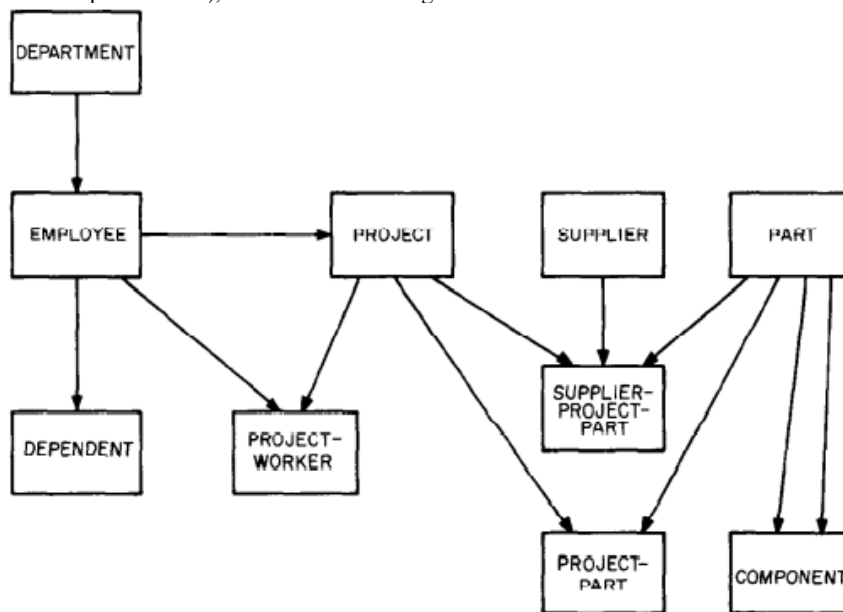


Fig. 20. The data structure diagram derived from the entity-relationship diagram in Fig. 11

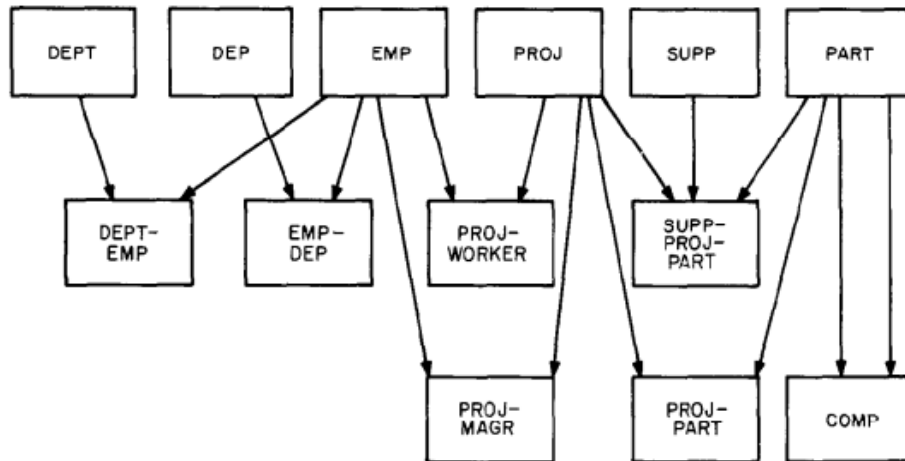


Fig. 21. The "disciplined" data structure diagram derived from the entity-relationship diagram in Fig. 11

En el modelo entidad-relación, se manejan todo tipo de asignaciones en las relaciones uniformemente.

El modelo de relación de entidad se puede utilizar como una herramienta en el diseño estructurado de bases de datos utilizando el modelo de red. El usuario primero dibuja una entidad-relación diagrama (Figura 11). Puede simplemente traducirlo en un diagrama de estructura de datos. (Figura 20). utilizando las reglas especificadas anteriormente. También puede seguir una disciplina que cada entidad o relación debe asignarse a un registro (es decir, "relación registros" se crean para todo tipo de relaciones sin importar que sean 1: no m: n asignaciones). Por tanto, en la Figura 11, todo lo que hay que hacer es cambiar los diamantes a los cuadros y agregar puntas de flecha en las líneas apropiadas. Usando este enfoque tres casillas más-DEPARTAMENTO-EMPLEADO, EMPLEADO-DEPENDIENTE y PROJECT-MANAGER-will agregarse a la Figura 20 (ver Figura 21). Las restricciones de validez discutidas en las Secciones 3.3-3.5 también serán útiles.

#### 4.3 El modelo de conjunto de entidades

4.3.1 Vista de conjunto de entidades. El elemento básico del modelo de conjunto de entidades es el entidad. Las entidades tienen nombres (nombres de entidades) como "Peter Jones", "azul" o "22". Los nombres de entidades que tienen algunas propiedades en común se recopilan en un conjunto de nombre de entidad, al que hace referencia el nombre de conjunto de nombre de entidad, como "NOMBRE", "COLOR" y "CANTIDAD". Una entidad está representada por el nombre-conjunto-nombre-entidad / nombre-entidad par como NOMBRE / Peter Jones, EMPLEADO-NO / 2566, y NO DE AÑOS / 20. Una entidad se describe por su asociación con otras entidades. La figura 22 ilustra la entidad establecer vista de datos. El Departamento" de la entidad EMPLEADO-NO / 2566 es el entidad DEPARTAMENTO-NO / 405. En otras palabras, "DEPARTAMENTO" es el papel que la entidad DEPARTAMENTO-NO / 405 juega para describir la entidad EM-PLREADO-NO / 2566. Del mismo modo, el "NOMBRE", "NOMBRE ALTERNATIVO", o "EDAD" de EMPLOYEE-NO / 2566 es "NAME / Peter Jones", "NAME / Sam Jones", o "NO DE AÑOS / 20", respectivamente. La descripción de la entidad EMPLEADO-NO / 2566 es una colección de las entidades relacionadas y sus roles (las entidades y roles encerrados en un círculo por la línea de puntos). Un ejemplo de la descripción de entidad de "EM-PLOYEE-NO / 2566" (en su forma completa, no factorizada) se ilustra con el conjunto de nombre-rol / nombre-entidad-nombre-conjunto / nombre-entidad trillizos que se muestran en la Figura 23. Conceptualmente, el modelo de conjunto de entidades difiere del modelo de entidad-relación en las siguientes formas:

(1) En el modelo de conjunto de entidades, todo se trata como una entidad. Por ejemplo, "DE COLOR NEGRO" y "NO-DE-AÑOS / 45" son entidades. En la entidad-relación- modelo de barco, "azul" y "36" generalmente se tratan como valores. Tenga en cuenta el tratamiento de los valores como las entidades pueden causar problemas semánticos. Por ejemplo, en la Figura 2'2, ¿cuál es la diferencia entre "EMPLOYEE-NO / 2566", "NOMBRE / Peter Jones" y "NOMBRE / Sam Jones"? ¿Representan diferentes entidades?

(2) Solo se utilizan relaciones binarias en el modelo de conjunto de entidades, 5 mientras que n-ary las relaciones pueden usarse en el modelo entidad-relación.

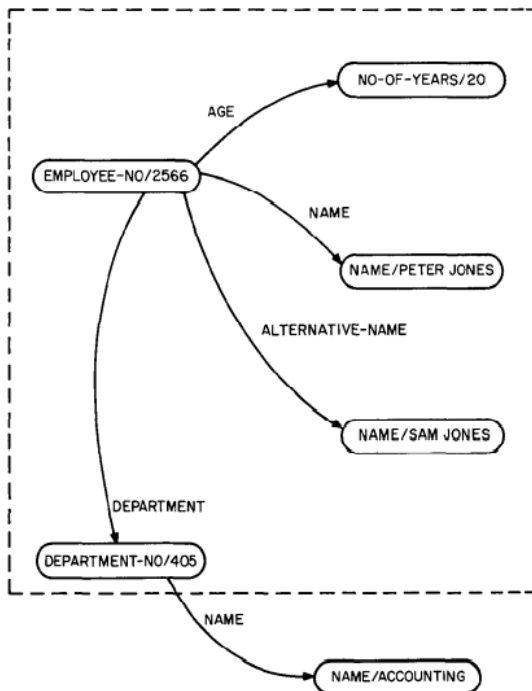


Fig. 22. The entity-set view

THE ENTITY- RELATIONSHIP MODEL TERMINOLOGY	ATTRIBUTE OR ROLE	VALUE SET	VALUE
THE ENTITY SET MODEL TERMINOLOGY	"ROLE-NAME"	"ENTITY-NAME- SET-NAME"	"ENTITY-NAME"
	IDENTIFIER	EMPLOYEE-NO	2566
	NAME	NAME	PETER JONES
	NAME	NAME	SAM JONES
	AGE	NO-OF-YEARS	25
	DEPARTMENT	DEPARTMENT-NO	405

Fig. 23. An "entity description" in the entity-set model

43.2 Derivación de la vista Conjunto de entidades. Una de las principales dificultades para comprender mantener el modelo de conjunto de entidades se debe a su visión del mundo (es decir, identificar valores con entidades). El modelo entidad-relación propuesto en este documento es útil para comprender Situar y derivar la vista de datos del conjunto de entidades. Considere las Figuras 2 y 6. En Figura 2, las entidades están representadas por *ei* (que existen en nuestras mentes o están apuntadas con los dedos). En la Figura 6, las entidades están representadas por valores. El conjunto de entidades, El modelo funciona tanto en el nivel 1 como en el nivel 2, pero explicaremos su punto de vista en el nivel 2.

(Figura 6). El modelo de conjunto de entidades trata todos los conjuntos de valores como NO-OF-AÑOS como "conjuntos de nombres de entidades" y todos los valores como "nombres de entidades". Los atributos se vuelven nombres de roles en el modelo de conjunto de entidades. Para relaciones binarias, la traducción es simple: el papel de una entidad en una relación (por ejemplo, el papel de "DEPART-MENT" en la relación DEPARTAMENTO-EMPLEADO) se convierte en el papel

nombre de la entidad al describir la otra entidad en la relación (ver Figura 22), para relaciones n-arias ( $n > 2$ ), debemos crear entidades artificiales para relaciones barcos para manejarlos en un mundo de relaciones binarias.

#### EXPRESIONES DE GRATITUD

El autor desea expresar su agradecimiento a George Mealy, Stuart Madnick, Murray Edelberg, Susan Brewer, Stephen Todd y los árbitros por sus valiosas sugerencias (la Figura 21 fue sugerida por uno de los árbitros). Este trabajo fue motivado por una serie de discusiones con Charles Bachman. El autor también está en deuda con EF Codd y ME Senko por sus valiosos comentarios y discusiones al revisar este papel.

#### REFERENCIAS

1. ABRIAL, JR Semántica de datos. En Gestión de bases de datos, JW Klimbie y KL Koffeman, Eds., Pub de Holanda Septentrional. Co., Amsterdam, 1974, págs. 1-60.
  2. BACHMAN, CW Software para procesamiento de acceso aleatorio. Datamation 11 (abril de 1965), 36-41.
  3. BACHMAN, CW Diagramas de estructura de datos. Base de datos 2, 2 (verano de 1969), 4-10.
  4. BACHMAN, CW Trends in database management-1975. Proc., AFIPS 1975 NCC, vol. 44, AFIPS Press, Montvale, Nueva Jersey, págs. 569-576.
  5. BIRKHOFF, G., Y BARTEE, TC Álgebra aplicada moderna. McGraw-Hill, Nueva York, 1970.
  6. CHAMBERLIN, DD, Y RAYMOND, FB SEQUEL: Un lenguaje de consulta estructurado en inglés. Proc. ACM-SIGMOD 1974, taller, Ann Arbor, Michigan, mayo de 1974.
  7. CODASYL. Informe del grupo de tareas de la base de datos. ACM, Nueva York, 1971.
  8. CODD, EF Un modelo relacional de datos para grandes bancos de datos compartidos. Comm. ACM 13, 6 (junio 1970), 377-387.
  9. CODD, EF Estructura de base de datos normalizada: un breve tutorial. Proc. ACM-SIGFIDET 1971, Workshop, San Diego, California, noviembre de 1971, págs. 1-18.
  10. CODD, EF Sublenguaje de una base de datos fundada en el cálculo relacional. Proc. ACM-SIGFIDET 1971, Workshop, San Diego, California, noviembre de 1971, págs. 35-68.
  11. CODD, EF Investigaciones recientes en sistemas de bases de datos relacionales. Proc. Congreso IFIP 1974, Pub de Holanda Septentrional. Co., Amsterdam, págs. 1017-1021.
  12. DEHENEFFE, C., HENNEBERT, H. Y PAULUS, W. Modelo relacional para base de datos. Proc. Congreso IFIP 1974, North-Holland Pub. Co., Amsterdam, págs. 1022-1025.
  13. DODD, GG APL: un lenguaje para el manejo de datos asociados en PL / I. Proc. AFIPS 1966 FJCC, vol. 29, Spartan Books, Nueva York, págs. 677-684.
  14. ESWARAN, KP Y CHAMBERLIN, DD Especificaciones funcionales de un subsistema de datos integridad de la base. Proc. Very Large Data Base Conf., Framingham, Mass., Septiembre de 1975, págs. 48-68.
  15. HAINAUT, JL Y LECHARLIER, B. Un modelo semántico extensible de base de datos y su lenguaje de datos. Proc. Congreso IFIP 1974, North-Holland Pub. Co., Amsterdam, págs. 1026-1030.
  16. HAMMER, MM, Y MCLEOD, DJ Integridad semántica en un sistema de base de datos de relaciones. Proc. Very Large Data Base Conf., Framingham, Mass., Septiembre de 1975, págs. 25-47.
  17. LINDGREEN, P. Las operaciones básicas sobre la información son una base para el diseño de bases de datos. Proc. IFIP Congreso 1974, North-Holland Pub. Co., Amsterdam, págs. 993-997.
  18. MEALY, GH Otra mirada a la base de datos. Proc. AFIPS 1967 FJCC, vol. 31, AFIPS Presione, Montvale, Nueva Jersey, págs. 525-534.
  19. NIJSEN, GM Estructuración de datos en el DDL y el modelo relacional. En la gestión de la base de datos ment, JW Klimbie y KL Koffeman, Eds., North-Holland Pub. Co., Amsterdam, 1974, págs. 363-379.
  20. OLLE, TW Tendencias actuales y futuras de los sistemas de gestión de bases de datos. Proc. IFIP Congress 1974, North-Holland Pub. Co., Amsterdam, págs. 998-1006.
  21. ROUSSOPOULOS, N., Y MYLOPOULOS, J. Uso de redes semánticas para la gestión de bases de datos. Proc. Very Large Data Base Conf., Framingham, Mass., Septiembre de 1975, págs. 144-172.
  22. RUSTIN, R. (Ed.). Proc. ACM-SOGMOD 1974-debate sobre modelos de datos. Ann Arbor, Michigan, Mayo de 1974.
  23. SCHMID, HA, Y SWENSON, JR Sobre la semántica del modelo relacional. Proc. ACM-SIGMOD 1975, Conferencia, San José, California, mayo de 1975, págs. 211-233.
  24. SENKO, ME Lenguaje de descripción de datos en el concepto de descripción estructurada multinivel: DIAM II con FORAL. En Descripción de la base de datos, BCM Dougue y GM Nijssen, Eds., Pub de Holanda Septentrional. Co., Amsterdam, págs. 239-258.
- ACM Transacciona en sistemas de bases de datos, vol. 1, N ° 1, Maroh 1976.

---

## Página 28

36

-

PP-S. Chen

25. SENKO, ME, ALTMAN, EB, ASTRAHAN, MM Y FEHDER, PL Estructuras de datos y

acceso en sistemas de base de datos. IBM Syst. J. 12, 1 (1973), 30-93.

26. SIBLEY, EH Sobre la equivalencia de los sistemas de bases de datos. Proc. Debate ACM-SIGMOD 1974 sobre modelos de datos, Ann Arbor, Michigan, mayo de 1974, págs. 43-76.

27. STEEL, TB Normalización de la base de datos: informe de situación. Proc. ACM-SIGMOD 1975, Conference, San José, California, mayo de 1975, págs. 65-78.

28. STONEBRAXER, M. Implementación de restricciones de integridad y vistas mediante modificación de consultas. Proc. ACM-SIGMOD 1975, Conferencia, San José, California, mayo de 1975, págs. 65-78.

29. SUNDGREN, B. Fundamento conceptual del enfoque infológico de las bases de datos. En datos Base Management, JW Klimbie y KL Koffeman, Eds., North-Holland Pub. Co., Amsterdam, 1974, págs. 61-95.

30. TAYLOR, RW Observaciones sobre los atributos de conjuntos de bases de datos. En la descripción de la base de datos, BCM Douge y GM Nijssen, Eds., North-Holland Pub. Co., Amsterdam, págs. 73-84.

31. TSICHRITZIS, D. Un marco de red para la implementación de relaciones. En la descripción de la base de datos, BCM Douge y GM Nijssen, Eds., North-Holland Pub. Co., Amsterdam, págs. 269-282