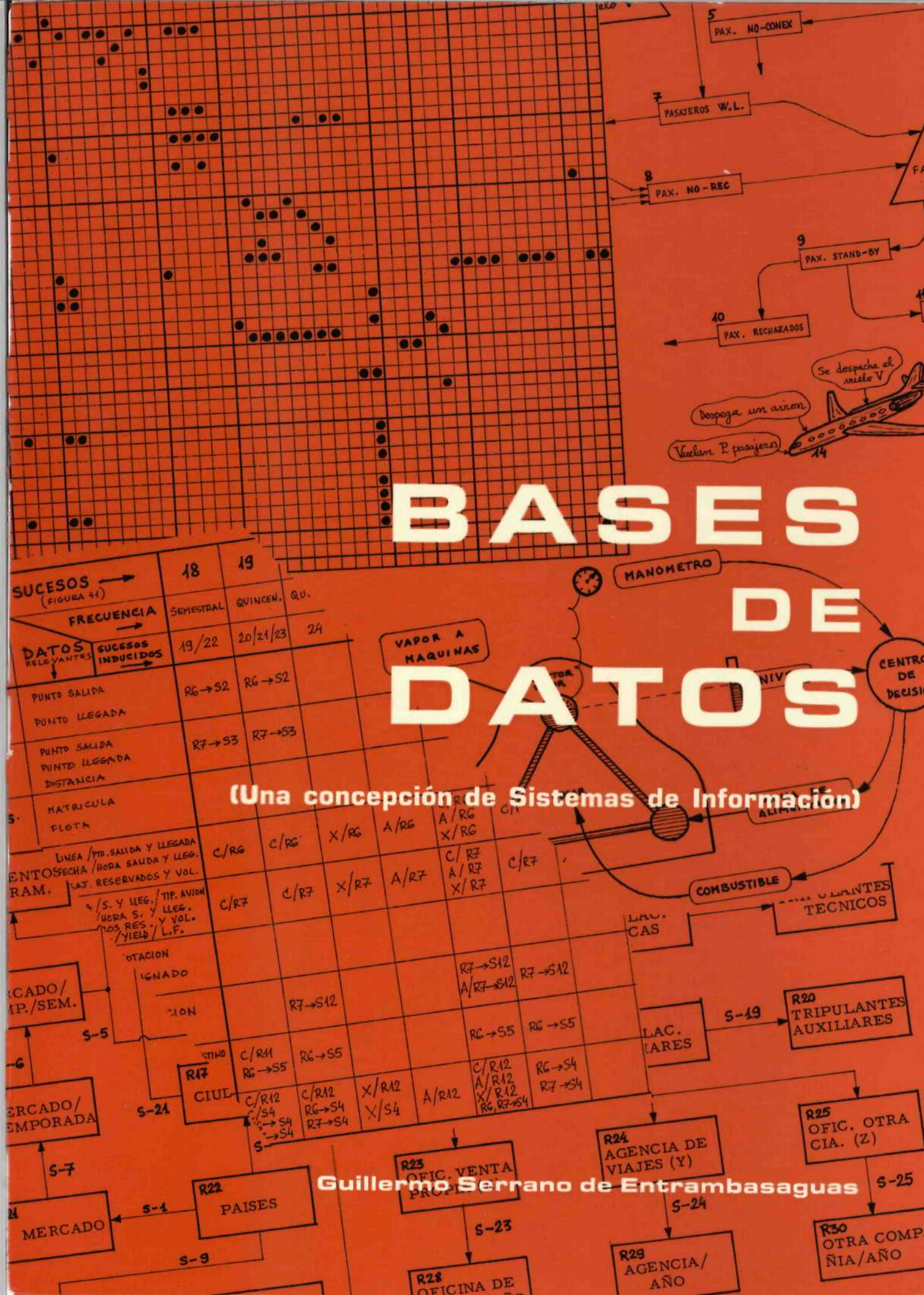


BASES DE DATOS

(Una concepción de Sistemas de Información)



SUCCESOS (FIGURA 41)	FRECUCENCIA	
	SEMESTRAL	QUINCEN. QU.
DATOS RELEVANTES	18	19
SUCCESOS INDUCIDOS	19/22	20/21/23 24
PUNTO SALIDA	RG→S2	RG→S2
PUNTO LLEGADA		
PUNTO SALIDA	R7→S3	R7→S3
PUNTO LLEGADA		
DISTANCIA		
MATRICULA FLOTA		

DATOS RELEVANTES	SUCCESOS			
	C/RG	C/RG	X/RG	A/RG
LINEA /PTD. SALIDA Y LLEGADA	C/RG	C/RG	X/RG	A/RG
FECHA /HORA SALIDA Y LLEG.	C/RG	C/RG	X/RG	A/RG
SAT. RESERVADOS Y VOL.	C/RG	C/RG	X/RG	A/RG
TIP. AVION	C/RG	C/RG	X/RG	A/RG
YIELD	C/RG	C/RG	X/RG	A/RG
OTACION				
IGNADO				
OTACION				

Guillermo Serrano de Entrambasaguas

BASES DE DATOS

(Una concepción de Sistemas de Información)

BASES DE DATOS

(Una concepción de Sistemas de Información)

por

Guillermo Serrano de Entrambasaguas

DOCTOR INGENIERO NAVAL
LICENCIADO EN INFORMÁTICA
PROFESOR AGREGADO DE ANÁLISIS DE SISTEMAS
DE GESTIÓN (4.º CURSO) EN EL INSTITUTO DE INFORMÁTICA

INSTITUTO DE INFORMATICA
MADRID, 1976

© SERVICIO DE PUBLICACIONES DEL MINISTERIO DE EDUCACION Y CIENCIA

Edita: Servicio de Publicaciones del Ministerio de Educación y Ciencia
Imprime: IMNASA. Menorca, 47. Madrid
Depósito Legal: M. 41.283-1976
I.S.B.N.: 84-369-0100-2
Impreso en España - Printed in Spain

*Mi agradecimiento a
Concepción Sauras Olivas,
sin cuya colaboración no hubiera sido
posible la realización del libro.*

INDICE

	<i>Págs.</i>
INTRODUCCION	13
1. LOS DATOS	17
1.1. LA RELACION DATOS/ACTIVIDAD	19
1.2. CAPTACION DEL DATO	23
1.2.1. Sucesos relevantes	24
1.2.2. Consecuencias relevantes de los sucesos	26
1.2.3. Tiempo de captación de los datos	33
1.2.4. Forma de obtener los datos	34
1.2.5. Soporte de los datos	38
1.2.6. Transmisión de los datos	40
1.3. EL CONTROL DE LOS DATOS	44
1.3.1. Control de producción del dato	46
1.3.2. Control de forma	47
1.3.3. Control de verosimilitud	48
1.3.4. Control de integración	49
1.3.5. Control de utilización	51
1.3.6. Control de mantenimiento	56
1.4. LAS ESTRUCTURAS DE DATOS	57
1.4.1. El registro, los grupos y los campos	59
1.4.2. Lista simple o cadena	60
1.4.3. Lista compuesta	61
1.4.4. Lista doble	63
1.4.5. Anillos	65
1.4.6. Arboles	68
1.4.7. Redes	70
1.4.8. El Set	72
1.4.9. El Esquema	73
1.5. CODIFICACION DE LOS DATOS	76
1.5.1. Tipos de códigos	78
1.5.2. Características de los códigos	79
1.5.3. Control de los códigos	82
1.5.4. Criterios para la elección de códigos	84

	<i>Págs.</i>
1.6. DEFINICION DE LA BASE DE DATOS	85
1.6.1. Identificación y características de los datos	88
1.6.2. Definición de la Estructura lógica	89
1.6.3. Definición del Almacenamiento físico	91
1.6.4. Método propuesto por Codasyl	91
1.6.5. Método propuesto por Guide/Share	94
1.7. UTILIZACION DE LOS DATOS	96
1.7.1. Usuarios informáticos	97
1.7.2. Usuarios no informáticos	97
1.7.3. La figura del Administrador	98
1.8. ADMINISTRACION DE LOS DATOS	98
1.8.1. La vida del Dato	99
1.8.2. Funciones de Administración	103
1.8.3. Dimensionado de la Base de Datos	105
1.8.4. Documentación de la Base de Datos	106
1.8.5. Imputación de Costes	106
2. LA EMPRESA COMO SISTEMA	109
2.1. LA ACTIVIDAD	112
2.1.1. El Sistema Productivo	114
2.1.2. Entidades componentes y sus características informáticas	120
2.1.3. Propiedades del Sistema Productivo	122
2.1.4. Umbrales de Actividad	123
2.1.5. Entorno del Sistema productivo	125
2.2. LOS SUCESOS	125
2.2.1. Relación Sistema Productivo/Producto	126
2.2.2. Integración de Sucesos Homogéneos	127
2.2.3. Redes de Sucesos	129
2.2.4. Relevancia de Sucesos	129
2.3. LA RED DE PRODUCCION	130
2.3.1. Flujo de Productos	134
2.3.2. Relación Sucesos Registros	139
2.3.3. Relación entre Sucesos y Relaciones lógicas entre Registros	145
2.3.4. Constitución de la Red de Producción	149
2.4. LA RED DE ADMINISTRACION	151
2.4.1. Flujos de Recursos	154
2.4.2. Inventario y Administración de Recursos	157
2.4.3. Bases para una Contabilidad Analítica	158

	<i>Págs.</i>
2.5. LA RED DE INFORMACION	159
2.5.1. La Red de Información. Flujos de mensajes	161
2.5.2. Redes Primaria y Secundaria	163
2.5.3. Matriz Estructural de la Red de Información	164
3. METODOLOGIA PARA EL DISEÑO DE UNA BASE DE DATOS	171
3.1. PLANTEAMIENTO DE OBJETIVOS Y MEDIOS	174
3.1.1. Lanzamiento del Proyecto	176
3.1.2. Análisis de Objetivos	179
3.1.3. Asignación de Recursos Materiales y Humanos	181
3.2. EL ESTUDIO DE LA ACTIVIDAD	181
3.2.1. Estudio de Sistemas Productivos	182
3.2.2. Delimitación de Entornos de Actividad	185
3.2.3. Aislamiento de Sucesos Relevantes	186
3.3. ESTUDIO DE LA RED DE PRODUCCION	190
3.3.1. La Matriz de Sucesos	193
3.3.2. Diagramas de Sucesos y análisis de la Red de Producción	196
3.3.3. Derivaciones e Integraciones de flujos	198
3.4. DISEÑO DEL ESQUEMA	199
3.4.1. Diseño de Registros	204
3.4.2. Diseño de Relaciones lógicas	211
3.4.3. Dimensionado previo del Almacenamiento de Datos	215
3.5. ESTUDIO DE LA RELACION DATOS/ACTIVIDAD	218
3.5.1. Expresión gráfica de la relación Datos/Actividad	218
3.5.2. Estudio de la dinámica de Registros	220
3.5.3. Influencia de los Sucesos y Captación de los Datos	222
3.5.4. Especificaciones de Acceso a los Datos	223
3.5.5. Reconsideración del Diseño del Esquema	224
3.6. ESPECIFICACIONES DE LOS DATOS	225
3.6.1. Definición de Registros	226
3.6.2. Definición de Sets	229
3.6.3. Definición de Areas Lógicas de Datos	231
4. CRITERIOS GENERALES DE DISEÑO DE LA BASE DE DATOS	235
4.1. EL CONCEPTO DE TIEMPO PRESENTE	238
4.2. UNICIDAD DE LOS DATOS	240

	<i>Págs.</i>
4.3. LA ORIENTACION MORFOLOGICA	242
4.4. CONOCIMIENTO DE LA ACTIVIDAD	244
4.5. LA DIMENSION DE FUTURO	245
4.6. LA RENTABILIDAD	246
5. EL SISTEMA DBMS DE CODASYL	249
5.1. OBJETIVOS DEL SISTEMA DBMS	252
5.2. DESCRIPCION GENERAL DEL SISTEMA DBMS	254
5.3. DDL O LENGUAJE DE DEFINICION DE DATOS	255
5.3.1. Entrada de Esquemas	257
5.3.2. Entrada de Area	258
5.3.3. Entrada de Registro	261
5.3.4. Entrada de Set	265
5.4. EL DML O LENGUAJE DE UTILIZACION DE LA BASE DE DATOS	269
5.4.1. Instrucciones DML en la Identification Division	270
5.4.2. Instrucciones DML en la Data Division	272
5.4.3. Instrucciones DML en la Procedure Division	273
6. BIBLIOGRAFIA	285

INTRODUCCION

Los Ordenadores se están aplicando en forma creciente a las diferentes áreas funcionales de las Organizaciones, y el conjunto de Datos que "pasan" por un Centro de Cálculo, constituye todo un Activo de Información que debe de rentabilizarse, como parte importante que es, dentro del inventario de recursos productivos.

La concepción de Base de Datos, contempla este Activo de Información con entidad propia, es decir, con independencia de las aplicaciones y programas informáticos que lo van a utilizar. Los Datos son extraídos de la dinámica productiva de la Organización, y representan a diferentes Sucesos que tuvieron lugar y a sus características y atributos. La Base de Datos, permite que este conjunto de Información, sea tratado como un Sistema, relacionando los Datos en una forma representativa de la conexión que dichos Datos en la realidad tienen, lo que nos permite configurar un Modelo de la actividad de la Organización, Modelo que representa el disponer en forma operativa procesable, de una potencia de correlación con la actividad real.

La orientación que se ha pretendido dar a esta obra, es la de incorporar la concepción de Bases de Datos a la metodología de Análisis de Sistemas, y su contenido en consecuencia, sigue una línea funcional. Como complemento se incluye en el último capítulo, un extracto de las Especificaciones propuestas por la Conferencia CODASYL en su versión de abril del 71.

Debo de advertir, especialmente a mis alumnos, que la presente publicación cubre solamente un ángulo de la materia "Bases de Datos", debiendo complementar el conocimiento con Bibliografía adicional y con los Manuales técnicos de los principales "paquetes" de gestión de Bases de Datos que existen en el mercado. Especial atención requieren las materias de Estructura de Datos y Recuperación de Información, sobre las que no se ha tratado en profundidad en esta publicación, porque por sí mismas son merecedoras de tratamiento exclusivo. Confío que la obra que estoy preparando sobre el particular, pueda ofrecer una alternativa para este complemento necesario.

1. LOS DATOS

1.1. LA RELACION DATOS/ACTIVIDAD

La Informática no es una disciplina nueva, como tampoco lo es la Teoría General de Sistemas. Sin embargo, es evidente que estas materias han dejado atrás el estado de reposo en el que se encontraban para inscribirse en un plano de máxima actualidad.

La capacidad de recopilar datos y de relacionarlos y procesarlos, era muy limitada antes de la aparición del ordenador, y únicamente en pequeños entornos, generalmente de intelectuales, se manejaban los criterios informáticos y la concepción de procesos en base a la Teoría de Sistemas. Solamente el hombre dotado de un nivel cultural elevado y con gran madurez mental, era capaz de “suplantar” al ordenador actual desarrollando su potencia analógica, y en consecuencia, las aportaciones de la Teoría de Sistemas sólo se contemplaban en círculos “ocultistas”.

Hoy en día el desarrollo tecnológico de la electrónica ha actuado de catalizador, y lo que antes pertenecía al mundo de las culturas minoritarias, se ha convertido hoy en soporte epistemológico que ampara a las ciencias aplicadas y al desarrollo de las organizaciones.

Allí donde existe un fenómeno productivo, donde se aplican recursos para lograr fines, o donde dicho de otras palabras existe *Actividad*, hay un entorno donde pueden ser aplicados los criterios informáticos, porque donde hay actividad hay también situaciones cambiantes y, por lo tanto, hay producción de *Datos*. Los datos vienen a constituir los puntos de referencia del desarrollo de la actividad, y en la misma forma en que la sucesión de los puntos luminosos determinados por los faros de un coche nos dibujan su trayectoria, la sucesión de *Datos* nos proporciona una película de la *Actividad*.

Para analizar más despacio esta idea, vamos a poner a título de ejemplo la actividad febril de un enfermo, es decir, la actividad de intercambio de calor del flujo sanguíneo de un enfermo con el entorno de sus canales circulatorios.

En la figura 1, se puede ver que la sucesión de puntos nos determina los valores del parámetro temperatura en los diferentes momentos en que dicha temperatura fue tomada. Este parámetro constituye el dato, dato en singular, porque en este caso se trata de un parámetro único, y el valor del tiempo que cada dato lleva asociado, constituye una característica muy importante del dato que es su nivel de actualidad o su coordenada histórica. Si el tiempo correspondiente fuera un día y hora perteneciente al futuro, el dato constituiría una previsión, previsión que lógicamente será posteriormente confirmada o desautorizada.

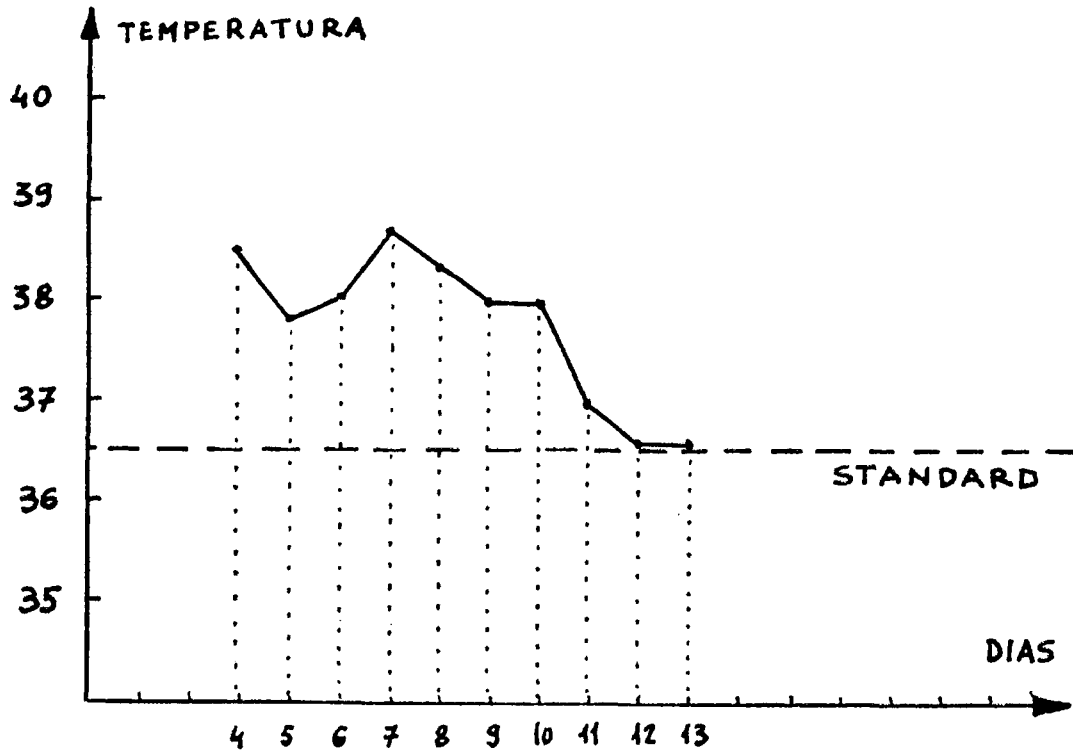


FIG. 1

Cabe distinguir dos tiempos diferentes:

- La coordenada temporal del fenómeno.
- La coordenada temporal del observador.

La primera coordenada es consustancial con la actividad y su valor no es dependiente de ningún factor, pues pertenece a la realidad pragmática. Por el contrario, la coordenada temporal del observador es subjetiva y su valor está asociado al momento en que el dato es contemplado.

Tengamos por lo tanto en cuenta la síntesis expresada en el siguiente cuadro.

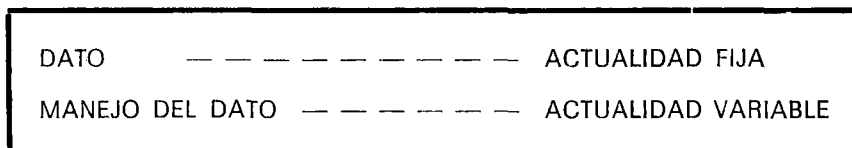


FIG. 2

Aunque la actualidad del dato sea un valor único, sin embargo no siempre es fácilmente determinable, o dicho en viceversa, para una coordenada temporal dada no siempre es determinable el valor o los valores paramétricos de los datos. Ello, porque, para que un dato, que como hemos dicho pertenece al mundo de la realidad

pragmática, entre a formar parte de un archivo, ha de ser recogido o captado de alguna forma y en algún momento. Volviendo a nuestro ejemplo, el dato "temperatura" es recogido mediante un termómetro y aunque el tiempo al que corresponde el dato sea el correspondiente a la aplicación del termómetro al paciente, el momento de captación del dato es el momento en que se realiza la lectura y se registra su resultado. La calidad del dato captado dependerá de la calidad del termómetro, y la calidad y demora de la lectura y registro.

Por lo tanto, de cada dato elemental hay que considerar:

- a) su valor,
- b) su actualidad,
- c) la forma de captación,
- d) el tiempo de captación,
- e) su registro,
- f) el tiempo de utilización.

Si no es adecuadamente establecido u observado el debido equilibrio entre estas seis características, el dato no es de utilidad o lo que es peor aún, el dato puede desorientar en lugar de informar. Una sucesión de datos desequilibrados producen información de una actividad irreal.

Hechas estas consideraciones sobre los datos elementales, vamos a examinar las sucesiones de datos elementales o registros de actividades. Volviendo al ejemplo de la figura 1, la sucesión de registros de lecturas del termómetro que se representa en el gráfico, ha de proporcionar una imagen de la actividad febril del paciente, y además el conocimiento que obtenemos de esta actividad, ha de servir para determinadas conclusiones acerca de la salud del paciente.

Los puntos elementales de información son como los fotogramas de una película y hay obviamente varios factores de equilibrio a tener en cuenta:

- a) homogeneidad entre datos elementales;
- b) homogeneidad de intervalos de tiempo;
- c) magnitud de los intervalos.

La homogeneidad de datos elementales, es requisito generalmente necesario y el grado de homogeneidad preciso será dado en la medida en que sea necesario evitar distorsiones en el conocimiento de la actividad.

La homogeneidad de intervalos de tiempo ha de ser paralela a la homogeneidad del fenómeno que se trata de modelizar y aquí entramos en la problemática de las estructuras de datos y estructuras de actividades de las que más adelante se tratará.

La magnitud de los intervalos establece la riqueza del modelo. Intervalos muy grandes, proporcionarán modelos sencillos y simples de la actividad. Intervalos muy pequeños, producen, por el contrario un modelo más preciso, aunque es importante tener en cuenta que un modelo demasiado preciso, amén de ser más "caro", puede en ocasiones proporcionar demasiada información obteniéndose "conocimientos" demasiado ricos en matices que no son operativos por ser caóticos. Volviendo al caso

de la medición de la temperatura de un enfermo, si se toman dos o tres mediciones al día, el médico obtiene una información por lo general suficiente de la trayectoria de su fiebre, pudiendo asociar dicha información a otros datos clínicos que en conjunto aporten una indicación de estado general de salud y de la reacción del organismo frente al proceso patológico.

Sin embargo, cincuenta mediciones a lo largo del día producen un modelo en el que cabe analizar la evolución del ritmo biológico intrínseco de la temperatura de la sangre del paciente. Si el objetivo fuera el estudiar dicho ritmo, el intervalo quizás hubiera sido acertado, pero si se trataba únicamente de seguir el proceso febril, está claro que hemos desplegado recursos innecesarios y además podemos confundir al médico con un exceso de información.

En cuanto al punto *b*), en el ejemplo puede comprenderse que lo conveniente es que los intervalos guarden una distancia homogénea, es decir, las temperaturas de varios días deben corresponder a horarios análogos en los diferentes días para que los datos sean comparables. La condición apuntada en *a*) se traduce en nuestro ejemplo a que las temperaturas sean tomadas con medios análogos, es decir, termómetros clínicos y aplicados a las mismas zonas corporales.

Establecido un "dato tipo" a considerar, es preciso por lo tanto determinar y asegurar el grado de homogeneidad adecuado entre los datos elementales que pertenecen al patrón del dato fijo considerado. Es preciso también la definición razonada de la magnitud de los intervalos de tiempo a que corresponden los datos de la serie, teniendo en cuenta que dicha magnitud nos da la medida de "nitidez" en el conocimiento de la actividad. Podríamos decir que hay que hacer dicha magnitud tan grande como sea posible siempre que obtengamos un "enfoque" de la actividad suficientemente "nítido", ya que a mayor nitidez manejamos mayor abundancia de datos, lo que consecuentemente encarece la información.

El grado de "nitidez" debe guardar un equilibrio con el resto de los datos que pertenecen a la misma estructura y a los mismos procesos. Pongamos como símil que viene al caso, el de una película obtenida a base de unir fotogramas en secuencia. De nada sirve que unos fotogramas tengan gran calidad de detalle, si hay otros que están desenfocados, y otro cuyo grano de película es muy grueso y contienen una baja definición de la imagen. El nivel medio de la película será un nivel producto de la media estadística de calidad de los fotogramas, tendiendo a separarse hacia abajo de la media en una medida relacionada con la dispersión de calidades de los fotogramas.

Hay que tener por lo tanto siempre en cuenta, que los conjuntos de datos que se relacionan en una estructura de datos o que participan en un mismo proceso informático, forman sistema, y las características y circunstancias de unos datos afectan a la utilidad de los demás, en la misma forma que las características de los fotogramas individuales afectan a la calidad de la película.

La homogeneidad de intervalos de tiempo es condición indispensable para que los datos elementales constituyan una serie válida del "dato tipo" al que corresponden. Téngase en cuenta que la actividad es un proceso continuo, mientras que las series de datos son discontinuas. Se logrará componer una imagen de la actividad

en la medida en que el umbral de discontinuidad sea homogéneo en la serie de datos. Pongamos un ejemplo:

Los sonidos que un ser humano percibe, corresponden a variaciones en el medio de una magnitud superior a 2 decibelios (depende de la frecuencia y de la persona, por supuesto). Las señales que captamos con nuestros oídos nos proporcionan un conocimiento parcial y útil de la actividad que estamos escuchando (el conocimiento realmente se complementa con otras percepciones). Imaginemos el grado de confusión que experimentaríamos si el umbral nuestro de audición no fuera más o menos constante. Unas veces “conoceríamos” ruidos de 1 decibelio y otras veces sin saberlo, se nos escaparían sonidos de 20 decibelios. Nuestra imagen sónica de la actividad estaría totalmente deformada.

Hay que tener muy presente que los procesos informáticos modelizan de alguna manera la actividad de la organización, y que el único conocimiento de la actividad que manejan los procesos son las series de datos obtenidas. Debe de estar muy estudiada, por lo tanto, la relación datos/actividad, relación que es un parámetro de calidad que condiciona los procesos informáticos. Dicho nivel de calidad ha de situarse en una posición suficiente que constituye un compromiso entre los factores económicos y los de buena técnica informática.

1.2. CAPTACION DEL DATO

La captación del dato es de importancia capital para los sistemas informáticos en general, porque el “dato” es la materia prima que se utiliza para la producción de información elaborada.

Si los datos no son apropiados, o no son fiables, o no representan suficientemente la actividad, la información producida a partir de ellos será necesariamente defectuosa, aunque los algoritmos de procesos sean correctos.

En un sistema de Base de Datos, la importancia de la captación de datos es mayor si cabe, porque cualquier dato parcial que penetra en la Base afecta a todos aquéllos con los que se relaciona, porque de hecho, en la Base de Datos se integran los datos en una estructura sistemática en base al significado lógico de la información. El efecto multiplicador es mayor en la medida en que el sistema de datos es más perfecto y, por lo tanto, más estructurado.

El estudio de la captación del dato requiere de un análisis previo de la actividad y de un planteamiento previo de los objetivos que se persiguen con el Sistema de Base de Datos, ello para que el problema pueda fragmentarse y abordarse con seguridad. El estudio puede hacerse por partes una vez que está claro el marco general en el que cada parte se inscribe y, por lo tanto, una vez conocida la incidencia de cada parte en otras partes del conjunto. Algo así como el etiquetado y numerado de piezas de un monumento que se quiere trasladar de localidad. No se puede abordar el problema del traslado individual de las piezas hasta que ha sido realizado el

trabajo previo de elaboración de un esquema director que identifica a las partes en el todo.

Dicho de otras palabras, antes de abordar el estudio de la captación de datos, la actividad ha sido examinada por completo (véase 3.2) y se dispone ya de un “Esquema” tentativo de la Base de Datos (véase 3.4). Sin embargo, del estudio de la captación de datos pueden surgir consecuencias que aconsejen incluso revisar el Esquema porque el diseño de una Base de Datos se realiza, como cualquier otro proceso creativo, por un método de aproximaciones sucesivas que siempre permite ir profundizando con mayor seguridad. Los ciclos de aproximación serán mayores en la medida que la Base de Datos esté más estructurada.

1.2.1. SUCESOS RELEVANTES

La actividad es un fenómeno dinámico en el continuo espacio-tiempo que determina una serie de modificaciones en los estados de las materias primas, productos, personas, máquinas, dinero, etc., que participan en el proceso productivo de una Organización.

Esta dinámica de cambio es de naturaleza continua y de acuerdo con las consideraciones que hicimos en el planteamiento de la relación datos/actividad, tenemos que sintetizarla en una serie discontinua de sucesos que nos proporcionen un modelo de la actividad suficiente y compatible con los objetivos de la Base de Datos y con el Esquema director.

Igual que la serie de números naturales, puede constituir un modelo informático del conjunto de números reales, debemos buscar un conjunto de sucesos que constituyan un modelo de la actividad.

Esta serie de sucesos es el conjunto de sucesos que tienen lugar y que son relevantes para el Sistema de Base de Datos.

Aunque más adelante (2.3), veremos con más profundidad cómo se realiza el estudio de la actividad, cómo se analizan los sucesos, y cómo se determina el Esquema director de la base de datos, vamos ahora a analizar un ejemplo que nos permita centrar el problema de la captación de datos, porque en ella radica el éxito de la Base de Datos.

Pongamos como ejemplo la actividad de aprovisionamiento de un almacén en una determinada empresa, actividad que ha sido fragmentada del conjunto general de actividades empresariales, y de la que se supone hemos de extraer un sistema de datos que permita alimentar una Base de Datos cuyo Esquema (véase 3.4) parcial, se representa en la figura 3.

Este Esquema es una parte del Esquema general de la Base de Datos, y determina el conjunto de registros tipo que van a almacenarse en la base, así como las relaciones entre ellos. El conjunto de campos que contendrá cada registro está ya definido de una forma tentativa, porque en realidad, y siempre con un criterio de aproximaciones sucesivas, fue ya examinada superficialmente la actividad a la hora de tantear los registros tipo que componen el esquema.

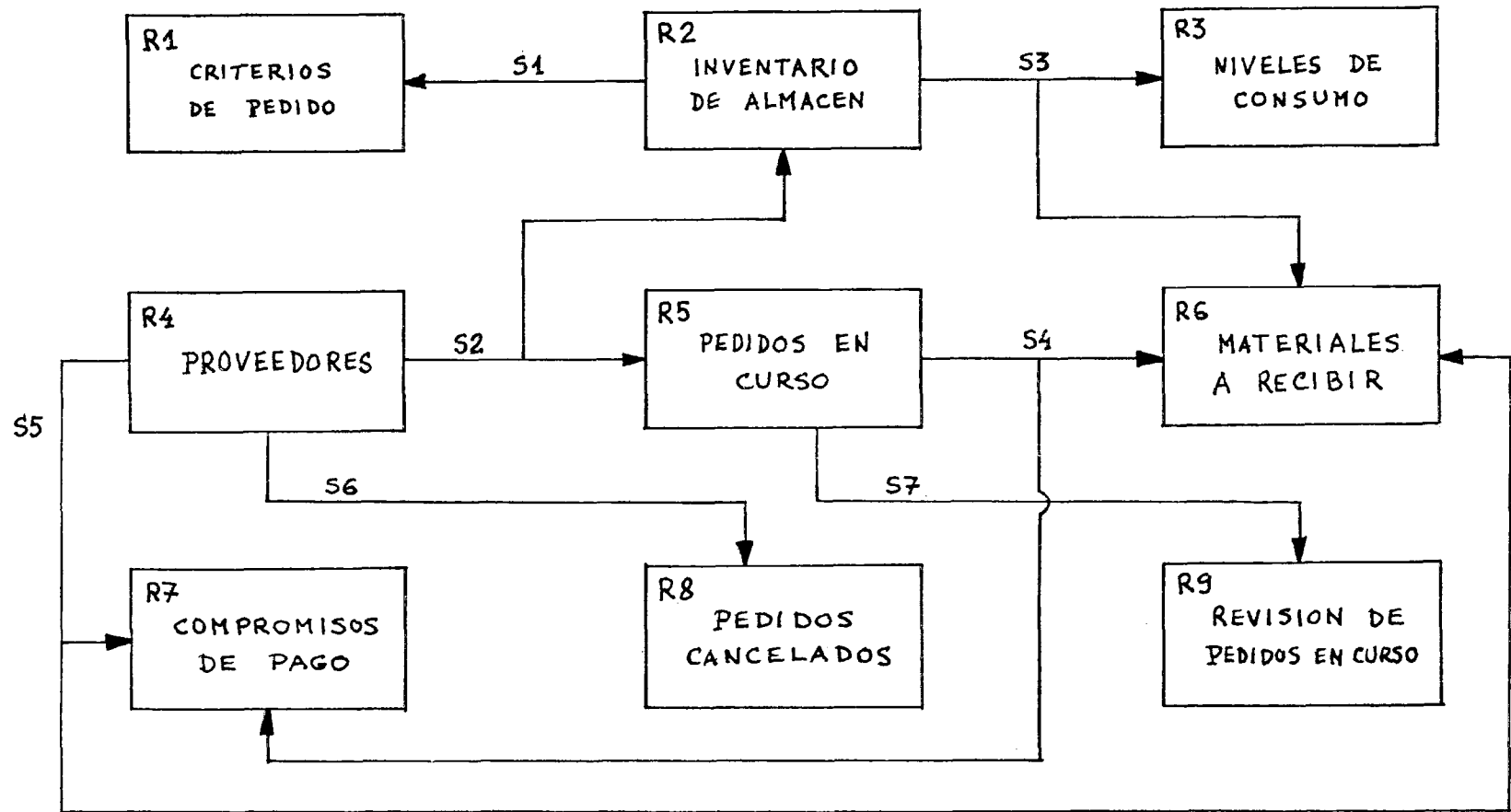


Fig. 3

El proceso continuo por medio del cual se hace posible el aprovisionamiento del almacén, puede modelizarse por el siguiente conjunto de sucesos relevantes:

- Lanzamiento de pedidos a proveedores.
- Reclamaciones de pedidos a proveedores.
- Anulación de pedidos a proveedores.
- Modificación de pedidos.
- Recepción de material pedido.
- Devolución de materiales a proveedor.
- Realización de pagos a proveedores.
- Liquidación de compromisos de pago.
- Altas en el inventario del Almacén.
- Calificación de comportamiento de proveedores.
- Clasificación de un nuevo proveedor.
- Eliminación de un proveedor.
- Determinación de criterios de pedido.

Este conjunto de sucesos se determina después de un repaso completo de la función de aprovisionamiento del almacén y de su actividad, y a la vista de la información que se produce durante el ejercicio de la actividad, teniendo presente las necesidades de información de la Base de Datos dentro del marco del Esquema e incluso en un marco más amplio si es que se vislumbra la conveniencia de ampliar el Esquema.

Existen también otros sucesos que no son relevantes a los objetivos de la Base de Datos y en este caso no se reseñan, aunque hay que tener presente que un proyecto de Base de Datos no ha de responder solamente a objetivos presentes, sino que ha de perfilarse con vocación de futuro. Así en el ejemplo no se ha reseñado como objetivo relevante el suceso “estibaje de materiales recibidos” por estimar que no es relevante. Sin embargo, si la naturaleza del almacén hiciera sospechar que en el futuro pudiera existir un sistema automático de estibaje y recuperación de los productos almacenados, sí deberíamos tener en cuenta dicho suceso para que la situación física de los productos almacenados, fuera un factor a utilizar en el futuro sin que la Base de Datos necesite de grandes modificaciones.

En la figura 4 se recoge un gráfico que muestra el flujo de los sucesos antes relacionados (véase 3.3.2) gráfico que es el resultado del análisis de las interrelaciones entre los sucesos. La figura 5 recoge la matriz de relación de los sucesos (véase 3.3.1).

1.2.2. CONSECUENCIAS RELEVANTES DE LOS SUCESOS

Una vez examinados los sucesos que tienen lugar en el ejercicio continuo de la actividad y que son relevantes para los objetivos que tenemos planteados para la Base de Datos, hay que examinar cada suceso en particular.

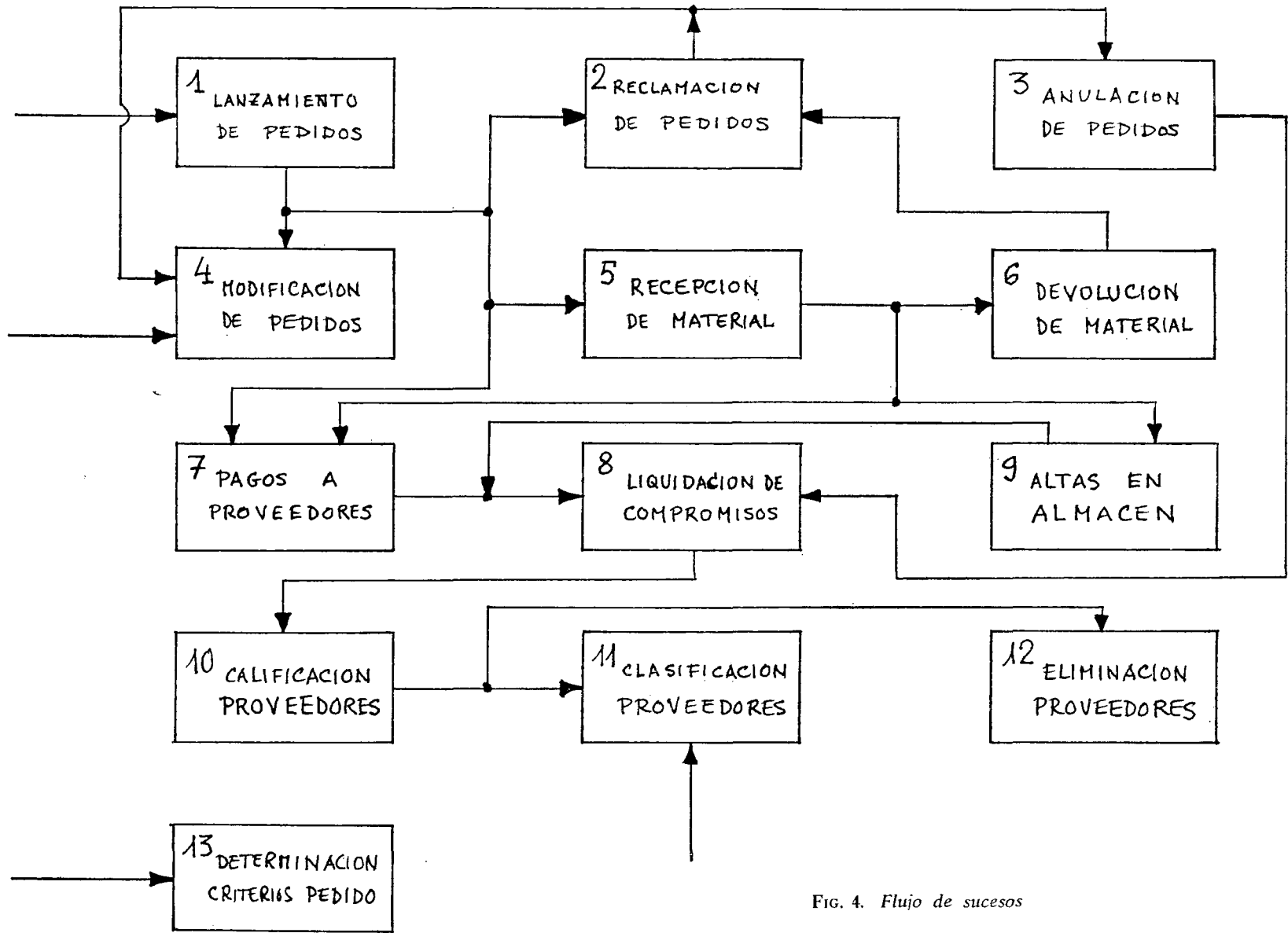


FIG. 4. Flujo de sucesos

	SUCESOS	1	2	3	4	5	6	7	8	9	10	11	12	13
1	LANZAMIENTO PEDIDOS		X		X	X		X						
2	RECLAMACION PEDIDOS			X	X									
3	ANULACION DE PEDIDOS								X					
4	MODIFICACION DE PEDIDOS													
5	RECEPCION DE MATERIAL						X	X		X				
6	DEVOLUCION DE MATERIAL		X											
7	PAGOS A PROVEEDORES								X					
8	LIQUIDACION COMPROMISOS										X			
9	ALTAS EN ALMACEN								X					
10	CALIFICACION PROVEEDORES											X	X	
11	CLASIFICACION PROVEEDOR													
12	ELIMINACION PROVEEDOR													
13	CRITERIOS PEDIDO													

Fig. 5. Matriz de sucesos

Cada suceso es un acontecimiento informático, y como hemos seleccionado los sucesos relevantes, podemos estar seguros que cada suceso afecta a la Base de Datos. Hay que analizar, por lo tanto, las relaciones de cada uno de los sucesos con la estructura de datos de la base, e incluso con los demás sucesos.

En el cuadro de la figura 6, se recoge una forma sistemática de sintetizar dicho análisis, método que no pretende ser único, pero que puede ayudar a realizar el análisis de los sucesos a quien no disponga de otra disciplina más adaptada a su estilo. No hay que olvidar que las metodologías no constituyen doctrina en sí mismas y, por tanto, las normas ofrecidas son siempre opciones que se proponen como recomendables por defecto a falta de un esquema propio, esquema que sólo se acaba destilando con la experiencia continuada del ejercicio de la profesión.

En el cuadro, reseñamos en la parte izquierda los “registros tipo” de la parte del “Esquema” que está afectada por la actividad en estudio, en este caso, la actividad de aprovisionamiento de un supuesto almacén. En la parte superior encabezamos las columnas con los sucesos que previamente hemos diferenciado como relevantes, sucesos que encadenados en una determinada secuencia geográfico-temporal, vienen en definitiva a modelizar el desarrollo de la actividad en estudio.

Para cada uno de los sucesos, hemos de contemplar en primer lugar la frecuencia con que se producen para ir ya tomando conciencia del nivel de actividad que dicho suceso va a representar para la dinámica de evolución de la Base de Datos. Un suceso con periodicidad de alta frecuencia, es decir, un suceso que se produce de forma continua con intervalos de tiempo pequeños, será de más relevancia para la Base de Datos que un suceso cuya frecuencia de producción es pequeña, ello en cuanto al simple análisis de frecuencias, porque de hecho puede haber sucesos esporádicos de frecuencia baja que sean, sin embargo, verdaderos umbrales para la evolución de la Base de Datos.

Un suceso tal como el “lanzamiento de pedidos” es realmente relevante para la actividad de “aprovisionamiento de almacenes” que estamos analizando en el cuadro, porque la función aprovisionamiento se realiza con base en el lanzamiento de pedidos. Estas consideraciones deben siempre estar presentes para que a lo largo del desarrollo del proyecto, se vaya adquiriendo un conocimiento “analógico” del sistema, conocimiento que es imprescindible para que el proyecto sea una concepción coherente que pueda ser enriquecida en su evolución y para que la Base de Datos sea un “sistema” y no simplemente una agrupación de ficheros de datos.

El suceso “lanzamiento de pedidos” viene a “cristalizar” un conjunto de actividades que se producen de manera continua, y dicho suceso, a la vista de una determinada naturaleza del almacén, supongamos tenga lugar con periodicidad diaria. Reseñamos, por lo tanto, la indicación “diaria” en la casilla que vemos en el cuadro destinada a la frecuencia del suceso. También en este momento podemos informarnos de forma aproximada del volumen de pedidos que se realizan a diario por término medio, porque de esta forma vamos obteniendo un conocimiento de las dimensiones y volúmenes de datos que se manejarán almacenándolos o procesándolos, de manera que a la hora de estudiar a fondo el dimensionado de la base de datos, realicemos un trabajo más perfecto con un criterio director ya casi formado.

Tengamos en cuenta que para una mentalidad informática madura, una imagen

REGISTROS	SUCESOS	LANZAM. PEDIDOS	RECLAM. PEDIDOS	ANULAC. PEDIDOS	MODIFIC. PEDIDOS	RECEP. MATERIAL	DEVOLUC. MATERIAL	PAGOS A PROVEED.	LIQUIDAC. COMPROM.	ALTA EN ALMACEN	CALIFIC. PROVEED.	CLASIFIC. PROVEED.	ELIMINAC. PROVEED.	D. CRITER. PEDIDO
		DIARIA	SEMANAL	OCCASIONAL	OCCASIONAL	DIARIA	OCCASIONAL	SEMANAL	SEMANAL	DIARIA	DIARIA	OCCASIONAL	OCCASIONAL	MENSUAL
	FRECUENCIA	2,4,5,7	3,4	8	6,7,9	2	8	10	11,12	8	10	12		
	DATOS	1	2	3	4	5	6	7	8	9	10	11	12	13
R1 CRITERIOS DE PEDIDO	CODIGO ARTICULO FORMA DE DECISION PUNTO DE PEDIDO PUNTOS HISTORICOS	R6→S3								A/R2 C/R2 C/S1,S3		R2→S2		A/R1 C/R1
R2 INVENTARIO	COD. ARTICULO / DESCRIP. CANTIDAD EN STOCK VALOR MONETARIO STOCKS MEDIOS HISTORICOS													
R3 NIVELES DE CONSUMO	COD. ARTICULO CONSUMO MENSUAL CONSUMOS ACUMULADOS CONSUMOS HISTORICOS													
R4 PROVEEDORES	PROVEEDOR / DIRECCION CALIFICACION DATOS HISTORICOS VOLUMENES DINERO	R5→S2 R6→S5 R7→S5		R8→S6							A/R4	C/R4 C/S2,S5 C/S6	X/R4	
R5 PEDIDOS EN CURSO	Nº PEDIDO / PROVEEDOR ARTICULOS PEDIDOS / INFORM. RECEP. / CANT. / PAGOS FECHAS / PRECIOS	C/R5 C/S4,S7 R6→S4 R7→S4		X/R5 X/S4	R9→S7	A/R5			X/R5					
R6 MATERIALES A RECIBIR	ARTICULOS / Nº PEDIDO CANTIDAD / PROVEEDOR FECHA CONTRATADA FECHA ACTUALIZADA	C/R6	A/R6	X/R6	A/R6	X/R6								
R7 COMPROMISOS DE PAGO	Nº PEDIDO / FECHA CONDICION DE PAGO CANTIDAD A PAGAR CANTIDAD REMANENTE	C/R7	A/R7	X/R7	A/R7	A/R7		X/R7	X/R7					
R8 PEDIDOS CANCELADOS	Nº PEDIDO / FECHA PROVEEDOR FECHA CANCELACION MOTIVO CANCELACION			C/R8										
R9 REVISION PEDIDOS	Nº DE PEDIDO FECHA REVISION ACTUALIZACION MAT. ACTUALIZACION FECHAS				C/R9				X/R9					

FIG. 6 Actividad aprovisionamiento

de la periodicidad de un suceso y del volumen de datos producidos, proporciona un conocimiento analógico aproximado del nivel de costes de almacenamiento y de proceso, y permite tener presente el criterio de rentabilidad a la hora de considerar las relevancias de la información y las utilidades que reporta.

En la casilla inmediatamente inferior, donde figura como leyenda “sucesos consecuentes”, reseñamos los números de los sucesos que serán o podrán ser inducidos directamente por cada suceso en particular.

Para ello, como puede verse, hemos numerado previamente los sucesos de acuerdo con las figuras 4 y 5. Se recoge aquí una síntesis de la matriz de sucesos para tener presente las relaciones de causalidad entre sucesos que son circunstancias determinantes de la causalidad entre los datos.

Para el suceso 1 se han reseñado como sucesos consecuentes los sucesos 2, 4, 5 y 7 que pretenden expresar que el hecho de que se lancen pedidos supone el que a consecuencia de ello se produzcan posteriormente reclamaciones (no tiene por qué suceder siempre), se produzcan modificaciones (el “pedido” como cualquier otra naturaleza está sujeta a posibles cambios), se reciban materiales de los proveedores y se realicen pagos en contraprestación.

Estos sucesos 2, 4, 5 y 7 son los sucesos inmediatamente consecuentes, porque a su vez éstos inducen otros sucesos, como puede verse en la figura 4 con más detalle.

Continuando con el cuadro de la figura 6, en la parte izquierda y a la derecha de cada encabezamiento de “registro”, reseñamos los campos de datos de cada uno de los registros que sean relevantes en relación con los sucesos. De esta forma tendremos presentes las relaciones de sucesos y registros a un mayor nivel de análisis.

En la columna correspondiente al suceso 1, se relacionan todas las “incidencias” que tendrán lugar en la Base de Datos como consecuencia de haberse producido el suceso 1. Vamos a relacionarlas a continuación, y para mejor comprensión, téngase a la vista el gráfico de la figura 3.

- C/R5, S4, S7. La producción del suceso 1 tiene como consecuencia la creación de un nuevo registro del tipo R5. Ello a su vez supone la creación de un set (véase Esquema) del tipo S4 y de otro del tipo S7. Esto se reseña en la posición correspondiente al registro R5 por afectar a dicho tipo de registros.
- C/R7. Tiene lugar la creación de nuevos registros del tipo R7 por generarse compromisos de pago a proveedores. Pueden ser más de uno.
- C/R6. Tiene lugar la creación de nuevos registros del tipo R6, porque el lanzamiento de un pedido implica una posterior recepción de materiales. Pueden ser más de uno.
- R5 → S2. El nuevo registro R5 creado por haberse producido un nuevo pedido, ha de ser insertado en el set del tipo S2 que corresponde al proveedor que recibe el pedido. Ello se reseña a la altura de los registros R4, por ser los registros R4 los correspondientes a los proveedores y los registros principales de los Set tipo S2.

- R6 → S5. El nuevo registro R6 creado ha de ser insertado en el set tipo S5 correspondiente al proveedor en cuestión. El set S5 nos agrupa los materiales pendientes de recibir de cada proveedor.
- R7 → S5. Inserción del nuevo registro R7 en el set tipo S5 que corresponda, análogamente a la operación anterior.
- R6 → S4. Inserción del registro R6 creado en el set S4 también de nueva creación. Esta es la relación lógica entre el registro de pedido y los registros que contienen los materiales a recibir.
- R7 → S4. Inserción del registro o registros tipo R7 creados en el set tipo S4 que encabeza el registro R4 del proveedor receptor del pedido. En este set agrupamos de esta forma todos los compromisos de pago que tenemos con cada proveedor.
- R6 → S3. Inserción de los registros tipo R6 creados en los set S3 correspondientes a los ítems de almacén que tenemos inventariados identificando los tipos de materiales a recibir. De esta forma, para cada ítem de almacén tenemos agrupado el conjunto de remesas pendientes de recibir.

Análogamente, se relacionan las operaciones que tienen lugar en la Base de Datos como consecuencia de los demás sucesos, operaciones que se reseñan en las correspondientes columnas según sucesos generadores, y en las correspondientes filas, según cuáles sean los registros afectados.

Así por ejemplo, en la posición relativa al suceso 8 y al registro R5, se reseña X/R5 que expresa que la liquidación de un compromiso (suceso 8) trae como consecuencia la cancelación de un registro de pedido en curso (registro R5).

También como ejemplo, en la posición determinada por el suceso 4 y el registro tipo R7, se indica A/R7 que expresa que la modificación de un pedido (suceso 4), trae consigo una posible actualización de los registros tipo R7 que recogen los compromisos de pago generados por ese pedido, por haber podido variar dichos compromisos.

De esta forma, para cada suceso se relacionan todas las operaciones consecuentes al suceso que suponen alteración de la base de datos.

Igualmente, en la fila correspondiente a cada tipo de registro, se recogen todas las operaciones que tendrán lugar en la familia de registros de ese tipo, con diferenciación de los sucesos que las producen.

Así, por ejemplo, para los registros tipo R6, que constituyen los registros de materiales pendientes de recibir, pueden producirse las siguientes operaciones:

- C/R6. Creación de los registros motivada por el suceso 1.
- A/R6. Actualización de los registros tipo R6 que puede ser producida por el suceso 2 (actualización fecha de recepción), o por el suceso 4 (alteración del pedido).
- X/R6. Cancelación de registros que puede tener lugar como consecuencia

del suceso 3 por anularse el pedido o como consecuencia del suceso 5, por recibirse ya el material.

El examen de este cuadro de la figura 6, nos proporciona el conocimiento de las “consecuencias informáticas” de la actividad empresarial, lo que nos permite un cuidadoso examen de cómo ha de orientarse la captación de datos, porque estamos contemplando los datos, los sucesos que los producen, sus consecuencias y las interrelaciones entre los propios sucesos y entre datos y sucesos.

1.2.3. TIEMPO DE CAPTACION DE LOS DATOS

El tiempo de captación del dato nos va a condicionar de forma directa el grado de actualidad de la información que tenemos en la Base de Datos. Al examinar la relación entre datos y actividad (1.1), veíamos que los datos en sí corresponden a la realidad pragmática y, por tanto, su actualidad es fija. Sin embargo, hay otra actualidad operativa condicionada por el manejo del dato, cuyo valor es variable. Este valor dependiente del manejo del dato, viene condicionado por el tiempo de captación.

El tiempo de captación puede ciertamente oscilar en una gama continua que parte del valor actual de producción. Si la captación se realiza en el momento en que el dato se produce, estaremos captando el dato en tiempo real. Como puede suponerse, aun cuando la captación en tiempo real nos permite ajustarnos con bastante fidelidad a la actualidad real del dato, se trata de un procedimiento caro, por lo que lo utilizaremos en caso necesario cuando ello sea compatible con los criterios de rentabilidad.

El tiempo de captación debe entonces considerarse a la vista de los procesos que vayan a utilizar el dato, y a la vista de los objetivos presentes y futuros que el sistema de base de datos deba de cumplir. Se trata por lo tanto de un estudio que habrá que revisar varias veces según avanza la realización del proyecto, e incluso estando el sistema ya en operación.

Podríamos establecer como criterios generales a la hora de considerar el tiempo de captación:

- TIEMPO SUFICIENTE. Quiere decir que no sobrepasemos un umbral de coste si ello es innecesario o simplemente no es conveniente por criterios de rentabilidad. Dicho de otra forma, limitémonos a una “calidad” o grado de actualidad suficiente en compatibilidad con los objetivos generales y particulares del sistema de Base de Datos, llevando, sin embargo, la calidad al límite superior posible sin que ello suponga aumentos sensibles de coste. Un aumento sensible sería un umbral.

Por ejemplo, siguiendo con el cuadro de la figura 6, la captación de los datos de los registros R5, R6 y R7 creados como consecuencia del suceso 1, se ha de hacer con tiempos suficientes para que los sucesos 2, 4, 5 y 7 no resulten condicionados ni afectada su consecuencia informática.

Asimismo, otros sucesos externos a esta zona de la base, como pudiera

- ser la subbase de datos que agrupa las actividades financieras, no deben ser condicionados por una actualidad insuficiente de los registros R7.
- **CONDICIONANTES TECNICOS.** Las circunstancias en que se producen los sucesos, determinan en ocasiones, condicionantes técnicos que deben ser tenidos en cuenta. Por ejemplo, el parte de ruta de un vehículo de transporte, no podrá ser obtenido hasta que el vehículo rinda viaje o pase por el punto de control, a menos, por supuesto, que se vayan reportando las posiciones o posibles incidencias vía radio.
 - **CONDICIONANTES ECONOMICOS.** El criterio de rentabilidad siempre ha de estar presente, como ocurre en todo proyecto.
 - **CONDICIONANTES PROPIOS DE LA ACTIVIDAD.** Debe de evitarse una anteposición de los objetivos informáticos a los propios de la actividad. Es decir, la actividad no debe resultar afectada de forma relevante por el hecho de que tengamos que recoger datos. Por lo menos ello no debe de ocurrir innecesariamente. Este es un error que se produce frecuentemente.
 - **EQUILIBRIO DE LA ESTRUCTURA DE DATOS.** Dado que los datos están relacionados entre sí, ha de guardarse un equilibrio en lo que se refiere a los valores de actualidad de los datos. Para ello hay que examinar varias veces a lo largo del proyecto para cada tipo de registro, el paquete de operaciones que tiene lugar con los registros de la familia atendiendo al grado de equilibrio proyectado en los tiempos de captación. Ello puede realizarse a la vista del cuadro de la figura 6 y repasando las soluciones de captación de datos que se hayan dispuesto. Es obligado también tener presente el proyecto de aplicaciones que van a utilizar los datos y las informaciones que habrán de producirse.

1.2.4. FORMA DE OBTENER LOS DATOS

La obtención de los datos casi siempre supone una perturbación para el ejercicio de la actividad, porque la captación del dato casi nunca se realiza con la “limpieza” con que el fotógrafo obtiene una placa. Quiere ello decir que la captación del dato es de hecho una nueva actividad que hemos de integrar en toda la cadena de sucesos.

La actividad bajo estudio, resultará afectada por la nueva actividad de captación de datos en la medida en que ambas estén interrelacionadas. Si la actividad de captación de datos pudiera diseñarse con sucesos independientes de los estudiados en la matriz de sucesos, la actividad permanecería inalterada.

Realmente, en muchas ocasiones, la captación de datos invita a reconsiderar todo el planteamiento de la actividad, y bien puede suceder que la alteración de los sucesos sea, por lo tanto, una acción deliberada.

Podemos por lo tanto distinguir 3 casos diferentes:

- a) *Obtención de datos como subproducto de la actividad.* Pongamos por ejemplo el caso del registro de las entradas y salidas de almacén, o el de

cualquier otra actividad de este tipo que supone el que una persona registre las entradas y salidas utilizando una máquina. Si esta máquina fuera capaz de trasladar la información de registro que introduce la persona a un soporte procesable, tal como banda de papel perforada o cassette, estaríamos recogiendo los datos de la actividad como un subproducto, sin necesidad de alterar básicamente el trabajo y procedimiento de las personas.

Por supuesto, el subproducto procesable obtenido, recoge la información con un tiempo de captación que hay que situarlo en el tiempo de registro de la entrada o salida del almacén, teniendo en cuenta que dicho soporte de información, todavía ha de ser trasladado al Centro de Cálculo, o a un terminal de entrada de datos, donde habrá de pasar por los oportunos controles (véase 1.3), antes de que los datos que contienen, pasen a integrarse en la base.

- b) *Obtención de datos como una actividad independiente.* Pongamos por ejemplo la obtención de los datos correspondiente a los pedidos de materiales. Si una vez analizadas las circunstancias, nos decidimos por recoger los datos de pedidos de materiales a proveedores recogiendo simplemente una copia del documento que hacemos llegar al proveedor, el proceso de captación del dato, podrá plantearse prácticamente como una actividad independiente de la del lanzamiento de pedidos, permaneciendo ésta inalterada.

Determinada la forma de obtener los datos como una actividad independiente, debe diseñarse todo el proceso consecuente de transmisión de información, transcripción, control, depuración, ... etc., hasta que los datos pasan a integrarse a la Base de Datos.

- c) *Obtención de los datos mediante una actividad condicionante.* La naturaleza de los datos que se van a obtener, puede requerir que la actividad generadora de los datos sea replanteada, resultando por lo tanto el proceso de obtención de datos una actividad condicionante.

Ello es muy frecuente que suceda, generalmente trae consigo una integración o un desdoblamiento de sucesos, la aparición de sucesos nuevos que antes no se tenían en cuenta porque eran irrelevantes para los anteriores objetivos, la modificación de documentos y la alteración de los procedimientos en ocasiones puede ser incluso drástica, especialmente cuando en el trabajo se introduce la utilización de terminales de teleproceso, o se utiliza el ordenador en tiempo real permitiendo realizar operaciones que previamente eran inasequibles.

En estos casos, la actividad generadora resulta totalmente modificada, siendo preciso replantearla de nuevo integrado en la nueva actividad el mecanismo de obtención de los datos.

Debe de tenerse en cuenta, como citábamos en un punto anterior, que al replantear una actividad, no deben de predominar los objetivos informáticos sobre los demás objetivos, evitando un error muy común que conduce a una captación de datos muy buena desde el punto de

vista informático, pero a un ejercicio de la actividad malo desde el punto de vista empresarial, bien porque resulte una imagen mala de servicio ante un posible público, una ejecución de trabajos con mal aprovechamiento de la plantilla, o cualquier otra circunstancia que a medio plazo tiene también repercusión informática, deteriorándose el sistema.

Por ello, siempre que la actividad generadora resulte condicionada de forma sensible, el nuevo diseño de métodos y procedimientos debe de ser realizado en colaboración con los responsables orgánicos de la actividad empresarial afectada, siendo también deseable que el analista de sistemas conozca la actividad empresarial bajo estudio, o al menos esté debidamente informado de sus aspectos más fundamentales.

En ocasiones, puede resultar conveniente que intervengan técnicos especializados en organización del trabajo porque puede ser conveniente un estudio de organización previo al nuevo planteamiento informático. Con esta precaución, podemos asegurarnos mejor que el nuevo planteamiento de la actividad ha sido contemplado con una óptica objetiva y, por lo tanto, sin deformar por las necesidades técnicas.

Independientemente de estas consideraciones, a la hora de plantear la forma en que vamos a obtener los datos, deberán tenerse en cuenta los siguientes criterios:

- *Condicionantes técnicos.* Las circunstancias de orden técnico en las que el dato ha de ser obtenido, suponen un condicionante cuya incidencia debe ser examinada. Debe de analizarse si las transformaciones consecuentes de los datos son viables de cara a obtener determinados niveles de calidad. Debe examinarse si las condiciones en las que se produce el dato inciden en la calidad final de dicho dato como producto.

Para estos efectos hay que suponer que la “entrega” del dato a la Base de Datos es el producto terminado de todo el “proceso” de fabricación. Dicho proceso comienza con la obtención del dato, siendo la forma de obtención un procedimiento de fabricación que necesariamente ha de condicionar al producto final.

- *Condicionantes económicos.* El criterio de rentabilidad siempre ha de estar presente, y es el que básicamente nos va a centrar de toda la gama de posibilidades informáticas, en qué niveles de calidad vamos a situarnos. Es decir, si se va a recoger datos en tiempo real, si vamos a instalar terminales de teleproceso, si la entrada de datos en la base va a ser “batch” ... etc.

La rentabilidad debe ser examinada aquí con criterios amplios y de acuerdo con los objetivos generales presentes y futuros de la Base de Datos. El concepto de “umbral” debe ser tenido en cuenta, porque hay soluciones que son mejorables de cara al futuro en etapas sucesivas, y por el contrario, hay otro tipo de soluciones que admiten difícilmente el perfeccionamiento.

- *Condicionantes geográficos.* El dato se produce bien en un punto geográfico determinado, o bien en todo un conjunto de localidades fijas o

variables. Si se ha contemplado el tiempo real, el teleproceso, "time sharing", o cualquier otro procedimiento que requiera comunicación directa con el ordenador, es preciso establecer todo un planteamiento de una red de comunicaciones. Como necesariamente el criterio de rentabilidad estará presente, lo más probable es que el conjunto de puntos geográficos donde se producen los datos, sea segregado, disponiéndose diferente tipo de soluciones para diferentes localidades. Habrá puntos geográficos que justifiquen disponer líneas telefónicas de alta velocidad, pero sin embargo, habrá otros puntos cuyo volumen de datos y cuya relevancia en el conjunto general, aconsejen disponer soluciones más convencionales.

Es importante advertir que para un mismo tipo de dato puede haber por lo tanto diferentes formas de captación dependiendo del lugar donde se produce, e incluso puede ello traer consigo el que la actividad sea ejercida de forma diferente en diferentes localidades.

- *Tiempos de proceso.* Como se ha apuntado antes, desde que el dato se produce hasta que el dato es entregado a la Base, tiene lugar todo un proceso que puede ser contemplado como un proceso de "fabricación". Durante este proceso existen unas latencias de tiempo que suponen necesariamente unas demoras, por lo que la Base de Datos "tarda" un determinado tiempo en disponer del dato desde que el dato fue producido. Hay que distinguir éste tiempo del tiempo de captación que fue previamente analizado, aunque necesariamente ambos tiempos están ligados y debe de existir un ritmo determinado entre sus valores.

De nada sirve un proceso muy ágil en llevar los datos a la Base, si el tiempo de captación es elevado y los datos corresponden por lo tanto a una actividad que tuvo lugar con mucha anterioridad.

Los tiempos de proceso y los tiempos de captación son valores que condicionan de forma muy directa las utilidades del sistema de la Base de Datos en lo que a la utilización de los datos, cuya captación está en estudio, se refieren. Por ello, debe de existir un "equilibrio" entre las demoras que introducen estos tiempos y los niveles de calidad que se consideran suficientes para los objetivos de la Base de Datos.

- *Frecuencias y volúmenes de producción de datos.* La frecuencia de producción de los datos condiciona la frecuencia de actualización de la Base de Datos. Por otra parte, la frecuencia de producción de datos determina un conocimiento "rítmico" de la realidad del ejercicio de la actividad empresarial. Si la frecuencia de producción de datos tiene un ritmo muy lento e intervalos homogéneos, hay que distinguir si ello corresponde realmente a la naturaleza de la actividad, o bien el "modelo" que por los datos se obtiene, sólo ha tenido en cuenta los primeros "armónicos" componentes de un "ritmo" más complejo.

La actividad debe de estar por lo tanto modelizada con una frecuencia de producción de datos adecuada y consecuente con los objetivos que se persiguen, para que la dinámica de actualizaciones de registros de la Base

de Datos produzca una actividad de cambios en la base que tenga un cierto paralelismo necesario con la actividad real.

La frecuencia de producción de datos es complementada con los volúmenes de datos producidos, y si la frecuencia la hemos asociado al ritmo, el volumen de datos es la intensidad del ritmo.

1.2.5. SOPORTE DE LOS DATOS

Desde el momento en que se producen los datos se dispone de una “materia prima” que habrá de ser “elaborada” para conseguir el producto final que es la actualización de la Base de Datos.

La obtención de los datos se materializa en un determinado “soporte” informático que es el vehículo en el que los datos están contenidos. A partir de ahí, puede suceder que ese mismo vehículo sirva para llevar los datos a la Base, o bien se producen transformaciones de soporte y, en consecuencia, transcripciones de los datos.

Siguiendo la analogía que hemos casi perfilado, según la cual la producción de datos es una generación de “ondas informáticas”, pudiéramos profundizar contemplando el “soporte” informático como la onda portadora de información, estando contenida en esta onda una modulación de frecuencia e intensidad constituida por los datos contenidos en el soporte.

Como onda portadora que es, el soporte informático es perturbado por el medio en el que se maneja, pudiendo producirse pérdidas, deterioros, ... etc. El número de veces que se cambia de soporte informático incide también en el proceso por poder tener lugar pérdidas en la transcripción.

Los soportes informáticos son con seguridad conocidos por el lector: documento manuscrito, ficha perforada, ficha magnética, cinta perforada, banda magnética, cassette, línea de transmisión, ... etc., por lo que no vamos a describir sus características. La integridad de los soportes ha de contemplarse en el sistema porque supone condición necesaria para la integridad de los datos.

Vamos a señalar algunos criterios a tener en cuenta a la hora de pensar en los soportes a utilizar:

- *Criterio geográfico.* Como complemento de los aspectos de rentabilidad y de lejanía ya contemplados, hemos de fijarnos en las condiciones físicas en las que el soporte va a desplazarse. Si se trata de una línea de transmisión, tenemos que pensar las calidades de transmisión de datos que cabe esperar de esa zona geográfica, en función de otras experiencias o del tipo de garantías que nos ofrezca el servicio telefónico, determinando en cualquier caso los tipos de pruebas que habremos de realizar en previsión.

Si el soporte es un documento escrito o perforado, debe igualmente estudiarse su transporte desde el punto de vista físico para asegurar el flujo de llegadas de información con los tiempos y volúmenes previstos y sin que los soportes se deterioren.

Si el soporte es de material magnético, tal como cintas magnéticas, discos magnéticos, cassettes, deben analizarse cuidadosamente la posible “agresión” del medio ambiente. A título anecdótico, cabe citar los controles de seguridad que se practican hoy en día en los aeropuertos, porque un detector de metales puede ser causa del deterioro de los datos contenidos en la cinta magnética.

- *Preparación de los datos.* El soporte primitivo puede no ser directamente procesable, por lo que puede exigir determinadas preparaciones previas para obtener un soporte que pueda ser “leído” por el ordenador.

Toda preparación de datos supone una transcripción, una manipulación, y generalmente, un cambio de soporte. Estas operaciones han de ser contempladas desde el principio, aunque sólo lleguen a determinarse en detalle por aproximaciones sucesivas en la etapa final del proyecto, porque estas transcripciones pueden afectar a las tasas de calidad del dato y a las latencias de tiempo de proceso.

- *Introducción de datos en el sistema.* Determinado el soporte de los datos y el flujo de producción de datos con su frecuencia y con su volumen, hay que plantear la operación última de “entrega” de los datos a la Base. Supone establecer el procedimiento de actualización, procedimiento que ha de estar integrado en el plan general de explotación de la Base de Datos, porque supone de hecho una actividad concreta dentro del plan de explotación.

Dependiendo del soporte que va a ser utilizado como vehículo, la introducción de datos en el sistema será planteada de una u otra forma. Si los datos vienen por línea de transmisión, la actualización de la Base se realizará en la forma y tiempos asignados para atender a los mensajes que proceden de esa línea de comunicación. Si la captación se realiza en tiempo real, la línea de comunicación estará permanentemente abierta y será el “manejador” de comunicaciones, el procedimiento que determina el proceso de actualización.

Si los datos están soportados por cinta magnética, fichas, ... etc., la actualización de los registros de la Base se incluirá dentro del conjunto de actividades de mantenimiento, asegurándonos que los datos llegan a la Base con las frecuencias y latencias de tiempos adecuados.

- *Corrección de errores.* El hecho de que existan soportes de datos, y dichos soportes hayan de ser manipulados y preparados, supone el que necesariamente va a existir una fuente de errores, porque incluso cuando se opera en tiempo real, las líneas de comunicación, también incurren en errores de transmisión.

Ello quiere decir que debe de contemplarse los errores que pueden producirse y las tasas de volumen de errores que pueden llegar a alcanzarse, contemplando desde el momento en que se estudia la captación de datos el tipo de procedimientos que van a disponerse para la corrección de errores, estudiando la forma de arreglarlos, y el flujo de vuelta al origen que ha de seguir la información errónea para que sea depurada.

La depuración del dato puede, en ocasiones, requerir la intervención de las personas o secciones de la Empresa que produjeron dicho dato, por lo que es preciso contemplar todo un nuevo ciclo informático, no pudiendo introducirse el dato en la Base hasta que haya pasado de nuevo por la fuente de producción, y vuelva a llegar al ordenador una vez preparado en el soporte adecuado. Ello tiene incidencia directa en las latencias de tiempo de captación y proceso, y dependiendo del volumen de errores, puede degradarse por completo el procedimiento de captación del dato, o bien simplemente, resultar afectado por un factor de calidad o de rendimiento. En muchos casos, pueden establecerse determinadas redundancias en la captación de información que sirvan de compensación a la degradación de la calidad que hemos apuntado

En otras ocasiones, no es preciso que sea la propia organización productora del dato la que tenga que corregirlo, pudiendo disponerse de soluciones alternativas que pueden ser más convenientes. Ello suele hacerse en los sistemas de tiempo real donde los errores que se producen pueden ser debidos a deficiencias en la transmisión que son fácilmente identificables por un experto, sin que el dato que se comunica resulte sensiblemente afectado, con lo que se agiliza el proceso de corrección, independizándolo de la zona de producción del dato, al menos en una componente importante.

Los flujos y tiempos de corrección de errores, tienen una importancia muy superior a la que generalmente se le concede y en muchas ocasiones, hemos podido conocer sistemas informáticos de mucha calidad, que luego en la práctica fracasan porque su explotación se degrada por las circunstancias que apuntamos.

- *Criterio económico.* El criterio de rentabilidad siempre ha de estar presente y su aplicación es obvia, y no es preciso comentarla.

1.2.6. TRANSMISION DE LOS DATOS

Una vez que los datos son captados, la información que hemos recogido acerca de la actividad empresarial a la que dichos datos corresponden, es independiente de la actividad en sí. Estos datos que nacieron del desarrollo de la actividad, se han desligado de ella, y ahora su papel es el de incorporarse a la Base de Datos para que ésta disponga de un conocimiento “suficiente” de la actividad.

La figura número 7 trata de resaltar la etapa de transmisión del dato *desde su captación hasta su integración* en la Base de Datos. Durante esta etapa de transmisión del dato, se pueden producir las siguientes circunstancias:

- *Se cubre una distancia geográfica.* Se trata de la distancia que media del lugar de captación hasta el Centro de Cálculo donde los datos acabarán integrándose.
- *Transcurre el tiempo.* Es el tiempo de demora existente entre el momento en el que el dato se captó y el momento en que el dato pasa a integrarse a la Base de Datos.

- *Se producen manipulaciones.* Es decir, puede haber intervención humana en la etapa de transmisión del dato, interviniendo por lo tanto procedimientos llevados a cabo por el hombre.
- *Tienen lugar transcripciones.* Los datos pueden cambiar de soporte, es decir, pueden pasar de papel a cinta magnética, o bien puede haber transcripciones de unos documentos a otros, ... etc.

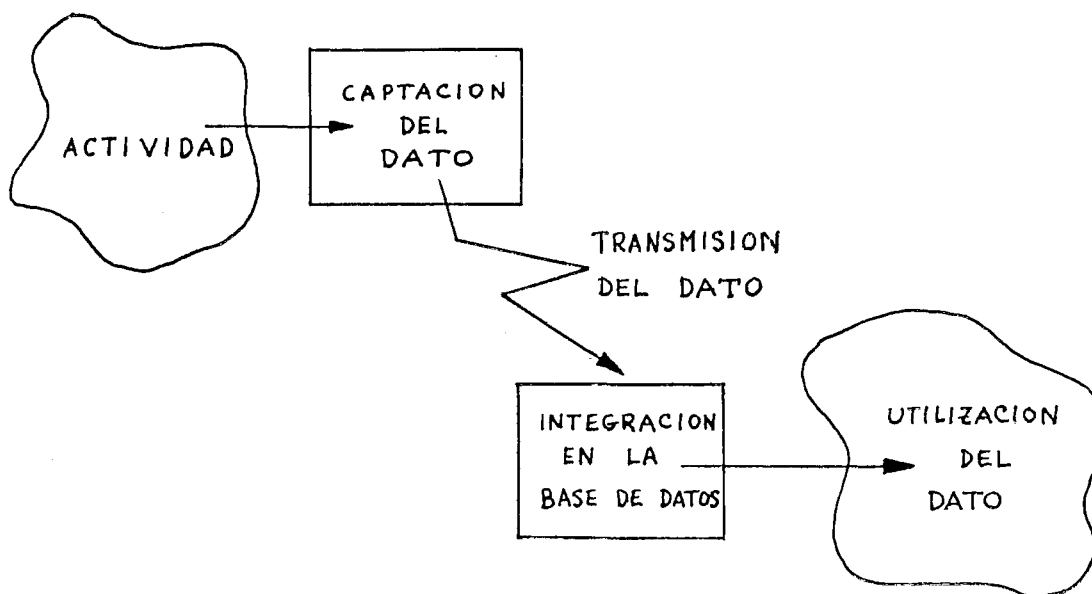


FIG. 7

- *Los datos pueden también ser procesados.* Además de los cambios de soporte, pueden tener lugar procesos, realizándose síntesis de datos, o determinadas elaboraciones previas a la entrada de los datos a la Base.
- *Ciclos de corrección de errores.* Los datos que una vez transmitidos son rechazados por la Base de Datos por no ser correctos, han de sufrir procesos de corrección de errores que introducen nuevas demoras, nuevas transcripciones e intervención humana.

Dentro de todas estas circunstancias, hay que tener además en cuenta que el fenómeno expresado en la figura 7 es dinámico, y la transmisión del dato se produce con unos determinados ciclos de frecuencia y con unos determinados volúmenes de información transmitida.

Hay que tener siempre presente que la Base de Datos constituye “un modelo” de la actividad y debe conocerse qué nivel de “actualidad” de información requiere el paquete de programas de utilización que van a operar con la Base de Datos. No puede darse ninguna regla de cómo ha de producirse la transmisión del dato ni con qué ritmos de tiempo, sino únicamente hay que asegurar que la transmisión del dato no debe de ser un condicionante de los niveles de eficacia y de calidad.

Los tiempos y las frecuencias de transmisión deben de tener unas determinadas cotas de proyecto, cotas cuyo límite superior está determinado por el umbral del nivel de "actualidad" de la Base de Datos que no debe traspasarse para no comprometer los objetivos generales de proyecto de utilización de la Base de Datos. Las cotas de tiempos y frecuencias de transmisión deben de situarse lo más alejadas posible del umbral indicado, en compatibilidad con un criterio general de rentabilidad.

Es decir, el tiempo de transmisión tiene un límite superior que no debe de traspasarse, y por otra parte, deberá ser el mínimo posible sin traspasar un límite inferior que pudiera constituir un umbral en el orden económico. Para mayor claridad, véase la figura 8.

Igualmente, la frecuencia de transmisión debe de ser lo más elevada posible. Si examinamos la figura 9, pudiéramos decir que la frecuencia de transmisión de los datos, debe de situarse entre los niveles *A* y *B*, correspondiendo el nivel *A* a un límite inferior que constituye un umbral que no debe traspasarse, pues de lo con-

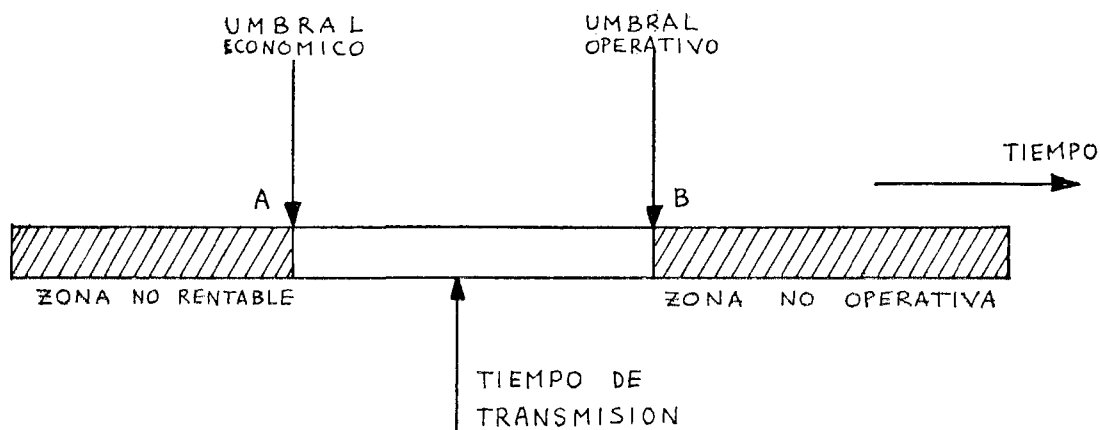


FIG. 8

trario estaríamos condicionando la calidad de la explotación de la Base de Datos. Por otra parte, si traspasamos el umbral *B*, que es un límite superior, podemos llegar a soluciones que en el orden económico comprometen la rentabilidad general del proyecto.

Obviamente, el tiempo real corresponde a valores mínimos del tiempo de transmisión (escala de la figura 8), y a valores altos de la frecuencia de transmisión (escala de la figura 9). En un proyecto en el que adoptáramos la captación de determinados datos en tiempo real, manteniéndonos dentro de un criterio de rentabilidad, querría decir que el límite inferior de la figura 8 y el límite superior de la figura 9, estarían prácticamente en el cero y en el infinito, respectivamente, lo cual querría decir que podríamos incurrir en soluciones de coste elevado, porque en este caso es muy importante para el proyecto que la Base de Datos tenga un nivel óptimo de actualidad con respecto a los datos que estamos considerando.

Si, por ejemplo, los datos cuyo tiempo de transmisión estamos contemplando, van a alimentar una Base de Datos, y dichos datos van a ser procesados a diario, el límite de un día es ciertamente un umbral para la frecuencia de transmisión, porque si estos datos se utilizan a diario, el ciclo de actualización de la Base en lo que se refiere a estos datos, no puede ser superior a un día. Si además, el proceso diario que se realiza requiriera que la información acerca de la actividad no fuera en ningún caso de una actualidad inferior a 48 horas, el valor de 48 horas sería un límite superior para el tiempo de transmisión de este tipo de datos.

El nivel de calidad de la explotación de la Base de Datos exige también determinados niveles de calidad de los datos que se integran en la Base, además, de los factores de eficacia y eficiencia antes apuntados. La calidad del dato, no debe por lo

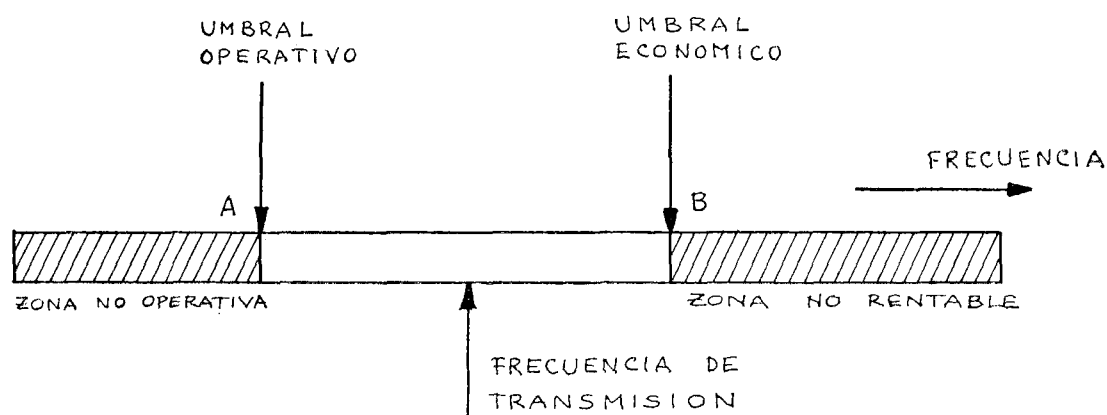


Fig. 9

tanto ser comprometida por la fase de transmisión del dato, y a estos efectos, hay que revisar la cadena de transcripciones, manipulaciones, procesos y ciclos de corrección, a los que el dato es sometido durante la etapa de transmisión.

En cuanto a los procesos, y dentro de estos procesos vamos a incluir las manipulaciones y transcripciones, hay que mantener como actitud mental la idea de que todos estos procesos son los eslabones que constituyen una cadena. Un extremo de la cadena arranca del ejercicio de la actividad y el primer eslabón será el proceso de captación del dato.

El otro extremo de la cadena corresponde a la explotación de la Base de Datos, y los últimos eslabones son los procesos de los datos en el ordenador.

En toda cadena basta que falle un eslabón para que la cadena se rompa, es decir, basta que la calidad de un proceso sea deficiente para que todo el sistema se venga abajo, aunque los demás procesos sean correctos. No tiene sentido entonces el tener unos "últimos eslabones" muy perfectos, pues se trata de procesos realizados en el ordenador, si "el final de la cadena", es decir, los objetivos finales de la Base, no se cubren por deficiencias en los procesos de transmisión de los datos.

Por lo tanto, las transcripciones, manipulaciones y procesos que se diseñan para la etapa de transmisión deben de estar diseñados con una gran seguridad.

Otro factor a tener en cuenta es el factor de “pérdida” o de “rendimiento” de la transmisión del dato, fenómeno que está presente en toda actividad real. Efectivamente, las circunstancias que concurren en la transmisión de los datos, distancia, tiempo, transcripciones, ... etc., son motivo suficiente para que en la transmisión de datos se produzcan determinadas pérdidas de información o deterioros de la misma. Este hecho debe de ser tenido en cuenta y para contrarrestarlo, la cantidad de información transmitida debe ser superior a la necesaria en la Base de Datos. Ello se materializa mediante el concepto de “redundancia”. La redundancia de información o información duplicada, suele disponerse en forma de “cuadros” de información, de información de control, “check-digits”, información de secuencia, “check-points”, sumas parciales, ... etc. Las redundancias habrán de ser mayores en tanto en cuanto la etapa de transmisión del dato tenga por su naturaleza mayores pérdidas.

1.3. EL CONTROL DE LOS DATOS

A lo largo de la obra se sigue como un concepto de base la afirmación de que un sistema de Base de Datos debe ser diseñado como una entidad viva que viene a suponer una “representación” simplificada y esquemática del ejercicio real de la actividad empresarial, al menos en aquellos aspectos que son más relevantes.

Quizás la diferencia fundamental de las explotaciones en concepto de Base de Datos de los sistemas convencionales estriba en que los archivos de información son modelos vivos de la actividad, dejando atrás anteriores concepciones de simple acumulación pasiva de información.

El sistema de Base de Datos es, por lo tanto, un sistema vivo. Su “subsistema locomotor” que regula los ciclos de explotación de la Base de Datos, está dirigida por un sistema operativo, y los resultados de esa actividad motora, se apoyan en todo el contenido de datos de la Base. En cuanto más rica sea la Base de Datos en información, y en cuanto más estructurada e interrelacionada esté dicha información, el sistema tendrá una mayor “conciencia” para permitir mayores utilidades y servir, por lo tanto, a superiores objetivos. Aunque entremos en un plano excesivamente abstracto, considero conveniente establecer una analogía entre el sistema de Base de Datos y un organismo vivo. En el organismo vivo, la información procedente del exterior, penetra en el sistema a través de unos organismos de entrada que son los sentidos, y la información acaba siendo almacenada a través de las redes neuronales en los dispositivos de almacenamiento que tienen naturaleza biológica. Existe un organismo cuyo papel interesa aquí destacar, que es la conciencia exterior, que es un filtro que protege al organismo viviente de influencias externas no deseables. Es decir, la conciencia exterior tiene unos criterios más o menos perfeccionados de los objetivos deseables para su esencia, y tamiza toda la información procedente del ex-

terior, contrastándola con dichos objetivos, y rechazando todos aquellos datos considerados como no deseables o perjudiciales. Solamente llega a ser almacenado en el interior, y solamente, por lo tanto, será utilizada en el futuro, aquella información que cuenta con el visto bueno de la conciencia exterior. Sin embargo, si los criterios de la conciencia exterior son defectuosos, el filtro deja pasar información perturbadora, información que acaba polarizando “sugestionando” los mecanismos internos de selectividad, resultando afectado el futuro vivencial del organismo.

En el sistema de Base de Datos, hay que situar igualmente los mecanismos adecuados de filtraje de información que preserven la dinámica de cambio de la Base de Datos de las influencias no deseables. Estos mecanismos de control, constituyen la “conciencia exterior” de la Base de Datos y en cuanto más perfeccionados sean los controles, más perfecta será la información que pasa a través de ellos.

Se ha establecido la anterior analogía, porque deseamos destacar el hecho de que la información en una Base de Datos está fuertemente interrelacionada, y forma un sistema integrado, y en este tipo de sistemas, pequeñas causas, pueden producir grandes efectos. Es más, en los sistemas complejos, la degradación siempre se produce por efectos de segundo orden, porque los efectos de primer orden, suelen ser detectados. Entonces, pequeñas deficiencias de la información que pasan a integrarse a la Base de Datos, puede producir grandes “envenenamientos”.

A lo largo del punto 1.3, vamos a ver los diferentes tipos de control que hay que practicar o que al menos deben tenerse siempre presentes en el diseño. Pertenecen a tres diferentes naturalezas:

- *Controles de producción.* Son los controles destinados a detectar si los datos llegan a la Base de acuerdo con el programa establecido, es decir, con la frecuencia adecuada, y volúmenes y períodos de tiempo válidos.
- *Controles de calidad.* Constituyen las pruebas de aceptación del dato y vienen a ser lo que en términos industriales se denomina “pasa/no pasa”. Vamos a considerar tres tipos diferentes: control de forma, control de verosimilitud y control de integración.
- *Controles de utilización y mantenimiento.* Son los controles de tipo permanente, internos de la Base de Datos, que garantizan por un lado que los datos van a continuar conservando durante su estancia en la Base los niveles necesarios de calidad, y por otro lado, van a asegurar que la utilización de los datos sea correctamente realizada. Si los controles de aceptación, los hemos comparado analógicamente con la conciencia exterior, los controles de utilización, los podíamos, en cierto modo, comparar con los códigos morales de conducta que en los seres humanos supone un control a nivel biológico.

Hay otro tipo de control que será descrito más adelante, que tiene ya una perspectiva futura que se sale incluso del ámbito de la Base de Datos y que corresponde al control del recurso información como una parte más del activo de la empresa. Son los controles de la vida de la información (véase punto 1.7).

1.3.1. CONTROL DE PRODUCCION DEL DATO

El control de la producción del dato, viene a ser un control de la captación de datos y un control de la transmisión, denunciando dichos procesos cuando el momento final de la integración de los datos a la Base, no se produce.

Es, básicamente, un control de la ausencia de información acerca de las actividades. Este control se ejerce, con el conocimiento de la cadencia o ritmo proyectado en la captación y transmisión de información. En la Base de Datos, integrados en la Biblioteca de programas, se disponen programas que manejen los datos de volúmenes de datos, frecuencias de datos y grados de actualidad indicados en estos valores mediante parámetros que puedan ser variados a lo largo de la vida del sistema, estableciéndose tolerancias igualmente susceptibles de variación.

Estos programas son llamados por el sistema operativo para su proceso de acuerdo con unas cadencias y con una frecuencia determinada en el plan de explotación de la Base de Datos, de acuerdo con un diseño que cuenta con el visto bueno del órgano administrador de la base de información.

Cuando estos programas se procesan, se detectan las lagunas de información y, por lo tanto, aquellas actividades de cuyo ejercicio no ha llegado información a la Base de Datos. El diagnóstico que estos programas producen, sirve para acudir a los procedimientos de captación del dato y subsiguiente transmisión a la Base, para detectar dónde se ha roto el “eslabón” y proceder a subsanar la deficiencia. Si la deficiencia acusada impide necesariamente el proceso de determinados programas de explotación, los oportunos mecanismos de alerta deben ser excitados, adoptándose los caminos alternativos de emergencia proyectados al efecto o en caso de no estar proyectados, serán adoptados en ese momento por los órganos de decisión de la explotación.

En algunos programas de explotación de altas frecuencias, como pueden ser procesos de tiempo real, time sharing, o teleproceso en general, los mecanismos de alerta deben funcionar de forma automática, pudiendo tener incluso diferentes niveles de alarma, según los niveles de falta de información que se registren.

Si el procedimiento de alarma no es automático, por no ser necesario o posible, la alarma puede simplemente requerir de una decisión humana para abortar o modificar oportunamente el curso de la explotación del sistema.

Este tipo de controles de producción del dato, se hace más necesario en la medida en que el sistema de Base de Datos está más integrado, es de mayor complejidad, y es, por lo tanto, en mayor medida, un modelo vivo de la actividad empresarial.

Pongamos a título de ejemplo, un sistema que regule la explotación de una gran flota de vehículos de transporte (taxis, camiones, autobuses de línea, etc.). El hecho fundamental que determina la actividad empresarial, es el desplazamiento de los vehículos y la carga de mercancías o de personas que transporta. El sistema puede estar concebido para que la explotación de los vehículos, se optimice en función de la posición geográfica y el estado de carga de los vehículos en cada momento. La actualización de la Base de Datos que “soporta” toda la explotación del sistema, puede estar realizada en base a unos soportes previamente establecidos para cada tipo de

vehículo: Por ejemplo, un vehículo puede reportar por mensaje vía radio, el momento en el que inicia un desplazamiento, indicando a dónde se dirige y la hora estimada de llegada. El sistema puede “decidir”, reclamar información acerca de la posición si transcurrido el tiempo estimado de llegada, más una determinada tolerancia, no ha reportado el vehículo el fin del viaje.

Si el vehículo permanece inactivo, puede convenirse una recapitulación periódica de su situación, revisando con una determinada cadencia el estado de disponibilidad de los vehículos, etc. Un sistema de este tipo, requiere un ejercicio de control de producción de datos, precisamente porque los datos de posición son en este caso, base fundamental de la explotación de la empresa.

Otra forma de ejercer el control de producción, puede ser en base a registros de datos en lugar de en base a programas. En lugar de existir una cadena de programas, que rastrean periódicamente el sistema para detectar las lagunas de información, pueden disponerse de unos registros especiales de información que vayan recogiendo “la historia” de las cadencias de frecuencias y volúmenes de los datos que llegan al sistema.

De esta forma, en vez de existir unos programas que se procesan periódicamente, que es una solución de mucha calidad pero muy cara, existen unos registros que pueden ser consultados por los programas, a la hora de utilizar los datos, dándonos estos registros el estado de actualidad y los volúmenes de actualizaciones, registrados últimamente en la información. Si los niveles no son los adecuados, se producirá entonces la excepción y se excitarán los oportunos niveles de alarma, requiriendo de la decisión humana si es preciso para decidir sobre la continuación de la ejecución de los programas o de la disposición del proceso de emergencia.

Volviendo al caso de la empresa de vehículos de transporte, en lugar de existir un rastreo permanente acerca de la posición geográfica y disponibilidad de los vehículos, se dispondría del mapa formado por las últimas posiciones reportadas, complementado por el contenido del último mensaje que se recibió y por el momento en que se recibió. De esta forma, se comprobaría a la hora de enviar una nueva orden a un vehículo, si existe alguna laguna de información que hiciera dudar de la disponibilidad de dicho vehículo.

El control de producción de datos, se hace en gran parte cuando se diseñan unos procedimientos adecuados de captación y transmisión de datos y se establecen unos métodos viables y seguros para el proceso previo de transmisión de datos y soportes. Si estos procedimientos tienen unos niveles de calidad y eficacia adecuadas y si hay una vigilancia e inspección de estos procedimientos, podremos estar seguros que el control de producción del dato, en buena medida, esta siendo ya realizado.

1.3.2. CONTROL DE FORMA

El control de forma es un proceso mediante el cual se analiza si los datos que llegan a la Base, cumplen con las especificaciones de diseño para ese tipo de datos. Se comprueba si el “mensaje” es completo en primer lugar y, en segundo lugar, si su forma es válida.

Los registros de información de entrada son analizados por programas de entrada de datos que son independientes de la Base de Datos, pudiendo incluso existir intervención humana en el proceso. Se comprueba para cada uno de los campos, si la información de entrada responde a las especificaciones previstas analizándose los tipos de caracteres, la longitud de los campos e incluso si se traspasan determinados valores absolutos. Es, por lo tanto, un control físico y no un control lógico.

Si los datos son rechazados por el control de forma, los mensajes son devueltos, generalmente con una indicación del error que contienen. Si la transmisión de los datos se realiza en tiempo real o en teleproceso, y el error en el que se incurre se identifica como debido al medio de comunicación, generalmente ya se establece en el sistema de comunicaciones el procedimiento automático para que la transmisión sea intentada de nuevo. Si el error es imputable al mensaje en sí, se devuelve al punto de origen o bien a un punto de conexión que se ha previsto al efecto, iniciándose un ciclo destinado a la corrección del error.

Los errores de formato, se producirán siempre en cualquier sistema, pero las tasas de error alcanzadas, han de ser siempre muy reducidas. Si se alcanzan tasas elevadas, por encima de lo normal, es que ocurre que los procedimientos son malos o las personas los aplican mal y se deberá distinguir si es viable mejorar la intervención humana o es preciso cambiar de procedimiento. En el último caso, se impondrá un adecuado programa de entrenamiento y motivación, y en el segundo caso habrá que pensar en utilizar otros procedimientos alternativos.

Las tasas normales de errores por control de forma, dependen de los tipos de soluciones técnicas que se manejen para los procesos de captación y transmisión del dato. Por ejemplo, para una captación en tiempo real desde un terminal de alta velocidad, con el personal debidamente entrenado, la tasa de errores de formato ha de ser despreciable. Sin embargo, en un mismo sistema de tiempo real, unos datos que sean captados por un teletipo convencional, con unos operadores no muy familiarizados con la gama de errores que pueden cometerse, pero sin que su dedicación sea completa a la producción de esos mensajes, pueden alcanzarse tasas de error de un 15 por 100.

El control de forma, es muy fácil de aplicar, pues pertenece al nivel de procedimientos que siempre han sido utilizados en las aplicaciones informáticas. Sin embargo, todo el paquete de controles de forma en una Base de Datos, pueden constituir una barrera de control que entorpece la buena marcha de la explotación del sistema en la medida en que los procedimientos de captación y transmisión de datos, estén imperfectamente diseñados o defectuosamente puestos en aplicación.

1.3.3. CONTROL DE VEROSIMILITUD

El control de verosimilitud, es complementario del control de forma, y supone un análisis del significado lógico de los datos.

A diferencia del control de forma, el control de verosimilitud, no sólo responde a unas especificaciones de diseño, sino que además se va enriqueciendo con la práctica.

A efectos del control de verosimilitud, la dinámica de producción de datos, constituye un “conocimiento” histórico de los niveles entre los que el dato oscila.

Las tendencias que acusen la historia del dato, nos irán enriqueciendo el nivel de control de verosimilitud ejercido.

Por ejemplo, el campo correspondiente a una cifra bancaria de cuenta corriente a la vista, que tenga más de cinco cifras, pasa correctamente un control de forma, pero un control de verosimilitud, debe advertir que si bien la transacción puede ser válida, corresponde a un caso excepcional. En este caso concreto, el procedimiento de control podría requerir una supervisión bien para advertir un posible error, o bien para que el caso por su excepción, sea conocido a nivel supervisión.

Hay que tener en cuenta que el control de verosimilitud, más que ser un control de tipo “pasa no pasa”, es un procedimiento que puede establecer diferentes niveles de alarma de acuerdo con contenidos lógicos excepcionales.

Es importante este tipo de control, en todo lo que se refiere en general a la medida de cantidades en unidades físicas o en unidades monetarias. Recordemos a título anecdótico, un error de datos en el computador de control logístico del Ejército de los Estados Unidos que “produjo” un envío de diez millones de pares de botas al Vietnam.

No hay regla fija para determinar cómo debe ejercerse este control, a diferencia del control de forma en el que claramente se responden a unas especificaciones de diseño del dato. El control de verosimilitud, se ha de plantear en base al contenido lógico de la información y, por lo tanto, es la experiencia, la historia y la buena administración de la Base de Datos, la que irá determinando la correcta aplicación.

1.3.4. CONTROL DE INTEGRACION

Los datos no solamente contienen una información lógica que proporciona un significado informático acerca de un fenómeno. También pueden incluir determinados códigos que establecen la relación de los datos con otros datos, o lo que es lo mismo, información de relación entre las actividades que los datos representan.

La concepción de sistemas en base de datos, se apoya en un sistema de datos muy estructurado, donde se establecen muchas interrelaciones entre los registros. En cuanto más integrado esté un sistema y mayores sean las interrelaciones lógicas, mayor es el área de influencia de cada dato en particular.

En las aplicaciones convencionales, la calidad de los datos que pasan a incorporarse a los ficheros, afectaba a dichos ficheros en forma individual y como consecuencia, a las aplicaciones que los utilizaban. En un sistema de base de datos, los datos afectan a la sección de la base de datos a la cual corresponden y a todas las demás secciones con las que el dato tiene definida una relación lógica estructural. Esta relación lógica estructural, se extrae del propio registro de información y debe ser controlada para asegurar que el dato se integrará en la base en la forma prevista

por los diseños de estructura lógica de la Base de Datos. Como sistema integrado y complejo que es una base de datos, resulta también más vulnerable y resulta casi más importante la calidad de la información de la relación de los datos con el resto de la Base, que la propia calidad del dato en sí. Podemos hacer notar la diferencia en términos vulgares, hablando de la “enfermedad de un dato” y de la “epidemia” que pueden provocar.

Los datos de relación, pueden venir expresados por el contenido lógico de determinados campos contenidos en el propio registro, o bien por campos específicos incluidos en el registro donde se especifique direcciones o formas de localizar los otros datos con los que ha de relacionarse. Los programas que monitorizan la entrada de datos a la Base, deberán de comprobar que la información de relación existe, que es la adecuada, y que situará al dato en la Base de acuerdo con la estructura lógica prevista. Al propio tiempo, deberá comprobarse que los datos con los cuales el dato que entra va a relacionarse, son datos que existen previamente y datos que tienen los valores adecuados.

También deberá asegurarse, que las modificaciones realizadas en la Base para insertar el nuevo dato, son las correctas. Estas modificaciones consisten generalmente en alteración de las direcciones contenidas en “pointers”, alteraciones que es preciso realizar para encadenar los nuevos datos en el lugar adecuado de la Base. Esta operación, en muchas ocasiones es realizada por el software básico del sistema de base de datos utilizado, mientras que en otros casos, son instrucciones de un programa de usuario.

Si la inserción del dato en la Base responde al primer caso, es decir, es realizada por instrucciones del paquete de software básico, se puede en general tener mayor seguridad de que la integración del dato ha sido correcta, porque el procedimiento utilizado son instrucciones de programa muy sancionadas por la práctica. Sin embargo, el diseñar la Base de Datos para que estas operaciones siempre se hagan a nivel sistema operativo, puede ser bastante oneroso, porque requiere prever mayor número de espacio para “pointers”, cuando muchos de ellos no van a ser utilizados, lo cual amplía el volumen de almacenamiento con el consiguiente coste.

Por ello necesariamente, la “inserción” de registros ha de plantearse en muchos casos como procedimientos de programas realizados al efecto, siendo preciso en este caso extremar, por lo tanto, el control.

Una vez almacenado el dato en la base, pudiera realizarse un procedimiento inverso de lectura como control del proceso anterior, antes de dar por finalizado el almacenamiento del registro, aunque esto puede encarecer la operatividad.

Todos los programas o rutinas que incorporen procedimientos de control de integración de datos en la Base, deben ser aprobados por la unidad de Administración de la Base de Datos, porque afectan en forma muy directa a la función de administración que más adelante examinaremos. También, el Administrador de la Base de Datos, puede especificar para determinados programas, los controles que necesariamente habrán de disponerse para garantizar la buena administración de la Base.

1.3.5. CONTROL DE UTILIZACION

Una vez que el dato ha entrado a formar parte de la Base, está a disposición de ser utilizado dentro del régimen general de explotación de la Base de Datos, lo que no quiere decir que este dato pueda ser accedido por cualquier programa.

El conjunto de datos contenidos en la Base de Datos, es un sistema dinámico muy estructurado, y la utilización de los datos, es decir, la recuperación de registros, la cancelación de registros, la actualización de registros, las modificaciones de relaciones lógicas entre registros, etc., son operaciones que deben ser realizadas de acuerdo con un criterio preestablecido de control a nivel operativo, siendo los objetivos principales de este control de utilización, los siguientes:

- *Asegurar la vida del dato.* Mientras el dato permanece en la Base, ha de ser preservado de operaciones que puedan afectar a la integridad del dato innecesariamente.
- *Asegurar la utilidad del dato.* El dato debe estar en condiciones de reportar la máxima utilidad a las unidades de tratamiento que van a utilizarlo.
- *Evitar la corrupción de los datos.* Además de controlar la integridad del dato, debe controlarse la correcta utilización de las relaciones lógicas entre datos, para evitar que la estructura lógica de relación se deteriore.
- *Proteger la confidencialidad de los datos.* La protección de información, debe también entenderse en el sentido de limitar una difusión indiscriminada de los datos, cuando éstos son de naturaleza confidencial, con respecto a determinadas unidades de tratamiento.

El control de utilización, generalmente está constituido por procedimientos integrados a nivel del sistema operativo de manejo de la Base de Datos y, por lo tanto, el grado de control que podemos ejercer, está en relación con las “facilidades” básicas que nos ofrezca el “paquete” de software de base de datos que utilicemos, dado que en general, a la hora de diseñar una base de datos, nos centramos en plantear el problema en su vertiente informática, de estructurar datos y diseñar aplicaciones de utilización, adoptando un paquete de software básico de los disponibles en el mercado, porque el diseño de software básico para la gestión de una base de datos, constituye un problema de informática fundamental, de naturaleza ajena a la utilización del software simple, lo que no quiere decir que las especificaciones para diseñar el software básico, no las realicen los expertos en resolución de aplicaciones informáticas bajo el concepto de base de datos.

Por lo tanto, como se ha apuntado, el control de utilización se debe de orientar en base a las características que ofrezca el software que vamos a utilizar. No obstante, pueden complementarse con unidades de tratamiento específicas para controlar la utilización de los datos, que sean diseñadas por el órgano administrador de la Base de Datos o bien por analistas de sistemas que diseñan aplicaciones que utilizan la Base de Datos, necesitando dichas aplicaciones disponer de mecanismos de control más afinados que los que ofrece el software básico disponible.

Si sucediera que dichos programas de control que se integran en la Biblioteca

básica de programas del sistema, representarían un conjunto importante, sería evidente que o bien estamos utilizando un software básico poco potente para las utilidades que necesitamos cubrir, o bien, nuestra Organización informática, tiene un grado de madurez tal, que se permite el lujo de construirse el software básico a su medida.

Esta última posibilidad, no debe ser nunca excluida si se trata de una Organización informática de dimensión importante, porque una base de datos compleja, supone una movilización de recursos humanos, económicos y técnicos importantes. En ese caso, puede ser muy conveniente el no ceñirse exclusivamente a las características que pueden encontrarse en un software básico del mercado. Un equipo de programadores de sistema operativo, nos pueden adaptar convenientemente los programas de control para cubrir las necesidades de una correcta administración de la Base de Datos.

Los procedimientos de control de utilización, deben de controlar:

- Qué programas van a utilizar cada tipo de dato.
- Con qué “autoridad” va a utilizar los datos las unidades de tratamiento, es decir, qué restricciones en la utilización de los datos, deben de imponerse en cada programa.
- Uso simultáneo de un mismo dato por parte de varias unidades de tratamiento.
- Protección durante el proceso de actualización de un registro.
- Registro histórico de la utilización de los datos. Hay algunos datos que por su importancia, puede ser necesario recoger una historia que reúna las veces, condiciones y momentos que el dato fue accedido, actualizado o cancelado.
- Discriminar la autoridad o acceder a los datos según la persona peticionaria. En sistemas de bases de datos con facilidades de interrogación por parte de personas usuarias a través de terminales, puede ser preciso disponer de claves personales e intransferibles, que sean decodificadas por el sistema, identificando el nivel de autoridad que a dicha clave le está concedido.
- Seguridad en la operación. Debe de existir un procedimiento que garantice la integridad de los datos cuando han sido modificados sólo parcialmente y se produce una terminación anormal de los programas de proceso.

El control más general de la utilización de los datos por parte de los programas, se establece a nivel de la propia definición de los datos, realizándose de diferentes formas dependiendo del sistema de base de datos de que se trate. Los sistemas que utilizan un “host language”, del que es ejemplo las especificaciones propuestas por el “Data Base Task Group” de la Conferencia CODASYL, establecen una clara diferenciación entre el software de definición de datos y el software de utilización. La descripción de datos se realiza independientemente, “escribiéndose” dicha definición en la Biblioteca del sistema, pudiéndose establecer también definiciones de tipo parcial que constituyan sólo descripciones parciales de la Base de Datos.

Los programas de utilización que vayan a acceder a la Base de Datos, han de realizar previamente una llamada a la definición de datos que vayan a utilizar, para que dicha definición sea extraída de la Biblioteca del sistema, y utilizada por el programa. Como hemos dicho que pueden establecerse diferentes definiciones parciales, se controla ya a este nivel, el sector de base de datos que puede ser accedido por cada aplicación informática.

También puede introducirse el concepto de sensibilidad entre datos y programas, estableciendo para los datos niveles de sensibilidad o autoridad que los programas pueden ejercer. Los programas de utilización “declaran” su autoridad para el manejo de los diferentes datos, controlándose de esta forma la relación entre datos y programas.

El sistema operativo que dirige la explotación de la Base de Datos, suele contener la información de los programas que en cada momento se encuentran en la memoria principal, y han declarado además su intención de acceder a la Base de Datos, de forma que dicho sistema operativo “conoce” qué programas pueden acceder en cada momento, y qué sector (definición de datos correspondiente) puede ser accedido por cada programa en particular. Para hacer ello posible, cuando una unidad de tratamiento que está procesándose, necesita acceder a la Base, ejecuta una instrucción de interés para el sistema operativo, que es la que le proporciona la información antes indicada, a la vez que autoriza al programa la utilización de la Base de Datos. Cuando el programa deja de necesitar el acceso a la Base, se ejecuta la instrucción opuesta para el cierre de la comunicación.

Además de esta regulación de diferentes definiciones de datos a nivel lógico y diferentes relaciones entre dichas definiciones y los programas de proceso, existe, por lo general, otra compartimentación en zonas de la Base de Datos atendiendo a las posiciones físicas que ocupan en las unidades de almacenamiento. Así, los datos están alojados en diferentes Areas físicas de almacenamiento, disponiéndose para dichas Areas, mecanismos de apertura y cierre análogos a las aperturas y cierres de ficheros convencionales.

Antes de que una unidad de tratamiento pueda acceder a un Area de almacenamiento, debe de “abrir” dicha Area para su utilización, declarando el tipo de autoridad con que va a operar en el Area, ello además de la apertura previa que el programa realiza para la utilización de un sector lógico de la Base de Datos.

A partir de ese momento, el programa puede acceder al Area de almacenamiento físico abierta al uso, pero únicamente a los registros que están contenidos en dicha Area, y pertenecen además al sector lógico de la Base de Datos cuya definición está siendo utilizada por el programa en cuestión.

Las autoridades que el programa puede tener respecto del Area abierta para utilización, son las siguientes:

- Carga inicial de datos.
- Recuperación de registros.
- Actualización de registros.
- Recuperación protegida de registros.

- Recuperación de registros en exclusiva.
- Actualización protegida de registros.
- Actualización de registros en exclusiva.

Además de estas autoridades que el programa puede ejercer según sea la declaración efectuada con la instrucción de apertura del Area, el sistema operativo debe de garantizar la correcta resolución de un posible “conflicto” de utilización de un mismo dato por parte de varios programas.

A estos efectos, el sistema operativo lleva un control dinámico de las autoridades concedidas en cada momento a los programas sobre las Areas de almacenamiento. Así, un Area que ha sido abierta por una unidad de tratamiento para una operación de carga inicial de datos, se cierra automáticamente a cualquier otro tipo de utilización por parte de otro programa. Si en ese momento, en el proceso de un programa se ejecuta una instrucción que pretende abrir ese mismo Area para operaciones de actualizaciones, pongamos por ejemplo, la ejecución del programa se interrumpe pasando la Unidad Central de Proceso a ejecutar otras tareas. El programa que está en estado de espera, solamente se vuelve a activar en el momento en que el programa que estaba efectuando una carga inicial de datos, cierra el Area, y la limitación de acceso a dicha Area, queda, por lo tanto, levantada. A partir de ese momento, el programa se activa nuevamente y podrá abrir el Area, aunque por supuesto, para volver a procesarse el programa, tendrá que aguardar el turno de prioridades que le corresponda dentro del régimen general de explotación.

Análogamente sucede:

- Si ha sido abierta un Area por un programa para recuperación de registros, durante el tiempo de apertura, quedará cerrado el uso de dicha Area para actualización por parte de cualquier otro programa, pues obviamente un programa no puede tener garantías del grado de actualidad necesario en la recuperación de registros, si otro programa simultáneamente, está realizando operaciones de actualización. Se trata de recuperación protegida, naturalmente.
Por supuesto, no pueden tampoco concederse autoridades simultáneas a un segundo programa de recuperación exclusiva.
- Si un Area ha sido abierta para actualización de registros, ese mismo Area no podrá abrirse simultáneamente por parte de otro programa, para recuperación protegida, pues obviamente se pueden estar produciendo alteraciones en los registros, no pudiendo garantizarse, por lo tanto, la protección de la información que podría ser recuperada.
Tampoco podrán concederse autoridades simultáneas para actualización protegida o exclusiva a otros programas.
- La apertura de un Area para utilización exclusiva, bien sea recuperación o actualización, cierra el uso simultáneo a cualquier otro programa para el mismo uso.
- Si se abre un Area para recuperación protegida, queda limitado el que cualquier otro programa simultáneamente pueda actualizar registros.

Algunos registros, tienen un interés de excepción bien por su condición de información clave de la Organización, o por el valor jurídico que pudiera representar de cara a relaciones contractuales o administrativas. En estos casos, puede disponerse un procedimiento de control que recoja la dinámica de evolución de dichos registros en la Base de Datos, de forma que en un momento dado, sea recuperable las diferentes actualizaciones que ha ido experimentando la información, así como las intervenciones que dieron lugar a ellas. Este procedimiento, también sirve de inspección a posteriori, en aquellos sistemas de base de datos, que están abiertos a utilización en teleproceso a muchos usuarios, permitiendo ejercer una auditoría en frío de la utilización del sistema, auditoría que practicada con una base estadística, puede sustituir a muchas de las operaciones de control que no siendo de índole fundamental, encarezcan innecesariamente la explotación general del sistema.

En los casos de sistemas con gran utilización de operaciones de pregunta-respuesta, en los que además se puedan producir actualizaciones por parte de gran número de operadores de terminales, es conveniente el establecer un régimen de "autoridad" de los usuarios. El usuario tiene una clave que le identifica en el sistema como utilizador, y a su vez existen unas determinadas autoridades, que han sido conferidas al poseedor de dicha clave. El utilizador introduce su clave como primera comunicación con el sistema de Base de Datos, y de esta forma, abre su diálogo estando limitadas sus posibilidades por un nivel de autoridad conferido a su clave. La clave puede ser también utilizada por el sistema para reseñar a título de firma cualquier operación de actualización o cancelación que el usuario de la clave practique.

Esta técnica de control, permite una responsabilización directa de las operaciones de modificación de datos, que se realicen en el sistema, lográndose por la vía de la responsabilidad, la disposición de una excesiva cadena de controles, que restarían operatividad y, por lo tanto, eficacia al sistema, máxime si se trata de una explotación en tiempo real.

Por supuesto, debe de estar garantizada la exclusividad e intransferibilidad de las claves personales, existiendo para ello diferentes técnicas. Una puede consistir en que la clave esté dividida en 2 partes relacionadas biunívocamente por un criterio único que puede ser muy sofisticado y que es mantenido únicamente a nivel sistema operativo, permaneciendo secreto. La segunda parte de la clave, sólo se utilizará para introducirla el operador como identificación personal para que el sistema operativo reconozca la "casación" entre las dos partes de la clave. La primera parte de la clave por el contrario, puede divulgarse, y es la que se asocia a cada operación de actualización.

La segunda parte de la clave, sólo la conoce el empleado en cuestión y la máxima autoridad en la administración de la Base de Datos. El empleado, si es debidamente instruido, tendrá un cuidado en introducir en el sistema la segunda parte de la clave con la máxima discreción, porque en caso de que su clave caiga en poder de otra persona, pueden ser relacionadas violaciones de la información contenida en la Base de Datos, que serán imputadas al operador propietario de la clave.

Para garantizar mayores seguridades, pueden realizarse cambios periódicos aleatorios de las claves.

En sistemas de tiempo real, de vocación operativa, se impone establecer muchas flexibilidades cuyos objetivos están encontrados con los criterios de control. En estos casos, el sistema de claves de firma, permite un seguimiento a posteriori de las prácticas de cada operador, determinando por un lado su grado de fiabilidad y, por otro lado, obteniendo conclusiones acerca de su adiestramiento como operador que actúa en el sistema.

Además de estos controles de proceso que afectan a las autoridades de usuarios y programas y a posibles conflictos de uso simultáneo de los datos, existen otros controles para garantizar la correcta dinámica de utilización de los datos.

Son controles que velan, porque el paquete de operaciones de acceso que una unidad de tratamiento realiza desde que abre su diálogo con la Base de Datos hasta que lo cierra, forme un conjunto único y coherente. Estos controles han de preservar a la base de datos de las operaciones abortadas de los programas.

A tal efecto, se suele disponer de un dispositivo de copia de Areas o páginas de la Base, previa a su actualización, pudiéndose de esta forma restaurar dichas Areas o páginas cuando se produce una terminación incompleta de los procesos, quedando de esta manera la base de datos en el estado que se encontraba antes de su actualización.

De una forma más general, han de estar previstos los procesos generales de respaldo de la Base de Datos ante posibles circunstancias imprevistas de averías hardware e inoperaciones de la instalación. Este tipo de controles, pertenecen más bien a las funciones generales de mantenimiento y de administración, por lo que serán examinados más adelante.

1.3.6. CONTROL DE MANTENIMIENTO

La actividad de control, no debe limitarse a la garantía de calidad del dato que pasa a integrarse a la Base de Datos y al posterior control de su utilización. Son muchas las circunstancias que pueden presentarse a lo largo de la vida de los datos en la Base que produzcan su corrupción.

Aun cuando veremos de forma más detallada y completa las funciones de administración de la Base de Datos, vamos a resaltar algunos de los controles que pueden practicarse en la Base de Datos para garantizar un correcto mantenimiento de la información.

Las operaciones generales de administración "mecánica" de la Base de Datos, tales como condensación de páginas o Areas, información, expansión de páginas, inicialización, parcheo de registros, etc., deben de estar conducidas por unos procedimientos exhaustivos de control tanto a nivel software como a nivel humano, por razones obvias de seguridad, ya que son operaciones que pueden afectar directamente a la información almacenada aun cuando la naturaleza de sus objetivos, sea de índole administrativa interna.

El control estadístico de los datos, debe ser siempre viable, para poder hacer en cualquier momento, cualquier análisis de conjunto o cualquier investigación parti-

cular, bien por atender a un procedimiento establecido o por ser necesaria ante una incidencia que se haya podido producir. Para que este control estadístico sea viable, deberá disponerse en la Biblioteca de programas, de las correspondientes unidades de tratamiento que permitan listado de Areas, páginas, conjuntos de registros, etc. También deberá poder listarse los controles de seguimiento que se hayan podido disponer en la utilización de los datos, tales como historia de registros, etc. Por supuesto, el paquete de estadísticas, no sólo sirve para controlar el mantenimiento de la información, sino que su vocación fundamental es la de ser instrumento de gestión en la administración de la Base de Datos (véase 1.8).

En el mantenimiento correcto de registros de páginas y Areas, también colaboran los procedimientos generales de control operativo mediante el empleo de "check-points", y "copiados" de páginas de almacenamiento, utilizándose generalmente una cinta magnética donde se va recogiendo en forma precautoria, el contenido de páginas que van a sufrir modificación, etc., lo que permite asegurar debidamente la protección de la información contenida en la Base.

También deben practicarse chequeos periódicos de la estructura lógica de los datos, verificando que los árboles, cadenas, anillos y demás estructuras lógicas establecidas en la Base (véase 1.4), se mantienen de acuerdo con los planes previstos, siendo las interrelaciones de registros las adecuadas. A tales efectos, pueden diseñarse programas de control de mantenimiento que sigan detectando "loops", que verifiquen que los anillos están correctamente formados y que los registros miembros que los integran, corresponden realmente al que pertenezca a ellos.

1.4. LAS ESTRUCTURAS DE DATOS

El concepto de Base de Datos, viene a aportar a la Informática, una característica muy potente que es la independencia entre las estructuras lógicas y físicas de almacenamiento. La problemática del almacenamiento físico de la información, es afrontada por la Administración de la Base de Datos, ayudada por el paquete del software básico y por analistas a nivel sistema operativo.

La estructura lógica de la información es, sin embargo, área de interés para los utilizadores de la información, y también, por supuesto, para el órgano administrador de la Base de Datos.

Los analistas de aplicaciones, utilizan de forma exclusiva, estructuras lógicas de almacenamiento, siendo para ellos transparente la estructura física. Ello, incluso cuando utilizan facilidades de acceso directo, porque los accesos directos en las bases de datos, suelen realizarse en base a direcciones lógicas absolutas que son traducidas por el sistema operativo de la Base de Datos, obteniéndose la dirección física que está asociada de una forma biunívoca a la dirección lógica. No son, por lo tanto, auténticos accesos directos.

Las estructuras de almacenamiento físico, son por supuesto las que se usan en los sistemas informáticos en general, y nos vamos a centrar, a lo largo del punto 1.4, en el examen de las estructuras lógicas de utilidad en los sistemas de bases de datos.

Previamente, vamos a centrar el conjunto de estas estructuras de una forma general, utilizando como algoritmo de expresión el grafo.

El registro (véase 1.4.1), será la unidad lógica utilizada, asociándose los registros a nodos de un grafo. La relación lógica entre dos registros, la asociamos a la arista del grafo, y generalmente estará materializada por un "pointer" que es un campo contenido en un registro y que contiene la dirección lógica del otro registro al cual de esta forma conecta.

Como vemos, el "pointer" es el factor que produce el "divorcio" entre el orden físico y el orden lógico, permitiendo que las ordenaciones lógicas de la información, sean independientes de su almacenamiento físico. Utilizamos la palabra inglesa pointer, por ser su uso muy generalizado en el ámbito informático, aunque también podríamos denominarlo apuntador o conector.

Una *Lista* de registros es un conjunto de registros tal que están conectados en secuencia mediante pointers, existiendo en cada registro un pointer que lo conecta al registro siguiente de la lista a excepción del último, y siendo cada registro, a excepción del primero, conectado al registro anterior por medio de un pointer incluido en dicho registro anterior. Si el primer registro de la lista coincide con el último, decimos que se trata de una *lista cerrada*, y si el primer registro y el último de la lista son diferentes, la lista se denomina *abierta*. En una lista, un mismo registro puede figurar más de una vez.

Si los registros de una lista son todos distintos y si la lista es abierta, la lista toma entonces la denominación de *Cadena*. En términos informáticos usuales, muchas veces se la denomina Lista a una Cadena, pero como vemos, la Cadena es un tipo específico de Lista.

Si en la Lista de registros todos los registros son distintos, y se trata además de una lista cerrada, la lista toma la denominación de *Anillo*.

Si la estructura lógica de registros puede asociarse a un grafo conexo, donde cada par de registros está unido al menos por una cadena, diremos que la estructura es conexa.

Un *Arbol*, es un caso particular de la estructura conexa en la que no existe ningún *Anillo*.

Si en la estructura existe un registro que contiene un pointer, el cual contiene la dirección del propio registro, diremos que existe "loop". Una *Red* es una estructura lógica conexa que no tiene "loops".

En los puntos siguientes vamos a ver en detalle cada una de éstas estructuras que hemos examinado en términos generales, finalizando con la exposición del concepto de *Set* y *Esquema*, según las especificaciones establecidas por el "Data Base Task Group" de CODASYL, por interpretar que corresponden a estructuras standard

de uso universal que pueden adoptarse como especificaciones de base para cualquier planteamiento de base de datos.

1.4.1. EL REGISTRO, LOS GRUPOS Y LOS CAMPOS

El *Campo* es un dato elemental que constituye la cantidad más pequeña de información que puede ser referenciada y que tiene una significación lógica susceptible de ser interpretada. El *Registro* está formado por un conjunto de campos que configuran una entidad lógica con "personalidad" informática unitaria. El registro es una unidad informática que constituye la célula básica de las estructuras lógicas.

Hay que distinguir, a la hora de expresarnos y entendernos correctamente, entre un registro *Tipo* y un registro *Específico*, o simplemente registro, diferencia que reflejamos en la figura número 10. El registro tipo es un patrón o una especificación de registros, mientras que los registros en general son entidades que se ajustan a la definición de un determinado registro tipo, y constituyen por lo tanto los datos que serán almacenados en la Base.

Generalmente, recibe la denominación de *Fichero* un conjunto de registros que corresponden a un mismo registro tipo.

En la figura 10 se señala también la composición de campos dentro de un registro, composición que obedece a una estructura determinada, estructura que es interna al registro y que se denomina *Estructura de Fichero*. Dicha estructura interna puede tener organizaciones jerárquicas potentes, tales como las utilizadas en el lenguaje **COBOL**.

Los campos de los registros corresponden a informaciones elementales, aunque en algunos casos puede descomponerse a su vez en otros campos más elementales, como es el caso del campo "Matrícula" del ejemplo de la figura 10, el cual puede subdividirse a su vez en tres parte, la primera conteniendo un código alfabético de

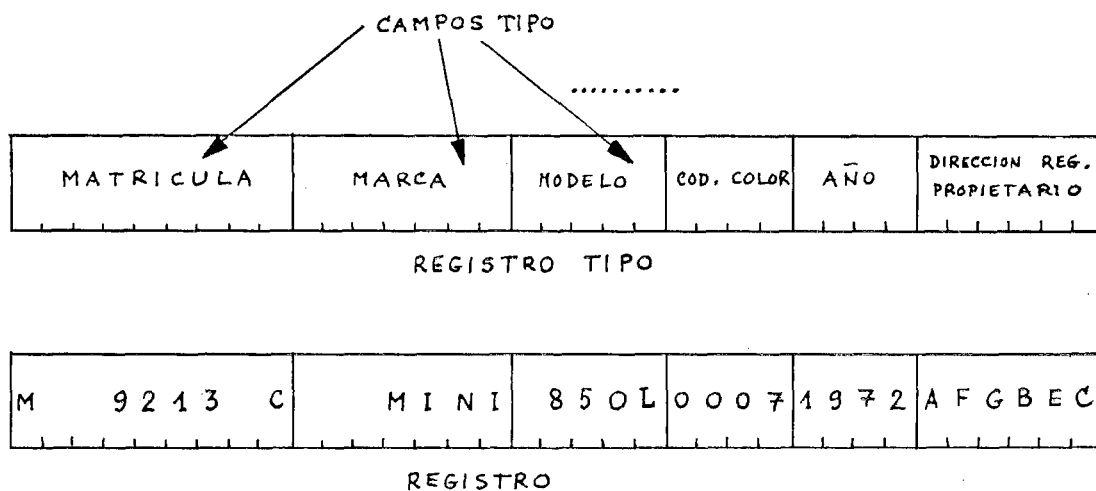


FIG. 10

identificación de la ciudad, la segunda parte un código numérico secuencial, y la tercera, un código alfabético secuencial. Estos campos compuestos reciben más propiamente el nombre de *Grupos*.

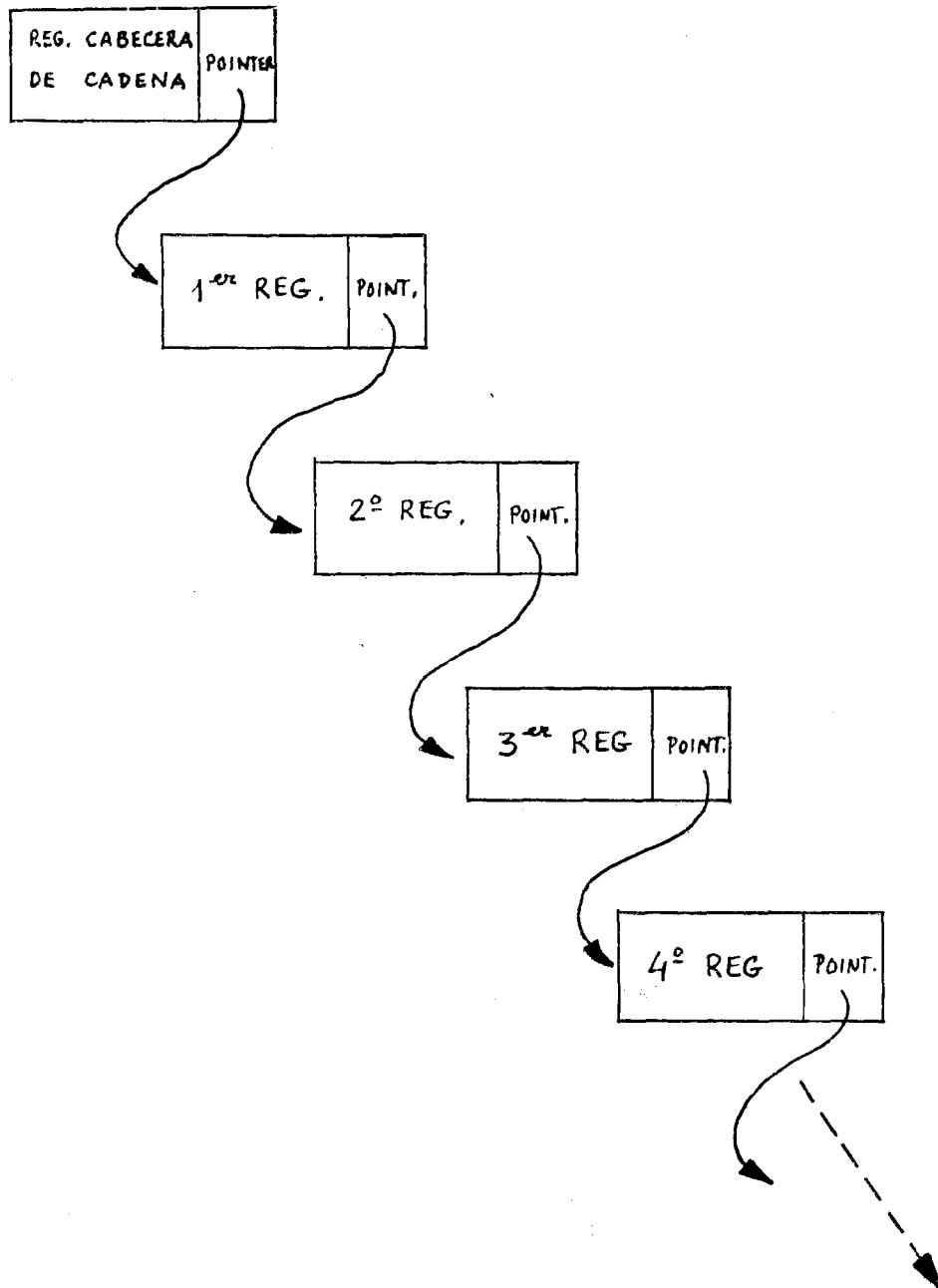


Fig. 11. Lista simple o cadena

1.4.2. LISTA SIMPLE O CADENA

En la figura 11 se representa una lista simple o cadena que corresponde a una lista abierta de registros distintos.

El primer registro de la cadena se denomina “cabecera de cadena”, y en el pointer de dicho registro de cabecera estará contenida la dirección absoluta lógica del primer registro de la cadena.

El primer registro de la cadena dispone de un pointer que contiene la dirección del segundo registro. El segundo registro dispone de un pointer que conecta al tercer registro y así sucesivamente, hasta llegar al último registro de la cadena que no dispone en su pointer de una dirección lógica válida.

Los registros de la cadena, son procesados de acuerdo con la secuencia lógica que determina su estructura, siendo los pointers los elementos que materializan dicha secuencia. El proceso de la cadena de la figura 11 sólo puede realizarse en un sentido único, comenzando en la cabecera de cadena y terminando en el último eslabón.

El acceso a un registro determinado exige una exploración lógica secuencial de la cadena hasta buscar el registro deseado.

La inserción de un nuevo registro en la cadena puede realizarse en cualquier posición de la secuencia. Si el registro nuevo se añade en la última posición, basta con obtener la dirección del nuevo registro que queremos encadenar y situarla en el pointer del registro que previamente ha sido identificado como el último registro “actual” de la cadena. Si por el contrario, no deseamos insertar el registro en la última posición, sino en la posición N de la secuencia lógica, la operación a realizar será el escribir la dirección del nuevo registro en el pointer del registro que ocupaba la posición $N-1$, escribiendo además en el pointer del nuevo registro la dirección del registro que antes ocupaba la posición N , y que ahora pasa a ocupar la posición $N + 1$.

Si el pointer del registro de cabecera no señala ninguna dirección válida, quiere decir que la cadena carece de eslabones. Si por error es destruido el pointer de un determinado registro de la cadena, y dicho pointer pasa a no señalar una dirección válida, la cadena queda truncada, y ese registro pasa a ser el último eslabón de la cadena.

Si el pointer de un registro de la cadena señalara bien la dirección del propio registro que contiene el pointer, o bien la dirección de cualquier otro registro anterior en la secuencia lógica, o la dirección del propio registro de cabecera, la cadena tiene un “loop” y por lo tanto, el proceso secuencial entra en un círculo vicioso. Esta es una característica no deseable a no ser que esté prevista previamente, en cuyo caso no estaríamos tratando de cadenas, sino de anillos, como más adelante veremos, para los cuales se prevén otros tipos de proceso.

Las cadenas, en general, cumplen las propiedades asociativa y distributiva, pudiéndose agrupar varias cadenas en una o dividir una en varias.

1.4.3. LISTA COMPUESTA

El algoritmo lógico de la cadena o de la lista simple permite establecer estructuras compuestas como la que se presenta en la figura número 12, en la que puede

contemplarse la cadena de registros $A, B, C, D, \dots N$. Dichos registros son a su vez cabeceras de otras cadenas, $A, A_1, A_2, A_3, \dots A_n, B, B_1, B_2, \dots B_n, \dots$ etc.

Para cada una de las cadenas secundarias, cabe hacer las mismas consideraciones del punto 1.4.2. Para la cadena principal $A, B, C, \dots N$, las posibles operaciones o procesos, son también análogos, aunque debemos tener en cuenta que cualquier operación que hagamos con un registro de la cadena principal, es una operación que afecta a toda la cadena que de dicho registro depende. El truncamiento de la cadena principal, provoca la pérdida de todas las cadenas secundarias a partir del registro cuyo pointer fue destruido.

La inserción de un nuevo registro en la cadena principal no requiere necesariamente que dicho nuevo registro lleve asociado una cadena, sino que pueda integrarse como cabecera de una cadena vacía.

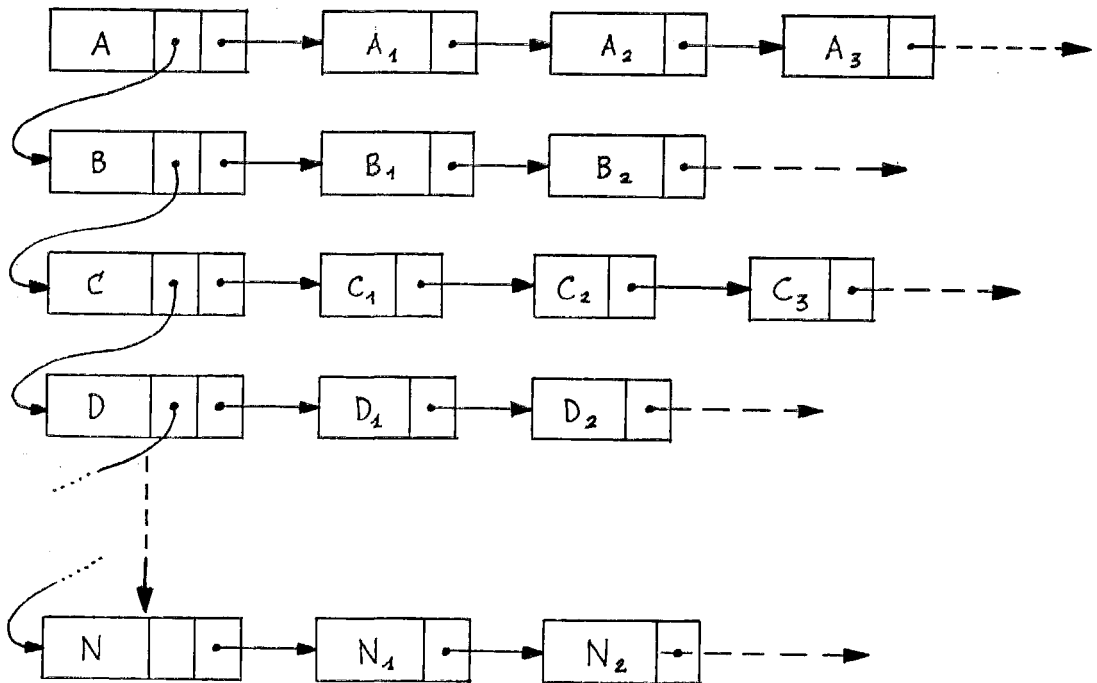


FIG. 12. Lista compuesta

Obsérvese que los registros de la cadena principal precisan de dos pointers diferentes, uno de ellos para su integración en la cadena principal y el segundo para servir de cabecera de una cadena secundaria.

Las listas compuestas permiten lo que se denomina proceso de “*inversión de listas*”, que en realidad consiste en sustituir una cadena que puede ser muy larga, y su proceso por lo tanto es lento, por un conjunto de cadenas compuestas que permiten optimizar el tiempo de acceso a los registros.

Este proceso de inversión de listas puede llevar a la situación extrema de tantas cadenas como registros haya, con lo que habría un registro único de cabecera, el cual contendría una matriz de pointers en la que estarían direccionados todos los

registros del conjunto. Esto viene a ser una forma de acceso directo a partir de dicho registro de cabecera, que muchas veces se utiliza, siendo concretamente una forma de organizar el Set (véase punto 1.4.8), definido por D.B.T.G. de CODASYL.

1.4.4. LISTA DOBLE

Cuando describimos la cadena, vimos que tenía como característica la necesidad de ser procesada secuencialmente en una sólo dirección a partir de la cabecera de cadena, de forma que para localizar un determinado registro, es preciso acceder previamente a los anteriores.

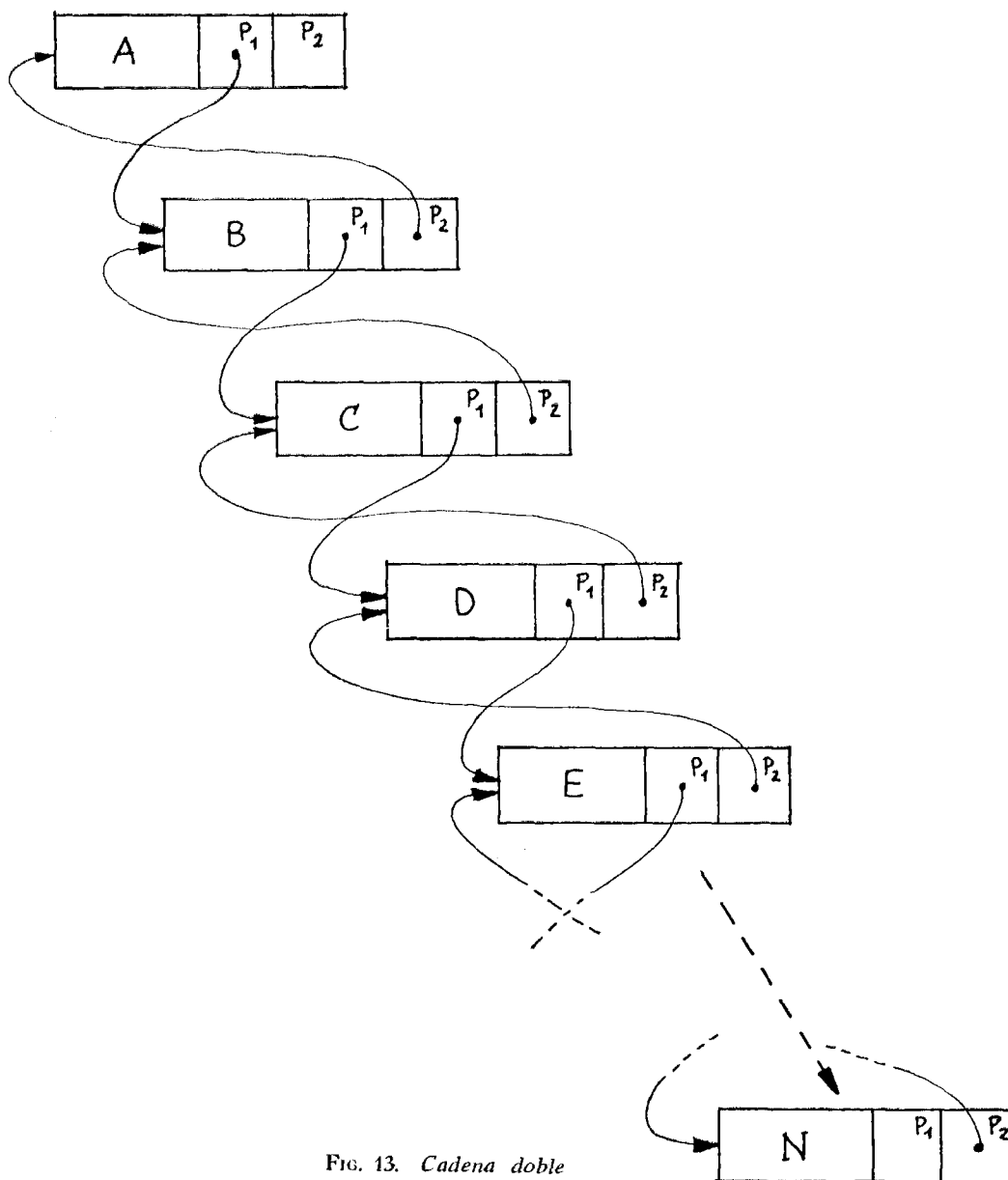


FIG. 13. Cadena doble

La cadena doble que se representa en la figura número 13, permite la posibilidad de ser procesada secuencialmente en los dos sentidos, es decir, desde la cabecera de cadena y hacia la cabecera de cadena. Para hacer ello posible, se dispone de un doble encadenamiento de los registros, de manera que cada registro, además de tener un pointer con la dirección del registro siguiente, dispone de otro pointer con la dirección del registro antecedente.

Así, siguiendo la imagen de la figura 13, un registro cualquiera, por ejemplo el C , incluye dos pointers, P_1 y P_2 . El pointer P_1 contiene la dirección del registro D , que es el siguiente, mientras que el pointer P_2 señala la dirección del registro B , anterior

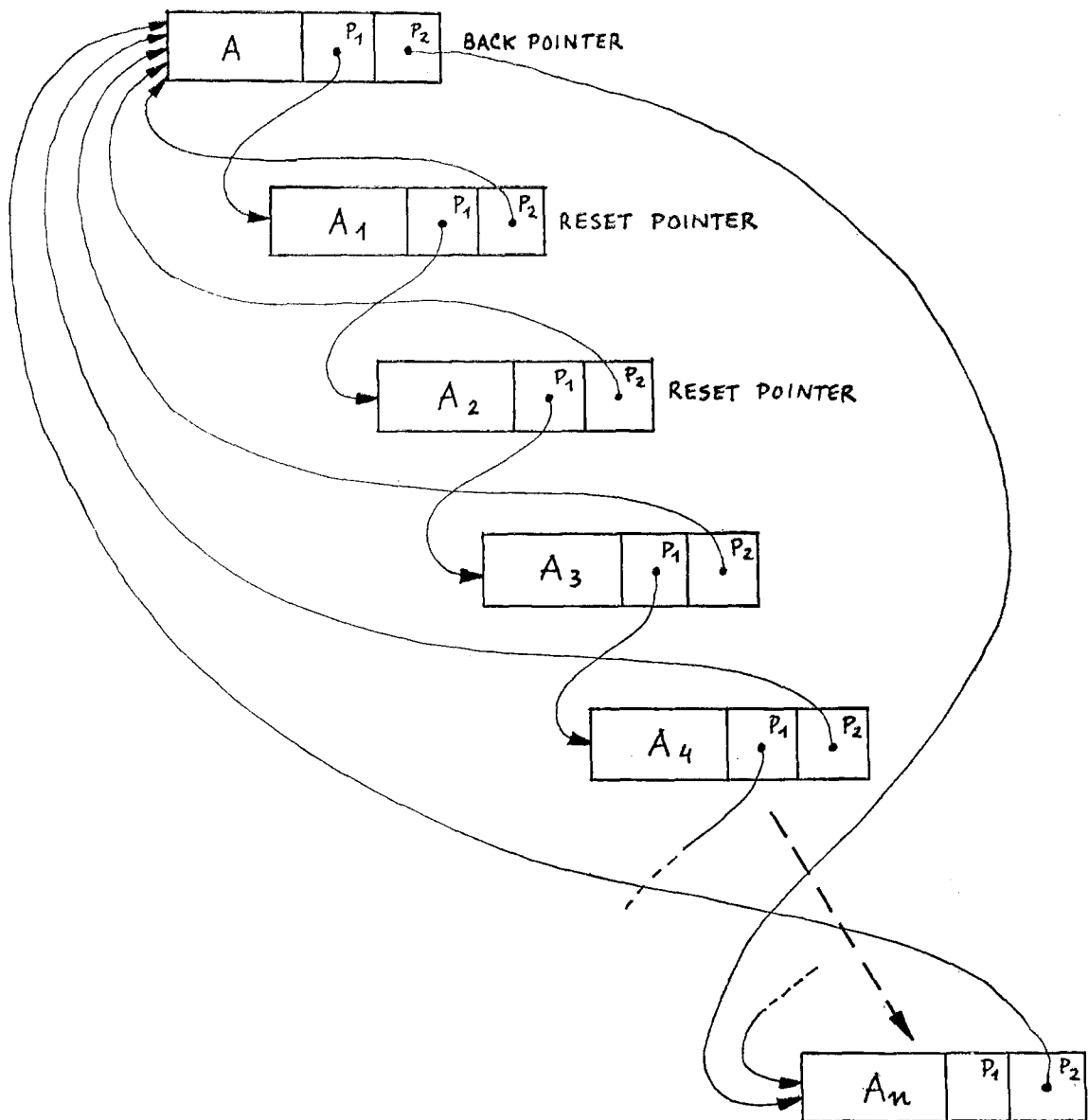


FIG. 14. Cadena con pointers "back" y "reset"

en el orden lógico. Naturalmente, el registro de cabecera, sólo utiliza un pointer para enlazar el primer registro de la cadena, ya que carece de elemento antecedente. Igualmente, el último registro de la cadena tendrá su pointer P_2 señalando al registro anterior, mientras que el pointer P_1 no contiene ninguna dirección válida.

Se pueden realizar las mismas operaciones de proceso descritas para la cadena simple, y habrá de tenerse en cuenta que la inserción de un nuevo registro en un punto intermedio de la cadena, requiere realizar un encadenamiento doble, escribiendo la dirección del nuevo registro en el pointer P_1 del registro anterior y en el pointer P_2 del registro siguiente, situando también las direcciones de dichos registros anterior y siguiente, en los pointers P_1 y P_2 , respectivamente, del nuevo registro insertado.

Además de hacer posible el acceso de registros en las dos direcciones, ascendente y descendente, en la secuencia lógica, el doble encadenamiento ofrece obviamente mayor seguridad ante un posible deterioro de los enlaces lógicos, ya que dichos enlaces están duplicados.

También puede adoptarse todavía una mayor seguridad en esta estructura de datos añadiendo nuevos enlaces lógicos adicionales. Los más utilizados son los que se representan en la figura 14. El "back-pointer" es un pointer único situado en el registro de cabecera que siempre contiene la dirección del último registro de la cadena. El "reset pointer", es un pointer contenido en cada uno de los registros miembros de la cadena, conteniendo todos ellos la dirección del registro de cabecera.

Estos pointers adicionales permiten agilizar el proceso de la cadena y por lo tanto, los accesos a los registros. El "back pointer" permite direccionar el último registro a partir de la cabecera, lo que facilita el proceso secuencial inverso de la cadena, si está doblemente enlazada. Los "reset pointers", permiten pasar desde cualquier registro de la cadena a la cabecera con una sola operación de acceso, sin necesidad de volver a recorrer el camino secuencial inverso, procedimiento que agiliza determinados tratamientos de la información, que requieren exploraciones sucesivas de registros, sin necesidad de conservar en memoria referencias de cabeceras de registros, que pueden ser muchas y estar repartidas en diferentes áreas y páginas del almacenamiento.

1.4.5. ANILLOS

El anillo es un lista cerrada de registros, todos ellos diferentes. Se trata por lo tanto de una cadena que ha sido "cerrada" uniendo el último eslabón a la cabecera. Se representa en la figura 15.

El registro R se denomina *registro principal*, y todos los demás registros del anillo se denominan *registros miembros*. El registro principal contiene un pointer que direcciona el primer registro miembro del anillo. Este, a su vez, tiene un pointer que direcciona el segundo miembro, y así sucesivamente hasta el último registro del anillo, cuyo pointer contiene la dirección del registro principal.

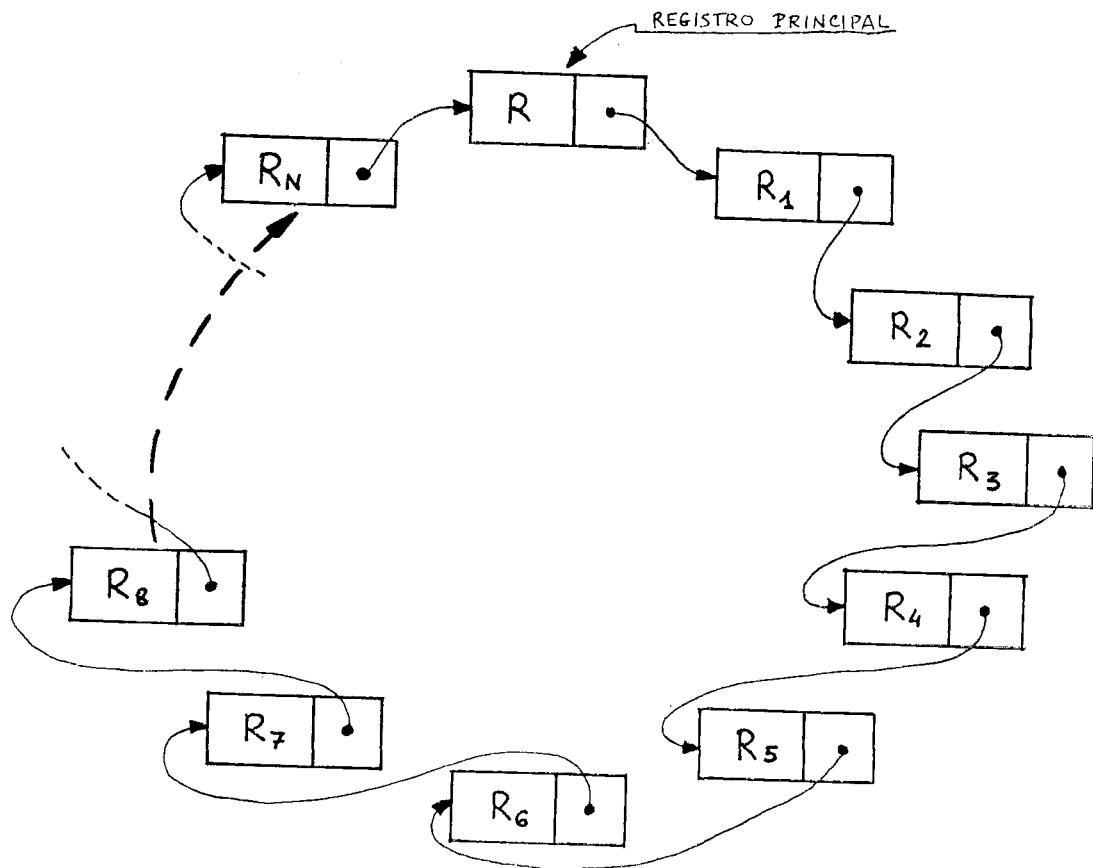


FIG. 15. Anillo

El proceso del anillo es análogo al proceso de la cadena de registros, debiendo seguirse un camino secuencial para direccionar un determinado registro miembro. La exploración secuencial finaliza cuando se llega de nuevo al registro principal, que fue la dirección de partida. Dentro del anillo, se pueden seguir diferentes criterios lógicos de ordenación de registros, utilizándose las mismas reglas lógicas indicadas para la cadena en lo que se refiere a la inserción o supresión de nuevos registros, o por superposición de ambas operaciones, para la alteración del orden de los registros miembros en la secuencia lógica.

Al igual que en el caso de la cadena, si el pointer, en vez de direccionar el registro siguiente, direcciona uno de los registros miembros precedentes en la secuencia lógica, existe entonces un "loop", lo que destruye la posibilidad de procesar el anillo por provocar una ruptura de la secuencia. Sin embargo, a diferencia de la cadena, las operaciones asociativas o distributivas entre diferentes anillos, si bien obviamente son posibles en el plano lógico, presentan en la realidad una tipología diferente.

El anillo puede ser también doble, como se muestra en la figura 16, donde cada registro miembro tiene dos pointers para direccionar sus registros antecedentes y consecuentes. El pointer P_2 del primer registro miembro R_1 , direcciona a su antecedente que es el registro principal, mientras que el último registro miembro R_n ,

direcciona también al registro principal como registro consecuente por medio del pointer P_1 . Los pointers del registro principal direccionan al primero y último registro miembro.

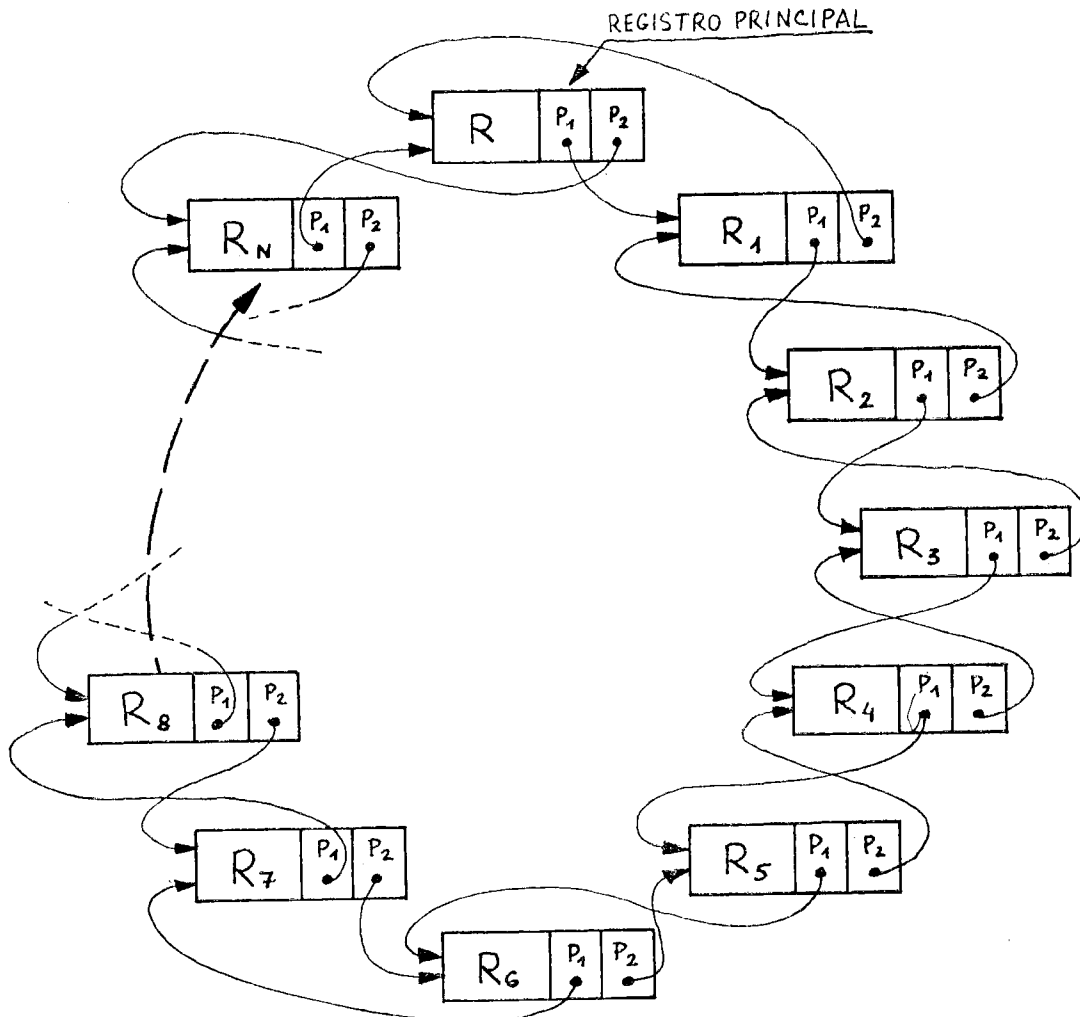


FIG. 16. Anillo doble

El anillo doble permite el proceso de registro en los dos sentidos de la secuencia, lo que agiliza determinados procesos y ofrece una mayor seguridad a la estructura de datos. Los registros miembros de los anillos, también pueden contener un "reset pointer" que direccionen al registro principal, con lo que puede obtenerse una flexibilidad y rapidez mucho mayor de los procesos. Obviamente, en el caso de los anillos, no existe un "back pointer".

En determinados casos en los que se precisa una agilidad todavía mayor en el acceso a registros, conservando sin embargo las estructuras del anillo, puede utilizarse el *anillo con matriz de pointers*, que consiste en integrar en el registro princi-

pal del anillo una matriz de pointers que contiene las direcciones de todos los registros miembros y que determina el criterio de ordenación.

La estructura de anillos, ofrece características muy potentes, cuando es utilizada como "célula lógica" para definir estructuras más complejas, habiendo sido adoptada como tal en las especificaciones propuestas por CODASYL.

1.4.6. ARBOLES

El Arbol es una estructura lógica que puede ser representada por un grafo conexo sin anillos. Es una estructura de tipo jerárquico que a título de ejemplo hemos representado en la figura 17, y que en líneas generales presenta la misma configuración que las estructuras jerárquicas que pueden especificarse en los programas Cobol, con la diferencia de que en el Cobol se determina una estructura jerárquica que va a formar un bloque compacto en un registro, mientras que en nuestro caso la estructura de relación trasciende de nivel registro, y además, mediante la utilización de pointers, el almacenamiento de los registros es completamente independiente.

En el Arbol existe un registro "raíz" del que parte toda la estructura jerárquica, siendo por lo tanto llave de entrada para el acceso de registros en general. El registro raíz representado por el registro *A* en la figura 17, incluye pointers que direccionan registros de nivel inmediatamente inferior a jerarquía, *B1*, *B2* y *B3* en la figura. Cada uno de estos registros tiene a su vez varios pointers, o solamente uno, o ninguno, que direccionan y encadenan los registros del siguiente nivel, *C1*, *C2*, *C3* ... etc., y así sucesivamente. Por lo tanto, cualquier registro del Arbol es direccionado siempre por otro registro de nivel superior, a excepción del registro raíz. Para todos y cada uno de los registros miembros del Arbol, existe una cadena única que partiendo del registro raíz, enlaza el registro. Por ejemplo, el registro *E4* (figura 17), es enlazado por la cadena *A*, *B1*, *C3*, *D2*, *E4*.

Para cada registro, por lo tanto, existe una cadena asociada por medio de la cual dicho registro puede ser accedido a partir del registro raíz. El Arbol está en realidad constituido por un conjunto de cadenas unidas, de forma que no se formen anillos, y al Arbol pueden, por lo tanto, seguir anexionándosele nuevas cadenas sin necesidad de modificaciones en su estructura básica imputables a dicho cambio. Por ello, es una estructura lógica muy interesante para las aplicaciones informáticas que implican el mantenimiento de grandes "diccionarios" para el acceso a la información. Las aplicaciones de recuperación de información (Information retrieval), utilizan preferentemente este tipo de estructuras, porque permite identificar la dirección de un registro por un código significativo o por un conjunto de palabras clave, siendo dicho código o el conjunto de palabras clave la información que nos identifica la "cadena" de acceso a partir del registro raíz.

Los Arboles pueden ser de varios tipos, y aunque no vamos a profundizar estableciendo la tipología, señalaremos como estructura muy utilizada, el Arbol binario, que es aquél en el cual cada registro direcciona siempre o bien dos registros del nivel inmediato inferior o bien ningún registro. En un Arbol binario, cada registro es, o bien final de cadena, o bien eslabón de una cadena y cabecera de otra, a excepción del registro raíz, que es cabecera de dos cadenas.

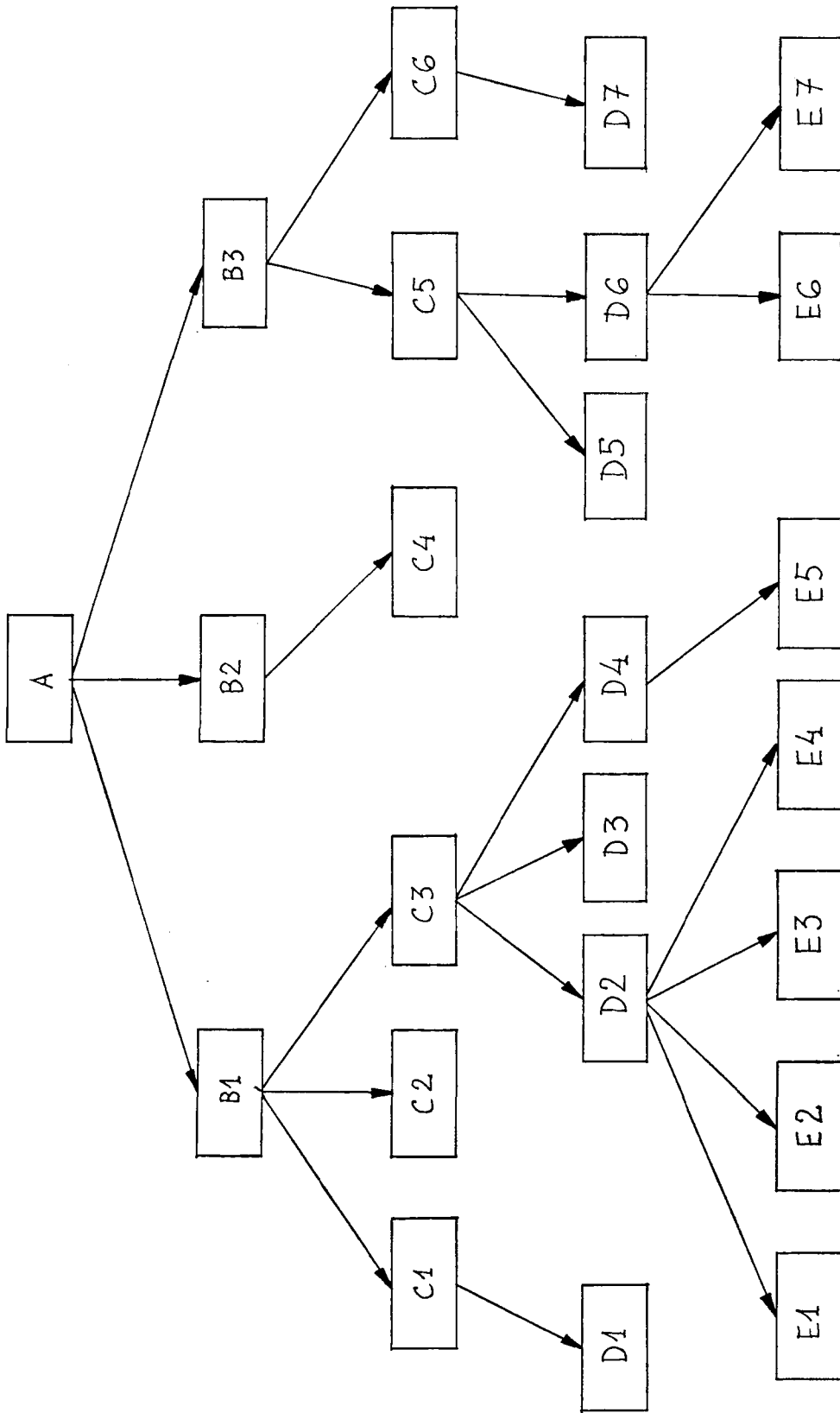


FIG. 17. Arbol

Conviene señalar, que en un Arbol binario pueden preverse en los registros de forma general dos únicos campos para los dos únicos posibles pointers que cada registro utilizará para integrarse en la estructura, y sin embargo, por el contrario, en un Arbol múltiple, cada registro puede requerir un número indeterminado de pointers según el número de ramas que eventualmente puedan llegar a enlazar dicho registro con registros del nivel inmediatamente inferior. En los paquetes de software de base de datos, pueden incluirse rutinas que en función de la información que ha de servir de base para insertar un registro en la estructura, establecen el encadenamiento de pointers por un procedimiento standard de programa. En el caso de estructuras de Arbol múltiple, esta facilidad puede encarecer la longitud de los registros, por prever para cada uno de ellos un conjunto de campos para escribir direcciones en previsión de todas las situaciones que pueden presentarse por aplicación de un procedimiento general.

En los casos de estructuras de Arboles múltiples, donde las multiplicidades del Arbol vienen dadas por la naturaleza de los datos a relacionar, es preferible que los algoritmos de encadenamiento, sean realizados por los analistas que proyectan el almacenamiento y utilización de esos datos, pues conocedores de la tipología del Arbol que manejan, podrán diseñar la estrategia de utilización de pointers minimizando posiciones de almacenamiento. En el caso del Arbol binario, o ternario, etc., la estrategia de direccionado puede ser universal, porque el tipo de estructura es obligado e independiente de la naturaleza de los datos.

1.4.7. REDES

Las Redes, son estructuras lógicas que pueden representarse por un grafo conexo, en el cual no existen loops, pero a diferencia del Arbol, pueden existir "circuitos", o dicho de otra manera, no existe ningún registro cabecera, y cada uno de los registros ha de estar necesariamente "conectado", pudiendo ser varios, tanto el número de registros a los cuales conecta, como el número de registros por los cuales es direccionado. Se representa en la figura 18 un ejemplo de una estructura de Red. Veamos que a diferencia del Arbol, el registro *C* no solamente direcciona a los registros *D*, *F*, *G* y *H*, sino que además es direccionado por los registros *A* y *B*.

La Red, es una estructura que puede llegar a ser tan sofisticada como sea preciso según las necesidades, y por reunir todas las cualidades de las estructuras más elementales que hemos visto, ya que de hecho es una generalización o un conjunto de todas ellas, presenta además una gran flexibilidad para su modificación y readaptación, podemos presentarla como la estructura básica general que se utiliza en las bases de datos. Tomando como patrón una estructura de Red, podemos relacionar conjuntos de datos por complejos que sean.

Basándonos en la Red como soporte estructural de relaciones lógicas entre los datos, vamos a presentar en los puntos siguientes, el concepto de *Set* y el concepto de *Esquema*, que han sido estudiados por el Data Base Task Group de la Conferencia CODASYL, y propuestos como especificaciones básicas para diseño de las bases de datos. Estas definiciones, son hoy en día universales, y muchos productores de software de base de datos, se han acogido a ellas. Nosotros a lo largo de la obra, pre-

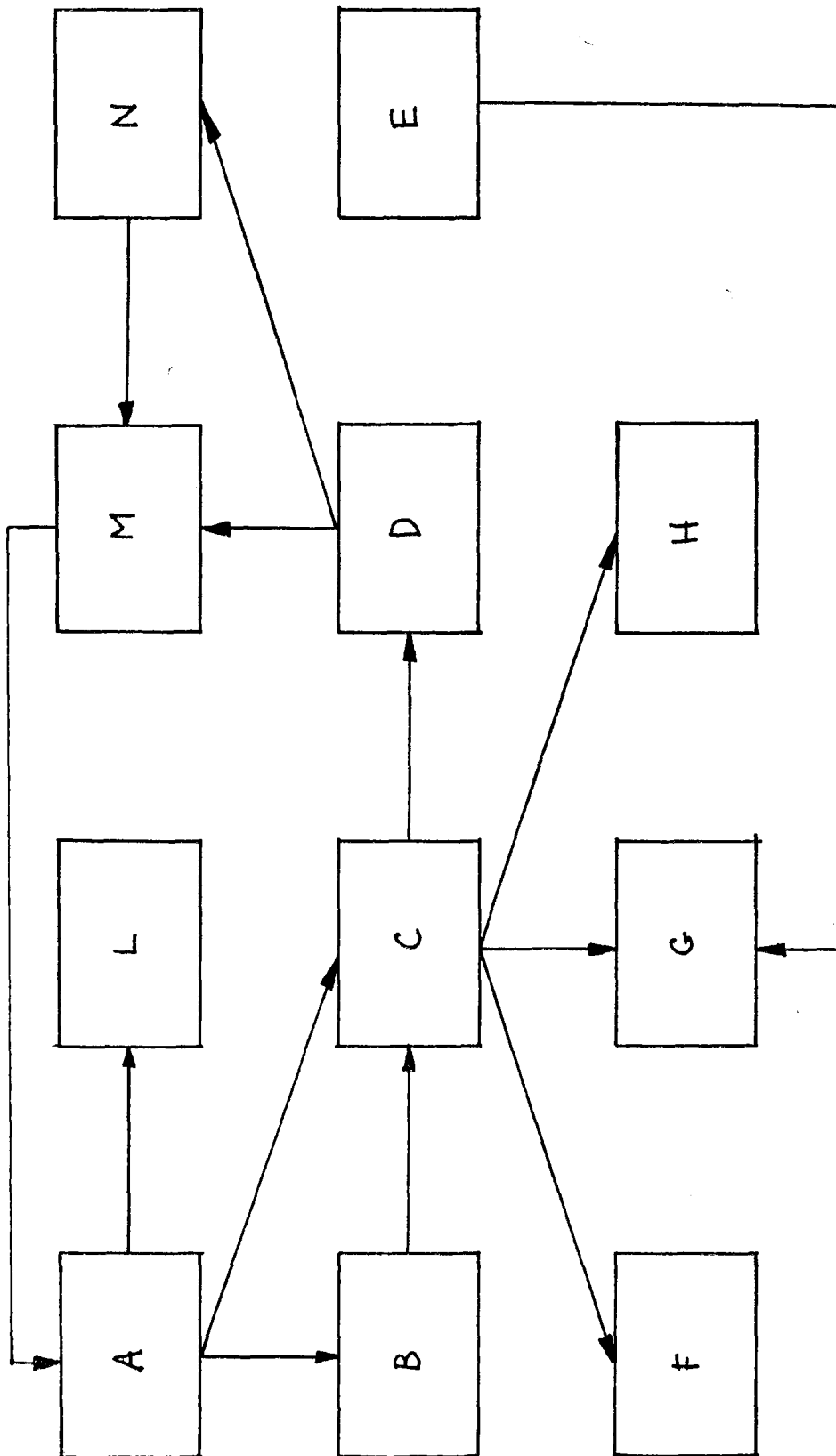


FIG. 18. *La red*

tendemos tomarlas también como patrón básico, y ello por dos razones, en primer lugar porque considero como lícitos y deseables, los objetivos de CODASYL, de buscar lenguajes y principios comunes entre diferentes usuarios, con independencia del software y de los ordenadores utilizados, y en segundo lugar, porque consideramos las especificaciones básicas del Data Base Task Group, las más completas hoy en día disponibles a nivel operativo, y las de uso más universal.

1.4.8. EL SET

El Set es un anillo tal y como se ha definido en el punto 1.4.5, y tiene, por lo tanto, un registro principal (en inglés suele denominarse *owner*) y unos registros miembro.

El "Set tipo", establece la regla lógica para la formación de anillos, y se representa en la figura 19. En dicha figura, se recoge de forma simbólica, la estructura lógica de relación que se determina de esta forma y puede existir entre los registros del tipo *A*, *B*, *C*, *D*. Esta simbología expresa que de todos los posibles Sets que pueden formarse ajustándose a ese patrón, es obligado que el registro principal sea un registro de la clase *A*, o lo que es lo mismo, del fichero *A*, y que los registros miembro, cuyo número puede ser desde 0 hasta un número ilimitado, pertenezcan a los ficheros *B*, *C* y *D*. En la figura 20, representamos un Set que puede existir en la Base de Datos, ajustándose a la definición lógica que en forma simbólica se representa en la figura 19. A estos Sets, los llamaremos Sets en general, denominando Sets Tipo al patrón al cual se ajustan (figura 19).

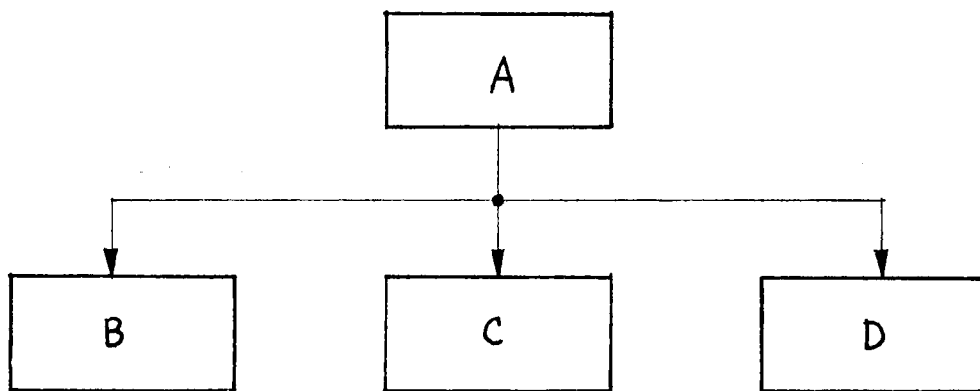


FIG. 19. El set tipo

El Set incluye además una información complementaria que veremos al describir el sistema DBMS (Data Base Management System), que determina el criterio de ordenación de los elementos que forman los anillos, y otras descripciones que completan los criterios directores a los que dichos anillos deberán siempre ajustarse.

Un Set puede ser "vacío", porque sólo contenga el registro principal y carezca de miembros, sin embargo, ningún Set puede prescindir del registro principal, por-

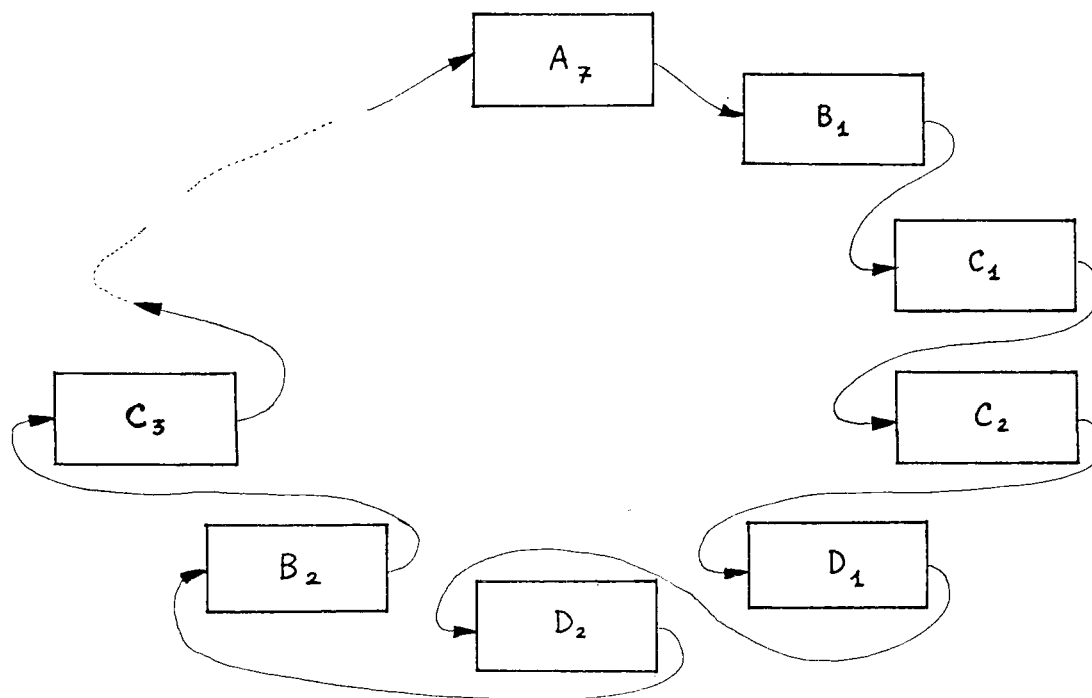


FIG. 20. El set

que un anillo, como ya vimos, no puede prescindir de su registro principal. Por lo tanto, cuando en la Base de Datos tengamos incluida una definición lógica como la expresada en la figura 19, y almacenemos un registro del tipo *A*, automáticamente dicho registro provoca la creación de un Set, en ese momento vacío, que tiene por registro principal, dicho registro del tipo *A*. Posteriormente, podrán ser almacenados registros del tipo *B*, *C* y *D*, que eventualmente podrían integrarse en dicho anillo creado, o, por supuesto, en cualquier otro anillo que tenga por registro principal a uno del tipo *A*. La figura 19, y las especificaciones complementarias citadas, establecen la reglamentación obligatoria, por tratarse de normativa de uso general para la Base de Datos. Sin embargo, criterios particulares serán los que establezcan para cada registro del tipo *B*, *C* o *D*, cuál es el anillo de todos los encabezados por un registro del tipo *A* en el que el registro debe encuadrarse, y una vez seleccionado el anillo, cuál es la posición que en el interior de dicho anillo, el nuevo registro debe de ocupar.

El Set, es la célula básica de una base de datos tipo DBMS.

1.4.9. EL ESQUEMA

El Esquema, es la definición más completa que puede establecerse de una base de datos, y constituye el modelo lógico por el que se ha de regir todo el almacenamiento de la información.

El Esquema, incluye, por lo tanto, una definición completa de los registros que van a integrarse en la Base de Datos, y una definición completa asimismo de los Sets que en dicha Base de Datos van a existir. La representación simbólica del Esquema, se suele presentar como el ejemplo expresado en la figura 21. En dicha figura, se representan todos los registros tipo que la Base de Datos reconoce, estableciéndose las posibles relaciones que pueden presentarse. La Red representada en la figura 21, constituye, por lo tanto, un modelo de cualquier estructura que la Base de Datos vaya a contener, no permitiéndose el establecimiento de relaciones lógicas no previstas en este Esquema.

En el Esquema, representado en la figura, se incluyen por lo tanto todas las definiciones de Set previstas para la Base de Datos. Un Set tipo, es el encabezado por el registro tipo *A*, incluyendo como miembros a los registros tipo *B*, *C*, *D*. Otro Set tipo sería el formado por el registro principal *D* con los registros *G* y *H* como miembros. Otro Set sería el encabezado por el registro principal *J* y los registros *D* y *H* como miembros... etc.

Como podemos ver, un registro puede ser registro principal de un Set y a la vez figurar como miembro de otro Set o de varios Sets, como es el caso del registro *D*.

El Esquema simbólico representado, naturalmente se complementa con un conjunto de definiciones que más adelante examinaremos y que representan el "estatuto jurídico" al que la Base de Datos ha de ajustarse. Dicho conjunto de definiciones, pueden ser objeto de modificación a lo largo de la vida de la Base de Datos, porque por propia definición, la Base de Datos debe de ofrecer flexibilidades. El que oportunamente se modifique el Esquema de la Base de Datos, no quiere decir que la Base de Datos deba de ser construida de nuevo, pudiendo en muchas ocasiones los registros, permanecer almacenados en sus mismas localizaciones, y sin embargo, establecer nuevas estructuras de relación actualizando los encadenamientos lógicos.

Cada programa que va a almacenar o recuperar registros de la Base de Datos, necesita incluir en sus instrucciones declarativas, la tipología de registros que va a utilizar, así como las estructuras lógicas entre dichos registros que están previstas en el sistema. Esto se realiza teniendo almacenada en la librería del sistema, una definición del Esquema, definición, la cual el programa puede invocar.

Sin embargo, sería engorroso y muchas veces inconveniente, el que cualquier programa de utilización, por elemental que fuera, utilizara la definición total de la Base de Datos. Para evitar esto, se pueden establecer definiciones parciales de la Base de Datos que son sub-esquemas incluidos en el Esquema general, pero que los llamamos igualmente Esquemas por constituir esquemas que pueden tener vigencia independiente, aunque constituyan una parte del Esquema total.

En la figura 21, se resalta un sub-esquema, que está formado por los registros tipo *A*, *B*, *C*, *D* y *F* y por los Sets *A/F* y *A/B*, *C*, *D*.

Pueden definirse tantos sub-esquemas como sea preciso, con la única limitación de que dichos sub-esquemas estén inscritos en el Esquema global. La utiliza-

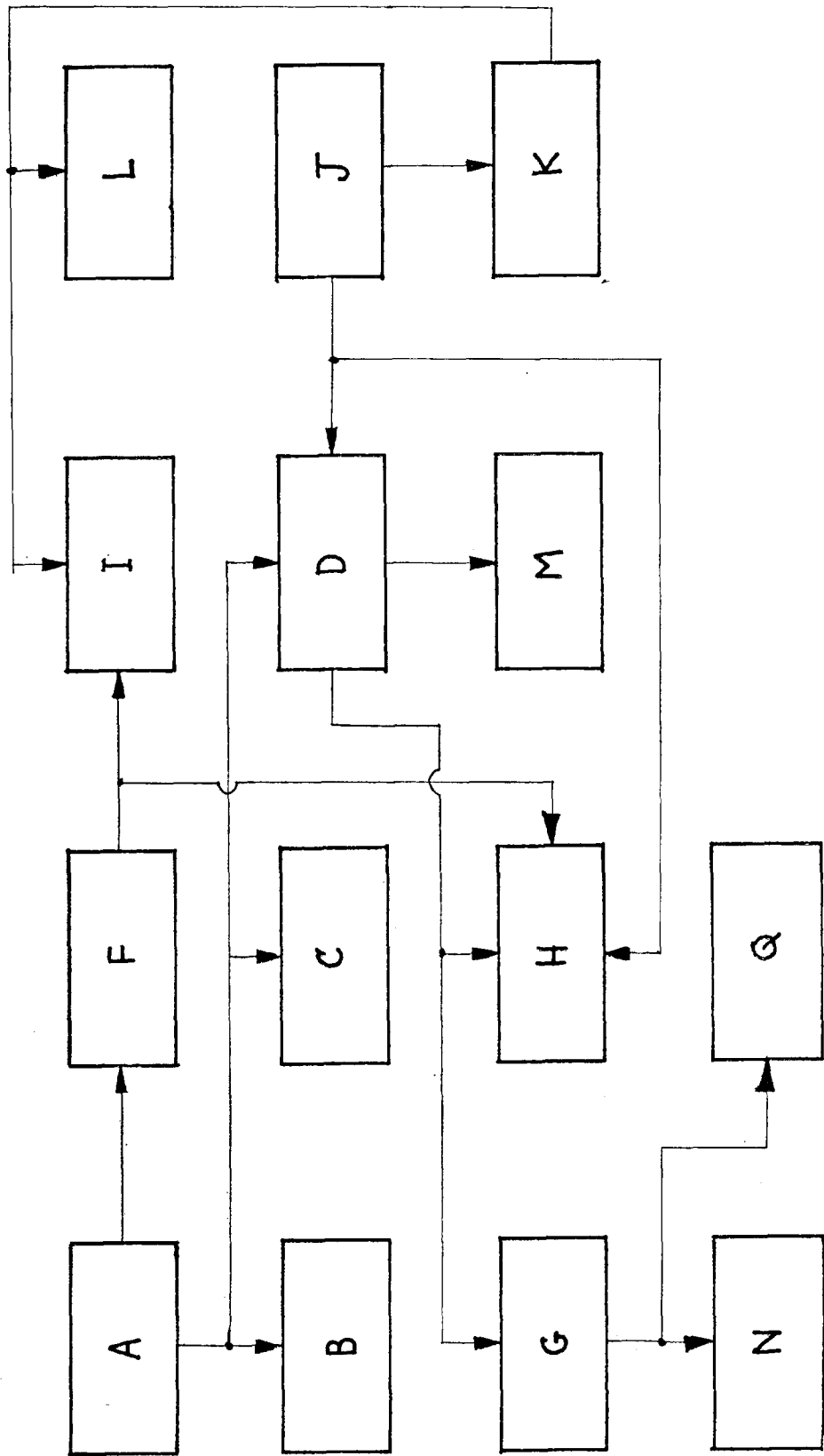


FIG. 21. El esquema

ción de sub-esquemas, permite mayor agilidad en la explotación de programas, y ofrece mayores seguridades de protección de la información, porque ningún programa podrá deteriorar aquella información a la cual no le es permitido acceder, característica que por otro lado, permite establecer seguridades de confidencialidad. Así, podemos utilizar informaciones muy reservadas que por un lado conviene que estén en la Base de Datos porque constituyen piezas importantes de relación lógica, sin que por otra parte, se ofrezcan a ser manipuladas indiscriminadamente.

1.5. CODIFICACION DE LOS DATOS

La codificación, es un paso previo obligado para el tratamiento mecanizado de la información. Aquellos campos de información que han de ser procesados, deben de estar codificados respetando criterios preestablecidos que son “conocidos” por los programas que utilizan los datos, pudiendo por lo tanto realizarse procesos en base al contenido lógico de significación que los códigos comportan.

Codificar consiste en asignar símbolos a entidades, utilizando alfabetos de símbolos previamente convenidos. El conjunto de símbolos obtenidos después de codificar una entidad, se denomina código, y el conjunto de códigos, de esta forma obtenidos, constituye un diccionario.

La codificación nos es a todos familiar, pues de hecho se utiliza constantemente en todos los órdenes de la vida. Todas nuestras percepciones utilizan en una u otra forma, diversos alfabetos de símbolos como elementos de soporte para la transmisión de información. El lenguaje natural, es de hecho una codificación. La gramática constituye el conjunto de criterios y normas para aplicar correctamente el alfabeto de símbolos a nuestras expresiones.

Por lo tanto, al codificar los datos para ser procesados por el ordenador, no estamos aplicando nuevas técnicas, sino únicamente unas gramáticas más o menos restringidas que han de servirnos a unos objetivos precisos.

Generalmente, se utilizan codificaciones cuya gramática sea fácilmente programable, ello en parte porque todavía arrastramos una práctica que era obligada en los primeros ordenadores, debido a su escasa capacidad de proceso y almacenamiento de información. En parte, también utilizamos codificaciones sencillas porque una gramática compleja, es mucho más cara de programar y utilizar, y debe por lo tanto evitarse siempre que no sea imprescindible o la naturaleza de la aplicación no lo exija. De hecho, hoy en día existen compiladores que utilizan subconjuntos bastante amplios del lenguaje natural y por lo tanto, técnicamente ya es posible el que los datos procesados en un ordenador, tengan una potencia de significación análoga a la que utilizamos nosotros mismos para expresarnos.

Sin embargo, codificaciones tan complejas, sólo se utilizarán en aplicaciones muy específicas o bien en un futuro quizás no muy lejano cuando nuevas tecnologías proporcionen mayores capacidades de proceso y almacenamiento a costes unitarios mucho más reducidos.

A la codificación en general, no le ha concedido el informático, toda la importancia que realmente tiene, aun cuando muchas aplicaciones han requerido ser reprogramadas de nuevo por falta de flexibilidad de las codificaciones utilizadas. En las aplicaciones de base de datos, la codificación adquiere una transcendencia mayor, porque en realidad codificar es estructurar y la Base de Datos es esencialmente un sistema muy estructurado que se apoya esencialmente en la disponibilidad de datos que necesariamente han de estar a su vez muy estructurados.

En la Base de Datos, toda la estructura lógica de relación, está concebida y soportada en base a códigos y por lo tanto, el "sistema de datos" será tan potente como la potencia de significación de los códigos, permita. Las relaciones lógicas que se establecen en la Base de Datos, descansan en códigos, porque en general los criterios de decisión para la determinación de los pointers de conexión, se aplican en base al significado de campos de registros que obviamente han de estar codificados.

Además, la Base de Datos, tiene que ser muy flexible porque debe permitir adaptaciones a lo largo de su vida sin necesidad de grandes reorganizaciones, lo que se consigue con un buen diseño y una buena administración, en los que la codificación siempre supone cuando menos, un conjunto de limitaciones. Tengamos en cuenta que hemos de servir a dos objetivos encontrados, estructuración y flexibilidad.

Vamos a tratar solamente los aspectos lógicos de la codificación, dejando a un lado el problema técnico de la transcripción de los códigos lógicos, utilizados en las aplicaciones informáticas, a los códigos binarios que necesariamente requiere la unidad central de proceso de un ordenador.

Tampoco examinaremos la problemática técnica de la transmisión, transcripción y proceso, de dichos códigos binarios. Nos ceñiremos a obtener una panorámica de la codificación que se realiza a nivel lógico, ocupándonos solamente de los códigos que utilizan los programas, y que representan a la información externamente al entorno del ordenador.

Unas veces, la codificación es un problema exclusivo para el proceso de los datos, pero generalmente, los códigos que manejamos van a ser utilizados en mayor o menor parte por entornos humanos de la organización, ajenos a la informática. Debe por ello analizarse con mucha seguridad, cuál es la información que vamos a codificar, y cuál es la información que no precisa de códigos específicos porque los procesos que va a experimentar, son solamente transcripciones. Optimizando la codificación, optimizamos la potencia de significación que es un factor importante en el rendimiento.

1.5.1. TIPOS DE CODIGOS

Vamos a repasar los tipos de códigos que generalmente se utilizan en informática, además del código binario en el que necesariamente, siempre se transforma la información para ser almacenada y procesada.

Atendiendo a la naturaleza de los símbolos utilizados, podemos clasificar los códigos en:

- *Alfabéticos*. Los símbolos utilizados son las letras del alfabeto del lenguaje, normalmente el alfabeto latino. Los códigos alfabéticos, poseen gran amplitud por ser relativamente grande la variedad de las letras. Las posibilidades nemotécnicas de estos códigos, son aceptables cuando se utilizan, además, para significaciones externas y se relacionan con códigos simplificados que extraemos del lenguaje, tales como siglas ... et-cétera. Como inconveniente de los códigos alfabéticos, cabe citarse que el control de errores en las transcripciones, es complicado.
- *Numéricos*. En estos códigos se utilizan los números naturales (números árabes). Las posibilidades de control de estos códigos, es superior a la de los códigos alfabéticos, pero en contrapartida, las posibilidades nemotécnicas son menores.

Los códigos numéricos pueden estar representados en un determinado sistema de numeración, decimal, hexadecimal, binario, ... etc., y pueden experimentar determinadas traducciones como es la condensación o empaquetamiento, ... etc.
- *Alfanuméricos*. Se trata simplemente de la combinación de números y letras. Por extensión, también pueden incluirse otros símbolos convenidos, como \times , $+$, $-$, $/$, ... etc.
- *Mixtos*. Son los códigos compuestos por varios campos de códigos.
- *Simbólicos*. No nos ocuparemos de estos códigos, por no ser digitalizables.

Atendiendo al criterio de asignación de los códigos, podemos también clasificarlos de la siguiente manera:

- *Códigos referenciales*. Son códigos que carecen de significación intrínseca y su única relación con la entidad que representan, es la correspondencia biunívoca, que establecemos mediante su asignación. Su aplicación es por lo tanto sencilla, y puede ser también muy económica porque la longitud del código, puede reducirse a un mínimo.

Sin embargo, la utilización de este código, puede no ser adecuada, porque el acceso a un registro de información identificado por un código de este tipo, requiere la utilización de un diccionario, lo que nos puede llevar a procesos especiales.

Además de referencial, el código puede ser secuencial, es decir, ser asignado a los correspondientes ítems en forma correlativa. A veces, se desea una cierta ordenación de las entidades representadas, lo que exige la posibilidad de intercalar ítems en la secuencia, disponiéndose en estos casos, de códigos de reserva intercalados en la secuencia, que no están asignados a ningún ítem y que se utilizarán posteriormente para intercalaciones. Se suelen llamar entonces, códigos secuenciales con reservas.

- *Significativos.* Este código, contiene una significación intrínseca, es decir, del examen de los símbolos por los que está compuesto, se extrae la significación de la entidad representada. Son por lo tanto códigos muy estructurados y muy potentes, pero en contrapartida, contienen por lo general mayor número de símbolos que los códigos referenciales, y pueden llegar a ser muy caros, poco manipulables por personas, si se pretende descender a muchos niveles de significación.

La significatividad de los códigos, facilita ciertamente los accesos de la información, porque la estructura del código, permite seleccionar los caminos lógicos adecuados para los accesos. Los códigos significativos, aportan un contenido estructural a la Base de Datos y no sólo facilita determinados procesos de extracción de información, tales como la obtención de estadísticas, sino que el propio concepto de “red lógica de datos”, se aplica con la utilización de la potencia de significación interna de determinados códigos. Invito al lector a que tenga presente esta idea a la hora de realizar algunos ejercicios de diseño de un Esquema.

Un código muy significativo y muy estructurado, exige que sea diseñado por un buen conocedor de las utilidades que el código ha de reportar para que los criterios aplicados sean los correctos, ya que a mayor estructura hay mayor potencial, pero si el diseño no es el adecuado, también hay muchas más limitaciones y falta de flexibilidad, no estando tampoco ausentes los criterios de economía.

- *Intermedia.* Designamos genéricamente como intermedia, a la codificación más frecuente que contiene símbolos significativos y símbolos simplemente referenciales. También puede componerse por un conjunto de campos, pudiendo ser la codificación de cada campo, secuencial, significativa o mixta.

Cuando el código está compuesto de varios campos, hay que tener en cuenta que la falta de flexibilidad de uno solo de los campos, puede producir la rigidez de todo el conjunto.

1.5.2. CARACTERÍSTICAS DE LOS CODIGOS

Examinemos las características fundamentales de los códigos:

- *Longitud.* La longitud del código se expresa por el número de símbolos que contiene. La cantidad de información contenida en cada símbolo es

el número de bits necesarios para codificar en binario cada símbolo (primer teorema de Shannon).

Si para codificar los N símbolos de un alfabeto son necesarios “ K ” “bits” podríamos diseñar una transformación de nuestra codificación lógica a la codificación binaria, de uso posterior obligado, que permitiera optimizar la economía de unidad central de proceso y posiciones de almacenamiento. Sin embargo, esto es un problema de software de base, y se utilizan sistemas internacionales para la traducción a binario, que compatibilizan la economía, con la modularidad y universalidad necesarias.

A nivel lógico, también podemos minimizar la longitud del código si utilizamos asignación secuencial de códigos referenciales. Esto nos daría la longitud mínima necesaria de un código para segregarse un conjunto de entidades. A partir de esa longitud, determinamos la longitud óptima del código que será mayor en la medida en que queramos incluir en él mayores niveles de significatividad intrínseca. Se selecciona generalmente una longitud $L > L$ mínima, suficiente para referenciar un conjunto presente y futuro de entidades, de acuerdo con los objetivos que el código haya de cumplir.

En códigos de longitud variable, la longitud media $\bar{L} \equiv \sum P_i \bar{L}_i$, es la suma de las diferentes longitudes posibles con sus diferentes respectivas probabilidades. Por ello en estos casos es aconsejable dar las mayores longitudes de código a las entidades de menos probabilidad de proceso y/o almacenamiento, con lo cual la longitud media del código será lo menor posible, con la consiguiente economía.

Las longitudes de los códigos, como veremos, casi siempre hay que incrementarlas con respecto a las estrictas necesidades lógicas, ello por necesidades de redundancia y de control.

- *Capacidad.* El coeficiente de capacidad de un código puede expresarse por el cociente entre el número de elementos a codificar y el número de combinaciones obtenibles.

El coeficiente de capacidad debe de permitir la incorporación de nuevos elementos al diccionario de códigos, con la suficiente proyección de futuro, por lo que dicho índice debe seleccionarse a la vista de la tasa demográfica de la familia de ítems referenciados.

Hay que tener en cuenta que la capacidad queda mermada por la aplicación de códigos significativos, siendo necesarios los excesos de capacidad no sólo para permitir volúmenes de nuevos códigos, sino para posibilitar también la segregación de los elementos.

Cuando un código se compone de varios campos, la capacidad del código está limitada por la capacidad de cada uno de sus campos.

- *Redundancia.* En toda transcripción, transmisión o proceso de informa-

ción, está presente el principio de Entropía y, en consecuencia, es preciso disponer redundancias en los datos.

Por lo tanto, si la longitud de un código es N por estrictas necesidades informáticas, se pueden disponer K símbolos de forma que existan " r " = $K-N$ dígitos redundantes, ello si se desea contrarrestar el efecto de las "pérdidas" de información.

Podemos designar por coeficiente de redundancia $R = r/K = 1 - N/K$, siendo N/K un índice de la eficacia del código. Para un mismo nivel de calidad en las transcripciones, transmisiones y procesos, en cuanto más eficaz es un código, menor redundancia se requiere.

Las redundancias se establecen en forma de campos de información que permitan realizar cuadros de control, la utilización de dígitos de control, etcétera.

— ESTRUCTURA O SIGNIFICACION

Esta característica es la que determina el nivel de significación del código, no puede medirse de una forma efectiva, aunque una idea de lo estructurado que está el código nos la da el excedente de capacidad que es necesario disponer en el código por necesidades de diferenciación de elementos de la familia de entidades representadas.

El nivel de estructura del código ha de ser cuidadosamente establecido porque condiciona el futuro como estructura que es. Su relación con las necesidades de la Base de Datos es estrecha porque la significación de los códigos es una potencia de relación lógica.

— SIMPLICIDAD Y SEGURIDAD DE USO

Esta característica tiene interés cuando los códigos han de ser manipulados por personas externamente, al ser mecanizado. Los "procesos" que realizan las personas están situados "en serie" con otros procesos que el ordenador realizará con esos mismos datos. Por eso, en la medida en que racionalicemos estos códigos tendremos menos errores en la manipulación y por lo tanto mayor economía, al poder reducir las redundancias y controles.

Por eso se suelen fraccionar los códigos en varios bloques cuando han de ser manipulados y transcritos. La demostración más elocuente de esta conveniencia es el gráfico de la figura 22, en el que se recoge el resultado de una muestra estadística (realizada, según creo, en el Instituto de la Percepción del Reino Unido), que recoge porcentajes de códigos transcritos correcta e incorrectamente, en función del número de símbolos del código. Como podemos ver, a partir de seis dígitos el rendimiento comienza a caer. En términos vulgares. "el número de teléfono" o la

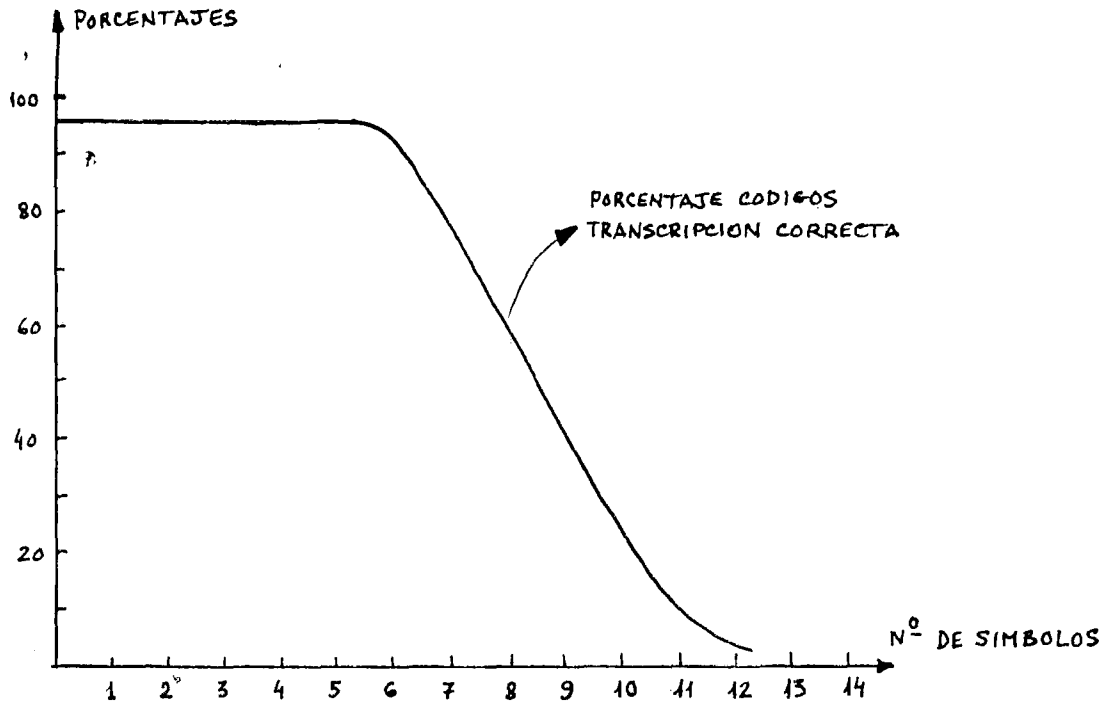


FIG. 22

“matrícula del coche” vienen a representar, más o menos, umbrales a estos efectos.

- *Uniformidad.* La característica de uniformidad es la de igualdad entre longitudes de los campos de los códigos. Esta uniformidad, a primera vista resulta onerosa por exigir que el código tenga longitud suficiente para albergar el código específico de mayor longitud que ocasionalmente pudiera presentarse.

Sin embargo, manejando bien las oportunidades de relación lógica que ofrece la Base de Datos, los códigos pueden fraccionarse en sub-códigos oportunamente relacionados, de forma que un campo de información pueda ser complementado cuando el caso lo requiera, por otro conjunto de campos, a los que él mismo da acceso.

Con estas precauciones pueden utilizarse en las bases de datos, registros de longitud fija, sin que ello suponga una penalización económica. En las bases de datos, es muy deseable utilizar registros de longitud fija, poniendo todo el mosaico de información lógica a partir de registros básicos tipificados de una manera uniforme.

1.5.3. CONTROL DE LOS CODIGOS

Como hemos visto, la transcripción, transmisión y proceso de los códigos, no escapa al principio de la Entropía y, por lo general, se aplican controles a los códi-

gos para detectar y depurar la tasa de errores que necesariamente habrán de producirse. La información redundante se utiliza tanto como información de control intrínseca contenida dentro del código como para superar posibles pérdidas en campos de información. Además de esta información redundante que se incluye en los registros lógicos, se prevén redundancias internas de los sistemas de transmisión y manejo de la información, tales como bits de paridad, protecciones de memoria, etc.

Las formas más comunes de controles directamente asociados con la información codificada, son:

- *Control de composición del código.* Se trata de comprobar como primera medida, la validez del código mediante una comprobación de si cumple o no las reglas establecidas. Estos controles se realizan a nivel programa de proceso, y entran dentro de lo que hemos denominado controles de forma, al tratar del control de la información de entrada a la Base de Datos.
- *Autocontrol.* Es un control que se ejerce por un programa de proceso apoyándose en la propia información intrínseca del código. Dicha información redundante de control puede consistir en campos de información que se utilizan para practicar “cuadros” con el resto de la información del código. La forma más general de autocontrol se realiza en base a dígitos claves o dígitos de control, existiendo muchos procedimientos de aplicación, que varían según el grado de certeza que se requiera en la validación del código. Son sobradamente conocidos en la buena práctica informática.
- *Control por incompatibilidades de campos.* Se trata de un control lógico al que hemos denominado anteriormente control de verosimilitud y del que ya hemos tratado.

En el manejo de los códigos debe de tenderse en lo posible a “descentralizar” el control, debiendo de entenderse este concepto como la asignación de la responsabilidad de la calidad del código a la entidad que realiza una transcripción. Tengamos en cuenta que el ejercicio del control supone también el reacondicionamiento de la información y, generalmente, es la fuente que incurrió en los errores, la entidad más capacitada para subsanarlos. Realmente, el especializar personas para realizar reacondicionamientos de información errónea, no tiene demasiado sentido y significan, por lo general, crear gremios laborales de baja productividad. Así están indicados estos especialistas, cuando han de realizar interpretaciones programadas para transcripción de determinados códigos confidenciales, pero en este caso ello supone ya una función diferente.

En la actualidad se ha seguido esta idea en las máquinas contables que se utilizan para multitud de procesos administrativos. A estas máquinas se les dota de un mecanismo de verificación de las operaciones efectuadas, con lo cual los documentos producidos o incluso los soportes informáticos obtenidos como subproducto (cintas perforadas, cintas magnéticas, etc.) han experimentado ya una depuración en origen. Con ello, por un lado se descarga a los ordenadores centrales del ejercicio

de estas tareas y por otra parte, se minimizan los ciclos de corrección de errores, mejorándose por lo tanto el grado de actualidad de la información.

1.5.4. CRITERIOS PARA LA ELECCION DE CODIGOS

Si se ha destacado la problemática de la codificación como tema de interés, dentro del concepto de explotación de bases de datos, es porque se trata de una estructuración. La Base de Datos se apoya en un concepto de “estructura flexible”, que constituye un cierto contrasentido racional. La estructuración viene a ser una especie de cristalización o modelación que en gran parte es irreversible, pues las modificaciones sensibles pueden obligar a comenzar de nuevo el planteamiento. Los grandes reptiles de la era terciaria acabaron siendo exterminados por su propia estructura.

La elección de los códigos y sus reglas semánticas forman parte de todo el planteamiento de la estructura lógica y han de dedicarse a ello los mejores momentos y recursos de estudio dentro de la etapa de diseño en el proyecto. Es una tarea por lo tanto a ser dirigida en detalle por el responsable del proyecto, quien debe poseer la suficiente madurez y práctica informática para aplicar el criterio básico, que no es otro que el del sentido común.

El planteamiento de la codificación, es diseñar una estructura y habrá que desarrollarla hasta darle el punto necesario de equilibrio. Ello se realiza determinando el nivel de significatividad de los códigos y el nivel de complejidad de las respectivas gramáticas. Habrá que tener presente los siguientes criterios:

- *Volumen de datos.* Para cada familia de datos, la longitud del código estará dada por el grado de amplitud que determine su volumen. Por otro lado, a mayor significatividad del código, hay una mayor segregación de entidades codificadas. La longitud de la parte no significativa del código ha de ser suficiente para albergar a la mayor población posible de las entidades segregadas.
- *Dinámica de cambio.* La codificación inicial debe prever futuras ampliaciones, sin que ello implique reorganizaciones de la Base de Datos. Por ello, o bien el código tiene una proyección futura que le permita ampliarse sin necesidad de reorganizar los datos, o bien hay que prever un código de suficiente amplitud para albergar futuras ampliaciones. Ello quiere decir que o bien diseñamos un código flexible o bien un código suficientemente amplio.
- Si la familia de datos es numerosa y la tasa demográfica es muy elevada, el prever códigos muy amplios para necesidades a largo plazo puede resultar oneroso a corto plazo, y en este caso habrá que disponer un código que sea ampliable por etapas. Lo general será que el código pueda preverse desde un principio con un grado de amplitud suficiente para las necesidades futuras. Para determinarlo es preciso calcular con suficiente preci-

sión los volúmenes de datos de las familias, realizando previsiones suficientemente fiables de las tasas de nacimiento y mortalidad de los registros de cada familia.

- *Estabilidad de los datos.* El grado de estabilidad de los datos viene a darnos una indicación del “nivel de compromiso” que tenemos al codificar. Los datos permanentes son piezas firmes del modelo de empresa que constituye la Base de Datos y, según su relevancia en la estructura, deberán estar tratados con mayor precisión los pilares básicos que sujetan un edificio.
- *Tipos de acceso y niveles de uso.* Hemos visto que los códigos referenciales han de ser traducidos por un diccionario, y conducen a procesos secuenciales, mientras que los códigos significativos permitían una mayor o menor selectividad agilizando los accesos a los registros. Para cada familia de datos deberá analizarse qué tipos de acceso son los convenientes y qué volúmenes de acceso se prevén, para diseñar unos códigos compatibles con el criterio más general de rentabilidad y eficacia en la explotación de la Base de Datos. Obviamente, para altos niveles de acceso habrá que emplear códigos más significativos mientras que para datos poco utilizados podrán determinarse códigos más referenciales, que siempre son más económicos. El equilibrio en la economía y en la eficacia, deberá conducirnos a no “agrandar” excesivamente las áreas de almacenamiento, pero sin embargo mantener los tiempos de acceso dentro de niveles suficientes de eficacia.
- *Utilizaciones externas.* Los códigos no son privativos del tratamiento mecanizado, sino que de ellos se hace un uso más o menos grande por parte de otras áreas de la Organización, externas al Centro de Cálculo. Los procesos externos que las familias de datos vayan a experimentar es un criterio más a la hora de diseñar los códigos. Recordemos la figura 22.

1.6. DEFINICION DE LA BASE DE DATOS

En las aplicaciones convencionales, cada aplicación informática utilizaba unos determinados ficheros de información, que tenían una entidad independiente. Estos ficheros únicamente entraban en comunicación por el hecho de que un programa los procesara a la vez y por el hecho de que dicho programa contuviera por lo tanto su definición en las correspondientes instrucciones declarativas.

El concepto de base de datos aporta una red de ficheros integrados en una estructura lógica de tal forma que los ficheros sean independientes de los programas. Un fichero ya no nace porque nazca un programa que necesita de él, sino por-

que existen unas informaciones relevantes en la Organización, que constituyen piezas estimables en toda la estructura lógica de datos. Después, los ficheros serán más o menos complejos y tendrán mayor o menor riqueza de detalles en función de las aplicaciones que vayan a plantearse.

La Base de Datos es una entidad que surge con independencia en los programas, y por lo tanto ya no es definida por las instrucciones declarativas de éstos, sino que es preciso establecer una definición propia. Los programas de utilización seguirán requiriendo instrucciones declarativas, pero éstas serán utilizadas para invocar el uso de las partes precisas de la Base de Datos con las que la definición de datos debe coincidir.

La tarea de definición de la Base de Datos consiste en la definición por medio de instrucciones declarativas de todo un modelo de empresa que es la estructura lógica diseñada. Esta definición se realiza utilizando lenguajes de programación “ad-hoc” que forman parte del entorno operacional del sistema de base de datos.

La potencia de definición que podemos conseguir dependerá del sistema general de software y de los ordenadores, que hemos seleccionado para nuestro sistema de base de datos, aunque naturalmente, siempre puede diseñarse un software “a medida”, cosa que sólo está justificada en aplicaciones muy concretas y de mucha envergadura.

Generalmente será un sistema de los disponibles en el mercado el que hayamos seleccionado, y en base al lenguaje de definición de datos que el mismo contenga estableceremos la correspondiente definición. El diseño previo del esquema, de los registros y de las relaciones lógicas ya se habrá realizado de acuerdo con las potencias que ofrezca el sistema de software elegido.

Bien sea nuestro sistema un sistema “host-language” o un sistema “self-contained”, normalmente contendrá un medio para definir la estructura de datos. En los puntos 1.6.4. y 1.6.5 examinaremos cómo se presenta la definición de datos en los sistemas propuestos por CODASYL y por “GUIDE/SHARE”, aunque dichos sistemas se examinarán después en conjunto.

Los objetivos fundamentales de la definición de datos son:

- Conseguir una independencia entre datos y programas, es decir, poder cambiar las aplicaciones sin necesidad de reorganizar la Base de Datos y, a su vez poder reorganizar la Base de Datos sin necesidad de modificar los programas, al menos de forma sensible.
- Conseguir una independencia entre la estructura lógica y la estructura física, es decir, la problemática del almacenamiento físico de la información, está solucionada por la definición de la Base de Datos, siendo transparente la estructura física para los programas de aplicación.
- Integración de datos, objetivo suficientemente comentado, que evita redundancias, permite mayor economía y potencia las posibles aplicaciones.
- Administración de la Base de Datos. El sistema de definición de datos es pilotado por el administrador de la Base de Datos, ejerciéndose un control

de este recurso informático porque es patente que bajo este concepto de explotación, los datos constituyen un recurso que es aplicado en todo el seno de la Organización a través de los correspondientes sistemas informáticos. Los datos ya no deben concebirse como patrimonio de un determinado departamento ni su control deberá ser ya ejercido por aplicaciones informáticas específicas.

La definición de los datos se desarrolla utilizando un lenguaje independiente de las aplicaciones, lenguaje que utiliza instrucciones de tipo descriptivo. Las definiciones "fuente" son compiladas y pasan a formar parte en su versión "objeto" de la biblioteca de la Base de Datos, quedando almacenadas a disposición del sistema operativo, para pasar a integrarse a la memoria principal cuando las unidades de tratamiento las invocan durante el proceso de explotación de la Base de Datos.

Las aplicaciones informáticas que utilicen la Base de Datos deberán invocar en cada una de sus unidades de tratamiento una definición de la Base de Datos, pudiendo existir diferentes definiciones que pertenecen a diferentes subconjuntos de toda la estructura de datos. Cuando una unidad de tratamiento invoca una definición de la Base, esta definición, que debe, naturalmente, existir, es "leída" en su versión objeto de la posición que ocupaba en el almacenamiento, para ser llevada a la memoria como soporte para el proceso de la unidad de tratamiento correspondiente. La unidad de tratamiento debe de contener también en sus propias instrucciones declarativas una identificación y descripción de los registros y campos lógicos de información que vaya a utilizar, para "superponer" dicha definición con la que encuentra en la memoria, como perteneciente a la Base de Datos.

De esta forma se realiza la identificación de los ficheros y campos de información a utilizar. Además, se adopta de esta forma una reglamentación específica para la unidad de tratamiento, porque esta unidad de tratamiento, utiliza los registros de información de la Base de Datos, pero emplea la subdivisión en campos e identificaciones lógicas que se encuentren declaradas en la unidad de tratamiento, aunque esta posible "sustitución" de la descripción de los datos de los ficheros deba de estar en parte prevista en la definición de la Base de Datos y siempre ajustándose a unas reglas lógicas que constituyen especificaciones básicas del sistema de base de datos.

De esta forma, unos mismos campos de información pueden ser contemplados por los diferentes programas de proceso como entidades diferentes tratándose a diferentes niveles de detalle. Un fichero muy complejo puede ser declarado por una unidad de tratamiento identificando y describiendo únicamente las partes del registro tipo que el programa va a procesar, delimitando y simplificando el entorno de influencia en la Base de Datos de dicho programa y los analistas que lo producen. Además, se preserva convenientemente las zonas de fichero excluidas de una eventual corrupción o mal uso de la unidad de tratamiento.

Por supuesto, las relaciones lógicas de la Base de Datos y la correspondencia entre la estructura lógica y la estructura física de almacenamiento constituyen una unidad “no opinable” por las unidades de tratamiento, utilizándose a estos efectos la especificación única expresada en la correspondiente definición de base de datos que ha sido invocada por la unidad de tratamiento.

Los sistemas “self-contained” o los propios sistemas “host”, que disponen de software correspondiente para procesar aplicaciones interactivas con usuarios no necesariamente impuestos en las técnicas informáticas, precisan disponer de unos programas previamente almacenados, que son activados por las funciones o transacciones que el usuario realiza. Estas unidades de tratamiento, que están integradas en el sistema de base de datos, forman parte del software básico y precisan a su vez de las oportunas definiciones de las zonas de la Base de Datos a utilizar.

Estas unidades de tratamiento pueden ser consideradas como un complemento en la definición de la Base, ya que, de hecho, son utilizadas en respuesta de las funciones de interrogación que un usuario plantea desde un terminal, siendo para ese usuario transparentes dichas unidades de tratamiento, ya que él se dirige a la Base de Datos a través de un paquete de transacciones previamente convenido.

El que estos programas constituyan un “interface” o bien estén integrados en la propia definición de la Base de Datos mediante un lenguaje “ad-hoc” de procedimiento es algo que pertenece a la arquitectura del sistema.

1.6.1. IDENTIFICACION Y CARACTERISTICAS DE LOS DATOS

El conjunto básico de definición es la identificación de descripción de los registros y campos de información que van a contener los ficheros que integran la Base de Datos.

La forma de definición es similar a las establecidas en las aplicaciones informáticas convencionales, con la diferencia de que esta definición se realiza de forma independiente a cualquier aplicación. Se utiliza para ello un lenguaje descriptivo (non procedural), es decir, un lenguaje cuyas instrucciones son exclusivamente declarativas.

Mediante el uso de estas instrucciones se especifica la descripción de la composición lógica de cada uno de los ficheros considerados como unidades independientes.

La unidad a describir para cada fichero es el registro tipo, que constituye el “molde” o matriz de todos los registros miembros del fichero. Debe de incluirse por lo tanto:

- Nombre identificativo del tipo de registro o fichero, es decir, nombre simbólico bajo el cual va a ser denominado por las aplicaciones.
- Nombre simbólico o identificaciones de cada uno de los campos del registro que van a tener una personalidad individual.
- Nombres simbólicos alternativos, denominados Alias, de los registros o de campos individuales, que van a ser manejados por las unidades de tratamiento, de acuerdo con esta capacidad de manejar los datos, de acuerdo con los diferentes puntos de vista requeridos por los programas o por conveniencia de administración de la Base.
- Descripciones de las longitudes y tipos de caracteres que constituyen cada campo, en forma similar a como se realiza en las instrucciones declarativas de cualquier programa.
- Especificación de campos de información contenidos en el registro y destinados a pointers o señaladores de dirección. Estos campos van a servir para soporte de toda la estructura lógica de la Base de Datos, y su espacio debe ser previsto aquí como parte constitutiva de los registros. Según el sistema utilizado, estos campos necesitan ser previstos por los diseñadores de la Base de Datos, o bien son calculados automáticamente por el software de definición de datos en función de las especificaciones de relación lógica declaradas. En cualquiera de los casos, son campos que forman parte de los registros.
- Formas de acceso de los registros. Análogamente a los ficheros convencionales, debe declararse para cada fichero de la Base de Datos el tipo de organización, o lo que es equivalente, la forma de acceso de los registros del fichero, es decir, acceso directo, secuencial indexado, etc.

Esta definición es específica de cada fichero individual, y a nivel lógico cabe contemplarla como independiente de la Base de Datos, aunque sin embargo las organizaciones de los ficheros pertenecen al conjunto de la estructura lógica. Aprovechemos para indicar que el acceso directo en estos casos es un acceso directo a nivel lógico, porque las unidades de tratamiento no manejan localizaciones directas de la posición física de un registro, sino unas direcciones directas lógicas, que después habrán de ser traducidas por una tabla. Estas direcciones lógicas se denominan localizadores, identificadores, o también claves de base de datos.

1.6.2. DEFINICION DE LA ESTRUCTURA LOGICA

La estructura lógica está formada por un conjunto de enlaces de unión entre registros de los diferentes ficheros que integran la Base de Datos. Esta estructura de relación no existía en las aplicaciones convencionales como característica especificable a nivel de definición de datos, por lo que las relaciones lógicas habían de ser establecidas por algoritmos de proceso.

La definición de las relaciones lógicas se realiza mediante unas instrucciones declarativas que definen las características de las estructuras lógicas que van a ser utilizadas, y que, en general corresponden a entidades de tipo de las que ya hemos descrito en el punto 1.4. Si se trata de definir bien un set, un anillo, una cadena, se especifica cuál es el registro tipo correspondiente al registro principal y cuáles son los tipos de registros correspondientes a los registros miembro. También se definirá la secuencia lógica de ordenación de los registros miembro, dentro de la estructura, lo que constituye una especificación de cuál es el lugar apropiado para insertar un nuevo registro en la estructura cuando haya de ejecutarse una instrucción de almacenamiento.

Las definiciones de la estructura lógica difieren de unos sistemas de bases de datos a otros, pero en definitiva, las instrucciones de definición vienen a establecer una normativa o reglamentación de acuerdo con la cual se utilizarán los pointers previstos en la definición de los registros, formándose los enlaces de unos registros con otros, escribiéndose la dirección lógica del registro enlazado en el pointer correspondiente del registro que ha de direccionarlo.

Los “caminos” lógicos que se establecen de esta forma constituyen también canales de acceso a los registros, siendo una nueva forma tipificada de acceso de registros, la que se realiza a través de la participación del registro en la estructura lógica.

Las estructuras lógicas tipificadas en la definición constituyen también familias en la misma manera que los registros están agrupados en ficheros, y debemos distinguir entre una estructura lógica tipo y una estructura lógica concreta que, ajustándose al tipo preestablecido, tengamos realmente almacenada en la Base de Datos. Así, si se ha definido un anillo determinado como estructura tipo, existirán en la Base de Datos tantos anillos como registros miembro existan en el fichero correspondiente al registro tipo especificado para el registro principal del anillo. Cada uno de estos anillos de la familia podrá tener un determinado número de registros miembro (el anillo puede ser también vacío), correspondientes a los ficheros formados por los registros tipo declarados en la definición como miembros permitidos para el anillo.

Por lo tanto, a la hora de acceder a un determinado registro a través de un camino lógico contenido, por ejemplo de la estructura constituida por esa familia de anillos, deberá primero ser accedido el correspondiente anillo de la familia, representado por la cabecera del anillo que es un registro miembro de un fichero que, a su vez, deberá ser accedido en la forma que la definición de datos especifique. Una vez seleccionado ese anillo miembro de la familia, el registro que buscamos se seleccionará de entre un conjunto de registros miembro en función de la secuencia lógica de ordenación que esté prevista en la definición.

Para cada estructura lógica definida deberá especificarse si van a utilizarse enlaces dobles, si se va a disponer de un “reset pointer”, etc.; es decir, todas aquellas potencias lógicas que prevea el lenguaje de definición utilizado.

1.6.3. DEFINICION DEL ALMACENAMIENTO FISICO

Constituye la parte de definición que debe permanecer transparente para los usuarios de la Base de Datos. Esta parte de la definición viene a establecer una correspondencia biunívoca entre la dirección lógica de un registro y su dirección física del almacenamiento.

En las aplicaciones convencionales, esta definición se podía establecer en parte en las instrucciones declarativas de los programas y en parte en instrucciones externas a los programas ejecutadas por el sistema operativo de la instalación.

Deberá especificarse los tipos de dispositivos de almacenamiento que se van a utilizar y las reglas de asignación de registros a las correspondientes posiciones de almacenamiento. En las bases de datos los almacenamientos físicos se dividen en Areas que constituyen un conjunto de fracciones que se completan y se excluyen del almacenamiento total previsto. Estas Areas, a su vez, se subdividen en páginas, y la página suele ser la unidad de almacenamiento transferible a la memoria principal, y viene a ser el bloque de registros que se transfieren en las aplicaciones convencionales.

La página se compone de registros que están almacenados dentro de ella, y de un directorio de control de la página con información de control.

La definición establecerá qué tipos de registros van a ser almacenados en cada Area y qué limitaciones existen en la utilización de las páginas. Asimismo se indicará qué páginas van a destinarse a posiciones de "overflow".

Las Areas de datos deberán también estar asignadas a las correspondientes Areas físicas, es decir, pistas de discos o cilindros magnéticos, aunque generalmente esta asignación se realiza a nivel sistema operativo, con independencia del lenguaje de definición de datos, porque de esta manera las definiciones pueden ser invariantes respecto de eventuales cambios en la configuración física.

1.6.4. METODO PROPUESTO POR CODASYL

En el sistema de base de datos, diseñado por el DBTG de la Conferencia CODASYL, se utiliza un lenguaje de definición de datos llamado DDL (Data Definition Language), lenguaje que tiene una estructura análoga al lenguaje COBOL, porque no en vano es el COBOL el lenguaje "host" previsto para la utilización de la base de datos y porque realmente COBOL fue también creado por la Conferencia CODASYL.

Aunque más adelante veremos en detalle el sistema DBMS propuesto por CODASYL, vamos a examinar la parte correspondiente a la definición de datos para completar esta primera parte dedicada a los datos, que considero de interés previo para conservar una unidad en la presentación.

La estructura de una definición de datos expresada en lenguaje DDL, está separada en dos divisiones. La "Identification Division" y la "Data Division". La

“Identification Division” contiene exclusivamente el nombre simbólico que se da a la definición de la red de datos o Esquema, y el código numérico de uso interno que corresponde a dicho nombre. Los lenguajes de utilización utilizarán el nombre expresado en la Identificación División, para invocar la utilización de esa definición de base de datos.

La Data División consta de tres secciones: la Area Section, la Record Section y la Set Section.

En el Area Section se expresa para cada una de las áreas contempladas en la estructura las siguientes características:

- Nombre simbólico del Area y código numérico de uso interno.
- Tipo de Area, pues se prevén varios tipos según la naturaleza de los datos que vaya a contener.
- Distribución de páginas dentro del Area, indicando la longitud de las páginas y cuáles de ellas van a ser páginas de overflow. También se prevén expansiones futuras de las Areas que declarándolas aquí permitirán la ampliación sin necesidad de reorganizar la Base.
- Especificación de las copias de páginas y/o Area que se requieren realizar para protección de la información y/o después de los procesos.
- Factor de carga con la que regula la distribución de registros dentro del Area en una carga inicial de datos.
- Pueden especificarse un determinado número de estructuras de cadena simples o dobles, previéndose en cada página del Area los correspondientes registros cabeceras de las posibles cadenas.

La especificación de copia preventiva de páginas antes y/o después de los procesos puede especificarse también de una manera general, para ser aplicada a todas las Areas, siendo esta información aplicable por defecto cuando no se ha establecido especificación para un Area en particular, ya que ello es opcional.

La Record Section contiene una definición del registro tipo de cada uno de los ficheros contemplados en la Base de Datos. Para cada registro tipo se indica un nombre simbólico que será el nombre del fichero, y un código numérico asociado a él, que es de uso interno.

A continuación se establece cuál va a ser la forma de acceso y almacenamiento del registro, pudiendo especificarse:

- *Acceso directo.* Es un acceso directo lógico, existiendo un identificador o clave asociado al registro. Esta clave es un código significativo que expresa el número de Area, el número de página y el número de orden dentro de la página. Existe un procedimiento en el sistema operativo de la Base de Datos para encontrar la dirección física en función de dicha clave.
- *Almacenamiento en base a un procedimiento de cálculo.* Se trata de almacenar los registros en las estructuras de cadena previstas en el Area

Section. La posición de almacenamiento del registro se determina en base a un procedimiento de cálculo, que puede ser standard del sistema o bien proporcionado por una aplicación usuaria. Este procedimiento utiliza como argumentos determinados campos propios del registro y calcula la cadena en la que dicho registro habrá de ser ubicado. Hay que especificar también en qué Area se desea almacenar.

- *Localización VIA Set.* Se trata de indicar que el registro en cuestión va a ser almacenado como miembro de un determinado Set. La búsqueda de su localización se realizará en base a la regla lógica declarada para el Set en la Section.

Para cada registro se especifica qué Areas y qué páginas de esas Areas se prevén como posibles zonas de almacenamiento, no quedando autorizada su ubicación fuera de las zonas especificadas. Como dicha definición es opcional, se entenderá por defecto que no se impone limitaciones al lugar de almacenamiento.

Opcionalmente pueden especificarse reservas de campos en el registro, para albergar pointers. Estos pointers serán utilizados para encadenar al registro en la estructura lógica por procesos de programa de utilización. Existen también otros pointers que no es preciso declararlos y que el sistema asigna de forma automática, por corresponder a declaraciones de estructura lógica que se especifican. Los pointers necesarios son calculados por el sistema y posteriormente también serán utilizados de forma automática.

Para cada registro se establece a continuación la definición jerárquica de todos los campos que lo constituyen en forma similar a como se realiza en la Data Division en cualquier programa COBOL.

La Set Section incluye una definición de cada uno de los Sets tipo que se prevén para la Base de Datos, indicándose el nombre simbólico de cada Set tipo y su código numérico asociado.

Para cada Set se especifica:

- El tipo de Set, es decir, si se trata de un anillo simple o doble, o si se trata de un Set organizado por matriz de pointers.
- La ordenación de los registros miembro dentro del Set es lo que constituye el criterio de clasificación que permite insertar nuevos registros a un Set.
- Registro principal del Set. El registro tipo aquí especificado determina el fichero al cual pertenecen todos los registros principales de los Sets que corresponden al Set tipo que se está describiendo.
- Registros miembros. Se indican los ficheros que nutrirán de registros miembros a los Sets.

Para cada registro tipo miembro se especifica si va a ser insertado en el anillo en forma automática o manual, es decir, por las rutinas correspon-

dientes del sistema operativo de la Base de Datos o por un procedimiento de una unidad de tratamiento. También se puede especificar para cada registro tipo miembro la reserva de un reset-pointer, para enlazar cada registro de ese tipo directamente al registro principal. Para cada registro miembro se puede indicar también opcionalmente un criterio relativo de clasificación para los registros de ese tipo que estén enlazados en un mismo anillo.

- Se define también si se permiten registros duplicados o no.
- Para cada Set se establece también el criterio de selección para segregar cada Set del conjunto o familia que pertenecen al Set tipo definido.

Como vemos, la definición del Esquema en el lenguaje DDL presenta un aspecto análogo a un programa Cobol que sólo tuviera instrucciones declarativas, porque el lenguaje DDL no es un lenguaje de procedimiento.

El Esquema de base de datos así definido recibe el nombre de Esquema Fuente, el cual habrá de ser procesado por el compilador de DDL, obteniéndose un Esquema Objeto que pasa a formar parte de la librería del sistema de base de datos. El Esquema Objeto pasará a la memoria principal cuando sea invocado por un programa de utilización, constituyendo el soporte de definición de base de datos que utilizará el programa en cuestión.

Pueden determinarse tantas definiciones de Esquema como se precisen de una misma base de datos, con la condición de que no exista contradicción física entre diferentes definiciones, pues el hecho, la Base de Datos, es única. Existen determinadas reglas que establecen el margen de maniobra a la hora de materializar una nueva definición de la Base de Datos, reglas que se aplican en relación con la definición total de la Base, pues necesariamente debe existir un Esquema más completo que los demás, que incluya la definición total de la Base, siendo los demás Esquemas definiciones de subconjuntos de la Base de Datos.

1.6.5. METODO PROPUESTO POR GUIDE/SHARE

El planteamiento conjunto de las Asociaciones GUIDE y SHARE es el fruto de la experiencia de los usuarios de sistemas IBM, ya que dichas Asociaciones agrupan a utilizadores de dicha marca.

El trabajo realizado por GUIDE/SHARE es más bien la definición de un conjunto de especificaciones básicas que se considera deberán adoptarse para un futuro diseño de software de Base de Datos. Por ello no existe una definición precisa de los instrumentos, como en el caso del sistema DBMS de CODASYL, aunque también es cierto que, por otro lado, los objetivos que plantea son más ambiciosos.

GUIDE/SHARE propone como definición de la Base de Datos lo que denomina un DD/D (Data Dictionary/Directory), que constituye una colección ordenada de todas las *Entidades* referenciadas en la Base de Datos, con un conjunto de *Descriptores* que establecen las oportunas descripciones de las Entidades.

El DD/D es por un lado un diccionario, porque contiene las definiciones y características de los datos, y por otro lado es un directorio, expresándonos el algoritmo de búsqueda para acceder a un determinado dato. También muestran el conjunto de relaciones lógicas entre los datos, señalando por lo tanto cualquier camino lógico de acceso que un programa pueda necesitar.

El DD/D integra y define la estructura de datos y será siempre el “Interface” de comunicación entre los programas y los datos del almacenamiento, controlando todos los accesos y actualizaciones.

La descripción de datos, de acuerdo con las especificaciones DD/D, se realiza por seis diferentes conjuntos de atributos de dichos datos:

- *Atributos de Identificación.* Incluye un identificador y un nombre del dato, claves alfabéticas para identificación externa, descripciones, tanto para manejo humano como para operaciones de edición, tipo de dato de que se trata (nombre, código, cantidad, fecha...), función que va a cumplir el dato (identificación, control, índice, total...), estado de elaboración, fechas de efectividad, copias para respaldo y protección, y clasificación del dato definido, indicando si se trata de un elemento, de un grupo, de un registro, de un fichero, de una Base de Datos, de un dato de salida, de un Alias, etc.
- *Atributos de representación.* Son todas aquellas características relativas a la simbología del dato y contiene la definición del tamaño máximo (número de bits o bytes), tipo de caracteres (binario...), precisión, indicación de si se trata de un registro de longitud fija o variable, criterio de justificación del dato (derecha, izquierda...), estructura de códigos utilizada (EBCDIC, USASCII, etc.), la descripción de los campos en forma análoga al COBOL, las reglas de dicción, la escala, valores de relleno y unidades de medida.
- *Atributos de localización.* Son los datos definitorios de las características físicas y claves de acceso a la información. Se determina el tipo de dispositivo y número de volumen de almacenamiento, la organización del fichero (secuencial, encadenado...), el método de acceso, el algoritmo de direccionado, la dirección física, el tamaño del bloque o de la página, un indicador de secuencia física y el directorio de Alias donde se relacionan los nombres lógicos alternativos por los que diferentes programas pueden reconocer al mismo dato.
- *Atributos de relación.* Expresan las relaciones lógicas entre diferentes datos, pudiendo describirse enlaces de relación con otros datos, expresando la entidad con la cual se relaciona y el tipo de relación. Es por lo tanto un directorio de caminos lógicos entre datos.
- *Atributos de control.* Definen unos parámetros de utilidad para el Administrador de la Base de Datos: número previsto de registros, número máximo futuro, tasa de crecimiento, frecuencia de uso, decisión en caso de overflow, nivel de prioridad ante conflictos de espacio y de acceso,

y otros registros estadísticos que recoge el número de veces que se accede al dato, las identidades de los usuarios que lo recuperan y la cuenta de tasa de datos de crecimiento en la familia de datos.

- *Atributos de seguridad.* Establecen claves de seguridad que habrán de ser identificadas con las aportadas por los programas, niveles de seguridad, tipo de acceso permitido (recuperación o actualización), fuente del dato, responsable de su utilización, responsable de su definición, entidad máxima que controla las autoridades para acceso, algoritmo de variación periódica de la clave de seguridad, frecuencia de dicha variación, número de veces que deberá intentarse de nuevo un acceso erróneo y lista de usuarios autorizados.

El DD/D es más bien una especificación básica que un lenguaje de definición, porque realmente todavía no se ha perfilado como tal. Está orientado para actuar siempre de interface con la Base de Datos y permite construir una base de datos universal, tanto para utilizarla con lenguaje host como por funciones self-contained, siendo en cada caso complementada la definición por otros “interfaces” que pudieran ser necesarios.

1.7. UTILIZACION DE LOS DATOS

La Base de Datos, se orienta fundamentalmente como modelo vivo de la Organización, estando “animado” por la estructura lógica que viene a constituir en cierto modo un reflejo de la estructura de actividades de la Empresa o parte de la Empresa involucrada, y por otro lado, por una constante aportación de datos a la Base, que pasan a enriquecer a ésta en función de la relación datos/actividad.

Por ello, el usuario de una base de datos, no es un determinado departamento de la Empresa, sino que debiéramos fijar como concepto, que de haber un usuario preponderante, éste es la propia Empresa. Los datos no son patrimonio de la unidad que los produce, sino que han de integrarse en el activo del recurso información de la Empresa.

Entonces, los usuarios de la Base de Datos, serán todas aquellas entidades, que requieran recuperar, actualizar o almacenar, datos de la Base.

Por necesidades de control operativo, de seguridad de la información, y de simplicidad en la operación, los diferentes usuarios, tienen diferentes “autoridades” y diferentes prioridades, por lo que existe una escala de valores que ha de ser coordinada y controlada.

Por otro lado, unos usuarios, serán conocedores en detalle del contenido y estructuración de los ficheros, y podrán concretar sus necesidades, con el oportuno análisis funcional y utilización consiguiente de un lenguaje de procedimiento, mien-

tras que otros usuarios, pueden precisar manejar datos de la Base, incluso en tiempo real, sin que posean el conocimiento de su estructura.

Por ello, vamos a analizar los grupos de usuarios en función del tipo de utilización que van a ejercer en la Base de Datos.

1.7.1. USUARIOS INFORMATICOS

Es el conjunto de usuarios, cuya relación con la Base de Datos, es explicitada por analistas funcionales, concededores de la estructura lógica de la parte de Base de Datos que van a utilizar. Los denominamos usuarios informáticos porque, o bien conocen las características informáticas de la Base de Datos, o bien van a ser representados por unos analistas de sistemas, poseedores de dichos conocimientos.

El servicio a estos usuarios, se realizará con programas de proceso escritos en general, en un lenguaje host, que es complementado por las instrucciones específicas para el acceso a la Base de Datos. Los programas Fuente escritos en lenguaje host, son tratados por un preprocesador, antes de ser compilados por el compilador standard.

La programación es muy flexible porque los analistas y programadores se desentienden por completo de la problemática física de la Base de Datos y, ellos únicamente han de conocer en detalle, la definición lógica o Esquema del sub-conjunto de Base de Datos que utilicen, sin que exista peligro de "agresión" a zonas en la Base de Datos por ellos desconocidas, porque el sistema prevé el oportuno control.

Para aplicaciones muy específicas, donde se requiera optimizar procesos y se disponga de programadores más experimentados, puede también utilizarse un lenguaje ensamblador, con lo que la utilización de la Base es más complicada, pero la potencia de uso es superior.

1.7.2. USUARIOS NO INFORMATICOS

Las posibilidades de un sistema de base de datos, sugieren que puedan beneficiarse con su uso, determinados usuarios que no son concededores de las técnicas informáticas, ni de las características de la estructura de la Base de Datos, ni tampoco pueden ser auxiliados por analistas o programadores, porque necesitan interrogar a la Base de Datos en forma conversacional. Para estos usuarios, es preciso disponer funciones self-contained, que son procedimientos controlados por el sistema operativo, que pasan a ser ejecutados en la memoria principal, después de que el usuario se comunica con el sistema mediante unas transacciones de coloquio pre-establecidas, y el sistema operativo selecciona cuáles son los procedimientos o funciones a procesar.

El Usuario conoce los datos que contiene la Base y aplica determinadas transacciones para recuperar, actualizar, o almacenar información. Sin embargo, el Usua-

rio no conoce la estructura de relación de unos datos con otros, ni necesita conocer el conjunto completo de datos, sino solamente aquellas partes que constituyan su área de interés. Tampoco controla el Usuario, la secuencia de procedimiento de programa que va a ejecutarse para atender las transacciones solicitadas porque dicha atención, es cubierta por las funciones self-contained que son de uso standard y que constituyen un interface entre el Usuario y la Base de Datos.

1.7.3. LA FIGURA DEL ADMINISTRADOR

En este entorno en el que existe un conjunto de ficheros interrelacionados y una multiplicidad de utilizadores, es preciso que exista una entidad que pilote o dirija la gestión del recurso de información contenido en la Base de Datos.

Por ello, surge la figura del Administrador. Pensemos que en las aplicaciones convencionales, un fichero era creado por una determinada aplicación, que a su vez daba servicio a un determinado usuario. Este usuario, era siempre consultado por el Centro de Cálculo para aquellas actividades que pudieran afectar a la integridad de la información, como eran la expiración de ficheros, la acción sobre un fichero de una nueva aplicación, etc. En el caso de la Base de Datos, es preciso que exista una entidad única que represente a todo ese conjunto de usuarios y que a su vez, ejerza determinadas funciones adicionales que en la Base de Datos aparecen, y que exigen atención.

El Administrador de la Base de Datos, coordina la utilización de los datos, resolviendo los posibles conflictos entre diferentes usuarios (conflictos de accesos y de prioridades), asigna los criterios de prioridad, concede a los usuarios los oportunos niveles de autoridad, establece las oportunas seguridades para protección de la información, realiza las definiciones de la Base de Datos, da el visto bueno al diseño o bien lo dirige, determina las estrategias de almacenamiento inicial de los datos, y supervisa todas las funciones diarias de mantenimiento y control de actualización.

El Administrador de la Base de Datos, también ejerce un control estadístico de las condiciones de operación y de los niveles de uso, periodicidades y tiempos de acceso, obteniendo de una manera permanente consecuencias para introducir mejoras que optimicen las condiciones de operatividad, de lo cual es también responsable.

1.8. ADMINISTRACION DE LOS DATOS

La Base de Datos, sugiere en toda su potencia, el que el recurso información, sea contemplado como parte importante en los procesos productivos. Sin entrar en el plano filosófico, hay que tener en cuenta que el modelo de estructura de da-

tos que hayamos podido construir, constituye un indicador fehaciente del nivel organizativo de la Empresa. Hemos puesto de manifiesto, la relación entre los datos y la actividad, y una actividad podrá modelizarse más perfectamente en la medida en que las unidades organizativas que la ejercen, actúen en una forma sistemática u organizada.

Por esto, la Administración de los Datos, está emparentada con la Organización de la Empresa y las decisiones de administración de datos, deberán ser consecuentes con los criterios orgánicos y funcionales que gobiernan la Empresa. Por supuesto, del ámbito de la Empresa al que pertenezca el entorno de utilización de la Base de Datos.

La Administración de la Base de Datos, no debe por lo tanto, ser contemplada como una función técnica, aunque sí requiera el ejercicio de determinadas tareas técnicas, para las que puede utilizarse el personal ad-hoc. Se trata de una función de gestión y el perfil del responsable, ha de ser el de un gestor, aunque ciertamente precisa un importante bagaje técnico en sus conocimientos.

El poso de conocimientos e información que a lo largo de la dinámica del tiempo, proporciona la Base de Datos, es una secuencia histórica de la actividad de Empresa que representa, y forma parte de ese “know-how” que en muchas Empresas, constituye una parte importante del activo que poseen, porque los procedimientos, la organización, y la actividad empresarial en general, va perfilándose con el tiempo, poco a poco, en la medida en que las funciones y grupos laborales, se van estructurando, y sólo se llega a estructuras correctas válidas, para servir eficazmente objetivos, si el recurso información ha sido aplicado y distribuido en forma homogénea y dentro de un criterio general de equilibrio. Téngase en cuenta que la información y la energía, son entidades duales que siempre están presentes en todo proceso productivo que pertenece a la realidad pragmática.

La administración de la información, debe de contemplar la Base de Datos, como un activo a su cargo, cuya integridad debe asegurarse, buscando siempre la forma de enriquecer los niveles de utilidad de los datos y la potencia del modelo de Empresa que constituye. Esta concepción la ligaremos más adelante con el concepto de *Tiempo Presente*, que manejaremos como criterio general de diseño de una base de datos.

1.8.1. LA VIDA DEL DATO

El dato comienza a “vivir” como entidad de valor propio, desde el momento en que es captado, y comienza a ser utilizado desde el momento en que es aceptado por la Base de Datos e integrado en ella. Desde ese momento:

- Puede permanecer inalterado.
- Puede sufrir actualizaciones.
- Puede integrarse con otros datos, perdiendo características de segregación.

- Puede cambiar su estructura de relación.
- Pueden variarse las protecciones.

Después, puede llegar un momento en el que se decreta que ese dato ya no es válido y su permanencia en la Base, no justifica el coste correspondiente. En ese momento, el dato deja de ser accesible por los programas de utilización, aunque realmente puede recogerse en otros soportes informáticos de barato almacenamiento (cintas magnéticas), lo que todavía lo hace utilizable comercialmente. También puede destruirse cuando se considere que nunca más va a ser necesario.

La administración del dato, debe de contemplar a éste, con todo su valor de entidad individual, desde el momento en que es captado del ejercicio de la actividad, hasta el momento que el dato es destruido por alcanzar una utilidad nula. En todo este período de vida, se plantearán las etapas vitales en función de unos criterios que vamos a analizar:

- *Criterio de utilidad.* El examen detallado de las utilidades que un determinado dato va a proporcionar a las aplicaciones informáticas que a él accedan, constituye una regla lógica para establecer el “plan” de administración del dato. Las necesidades de manejo del dato en cuestión, nos determinarán por un lado la forma de almacenamiento y por otro lado, las especificaciones a aplicar para la operación de acceso a la información.

Fruto del examen de utilidad del dato, será la obtención de unas cifras previas de dimensión que establezcan la frecuencia de proceso del dato, el volumen de datos a almacenar, la tasa de recuperaciones por unidad de tiempo e igualmente la tasa de actualizaciones. Igualmente, el distinguir los tipos de procesos en los que el dato va a intervenir, es decir, si se trata de aplicaciones con diálogos interactivos, aplicaciones de teleproceso, aplicaciones batch ..., lo que nos proporcionará una indicación del nivel de rapidez que será exigible en los accesos, lo que viene a aconsejarnos la estructura lógica adecuada para adaptarse a las necesidades.

En cualquier caso, hay que distinguir la carga inicial de la Base de Datos, de las operaciones posteriores de mantenimiento y optimización, porque el primer planteamiento de almacenamiento y estructura lógica de los datos, es susceptible de futuras modificaciones a la vista de los informes periódicos que el Administrador ha de manejar, informes que le irán aconsejando mejorar las características de acceso, o bien “empeorarlas” si la utilidad es suficiente y con ello, se consiguieran otros objetivos como pudiera ser la economía de espacios de almacenamiento.

A la vista de los tanteos iniciales, se considerará además de la estructura lógica de relación, la composición de los registros y la diferenciación de datos en diferentes registros, determinándose los niveles de segregación de la información que sean necesarios.

- *Criterio de relevancia.* Mediante este criterio, debiéramos hacer abstracción de la utilidad técnica o funcional que cada dato proporciona a los programas de proceso y aplicaciones informáticas en operación o proyectadas. Debemos contemplar el dato con su propio valor objetivo, como formando parte del activo de información de la empresa, ya que la Base de Datos debe de constituir una entidad con legitimidad propia sin justificar necesariamente la existencia de cada dato por su simple aplicación directa.

El dato puede representar utilidad en el ámbito de la empresa, aunque no existan programadas aplicaciones informáticas periódicas que lo manejen, porque el modelo de actividad que la Base de Datos representa, debe estar siempre dispuesto en forma operativa y con la suficiente perspectiva histórica para aplicaciones esporádicas que siempre surgen en los procesos de decisión de la vida diaria de la empresa. En muchas ocasiones, hemos visto cómo no se han podido realizar evaluaciones importantes o documentarse debidamente expedientes, o soportarse eficazmente las decisiones, por no disponerse de un soporte de datos o aun disponiéndose de los datos, no ser éstos susceptibles de utilizarse por no estar estructurados ni elaborados.

La relevancia del dato, por un lado puede ser motivo suficiente para almacenar dicho dato, aun cuando su utilidad no sea inmediata, y por otro lado, en base a ella es de determinar los niveles de síntesis y análisis que tendrá el dato a lo largo de su permanencia en la Base de Datos, así como los períodos de retención que habremos de fijar para cada nivel de síntesis o análisis de los datos y para su permanencia física en la Base de Datos, y quizás también posteriormente, en soportes procesales ajenos ya al entorno de la Base. Incluso también en soportes no procesables a efectos de archivo histórico consultable.

Al referirnos a los niveles de síntesis o análisis, queremos establecer los diferentes “estados” bajo los que el dato puede manifestarse en la Base. Al avanzar el tiempo, y a medida que el dato va perteneciendo a una dimensión más lejana en la historia, pueden producirse condensaciones de forma que lo que en un pretérito inmediato tengamos representadas características referentes a períodos diarios, en un pasado lejano, estas mismas características, pudieran corresponder a períodos mensuales o anuales. El dato va sufriendo operaciones de síntesis, perdiendo capacidad de profundizar en el detalle, porque a medida que pasa el tiempo, también varía la relevancia del dato. El criterio de rentabilidad está aquí presente porque en cada operación de síntesis, se pueden cancelar los registros más analíticos, pasándolos a otros soportes (cintas magnéticas, por ejemplo). Las fechas hábiles para sintetizar los datos y cancelar consecuentemente los registros previos, deberá de fijarlas el Administrador de la Base, una vez que haya escuchado las argumentaciones de los posibles usuarios de la información en el ámbito de la Compañía. Igualmente, la destrucción de cintas magnéticas de datos antiguos, deberá

ser fijada colegiadamente con los usuarios afectados, decidiendo en su caso si procede conservar la información aun cuando no fuera procesable por eliminarse las cintas magnéticas.

- *Criterio de rentabilidad.* Este criterio, como hemos visto, será el que fundamentalmente nos motive a sintetizar datos y a cancelar registros de la Base, porque todo lo que sea acortar las fechas de expiración de los registros, contribuye a una economía de almacenamiento, aunque deberán adoptarse los términos medios prudentes que permitan que la economía obtenida en el almacenamiento de la información, no sea compensada por el “daño emergente” cuando sea preciso utilizar información que ha sido ya destruida.

El tiempo de acceso, es un valor determinante del coste de operación y como norma de rentabilidad, debiéramos de disponer unos tiempos de acceso suficientes, lo más cortos posible, pero sin llegar a traspasar umbrales económicos que nos hagan incurrir en unos costes no compensables por la rapidez de acceso obtenida. Hay que distinguir a la hora de estudiar los tiempos de acceso, que debe de haber un equilibrio entre la capacidad hardware de acceso, y las demoras imputables al software, porque pudiéramos incurrir en costes innecesarios para reforzar la configuración física de equipos, cuando las deficiencias de acceso se deben a una estructura lógica de datos inadecuada o a una programación de la aplicación, poco eficiente. Digamos que la potencia del hardware en especificaciones de acceso de canales, es una referencia de partida, o bien un punto a decidir si es que la dotación del hardware entra también dentro del proyecto. A partir de esa referencia, deberá considerarse a la estructura lógica, a la disposición física de registros y a las decisiones cristalizadas en software, como fuentes de las demoras del acceso. Naturalmente, también las decisiones a nivel sistema operativo basadas en las prioridades asignadas a los programas, son causas que inciden en el tiempo de acceso, por lo que el Administrador de datos, deberá estudiar el paquete de prioridades, de acuerdo con la relevancia de la información y de las aplicaciones informáticas en explotación.

- *Panorama futuro.* La perspectiva de futuro, es necesaria para el diseño de la administración de la Base de Datos, porque siendo consecuentes con el concepto de base de datos como entidad de valor propio, la utilidad debe dar servicio, tanto a las aplicaciones en explotación o planificadas a corto plazo, como a aquellas otras posibles aplicaciones que se prevén a medio o largo plazo o que incluso, simplemente se intuyen como posibles. La dimensión de futuro que contenga la Base de Datos, será el factor determinante de la flexibilidad y capacidad de adaptación del sistema.

La vida del dato, hay también que contemplarla “más allá” del entorno de la Base de Datos y a tales efectos, existen hoy en día medios para que los datos históricos o relativamente recientes, que ya no van a ser procesados en el ordenador, pue-

dan sin embargo, ser procesados por personas humanas, beneficiándose a la hora de acceder a la consulta de los datos, de los criterios de organización que el paso por un entorno de base de datos, les legó. El microfilm y la microficha, se presenta como solución muy adecuada a este respecto, e incluso mucha de la información contenida en la Base de Datos, pudiera ya presentarse también en soportes de microficha cuando eventuales necesidades de consulta, son por un lado patentes pero por otro no justificativas de aplicaciones de interrogación "online".

Consideramos que la administración de los datos a lo largo de su vida útil, no debe ceñirse al entorno del Centro de Cálculo y, deben ser contempladas utilidades adicionales como la esbozada basada en la microfilmación que proporcionan sistemas ágiles de explotación humana del activo de información, que viene a cubrir lagunas muchas veces existentes, a cifras de coste marginales. Un sistema de manejo de información microfilmada, permite también incorporar otro tipo de información que no necesariamente haya sido procesada y puede por lo tanto tener unas bases de existencia independientes de la mecanización, y en todo caso, complementarias.

1.8.2. FUNCIONES DE ADMINISTRACION

La figura del Administrador, nace sugerida por el concepto de explotación en base de datos, aunque algunas aplicaciones informáticas complejas, ya requerían de esta figura, porque la nueva función se aplica por necesidades de coordinación al existir muchos programas y usuarios utilizadores de un conjunto de datos complejo y estructurado.

La función de administración comprende un conjunto de tareas que no necesariamente ha de ser cubierto por una unidad orgánica ad-hoc, sino que más bien supone un conjunto de responsabilidades que hay que asegurarse que son debidamente ejercidas en el seno de la organización. No obstante, cuando se plantean proyectos de bases de datos complejas, es muy conveniente y casi podíamos decir imprescindible, que exista la figura del responsable de la administración de los datos como gestor de un importante Activo de información. El Administrador, puede estar auxiliado por una unidad orgánica a su cargo que reúna a los especialistas que se precisen.

El conjunto de funciones de administración, pudiéramos clasificarlo en cuatro grandes grupos: Las funciones de definición y organización de datos, las funciones de seguridad y protección de la información, la función de dimensionado y la función de documentación.

Dentro de las funciones de definición y organización de datos, podemos enumerar las siguientes tareas:

- Previsión de futuras necesidades.
- Establecimiento de autoridades y prioridades de usuarios y programas.
- Arbitraje de conflictos entre usuarios.

- Definición de la estructura lógica de los datos.
- Definición de las estructuras de los registros.
- Definición del almacenamiento físico.
- Determinación de las formas de acceso a la información en función de las especificaciones de utilización.
- Tarea continuada de optimización de la organización de datos, determinando y ejecutando las modificaciones de estructura o definición que sean precisas para mejorar la utilidad.
- Generación de Esquemas que se precisen para el proceso de las aplicaciones.

La administración de la Base de Datos, deberá tener la visión de conjunto de la Organización, que le permita tratar los problemas con objetividad, ofreciendo confianza a los usuarios, y representando en todo momento a la alta Dirección, para anteponer siempre los objetivos de empresa en sus decisiones.

Como tareas más importantes del concepto de seguridad y protección de la información, tenemos:

- Adecuada protección del acceso y proceso de registros, permitiendo la recuperación de información solamente a las aplicaciones autorizadas.
- Las actualizaciones estarán debidamente diseñadas para evitar la corrupción de la Base de Datos. Estará estudiada y definida, no sólo qué unidades de tratamiento estarán autorizadas para actualizar los datos, sino también la resolución de actividades de actualización concurrentes. Ello implica un control del uso de las facilidades de acceso que ofrece el software utilizado.
- Control de la cancelación de datos de la Base.
- Asegurar la integridad de la Base de Datos, para lo que se auxilia de un sistema de auditoría que generalmente ofrecen los sistemas de base de datos que consiste básicamente en una cinta magnética gobernada por el sistema operativo, que recoge copias de las transacciones de acceso al sistema, y copias de los registros y áreas antes de su actualización. Esta cinta sirve también para recoger “check-points” que indican el estado de actualización al que en cada momento hay que retroceder cuando se producen determinadas averías o excepciones que provocan el aborto de programas de proceso.
- Diseño de “respaldos” del sistema. Puede haber diferentes niveles de respaldo o “back up”, según una clasificación de problemas que puedan producirse. En el diseño de respaldos del sistema, habrá de tenerse en cuenta los tiempos de recuperación necesarios.
- Establecer tablas de decisión para uso de los operadores, que determinen las acciones a realizar ante errores de proceso, fallo de programa,

fallo de energía, inconsistencias de la configuración, corrupción de datos ... etc.

- Proceso periódico de programas de diagnóstico. Por un lado, deben permitirnos detectar posibles corrupciones en la Base de Datos y por otro lado, pueden realizarse simulaciones que nos aconsejen modificaciones.
- Identificar los casos que se produzcan de violación voluntaria o involuntaria de la normativa en vigor, disponiendo las soluciones oportunas para evitar posteriores repeticiones.

El Administrador de la Base de Datos, debe también velar para que la Biblioteca de programas de utilidad, sea adecuada y esté en relación con la dimensión del sistema.

1.8.3. DIMENSIONADO DE LA BASE DE DATOS

El dimensionado de la Base, es un problema inicial de diseño y una tarea posterior de mantenimiento. Aunque lo estudiaremos en detalle como etapa obligada de proyecto, vamos a enumerar los factores a dimensionar que deberán ser estudiados para cada uno de los datos contenidos en la Base.

- Tamaño del registro.
- Tamaño del registro incluyendo pointers.
- Volumen de registros de carga inicial.
- Tasa de crecimiento de registros por unidad de tiempo.
- Tasa de recuperación de registros.
- Tasa de actualización de registros.
- Tasa de cancelación.
- Requerimientos de tiempos de respuesta.
- Distribución de registros en Areas.
- Dimensionado de Areas y páginas.
- Dimensionado de Areas de overflow.

El Administrador, deberá partir de unas previsiones iniciales, y generalmente en un primer diseño, nunca logrará una carga inicial de la Base de Datos óptima, debido a la incertidumbre de posterior cumplimiento de las previsiones. Por eso, la función de dimensionado requiere una continua vigilancia de los factores antes relacionados, para lo que el Administrador deberá contar con los programas estadísticos necesarios que le indiquen el desarrollo a lo largo del tiempo de “load factors”, de ocupaciones de Areas y Páginas, volúmenes de registros, tiempos medios de acceso, ... etc.

Ni que decir tiene que la administración, deberá siempre anticiparse a todo problema de saturación, adoptando soluciones con la antelación precisa.

1.8.4. DOCUMENTACION DE LA BASE DE DATOS

La documentación debe estar dirigida a cuatro diferentes grupos de la Organización: la propia Administración, los Operadores, los Analistas y los Usuarios en general.

Para utilización de la Administración, deberá mantenerse documentación histórica y documentación actualizada de las Areas y su contenido, de los diagramas lógicos de estructura de datos, de las relaciones lógicas entre registros, y de los volúmenes de datos almacenados, así como las desviaciones habidas respecto a standards anteriormente previstos.

Para utilización de los Operadores, deberán estar perfectamente establecidos y documentados los procedimientos de mantenimiento, la guía de decisiones ante incidencias, y los "back-ups" del sistema con las prioridades de recuperación.

Para utilización de los Analistas de aplicaciones, deberá proporcionarse documentación de cada uno de los Esquemas disponibles de definición de la Base, con relación de las aplicaciones que los utilizan y autoridades de dichas aplicaciones.

Para los Usuarios en general, deberá proporcionarse un directorio del Activo de información que ofrezca un conocimiento preciso de la disponibilidad de datos. También deberá proporcionarse documentación de los informes que los usuarios pueden obtener a petición, independientemente de que las aplicaciones de uso periódico estén provistas también de la oportuna documentación.

De importancia general para el Centro de Cálculo y la Administración, son las estadísticas reales de frecuencia de utilización de datos por los distintos programas de aplicaciones, con una evaluación de las características de utilización de los diferentes programas respecto de cada fichero, tiempo de acceso, volúmenes de recuperación ... etc., lo que completa la información estadísticas a efectos de la permanente labor de dimensionado.

1.8.5. IMPUTACION DE COSTES

Aunque pertenece esta tarea a la administración económica de la Organización, sin embargo, vamos a referirnos aquí a ella brevemente.

En una Organización, se suelen manejar tres criterios a la hora de conceptuar y clasificar los costes de explotación:

- Gastos generales.
- Imputación a usuarios.
- En parte imputación a los usuarios y en parte conceptuación gasto general.

En el caso de la Base de Datos, quien realmente tiene mejor criterio para reajustar los coeficientes de imputación a usuarios, es la unidad de Administración de la Base de Datos. Sus determinaciones, han de ser extremadamente objetivas en el caso de que el criterio sea el de imputar siempre los costes, pues el propio desarrollo futuro del sistema, dependerá de lo objetiva que sea la imputación ya que lógicamente, los Usuarios tienen mayor capacidad de decisión en el sistema en función de su responsabilidad presupuestaria, y el Administrador no dispone en este caso, de más armas que la objetividad y la confianza que ella le confiere, tanto a la Alta Dirección como a los utilizadores de la base.

Por eso, en una base de datos compleja con multiplicidad de usuarios, conviene dosificar convenientemente el criterio de gasto general con el criterio de imputación, adoptando una solución mixta que por un lado conciencie a los usuarios de los costes y rendimiento, y por otro lado, garantice la autoridad suficiente para el pilotaje del sistema.

2. LA EMPRESA COMO SISTEMA

El analista informático, tiene que plantear necesariamente sus Esquemas en forma pragmática porque su producto es la información, y la información es materia prima para la gestión. Por otro lado, tiene que ser también sistemático porque los ordenadores reportan mayores utilidades en la medida en que incorporan mayor número de criterios y éstos de mayor estructura, dando por sentado por supuesto, que la técnica es correcta. La falta de método y sistema, da como consecuencia un producto informático estéril que establece puntos de rigidez en la Organización y no produce los efectos deseados y previstos.

Contra lo que pudiera parecer, ser sistemático y a la vez pragmático, no son cosas contrapuestas, sino que son perfectamente compatibles. No debemos confundir el actuar en forma sistemática con el pecar de teórico, ni tampoco debemos renunciar a plantear los problemas con el debido rigor metódico acuciados por esa gran disculpa que es la resolución de problemas “calientes” y la necesidad de cubrir objetivos a corto plazo.

Tengamos presente que para ganar batallas, hay que plantear buenos esquemas estratégicos y hay que disponer de un aparato logístico concebido en frío, con la consistencia que da un método concienzudamente estudiado, y además teniendo presentes las necesidades reales que se producirán en el campo de batalla. La Cibernética y la Teoría de Sistemas, han demostrado sus mejores resultados en aplicaciones militares y de hecho, nada pertenece a una realidad más profunda y desgarrada, que una guerra. Tomemos la debida enseñanza de las Instituciones Militares que no en vano son las más antiguas Organizaciones y las que mayor solera tienen.

Manejar datos, es manejar muestras de la realidad pragmática que tiene lugar en la Organización. Dichos datos o muestras, se complementan unos a otros y pueden llegar a formar un conjunto homogéneo representativo de la realidad, en cuyo caso decimos que ese conjunto de datos forman un Sistema.

Captar datos es por lo tanto un sinónimo de “tomar muestras”, y todo el proceso de captación ha de ser realizado en base a un método preconcebido en detalle, precisamente porque el producto ha de ser un Sistema de Datos, y un Sistema coherente y representativo de la praxis no surge por generación espontánea.

Procesar datos de una Base, equivale, por lo tanto, a manejar Esquemas representativos de una realidad y los resultados que obtengamos, es decir, los productos informáticos, serán “síntesis” o “consecuencias” de esa realidad.

Construir una Base de Datos, equivale entonces a construir un modelo de la realidad pragmática de nuestra Organización o de nuestra Empresa, modelo que

estará formado por un Sistema de Datos. Vamos a ver cómo podemos contemplar esa realidad y cómo vamos a traducirla en una Base de Datos.

Para ello, vamos a ofrecer un modelo de Empresa muy potente en ideas, muy flexible en su aplicación y de uso muy universal, porque con él nos podremos aproximar a cualquier planteamiento con independencia de las particularidades propias de la Organización que va a ser analizada

Para todo ello, vamos a ayudarnos de esa metadisciplina universal llamada Teoría de Sistemas, que no es otra cosa que una moderna versión del pensamiento filosófico de siempre, que siempre ha sido el soporte de las Ciencias y de su aplicación.

2.1. LA ACTIVIDAD

La Empresa u Organización, vamos a considerarla como un sistema abierto, es decir, como un sistema que intercambia energía e información con el entorno y que por lo tanto, están produciéndose en él transformaciones a lo largo de la variable tiempo. La realidad histórica se va construyendo paso a paso al transcurrir la dimensión tiempo, dimensión que establece la secuencia de actualidad histórica en la misma forma en que estarían ordenados los fotogramas de una película continua en la que recogeríamos la realidad de la Organización.

La actividad, es un fenómeno consustancial con el proceso de transformación del Sistema y supone una variación de los estados potenciales tanto energéticos como informáticos, de las partes constituyentes de la Organización. La actividad se produce con el intercambio de energía e información, entre el sistema y el entorno, y asimismo se produce con análogos intercambios a nivel interno entre distintas partes constituyentes del Sistema. Las leyes físicas e informáticas no están ausentes, y los fenómenos entrópicos se presentan con las consiguientes degradaciones o tendencia a igualarse potenciales.

La actividad, podemos asimilarla como la dinámica empresarial o como el “aparato propulsor” de la Empresa u Organización. El resultado físico del ejercicio de la actividad, es la consecución de productos, que denominamos productos finales en el caso de que estemos contemplando la actividad empresarial de forma global. Los denominamos productos intermedios, si corresponden a actividades parciales de la Empresa, y en este caso, dichos productos intermedios, serán generalmente materias primas utilizadas en otros procesos de actividad del entorno de la Organización. Los productos finales, podrán ser a su vez materias primas para otros Sistemas ajenos a nuestro Sistema de Empresa.

Para ejercer la actividad, se utilizan Recursos que básicamente se componen de personas, dinero, máquinas e instrumentos, y además se consume energía, energía que consiste en materias primas y bienes de consumo, los cuales son a su vez productos finales proporcionados por otros Sistemas o productos intermedios proporcionados por otras partes del Sistema.

Además, el ejercicio de la actividad, requiere la utilización de información y a su vez, hay que tener presente que la actividad en sí, también produce información,

y entonces por un lado hay que asegurar productos informáticos adecuados y también hay que captar los datos adecuados generados por la actividad. La Organización informática, habrá de formar parte del sistema global de Empresa, aunque también podemos contemplarla como un sistema separado y complementario, obligándonos entonces a considerar una actividad empresarial parcial.

La actividad, es en general una dinámica puesta en marcha para dar satisfacción a unas motivaciones o lo que es lo mismo, para cubrir unos determinados objetivos. Por ello, en la Empresa hay unos mecanismos de decisión que fijan objetivos, determinan motivaciones de Empresa y controlan el ejercicio de la actividad en el seno de la Organización. Estos mecanismos de decisión, constituyen lo que pudiéramos denominar conciencia externa del Sistema.

Sin embargo, siempre hay componentes de la actividad que pueden perseguir y que persiguen de hecho, motivaciones externas al Sistema, es decir, objetivos de rango superior a los fijados para el sistema, así como también hay entidades participantes en la actividad, que se guían por motivaciones de orden interno y de rango inferior al sistema que constituyen por lo tanto sub-objetivos. La persona mismo, en su calidad de parte integrante de un sistema organizativo, persigue por una parte objetivos empresariales, por otra parte objetivos personales e incluso objetivos de rango superior, ya que la persona depende por otro lado de actividades del sistema social al que pertenece y de manifestaciones del sistema biológico que encarna.

Del conjunto de actividades, hay unas actividades conscientes y otras inconscientes al "piloto" del Sistema, entendiéndolo por piloto al órgano superior de decisión. En la obligada comparación con el ser humano, pues humanos somos y como tales utilizamos en nuestras realizaciones las soluciones que encarnamos, la *fagocitosis* por ejemplo, es una actividad inconsciente, aunque sin embargo, puede responder a una motivación consciente como puede ser la voluntad de sanar de la persona. En ese caso, más que actividad inconsciente, se trata de una actividad transparente, ya que aun respondiendo a directrices de nuestra conciencia, la actividad se verifica a nivel celular, nivel que es excesivamente analítico para nuestra conciencia exterior. También esa actividad responde además a necesidades biológicas de distintas procedencias de la conciencia, pues cada célula como entidad individual, es a su vez un sistema de objetivos.

Los objetivos necesitan un sistema de información para llevar su conocimiento a niveles transparentes al sistema de decisión, es decir, la conciencia exterior necesita de alguna forma transmitir a cada célula unos objetivos finales, y análogamente, el centro de decisión del sistema, necesita que en él converja una red ascendente de información que le ofrezca síntesis bien construidas indicativas del estado de situación, información que obedece por lo tanto al concepto de "feedback".

La célula actúa sirviendo a objetivos generales, en la medida en que los conoce, en la medida en que dichos objetivos están en línea con los objetivos propios individuales y en la medida en que a su vez dicha célula tenga un sistema interno de decisiones, u órgano piloto, que sepa dirigir la actividad celular y cumplimentar el paquete de objetivos.

El aparato o sistema de información, es factor determinante de la coordinación general y soporte básico por tanto, de la Conciencia global del Sistema Organizativo y su Piloto.

Un Piloto u órgano de decisión con un sistema de información malo necesariamente tomará decisiones erróneas, sin embargo, la viceversa no se cumple puesto que si el Sistema de Información es bueno, el órgano de decisión pilotará bien en la medida en que buen piloto sea y la coyuntura lo permita.

Un buen pilotaje, dirigirá al Sistema en la coyuntura. Su capacidad de intervención en el entorno, estará condicionada por la buena o mala dirección y por las oportunidades que eventualmente se presenten. Sin embargo, si “no hay dirección”, el Sistema “navega a la deriva” atrapado por la coyuntura.

2.1.1. EL SISTEMA PRODUCTIVO

Antes de proseguir, vamos a sedimentar un poco las ideas, y para ello nada mejor que acudir al ejemplo y a la imagen. Veamos la figura 23.

En la figura, se representa en forma muy primaria y esquemática, una unidad organizativa o entidad en transformación, que es un barco. En dicho barco se desarrolla un conjunto de actividades, que dan lugar a transformaciones internas y también naturalmente, a intercambios con el entorno socioeconómico-geográfico-temporal al que el barco pertenece. Los intercambios consisten en carga a las bodegas o carga de las bodegas al exterior, y en la aportación de energía e información del exterior al barco, para alimentar el desarrollo de la actividad de transporte. Así, el barco embarca víveres, combustible, agua, ... etc., y recibe directrices, información meteorológica, contactos radio, ... etc. El barco también entrega energía al mar en forma de rozamiento al deslizamiento, formación de olas, y rozamiento con el aire, rindiendo tributo a las leyes físicas.

El objetivo final del buque, será el trasladar de lugar una mercancía en un tiempo determinado, para lo cual habrá de realizar una singladura siguiendo un determinado rumbo y manteniendo a lo largo de su trayectoria, unas determinadas velocidades, siendo rumbo y velocidad, parámetros objeto de continua vigilancia y decisión por parte del órgano de decisión del buque que es el Capitán del navío.

Sin embargo, si contemplamos la Empresa Naviera fletadora del buque como unidad organizativa, los objetivos del buque en particular, serán objetivos parciales de una parte componente de la Organización, ello si la Naviera es propietaria de la nave, porque sino lo es, el producto final del buque es simplemente un recurso sub-contratado, ocupándose la Organización simplemente de recoger la prestación contractual y sin constituir por lo tanto un sub-objetivo a pilotar. En cualquier caso, el producto final del buque, es solamente parte constituyente del producto final de la Organización.

Sin embargo, si con esta misma óptica penetramos en el interior del buque y consideramos como unidades organizativas sus partes componentes, estaremos contemplando objetivos y actividades de un nivel más analítico. Así, la turbina o tur-

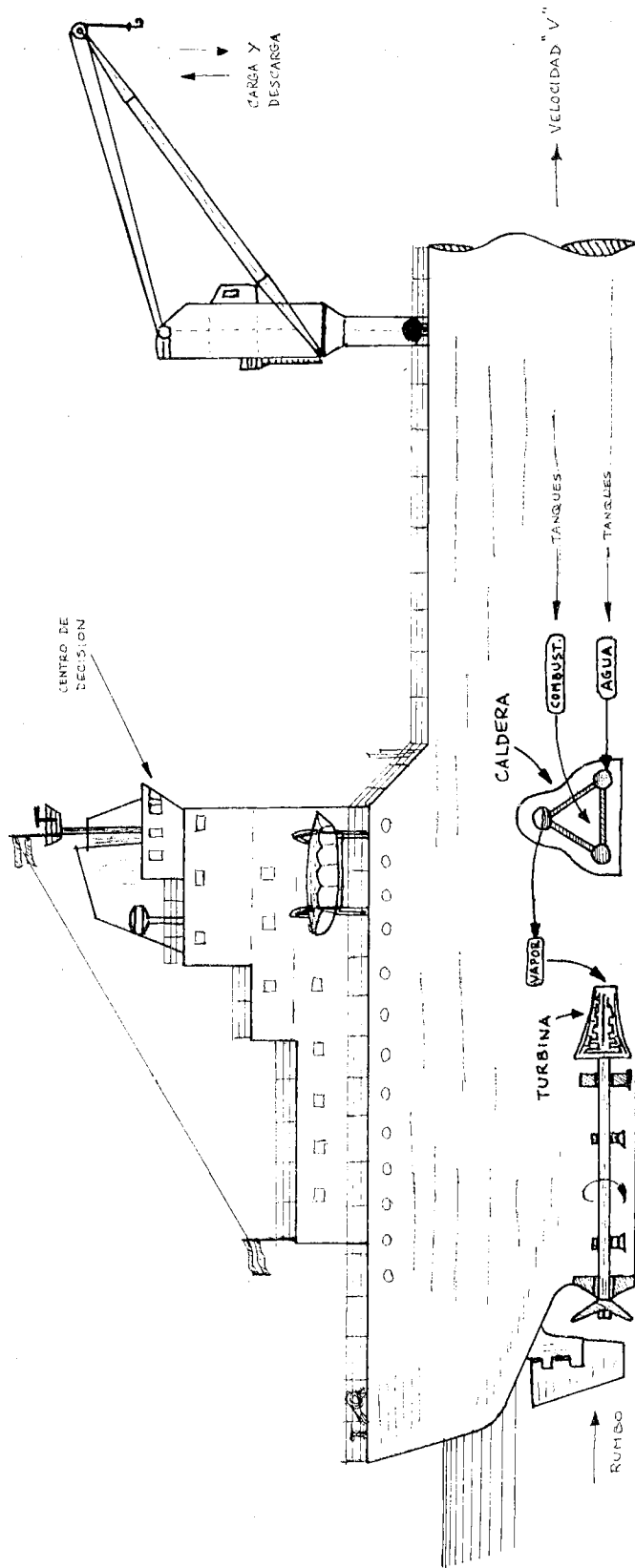


Fig. 23

binas del buque, tienen por objetivo comunicar unas determinadas revoluciones al eje propulsor que va a mover la hélice, revoluciones que constituyen un objetivo fijado en cada momento por el Capitán del buque. El órgano de decisión del conjunto de turbinas, empleará una mayor o menor cantidad de materia prima que es el vapor en este caso, para ir cubriendo su objetivo de obtener un determinado régimen de revoluciones. El vapor empleado por la turbina como materia prima, es a su vez el producto que manufactura otra unidad organizativa que es el conjunto de calderas. El Piloto de las calderas, tendrá como objetivo el mantener una determinada presión y una determinada temperatura en los colectores de vapor, objetivo que pertenece al diseño ingenieril de la instalación, pero cuya consecuente actividad para alcanzarlo, está en cada momento condicionada por la extracción de vapor del colector que realizan las turbinas. El proceso informativo se verifica para conocimiento del Piloto de las calderas, por medio de los correspondientes manómetros y termómetros, pudiendo incluso establecerse un mecanismo de control *a priori* conociendo las órdenes del Capitán de la nave e intuyendo por anticipado las necesidades de vapor que tendrán las turbinas en función de los objetivos marcados.

La actividad en la caldera, se desarrolla admitiendo mayor o menor cantidad de agua para cebar el proceso de evaporación, e introduciendo mayor o menor cantidad de combustible en el hogar, para tener un ritmo de fuego adecuado. El combustible y el agua, son materiales de consumo para la caldera, y son a su vez productos intermedios que nos entregan otros Sistemas organizativos componentes del buque, que son el sistema de agua de alimentación y el sistema de combustible. Así, a su vez podríamos continuar examinando los componentes del buque, categorizándolos con una óptica de sistemas en diferentes unidades organizativas a diferentes niveles de síntesis o análisis.

Para realizar esta descripción, en realidad lo que hemos hecho, es utilizar el concepto de sistema productivo, concepto que está íntimamente relacionado con el de actividad. Por eso, vamos a intentar aislar un sistema productivo dentro de la unidad organizativa del buque representado en la figura 23, y vamos a estudiar la actividad de ese sistema productivo, y ello antes de definir lo que en realidad es un Sistema Productivo.

Esta metodología de aproximarse y estudiar una cosa antes de definirla, es en realidad una permanente necesidad del hombre, y podemos compararla a la forma en que los biólogos se aproximan al concepto de vida. Los biólogos, no pueden definir qué es la vida, igual que no puede definirse lo que es un sistema, sin embargo, repasamos las circunstancias y atributos que observamos en lo que en conciencia reconocemos como organismos vivientes y posteriormente hacemos una abstracción adjudicando la denominación de Organismo con vida a un Sistema que entra dentro de los esquemas previamente observados.

En la figura 24, se representa de una forma simple y esquemática, la caldera del buque de la figura 23. Para aislar dicha caldera del buque, de forma que conserve una entidad propia de unidad productiva con título de legitimidad, es preciso efectuar la separación "física" en algún punto de la tubería de vapor de salida del colector de vapor, en algún punto de la tubería de alimentación de combustible y asimismo, en algún punto de la conducción de agua de alimentación. Habría también que considerarla como unidad seccionada de la estructura del buque, y si

el Sistema que vamos a estudiar se ocupara de problemas de resistencia mecánica de la estructura del buque o de vibraciones, habría de estudiarse también cuidadosamente, las zonas que delimitan la unión del buque a la caldera.

Con esto, hemos aislado la caldera como entidad, pero para que dicha entidad conserve su característica de unidad productiva, es preciso también que junto con la caldera, consideremos los mecanismos que regulan su funcionamiento y establecen las decisiones a lo largo del tiempo. De esta forma, se representa en la figura, un centro de decisión que no es otra cosa que el "piloto" de la caldera. Este piloto, puede ser una persona, puede ser un conjunto de personas, o simplemente puede ser un mecanismo automático como en la realidad sucede en los buques modernos. Su razón de ser se deriva del hecho de que el Sistema "caldera", tiene un objetivo productivo encomendado, que en este caso es el poner en la tubería de

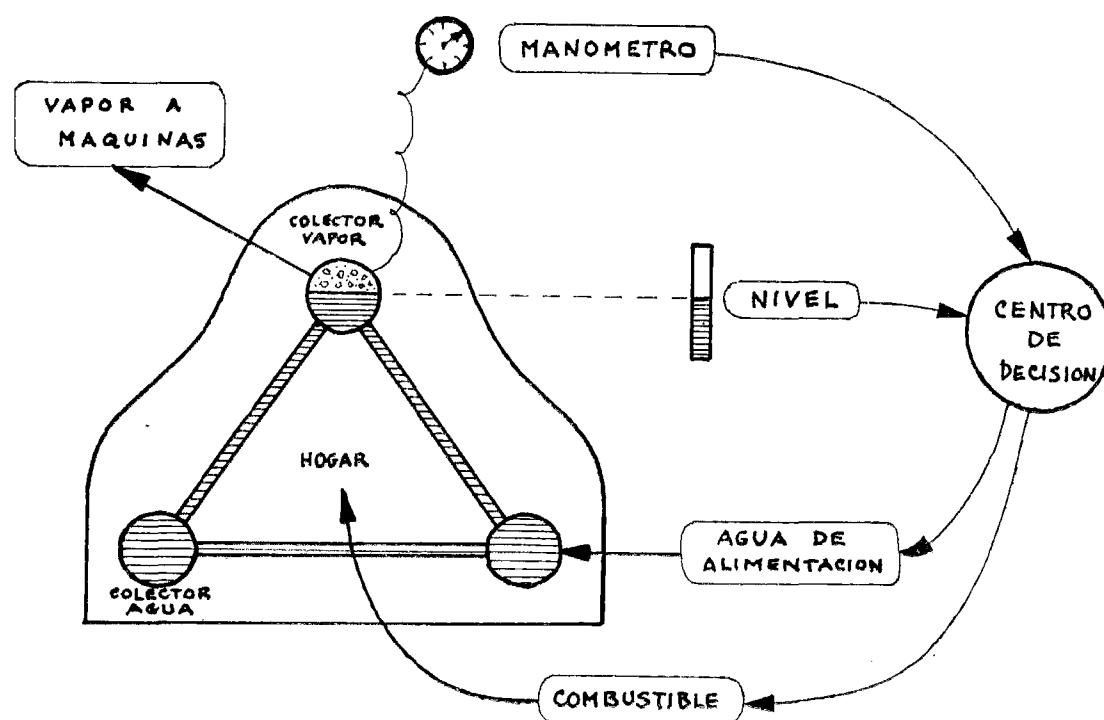


Fig. 24

salida de vapor, todo el vapor que precisen en cada momento las turbinas, en unas condiciones de presión y temperatura previamente establecidas. Por eso, el "piloto", si es una persona, tendrá que reaccionar ante las variaciones que observa en el manómetro que mide la presión en el colector de vapor o si es un mecanismo, habrá de reaccionar ante las variaciones de la señal emitida por el correspondiente presostato. De esta forma, el piloto tiene una indicación permanente de sus objetivos, ya que al aumentar las necesidades de vapor, se detecta una caída de presión y viceversa.

Igualmente, pueden existir otro tipo de objetivos que no están directamente asociados a la consecución de un producto, sino que son necesidades internas del Sistema, como en este caso es el objetivo de mantener el nivel de agua en el colector de vapor, dentro de una cota de niveles que permitan el correcto funcionamiento de la instalación.

Esta información de presión y nivel del colector de vapor, no es ni más ni menos que lo que denominamos como información de "feed-back". Mediante esta información, el Centro de Decisión establece las correspondientes acciones correctivas para adaptarse a los objetivos en cada momento, contribuyendo en este caso estas acciones, a aumentar o disminuir el caudal de combustible a los quemadores, y aumentar o disminuir el caudal de agua de alimentación de la caldera.

Con este ejemplo, pretendemos tener una imagen gráfica del Sistema Productivo, antes de hacer una abstracción, y establecer sus componentes y propiedades, para de esta forma, poderlo utilizar como método de aproximación al análisis de la actividad en cualquier tipo de Organización que se encuentre en transformación.

No vamos a profundizar en la teoría del control, diferenciando lo que es control y regulación, lo que es control por ciclo cerrado, control por ciclo abierto, la utilidad del control mixto, ... etc., por entender que todo ello pertenece a la Teoría de Sistemas en general. En todo caso, dado que la Teoría de Sistemas es un sopor-

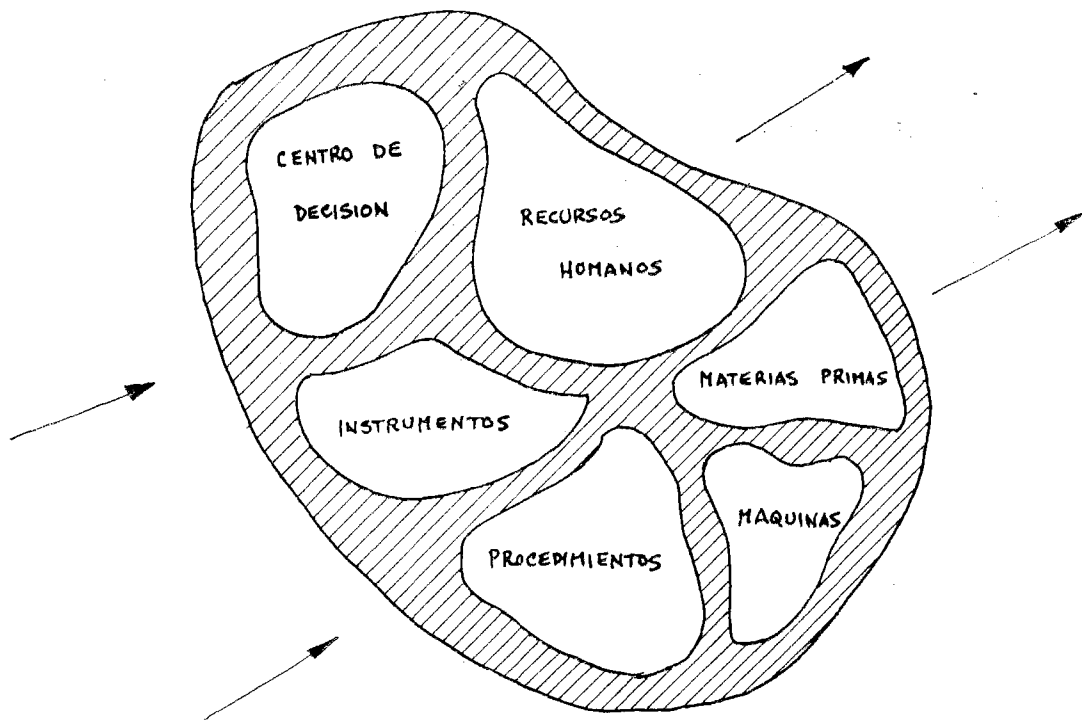


FIG. 25

te básico de nuestro planteamiento, pudiera convenir una pequeña inmersión previa en ella al lector que no esté familiarizado, por lo que sugiero acudir a un tratado específico al respecto.

Haciendo abstracción del Sistema Productivo de la figura 24, representamos en la figura 25 un sistema genérico que está constituido por los ingredientes que de forma evidente se utilizan en los ciclos de producción:

- Mecanismo de decisión.
- Recursos humanos.
- Materias primas.
- Instrumentos.
- Procedimientos.
- Máquinas.

Este conjunto de componentes se agrupan en dos grandes núcleos conceptualmente diferentes en cuanto a su naturaleza: El Centro de Decisión y el Centro de Operación (figura 26).

El Centro de Decisión es el "piloto" del Sistema Productivo, y es la unidad responsable de las relaciones del Sistema Productivo con otros Sistemas, y por lo tanto es la unidad responsable del comportamiento en definitiva del Sistema Productivo.

El Centro de Operación, representa todo el Sistema de Producción interno al Sistema, que es utilizado por el Centro de Decisión para llevar a cabo el cumplimiento de los objetivos establecidos.

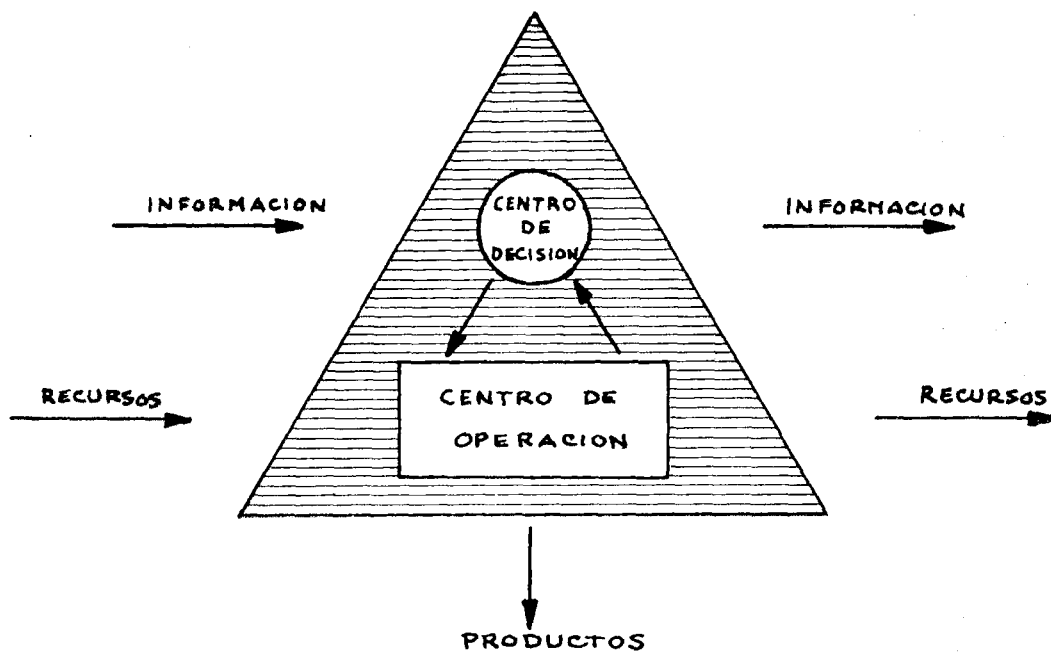


Fig. 26. *El Sistema Productivo*

El Sistema Productivo es por definición un Sistema abierto que intercambia energía e información con el entorno y siempre tiene un fin último que es la consecución de productos; productos que vienen a constituir materias primas o recursos de producción para utilización de otros Sistemas, lo que viene a ser “energía comprimida” al fin y al cabo.

Para observar el fenómeno informático y obtener consecuencias para el diseño de Sistemas, para la captación de información y para determinar las necesidades de información de las unidades organizativa, hay que aproximarse a ellas en forma abstracta e impersonal, permaneciendo fuera del problema, y de aquí la utilidad de hacer abstracción de la unidad organizativa, asimilándola a un Sistema Productivo que tiene una entidad propia, que tiene unos mecanismos internos a los que, a su vez, se puede aplicar la misma óptica, y que su conexión con el exterior es el intercambio de materias, productos, recursos, etc. Sin embargo, el plantear el análisis con abstracción no exime al analista de ser conocedor o de procurarse un buen conocimiento de la actividad específica que está en estudio, porque el primer criterio de análisis es y será siempre un correcto conocimiento de la actividad.

2.1.2. ENTIDADES COMPONENTES Y SUS CARACTERISTICAS INFORMATICAS

Hemos visto que en el Sistema Productivo aplicamos una vez más la lógica binaria que tantas veces utilizamos y lo hemos contemplado como formado por un Centro de Decisión y un Centro de Operación, con misiones respectivas de dirección de la actividad y de ejercicio de la actividad. En la figura 27 se presenta en un cuadro los componentes normales de los Sistemas.

En el plano físico o energético, un Centro de Decisión está formado por personas y/o por mecanismos. Generalmente los Sistemas de Decisión que vamos a analizar de cara a considerar la información que utilizan en el planteamiento de nuestra Base de Datos serán Centros integrados por personas, ya que aquellos Sistemas dirigidos por mecanismos pertenecen informáticamente a niveles demasiado analíticos a los que probablemente no descenderemos. Los componentes informáticos del Centro de Decisión, son: La información de gestión, criterios y directrices.

Información de gestión, que constituye la materia prima, es información de la que permanentemente ha de alimentarse. Esta información puede ser exógena o procedente del exterior del Sistema, o de tipo endógeno o procedente del Centro de Operación, y en este caso se contempla cómo el producto de un mecanismo de “feed-back”. El Centro de Operación dispone también de un paquete de criterios ya elaborados a lo largo de la experiencia, que en realidad no son otra cosa que conjuntos de decisiones cristalizadas y sancionadas por la práctica, actuando en realidad el Centro de Decisión “por excepción” cuando el conjunto de criterios no son aplicables y requieren el ser superados por el superior criterio del Piloto del Sistema. En Teoría de Sistemas, ese conjunto de criterios o mecanismos de decisión preestudiada se identifican con los reguladores por diferenciarlo del control en sí, que es el

mecanismo de decisión de nivel superior. Algo así como el tornillo de aproximación y el de ajuste de un microscopio o de un teodolito.

COMPONENTES	ENERGETICOS	INFORMATICOS
CENTRO DE DECISION (Pilotaje de la Actividad)	PERSONAS MECANISMOS	INFORMACION DE GESTION CRITERIOS DIRECTRICES
CENTRO DE OPERACION (Ejercicio de la Actividad)	RECURSOS HUMANOS MAQUINAS INSTRUMENTOS MATERIAS PRIMAS DINERO	INFORMACION DE GESTION INFORMACION OPERATIVA SISTEMA DE INFORMACION CRITERIOS PROCEDIMIENTOS INFORMACION DE CONTROL

FIG. 27. Componentes Sistema Productivo

El Centro de Decisión recibe del exterior directrices que corresponden a objetivos marcados por un Centro de Decisión de un Sistema de orden superior, y a su vez emite directrices para Centros de Decisión de orden inferior en la jerarquía lógica, internos al Centro de Operación.

El contemplar informáticamente un Centro de Decisión y, por integración, el contemplar el Sistema de Centros de Decisión de la Empresa, es contemplar las necesidades de información de gestión y la estructura organizativa, lo que es interesante para proyectar una estructura lógica de base de datos que aporte valores legítimos y sirva de soporte a aplicaciones integradas.

El Centro de Operación se compone de los elementos usuales que concurren al fenómeno de Producción: recursos humanos, máquinas, instrumentos, materias primas y de consumo y dinero.

En el plano informático, el tipo de información que se utiliza se puede englobar en tres familias: Información de productos, que es la relativa a los productos que el Sistema Productivo "manufactura", información de producción, que constituye por un lado el sistema de información interno al Centro de Operación y por el otro la aportación de información de feed-back al Centro de Decisión, y la información de administración, que constituye la base contable necesaria para el control económico de los recursos.

El Centro de Operación maneja criterios y procedimientos desarrollados y enriquecidos a lo largo del tiempo, y recibe directrices del Centro de Decisión. También recibe información exógena igual que del exterior recibe materiales necesarios para su actividad.

Especial interés tienen las características informáticas de dos componentes específicos, que siempre están presentes en cualquier actividad, que son las personas y el dinero.

Informáticamente pueden ser de interés de un Sistema Productivo los atributos o características de los componentes físicos y sus valores de medida. Cuando se utiliza el dinero como valor de medida estamos contemplando la contabilidad de la Compañía.

Cabe tener presente también dos ingredientes esenciales que aglutinan el Sistema Productivo, que son el tiempo y el espacio. El tiempo es una característica global de cualquier Sistema desde el punto de vista informático permite una ordenación secuencial de la actividad, cosa que vamos a ver al hablar de los Sucesos. El espacio es, sin embargo, una característica individual de cada componente. El espacio asociado con el tiempo, al ser contemplado como atributo de todo componente físico, nos proporciona trayectorias reales geográfico-temporales, que servirán de soporte para manejar los flujos de producción.

2.1.3. PROPIEDADES DEL SISTEMA PRODUCTIVO

La propiedad esencial de un Sistema Productivo es que tiene unos objetivos y que está destinado a la consecución de productos, entendiendo por producto el concepto económico, que es más amplio que el concepto físico, pues puede no tratarse de crear nuevos objetos, sino simplemente de aumentar la utilidad de objetos ya existentes. Si en un Sistema Productivo no logramos identificar sus objetivos debemos modificar nuestros puntos de vista porque no nos estamos aproximando correctamente al análisis de la actividad.

Otra característica es la globalidad, que podría definirse como “el todo es mayor que la suma de las partes”. Es la característica que confiere entidad y personalidad al Sistema. En el Sistema que mentalmente “aislemos” debemos observar de una forma clara que las partes que en él concurren se “enriquecen” al formar conjunto.

El Sistema está en permanente transformación, es decir, intercambia elementos físicos e informáticos con el entorno, bien sea con otros Sistemas Productivos de la propia Empresa, bien sea con otras Empresas, o con el marco jurídico-social en el que necesariamente la Empresa se desenvuelve. El Sistema Productivo es por lo tanto un Sistema abierto.

Otra característica del Sistema es los enlaces o interrelaciones que ligan a sus componentes. El cambio de características o atributos de una parte del Sistema afecta a otros componentes y a su vez es afectado por ellos, Como a su vez el Sistema tiene internamente otros Sistemas de nivel más analítico, existen también objetivos interiores o sub-objetivos. Dichos sub-objetivos muchas veces están encontrados, aunque como conjunto compongan un componente que son los objetivos del Sistema. Queda claro que pueden identificarse objetivos internos contrapuestos.

La propiedad de Entropía está presente como en todo fenómeno real. En el plano físico podemos decir que los potenciales energéticos tienden a aproximarse y por lo tanto se degradan, y en el plano informático que es el que más nos interesa, hay una tendencia a perderse capacidad de diferenciación o segregación en la información.

La propiedad asociativa y distributiva de los Sistemas nos proporciona la forma de descomponer un Sistema en sub-sistemas y la forma de integrar sistemas formando un supersistema.

Con la propiedad asociativa y distributiva el Sistema Productivo adquiere un rango de base celular de la Organización. Con la óptica del Sistema Productivo podemos contemplar el conjunto de la Organización y a la vez cada una de sus partes, descendiendo al nivel de análisis que en cada momento se requiera para el estudio de la actividad sin que por ello pierda coherencia el planteamiento.

2.1.4. UMBRALES DE ACTIVIDAD

El objetivo global de un Sistema Productivo es la consecución de productos, y a este fin se orienta todo el desarrollo de la actividad. El Producto no surge de forma espontánea, sino que se va constituyendo poco a poco, configurándose como tal producto en el momento en que dicho Producto tiene una entidad independiente del Sistema Productivo, abandonando el entorno de actividad que lo generó. En ese momento se produce un Umbral de Actividad (véase figura 28), caracterizado por el fenómeno de consecución del producto, fenómeno que pudiéramos denominar de “ovulación del Sistema Productivo”, utilizando un símil morfogénico.

La importancia del Producto dentro de la actividad general de la Organización y los volúmenes de producción son índices de la relevancia del Sistema Productivo. La relación de volumen de recursos aplicados en el ejercicio de la actividad y el volumen de producción nos proporcionará indicación del rendimiento o “grado de fertilidad” del Sistema Productivo. Un Centro de Actividad que no tuviera Umbrales sería por lo tanto improductivo, o siguiendo la analogía anterior, podríamos decir que su actividad es “estéril”. Esta atención a la relevancia del Sistema Productivo y a su grado de fertilidad es conveniente siempre tenerla presente cuando examinamos los flujos de actividad en la red de producción, porque suele ser general el que los Sistemas más improductivos son los más burocratizados, y son asimismo los que más rigideces introducen en los Sistemas de Información.

Téngase en cuenta que el fenómeno informático y su contenido de significación está directamente relacionado con los Umbrales de Actividad, y éstos son, por lo tanto, manantiales que aportan caudal a la “circulación” de datos. Por el contrario, la actividad en sí es un sumidero que absorbe información porque ésta es en definitiva materia prima para el ejercicio de la actividad. Obviamente, también, por otra parte, es interesante contemplar el grado de fertilidad de un Sistema Productivo, aun cuando no sea tarea del informático establecer juicios a este respecto.

Un Umbral de Actividad está asociado a la coordenada espacial y temporal de consecución del producto, lo que quiere decir que la separación de producto y Sistema Productivo se realiza en un determinado lugar, y en un instante determinado.

Sin embargo, no vamos a establecer como postulado el que el fenómeno productivo sea de naturaleza cuántica, aunque de hecho vamos a ver que la producción continua puede ser asociada igualmente a Umbrales.

Una producción de naturaleza cuántica o discontinua sería por ejemplo la de un Centro de Actividad que fuera una fábrica de muebles. Cada vez que es terminado y entregado un mueble se produce el Umbral, al que se asocian el tiempo de consecución y el lugar de entrega.

Una producción de naturaleza continua sería la producción de vapor de una caldera, contemplada como Sistema Productivo. La naturaleza del Producto es factor determinante de la naturaleza de la Producción, y en este caso el Producto vapor es entregado a una canalización de forma continua. Sin embargo, en estos casos podemos seguir contemplando la producción en base a Umbrales de Actividad utilizando para ello el concepto de "standard de producción". Este standard se medirá en parámetros del caudal continuo que por unidad de tiempo determina un

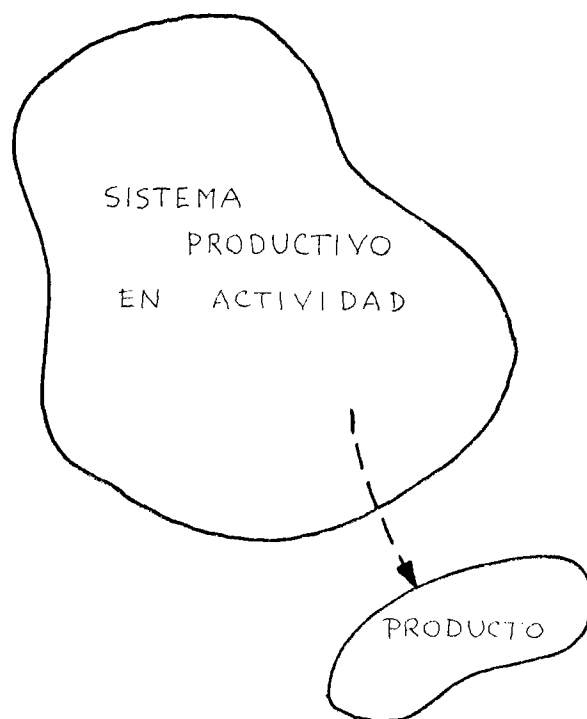


FIG. 28. Umbral de Actividad

cierto volumen de producción, considerándose como un Umbral el momento en el que se alcanza un determinado valor de desviación con respecto al Standard. Es decir, supuestos unos determinados volúmenes de producción continua, asociamos los Umbrales a la aceleración o deceleración del ritmo de producción. En el caso de la caldera, los Umbrales de Actividad corresponderán a las variaciones en el flujo de vapor. La discontinuidad y el concepto de Umbral está por lo tanto siempre presente en los Sistemas Productivos.

La utilidad que vamos a obtener del concepto de Umbral de Actividad, es la de poder contemplar al Sistema Productivo como una "caja negra", es decir, como

una entidad en cuya problemática interior no vamos a penetrar y de la que sólo vamos a examinar las aportaciones que entrega al entorno de la Organización. De esta forma, perfeccionamos la aplicación de las propiedades asociativa y distributiva de los Sistemas. La aplicación de la propiedad asociativa será el identificar una “caja negra” en cuyo interior albergue otras “cajas negras” y los productos que entre ellas se intercambian, aunque ciertamente sólo podrá realizarse con la condición de que dichas “cajas negras”, componentes, constituyan en realidad un Sistema. Realizada la asociación, podemos ignorar a partir de ese momento, los subsistemas interiores, a efectos de relación de la “nueva caja negra” con el exterior.

La aplicación de la Propiedad Distributiva la aplicaremos para examinar en mayor detalle un Sistema Productivo en su interior, y el ejercicio lógico consistirá en identificar un conjunto de “cajas negras” o subsistemas interiores, con unos determinados flujos interiores de productos originados por los Umbrales de actividad de los subsistemas.

La producción de un Sistema Productivo está determinada por el conjunto de Umbrales y los productos que en ellos se generan. La cadencia de Umbrales nos determina “el ritmo” de Producción.

2.1.5. ENTORNO DEL SISTEMA PRODUCTIVO

El Entorno de un Centro de Actividad es el Sistema formado por aquellas entidades del marco geográfico temporal, que pueden estar de alguna forma en relación con dicho Centro de Actividad. El Sistema Productivo y su entorno forman por lo tanto un “universo” o sistema cerrado para cualquier entidad integrada en el Sistema Productivo.

El Entorno de un Sistema es siempre obligado el tenerlo presente, simplemente porque el Entorno influye en el Sistema y el Sistema a su vez influye en el Entorno. Contemplar un Sistema sin atender a su Entorno es operar “in vitro”, y las conclusiones serán siempre vulnerables.

Es importante observar que el Entorno de un Sistema trasciende muchas veces del ámbito de la Organización, porque las interacciones de una Organización o Empresa con el mundo exterior a ella tienen lugar en muchas partes o Sistemas de la Organización, y ésta influye y es influida obviamente y por el exterior.

Cada Sistema tiene por lo tanto su Entorno, y dicho Entorno puede ser contemplado a su vez como un Sistema dual del Sistema Productivo, por lo que a él son también aplicables las leyes que presentamos.

2.2. LOS SUCESOS

El Sistema Productivo responde a los fines para los que existe en sus Umbrales de Actividad, en los que nos revela su trascendencia. Desde el exterior del Sistema

contemplamos a éste como una “caja negra” de la que solamente percibimos significación cuando se produce un Umbral, lo que desde el punto de vista informático constituye un Suceso. Así, cuando un Sistema en actividad pasa por un Umbral diremos que se ha producido un Suceso.

El Suceso es una muestra de la realidad, y para nosotros va a ser la serie de Sucesos la que nos “reconstruya” la actividad del Sistema Productivo.

Conviene distinguir los dos diferentes aspectos de los conceptos de Umbral y de Suceso. El Umbral es una característica del Sistema Productivo, y tiene significación como fin de etapa de un proceso de actividad interno. El Suceso es sin embargo un componente informático cuya significación puede ser o no objeto de interés para los Sistemas Productivos.

El Suceso, como sujeto informático, tiene unos determinados atributos y características, y tiene una determinada coordenada espacio/tiempo. Conociendo el conjunto de Sucesos y sus atributos podemos reconstruir la consecución de objetivos de los órganos de la empresa, o lo que es lo mismo, podemos construir un modelo de la actividad empresarial.

2.2.1. RELACION SISTEMA PRODUCTIVO/PRODUCTO

Vamos a poner como ejemplo a continuación un Centro de actividad que consiste en la Oficina de Ventas de una Compañía Aérea. Permítasenos iniciar así una aproximación, que será continuada con sucesivos ejemplos, a un sistema de Base de Datos del sector operativo/comercial de una Compañía Aérea, ejemplo que ha sido seleccionado no tanto por su interés y conocimiento del autor en él como por la potencia informática y riqueza de matices que aporta como aplicación típica del concepto de explotación de Base de Datos. Se inicia de esta forma una serie de alusiones a la Base de Datos que más adelante plantearemos en toda su extensión.

En la figura 29 se representa una Oficina de Ventas de una Compañía Aérea como Centro de Actividad. La consecución del Producto (Suceso informático), tiene lugar cuando se realiza la venta de un pasaje aéreo a un cliente.

Cabe distinguir dos tipos de datos que son diferentes desde el punto de vista informático. Unos datos son los relativos al Centro de Actividad y son por lo tanto de orden estructural. Son estos datos los atributos de los recursos de la Oficina de Ventas, los costes de su actividad, los datos de su personal, etc. Otro tipo de datos son los relativos a la producción y a los productos, datos que son de naturaleza coyuntural.

Los datos relativos al Sistema son de interés para el estudio de la red de Administración como oportunamente analizaremos (2.4).

Los datos relativos a la producción son de interés para analizar el comportamiento del Sistema Productivo y sus circunstancias a lo largo del tiempo.

Los datos relativos a los Productos son de interés para constituir la red de producción (2.3).

Atendiendo a la cadena de Sucesos que a lo largo del tiempo tiene lugar en la Oficina de Ventas, podemos señalar como atributos de interés: el espacio/tiempo donde se produce la venta, los volúmenes monetarios que se generan con la venta de billetes, el ritmo de venta de billetes, el mercado; es decir, vuelos y días que se

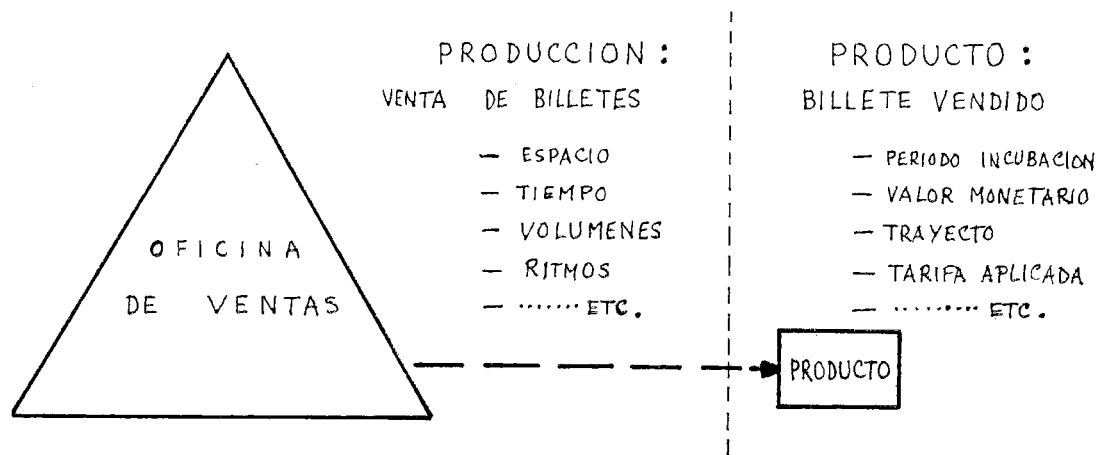


FIG. 29. Suceso

venden, con sus respectivos volúmenes, etc. Estos datos de la producción son atributos informáticos de los Sucesos que nos van a proporcionar información necesaria para la Organización acerca de la canalización de productos que esta Oficina aporta en el conjunto de la producción total, debiendo obtenerse todo el nivel de detalle preciso para el grado de control que nos impongamos como objetivo. La captación de estos datos debe hacerse en tiempo de Suceso, aunque el conjunto de datos de producción, como veremos en el punto siguiente, se obtiene a lo largo de la cadena de Sucesos por integración.

Los datos relativos a un Producto (billete vendido) puede referirse a características del período anterior a la venta, al valor monetario del billete, al trayecto de transportes objeto de la venta, a la tarifa aplicada, etc.

2.2.2. INTEGRACION DE SUCESOS HOMOGENEOS

El Centro de Actividad como Sistema de Producción entrega a su entorno determinados productos a lo largo del tiempo. Estos productos, que en el caso de la Oficina de Ventas de la figura 29 corresponden a billetes vendidos, no son únicos en la red de producción, sino, por el contrario, existirán por lo general otros puntos de venta o Centros de Actividad que proporcionen productos análogos; por lo tanto, si integramos Centros de Actividades y obtenemos conjuntos que formen sistema estaremos contemplando igualmente los correspondientes conjuntos o Sistemas de Sucesos y productos generados por los Sistemas Productivos integrantes.

Siguiendo el ejemplo iniciado en la figura 29, supongamos que agrupamos todas las Oficinas de Ventas que pertenecen a un determinado territorio. Habiendo previamente asociado, desde el punto de vista informático, los billetes vendidos y todos sus atributos relevantes a las Oficinas de Ventas generadoras, podremos posteriormente establecer el mismo tipo de atributos y matices para la producción de toda la zona territorial por integración de los datos relativos a las producciones de cada una de las Oficinas de Ventas del territorio considerado.

Las características informáticas consideradas del producto de cada Centro de Actividad nos permitirá contemplar el producto global de todo un territorio al mismo nivel de características. Así, por ejemplo, obtendremos para el territorio:

- Billetes totales vendidos por cada día.
- Billetes totales vendidos para cada destino.
- Billetes totales vendidos a crédito.
- Producción a crédito para cada destino.
- Etcétera.

Ello suponiendo que el dato correspondiente al destino del pasaje, a si la venta era a crédito o no, etc., hubiera sido recogido como información relevante al captar los datos de los productos.

Por otro lado, la integración de Sucesos producirá informaciones de síntesis sin que por ello la riqueza de matices se vea comprometida porque la potencia informática de la Base de Datos queda determinada en gran parte por la captación de datos. En cuanto mayor segregación imponamos a los datos relativos a los Sucesos, mayores productos informáticos podremos obtener al integrar la información correspondiente a un conjunto de Sistemas Productivos.

En el ejemplo que nos ocupa, en cuanto mayor es el número de detalles que hayamos recogido de la emisión de cada billete mayor es el número de aspectos bajo los que se puede proporcionar información, segregando las influencias relativas de las distintas Oficinas dentro del territorio.

Sí hemos recogido datos tales como:

- Antelación de venta del billete respecto del día del viaje.
- Edad aproximada del cliente.
- Destino final del viaje.
- Reserva efectuada en otra Compañía.
- Tarifa aplicada.
- Etcétera.

podremos determinar por un lado standards diferentes acerca de dichas características para los distintos conjuntos de oficinas que queramos agrupar, y a su vez estaremos en condiciones de ofrecer información acerca de la participación de las diferentes oficinas en los resultados globales o, dicho en otras palabras, podremos establecer "rankings", diferenciando las producciones de las oficinas en base a las características segregadoras antes apuntadas.

Resumiendo:

- Los Sucesos que tienen lugar en un Centro de Actividad son homogéneos con los Sucesos generados en otros Centros de Actividad de la misma naturaleza.
- Todas las características informáticas de los datos recogidos en los Sucesos serán obtenibles por integración al agrupar los Centros de Actividad en sistemas de orden superior.
- El conjunto de Sucesos de los sistemas componentes constituirá el conjunto de Sucesos del sistema integrado.
- Los sistemas homogéneos que han sido integrados podrán segregarse dentro del sistema global bajo el ángulo de aquellas características que hayan sido segregadas al captar datos de los Sucesos elementales.

2.2.3. REDES DE SUCESOS

Los Sucesos tienen una característica constante asociada a ellos que es la coordenada temporal en la que se han producido, por lo que los Sucesos pueden ordenarse cronológicamente. Además, la óptica de sistemas que vamos presentando relaciona unos Centros de Actividad con otros en base al intercambio de productos que practican, por lo que los Sucesos que son los actos “de entrega” de los productos pueden ser relacionados unos con otros con idéntico criterio. Ello nos lleva a la posibilidad de construir unas Redes de Sucesos como la que presentamos en la figura 42. En esta Red de Sucesos nos permite centrar a nivel lógico, dentro de la realidad que tiene lugar, el encadenamiento de Sucesos y las influencias de unos Sucesos en otros, lo que en definitiva, apoyándonos en la cronología de Sucesos, viene a representar un modelo de la forma en que se desarrolla la realidad de la Organización.

Como venimos repitiendo, la Base de Datos ha de constituir un modelo de la actividad, y por ello el análisis profundo de la Red de Sucesos nos ayudará a perfeccionar el planteamiento que hemos de dar a las estructuras de datos. Si los datos van a representar a la realidad y los datos son obtenidos de los Sucesos, el mosaico de Sucesos es relevante ciertamente en la etapa de diseño.

Consecuencia directa de la construcción de la Red de Sucesos es la obtención de la matriz de Sucesos (véase figura 41), que viene a ser una abstracción o digitalización de la Red y un producto más operativo cuando se estudie en profundidad la relación datos/actividad, como veremos a lo largo del planteamiento de la metodología de diseño de una Base de Datos.

2.2.4. RELEVANCIA DE SUCESOS

La relevancia del Suceso está en función de la trascendencia del producto obtenido, trascendencia que cabe contemplarla en función del valor del producto, su utilidad, tiempo de consecución, etc.

En lo que se refiere al punto de vista del Sistema Productivo, la relevancia de los Sucesos puede categorizarse en la misma forma que los objetivos del Sistema Productivo, porque en definitiva los objetivos se alcanzan en los Sucesos. Para el Sistema Productivo la relevancia de un Suceso estará en línea con la relevancia del objetivo alcanzado.

Desde el punto de vista de la Empresa como conjunto, las relevancias relativas de los Sucesos pueden contemplarse analizando los puntos claves y cuellos de botella de la jerarquía natural que representa la Red de Sucesos. En este caso la Red de Sucesos no debe limitarse al ámbito en el que vayamos a diseñar una Base de Datos, sino que habrá que analizar un contexto lo más general posible.

Para el sujeto informático, que tiene que captar atributos de los Productos y características de los Sucesos, deberá ser analizada la relevancia de los Sucesos en el contexto de la Red, teniendo en cuenta todas las cadenas en las que cada Suceso participa, y asociando a ellas los volúmenes de productos, las frecuencias de producción, etc., lo que nos proporciona ángulos de comparación que permitirán calibrar las importancias relativas de los Sucesos y la potencia informática que cada uno ofrece. Ello permite valorar el nivel de síntesis o análisis en el que nos debemos situar a la hora de captación de los datos, buscando punto de compromiso entre las motivaciones económicas y la consiguiente menor significación de los datos y los almacenamientos voluminosos, y por lo tanto caros, que nos lleven a recoger datos irrelevantes para los objetivos informáticos que pretendemos alcanzar.

Los grados de relevancia que concedamos a los Sucesos vienen a aconsejarnos el nivel de análisis con que debemos tratar cada uno de ellos en el conjunto. A partir de este grado de síntesis o análisis que pudiéramos considerar como adecuado, deberemos tener en cuenta los Sucesos en los que, aun rompiendo el equilibrio del conjunto informático, habremos de profundizar en mayor medida por interpretar a la Organización una mayor riqueza de matices por la circunstancia que fuere. En estos casos, a mayor profundidad de análisis requerida deberemos de descender al estudio de Sucesos de menor relevancia.

2.3. LA RED DE PRODUCCION

La Red de Producción es, como veremos, un instrumento que permite al analista de Sistemas contemplar la actividad de la Organización de una forma homogénea y equilibrada. Los componentes de la Red de Producción son:

- Los Centros de Actividad o Sistemas Productivos y sus partes integrantes y sus entornos.
- Los Umbrales y Sucesos.
- Los flujos de productos derivados de los Sucesos.

Los Sucesos vienen a constituir los agentes que distribuyen flujos de productos y, como los Sucesos, deben su legitimidad a la actividad de los Sistemas Productivos, podemos representar la Red de Producción como un conjunto de flujos o trayectorias espacio-temporales, que afluyen a los Sistemas en actividad o se derivan de ellos.

En esta representación vamos a considerar como elemento dual de un Suceso la absorción de un producto por parte de un Sistema Productivo. Con este aspecto dual del Suceso toda trayectoria espacio-tiempo representada en la Red de Producción, tendrá su nacimiento en un Suceso y finalizará en el punto donde se produce el Suceso dual correspondiente a la entrega del producto.

Recordando la analogía genética a la que en otros momentos hemos aludido, si el nacimiento de un producto lo asociábamos a una ovulación, la entrega del producto en el Suceso dual, puede contemplarse como una fecundación. Bajo esta analogía tienen considerable fuerza las características de fertilidad o esterilidad que presenten los Sistemas Productivos.

Vamos a examinar la Red de Producción que se presenta en la figura 30. Corresponde a la actividad de una Compañía aérea en su aspecto operativo, relacionado con el tráfico de pasajeros. Vamos a describirla a continuación advirtiendo de antemano que vamos a emplear tantos términos técnicos propios de la naturaleza de la actividad como sean necesarios, porque a lo largo de la obra se pretende ir construyendo una Base de Datos que pueda aplicarse realmente y que corresponda a la vida práctica y no al terreno de las ideas especulativas.

Los Centros de Actividad denominados como "X" corresponden a Oficinas de Ventas de la propia Compañía aérea. En estos Centros de Actividad es donde tiene lugar la formalización de la venta del producto terminando el contacto que en algún momento se inició con el cliente, con la venta del correspondiente pasaje aéreo.

Los Centros de Actividad de naturaleza Y corresponden a las Agencias de Viajes autorizadas por la Compañía aérea para formalizar la venta de sus billetes. Estos Sistemas productivos de tipo Y tienen por lo tanto una actividad análoga a la de los Centros X, aunque hay que señalar que en vez de pertenecer a la organización propia pertenecen a su entorno, y por lo tanto no sólo se producen en ellos Sucesos productivos para la Compañía aérea considerada, sino también para la competencia.

También debe observarse que entre los Centros X y los Y hay en realidad un intercambio de productos porque los Centros Y son unas prolongaciones de las Oficinas X, motivadas para el logro de una mayor cobertura de mercado, y que han de utilizar formularios de billetes administrados por las Oficinas X.

Los Centros de Producción del tipo Z corresponden a Oficinas de otras Compañías aéreas competidoras, las cuales en virtud de las posibilidades de intercambio de servicios previamente establecidos, canalizan ciertos pasajeros hacia los servicios programados por la Compañía protagonista de nuestro ejemplo. Estos Centros Z utilizan formularios de billetes propios que son susceptibles de utilizarse en vuelos de otras Compañías en virtud de determinados acuerdos existentes y dentro de determinadas limitaciones que no vamos a analizar.

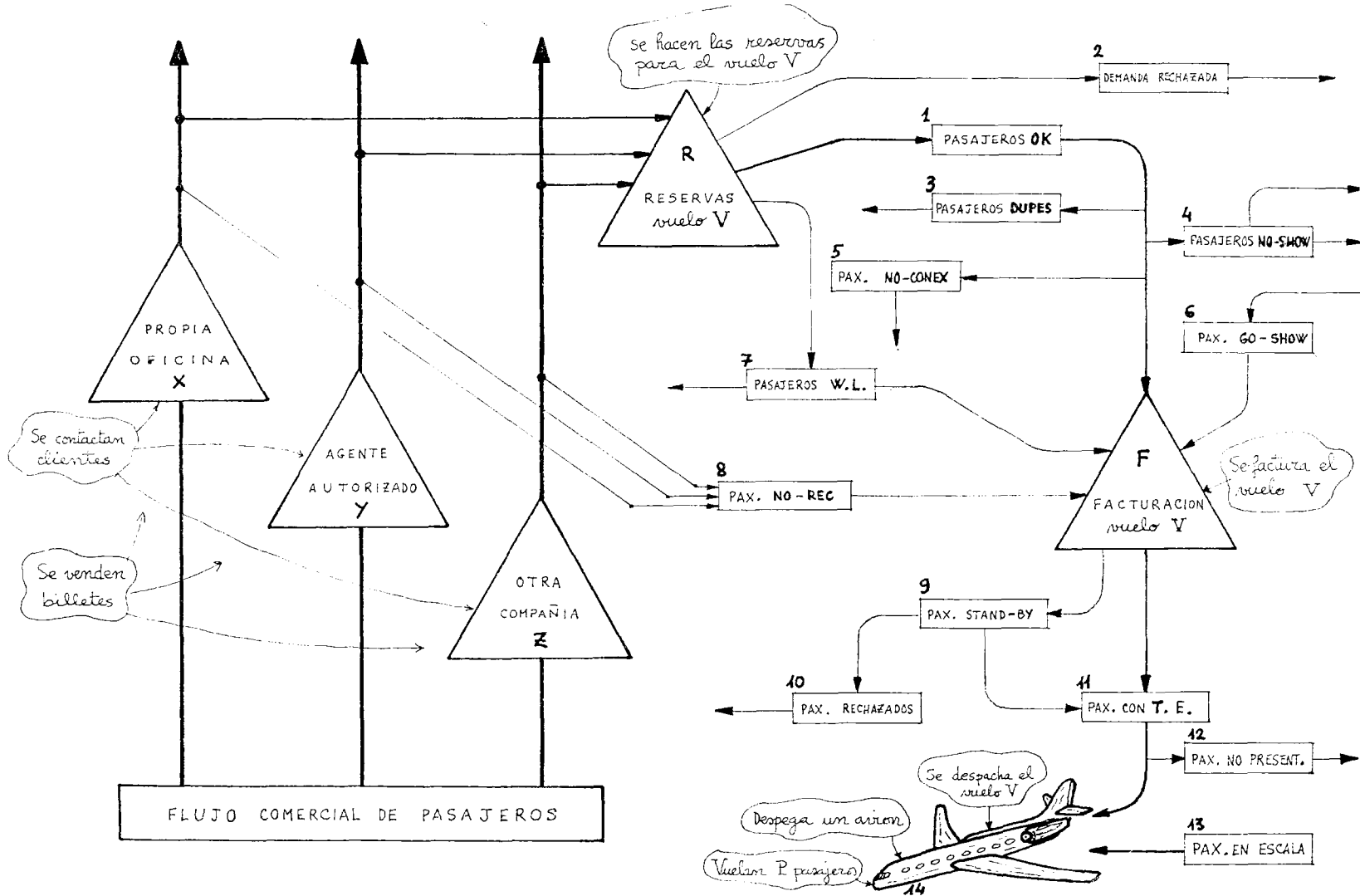


FIG. 30

Los Sistemas Productivos X , Y y Z modelizan al conjunto de Centros de Actividad que activan el flujo comercial de pasajeros que es de esta forma captado por la Compañía aérea.

En la distribución de este flujo comercial de pasajeros a los diferentes vuelos programados por la Compañía interviene el Centro de Reservas R como agente básico distribuidor.

El Centro de Reservas R tiene como misión la distribución de los pasajeros que componen el flujo comercial, en los diferentes vuelos que va a operar la Compañía. El Centro R produce un flujo básico de pasajeros, denominados en el argot aeronáutico pasajeros OK (1), que es el flujo básico de entrada para el Centro de Actividad F , que es el Centro destinado a la gestión final de aceptación de pasajeros a la aeronave para su transporte, Centro de Actividad que obviamente se encuentra ubicado en un Aeropuerto. Su actividad se denomina generalmente Facturación.

Examinando la Red de Producción de la figura 30, observamos que los Centros de Actividad, constituyen los elementos básicos de soporte de la Red, y los flujos de Productos, constituyen los elementos dinámicos. Estos flujos de Productos, tienen las mismas propiedades generales de cualquier sistema circulatorio de distribución, por lo que se puede examinar en detalle recorriendo las trayectorias espacio-tiempo, las extracciones de flujo y las adiciones de flujo que ciertamente existen en posiciones intermedias a los Centros de Actividad. Una extracción de flujo, puede derivar flujo al Entorno, o bien a otro Centro de actividad, lo mismo que una adición o captación intermedia de flujo, puede proceder de otro Centro de Actividad o igualmente del Entorno. Por ejemplo, en la figura 30, la extracción de flujo correspondiente a pasajeros “dupes” (duplicados) (3), corresponde al hecho de que existían ciertos pasajeros reservados contabilizados por partida doble. Esta extracción de flujo, conduce al Entorno y representa un flujo improductivo motivado por un determinado error de operación. Análogamente, el flujo de pasajeros “go-show” (6), que corresponde a pasajeros que entran en contacto por primera vez con la Compañía aérea en el Aeropuerto, en donde declaran su intención de volar, es una adición de flujo, en este caso productivo, que procede del Entorno, y que alimenta al Centro de Actividad F .

La subdivisión de flujos en sub-flujos y la consideración de derivaciones y aportes de flujo, puede llevarse a un nivel de análisis tan profundo como sea preciso, segregando los flujos en tantas características como sean necesarias. La riqueza y diferenciación de datos que hayamos contemplado en los Sucesos, está en relación directa con la riqueza de diferenciación que asociamos a los flujos. La tarea informática, viene a constituir un trabajo sistemático de “colocación” de contadores de flujos en las diferentes derivaciones, y la “lectura” de dichos contadores, en diferentes valores de tiempo.

En los Sistemas Productivos, puede establecerse una división de la actividad, atendiendo a la naturaleza cuántica de los procesos, lo que en determinadas ocasiones simplifica, y por lo tanto enriquece, la tarea de análisis. Por ejemplo, el Centro de Actividad de Reservas (R), produce un Suceso, cada vez que se negocia una reserva con un cliente. Como sucede siempre que una reserva se realiza para un determinado vuelo, puede establecerse el Suceso global “reservas para el Vuelo V ”,

como el conjunto de los Sucesos que producen reservas para dicho vuelo. Ello viene a suponer una división de la actividad del Sistema Productivo R , en tantas partes como vuelos existan programados. Procediendo de esta forma, se realiza el estudio de la Red de Producción para un determinado vuelo, lo que resulta más simple, obteniéndose después todo lo relativo a la actividad global por el proceso de integración de Sucesos homogéneos.

Este planteamiento de Red de Producción, permite análisis tan sintéticos o globales o por el contrario, tan profundos o detallados como sea necesario ya que, se puede fraccionar el análisis de la actividad en tantas partes como se desee, con la facilidad de profundizar en detalle en las partes en que sea necesario, sin que ello obligue a hacer lo propio en los demás componentes. Un ejemplo elocuente veremos más adelante en la figura 31.

Hay que tener en cuenta que la Red de Producción es un modelo de la realidad pragmática de la Organización, y conviene tener muy claro si lo que se está construyendo es un modelo real o por el contrario es un “posible” modelo deseado por el analista. Por ello, queremos dejar muy claro que el hecho de proponer modelos alternativos equivale a la tarea de organizar que siempre es complementaria de toda labor informática, pero que sin embargo no vamos a tratar en la obra por considerar que pertenece al ámbito general del análisis de Sistemas. Nos vamos a centrar en la tarea propia del diseñador de una Base de Datos, cuya labor consiste en “destilar” un modelo informático, haciendo abstracción de si se desean introducir modificaciones o no en la organización y aun cuando, y en otro momento nos referiremos a ello, sea aconsejable mantener vigente una actividad perfeccionista a la hora del planteamiento de la Base de Datos.

El objetivo de la construcción y utilización de la Red de Producción, es en definitiva el obtener conclusiones de qué registros o ficheros han de ser contemplados en la Base de Datos, de qué información han de contener, cómo han de estructurarse y cómo interesa relacionarlos.

La Red de Producción inicia un proceso sistemático de análisis que como veremos, nos conduce al diseño del Esquema de la Base de Datos.

2.3.1. FLUJOS DE PRODUCTOS

Los flujos de Productos, son las trayectorias espacio-temporales de los Productos, con la concepción del Producto que hemos presentado anteriormente. Estas trayectorias tienen un punto de nacimiento y un punto de llegada.

Los flujos de Productos, se inician en un determinado Sistema Productivo, con una coordenada espacio-tiempo, que viene determinada por el correspondiente Suceso generador. Los flujos pueden también iniciarse en el Entorno del conjunto de Sistemas Productivos, que estamos contemplando, y cabría también analizar, si interesa, el Centro de Actividad de procedencia aunque éste pertenezca al Entorno.

Los flujos de Productos, conducen a Sistemas Productivos, donde finalizan, y para los que constituyen una aportación al mantenimiento de su actividad. Los flu-

jos pueden transpasar la frontera con el Entorno si el producto correspondiente que representan, trasciende del área de la Organización. Análogamente al caso del nacimiento de los flujos, también puede considerarse la tipología de los Centros de Actividad que alimentan, lo que en conjunto viene a representar la influencia del sistema organizativo en el Entorno.

Un Sistema Productivo, tiene por lo tanto trayectorias afluyentes de productos cuya segregación y tipología dependen de los Sucesos que los originaron “aguas arriba”. Asimismo, el Sistema produce un conjunto de flujos derivados que dependen del conjunto de Sucesos que tienen lugar durante el ejercicio de la actividad del propio Sistema.

Un Sistema Productivo, es influido por los flujos afluyentes que integra, y a su vez influye en su Entorno por medio de los flujos derivados que produce.

Sin embargo, un Sistema Productivo, asimilado al concepto de “caja negra” al considerar los flujos afluyentes y los flujos derivados, puede ser analizado en mayor profundidad, penetrando dentro del sistema, descomponiendo dicho sistema en “cajas negras” más elementales y analizando los flujos internos al sistema que conexionan los subsistemas interiores. Vamos a ver como ejemplo en la figura 31, la forma en que examinamos la Red de Producción interna al Centro de Actividad R de la figura 30.

El Sistema Productivo R , considerado como una “caja negra” (fig. 30), intercambia con los Sistemas X , Y , Z , las peticiones y contestaciones acerca de reservas de pasaje. A los Centros de Actividad F , proporciona el producto final que básicamente está constituido por los pasajeros reservados para cada vuelo. También tiene otro tipo de intercambios con otras unidades productivas y con el Entorno, que no vamos a detallar. Si penetramos dentro del Sistema R , para establecer un análisis más profundo de los flujos de producto, y para considerar otros flujos internos del Centro R , obtendremos el esquema de la figura 31, en la que se han referenciado por números y letras, los flujos y los Centros de Actividad internos para mejor comprensión.

El flujo 1, es el correspondiente a las peticiones de reservas que una Agencia de Viaje X realiza a un Centro de reservas RV de ventas por teléfono. Las peticiones son atendidas por unos agentes que a su vez generan el flujo núm. 11, enviando mensajes de petición al inventario de control de reservas que generalmente está manejado por un ordenador (RO en la figura). El flujo 12, corresponde a los productos de respuesta que el centro RO entrega al centro RV , quien a su vez proporciona al centro X la respuesta pertinente mediante el flujo 2.

El centro Y , que corresponde a una oficina de billetes propia de la Compañía, intercambia en la misma forma descrita, los flujos 3 y 4 con el centro R . Dentro del centro R , el flujo 3 lo podemos considerar subdividido en los flujos 7 y 9, correspondiendo el número 7 a relación telefónica entre el centro Y y el RV , y el flujo 9 a una comunicación directa del centro Y con el RO , es decir a un acceso directo al inventario de reservas. Análogamente, los flujos 8 y 10 de respuesta, componen el flujo 4 de contestación del centro R al centro Y .

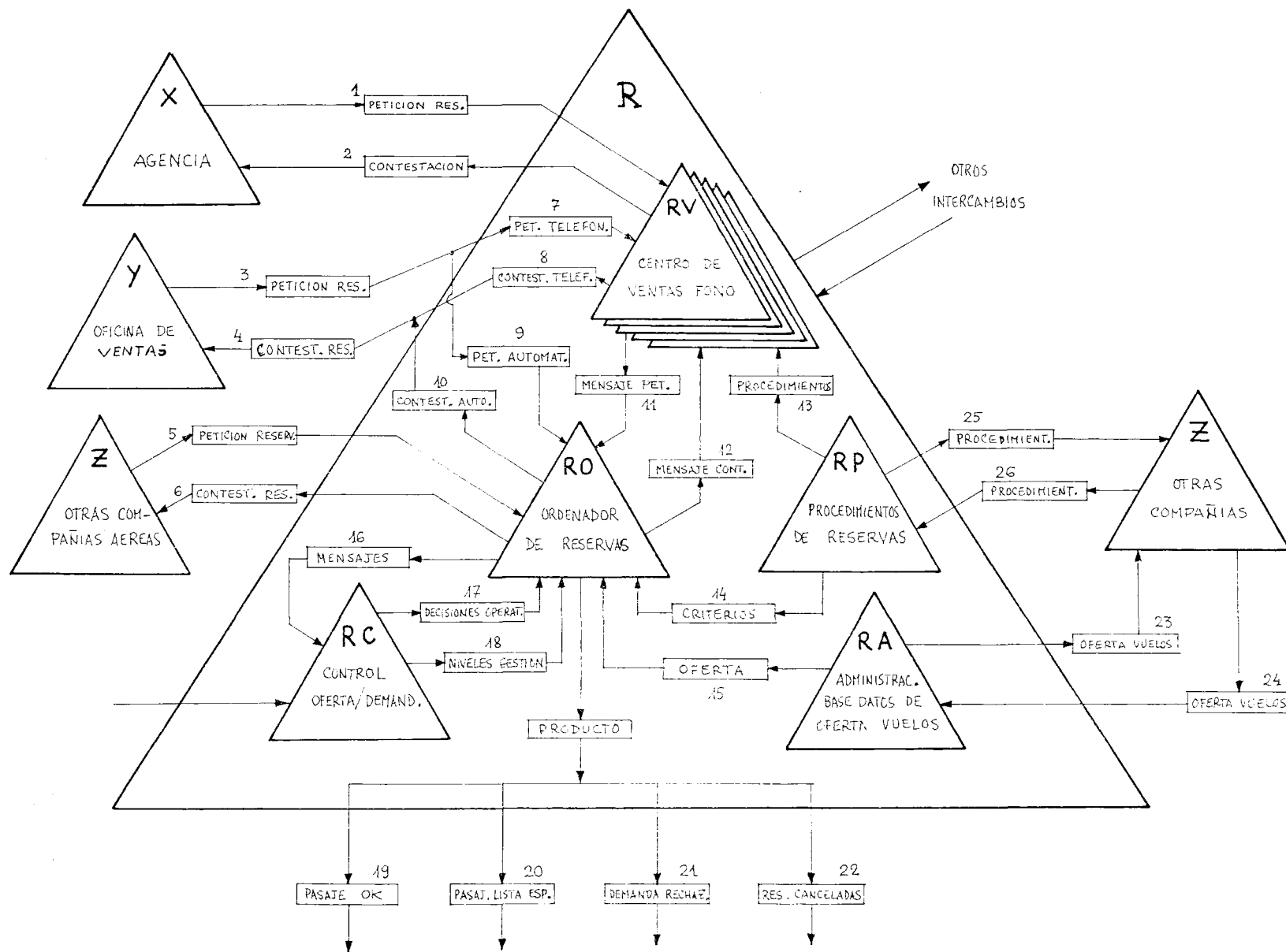


FIG. 31

Hay que señalar que una Compañía aérea de cierta entidad, dispone en la actualidad de un sistema de operación de reservas en tiempo real, soportado por una red de comunicaciones que permite el acceso directo a los inventarios de reservas desde puntos geográficos diversos donde los volúmenes de operación así lo justifiquen. Aun cuando no se ha representado en la figura, el sistema electrónico puede enlazar directamente a Agencias de Viaje.

Las otras Compañías aéreas, representadas en la figura 31, por Centros de Actividad del tipo *Z*, intercambian productos a través de los flujos 5 y 6 con nuestro centro de reservas. En este caso, la comunicación es directa con el centro *RO* merced a determinados procedimientos internacionales que permiten el diálogo recíproco de ordenadores, procedimientos que son intercambiados con las Compañías por medio de los flujos 25 y 26 representados en la figura. Este intercambio de procedimientos, también permite que el centro *Z*, pueda en ciertos vuelos, realizar la reserva de forma inmediata sin esperar a la comunicación con el centro *R*, de la misma forma que el centro *R* puede producir reservas en vuelos de la Compañía *Z* de forma inmediata reportando su realización *a posteriori*. Ello tiene lugar a través del intercambio de productos entre los centros *RA* de administración de Base de Datos de la oferta de vuelos por medio de los flujos 23 y 24. Los productos intercambiados son disponibilidades de plazas de avión a la venta, para incorporar en el caso de nuestra Compañía, al inventario central *RO*, por medio del flujo 15 canalizador de la oferta.

Adviértase que dentro del centro *R*, existe una unidad *RA* de administración de una Base de Datos operada en el ordenador *RO*, y téngase en cuenta que esta Base de Datos puede ser independiente de otro sistema más general de Base de Datos de gestión que modelice redes productivas de la empresa de mayor nivel de síntesis, o bien, pueda estar integrada en ella. En cualquier caso, ambas estarán relacionadas y existirán los oportunos interfases.

El centro *RO*, eminentemente operativo, es pilotado por el centro *RC*, denominado centro de Control de Espacio en el argot aeronáutico, que está practicando un permanente análisis de la relación oferta-demanda, perfeccionando en cada momento los criterios de operación del centro *RO* por medio de los flujos 17 y 18 que corresponden a las variaciones oportunas de los parámetros de control, decisiones que son elaboradas a partir del flujo 16 de feed-back y de la información exógena general de que dispone el centro *RC*.

El flujo de productos que se deriva del centro *R*, a lo largo del proceso productivo, lo hemos considerado subdividido en 4 componentes (flujos 19, 20, 21 y 22), que corresponden a:

- Pasajeros reservados para un determinado vuelo *V*.
- Pasajeros en lista de espera para un determinado vuelo *V*.
- Pasajeros cuya reserva ha sido denegada para el vuelo *V*.
- Pasajeros que han cancelado su reserva para el vuelo *V*.

Como iremos viendo, a la hora de ir elaborando el diseño de los registros, y estructurando por lo tanto los campos de información que va a contener la Base de

Datos, iremos advirtiendo las necesidades que se nos puedan presentar de profundizar más en un proceso de análisis de la actividad. Tomemos como ejemplo la regla distributiva que se ha desarrollado de forma práctica en la figura 31. La Red de Producción de la figura 30, es compatible y complementaria de la Red de la figura 31, sin embargo, los niveles de síntesis son diferentes. De esta forma, hemos profundizado en la Red de Producción interior al centro R , sin perturbar el equilibrio de la Red de la figura 30. Es por tanto tarea esencial en el proceso de modelización para nuestra Base de Datos, el fijar los niveles de análisis convenientes, lo que conduce al grado de diferenciación adecuado de los flujos.

Téngase también en cuenta que un determinado flujo puede considerarse como la suma de un conjunto de flujos componentes, por ejemplo el flujo 22 (fig. 31), que corresponde a los pasajeros que han cancelado su reserva para el vuelo V , pudiera considerarse subdividido por los siguientes componentes:

- Pasajeros cancelados con más de un mes de antelación a la fecha del vuelo.
- Pasajeros cancelados entre 30 y 15 días de antelación.
- Pasajeros cancelados entre 15 y 3 días de antelación.
- Pasajeros cancelados en los últimos tres días.

Asimismo, pudiéramos considerar el flujo 22 subdividido por otra naturaleza de componentes:

- Pasajeros cancelados que previamente fueron reservados a través de una Agencia X .
- Pasajeros procedentes de una oficina de billetes Y .
- Pasajeros reservados por otra Compañía Z .

También cabe compaginar ambos tipos de naturalezas en la subdivisión de flujos, y cabe el plantear una matriz de flujos que en conjunto corresponderían al flujo 22. Algo así como suponer a un flujo como un “determinado cable” que al seccionarlo advertimos una gama de cables componentes de inferior tamaño.

Lo que no puede faltar nunca es la atención a la mejor consecución de los objetivos finales, porque la consideración de las diferenciaciones y sub-divisiones que habremos de establecer en los flujos, nos será dada por el criterio que siempre debemos tener presente de mejor servicio a los objetivos planteados.

Podemos quizá enunciar como regla, que para conseguir un objetivo informático, debemos de introducir un contador o medidor de flujo en un determinado punto de nuestro modelo. De acuerdo con ello, deberemos diferenciar los flujos en sus componentes, en la medida en que necesitemos poner diferentes contadores de flujo.

Como los flujos son trayectorias espacio-temporales, los datos que nos ofrezcan los contadores, serán los elementos informáticos que una vez recogidos en la Base de Datos, nos proporcionarán el conjunto de “fotogramas” de la “película” de la actividad.

Resumiendo, el estudio de la Red de los flujos de productos, nos permite dentro del ciclo de aproximaciones sucesivas obligado en un proyecto de este tipo, ir determinando las diferenciaciones necesarias, los nuevos niveles de profundidad a los que es preciso llegar, y de esta forma, ir situando en el modelo los oportunos contadores de flujo, que después continuaremos hasta agruparlos en un Esquema de Base de Datos que ordene el proceso informático de las “lecturas de contadores”.

2.3.2. RELACION SUCESOS/REGISTROS

El estudio de la Estructura de la Actividad y de la Estructura de los Sucesos, nos tiene que conducir a la concepción de una Estructura de Datos que constituya el soporte de la Base de Datos y por lo tanto, el modelo de la Organización que nos proponemos elaborar como objetivo fundamental de diseño.

La Actividad es ejercida, como hemos visto, por los Sistemas Productivos, Sistemas que como tales tienen una vida propia y constituyen un ámbito dinámico de información, en gran parte, información interna al Sistema. Esta información interna, es una Información de Estructura ligada a los Centros de Actividad y los datos informáticos que de esta naturaleza manejamos, serán por lo tanto, datos asociados a los Sistemas Productivos.

Por otro lado, en el desarrollo de la Actividad, se producen Umbrales que originan como consecuencia, los Sucesos Productivos a los que nos venimos refiriendo. En estos Sucesos, tiene lugar la consecución de Productos y la alimentación de la Red de Flujos de producción. Los datos que manejamos de los Flujos de Producción, constituirá una Información de Coyuntura que está ligada a los Sucesos, ya que son datos producidos o asimilados por estas “cajas negras” que constituyen los Centros de Actividad, que son por lo tanto los elementos que aportan la componente dinámica al Sistema.

Esta dicotomía configurada por la Información de Estructura y la Información de Coyuntura, nace del propio concepto de Umbral, y por ello, el Suceso constituye el elemento “gozne” de todo nuestro planteamiento. Por ello, a la hora de considerar cuáles son los datos informáticos que vamos a incluir en nuestro modelo de la Organización, es decir, en nuestra Base de Datos, ha de ser el Suceso necesariamente el elemento básico a contemplar. Conviene aclarar, sin embargo, que la Información de Estructura y la Información de Coyuntura, son dos extremos, y realmente los conceptos que encarnan, serán utilizados en términos comparativos puesto que en cualquier Registro de Información, Podemos generalmente identificar tanto componentes estructurales como coyunturales.

La Información de Estructura, sirve de marco a los Sucesos. Por ejemplo, las figuras 30 y 31. En ellas el Centro Productivo R (Reservas), aporta a la Red de Producción pasajeros confirmados (Flujo 19) para cada uno de los vuelos que constituyen la oferta de servicio. Analizando un determinado vuelo V bajo el prisma de los Sucesos que producen el Flujo 19 de pasajeros confirmados, el vuelo V se

constituye en Recurso Productivo para el Centro de Reservas, y como tal Recurso, es algo que forma parte de la Estructura de los mecanismos de Producción.

Supongamos que el vuelo V considerado, une las ciudades de Madrid y París. Los datos correspondientes a la identificación de dichas ciudades, serán obviamente datos muy estructurales, pues corresponden a la naturaleza geográfica que es estable, así como la distancia kilométrica entre ambas ciudades. El sector de mercado en el que dicho vuelo está inscrito, es también un dato estructural, al menos en la misma medida en que estructural sea nuestra estrategia de distribución de mercados.

De estos datos y otros más que se derivan del Suceso, habremos de ir tomando aquellos que consideremos de suficiente relevancia para el contexto general de nuestros objetivos. Fundamentalmente extraeremos en primer lugar aquellos datos que nosotros podamos utilizar como “marco” de la producción de Sucesos. Así, en nuestro ejemplo, seleccionaremos:

- El número de la línea a la que pertenece el vuelo V .
- Los días en que dicho vuelo opera.
- Los horarios de salida y de llegada.
- El tipo de avión.
- La capacidad de asientos.

Estos datos son de naturaleza Estructural con relación a la Actividad operativa de Reservas porque constituyen recursos básicos para el ejercicio de la Actividad, aunque sin embargo, serían de naturaleza menos Estructural y más Coyuntural, si los contempláramos en una planificación a medio o largo plazo o simplemente si manejáramos series históricas de datos.

Dependiendo de los objetivos de la Base de Datos, también podrían ser relevantes los datos ligados a la propia Estructura del Centro de Actividad, es decir, la Información correspondiente a la propia “fisiología” del Centro de Actividad y los elementos informáticos asociados a las entidades físicas que constituye el inventario de recursos del Centro de Actividad.

Así como la Información de Estructura, sirve esencialmente de marco a la generación de Sucesos, la Información de Coyuntura es un producto derivado de dichos Sucesos. Siguiendo con el mismo ejemplo, profundizando en el análisis del Flujo 19 de la figura 31, podemos considerar en detalle, el conjunto de datos que se producen en los Sucesos generadores de dicho Flujo. Serán datos componentes:

- El nombre del pasajero que en cada Suceso es confirmado.
- El vuelo de procedencia anterior al punto de embarque, si existe.
- El vuelo de continuación, si existe.
- La dirección donde se puede contactar al pasajero.
- El punto de origen de la petición de reserva.
- La hora y fecha en que el pasajero fue confirmado.
- Etcétera.

Estos datos, nos definen los atributos del Producto, y los recogeremos en los ficheros de nuestra Base de Datos, en la medida en que sean relevantes para nuestros objetivos bien porque los necesitemos por propio interés de la Base de Datos como Sistema, como por necesidades operativas de los propios Sistemas Productivos. Estos datos, constituyen la determinación cualitativa y cuantitativa de los Flujos de la Red de Producción.

La dinámica de consecución de Sucesos, trae consigo el que se produzcan a su vez determinados Sucesos inducidos, motivados por la consecución de Umbrales inducidos de Actividad. Por ejemplo el Suceso “pasajero reservado en un determinado vuelo V ”, se repetirá muchas veces. Este proceso dinámico, puede traer consigo la inducción de un Umbral de naturaleza superior cual sería en este caso el hecho de estar ya completo el vuelo V y no poderse realizar una nueva reserva. Este Suceso inducido, viene a aportarnos una Información altamente relevante cual es el hecho de que la Actividad de reservar pasajeros en el vuelo V , no puede seguir siendo ejercida. Los datos generados por estos procesos inducidos, suelen recogerse en los Registros como Información de Estado y constituyen Indices indicadores de Coyuntura.

Otro tipo de Información es la Información Histórica y de Síntesis, que constituye el marco de referencia que da soporte a la dinamicidad del modelo. La dinámica de Sucesos, viene a comportarse como un Agente calificador de los elementos estructurales de producción, añadiendo a estos elementos nuevos atributos. En nuestro ejemplo del vuelo V de Madrid a París,

- Los factores de ocupación mensuales históricos de las líneas Madrid-París.
- Las tarifas medias mensuales históricas resultantes.
- El número de pasajeros transportados en cada mes.
- Etcétera.

constituyen extractos sintéticos de la Información de Coyuntura generada en la Red de Producción. Estos extractos sintéticos o “comprimidos de información”, se convierten en datos estructurales, determinando ni más ni menos que la trayectoria histórica del fenómeno productivo.

Este tipo de Información, tiene que irse destilando con la vida de la Base de Datos, en forma paralela, porque la Base de Datos, como modelo que es de la empresa y de su actividad, necesita de la panorámica temporal para ser operativo. Este proceso de “maduración” o “destilación” de los datos de Coyuntura, viene a configurar el concepto que he dado en denominar “Tiempo presente”, concepto que será objeto de atención en el punto 4.1.

Como es usual en la práctica informática, vamos a manejar el Registro como unidad lógica de la Base de Datos. El Registro, constituye el patrón lógico en torno al cual se construye una familia de Registros homólogos que define lo que generalmente denominamos como Fichero.

Tomamos el Registro como unidad lógica, porque es el elemento frontera entre su propia estructura interna, estructura de campos de información, y la estruc-

tura externa de relación entre diferentes Ficheros. Internamente, el Registro tiene una Estructura, estructura que suele presentar una composición de tipo jerárquico tal y como se define cuando se utiliza el lenguaje COBOL. Sin embargo, el Registro en una unidad lógica útil porque como una sola unidad, tienen lugar los Procesos Informáticos fundamentales que tendrán lugar en la Base de Datos, es decir, será todo el Registro como elemento único que engloba a su estructura interna, el objeto formal de cualquier operación de almacenamiento, transferencia, actualización, etc. Solamente cuando la unidad lógica sea integrada en la memoria principal con una Unidad específica de Tratamiento, serán separados los elementos constitutivos de su estructura interna.

Conviene resaltar que hablar de un diseño de un Registro, es lo mismo que hablar del diseño de un Fichero, y para diferenciar nuestras referencias al diseño del Fichero o a un determinado Registro que forma parte de dicho Fichero, convendremos en denominar como Registro Tipo, al Registro "patrón" definitorio de la estructura interna del Fichero, y denominaremos como Registro Específico a los Registros "concretos" que están integrados en el Fichero.

Será un Registro Tipo: N.º VUELO/TRAYECTO/HORA SALIDA/HORA LLEGADA.

Un Registro Específico perteneciente al Fichero definido por este Registro Tipo, será: 914/MALAGA-MADRID/1230/1445.

Generalmente, nos referiremos en mayor medida al Registro Tipo, que es la Estructura lógica a configurar para el diseño de la Base de Datos. A la hora de diseñar Registros, es preciso tener presente, la presunta Actividad de los datos sobre los que estamos considerando una posible integración en el Registro. Aunque esta particularidad, será siempre examinada posteriormente en profundidad, ya que el desarrollo de una Base de Datos, como cualquier otra Actividad de proyecto, se realiza por un proceso de aproximaciones sucesivas. Aunque en otro capítulo se analice esta circunstancia, aclaremos que nos referimos a la Actividad de los datos en el sentido de su dinámica de utilización en el proceso informático de la Base de Datos.

Dentro de un mismo Registro, pueden incluirse datos estructurales y datos coyunturales, pero sin embargo, el principio lógico a preservar, es el de buscar una "homogeneidad suficiente" entre las tasas de Actividad de los datos que constituyen el Registro, sin considerar naturalmente aquellos campos del Registro que se utilizan para establecer referencias cruzadas, cuya Actividad dice más de la dinámica de cambio de la Estructura externa al Registro, que de la interna. Naturalmente, la regla lógica, es una guía a seguir hasta el límite razonable en compromiso con los demás factores de proyecto, siendo el ejercicio continuado de la práctica profesional, el que nos proporcionará un criterio analógico a este respecto.

Por lo general, los tipos de Registros que manejaremos, serán:

- Registros Estructurales relativos a Centros de Actividad y a sus recursos. Estos inducirán, o se asociarán a ellos, series temporales en las que se cuantifican parámetros de actividad. Si contemplamos los recursos

básicos, “personas y dinero”, estaremos recogiendo datos de la Red de Administración que será tratada en el punto 2.4.

- Registros estructurales que servirán de soporte a los datos generados por los Sucesos.
- Registros asociados a Sucesos, en los cuales se incluirán datos estructurales y coyunturales.
- Registros asociados a productos, incluyendo datos coyunturales derivados de los Sucesos.
- Registros de síntesis históricas. Incluirán datos asociados a las producciones, así como datos asociados a los Centros de Actividad, datos en este caso muy ligados a los datos estructurales de Recursos.

Como regla general para la determinación de los Registros, hay que contemplar el Suceso como fenómeno generador y en el Centro de Actividad como entidad que va a dar vida al Registro. La estructura de los Registros, corresponderá a la estructura de Flujos derivados de los Sucesos y a la composición de los Flujos afluyentes a los Centros de Actividad.

Dentro de los Registros, tendremos que segregar datos, en la medida en que necesitemos segregar Flujos derivados de los Sucesos, y así los Registros contendrán tantos datos como “contadores de Flujo” situemos aguas abajo o aguas arriba del Centro de Actividad. Por ejemplo, remitiéndonos al producto del Centro *R* (figura 31), si para cada vuelo nos interesa recoger Registros estadísticos con valores cuantitativos de los Flujos 19, 20, 21 y 22 correspondientes a pasajeros confirmados, pasajeros en lista de espera, pasajeros a los que se ha denegado la reserva, y pasajeros cuya reserva ha sido cancelada, este Registro estadístico, deberá contener 4 campos con los valores cuantitativos de estos 4 Flujos, además del Campo de Cabecera que identifica a la entidad vuelo.

Si adicionalmente, los flujos 19, 20, 21 y 22 de la figura 31, quisiéramos descomponerlos en sub-flujos, segregando la información por un lado, atendiendo a la procedencia de la petición de reserva, es decir, según la petición proceda de una Agencia, de la propia Compañía, o de otra Compañía, y por otro lado segregando en función de la antelación con que se ha producido el Suceso con relación a la fecha del vuelo (hasta 30 días, entre 30 y 15 y de 15 días a la salida), los campos del Registro habrán de tener una tipología análoga a la estructura natural de estos flujos y sub-flujos, estructura que se representan en la figura 32.

En muchos casos, puede interesar el escindir el Registro en dos o más Registros, o lo que es lo mismo, dividir un Fichero en varios Ficheros. Respecto a esta posibilidad, la regla lógica a seguir es el ajustarse al objetivo del rendimiento óptimo de la Base de Datos, lo que quiere decir en definitiva que la flexibilidad del modelo, habrá de ser contemplada con prioridad. Si con la escisión del Registro Tipo en dos o más Registros, se consigue una economía de almacenamiento, o unas técnicas de acceso más eficaces y además no se perjudica esencialmente a ninguna otra característica, el Registro ha sido correctamente escindido.

También hay que tener en cuenta como regla, que el Registro es una unidad lógica básica y que tal unidad tendrá sentido el diseñarla, cuando los datos integrantes del Registro vayan a ser necesitados en conjunto o al menos el hecho de manejarlos en conjunto, tenga una justificación lógica.

Tengamos en cuenta que en cuanto menos Ficheros tenga la Base de Datos y en cuanto más datos contenga cada Registro, la Base de Datos es más rígida. Por el contrario, a mayor número de Ficheros y menor cantidad de datos contenidos en cada Registro, la Base de Datos es más flexible pero en contrapartida, requiere mayor capacidad de almacenamiento. Como siempre, podríamos decir que en el punto medio está la “virtud” y el dar con la posición de equilibrio, es algo que conseguiremos con el sentido profesional que da la práctica.

Lo que sí debe tener cada Registro, es decir, cada Fichero, es:

- Una entidad que viene a constituir el sujeto informático del Registro.

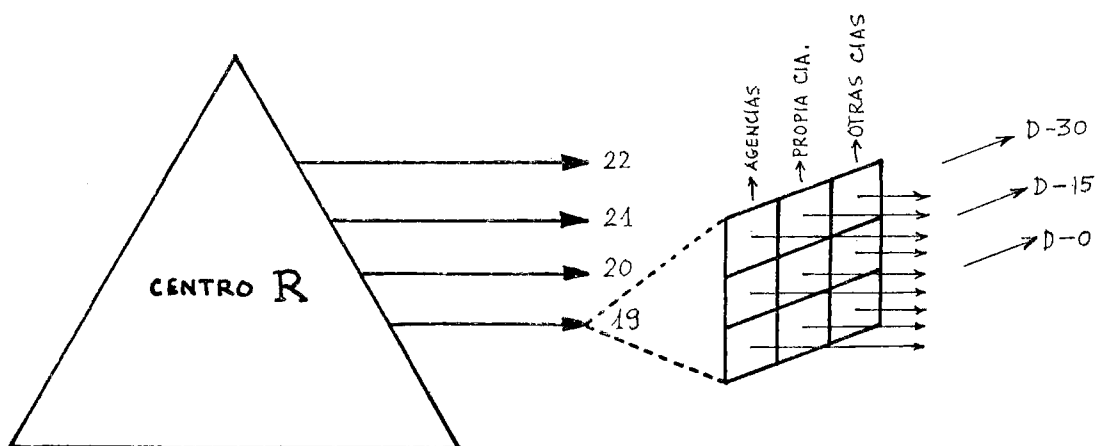


FIG. 32

- Un objetivo o razón de ser que habrá de responder a utilidades del Sistema de Base de Datos.
- Lo demás, son atributos, sus medidas, referencias cruzadas.
- Etcétera.

En cuanto a la estructura interna de los Registros, es decir, la determinación de hasta qué profundidad se ha de llegar en la segregación de Flujos, y por lo tanto hasta qué grado de detalle descenderán los Registros, es más bien una consecuencia de la “potencia” informática que pretendamos para nuestra Base de Datos, lo que equivale a tener siempre presente como elemento director básico el conjunto de objetivos del proyecto. Por un lado, el exceso de detalles encarece la Base de Datos porque es mayor el almacenamiento necesario, y en esto influye mucho la segregación de Ficheros que hayamos dispuesto. Por otra parte, el

sintetizar la información en exceso tiene como consecuencia el hecho de que las síntesis son irreversibles y aquellos atributos cuya información no haya sido segregada, son aspectos informáticos que se pierden para el futuro. El punto de equilibrio es nuevamente algo a buscar.

2.3.3. RELACION ENTRE SUCESOS Y RELACIONES LOGICAS ENTRE REGISTROS

Los registros, una vez configurados como unidades informáticas, contendrán información de diferentes naturalezas:

- Información de estructura de la Entidad o Entidades a las cuales dicho registro está asociado.
- Coordenadas espacio-temporales de producción de Sucesos.
- Cuentas de flujos derivados de los Sucesos.
- Características y atributos de los Sucesos.

La información de naturaleza estructural está relacionada en la red de producción en base a la Jerarquía orgánica real en la cual están integradas las Entidades que representan. Al examinar la Red de Administración nos ocuparemos de estas relaciones.

La información de Coyuntura, o información ligada a los Sucesos, constituye el conjunto de información de las relaciones establecidas en la Red de Producción, ya que el entramado de la Red de Producción está constituido como consecuencia de los Sucesos.

Por otra parte, la Red de Producción es simplemente un instrumento analógico que vamos a utilizar como marco de referencia para el estudio de la relación entre los Datos y la Actividad, y como tal, no tendrá después un protagonismo físico en la dinámica de explotación del sistema de Base de Datos. Sin embargo, ya hemos indicado que el proceso continuo de la Base de Datos viene a suponer una permanente reconstrucción de la Actividad y la estructura de datos que manejamos ha de ser una modelación de dicha actividad. Por ello la estructura de datos tendrá necesariamente que “representar” a la Red de Producción que así de esta forma, seguirá siendo el soporte básico del sistema, aun cuando no tenga una realidad física.

La Red de Producción nos proporciona sugerencias directas de cómo configurar la estructura de datos, tanto en lo relativo a la estructura interna de los registros como en lo que se refiere a las relaciones lógicas que entre los registros va a convenirnos establecer.

La estructura interna de los registros, que es la estructura de los ficheros representativos de los Centros de Actividad, corresponderá de una forma dual a:

- La estructura de los flujos derivados de los Sucesos.
- La composición de flujos afluyentes a los Centros de Actividad.

Y comportará un conocimiento preciso de la relación que el Centro de Actividad (que el registro representa) mantiene con la red de Producción.

Adicionalmente vamos a relacionar unos registros con otros, utilizando para ello campos de datos contenidos en los propios registros, campos que determinan las relaciones lógicas que compondrán la estructura de datos del sistema. Son los vínculos de unión que integran los diferentes ficheros en una misma estructura y constituyen el “esqueleto” del modelo de actividad, esqueleto sobre el que la dinámica de Sucesos conformará un órgano “vivo” que es la Base de Datos, órgano que tendrá una vida paralela a la propia actividad de la Organización.

Conviene detenernos a considerar la posible evolución que el sistema de Base de Datos, como órgano vivo, va a experimentar en el futuro:

- La estructura de datos será en general poco cambiante, porque la red de producción tiene una tipología más o menos determinada por el tipo de actividad de la Empresa u Organismo. No obstante, podrá evolucionar enriqueciéndose el modelo, incluyendo nuevas relaciones lógicas.
- La estructura habrá de ser ampliable y de hecho se extenderá en la medida en que precisemos abarcar nuevas actividades de la Empresa u Organismo en el sistema informático.
- El propio “esqueleto” de la Base de Datos que está constituido por todo el conjunto de registros tipo y de sus estructuras lógicas puede también ser variado, porque en base a la vocación morfológica que el sistema debe perseguir, sus estructuras más básicas deben ser también perfeccionables. Sin embargo ello habrá de constituir muy contadas excepciones y responderá a motivaciones muy precisas. En un animal vivo pueden generarse estímulos de osificación que procedan a alterar en alguna medida su estructura ósea base de su organismo físico-biológico, aunque ello obedecerá a motivos muy especiales, y si ello sucede de forma drástica, estaremos ante una mutación. En nuestro caso de sistema de base de datos la alteración drástica de la estructura lógica de los datos, será índice de dos causas posibles:
 - Cambio total de la Actividad empresarial.
 - Diseño deficiente de la estructura.

De esta segunda causa, el Informático debe situarse siempre a salvo.

Las relaciones lógicas entre registros son creaciones que obtenemos de un estudio razonado de la actividad y de la red de producción. De este conocimiento se nos sugieren los conexiones que habremos de disponer en el modelo para obtener un conjunto operativo que cubra los objetivos informáticos. Por ser un producto lógico, es la lógica la regla maestra a seguir para completar la estructura, sin

embargo podemos citar a título de guía los siguientes criterios que pueden resultar de ayuda.

- Contemplar todas las posibles relaciones, es decir, tantear todos los posibles lazos lógicos entre ficheros y razonar acerca de ellos enjuiciando su relevancia. Percibiremos claramente las relaciones lógicas que dan consistencia a nuestro modelo.
- Enjuiciar la cantidad de relaciones que hemos establecido. Una multiplicidad de relaciones lógicas puede ser índice de una defectuosa definición de los registros y pudiera sugerirse una mayor compactación de los datos. Por el contrario, una escasez de relaciones nos determina un modelo rígido y operativamente pobre.

Una reflexión sobre la matriz asociada al grafo que representa la red lógica, nos muestra el equilibrio entre las relaciones. En ella vemos en forma abstracta las relaciones de dependencia lógica de los registros y podremos comprobar:

- Registros muy relevantes no deben ser excesivamente dependientes.
- La consistencia entre las excesivas autoridades de registros cabeceras de estructura y la posterior dinámica de actualización. Tengamos en cuenta que en cuanto más “raíz” es un registro tipo, es decir, en cuanto dependen de él mayor número de registros tipo, más condiciona al conjunto tanto en los aspectos positivos como en los negativos.
- Los registros estructurales deben estar situados en las “raíces” o próximos a ellas.

También tendremos que tener presente la dinámica de proceso operativo posterior del sistema, para comprobar la estabilidad del conjunto, pero para estos fines ya enumeraremos los procedimientos adecuados al exponer la metodología de diseño de la Base de Datos.

Como en toda actividad creativa, el planteamiento de la estructura lógica más conveniente es un proceso inductivo de síntesis y hasta tanto no tengamos conclusiones sólidas debemos mantener una actitud de revisión, sometiendo siempre a nuevas consideraciones la unidad de ficheros, es decir, la composición de los registros tipo. En realidad los registros tipo y las relaciones lógicas entre ellos son aspectos duales de un mismo conjunto y como tales se complementan y siempre es obligado manejarlos en paralelo a un tiempo.

Veamos el ejemplo de la figura 33, en donde se representa un registro específico que pertenece al “fichero de vuelos programados”. El registro tipo definidor de la estructura interna del fichero contiene los siguientes campos de información:

- N.º de línea.
- Trayecto.
- Fecha de operación.
- N.º de pasajeros.
- Subcampos que relacionan los mercados de operación.

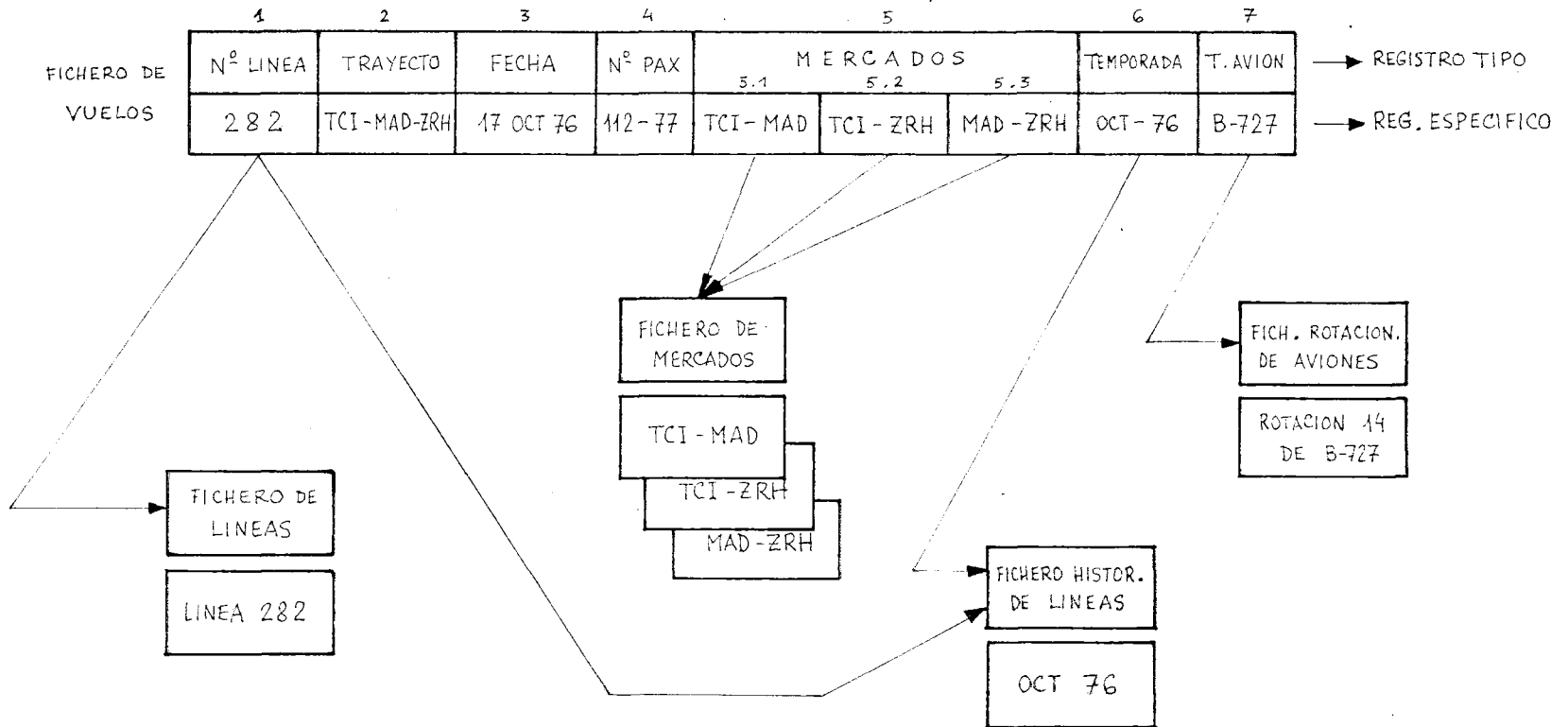


FIG. 33

- Temporada.
- Tipo de avión.

Las relaciones lógicas con otros ficheros se han representado en la figura con fechas, sin que se pretenda de momento dar un significado al sentido de la flecha. Las conexiones lógicas establecidas responden a las siguientes conveniencias:

- Conectar el fichero de líneas con el fichero de vuelos programados en base al código de la línea. De esta forma tendremos agrupados en una misma familia lógica todos los vuelos que en las diferentes fechas se encuentran programados o han sido ya realizados y pertenecen a una misma línea o ruta. La relación lógica que une ambos ficheros se deriva directamente de la jerarquía de flujos de la red de producción y ciertamente se comprende que la conexión es relevante.
- La relación entre el fichero de vuelos y el fichero de mercados es un ejemplo de la unión de flujos de productos a soportes estructurales como son los mercados que dan base de referencia al conjunto de vuelos. Cualquier otro fichero que se relacione con el fichero de vuelos, podrá ser referido a la estructura de mercados.
- La conexión con el fichero de rotación de aviones es una relación lógica de la misma entidad que la citada en primer lugar, ya que corresponde a la jerarquía de flujos de la red de producción.
- La relación lógica entre el fichero de vuelos y el fichero histórico de líneas es muestra de enlace con registros históricos de síntesis o de estado. Los datos agrupados bajo la referencia histórica pueden ser objeto de procesos sintetizadores en la Base de Datos, que presenten la trayectoria histórica operativa necesaria eliminando los detalles analíticos que el transcurso del tiempo convierta en obsoletos.

La potencia operativa del sistema de Base de Datos está fuertemente condicionada por la riqueza lógica de ese "esqueleto" que forman la definición de los registros tipo y las relaciones entre ellos existentes.

2.3.4. CONSTITUCION DE LA RED DE PRODUCCION

La Red de Producción va a constituir el Marco de Referencia para todo el desarrollo del proyecto de Base de Datos y va a ser por lo tanto a partir de su creación el elemento número uno a consultar durante la etapa de diseño e incluso ante planteamientos futuros que impliquen una revisión de la estructura de datos o simplemente de una ampliación de la misma. Será también un elemento de ayuda en los problemas que plantea la explotación de la Base de Datos o el desarrollo de aplicaciones que la utilizan.

Los componentes básicos de la Red de Producción son:

- Los Centros de Actividad.
- Los Sucesos.
- Los flujos de productos.

Los Centros de Actividad hemos visto constituyen Unidades individuales y como tales pueden ser analizadas y estudiadas. Sin embargo, resulta necesario el análisis de la jerarquía natural existente entre los Centros de Actividad para poder considerar las relaciones que entre ellos existen, ya que además del estudio de los Centros de Actividad en forma individual es preciso estudiar también la interrelación que cada Centro tiene con los demás, para poder seleccionar aquellas interrelaciones que estimemos sean relevantes para el Sistema.

La dinámica de la actividad, y por lo tanto la estructura de los Sucesos, son los elementos que van a completar a la jerarquía natural de los Centros de Actividad para aportarnos el conexionado de la Red de Producción. La estructura de los Sucesos nos indica los hechos reales relevantes que se producen en los Centros de Actividad, y por tanto constituyen las piezas básicas para que la Red de Producción sea un conjunto conexo. Los Sucesos son, por lo tanto, ingredientes básicos con los que vamos a construir la Red, y las características de dichos Sucesos habrán sido contempladas a la hora de estudiar la actividad de los Sistemas Productivos.

Así, la constitución de la Red de Producción requiere que los Centros de Actividad sean analizados atendiendo a las dos aproximaciones duales: estructural y coyuntural, o lo que es lo mismo, al análisis de la composición orgánica interna de los Centros de Actividad y al análisis de la influencia de cada Centro ante el entorno.

La Red de Producción ha de utilizarse primordialmente para diseñar la lógica de los registros de datos y sus interrelaciones. El examen de la Red de Producción ha de aportarnos un conocimiento suficiente acerca de la tipología de la información que generan los Centros de Actividad, lo que permite al analista el enjuiciar el conjunto de datos que son de interés para el panorama de objetivos del proyecto. Es decir, es la Red de Producción la que aporta al analista la necesaria visión de conjunto para poder extraer la información relevante para el Sistema mediante el diseño de los registros de datos.

Este objetivo principal que pedimos a la Red de Producción de facilitarnos el conocimiento de la información generada, su naturaleza, su relevancia y sus componentes, obliga por otra parte a contemplar la Red de Producción a diferentes niveles de profundidad o de detalle, dependiendo de los Centros de Actividad examinados. Así, por ejemplo, en la figura 30 se representa una Red de producción que incluye al Centro R como uno de los Sistemas productivos participantes en la Red. Sin embargo, si en la tarea del análisis fuera preciso un mayor nivel de profundidad a la hora de examinar la información que llega al Centro R , la información que el Centro R genera, o bien la información interna que la actividad del Centro R maneja, es preciso profundizar más en el conocimiento de esa zona de la Red de Producción, análisis del que es ejemplo la figura 31, sin que por ello tengamos que abandonar la visión de conjunto que la Red de la figura 30 nos ofrece.

La Red de Producción es por lo tanto flexible, y se puede descomponer en sub-redes, en la misma forma que un sistema puede descomponerse en sub-sistemas. Las sub-redes podrán ser estudiadas a diferentes niveles de profundidad, conservando en todo momento la propiedad asociativa de poder ser contemplado su conjunto como una sola Red. Asimismo puede ocurrir que una determinada sub-red sea preciso descomponerla a su vez en redes más elementales, lo que es posible sin perder generalidad el conjunto, ya que las propiedades asociativa y distributiva de los Sistemas pueden ser aplicadas en forma sucesiva.

La regla lógica a seguir para la constitución de la Red de Producción, consiste en llegar a profundizar en el análisis en cada zona de la Red, hasta un nivel que resulte suficiente para la determinación de la lógica de los registros de datos, es decir, del diseño de los registros y del conocimiento de sus interrelaciones.

Merece también tener presente que, dado el carácter de modelo estructural de la Red de Producción, puede considerarse ésta como un marco adecuado para la Codificación de los datos. La Codificación, cuya importancia ya hemos resaltado, es una componente de estructura básica para el sistema de Base de Datos y debe de establecerse como una tarea adicional al diseño de los registros y sus relaciones, tarea que puede realizarse en paralelo pero que resulta de interés proceda del mismo proceso de creación.

2.4. LA RED DE ADMINISTRACION

En todos los Sistemas Productivos, configurados según el modelo que hemos presentado, tiene lugar un fenómeno común: el desarrollo de Actividad. Este ejercicio de la Actividad productiva tiene a su vez una característica invariable, cual es la utilización de Recursos cuya concurrencia hace posible el fenómeno productivo.

Estos Recursos pueden ser materiales producidos por otro Centro de Actividad de la Organización, en cuyo caso generalmente han sido tenidos en cuenta en la Red de Producción, porque en dicha Red estarán reflejados como productos del Centro de Actividad proveedor.

Otro tipo de Recursos puede consistir en la concurrencia de materiales de consumo o materias primas procedentes del Entorno exterior a la Organización, en cuyo caso podremos considerarlos indistintamente como aportaciones a la Red de Administración o aportaciones a la Red de Producción, conviniendo quizás conceptualmente más lo segundo, pero en ocasiones más lo primero desde el punto de vista administrativo.

Además de este tipo de Recursos que pudiéramos englobarlos en un concepto de "materias primas", existen otras tres naturalezas de Recursos básicos para el ejercicio de la actividad:

- El Equipo instrumental.
- El Personal.
- El Dinero.

El Equipo instrumental está formado por el conjunto de máquinas, herramientas, enseres, mobiliario, etc. que forman parte del Activo del Sistema productivo. Dicho Activo es dinámico puesto que evoluciona, y su identificación y valoración está generalmente recogida en un inventario. Si los elementos constituyentes de este conjunto han sido adquiridos en propiedad y pertenecen por lo tanto a la Organización, existirán unos determinados esquemas de amortización a considerar. Si por el contrario, el conjunto o alguna de sus partes han sido alquilados, no interesará considerar tanto su valor monetario o su período de vida como los conceptos de coste en los períodos de alquiler. Este tipo de Recursos está prácticamente siempre presente en el fenómeno productivo.

La plantilla de Personal del Centro de Actividad constituye el conjunto de recursos humanos que ejercen la actividad. Las personas pueden estar integradas en la Organización o bien pueden haber sido contratadas temporalmente o su concurrencia puede estar originada en base a un contrato de prestación de servicios suscrito con otras Entidades. En uno u otro caso, al igual que en el caso del Equipo, los aspectos relevantes a considerar pueden ser diferentes. Las personas siempre están igualmente presentes en las tareas de producción, aun cuando a un nivel profundo el análisis pudiera llegarse a Sistemas Productivos muy elementales que no requirieran de una intervención continua de personas por estar automatizados, si bien a dicho grado de detalle, generalmente no llegaremos, al menos en el desarrollo de Sistemas informáticos de gestión.

El Dinero es, obviamente, un recurso básico indispensable que siempre va a estar presente, porque sin su concurrencia la actividad productiva no es posible. El método normal de canalizar la aportación de recurso financiero a los Centros de Actividad es en base a las cuentas presupuestarias, que en conjunto constituirán el presupuesto de la Organización.

Cabría referirnos a otro recurso básico también presente siempre, que es el recurso Información, pero por pertenecer a otra dimensión en el modelo que estamos perfilando será considerado por separado en el punto 2.5.

La Red de Administración la presentamos como elemento integrador de todos los recursos de la Organización y de su evolución y circunstancias a lo largo del tiempo, lo que en definitiva nos permitirá plantear, bajo un concepto de sistema, una Base de Datos para la gestión de personal, o la contabilidad por citar las dos aplicaciones más generalizadas.

Como se da la circunstancia de que en cada Centro de Actividad es necesaria la aplicación de recursos, al igual que es necesaria la existencia de sucesos productivos, siempre se producirá en cada Centro de Actividad la intersección de la Red de Producción con la Red de Administración (fig. 34), por ello ambas redes serán contempladas como ortogonales, es decir, se examinarán por separado en dos pla-

nos diferentes, pero siempre tendrán ambas redes los mismos nodos, que son los Centros de Actividad en estudio.

Al igual que en el caso de la Red de Producción, se pueden examinar los flujos de recursos bajo diferentes grados de síntesis o análisis, puesto que aplicamos las mismas reglas asociativas y distributivas válidas para los sistemas productivos. Generalmente, las necesidades de profundizar a niveles de detalle o bien sintetizar a niveles de generalidad, serán las mismas para la Red de Administración que para la Red de Producción. Sin embargo, en caso de que las necesidades sean diferentes,

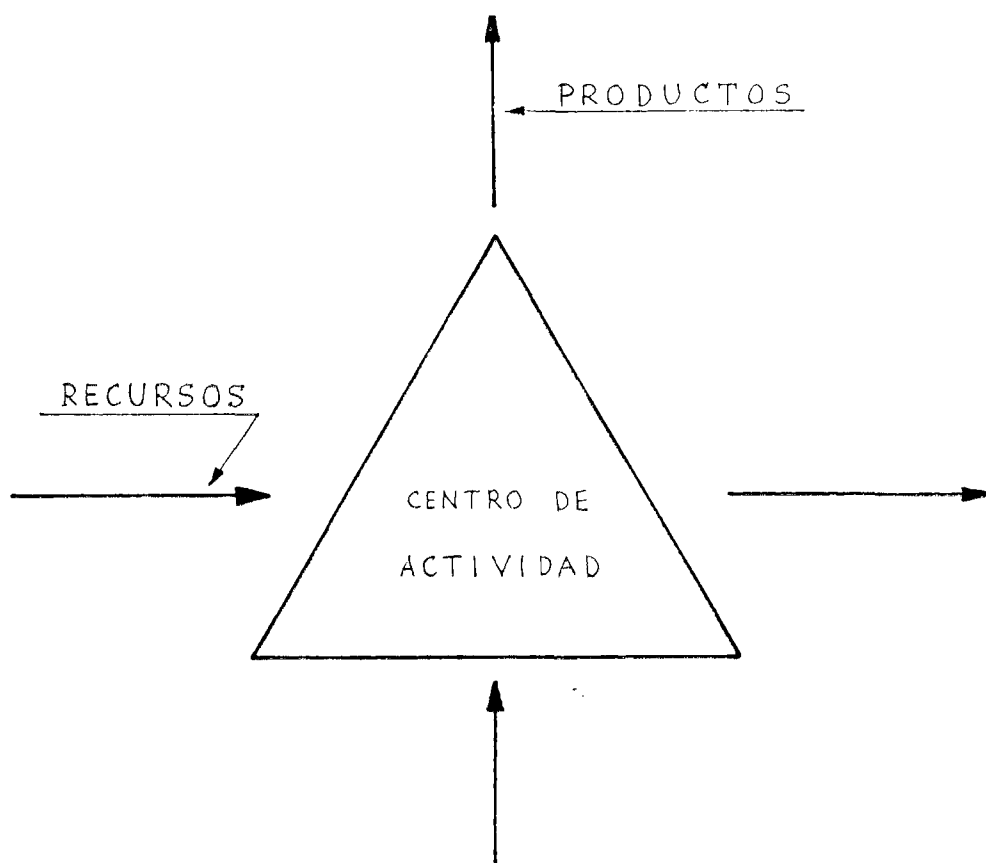


FIG. 34

ello podrá obedecer a que a efectos de administración de recursos hemos fraccionado el Centro de Actividad en diferentes secciones, y ello podrá ser realizado penetrando dentro de la "caja negra" que representa dicho Centro, sin que necesariamente hayamos de tener en cuenta los intercambios internos de productos entre las partes del Centro. Ello quiere decir que en caso de requerirse diferentes niveles de análisis será en la Red de Administración donde se alcance mayor profundidad porque a la inversa no tendrá sentido el que consideremos unos flujos de productos entre Centros de Actividad sin atender a los recursos que hacen posible dichas actividades parciales.

No obstante, el estudio de la Red de Administración, que es objeto de análisis

necesario para los sistemas de gestión, pudiera no llevarse a cabo cuando la Base de Datos que pretendemos diseñar tiene objetivos exclusivos de orden operativo sin pretender abarcar el control de los recursos.

Presentamos la Red de Administración como una zona más de nuestro modelo de empresa, para abarcar el caso más general de que el sistema de Base de Datos a desarrollar constituya un sistema integrado total. De esta forma nuestro modelo es de uso universal, porque aplicando la óptica de sistemas bajo la que lo hemos construido, es igualmente válido para utilización parcial o, lo que es lo mismo, para el planteamiento de una Base de Datos para una aplicación específica y acotada dentro de la Organización.

2.4.1. FLUJOS DE RECURSOS

En la Red de Producción los flujos de productos emanaban de unas Entidades que hemos denominado Sucesos. Al producirse los Sucesos se generaban unos productos que abandonaban el Centro de Actividad, entrando en juego unas coordenadas geográfico-temporales. En la Red de Administración, los flujos de recursos emanan de las Entidades que forman el Activo de cada Centro de Actividad. El flujo de recursos no tiene en este caso una realidad material salvo en los casos de incorporación de nuevos recursos a los Centro de Actividad, en los que se produce una “entrada” del entorno al Centro de Actividad. Quiere decirse que, generalmente, los recursos están encuadrados en el interior del Centro de Actividad y que el concepto de flujo ha de ser contemplado en abstracto como si la Red de Administración aportara en cada momento dichos recursos para el ejercicio productivo. Esta abstracción tiene por objeto el análisis del factor dinámico, es decir, la consideración de las circunstancias temporales por las que en cada momento atraviesan los elementos constitutivos del Activo de cada Centro Productivo.

En la figura 35 se presenta un esquema que puede utilizarse como ideograma de referencia a la hora de plantearnos la Red de Administración.

En dicha figura se presentan las diferentes Entidades que por su relevancia pueden generalmente ser incluidas en nuestra hipotética Base de Datos. De estas Entidades se hace una separación entre las referidas a la propia personalidad de los Recursos y las referidas a la aplicación de dichos Recursos. Entre las primeras se consideran los elementos del Activo propiamente dichos, contemplando no sólo el Equipo inventariable, sino las personas y las partidas presupuestarias o consideraciones monetarias e incluso el material de consumo como materias primas. Como Entidades referidas a la aplicación de estos recursos presentamos las funciones, standards de ocupación de recursos, standards de producción, Sucesos y Productos, estos últimos con un entronque con el conjunto de datos referentes a la Red de Producción, porque el Suceso tiene lugar en el umbral en el que los productos pasan a ser elementos constituyentes del Activo del Centro de Actividad, a ser elementos de aporte a la Red de Producción.

Merece destacarse como Entidades muy útiles las Funciones, cuyo conjunto viene a constituir la actividad global y cuyas individualidades vienen a representar di-

CENTRO DE ACTIVIDAD		RECURSOS	ACTIVIDAD
ENTIDADES		<p>Equipo inventariable</p> <p>Material de Consumo</p> <p>Personas</p> <p>Presupuesto</p>	<p>Funciones</p> <p>Standards de ocupación de recursos</p> <p>Standards de Producción</p> <p>Sucesos</p> <p>Productos</p>
ATRIBUTOS FISICOS		<p>Nombre y código</p> <p>Cantidad potencial</p> <p>Cantidad ocupada</p> <p>Atributos específicos del personal</p>	<p>Medida de la producción</p> <p>Desviación respecto standards</p> <p>Nombre y código de funciones, sucesos y productos</p> <p>Recursos asignados a funciones</p>
ATRIBUTOS FINANCIEROS		<p>Valores de inventario</p> <p>Nómina</p> <p>Esquemas de amortización</p> <p>Costes inertes</p> <p>Costes de ocupación e inocupación</p>	<p>Costes de producción por funciones</p> <p>Valor de los productos</p> <p>Costes de materiales</p> <p>Amortizaciones</p>
SERIES TEMPORALES	HISTORIA	<p>Síntesis por períodos de todo el conjunto histórico.</p> <p>Datos analíticos relevantes de la historia más reciente.</p> <p>Datos completos del ejercicio económico en vigor ya realizado.</p>	
	PRESENTE	<p>Medidas físicas y monetarias de las Entidades que componen los Recursos y la Actividad, así como las de sus atributos</p>	
	FUTURO	<p>Previsiones, Presupuestos, Objetivos de Producción</p>	

FIG. 35

ferentes sectores de la actividad. La descomposición de la actividad en diferentes funciones habrá de ser manejada en forma análoga a la descomposición de un Centro de Actividad en diferentes Centros, es decir, con las reglas de los Sistemas, aplicando las propiedades asociativa y distributiva en la forma que venimos presentando.

Estas Entidades, tanto las referentes a los recursos propiamente dichos como las relativas a su aplicación, o lo que es lo mismo, a la actividad, habrán de ser examinadas mediante unidades de cuenta tanto en sus aspectos físicos como en sus aspectos financieros, y ello en una perspectiva de tiempo que recoja la historia, el presente y el futuro, de manera que la Base de Datos resultante sea un soporte adecuado e incluya los necesarios factores de dinamicidad.

Entre los atributos físicos de los recursos, y además de todos los relativos al personal, que no vamos a enumerar, destacamos la medida física de la cantidad de recurso potencialmente disponible, lo que representa su "techo", así como el valor físico de la cantidad ocupada en cada momento. El recoger estos datos permitirá analizar en unidades físicas la evolución de la ocupación de los recursos, y ello en cada Centro de Actividad a título individual y en los diferentes conjuntos, englobando los componentes de una misma naturaleza de recurso aplicados en diferentes Centros de Actividad.

Entre los atributos físicos de la Actividad, resaltamos la medida de la producción, las desviaciones respecto de standards, etc., lo que recogerá las necesarias unidades de cuenta que en cada caso consideremos son convenientes a nuestros objetivos.

Como atributos financieros, manejaremos las partidas de coste, los valores monetarios de los recursos, los esquemas de amortizaciones, el valor de los productos, la nómina, etc.

Todo este panorama de datos que presentamos en la figura 35, aunque no de una forma exhaustiva, habrá de ser tenido en cuenta a la hora de diseñar los registros de nuestra Base de Datos en dos diferentes circunstancias:

- Para repasar a título de "lista de comprobación", la tipología de datos que habremos de incluir en registros ya diseñados, esto es, para complementar nuestra base informática ya emanada de la Red de Producción.
- Para diseñar nuevos registros o una nueva sección de la Base de Datos que atienda a aplicaciones específicas de administración, sección de la Base de Datos que naturalmente estará integrada en el resto aun cuando su proceso pudiera ser individual.

Las series temporales de datos habrán de contemplar una perspectiva histórica, completa en detalle de datos para la historia inmediata, y sintetizando y consolidando datos para tiempos más pretéritos. Asimismo podrán contener datos previstos para el futuro referentes a presupuestos, objetivos de previsión, etc., que sirvan de soporte para contraste posterior con las realizaciones realmente ocurridas.

2.4.2. INVENTARIO Y ADMINISTRACION DE RECURSOS

El propósito básico de la Red de Administración como instrumento de análisis, es el planteamiento de una Base de Datos o de una sección de ella que constituya el soporte de proceso de las aplicaciones de naturaleza administrativa de la Empresa u Organismo.

Esta Base de Datos o "sistema" de ficheros interrelacionados constituirá por lo tanto la base jurídica donde las aplicaciones habrán de encontrar todas las circunstancias que han tenido lugar en la vida real de la Organización y que constituyen materia susceptible de administración. La orientación de la Base de Datos es la de "registro notarial", donde se han de registrar los hechos físicos que se producen y que afectan a las Entidades que hemos considerado en los flujos de recursos, valorando sus atributos en medidas físicas y monetarias.

Especial importancia reviste el factor tiempo en las tareas administrativas, por lo que el grado de actualidad de la información es necesario que en primer lugar sea homogéneo para las diversas Entidades y Centros de Actividad y en segundo lugar sea suficiente para el plazo de demora que en términos de proyecto determinamos como admisible. La Captación de datos es en estos casos de importancia decisiva.

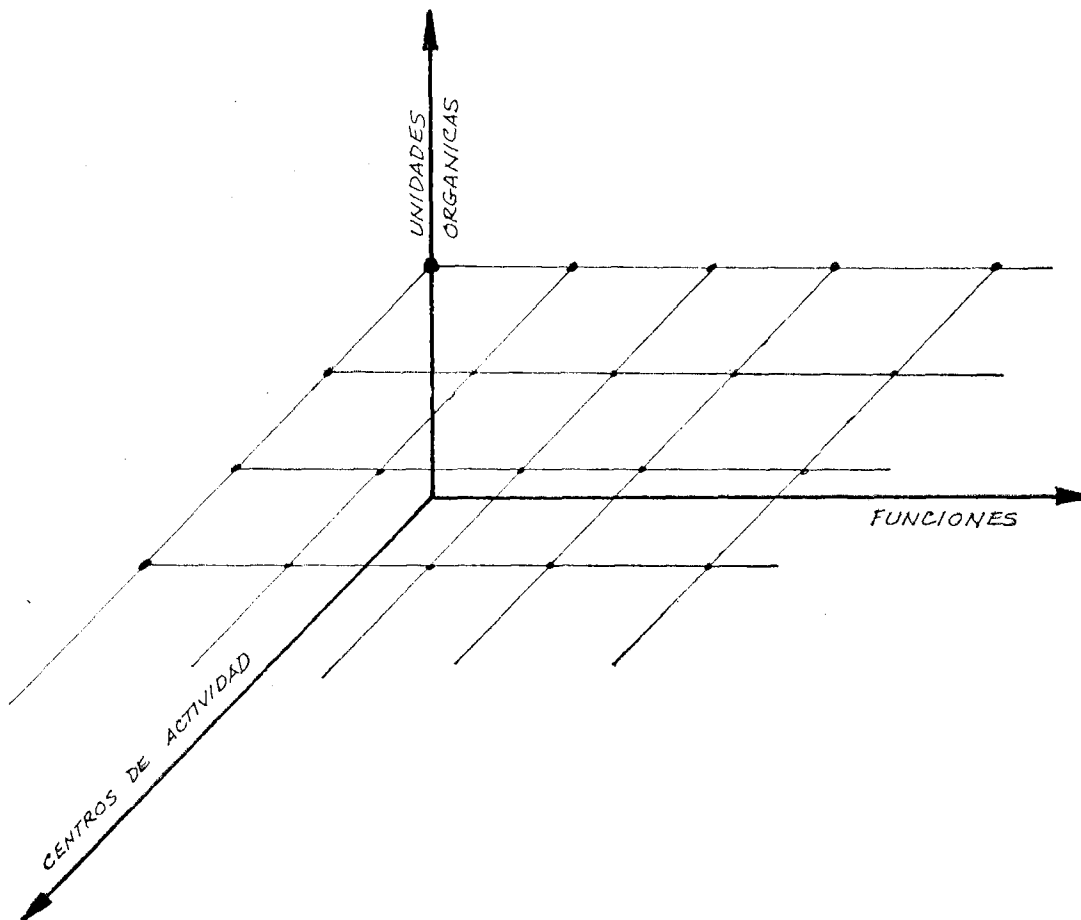


FIG. 36

Para plantear las aplicaciones de Gestión y Administración de Recursos, es también necesario disponer en la Base de Datos de la información de estructura de la Organización, para que en los procesos puedan agruparse las distintas Entidades, Funciones y Centros de Actividad, atendiendo en cada caso al nivel de síntesis o análisis que pueda interesar, sin que por ello la Base de Datos tenga que preverlos. La figura 36 expresa el tipo de matriz de datos que habrá de disponer, agrupando los Centros de Actividad constituyentes de cada Unión Orgánica y recogiendo las funciones en las que clasificamos la actividad de cada Centro. Esta matriz de datos puede también estar implícita en las relaciones lógicas que determinamos entre los registros que contienen datos de las Entidades.

La codificación de los datos tiene mucha importancia sea realizada con la visión global de la Red, ya que supone una estructura general de la Organización y como tal tiene que contar con el grado adecuado de flexibilidad y operatividad que no comprometa la evolución de la Base de Datos.

La Codificación de las Entidades y Centros de Actividad habrá de ser realizada por tanto en compatibilidad con la Red de Administración, manejándose a la vista de ella los criterios que establecíamos en el capítulo 1.5 de forma que haya una relación armónica entre la estabilidad de las estructuras de la Red y la significación de los códigos.

De entre los tipos de recursos a considerar, uno de ellos, que es el Personal, habrá de ser objeto de estudio adicional en otro marco más amplio que el Modelo de Empresa, ya que la gestión de personal tiene connotaciones de orden jurídico-social. Por ello, al analizar la sección de la Base de Datos que cubra los datos del personal de la Organización, habremos de tener presentes las necesidades informáticas de este tipo de Aplicaciones, en las que no vamos a entrar por estar suficientemente estudiadas en trabajos monográficos.

2.4.3. BASES PARA UNA CONTABILIDAD ANALÍTICA

Las Aplicaciones informáticas de gestión económico-financiera son quizás las de mayor trascendencia para la Empresa y las que en primer lugar justifican generalmente la incorporación de un ordenador electrónico. La Contabilidad es de entre estas aplicaciones, la más importante, porque supone el usual procedimiento de control del recurso Dinero.

En un sistema de Base de Datos de apoyo a la gestión económico-financiera, la contabilidad puede ser contemplada como el estudio del “metabolismo del dinero” en ese organismo dinámico y “vivo” que es la Empresa.

Una adecuada coordinación entre los especialistas en organización, en temas económicos y en informática, puede conseguir la adecuada tipificación de las Entidades que para cada Centro de Actividad vayamos a considerar. Recordemos que el grado de análisis que con las aplicaciones informáticas podremos cubrir, viene determinado por la segregación de Entidades y sus atributos que en la Base de Datos hayamos dispuesto.

El concepto de Base de Datos permite construir un sistema de contabilidad analítica de gestión que con las aplicaciones convencionales resultaba poco efectiva. Los datos elementales no están simplemente recogidos en los registros, sino, además, oportunamente relacionados, por lo que puede operarse a diferentes niveles de síntesis-análisis tanto para la elaboración de informes muy detallados para los especialistas como de resúmenes comparativos de diagnóstico para la Dirección.

Recogiendo datos elementales para cada Centro de Actividad tales como: costes inertes, costes de ocupación e inocupación de recursos, costes de producción por funciones, esquemas de amortización, costes de materiales, etc., que se presentan esquemáticamente en la figura 35, habremos planteado una Contabilidad Analítica en potencia, Contabilidad que después será tan útil como hábiles hayamos sido al proyectar la correspondiente aplicación. La Base de Datos es aquí, una vez más, la infraestructura.

Los modernos planteamientos de Contabilidad Matricial parten, en definitiva, de un concepto de infraestructura contable que podemos aplicar con un buen diseño de una Base de Datos. Esta estructura de base permite desarrollar las aplicaciones informáticas, enriqueciéndose con el tiempo mediante tablas "input-output" que ofrezcan diagnósticos previsorios de resultados ante variaciones porcentuales de las medidas físicas y monetarias de los atributos de las Entidades.

Sin penetrar más en este área, sugerimos al lector tenga presente el concepto de explotación en Base de Datos cuando tenga que enfrentarse con una aplicación informática de contabilidad de gestión.

2.5. LA RED DE INFORMACION

Los Centros de Actividad utilizan otro recurso básico que es la Información y que tiene una naturaleza diferente a los demás. Si bien nuestra tarea de analistas de sistemas de información nos lleva a plantear la Red de Producción y la Red de Administración como elementos de estudio con el fin de diseñar un sistema de información, nuestro producto final, la Base de Datos, supondrá la mecanización de ciertos procesos de información, subsistiendo muchos otros flujos de información ajenos al Centro de Cálculo, pero igualmente relevantes para el ejercicio de la actividad e igualmente necesarios de analizar, para el informático.

El técnico en informática, en tanto en cuanto deja de realizar simples aplicaciones de proceso de datos, y comienza a proyectar sistemas integrados, tiene que contemplar los circuitos de información con independencia de si pertenecen a la zona mecanizada o bien son procesos manuales. A priori, el sistema de información de la Organización puede considerarse como un Todo. El hecho de que unas partes del conjunto, es decir, algunos subsistemas, se vayan mecanizando es realmente un cambio de procedimientos y una modificación del sistema, pero no una ruptura del Conjunto. Además el progreso que pueda suponer la mecanización y el desarro-

llo de ésta hay que enmarcarlo en el sistema total para poder evaluarlo y dirigirlo convenientemente.

Por eso vamos a incluir en nuestro Sistema de Empresa, es decir, en el modelo que estamos perfilando, la Red de Información. Como la Información la hemos distinguido de los demás Recursos y como también la aislamos de la Red de Producción, aunque en ésta se genere realmente dicha información, la Red de Información vamos a considerarla como un conjunto de extracciones de flujo que tienen lugar en los Centros de Actividad, y que van a alimentar a otros Centros de Actividad. Este flujo de información llega a todos los Centros de Actividad, y en ellos se produce la intersección con los otros flujos de Recursos y de Productos, por lo que consideraremos las tres naturalezas de flujos apuntadas como flujos ortogonales, según se representa en la figura 37.

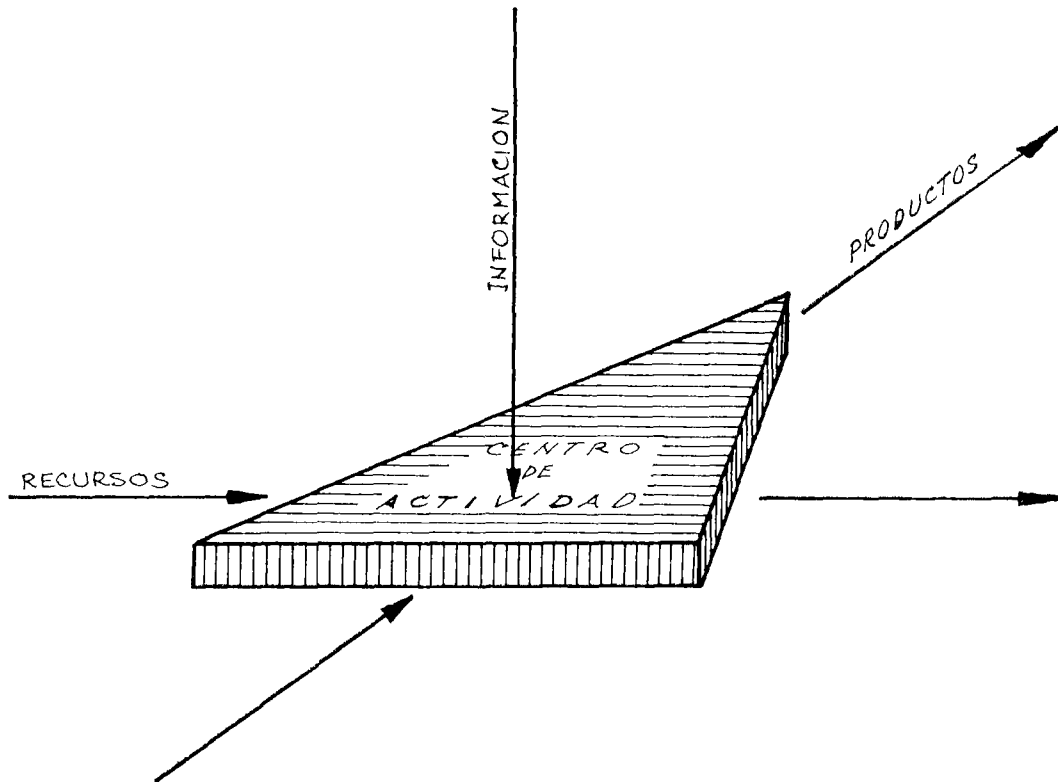


FIG. 37

Los flujos de información nacen de los Centros de Actividad y a su vez se extinguen en ellos, por lo que la información puede ser considerada como un producto para los Sistemas productivos emisores y también como un recurso para los Centros receptores.

La información que un determinado Centro de Actividad recibe y emite, considerando los dos aspectos, constituye lo que normalmente llamamos las "Entradas" y "Salidas" de información, siguiendo una vez más el concepto de "caja negra" o "sistema abierto", y podemos clasificarlas en cuatro grandes grupos que se completan y se excluyen, con respecto a un supuesto sistema mecanizado en estudio:

- Información mecanizada directamente relacionada en la actualidad con el sistema mecanizado en estudio.
- Información mecanizada o no, que aun no estando actualmente relacionada con el sistema en estudio, influye o es influida por él.
- Información externa al sistema en estudio.
- Información de directrices.

La información directamente relacionada con nuestro sistema de Base de Datos en estudio, es preciso tenerla en cuenta porque es el punto de partida ya existente, aunque bien pudiera suceder que fuera inexistente, bien porque nuestro sistema aborde problemas no mecanizados previamente o por no estar la Empresa mecanizada. La información del segundo grupo, obviamente, siempre existe, y es el área fundamental de interés para el planteamiento de nuestra Base de Datos. Esta información pertenece a la sección de la Empresa para la cual estudiamos la Red de Producción que será nuestro marco principal de análisis.

La información externa no será tratada en principio, aunque conviene examinarla someramente a la hora de analizar los Centros de Actividad de la Red de Producción por si hubiera alguna relación no advertida anteriormente que aconsejara ampliar el conjunto del segundo grupo.

La información de directrices es de otra naturaleza y no vamos a considerarla, pues es campo de estudio de analistas de organización; simplemente señalemos de paso la eterna polémica de si se debe organizar antes de mecanizar o, por el contrario, es factible realizar ambas tareas en paralelo, lo que siempre debe ser meditado.

2.5.1. LA RED DE INFORMACION. FLUJOS DE MENSAJES

Las comunicaciones de Información entre los diferentes Centros de Actividad pueden realizarse de formas diferentes. El vehículo más usual de comunicación es el documento, aunque hoy en día hay que tener en cuenta muchas otras formas de contacto, como puede ser la línea telefónica, télex, microfilm, cinta de papel perforada, cintas magnéticas, etc.; en cualquier caso, atenderemos al hecho fundamental de que se está transmitiendo un conjunto de datos de un Centro a otro, y que dicho conjunto de datos significativos están soportados en un mensaje, pudiendo ser dicho mensaje un impreso, un telegrama, un juego de fichas perforadas, etc.

Para el estudio de la Red de Información vamos a considerar que la Información es transportada en mensajes, y que dichos mensajes tienen un Continente que constituye el vehículo de transmisión, y un Contenido que es el conjunto de datos significativos que el mensaje transporta.

La dinámica de mensajes configura la Red propiamente dicha, y cabe asociarla íntimamente con la Red de Producción porque como ya hemos apuntado antes, los Centros emisores y receptores son los mismos Centros de Actividad que en la Red de Producción considerábamos. El mensaje es un elemento derivado del propio fe-

nómeno productivo, puesto que su razón de ser pertenece a una necesidad surgida en algún momento para asegurar los flujos de producción, aunque cabe aplicar al fenómeno de la Información las consideraciones que aplicábamos a los productos cuando hablábamos de “fertilidad” o “esterilidad”.

De hecho, el propio mensaje puede ser considerado como un producto y en la Red de Producción puede ser considerada la emisión del mensaje, como un Suceso productivo, porque parte de la actividad ha de dedicarse a proporcionar información o incluso porque, aunque a veces la información es un subproducto, puede resultar conveniente tenerla en consideración en la Red de Producción.

Por ejemplo, el Centro de Cálculo será siempre un Centro de Actividad cuya mayor parte de Sucesos productivos corresponda a la emisión de mensajes, y, obviamente, así deberá contemplarse su actividad cuando dicho Centro sea analizado dentro de la Red de Producción.

Esto no va en contra de que, por otra parte, estudiemos todo el conjunto de mensajes intercambiados entre los Centros de Actividad, formando la Red de Producción, haciendo abstracción del hecho de que en algunas partes de la Red de Flujos de Información hayan sido tenidos en cuenta como flujos de la Red de Producción. Los flujos de ambas Redes no tienen por lo tanto necesidad de complementarse y excluirse, y aunque en algún caso exista un flujo común, éste será contemplado en dos planos diferentes, según hemos apuntado anteriormente.

Teniendo en cuenta que el objeto de considerar la Red de Información es el disponer de un instrumento más para el análisis a la hora de plantear un sistema mecanizado de información como pudiera ser una Base de Datos, procede considerar la Red de Información distribuida en tres subredes.

- Red Primaria.
- Red Secundaria.
- Red Externa.

La Red Primaria o circuito Primario de información será la sub-red que abarca el tráfico de mensajes entre los diferentes Centros de Actividad y el Centro de Cálculo y viceversa, o lo que es lo mismo, la Red Primaria estará formada por todos los mensajes que el Centro de Cálculo recibe y emite. Es en definitiva la Red de Información mecanizada del Sistema.

La Red Secundaria o circuito secundario de información es el subconjunto de la Red de Información que soporta el tráfico de mensajes cuyo contenido informático influye o es influido por la Red Primaria, siendo dichos mensajes intercambiados por Centros de Actividad excluido el Centro de Cálculo. La Red Secundaria viene a constituir el Entorno de la Red Primaria.

La Red Externa está formada por los flujos de información restantes contemplados en la Red de Información, y en principio no tienen relación posible con el Centro de Cálculo. Sin embargo, puede considerarse muy detalladamente esta

Red ante un nuevo desarrollo informático, es decir, a la hora de pensar en mecanizar nuevas aplicaciones en la Empresa.

2.5.2. REDES PRIMARIA Y SECUNDARIA

Generalmente, centraremos más la atención en la Red Primaria y en la Red Secundaria, puesto que la Red de Información será utilizada como instrumento que nos aporte una visión de conjunto del equilibrio de flujos de información en la Empresa, y para elaborar un juicio es preciso partir de una situación dada de mecanización o, lo que es lo mismo, de una Red Primaria determinada bien sea existente o bien sea simplemente un proyecto. Por ello, cuando examinemos nuevas zonas a mecanizar, lo que estaremos haciendo es pasar flujos de información de la Red Externa a las Redes Primaria y Secundaria como forma tentativa de diseño, y de ahí que consideremos estas dos Redes como el objeto formal de análisis de la Red de Información.

En la figura 38 se presenta a título de ejemplo una expresión gráfica de una supuesta Red de Información en la que intercambian mensajes ocho Centros de Actividad:

- A) Proveedores. Centro de Actividad que obviamente pertenece al Entorno de la Empresa.
- B) Almacén.
- C) Servicio de compras.
- D) Talleres de fabricación.
- E) Oficinas de distribución.
- F) Centro de Cálculo.
- G) Dirección Gerencia.
- H) Departamento Administrativo.

Centros de Actividad que componen un sub-sistema de Empresa que pertenece a la ficción y con el que por lo tanto, haremos abstracción de su significado real.

En este Esquema de flujos de información, podemos comenzar a examinar la circulación de documentos en la Empresa, y habremos de ayudarnos de inventarios de mensajes de los generalmente utilizados en las metodologías clásicas de análisis, es decir, para los flujos de información habremos de considerar:

- Los tipos de mensajes que circulan, lo que determina su relevancia, aspecto siempre primordial a contemplar.
- Los Centros emisores y receptores.
- Periodicidad de los mensajes.
- Volumen de mensajes en la unidad de tiempo.
- Las acumulaciones de información que se producen en los Centros de Actividad.

- Los desfases entre la emisión y recepción de mensajes.
- Desfases entre los datos contenidos en los mensajes y los tiempos en los datos fueron producidos.

El contraste analítico entre la Red de Información, con todas las matizaciones de inventario de mensajes, y la Red de Producción con todos los detalles relativos a los Sucesos, constituye una valiosa aportación para advertir la fertilidad o esterilidad de cada uno de los circuitos informáticos, y poder considerar en cada momento circuitos sustitutivos.

La implantación de una Base de Datos, supone una gran acumulación en el Centro de Cálculo de información en estado operativo y de información, que será por lo tanto punto de interés para muchos Centros de Actividad. Por ello, el concepto de Base de Datos, sitúa al Centro de Cálculo como punto receptor de muchos mensajes de diversas procedencias y también como viceversa, punto proveedor de información a muchos destinos.

La implantación de una Base de Datos, supone por lo tanto una considerable potenciación de la Red Primaria.

2.5.3. MATRIZ ESTRUCTURAL DE LA RED DE INFORMACION

El grafo de la figura 38, por el cual hemos representado la Red de Información, puede analizarse en abstracto utilizando la matriz asociada que representamos en la figura 39.

La posición ij indica el número de diferentes tipos de mensajes que emite el Centro de Actividad i y que recibe el Centro de Actividad j . Es decir, la posición $i = 3, j = 8$ corresponderá en la figura 39 a los tipos de mensajes que emite el Centro C y recibe el Centro H , en el caso de nuestro ejemplo, solamente uno.

La fila y la columna de la matriz correspondientes al Centro de Cálculo, en nuestro caso el Centro F , corresponderán a los datos pertenecientes a la Red Primaria, y el resto de cifras de la matriz, aportarán la información correspondiente a la Red Secundaria.

A dicha matriz, podemos añadirles nuevas dimensiones que parten de cada uno de sus elementos, y que básicamente habrán de contener:

- Un factor de ponderación de la relevancia de cada tipo de mensaje.
- Datos de periodicidad y volumen.
- Desfases de tiempo.

La matriz contiene por lo tanto, información de la estructura de las Redes Primaria y Secundaria, matizándose para cada elemento de la estructura, la rele-

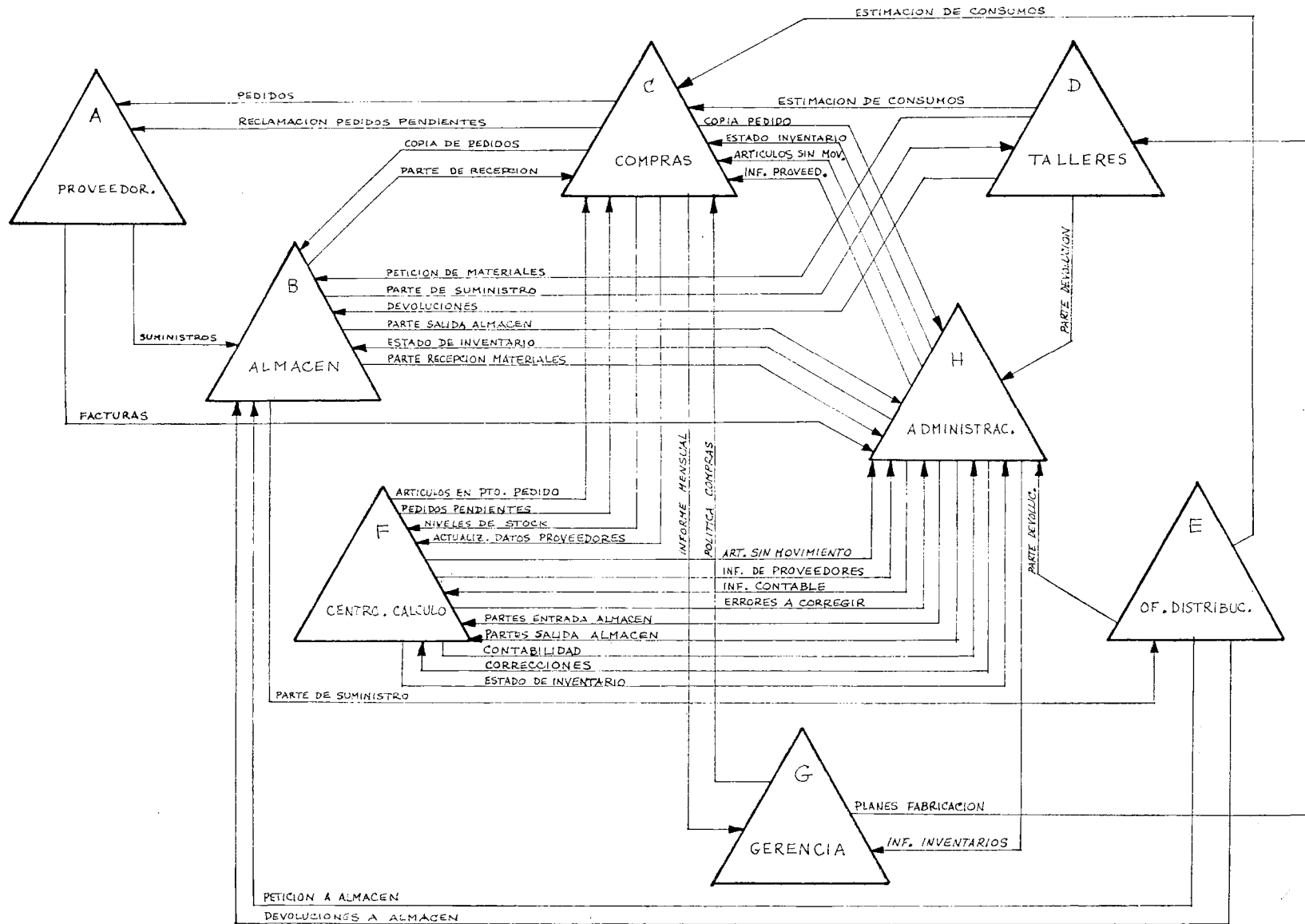


FIG. 38

vancia de cada tipo de mensaje, sus datos de periodicidad y volumen y la información de desfases de tiempo, con lo que se dispone de un modelo de la Red de Información en el que pueden analizarse los circuitos de información de la Empresa en una forma análoga a la que un investigador en neuroanatomía práctica para un mejor conocimiento del sistema nervioso.

	A	B	C	D	E	F	G	H
A	0	1	0	0	0	0	0	1
B	0	0	1	1	1	0	0	2
C	2	1	0	0	0	2	1	1
D	0	2	1	0	0	0	0	1
E	0	2	1	0	0	0	0	1
F	0	0	2	0	0	0	0	5
G	0	0	1	1	0	0	0	0
H	0	1	3	0	0	4	1	0

FIG. 39

Del examen de la matriz estructural, pueden deducirse conclusiones no sólo de conjunto, sino también de específicas. En principio, habrá de analizarse si existen determinados puntos de la Red Secundaria en los que se registre gran acumulación de información, comprobando en qué medida cuando dicha información llegue a la Red Primaria o bien influya en datos de mensajes de la Red Primaria, la información ha sufrido demasiadas transformaciones o se ha degradado.

En realidad, debemos polarizarnos para observar el fenómeno informático puro, de forma que cada Centro de Actividad será contemplado como un centro de pro-

ceso de información, deduciendo del examen de la matriz de estructura, la relevancia del proceso de información que tiene lugar en cada Centro de Actividad, en base a la influencia que ejerce en el conjunto de la Red. Desde el punto de vista informático, obtendremos conclusiones que nos conducirán a plantear modificaciones en el conexionado de la Red, lo que supone una mejora de la estructura. Ello se realiza en base a 3 tipos de operación:

- Supresión de flujos de mensajes.
- Sustitución por flujos alternativos.
- Creación de nuevos flujos.

Por ello, hay que tener muy presente que los elementos de información que manejamos, son datos discontinuos derivados del desarrollo de la actividad que es por el contrario, un fenómeno continuo y que además es el objeto formal que da razón de ser al Centro de Actividad. Quiere ello decir que el hecho de que contemplemos el Centro de Actividad, a efectos informáticos, como un Centro que procesa información, no deberá llevarnos a anteponer los objetivos informáticos a los propios de producción, por lo que no pueden disponerse modificaciones de flujos de información, sin examinar cuidadosamente el desarrollo de la actividad que conduce a ellos. Es decir, aun cuando el fenómeno de desarrollo de la actividad productiva sea también un fenómeno informático, la obtención de datos debe contemplarse como un sub-producto.

El proceso de modificación de estructura de la Red, se realizará por lo tanto después de un profundo análisis y una detallada crítica. A título de ejemplo, y volviendo a la figura 39, suponiendo que todos los mensajes fueran igualmente relevantes, se advierte que el Departamento Administrativo, es el principal Centro de Proceso de información en este ejemplo, puesto que es el que más información recibe y emite. Ello bien pudiera estar plenamente justificado y ser simplemente un hecho a destacar del que se debe tomar conciencia, aunque en este caso concreto se advierte del examen de la Red, que mucha de la información que el departamento administrativo recibe, es enviado posteriormente a otros Centros con poca o ninguna variación, actuando simplemente como intermediario.

Por ejemplo, el informe que el Centro de Cálculo produce de artículos de almacén que no han registrado movimientos de entradas y salidas, es recibido por el Departamento Administrativo y enviado al Departamento de Compras. La utilización que el Departamento Administrativo hace de esta información, no es quizás muy importante, salvo el conocimiento que de ella se deduce, de la cifra económica global invertida en materiales que no son de uso corriente, información que por otro lado, puede interesar más a la Gerencia y de hecho, deberá recibirla en alguna forma a través del extracto mensual que el Departamento de Compras le envía, bien para poder juzgar la conveniencia de dar nuevas directrices a dicho Departamento o para introducir modificaciones en la política de compras. Por otra parte, el Departamento Administrativo, poca o ninguna transformación hará de esta información de artículos en movimiento, pues únicamente es concebible su actuación en la corrección de errores en los datos, y ello simplemente por el hecho de ser un

órgano concentrador de información. Una estructura informática inadecuada a este respecto, determina que el punto informáticamente más capacitado para el control de las inmovilizaciones de stocks sea el Departamento Administrativo, dándose la paradoja de que otra Unidad más responsable al respecto como es el Departamento de Compras, puede recibir información menos fiable.

Puede sugerirse entonces la idea de que el informe de artículos de almacén sin movimiento, sea enviado directamente por el Centro de Cálculo al Departamento de compras, sin necesidad de que circule a través del Departamento Administrativo, desplazando el flujo de mensajes de la red secundaria a la red primaria. Sin embargo, hagamos notar que el hecho de pensar en esta alteración de flujos en forma aislada, puede tener graves consecuencias, pues si realmente el Departamento Administrativo actúa en este sistema como principal controlador y depurador de la información, podría suceder que al sustraer de su control datos parciales, se produjera como consecuencia un descenso del nivel general de calidad en la Red de Información.

Con esta pequeña reflexión sobre nuestro ejemplo, queremos poner de manifiesto que si bien parece advertirse en seguida conclusiones parciales en una primera toma de contacto con la Red de Información, el formalizar dichas conclusiones sólo es aconsejable realizarlo después de haber consolidado un contraste de la Red con los objetivos que tenemos planteados y en base a los cuales estamos analizando.

Así en el caso del ejemplo, después de perfeccionar ya un examen en profundidad, llegaríamos al conocimiento de que efectivamente se produce una concentración de control de la información en el Departamento Administrativo, probablemente debido a que el sistema de empresa en general, era en principio de un carácter mucho más "administrativo" que de "gestión", y si la mecanización se aplicó sobre la base existente, el ordenador fue utilizado más bien a la sustitución de procesos manuales pre-establecidos. Es decir, se introdujo un nuevo Centro de Actividad, el Centro de Cálculo, que vino a actuar de intermediario para perfeccionar determinados flujos de información, pero quizás no se pensó en un cambio cualitativo de la estructura de la Red de Información.

También puede haber sucedido que se tenía conciencia de que el Departamento Administrativo estaba bien dotado de recursos y que debido a su experiencia, era la Unidad más idónea para ejercer el control principal de la información del sistema, hasta tanto en cuanto el Centro de Cálculo, no probara su madurez. O bien, quizás en un principio, el Centro de Cálculo comenzó dependiendo del Departamento Administrativo, cosa muy común aún actualmente, pero desde luego, situación prácticamente generalizada en las primeras etapas de la mecanización, ya que el ordenador, la mayor parte de las veces, fue introducido en las Organizaciones, justificándose su rentabilidad en base a la simplificación de las tareas administrativas.

Lo anterior constituye una pequeña muestra de las consideraciones que debemos tener en cuenta para mejor estudio y comprensión de la Red de Información.

En nuestro planteamiento de una Base de Datos, vamos a producir en la Red de Información una verdadera mutación de su estructura porque vamos a potenciar considerablemente la Red Primaria en detrimento de la Red Secundaria. Como todo cambio drástico, el éxito de nuestra Base de Datos, sólo podrá consumarse si la Red de Información resultante es armónica, equilibrada y compatible con la actividad productiva.

3. METODOLOGIA PARA EL DISEÑO DE UNA BASE DE DATOS

El diseño e implantación de una Base de Datos, tiene dos vertientes básicas a considerar:

- Las actividades específicas de proceso de datos, cual es el poner a punto el Software y el Hardware.
- El planteamiento informático adecuado para que el sistema de Base de Datos se ajuste a unos objetivos de Empresa, y conduzca a una Red de Información racional.

El primer aspecto, es decir, el puramente técnico, comienza con la elección del Ordenador y del Software de base, en lo que no vamos a entrar por ser un problema de tipo general muy examinado. Solamente plantearémos a ese respecto, la conveniencia de adoptar un Software de Base de Datos disponible en el mercado o bien, efectuar un diseño ad-hoc de Software básico.

El diseño ad-hoc, puede ser requisito muy conveniente en sistemas de índole muy concreta para los que los paquetes disponibles en el mercado, no se ajusten con garantías, y estando justificado por otro lado en base a la envergadura del problema a resolver, ya que para diseñar y mantener un software de Base de Datos hecho a la medida, se requiere disponer de profesionales muy cualificados, lo que sólo ocurrirá en aplicaciones de mucho contenido.

Por otra parte, la adopción de sistemas privados de Base de Datos, tiene siempre el riesgo de que sólo podremos hacer mejorar el sistema con nuestras propias fuerzas sin beneficiarnos directamente por "ósmosis" del progreso tecnológico de toda una comunidad de usuarios, lo que a medio y largo plazo, pone en peligro la compatibilidad con los nuevos desarrollos informáticos.

Hoy en día, la industria informática en general, tiene planteado como actitud conveniente, el desarrollo de sistemas compatibles de uso universal y consecuencia de ello, han sido las especificaciones de la Conferencia CODASYL, propuestas por un Grupo de Trabajo formado por proveedores y usuarios para aplicación a sistemas de Bases de Datos, producto que parece estar consolidándose, pues no en vano ha nacido en la misma "cuna" en la que nació el COBOL. Quizá no sea mucho aventurar que para el futuro, habrá sistemas standard adoptados por diferentes fabricantes y con garantías de compatibilidad, pues ya en la actualidad hay muestras de ello, aunque los proveedores actuales que siguen las especificaciones de CODASYL, hayan introducido algunas modificaciones; en general, para compatibilizar el nuevo planteamiento, con los sistemas de ficheros que cada proveedor tenía en el pasado.

No hay que ocultar tampoco, que el utilizar software standard, tiene también sus desventajas, pues como herramienta universal que es, ha de prever estructuras

y soluciones de uso general que cubran no sólo las necesidades de nuestro sistema específico, sino de otros muchos, y ello siempre introduce penalidades de utilización de memoria, tiempos de acceso, ... etc. Aun cuando el planteamiento de una metodología ha de ser independiente del sistema de software utilizado, a la hora de acudir al ejemplo como paso obligado para fijar ideas, utilizaremos el sistema DBMS de CODASYL, por ser el conjunto de especificaciones que en este momento son más universales y estar disponibles para la mayoría de los ordenadores (*).

Continuando con los aspectos derivados de las actividades técnicas de proceso de datos, nos ocuparemos fundamentalmente de aquellas que tienen una relación con los problemas que la Base de Datos pretende resolver, aspectos tales como:

- Diseño de estructuras lógicas.
- Dimensión de almacenamiento de datos.
- Dinámica de procesos de registros.
- Especificaciones de acceso a los datos.
- Captación de datos.
- Especificaciones de protección de la información.
- Administración de la Base de Datos.

La segunda vertiente de diseño de la Base de Datos, es la relativa a su engarce en la Red de Información de la Organización, actividad que debe ser guía para todo el planteamiento y que es la que pretendemos básicamente abordar.

Para ello, el estudio informático puro, no puede plantearse en base a un procedimiento determinista muy perfilado que se aplique en forma mecánica, porque la tarea de análisis es un proceso creativo en el que no sólo se aplica una técnica, sino que tampoco están ausentes las aportaciones de tipo artístico. Por ello, la metodología que se presenta, pretende ser un abanico de sugerencias, una guía de aspectos a abordar, y un conjunto de instrumentos a manejar, más que un camino perfectamente definido. Dicho en otras palabras, no se considera conveniente guiar al analista, sino simplemente orientarlo para que se guíe con su propio criterio, puesto que solamente analistas con suficiente criterio, deberán abordar un proyecto informático de la complejidad de una Base de Datos.

3.1. PLANTEAMIENTO DE OBJETIVOS Y MEDIOS

El planteamiento de Objetivos y Medios, es el punto de partida racional de cualquier actividad informática y por consiguiente, habrá de marcar el comienzo de un

* *Nota del autor.*—Aunque en estos momentos IBM no disponga de productos propios ajustados al sistema Codasyl, es de hecho una de las firmas participantes en su creación y por otra parte, puede adquirirse en el mercado software para usar el sistema Codasyl en los ordenadores IBM.

proyecto de Base de Datos. Deberá quedar establecido de una forma clara y tan detallada como se estime necesario:

- Antecedentes relevantes.
- Las circunstancias que dan lugar al planteamiento del proyecto.
- Qué Unidad plantea el proyecto.
- Qué Areas de la Organización están involucradas.
- Una idea primaria de la envergadura económica.
- Detalle de los objetivos que se pretenden cubrir.
- Recursos informáticos a utilizar en principio.
- Panorámica tentativa de tiempos.

Ello deberá centrarnos el cuadro general en el que vamos a operar, por lo que es conveniente que todo este conjunto de puntos de partida, se agrupen en un Informe que tiene el doble propósito de aportar el debido conocimiento para el equipo de especialistas que van a desarrollar el trabajo, y de ser sometido a la Dirección de la Organización, para que se vincule en forma activa al proyecto o en su defecto, proponga de forma alternativa, nuevos enfoques o modificaciones.

Este estudio de partida, debe ser suficiente para que situemos nuestro proyecto en su justo nivel, obteniendo una idea clara de la relevancia que va a tener en la Organización y concretamente del impacto que en líneas generales, va a suponer para la Red de Información.

También debemos de formarnos una idea de cuál es el entorno de la actividad empresarial que está involucrada con el proyecto, es decir, desde el comienzo, habremos de conocer qué Unidades Organizativas van a ser usuarias de la Base de Datos y qué Unidades son las que en el desarrollo de su actividad, van a proporcionarnos datos para incorporar a la Base.

Este primer conocimiento, nos ha de aportar un criterio económico de partida, que nos sirva de orientación, al menos en principio, a la hora de profundizar en el estudio de la actividad, y determinar los grados de detalle a conseguir. Con el conocimiento paralelo necesario de la problemática de la Organización, a la hora de considerar cada dato elemental, podremos matizar el volumen de información de conjunto que será preciso tener almacenado, lo que enfrentado a la dimensión económica del proyecto, nos permitirá pronunciarnos sobre la conveniencia de manejar tal dato, o bien sintetizarlo o incluso segregarlo en componentes todavía a mayor grado de detalle.

El primer planteamiento, tiene en definitiva que expresarnos con claridad, las motivaciones del proyecto, lo que de él se espera conseguir y los medios que podremos movilizar. No vamos a detallar mucho más acerca del planteamiento de objetivos y medios, porque se trata en realidad de un tema general de aplicación a cualquier tipo de proyecto, sea informático o no informático.

3.1.1. LANZAMIENTO DEL PROYECTO

El proyecto de desarrollo de una Base de Datos, puede surgir por diferentes iniciativas:

- La Alta Dirección de la Empresa advierte por ella misma o ayudada por sus asesores, que la Base de Datos es un instrumento que integra la información desde sus estructuras naturales, y que es por lo tanto un camino conveniente de emprender para mejor coordinación de las Unidades participantes en las diferentes funciones y para la obtención de información de gestión a diferentes niveles de síntesis. Esta iniciativa, suele corresponder a Direcciones modernas y dinámicas, que acostumbran a informarse de los nuevos procedimientos, pudiendo suceder también que la persona o personas de Dirección, tengan conocimiento de Informática o procedan de esta área de actividad, lo que ya sucede en muchas ocasiones. De hecho, la Informática analiza desde un mismo prisma los diferentes sectores de la Empresa, y cuando un experto en informática ha desarrollado varios sistemas, puede llegar a conocer muy bien la Empresa, lo que le califica para un puesto de alta Dirección, si se diera el caso de cumplir algunas características personales adicionales obviamente necesarias.
- El proyecto puede surgir a propuesta de una o varias Unidades usuarias, generalmente una, porque la coordinación entre diferentes usuarios para estos menesteres, suele dejar bastante que desear, a no ser que esté constituido un Comité de Usuarios no demasiado frecuente en Europa hasta el momento, en cuyo caso, suele cumplir una función de Asesoría de Gerencia, y un planteamiento de dicho Comité equivale a una iniciativa de la alta Dirección.

Cuando el planteamiento parte del usuario, el proyecto propuesto está basado en un conocimiento detallado, pero solamente de un sector determinado de la Organización, y en ocasiones puede ser también un conocimiento deformado, a no ser de que la Unidad usuaria esté muy impuesta en Sistemas de Información o que por su natural función en la Organización, perciba la problemática de muchas otras Unidades y sea quien detecte en primer lugar la conveniencia de coordinar los Sistemas de Información.

En cualquier caso, cuando el planteamiento parte de un usuario, antes de que el proyecto se formalice, suele ser discutido en un ámbito más amplio de la Organización, porque una Base de Datos, suele tener implicaciones en un entorno más amplio que el del usuario que las plantea.

- El proyecto puede surgir como una propuesta del Departamento de Informática. En este caso, puede suceder que la Unidad de Informática tenga un conocimiento adecuado de la utilización de los Sistemas de Información en las diferentes áreas de la Organización, en cuyo caso, está perfectamente capacitada para plantear el proyecto en términos que a la

Empresa realmente interesen. Sin embargo, en caso de que la Unidad de Informática, no pueda “sustituir” a los usuarios, por falta de conocimiento de las funciones, el planteamiento puede surgir en base a motivaciones técnicas, las que normalmente habrán de ser consideradas como secundarias.

Sin embargo, resulta interesante que sea la Unidad de Informática, quien tiene la iniciativa porque en este caso, estamos ante una Organización que efectivamente cuenta con un equipo que vela por los Sistemas de Información y no solamente por la técnica hard/soft de Proceso de Datos y es precisamente un núcleo de personas con esa vocación, el que suele implantar con éxito este tipo de aplicaciones complejas e integradas.

- El proyecto puede surgir como una necesidad por pura presión de la competencia. Hay determinados tipos de Empresas y actividades dentro de ellas, que requieren la utilización de Bases de Datos, y la Empresa del ramo que de esta forma plantea sus problemas, puede llegar a adquirir determinadas ventajas en el mercado, lo que obliga a la competencia a seguir los mismos pasos.

Cuando el lanzamiento del proyecto esté motivado por presión de la competencia, tendremos que “copiar” lo más posible las prácticas de nuestros competidores, pero ejerciendo a la vez una serena actividad crítica que nos permita evitar las partes negativas y en todo caso perfeccionar nuestro proyecto respecto del de los competidores, porque la experiencia de éstos, debemos de tratar de hacerla nuestra. Hay por lo tanto que evitar el simple mecanicismo de adoptar procedimientos externos sin antes matizar lo que en sí mismos tienen de buenos o de malos, y la medida en que son aceptables a nuestra Organización.

- El lanzamiento del proyecto, puede ser producto de las sugerencias de los proveedores de Informática. Las aplicaciones que nos propongan siempre estarán estudiadas atendiendo a sus intereses comerciales, y ello aun cuando éstos también pudieran coincidir en gran parte con los intereses propios de nuestra Organización. Esto hay que tenerlo muy en cuenta, y nunca debemos utilizar en toda su amplitud los planteamientos de los proveedores como sustitutivos de nuestras propias ideas. El proyecto de Base de Datos, tiene una componente importante de integración de Sistemas, que obligadamente habrá de ser tratada en el seno de la Organización. El proveedor, aun cuando le mueva un sano interés y una acción comercial propia muy honesta, no estará debidamente impuesto en todas las circunstancias internas que obligadamente hay que considerar, a no ser que éste sea un consultor al que le hemos solicitado expresamente que nos realice tal estudio, en cuyo caso lo que viene a realizar es una función que de antemano sabemos no podemos ejercerla en propio.

En cualquier caso, sea cual fuere el génesis del proyecto de Base de Datos, a la hora de consolidarse como actividad de desarrollo, debe de encontrar un gran

apoyo en la Organización, tanto para asegurarnos de que existe un interés de todas las partes involucradas en llevarlo a término, como para evitar problemas de aceptación por parte de alguna Unidad involucrada que mantiene sus reservas sobre el proyecto.

La forma de evitar cualquier posible reserva que pueda perturbar o retrasar el desarrollo del proyecto, es única y podemos resumirla en las tres siguientes medidas complementarias:

- Ideas claras y planteamiento diáfano.
- Apoyo de la Alta Dirección.
- Responsabilidad clara de cada una de las partes involucradas.

Vamos a presentar a título de ejemplo, continuando otros ejemplos parciales expuestos en capítulos anteriores, un supuesto lanzamiento de proyecto de Base de Datos, que sin tratar de convertirse en una aplicación exhaustiva, sirva para fijar nuestras ideas de una forma más coherente a lo largo del estudio de la metodología. Con un ejemplo único integraremos más nuestras ideas, aun cuando el propio ejemplo habrá de ser solamente parcial por obligada concisión y aun cuando dentro de la parte a tratar, tampoco podemos descender demasiado a la casuística.

Vamos a plantear como área de estudio, una Base de Datos del tráfico de pasajeros, aviones y tripulantes en una Compañía aérea regular cuyos objetivos globales son:

- Registro del pasajero y de las circunstancias y características de su relación con la Compañía.
- Información de análisis de mercados, líneas, tarifas, ingresos y rendimientos en general del programa de servicios y de los recursos dedicados a realizarlo.
- Programación interactiva de vuelos y tripulaciones. La programación de vuelos de una Compañía regular de cierta dimensión, es extraordinariamente complicada. Se pretende dirigir la programación en forma interactiva, esto es, contrastando cualquier programa de vuelos o modificación del mismo, con diagnósticos automáticos de ayuda a la decisión.
- Formación de un modelo estructural de la actividad comercial, relacionando tipos de avión, puntos de operación, líneas, países, ciudades, mercados, temporadas, ... etc., de forma que cualquier tipo de información real o cualquier síntesis histórica o prospectiva que sea relevante, pueda obtenerse para el futuro de una forma coordinada con datos únicos y fiables.

Vamos a prescindir de exponer las motivaciones y génesis de nuestro ejemplo de proyecto, por no aportar conclusiones universales.

3.1.2. ANALISIS DE OBJETIVOS

Los objetivos que pretenden alcanzar con la implantación del proyecto, conviene que sean analizados al mayor grado de detalle posible, para ver la relevancia de cada uno de ellos y las interconexiones que existen, juzgando cuáles han de ser las prioridades cuando bajo algún determinado aspecto, haya objetivos contrapuestos. En cualquier caso, el tener claros los objetivos, es condición necesaria para tener las ideas claras, por lo que en caso de defecto, el Analista o Analistas, deberán familiarizarse previamente con el problema, sin avanzar más hasta que este conocimiento haya sido adquirido.

El conjunto de objetivos, tiene que ser estudiado en el ámbito temporal, porque la Base de Datos es una infraestructura que como tal, sirve no sólo a objetivos presentes, sino también a los de largo plazo, o al menos ha de ser compatible con ellos.

De este análisis, habremos de manejar ya un criterio suficiente para determinar el grado de detalle al que habremos de llegar en el manejo de los datos, lo que es un compromiso de rentabilidad.

El conjunto de objetivos constituye la finalidad que justifica la Base de Datos y por lo tanto, debe estar claramente expresada y comprendida antes de continuar. No se puede dar una regla lógica concreta para el análisis de los objetivos, sino solamente indicaremos que el análisis de objetivos ha de desarrollarse hasta un punto tal, que pueda expresarse dicho análisis por escrito y que el informe resultante, resulte inteligible, suficientemente claro, y convincente en ideas para un lector conocedor del área de la actividad, pero ajeno a nuestro proyecto. Esta es la mejor prueba de que el análisis de objetivos ha sido correctamente desarrollado. Para asegurarlo, puede resultar conveniente que este informe resultante, tenga que contar con la aprobación de la Dirección antes de proseguir su desarrollo.

A título de ejemplo y siguiendo con el desarrollo de nuestro caso, vamos a analizar a continuación uno de los objetivos: la programación de vuelos.

La programación de vuelos de una Compañía aérea, es realizada por un grupo de expertos que han de manejar los tiempos hábiles de utilización de las flotas de aviones para cubrir con ellos el servicio de líneas comerciales que convienen a la Empresa. Los aviones y las líneas que sirven, son intercambiables en los Aeropuertos en función de los horarios y tipos de avión, lo que hace que las rutas previstas durante una temporada para cada uno de los aviones físicos existentes, pueda reajustarse en función de las incidencias en la operación de los programas. Por ello, las "rotaciones" de los aviones, pueden ser intercambiadas, y en todo caso, la operación de una rotación, condiciona o puede condicionar el desarrollo de las demás.

Cuando se programa una nueva línea o se modifica un determinado servicio, todo el conjunto resulta afectado pues o bien existen holguras en las rotaciones, o hay que "abrir hueco" en nuestros recursos de vuelo, para poder realizar el nuevo servicio. Como además se da la circunstancia de que en la práctica, las necesidades de modificar la programación se producen a diario, los expertos que manejan

la programación, habrán de tener necesariamente un modelo analógico en su cabeza, que les proporcione un criterio suficiente para intuir cómo cada decisión que adoptan, influye en el resto de la programación.

Nuestra Base de Datos, no puede por lo tanto proporcionar un sistema que reemplace estos criterios de personal experto, sino que los complemente. Por ello, nuestro objetivo al respecto, ha de ser potenciar al responsable de la programación al máximo, descargándole de todas las tareas analíticas que un ordenador puede ofrecerle, pero procurando no programar decisiones implícitas que introduzcan rigideces en el sistema y limiten la creatividad del experto.

Como la programación puede ser compleja en exceso cuando se manejan cifras de 50 a 100 aviones, el primer objetivo del sistema, será el disponer en la Base de Datos un modelo operativo, es decir, procesable, de cuál es la programación en cada momento planteada, y cuáles son los tiempos disponibles, y los tiempos asignados para cada avión. Así, en primer lugar, el usuario podrá interrogar al sistema de cualquier detalle que precisen en un momento dado acerca de la programación sin necesidad de manejar demasiada documentación. Además, esta información constituye la definición de la oferta de servicios de la Compañía, lo que obviamente servirá de marco de referencia al cual se asociarán todos los datos de tráfico de pasajeros, ventas de billetes, ... etc.

Los responsables de la programación de vuelos, realizarán la programación, o modificarán la misma, operando con la Base de Datos en forma interactiva mediante programas de proceso que comparen los atributos y características del programa modificado que plantean con el ya existente, de forma que se tenga un diagnóstico de lo que con la modificación va a conseguirse. Es decir, el conjunto de datos que reflejamos en la Base, deberá ser suficiente para elaborar programas de proceso del siguiente tipo:

- Diálogo mediante transacciones del usuario con el sistema sobre holguras disponibles en la utilización de aviones y nuevos servicios a realizar con aportación de posibles reajustes de vuelos.
- Informes que evalúen los programas de vuelos que se introduzcan.
- Posibilidad de comparar programas históricos presentes y previstos para el futuro.
- Posibilidades de registro de las incidencias del programa operativo y sus causas, elaborando un “conocimiento” de las incidencias imputables a la propia estructura del programa.
- Diagnósticos analíticos de los programas en base a mercados, tráfico históricos, etc.
- Etcétera.

Con un análisis de objetivos de este tipo, detallando tanto como sea necesario, estaremos ya en condiciones de plantear el estudio de la actividad.

3.1.3. ASIGNACION DE RECURSOS MATERIALES Y HUMANOS

Es un requisito previo para el desarrollo de todo proyecto informático y no vamos a extendernos en comentarios. En cuanto a los materiales, ya hemos indicado hay que hacer consideraciones previas acerca del hardware y software básico que va a ser utilizado, punto de partida que damos por supuesto y que no analizamos.

En lo que se refiere a la asignación de recursos humanos, tengamos en cuenta:

- Es preferible que a nivel de Analistas y colaboradores, seleccionemos personas con ideas claras y prácticas más que grandes técnicos. Estos últimos pueden aportar un buen nivel científico, pero generalmente tienden a complicar los problemas en exceso.
- Los Analistas participantes, deben ser grandes conocedores de los problemas, o bien adquirir dicho conocimiento convenientemente. Sin embargo, debemos de procurar evitar el concurso de personas que aún conociendo los problemas, fueran parte de ellos. Es decir, personas conociendo los problemas, pero sin estar condicionados por ellos, para poder tener un espíritu de apertura necesario para posibles cambios de procedimientos.
- Por supuesto, los Analistas deben de tener un conocimiento suficiente de la buena práctica informática y de las técnicas de proceso de datos que van a utilizar.

3.2. EL ESTUDIO DE LA ACTIVIDAD

Los objetivos de esta fase preliminar de estudio de la Actividad son:

- Identificar los Centros de Actividad de la Organización y de su Entorno que influyen o son influidos directamente por el sistema de Base de Datos que pretendemos diseñar.
- Determinar el entorno de cada Centro de Actividad y del conjunto que constituye nuestra área de interés, asegurándonos que fuera del marco de actividad que vamos a manejar, no existen fenómenos productivos relevantes para los objetivos del proyecto.
- Identificación de los Sucesos productivos que tienen lugar en los Centros de Actividad y que son relevantes en el marco de nuestro proyecto.
- Detectar en un primer tanteo, posibles necesidades de modificar procedimientos de trabajo en la Organización.

Para centrar estas cuatro finalidades, es preciso considerar cada posible Centro de Actividad como un sistema abierto o “caja negra”, en el cual se integran determinadas unidades de recursos de la Organización. Por ello, es conveniente tener presente la estructura de Unidades Orgánicas conociendo las funciones que en la práctica cada Unidad realizó, y conviene recalcar que sea el conjunto de funciones real o pragmáticamente ejercidas y no el conjunto de tareas que pueden estar asignadas en un Manual de Organización, que muy comúnmente no se corresponden con la práctica.

Si se dispone de una Red de Información ya analizada, que recoja los flujos informáticos disponibles en los momentos actuales, ello servirá como instrumento valioso de ayuda para identificar nuestros Centros de Actividad en la forma que nos interese, incluso podremos advertir las lagunas, redundancias e inconsistencias que nuestro proyecto deberá subsanar.

Sin embargo, lo más probable es que no dispongamos de una Red de Información suficientemente estudiada puesto que su análisis se suele realizar con motivo de abordar nuevos proyectos. En este caso, y en este primer tanteo, nos limitaremos a analizar de forma individual para cada Centro de Actividad que identifiquemos, las entradas y salidas actuales de información que el Centro de Actividad tiene, como sistema abierto que es. Ello nos permitirá por un lado, confirmarnos a nosotros mismos que el Centro de Actividad está bien identificado, y por otro lado, servirá de contraste posterior con los Sucesos relevantes que en dicho Centro se produzcan. De dicho contraste, resaltarán las inconsistencias que desde el punto de vista informático, existan como punto de partida.

3.2.1. ESTUDIO DE SISTEMAS PRODUCTIVOS

Consiste este análisis en identificar los Centros de Actividad que componen el marco de operación en la Red de Producción. Para ello hemos de utilizar los conceptos expuestos anteriormente sobre el Sistema de Empresa, y servirnos como ayuda de las reglas normales de la lógica deductiva e inductiva y del conocimiento más o menos completo que ya hemos de tener de las áreas funcionales sobre las que vamos a proyectar.

Deberemos examinar con detalle las funciones de Organización en general que están involucradas con nuestros objetivos de Base de Datos, y para cada una de ellas, tendremos que repasar de forma sistemática qué Unidades orgánicas las realizan y cómo se desarrollan en la realidad pragmática. Con este repaso de las funciones y sus atributos a manera de listas de comprobación, y ayudados por los instrumentos antes citados, estaremos realizando un trabajo analógico que debe estar presidido por la creatividad, y ello habrá de producirse en la obtención de una síntesis de los Centros de Actividad primarios que vayan a componer nuestra Red de Producción.

Nuestro proceso mental, deberá en todo momento considerar cuáles son las unidades orgánicas que realmente existen en la Organización, pero por otro lado, deberá impregnarse de la idea de que la “función crea el órgano” y en nuestros pen-

samientos las “funciones” serán para nosotros los “paquetes de actividad” pragmáticamente ejercidos, y los “órganos” serán los Centros de Actividad que habremos de identificar. Es decir, con una contemplación sistemática de las funciones, extraídas de la más estricta praxis, habremos de destilar el marco orgánico que la teoría de sistemas “reconoce”, independientemente de que éste corresponda o no a la ordenación de Unidades operativas que esté vigente en la Organización.

Volviendo a nuestro ejemplo, y supuesto ejercido todo este proceso de análisis, habremos llegado a unas conclusiones que resumimos:

- *Función Ventas.* Es desarrollada en Organizaciones territoriales y el fin último de la venta se materializa en las oficinas donde tiene lugar el acto jurídico de la contratación del servicio. Estas oficinas de ventas, identificamos en nuestro caso que corresponden a tres grupos que para nosotros es relevante el diferenciar:
 - *X* — Oficina de Ventas propiedad de nuestra Organización.
 - *Y* — Agencias de Viajes.
 - *Z* — Oficina de Ventas pertenecientes a otras Compañías aéreas.

Para nosotros, cada Oficina de Venta será un Centro de Actividad que corresponderá a uno de los tres tipos de los Centros de Actividad expresados.

- *Función Reservas.* La función de Reservas la identificamos con un mecanismo de distribución de la oferta de servicios que canalizará el flujo de la actividad comercial y en principio, vamos a identificar un Centro de Actividad *R* y posteriormente ya entraremos en consideraciones de en qué forma aplicaremos, caso necesario, la propiedad distributiva de los sistemas, subdividiendo dicho Centro de Actividad en un conjunto de sistemas productivos de nivel inferior (véase ejemplo de la fig. 31).
- *Función de embarque de pasajeros.* Es la actividad destinada a que un pasajero que llega al Aeropuerto, sea embarcado en un avión o bien rechazado por el motivo que fuere. Esta función, nos lleva a identificar al centro *F* que denominamos Facturación, que lógicamente está integrado en un Aeropuerto, por lo que habrá un centro *F* al menos en cada Aeropuerto que consideremos, e incluso dentro de cada Aeropuerto podremos posteriormente aplicar la regla distributiva en la medida en que nos convenga.
- *Función que permite la salida de un vuelo de un Aeropuerto.* Identificamos el centro de actividad *D* que denominaremos “Despacho del vuelo” que análogamente al caso anterior, se trata de un Centro de Actividad localizado en un Aeropuerto.
- *Recepción de vuelos que llegan.* Esta actividad es la destinada a recibir un avión que llegue de viaje y canalizar convenientemente a los pasajeros de llegada. Identificamos los Centros de Actividad tipo *L* que están igualmente localizados geográficamente en Aeropuertos.

- El propio Aeropuerto, es de por sí un Centro de Actividad de suficiente relevancia para que lo consideremos como el sistema más general que está presente en la correspondiente localidad. Le denominaremos S y obviamente integrará a los Centros F , D y L en cada caso.
- *Función de Control de Red* Es la función que garantiza toda la mecánica general de la Compañía de movimiento diario de aviones, e identificamos un Centro C que denominaremos Control de Red, cuyo producto principal, serán las decisiones diarias de asignación de aviones físicos a rutas programadas, y asimismo, de asignación de tripulaciones a dichos aviones físicos. Naturalmente, estas decisiones están en gran medida ya programadas de antemano y este Centro de Actividad lo que en realidad efectúa es el “acabado” del producto con las necesarias decisiones de última hora en función de las incidencias que en cada momento pueden producirse que obviamente, no han podido contemplarse en la programación anterior.

Este Centro de Actividad, lo concebimos de momento como un Centro único.

- *Función de programación de la oferta de servicios.* Identificamos el Centro de Actividad P de Programación de vuelos, que habrá de realizar la planificación que acabamos de señalar. Identificaremos su producto en dos diferentes vertientes relevantes a nuestros objetivos de Base de Datos:
 - Definición de la oferta de servicios que los centros X e Y podrán vender, y que el centro R podrá distribuir.
 - Definición del esquema de movimiento de aviones que constituye el producto que recibe el Centro C como base de partida para su actividad permanente. Normalmente, en una Compañía aérea, la frontera entre el Centro P y el Centro C , se sitúa a estos efectos en veinticuatro horas anteriores a la salida de cada vuelo. Antes de pasar a dicha frontera, las decisiones de programación de aviones, deberá ejercerlas el Centro P . Después de dicha frontera, será el Centro C el sistema responsable.

El Centro P , lo identificamos como único Centro, aunque posteriormente podría convenirnos aplicar la propiedad distributiva de sistemas atendiendo a las diferentes flotas de aviones.

- *Programación de tripulaciones.* Identificamos el Centro T sobre el que cabe hacer análogas consideraciones que para el Centro P , aun cuando sólo en lo relativo al segundo aspecto de los dos considerandos.
- *Mantenimiento de aviones.* Esta actividad, de extraordinaria complejidad solamente nos interesa contemplarla bajo el aspecto del producto final que nos determina los días y horas en los que cada uno de los aviones está disponible para el servicio, y por consiguiente, nos limita las fechas en las que cada avión está disponible por entrar dentro del

proceso de mantenimiento. Como este producto final es relevante para nuestros objetivos, consideraremos en nuestra Red de Producción, un Centro *M* que designaremos como Mantenimiento de aviones.

Esta primera selección de Centros de Actividad, viene a determinar el “marco” de la Red de Producción, y posteriores análisis nos aconsejarán los diferentes niveles a los que en cada caso habremos de llegar aplicando la propiedad distributiva de los sistemas. Hay que señalar que dado que la actividad de proyecto siempre se realiza por aproximaciones sucesivas, bien puede suceder que tengamos que volver sobre estas consideraciones previas, modificando la determinación de Centros de Actividad por necesidades de eficacia del modelo. Nuestro modelo se comportará por lo tanto como un sistema abierto, dispuesto siempre a recibir “realimentaciones” que lo perfeccionen.

3.2.2. DELIMITACION DE ENTORNOS DE ACTIVIDAD

La identificación de Centros de Actividad que hemos realizado, vamos a utilizarla como “marco productivo” que nuestro proyecto va a tener en cuenta. Por ello, es necesario:

- Asegurarnos que el marco es completo y fuera de él, no existe ningún fenómeno productivo relevante para nuestros objetivos.
- Eliminar de los Centros de Actividad que integran nuestro marco, aquellas actividades que sean claramente independientes de nuestro proyecto.

El criterio principal a aplicar, es el distinguir qué es lo relevante y qué es lo accesorio en relación con nuestros objetivos. Teniendo presente este criterio, procederemos en la siguiente secuencia:

- Para cada Centro de Actividad previamente establecido, identificaremos cuáles son las actividades productivas que en el Centro se desarrollan y que influyen o serán influenciadas en forma directa por nuestro proyecto, o lo que es lo mismo, están relacionadas en forma suficientemente relevante con nuestros objetivos. Lo que haremos será seleccionar del total de las actividades que en dicho Centro tienen lugar, aquel conjunto que constituye nuestro área de interés.
- Repasar posteriormente, todos los Centros de Actividad uno a uno y las actividades que hemos considerado, asegurándonos que el conjunto de actividad es completo y que no falta ningún fenómeno productivo en la Organización que a primera vista consideremos relevante y que no lo hayamos incluido en nuestro marco.
- Repasar igualmente examinando los Centros uno a uno, que no hemos incurrido en duplicidades a la hora de encuadrar actividades y centros, para evitar que nuestro modelo tenga ambigüedades. Es decir, para cual

quier actividad en la que pensemos, el modelo deberá indicarnos dónde, cómo y quién la ejerce.

- A la vista de cada Centro de Actividad y de los aspectos productivos contemplados, podemos formarnos una primera idea tentativa en cuanto al nivel de profundidad al que habremos de llegar en cada Centro. Es decir, podemos aprovechar este repaso sistemático de los centros de actividad identificados, para empezar a formarnos nuestras ideas en cuanto a las sub-divisiones lógicas que para cada Centro habremos de considerar, descomponiendo éste en sistemas de inferior nivel.

Esta secuencia de cuatro etapas, es obligada al menos en lo que respecta a las tres primeras etapas, al término de las cuales tiene que estar perfectamente definido el entorno de actividad que vamos a tomar en consideración. La cuarta etapa, pudiéramos considerarla como opcional, aunque conveniente a nivel de ideas para proseguir con nuestro trabajo.

3.2.3. AISLAMIENTO DE SUCESOS RELEVANTES

Esta fase consiste en analizar a la vista de cada uno de los centros productivos y de las actividades con que contribuyen a nuestro entorno de interés, los Sucesos productivos que tienen lugar y que tienen relevancia suficiente en relación con nuestros objetivos del sistema de Base de Datos en proyecto.

Como consecuencia de este análisis deberá determinarse cuáles son los sistemas productivos que vamos a considerar en la Red de Producción. Esta definición de sistemas viene a definir las “salidas” de productos de los Centros de Actividad y supone por lo tanto el elemento básico para la construcción de la Red de Producción. La selección de Sucesos reviste una importancia muy grande, porque el Suceso es nuestro elemento central a la hora de diseñar registros. Por ello, en esta etapa hay que aplicar la lógica funcional con una extraordinaria claridad de ideas y presididos de un criterio jurídico respecto de las dimensiones espacio y tiempo.

Por esto, vamos a revisar qué es lo que un Suceso puede representar para nuestra Base de Datos, suponiendo por abstracción que la Base de Datos está ya construida y en operación, lo que nos ofrecerá una imagen de la relevancia del Suceso y por lo tanto un diagnóstico de la conveniencia de seleccionar el Suceso en cuestión. Veamos por lo tanto qué es lo que un Suceso puede producir:

- Creación de uno o más registros en la Base de Datos.
- Creación de estructuras lógicas en la Base de Datos.
- Integración de registros a estructuras lógicas existentes.
- Actualización de registros.
- Actualización de estructuras lógicas.
- Cancelación de registros.

- Cancelación de estructuras lógicas.
- Determinación de procesos de síntesis de información.

Considerando un determinado Suceso, y a la vista de qué actividad de las que hemos señalado va a desencadenar en la Base de Datos, supuestamente ya operativa, tendremos que formarnos una idea cualitativa de en qué medida dicho Suceso es relevante y por lo tanto de en qué medida debemos formalizarlo ya como Suceso “a aislar”.

Si repasamos las incidencias que hemos señalado puede un Suceso producir, cabe distinguir dos diferentes naturalezas:

- Las relativas a creación, actualización o cancelación de registros.
- Las consecuencias inducidas de dicha modificación de registros.

Estas dos naturalezas de las incidencias nos vienen a señalar dos diferentes medidas de la relevancia de los Sucesos:

- Un Suceso es relevante en la medida en que tenga en la realidad una correlación en la Base de Datos, produciendo actividad informática de modificación de registros.
- Un Suceso es relevante en la medida en que la modificación de registros que provoca en la Base de Datos desencadene efectos inducidos subsiguientes, modificando por un lado las relaciones lógicas y provocando por otro la ejecución de programas de proceso.

El primer criterio de relevancia podrá ser posteriormente reconsiderado al estudiar la relación Datos/Actividad y los efectos inducidos que en cada Suceso provoca, podrán volver a analizarse al determinar la Matriz de Sucesos.

En nuestro ejemplo, una vez realizado este proceso aislamos los siguientes Sucesos como relevantes para nuestros objetivos:

- Sucesos que tienen lugar en los Centros de Actividad tipo *X*, *Y* o *Z*.
 1. Venta de un billete. Momento en que se formaliza un contrato de transporte aéreo con la consiguiente transacción económica.
 2. Reembolso de un billete. El cliente devuelve un billete y recupera su dinero.
 3. Cierre en situación *OK* de un billete. La Oficina de Ventas incluye en el billete del pasajero la condición de reserva confirmada (*OK*) para un determinado vuelo programado. El billete pudo haber sido vendido en ese momento o con anterioridad.
 4. Cierre en situación *RQ* de un billete. Caso análogo al anterior, con la diferencia de que la situación *RQ* no supone la reserva confirma-

da, sino que recoge el hecho de que el pasajero está en lista de espera para un determinado vuelo.

— Sucesos que tienen lugar en el Centro de Actividad *R*.

5. Se realiza una reserva para un determinado cliente y para un determinado vuelo.
6. Se cancela una reserva previamente efectuada.
7. Se sitúa a un pasajero en lista de espera de un determinado vuelo.
8. Se pone un vuelo en situación de *CL*, situación en la cual solamente admite lista de espera.
9. Se pone un vuelo en condición *CC*, situación por la cual ni admite reservas ni admite lista de espera.
10. Se pone un vuelo en situación *AS*, lo que supone eliminar una situación previa de *CL* o *CC*.
11. Se reconfirma una reserva. Es un procedimiento que las Compañías Aéreas utilizan en determinadas rutas aéreas.
12. Se solicita al Centro de Actividad *P* o *C* una modificación de oferta programada para una determinada fecha y para unos determinados destinos.

— Sucesos producidos en los Centros tipo *F*.

13. Apertura del proceso de facturación en un vuelo determinado. A partir de ese momento son admitidos los pasajeros en el Aeropuerto para el vuelo en cuestión.
14. Aceptación de un pasajero al embarque. Al pasajero en cuestión se le perfecciona el contrato de transporte acreditándole para subir al avión mediante la denominada tarjeta de embarque.
15. Entrega de la tarjeta de stand-by. Al pasajero se le ha provisto de dicha tarjeta, que le acredita con una determinada prioridad como componente de la lista de espera.
16. Se deniega el embarque a un pasajero.
17. Cierre de facturación en un vuelo. A partir de ese momento el vuelo no admite ya nuevos pasajeros.

— Sucesos que tienen lugar en el Centro de Actividad *P*.

18. Programación de vuelos para una temporada. Las temporadas clásicas que las Compañías aéreas contemplan son dos anuales, una cubriendo el período de 1 de Abril a 1 de Noviembre y la otra la época restante.

19. Se programa una línea.
 20. Se cancela una línea.
 21. Se modifica una línea programada.
 22. Se modifica la programación de una temporada.
 23. Se programa un vuelo
 24. Se cancela un vuelo.
 25. Se modifica la programación de un vuelo.
 26. Se deniega una petición de modificación de oferta.
 27. Se determina una rotación de avión. Se designa como rotación de avión en la jerga del transporte aéreo a un avión que pudiéramos llamar “avión lógico”, que enlaza una sucesión de servicios en secuencia con la garantía de poder ser operados con una unidad de avión.
 28. Se modifica una rotación de avión.
- Sucesos que tienen lugar en el Centro de Actividad *M*.
29. Se planifican períodos de revisión de aviones. Estos Sucesos condicionan el producto a medio y largo del Centro *P*.
 30. Se modifica la programación de revisión de aviones.
 31. Se determinan los períodos operativos de aviones.
- Sucesos que tienen lugar en el Centro *T*.
32. Se constituyen rotaciones de tripulaciones. Análogamente al Suceso 27 se diseñan aquí unas “tripulaciones lógicas” que permitan componer una secuencia viable que posteriormente será asignada a personas físicas.
 33. Modificación de rotaciones de tripulaciones.
 34. Se asigna una tripulación a un vuelo.
- Sucesos que tienen lugar en el Centro *C*.
35. Se suspende un vuelo. Este Suceso sólo puede tener lugar con una antelación no superior a 24 horas de la fecha programada de salida.
 36. Se decide hacer un vuelo de situación. Estos vuelos tienen por objeto situar un avión en un determinado Aeropuerto para iniciar allí un servicio.
 37. Se cambia el tipo de avión en un vuelo. También con una antelación no superior a 24 horas.

38. Se decide hacer un vuelo en Ferry. Se trata de un vuelo sin pasajeros.
39. Se decide sobrevolar una escala programada.
40. Se decide el aterrizaje de un avión en un Aeropuerto alternativo al programado.
41. Se decide el regreso del avión al Aeropuerto de salida.
42. Se decide que un avión realice una escala no programada.

— Sucesos que tiene lugar en el Centro *D*.

43. Se confecciona la hoja de Carga y Centrado para un determinado vuelo. Esta hoja recoge las componentes de pasaje, carga, combustible, etc., que lleva un avión, suponiendo un visto bueno a la operatividad del avión en esas circunstancias.
44. Se despacha un vuelo. Centrémonos en el Suceso informático de la emisión de un mensaje de salida con la información de interés acerca de la salida del avión.

— Sucesos que se producen en el Centro *L*.

45. Recepción de un vuelo que llega. Centremos el Suceso informático de emisión de un determinado mensaje que recoja aspectos informativos relevantes.

Una vez compuesta esta lista de Sucesos resulta conveniente volver al punto 3.2.2 y reflexionar nuevamente repasando cómo una lista de comprobación, nuestro conjunto de Sucesos, confirmando que no existen actividades que sean de interés y que no estén contempladas o, en caso contrario, introduciendo las oportunas modificaciones en el modelo que poco a poco vamos perfilando.

3.3. ESTUDIO DE LA RED DE PRODUCCION

Una vez que han sido determinados en principio los Centros de Actividad y los Sucesos Productivos relevantes, estamos en condiciones de componer la Red de Producción como instrumento soporte del modelo que vamos a extraer. Para componer la Red de Producción conviene partir en primer lugar de los Centros de Actividad determinados, uniéndolos con trayectorias geográfico-temporales, y prescindiendo de momento de los Sucesos. El producto resultante nos ofrecerá la referencia de base de cómo los Centros de Actividad considerados se relacionan unos con otros bajo el punto de vista de la consecución de productos.

En la figura 40 se recoge la Red básica de producción del ejemplo que venimos exponiendo. Los Centros de Actividad tipo *X*, *Y* y *Z* son los que ponen en movimiento un determinado flujo de pasajeros. Independientemente de que los pasajeros decidan volar o compren su billete en estos Centros, existirá por lo general un mecanismo o función por medio del cual se administra la oferta, y se realizan reservas para los pasajeros, en función de la disponibilidad de plazas para los distintos vuelos y fechas de servicio, lo cual viene a constituir la actividad del centro *R*, el cual según se representa en la figura 40, produce listas de pasajeros reservados para cada uno de los vuelos programados, lo que constituye la distribución y aporte de pasajeros a los Centros *S* que son los Aeropuertos.

El Centro *R*, realiza la administración de la oferta de acuerdo con el producto que obtiene el Centro *P*, el cual es el órgano definidor de la programación de vuelos. El Centro *P*, como órgano responsable de la programación de vuelos, es el Centro que determina y eventualmente modifica la oferta, pero en el momento que para un determinado vuelo programado faltan exactamente 24 horas para que se realice su salida, dicho Centro *P* deja de ejercer poder de decisión en todo lo relativo a dicho vuelo, ya que en las últimas 24 horas, el control de la oferta de vuelos es ejercido por el Centro de Actividad *C*. El Centro de Actividad *C*, recibe análogamente del Centro *T* la programación de tripulaciones y asimismo recibe del Centro *M* los aviones en condiciones de iniciar los vuelos. Los Centros *P*, *T*, *M* y *C* son únicos.

Los Centros de Actividad tipo *S* son varios, existiendo tantos como Aeropuertos en los que la Compañía opere, y dentro de cada Centro *S* distinguimos otros tres Centros de Actividad de nivel inferior: el Centro *F*, destinado a la recepción de pasajeros; el Centro *D*, destinado a la actividad técnica de despacho del vuelo, y el Centro *L*, de recepción del vuelo a la llegada.

En la Red de Producción representada en la figura 40 se recoge por un lado el ciclo total de los pasajeros, y por otro lado el ciclo que siguen los aviones, y por tanto también las tripulaciones. El ciclo de pasajeros transportados tiene a primera vista cuatro puntos básicos en los que se producen “contactos” relevantes de los pasajeros con la Compañía:

- El contacto con el punto de venta *X*, *Y* o *Z*.
- El contacto con el Centro de Reservas *R*.
- El contacto con el Centro *F* de recepción de pasajeros en Aeropuertos.
- El contacto con el Centro *L* por medio del cual se completa el ciclo.

El ciclo de aviones lo suponemos controlado por el Centro de Actividad *C*, siendo los puntos importantes a tener en cuenta para todo el planteamiento de la Red de Producción, los que se derivan del contacto del avión con los Centros *D* y *L*, en los que se produce la salida y la llegada.

En la figura se reseña con líneas de trazos la indicación de la existencia de diferentes Aeropuertos o Centros tipo *S*, a través de los cuales se produce tanto el ciclo de pasajeros como el ciclo de aviones y tripulaciones.

Una vez compuesta la Red de Producción, relacionando los Centros de Activi-

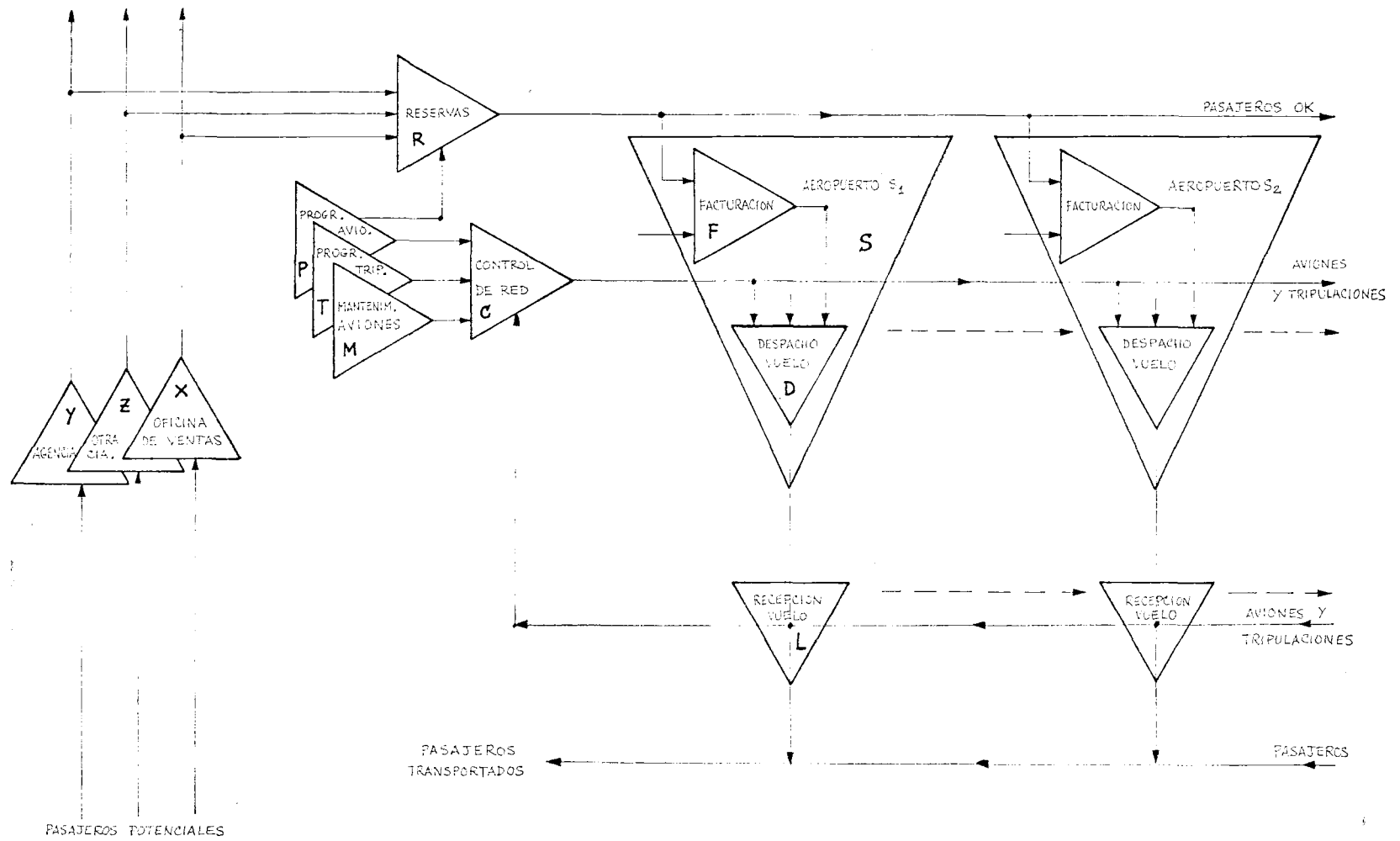


FIG. 40. Red de producción

dad, podemos comenzar a tomar en cuenta la relación de Sucesos que hemos determinado en un principio, aunque también conviene realizar un examen previo de dichos Sucesos y de las relaciones que entre ellos se producen, antes de profundizar más en el análisis de la Red de Producción.

3.3.1. LA MATRIZ DE SUCESOS

La Matriz de Sucesos es un instrumento de análisis de cómo los Sucesos están interrelacionados unos con otros, puesto que en la actividad productiva los umbrales no son independientes, produciéndose en la realidad un proceso de inducción por medio del cual el hecho de que tenga lugar un Suceso productivo supone el que posteriormente tengan lugar otros Sucesos inducidos. El estudio de estas repercusiones de Sucesos en cadenas es el objeto de la Matriz de Sucesos.

En la figura 41 se representa la Matriz de Sucesos de nuestro ejemplo. En la matriz representamos tantas filas y columnas como sucesos relevantes hemos considerado en un principio y tal como se muestra en la figura, se recoge la numeración que a cada suceso hemos asignado y el Centro de Actividad en el que cada suceso tiene lugar. Hemos representado las líneas de trazo grueso para delimitar los grupos de sucesos que corresponden a diferentes Centros de Actividad.

Una fila de la matriz corresponderá a un determinado suceso productivo, reseñándose en dicha fila, mediante un pequeño círculo, aquellos sucesos que pueden tener lugar con posterioridad al suceso del cual parte la fila, siendo inducidos por él. Los círculos se sitúan en las columnas correspondientes a los sucesos inducidos.

De esta forma, cada fila nos establece cuáles son los sucesos inducidos de uno determinado y viceversa. Cada columna nos indicará cuáles son los sucesos que pueden inducir a un suceso determinado. Así, por ejemplo, el suceso número 23, que corresponde a la programación de un vuelo, tiene como sucesos consecuentes los siguientes:

- 5. Se realizan reservas para dicho vuelo.
- 13. Se abre el vuelo a facturación en un determinado momento, a menos de que previamente fuera cancelado.
- 24. Eventual cancelación de dicho vuelo.
- 25. Eventual modificación en la programación de dicho vuelo.
- 34. Asignación de una tripulación para dicho vuelo.
- 35. Eventual suspensión del vuelo en las últimas 24 horas.
- 37. Eventual cambio del tipo de avión.
- 39. Posible sobrevuelo de una escala programada.
- 40. Posible aterrizaje en un Aeropuerto alternativo.
- 41. Posible regreso del avión al Aeropuerto de salida.

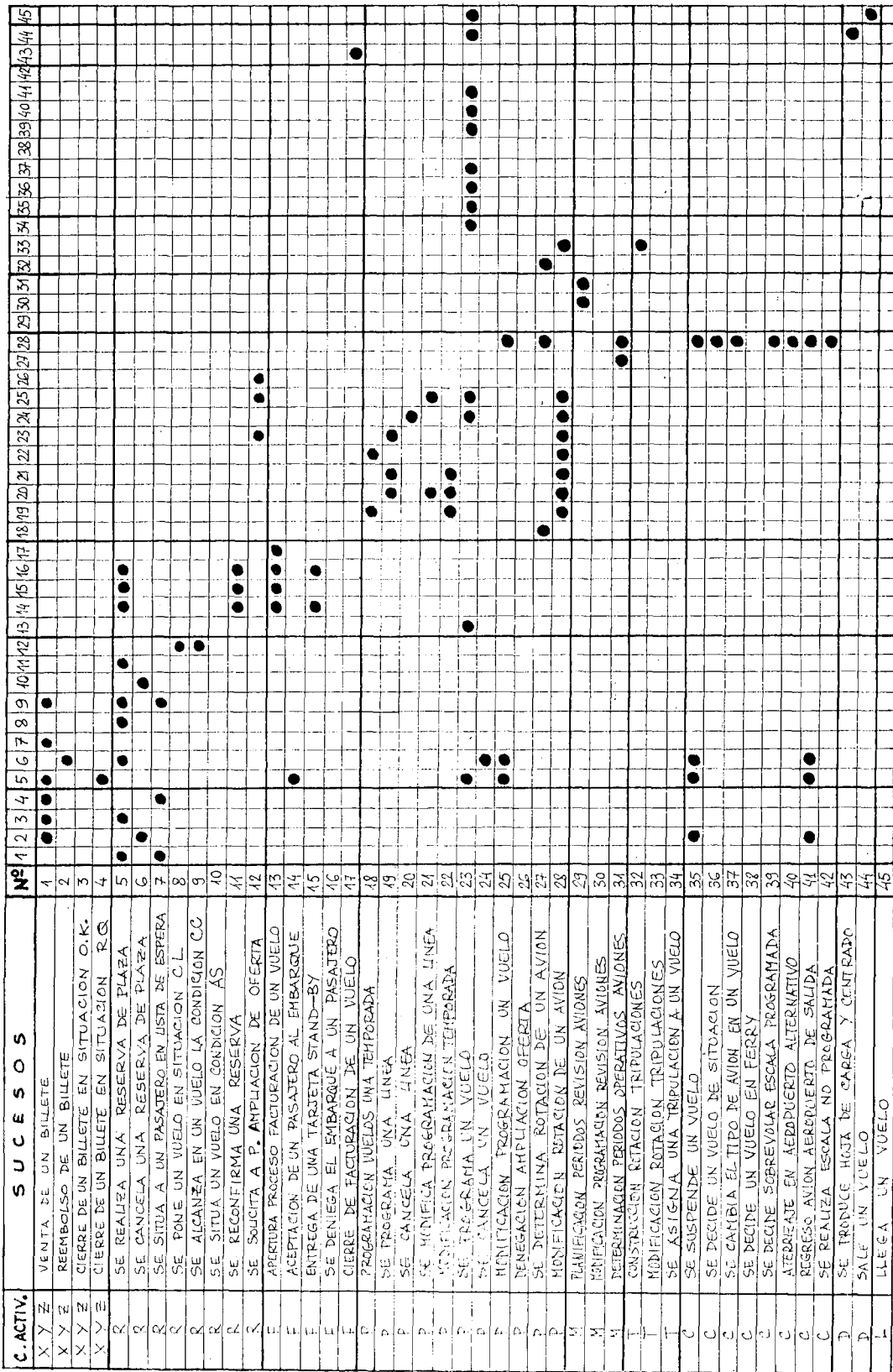


Fig. 41. Matriz de sucesos

- 44. Salida de un vuelo. Esto se producirá siempre, a menos que haya habido cancelación o suspensión.
- 45. Llegada normal de un vuelo.

Hay que distinguir las diferentes posibilidades a la hora de reseñar una relación en la Matriz de Sucesos:

- El suceso inducido puede ser obligatoriamente inducido, o bien tendrá una cierta probabilidad de producirse. Por ejemplo, siempre que se produce el suceso 13 se inducirá necesariamente el 17, pero sin embargo el 16 no habrá necesariamente de producirse, existiendo una determinada probabilidad de que ocurra, y pudiendo interesar el controlar el valor de dicha probabilidad en función de otros factores desencadenantes que pudiéramos identificar en la Red de Producción.
- El suceso inducido no tiene por qué ser directamente inducido. Podemos reseñar una inducción indirecta porque la consideramos relevante. Por ejemplo, el suceso 41 no es inducido directamente por el suceso 23, y sin embargo recogemos la relación 23-41 por identificar todos los sucesos derivados del suceso 23, que complementen el conjunto total de hechos que con un vuelo programado pueden ocurrir.
- La Matriz de Sucesos es un instrumento de análisis y podemos recoger relaciones entre sucesos de antecedente a consecuente, sin que existan claramente factores de inducción. Por ejemplo el suceso 14, no tiene por qué inducir en esencia el que se produzca el suceso 5, pero, sin embargo, dicha relación de antecedente a consecuente, bien pudiera ser relevante para nuestro modelo, recogiéndose la posibilidad del hecho de que un pasajero acuda al Aeropuerto directamente y a posteriori se le efectúe una reserva por el hecho de que tenga vuelos posteriores en conexión, cuyo registro interese mantener en la Base de Datos hasta que el pasajero complete su ciclo de viaje.
- El criterio de relevancia siempre estará presente, no siendo preciso el que sean recogidas todas y cada una de las relaciones que en la realidad se producen, sino todas aquellas que contienen significación de hechos que nos interesa analizar o modelizar.

La construcción de la Matriz de Sucesos es por lo tanto un suceso creativo y deberá realizarse aplicando buenos criterios, siendo insuficiente la estricta aplicación de la metodología.

La Matriz de Sucesos nos establece la lógica de relación entre los diferentes tipos de suceso que hemos considerado, y constituirá un criterio lógico a manejar a la hora de profundizar en el análisis de la Red de Producción, puesto que la Red de Producción está configurada por el flujo de productos que los Centros de Actividad se intercambian y estos flujos son generados a partir de los sucesos. Las relaciones entre sucesos será nuestra guía para analizar la relación entre flujos de productos.

3.3.2. DIAGRAMAS DE SUCESOS Y ANALISIS DE LA RED DE PRODUCCION

Una vez que disponemos de la red básica de producción y de la matriz de sucesos, estamos en condiciones de analizar al nivel de profundidad que precisemos, la fenomenología de la Red de Producción. Para ello seguiremos dos etapas:

- Estudio de la Red de Producción sirviéndonos del encadenamiento de sucesos que hemos establecido mediante la Matriz de Sucesos, pero sin aplicar la regla distributiva de los sistemas, profundizando en el análisis de Centros de Actividad de nivel inferior a los que en un principio hemos considerado. Se presenta un ejemplo en la figura 42.
- Estudio de las necesidades de derivación de flujos, lo que puede traducirse en penetrar en la Red de Producción a niveles más profundos de análisis, para lo cual habremos de aplicar la propiedad distributiva de los sistemas. Sirve de ejemplo la figura 31.

En la figura 42 representamos lo que podemos denominar como Diagrama de Sucesos. Manejando la Red básica de Producción y la Matriz de Sucesos, podemos construir diferentes Diagramas de Sucesos como el de la figura 42, pues con un diagrama responderemos a una determinada motivación de análisis. La figura 42 nos ha motivado a analizar el flujo de aviones y reflexionar sobre la problemática de información y control que dicho flujo de aviones plantea, para poder después determinar el modelo informático que de dicho fenómeno nos interesa manejar, y extraer por lo tanto los oportunos registros de información. Así se recoge la línea de sucesos 18, 19, 23, 43, 44 y 45 como línea básica que recoge el siguiente encadenamiento de acontecimientos relevantes:

- Programación para toda la temporada.
- Programación de una determinada línea.
- Programación de un determinado vuelo.
- Preparación técnica final para la operación del vuelo (hoja de carga y centrado).
- Salida del vuelo.
- Llegada del vuelo.

Naturalmente, esta línea básica de sucesos no es única, sino que existen otras derivaciones de flujo que hemos contemplado en la Matriz de Sucesos, por lo que se representan los sucesos 22, 21, 25 o bien los 20 y 24, ... etc., y, en general, todos aquellos flujos de productos derivados de los sucesos que tengan relación más o menos directa con el objeto formal de nuestro proceso de análisis. Por ejemplo, aun cuando nuestra motivación sea el modelizar el fenómeno del ciclo de aviones, no puede estar ausente la relación 23-5, puesto que aunque la situación en el tiempo del suceso 23, que corresponde a la programación de un vuelo, no sea relevante con relación al movimiento estricto de aviones, sí es de importancia para la actividad de distribución de la oferta que realiza el Centro de Actividad *R*. Dicha relación

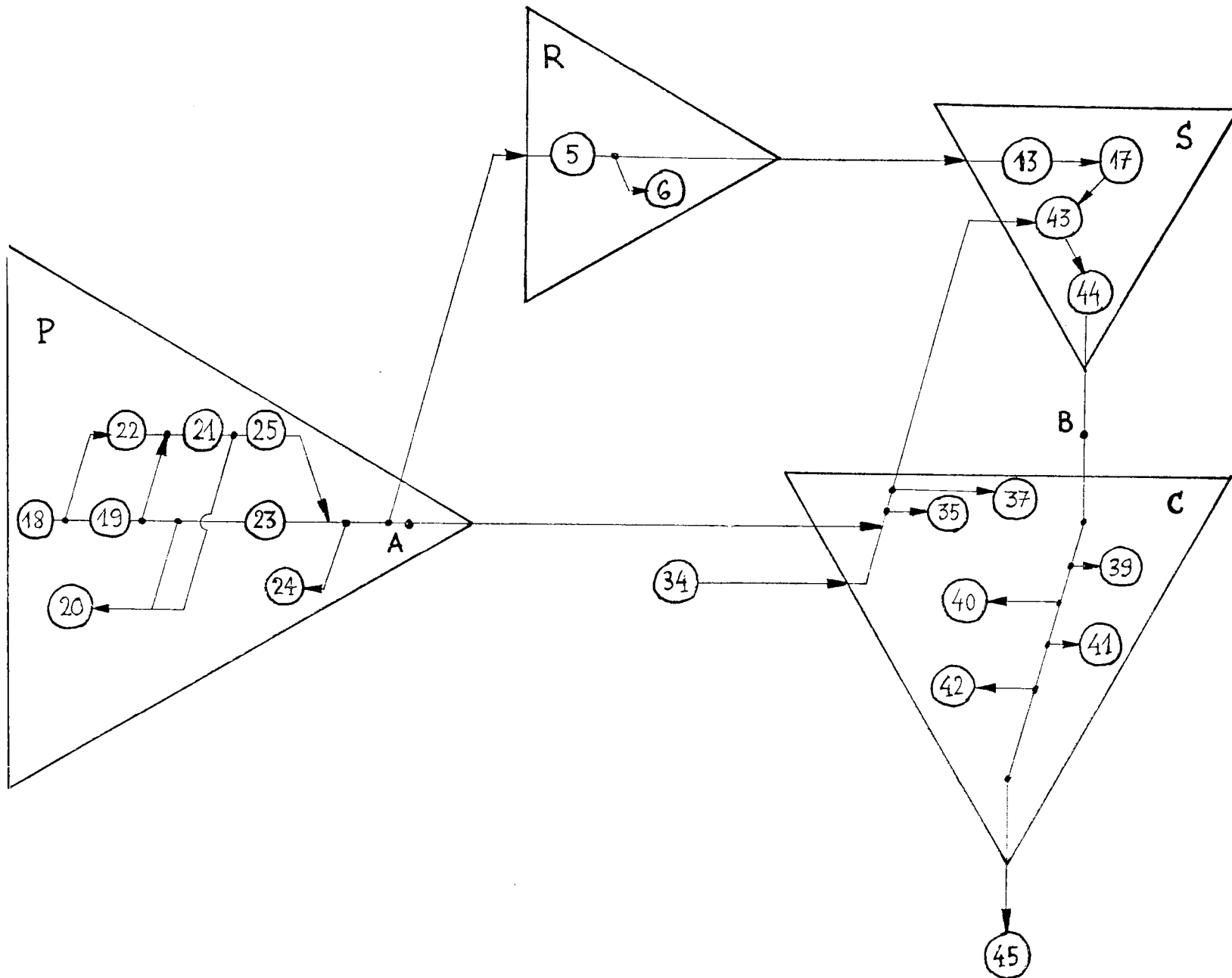


FIG. 42. Diagrama de sucesos

23-5 interesa sea siempre contemplada y en este caso será un factor condicionante en el análisis del ciclo de aviones y en las consecuencias que de él se deduzcan.

3.3.3. DERIVACIONES E INTEGRACIONES DE FLUJOS

Un Diagrama de Sucesos nos clarifica las cadenas de flujos de productos en el espacio y en el tiempo, pero a la hora de considerar la información asociada a los flujos de productos, los diagramas de sucesos resultan insuficientes.

Por ejemplo, remitiéndonos a la figura 42, por el punto indicado con la letra *B*, pasa el flujo de aviones en vuelo, que está generado por el suceso 44 mediante el cual se produce la salida de los aviones de los Aeropuertos.

Dicho flujo que pasa por el punto *B* tiene diferentes matices acerca de los cuales habrá que recoger datos para componer debidamente nuestro modelo, lo cual supone que consideremos el flujo subdividido en todo un conjunto de componentes que constituyen derivaciones paralelas de flujo, al igual que un cable eléctrico está generalmente subdividido en otros cables más elementales.

En el punto *B* habremos de contemplar:

- Coordenadas espacio-temporales de producción del suceso 44, lo que constituya los datos relativos al origen del flujo.
- Datos correspondientes a los pasajeros que van dentro del avión: número de pasajeros en primera clase, número de pasajeros en clase turista, número de niños, etc.
- Datos relativos a la tripulación.
- Datos relativos al avión como su matrícula, plan de vuelo, etc.

Estas derivaciones de flujo suponen el aproximarnos a un nivel de análisis más detallado, y habremos de profundizar siempre hasta un grado en el cual no dejemos ningún detalle relevante por considerar.

En este período de análisis podemos ir considerando también de forma preliminar cómo vamos a orientar la captación de la información. Así, por ejemplo, para el caso de la información asociada al flujo que pasa por el punto *B*, se nos sugiere inmediatamente la idea de emitir un mensaje telegráfico cada vez que se produce el suceso 44, con lo cual podrá reunirse en el ordenador, en tiempos mínimos, toda la información relativa al movimiento de aviones y la información asociada del tráfico de pasajeros en cifras cuantitativas.

También habremos de tener en cuenta a la hora de examinar los Diagramas de Sucesos las integraciones de flujo que informáticamente sean relevantes para garantizar que estamos contemplando todos los datos precisos para que dichas integraciones sean posteriormente viables. Así, por ejemplo, el flujo que pasa por el punto *A*, continuando con la figura 42, tiene diferentes matices informáticos:

- Tramos y segmentos que componen cada vuelo (los tramos son los recorridos entre escalas que realiza el avión, y los segmentos, los diferen-

tes pares de escalas que pueden constituir los puntos de embarque y desembarque de un pasajero potencial).

- Denominación de la línea.
- Fecha del vuelo.
- Etcétera.

El dato correspondiente a la fecha es, por ejemplo, un factor esencial a la hora de integrar diferentes vuelos correspondientes a una misma línea. Los datos correspondientes a los tramos serán de interés para integrar las operaciones que salen o llegan a un determinado Aeropuerto. Los datos que identifican los segmentos servirán para integrar los datos correspondientes a un determinado mercado o zona geográfica en la que operamos.

Por lo tanto, a partir del Diagrama de Sucesos, pensaremos por un lado en la segregación de información o, lo que es lo mismo, en la derivación de flujos, y por otro lado, en la sintetización de datos elementales o, lo que es lo mismo, en la integración de flujos.

A partir de aquí estaremos ya en condiciones de determinar dónde vamos a poner “contadores de flujo” o, lo que es lo mismo, de determinar qué datos elementales vamos a incorporar a nuestros registros, lo que nos sitúa en el momento de realizar el diseño de registros y sus relaciones, y por lo tanto en la confección del Esquema de la Base de Datos.

3.4. DISEÑO DEL ESQUEMA

El Esquema es el eje o elemento fundamental de la Base de Datos y sus características condicionan todo lo que del Sistema vayamos a obtener. Si realizamos un buen diseño del Esquema llegaremos a obtener un buen Sistema de Base de Datos en la medida en que la técnica informática de desarrollo del software que apliquemos sea igualmente buena. Por el contrario, si el Diseño del Esquema es defectuoso, el Sistema de Base de Datos resultante será igualmente defectuoso, con independencia de que la técnica de proceso de datos aplicada sea más o menos correcta.

Por estas razones, el planteamiento del Esquema general de la Base de Datos constituye la definición de la infraestructura general de la Base y habrá de abordarse con la mayor seguridad y en las mejores condiciones. Esto nos obliga a que llegado este momento tengamos muy claros los conceptos a manejar y hayamos estudiado la actividad de la Organización desde el plano más general y habiendo profundizado hasta los niveles de análisis más profundos que hayamos estimado como relevantes. Los objetivos de presente y de futuro, es decir, la utilidad que de la Base de Datos se espera, tienen que estar igualmente enfocados con total claridad.

En estas condiciones abordamos el diseño del Esquema como proceso fundamental de creación, proceso que una vez más habrá de ser dirigido a la luz de la lógica, manejando criterios más que mediante la aplicación estricta de un método. El Esquema será por lo tanto un fruto de nuestro conocimiento de la actividad enfrentado a las utilidades que hemos de obtener, y nuestro proceso creativo irá consolidando el diseño del Esquema por medio de aproximaciones sucesivas, manteniendo una actitud de crítica en cada aproximación de diseño mediante la consideración de los criterios fundamentales a manejar.

El buen sentido analógico y la buena práctica informática nos ayudarán a proponer un determinado Esquema y a continuación habremos de ejercer nuestra capacidad de crítica, e incluso someterlo a la crítica de los analistas que posteriormente hayan de desarrollar aplicaciones utilizando la Base de Datos. De esta forma, el Esquema se irá perfeccionando, aunque el producto que obtengamos en esta etapa del proyecto no habrá de ser considerado como definitivo. En primer lugar, porque el diseño del Esquema es una actividad abierta que va recibiendo nuevas críticas y, en consecuencia, está en condiciones de ser mejorado, a medida que el Sistema en conjunto está en desarrollo. En segundo lugar, porque aun cuando el Esquema deberá tener una estabilidad considerable durante la vida operativa del Sistema de Base de Datos, dicho Esquema es susceptible de evolucionar durante la vida útil de la Base de Datos, y más concretamente será normalmente ampliado con nuevas proposiciones de objetivos.

Los criterios fundamentales a tener presentes a la hora de mantener una actitud crítica para enjuiciar el nivel de suficiencia alcanzado por el diseño de Esquema que proponemos son los siguientes:

- *Modelización.* El diseño del Esquema es correcto en la medida en que modeliza la actividad real de la Organización, por lo que estará orientado a las funciones y sucesos que en la práctica tienen lugar. Para juzgar en qué medida el Esquema modeliza a la Actividad, habrán de tenerse presentes los conceptos de la Teoría de Sistemas y el conocimiento de las particularidades que hemos analizado mediante el estudio de la actividad. Para que el modelo sea correcto habrá de ser posible reconstruir a partir de él, todos aquellos aspectos de la Actividad que sean relevantes para nuestros objetivos.
- *Relevancia.* El Esquema diseñado habrá de ser “enfrentado” de una forma metodológica y exhaustiva a la utilización que de dicho Esquema vamos a realizar, al conjunto de aplicaciones o programas que tenemos planificados o pensamos desarrollar en el futuro utilizando la Base de Datos, o lo que es lo mismo, al conjunto de objetivos que tenemos planteados. De este contraste debemos de obtener conclusiones acerca de la relevancia de los datos y relaciones lógicas contenidas en el Esquema. Es decir, debe de detectarse la carencia de datos o relaciones que sean relevantes, y asimismo deben de detectarse los elementos o relaciones que de momento hemos incluido, pero que son innecesarios o están contraindicados.

- *Homogeneidad y consistencia.* Las agrupaciones de Datos habrán de estar realizadas de forma homogénea, atendiendo a las tasas de actividad de los datos, es decir, a sus frecuencias de diferentes tipos de utilización, existiendo al menos compatibilidad cuando no existe homogeneidad.

Es importante a este respecto la crítica para cada Fichero de donde hemos situado la frontera en cuanto a la agrupación de datos que dicho Fichero supone. Es decir, el hecho de configurar un Registro Tipo o Fichero equivale a diseñar una estructura interna de datos que son inseparables. Si al situar la frontera hemos descendido demasiado, tendremos mayor diversidad de registros y mayor cantidad de relaciones lógicas entre ellos. Si hemos agrupado datos en exceso puede ocurrir que por el contrario, diseñemos un Esquema con poca diversidad de registros y, por lo tanto, muy estructurado. Es esencial ejercer el sentido crítico y juzgar cuál ha de ser el punto justo a alcanzar en la agrupación de datos formando estructuras inseparables dentro de los registros.

- *Flexibilidad y profundidad del modelo.* El modelo deberá permitir distribuir y asociar las corrientes de flujo que hayamos determinado como relevantes al realizar el estudio de la Actividad y en todas esas características relevantes, el modelo no debe perder riqueza de matices respecto de la dinámica real de la Organización.

Habrá que juzgar con cuidado hasta qué punto estamos sintetizando la información o hasta qué punto habremos de conservar la información actualizada. Estos puntos habremos de situarlos en el ámbito geográfico y en el ámbito temporal, puesto que las síntesis de información consisten en agrupar flujos en el espacio y/o en el tiempo. El mantenimiento de los datos a nivel muy atomizado da ciertamente mayor flexibilidad al modelo, pero sin embargo está en oposición al criterio de rentabilidad por requerirse mayor espacio de almacenamiento. Por eso, para analizar este criterio conviene tener presente un tanteo previo del dimensionado, según indicamos en el punto 3.4.3, analizando para los datos en qué medida el sintetizar representa una economía clara frente a la falta de flexibilidad que produce.

- *Perspectiva temporal.* Una utilidad general de la Base de Datos es la capacidad que ofrece para poder establecer correlaciones en tiempo real o al menos en tiempos mínimos, con series temporales de datos o parámetros, e incluso con series de datos que se prevén para el futuro con parámetros que se plantean como objetivos a alcanzar. Por ello la dimensión histórica del modelo debe tener un horizonte suficiente para permitir toda la tipología de correlaciones que vayan a ser de utilidad en el uso presente y futuro del Sistema.

El nivel de detalle de la información será por lo general diferente en función de la actualidad del dato y así, dentro de esta perspectiva temporal habremos de criticar y reconsiderar hasta qué profundidad en la historia mantenemos los datos con su capacidad total de detalle y asimismo a qué profundidad en el tiempo realizamos los distintos procesos de

síntesis. Del manejo de este criterio reconsideraremos fundamentalmente los registros de síntesis y sus relaciones lógicas en el Esquema.

- *Rentabilidad.* Este criterio estará obviamente siempre presente porque no podemos disociar los objetivos que nos proponemos de los niveles de coste en los que prevemos situarnos, ya que el proceso creativo de diseño del Esquema ha de conducirnos a un producto viable. La aplicación de este criterio nos obligará ciertamente a renunciar a algunas características, ya que los recursos en la práctica siempre tienen limitación. En todo caso, si las consideraciones económicas fueran a descalificar alguna potencia del modelo que se considere esencial o especialmente relevante, siempre podrá en ese momento plantearse la correspondiente alternativa a los niveles superiores de decisión en el proyecto.
- Los criterios generales de la lógica de Sistemas habrán de estar por supuesto también presentes.

Dentro de este planteamiento, para cebar nuestra actividad creativa, repasaremos la lista de Sucesos que previamente hemos seleccionado, y consideraremos a todos estos Sucesos como fenómenos generadores de información, y como tales habremos de criticar, suceso por suceso, si nuestro modelo prevé la inclusión de todos los datos que el suceso genera. Este ejercicio viene a perfeccionar la aplicación del criterio de relevancia indicado anteriormente.

También debemos realizar un examen crítico atendiendo a la lista de Centros de Actividad que hemos considerado en nuestro análisis, contemplando cada Centro como elemento fecundador de la actividad. De aquí debemos obtener conclusiones de las lagunas de nuestro modelo en lo relativo a los datos de tipo estructural, asociados a estos Sistemas Productivos.

Las estructuras orgánicas de referencia que hayamos manejado como fundamentales a la hora de estudiar la actividad también deben de ser contempladas a la hora de enjuiciar nuestro modelo. Estructuras geográficas, de agrupaciones territoriales o estructuras jerárquicas de órganos de la empresa, etc. Ello equivale a confrontar la estructura de flujos de la Red de Producción, con las potencias informáticas de nuestro modelo.

Como un elemento más de juicio, se impone el realizar un tanteo de dimensionado de los volúmenes de almacenamiento a los que conduce el Esquema que proponemos, para que de esta forma, tanteemos desde el principio un ajuste de Esquema de tipo práctico. De este tanteo de los volúmenes de almacenamiento ejerceremos el criterio de rentabilidad y en cualquier caso tendremos valorada la incidencia de los diferentes tipos de datos en relación con este criterio, que de esta forma podrá ser ejercido de forma selectiva.

De acuerdo con todo este planteamiento, y siguiendo con nuestro ejemplo, nos hemos propuesto el Esquema cuyo gráfico se recoge en la figura 43. En él se reflejan desde R-1 hasta R-30, los 30 tipos de registros que hemos considerado, registros que han sido concebidos contemplando todo su contenido de información y

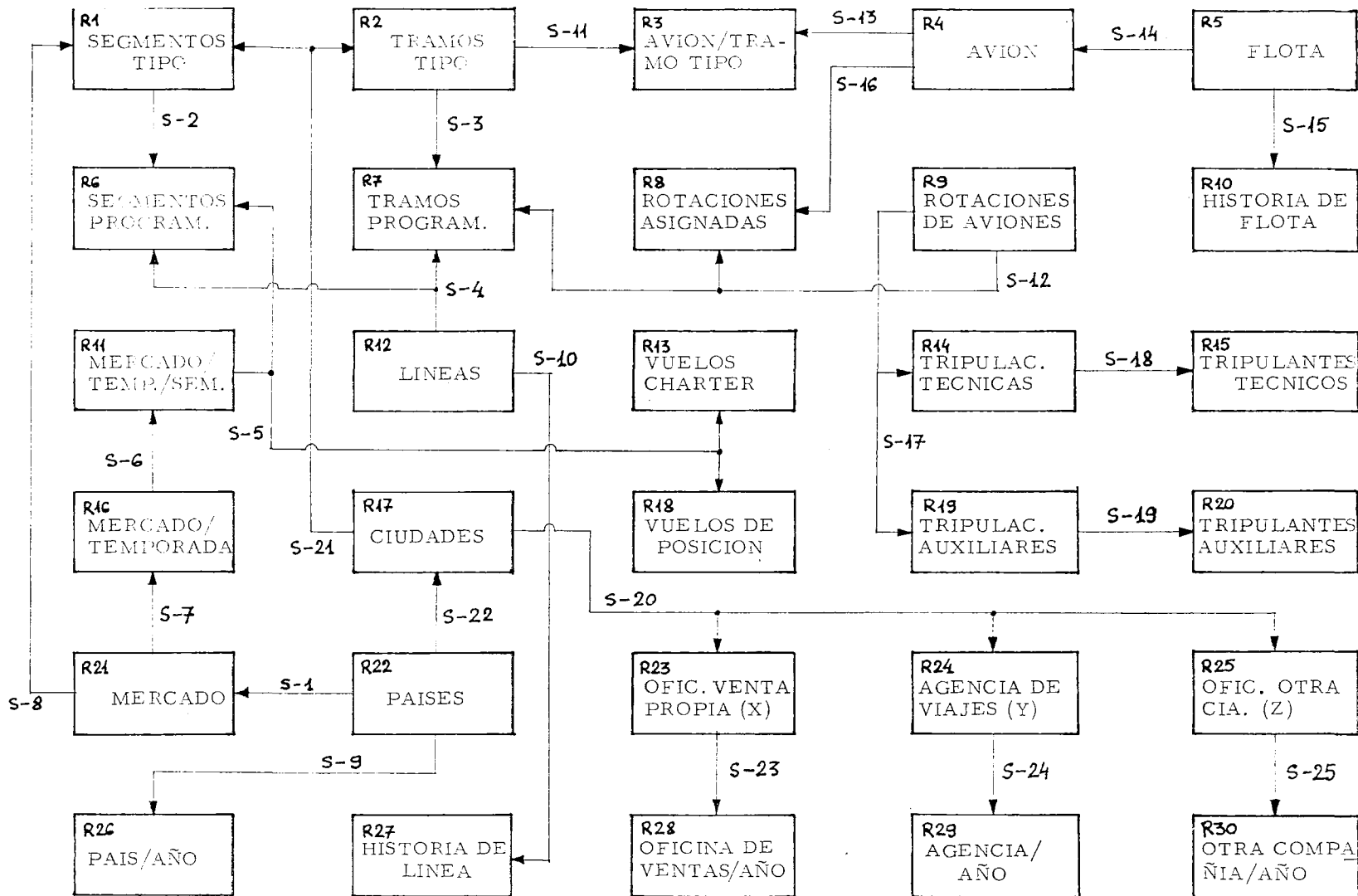


FIG. 43. Esquema

no simplemente la denominación que en dicho gráfico figura. El contenido cualitativo de cada Registro Tipo no se ofrece de forma exhaustiva por no juzgarlo de mayor interés frente a la extensión que representa, aun cuando sin embargo será más adelante cuando comentado en aquellos casos en los que nos aporte alguna forma de conocimiento generalizable.

De la misma forma se representa en el gráfico, desde S-1 hasta S-25, los 25 Sets Tipo que el Esquema contiene, algunos de los cuales serán comentados más adelante.

3.4.1. DISEÑO DE REGISTROS

Conviene tener presente que el diseño de un Registro Tipo es la definición de un elemento “frontera”. La parte interna del registro es una estructura rígida formada por un conjunto de datos que van a ser manejados siempre en bloque, pues solamente en operaciones internas de la memoria central podrá penetrarse dentro de esa estructura. El registro se comporta como una unidad para todas las operaciones de acceso entre la Base de Datos y la Memoria principal y asimismo para toda la problemática de estrategia de almacenamiento. Es decir, la estructura interna del registro es “transparente” a nivel del almacenamiento de datos y sólo es abordable a nivel de Memoria Central.

Como ya hemos apuntado en otros capítulos, las definiciones de los registros condicionan negativamente la flexibilidad del modelo o bien la potencian, en función de lo acertado que sea el diseño; e igualmente se condiciona el equilibrio entre las estructuras internas de los registros y las estructuras externas de relación lógica entre registros lo que incide en la economía del almacenamiento y de los procesos de acceso.

En lo relativo a las relaciones lógicas, que constituyen de hecho el elemento que da fertilidad y potencia a la Base de Datos, el Registro es precisamente el objeto de estas relaciones. La relación lógica hace transparente para los programas de proceso la “distancia” física entre los registros y podremos individualizar los datos tanto como creamos conveniente.

La creación de este elemento frontera que es el registro ha de tener lugar en base a una motivación directa. El formar un conjunto de datos que van a estar rígidamente enlazados en un bloque no es algo que simplemente se tantea o se prepara. Tiene que ser la consecuencia de un proceso creativo y ha de obedecer a un fin que proporcione consistencia y armonía al modelo que estamos configurando. En esta línea de pensamiento repasemos las diferentes naturalezas de registros que anteriormente habíamos expuesto:

- *Registros estructurales ligados a los Centros de Actividad.* Su motivación responde a la conveniencia de agrupar datos que son relativos a características de las entidades que constituyen dichos Centros; quizás porque la asociación entre estos datos y el Centro puede contemplarse

como una relación biunívoca entre dos unidades, el propio Centro y el Registro que agrupará a los datos. El acceso a este registro supondrá el obtener en una sola operación todos los datos relevantes de los elementos internos al Centro de Actividad.

- *Registros estructurales para soporte de los Sucesos.* Responden a la necesidad de tener tipificados y debidamente organizados los elementos que vamos a utilizar para aglutinar los conjuntos de datos que se generan en los sucesos. Los sucesos en sí, al producirse, conducirán a la creación de determinados registros, y éstos habrán de ordenarse con unos criterios. Estos criterios los consolidamos creando unos Registros estructurales que servirán de puntos de acumulación. Al comentar nuestro ejemplo lo pondremos de manifiesto.
- *Registros asociados a los Sucesos y a los Productos.* La consecución de un producto en la Red de Producción, o la generación de un suceso darán lugar a la captación de unos datos que atendiendo a la unidad que conforman en el tiempo y en el espacio pueden aconsejarnos el agruparlos en un bloque único o registro. Este tipo de registros constituye el subconjunto coyuntural del modelo y el elemento que va enriqueciendo la Base de Datos, llevando al almacenamiento las “noticias” que se producen con el ejercicio de la actividad.
- *Registros de síntesis históricas.* Su motivación consiste en dotar a nuestro modelo de la posibilidad de recoger datos sintéticos, referidos a sucesos alejados en el tiempo, que permitan trazar las trayectorias de evolución necesarias a efectos estadísticos y de control de la actividad, así como para cualquier tipo de correlación necesaria que haya de manejar series históricas de datos. Estos registros servirán para recoger datos básicos acumulados cuando se eliminan registros de la Base de Datos que ya no interesan a un nivel analítico debido a su antigüedad, pero que, sin embargo, en conjunto, proporcionan valores de síntesis que nos reflejan circunstancias del pasado que pueden todavía ser objeto de proceso.

Estas cuatro naturalezas de registro vienen a responder a las cuatro necesidades que resumimos, que se corresponden correlativamente.

- Recoger datos de la orgánica de los Centros de Actividad y de sus atributos y características.
- Estructurar adecuadamente la dinámica de Sucesos derivados del ejercicio de la actividad.
- Recoger los datos que en los Sucesos se generan.
- Dotar al modelo de un panorama temporal.

En cuanto a la estructura de los registros, recordemos la tipología de datos que van a constituir sus campos internos:

- Datos relativos a estructuras de entidades, bien sean Centros de Actividad, Sucesos o Productos.
- Coordenadas espacio-temporales de los Sucesos.
- Cuentas de densidad de flujos de la Red de Producción.
- Características y atributos de los Sucesos y Productos.

Siguiendo estas reflexiones vamos a comentar cómo hemos llegado a concebir los registros que se representan en el Esquema de nuestro ejemplo que recogemos en la figura 43, aunque hay que advertir que el diseño de registros no es independiente del diseño de las relaciones lógicas, sino que en realidad ambas definiciones se van realizando en paralelo, aun cuando las estemos exponiendo como etapas correlativas.

Los registros R23, R24 y R25 corresponden a un ejemplo claro de registros estructurales relativos a Centros de Actividad, en este caso correspondiendo a los Centros X, Y y Z. En nuestro proceso creativo, concebimos a cada uno de los Centros tipo X, Y o Z con importancia suficiente como para estar representados en la Base de Datos, con un registro por cada uno de los Centros de cada tipo, recogiendo en dicho registro la información de tipo estructural o los parámetros generales relativos a la actividad que desarrollan. Así, para cada uno de estos Centros, indicaremos los niveles de producción que obtienen, y los objetivos de producción que tenemos fijados, así como el inventario de recursos con que cada uno de estos Centros cuenta, al menos si se trata de recursos propios de la Organización.

Abundando en el significado del registro R23, puede agruparse en su estructura interna, los siguientes datos que relacionamos a título de ejemplo:

- Mercado al que pertenece la Oficina de Ventas.
- Inventario de recursos en unidades físicas y unidades monetarias.
- Volúmenes de producción.
- Objetivos fijados y su desglose y cuantificación.
- Agencias de Viaje cuyas cuentas controla.
- Datos de actividad en cuentas claves de clientes.
- Medidas de la actividad de sus vendedores.
- Volumen de tráfico generado en su área de acción.
- Volúmenes de ingresos producidos.
- Volúmenes de intercambio con otras Compañías.
- Etcétera.

Todos estos campos corresponden a entidades directamente asociadas con la Oficina de Ventas X que da razón de ser al registro y permite englobar el conjunto de información relevante de ese Centro. Por extensión, este mismo registro R23, que lo hemos presentado como registro ligado a un Centro de Actividad, pudiera servir a su vez de registro estructural que sirve para encuadrar otros registros aso-

ciados a los sucesos informáticos. Así, por ejemplo, si el control de la “actividad de los vendedores” entrara en los objetivos de la Base de Datos como una aplicación a desarrollar en mayor profundidad, existirían registros de datos más detallados correspondientes a los sucesos que se producen con la actividad de venta, que serían incorporados a la Base asociándolos al registro de tipo R23 del cual dependen. Habría sucedido que unos determinados campos de datos que formaban parte del registro R23, habrían cobrado una importancia, en este caso debido al objetivo de controlar la actividad de los vendedores, que aconsejaban proyectarlos fuera del registro constituyendo otros registros individuales pero en alguna forma dependientes del registro tipo R23. Este último se comportaría como un registro raíz frente a esa nueva estructura que potencialmente podría desarrollarse a niveles más profundos de análisis.

Análogas consideraciones cabe hacer cuando interesa diferenciar para cada Centro de Actividad los diferentes períodos de tiempo. Así por ejemplo, el registro R23 puede incluir datos globales de producción y objetivos fijados, pero puede convenir también desglosar a un nivel más detallado la información individualizando las mismas medidas de actividad para los diferentes años con lo cual estamos cualificando la información no sólo para un Centro del tipo X, sino también para un año determinado. Esta es la motivación del registro tipo R28 que permite constituir un fichero que agrupa los diferentes años de actuación bajo la dependencia directa del registro R23 que corresponde al Centro de Actividad raíz.

Así, los registros R28, R29 y R30, son registros asociados a los anteriormente indicados por medio de los Sets S23, S24 y S25, y permiten establecer la unidad de tiempo “año” como horizonte de síntesis y de diferenciación de la información que de los Centros X, Y y Z vamos a manejar. Así, en el Set S24, recogeremos para cada agencia, la evolución año por año de los parámetros de control que estemos relevantes, lo que nos permitirá disponer en el Sistema, de la trayectoria de actividad que cada Agencia tiene con la Empresa y por síntesis, podremos considerar además trayectorias de conjuntos de Agencias según los diferentes criterios de agrupación que precisemos utilizar.

Un criterio de agrupación, sería naturalmente las relaciones de negocio que entre las diferentes Agencias existen, criterio que es manejable en función de la información que al respecto incluyamos en cada registro de Agencia, como datos estructurales propios.

El registro R22, es un registro estructural básico, que sirve de soporte a los Sucesos y que en este caso incluirá información de cada país en que la Compañía opera, con definición del nombre, del número de Ciudades con Aeropuertos, y cualquier otro tipo de información que referida a cada país sea preciso el procesar. El registro R22 a su vez, servirá de raíz a otras entidades estructurales de menor nivel jerárquico, que son las Ciudades, y que motivan la existencia de registros del tipo R17 que nos proporcionarán el criterio de organización geográfica que la Base de Datos deberá tener para ser un modelo efectivo. De cada Ciudad se englobará en el fichero R17 la estructura de datos relevantes para las correlaciones que precisemos; número de habitantes, número de oficinas de venta propias, número de Agencias de Viaje, Centros Z de otras Compañías, ... etc.

El registro R22, puede sugerir a su vez la necesidad de diferenciar los datos relativos a cada País para cada uno de los años, ello con objeto de tener modelizada la evolución de los atributos y características del País y poder establecer correlación entre dicha evolución y la de nuestra propia actividad empresarial. Así, parece aconsejable el disponer de unos registros del tipo R26 que para cada País y para cada Año, nos asocien los datos de número de habitantes, producto nacional bruto, renta *per capita*, índice demográfico, número de ciudades superiores a 100.000 habitantes, número de Aeropuertos, ... etc.

La estructura geográfica que nos determinan los registros R17 y R22, constituirá un soporte estructural idóneo para encuadrar la actividad comercial de venta del producto, ya que esta actividad está asociada directamente a puntos geográficos. Sin embargo, la actividad del transporte, implica el enlazar pares de puntos geográficos puesto que el viaje del pasajero o el trayecto de un avión, es una ruta que enlaza puntos diferentes. Entonces, para establecer correlaciones entre la actividad de ventas y la actividad de transporte, habremos de diseñar unos registros del tipo R21 que hemos denominado como fichero de "mercados", que asocia un país de origen con otro país de destino, y que nos servirá de marco para agrupar las corrientes de transporte que entre dichos países se establezcan. En este registro englobaremos datos referentes a las características y atributos de cada uno de los mercados y siguiendo el criterio antes indicado de diferenciar períodos de tiempo y diferentes niveles de detalle, servirá de registro raíz de los registros R16 y R11, de los que nos ocuparemos al hablar de las relaciones lógicas en el siguiente capítulo.

Fijémonos que los registros del tipo R17, R22 y R21, son registros estructurales que corresponden a una organización geográfica como marco práctico de la actividad, mientras que los registros del tipo R26 y R16, e incluso el R11, corresponden a estructuras temporales.

El registro R4, está referido también a un Centro de Actividad en este caso correspondiente a un Avión y agrupará para cada uno de los aviones propios de la flota o aviones ajenos contratados en permanencia, la matrícula, la flota a la cual pertenece, la fecha de alta en el servicio y demás características de tipo físico. Este registro por un lado contendrá las características de cada uno de los aviones, mientras que por otro, podrá servir de registro raíz para agrupar toda una posible jerarquía de otro tipo de registros de mayor nivel de detalle tales como registros de revisiones, de averías, de rotables, ... etc., si entrara dentro de los objetivos de la Base de Datos la explotación de aplicaciones de control de mantenimiento de aviones. A través de este registro R4, estaría ligado de una forma estructurada y procesable, la evolución del mantenimiento del avión con sus medidas físicas y monetarias y por otra parte, la actividad de transporte que el mismo avión realiza, pudiendo analizarse en consecuencia, posibles condicionantes. De hecho, el programa de mantenimiento de un avión, condiciona a la actividad comercial y a su vez ésta es sugeridora de aquél.

El registro R5, tiene la misma vocación que el registro R4, pero corresponde a un nivel de síntesis mayor, pues agrupa a todos los aviones que corresponden a un mismo tipo y podría recoger en su estructura interna datos tales como el número de aviones de que consta la flota, la utilización media diaria, el retraso medio

por despegue, la longitud de la etapa media, el número de tramos que por término medio realiza un avión en un día, ... etc., lo que corresponde a índices globales de actividad de la flota, datos que incluso pueden extenderse para diferenciar por períodos de tiempo como en el caso del registro R22 del cual emanaban registros del tipo R26.

Los registros del tipo R15 y R20, tienen la misma vocación que el registro del tipo R4, y son definiciones de recursos que participarán en los sucesos productivos.

El registro tipo R2, es otro ejemplo de registro estructural y agrupa en un fichero todos los posibles Tramos que puede un avión realizar. Es decir, un registro del tipo R2 definirá un Aeropuerto de salida, un Aeropuerto de llegada, y una distancia kilométrica entre ambos y nos servirá fundamentalmente de soporte de los Sucesos 44 y 45 (véase fig. 41). El Suceso 44 implica el que posteriormente se produzca el Suceso 45 y ambos exigen como raíz, la existencia de un registro del tipo R2. El fichero R2 es por lo tanto un soporte para el movimiento de los aviones.

Igualmente, el registro tipo R1, agrupa los posibles pares de puntos de salida y llegada que un pasajero puede recorrer sirviendo de marco a los Sucesos 1, 5 y 14, que tienen una gran significación en la relación contractual entre el pasajero y la Compañía transportista. Cada vez que se produzca un Suceso del tipo 1 ó del 5 ó del 14 que corresponden a la venta de un billete, a la realización de una reserva o a la aceptación de un pasajero al embarque, un registro del fichero R1, estará siendo elemento raíz de la actividad. Los registros del tipo R1 servirán por lo tanto de soporte para todo el movimiento de los pasajeros.

El fichero R12, es un ejemplo de registros asociados a productos, pues cada registro del tipo R12 corresponde a una línea programada por la Compañía. En el registro R12 se recogerán todos los datos referentes a número de línea, número de tramos, definición de los tramos, frecuencias de operación, período de validez, ... etc., lo que en conjunto constituirá el fichero que nos define la oferta de la Compañía. Los registros del tipo R12 hacen su aparición en la Base de Datos cuando tienen lugar los Sucesos 19 y 23 que son los sucesos generadores de la oferta de vuelos.

Desglosando a mayor nivel de análisis la definición de la oferta de vuelos, llegamos a la conclusión de que es necesaria la creación de los ficheros R6 y R7. El registro R26 establece la definición detallada de cada uno de los Segmentos programados, entendiendo por segmento el "salto" que un pasajero puede dar embarcando en un punto de la línea y desembarcando en otro en una fecha determinada, es decir, se trata de la definición de los servicios concretos que un pasajero puede comprar y por consiguiente de la prestación que un pasajero puede recibir. El fichero R7 recogerá los "saltos" que los aviones van a dar en función de las líneas que han sido programadas y establece el compromiso de operación de vuelos que se deriva de la oferta definida por el fichero R12.

Los registros de tipos R6 y R7, son elementos de información avanzada de Sucesos que van a tener lugar y estos registros serán actualizados cuando dichos sucesos, es decir la operación de los vuelos y el transporte de los pasajeros, tengan

lugar. Así por ejemplo, un registro del tipo R6 incluirá la Línea al cual pertenece el segmento programado de que se trate, definirá un punto de salida, un punto de llegada, la fecha en que el transporte tendrá lugar, e incluso la hora de salida y la hora de llegada. Cuando se producen los Sucesos número 5, es decir la reserva de plazas para este determinado segmento, vamos aceptando pasajeros para ese servicio, y cuando llega la hora de salida y se produce el Suceso 17 por el cual se termina la aceptación de pasajeros al vuelo, y posteriormente el Suceso 44 que determina la salida del vuelo, estamos en condiciones de conocer el número de pasajeros que en dicho segmento han sido transportados. En ese momento, Suceso 44, se dispone por lo tanto de información que permitirá actualizar el registro de tipo R6 correspondiente y así, del ejercicio de esa actividad, tendremos determinados los pasajeros realmente transportados en cada uno de los segmentos programados, lo que nos sitúa en condición de poder reproducir con fidelidad cómo han sido canalizados y en qué cantidades, los flujos de pasajeros, lo que constituirá soporte informático suficiente para cualquier programa de aplicación que haya de manejar esta información.

De forma análoga, un registro del tipo R7 contendrá la línea al cual pertenece, el punto de salida del avión, el Aeropuerto de llegada, el tipo de avión previsto, la fecha de operación, la hora de salida y la hora de llegada. Cuando tiene lugar el Suceso 44 en el que se produce la salida real del vuelo, estamos en condiciones de actualizar el citado registro incluyendo en él unos campos de datos que recojan el número de pasajeros que había reservados, el número de pasajeros que realmente volaron, el factor de carga del avión, el ingreso en unidades monetarias que dicho transporte produce y la tarifa media, lo que nos proporciona por suma de todos los registros del fichero R7, una medida exhaustiva de la actividad de transporte, pudiendo reproducirse en cualquier aplicación informática, la actividad real que tuvo lugar. Adicionalmente, y de acuerdo con el criterio de ordenación que para los registros del tipo R7 establezcamos en la Base de Datos, disponemos de un marco director de la propia actividad de actualización de los registros del fichero, pues cada registro del tipo R7 sabemos que es un compromiso de transporte que habrá de ser formalizado cuando llegue la fecha de salida del vuelo. Ello nos sugiere como criterio de ordenación la variable tiempo, como veremos al analizar las relaciones lógicas entre registros.

Los registros del tipo R9 son definidores del encadenamiento de los diferentes tramos programados en el fichero R7, los tramos han de formar secuencia en función de las fechas disponiendo tantas cadenas de registros en paralelo, como aviones físicos tengamos para efectuar el servicio. Cada registro del tipo R9 es por lo tanto una cabecera de cadena que enlaza a registros del tipo R7 y la cadena constituirá la definición de "saltos" que un avión físico va a realizar en la práctica. Obsérvese que aunque el registro R9 encabece una cadena y defina el recorrido de un determinado avión, dos cadenas diferentes pueden intercambiar los aviones físicos que las sirven cuando se produce una simultaneidad de dos aviones en un mismo Aeropuerto, por lo que los registros R9 definen cadenas o rotaciones lógicas y definimos unos registros del tipo R8 para realizar la asignación de una rotación lógica a un avión físico determinado. El registro R8, viene a establecer la asignación dinámica de una rotación R9 con un avión R4 y esa asignación es-

tará en vigor hasta que un nuevo registro del tipo R8 se incorpore a la Base de Datos cambiando la asignación. De esta forma, se puede ir recogiendo el modelo dinámico de la actividad que realmente está teniendo lugar, con lo que nuestra Base de Datos será autosuficiente para reconstruir todo lo que a este respecto en la realidad sucedió. También sirve el modelo, por supuesto, para programar los sucesos con la debida anticipación en la forma más racional y rentable, pues una de las potencias de mayor interés de tener un modelo estructurado, es la de poder simular con supuestos reales ya que estamos en condiciones de ensayar con una aplicación que contemple el futuro, lo que cabe esperar de cualquier rotación de aviones que proyectemos. Obsérvese el contenido de dinamicidad que podemos modelizar manejando adecuadamente las estructuras de registros y sus relaciones.

Los registros R14 y R19 vienen a definir las rotaciones previstas para las tripulaciones técnicas y las tripulaciones auxiliares de los aviones, complemento necesario del fichero R9 de rotaciones de aviones, ya que los tres recursos han de seguir líneas paralelas y coincidir para que la actividad de transporte pueda tener lugar.

Los registros R13 y R18, constituyen la programación de vuelos charter y vuelos de posición que tienen un tratamiento de excepción por obedecer a necesidades coyunturales adicionales al programa de servicio regular y se han individualizado por considerar conveniente medir por separado la actividad charter por motivos obvios y la actividad de vuelos de posición por su característica de vuelos improductivos.

El registro R10 y el registro R27, constituyen ficheros de síntesis, donde se agrupan perspectivas históricas que nos aporten datos pretéritos de utilidad para aplicaciones de análisis. Un registro del tipo R27, podría resumir para una línea determinada, en un período de tiempo definido por una fecha de comienzo y otra de terminación, la frecuencia con que opera esa línea en el pasado, las frecuencias que tuvo, el factor de carga medio, las ocupaciones máximas y mínimas, número de vuelos que se realizaron, ... etc., valores de indudable interés para cualquier estudio de análisis de comportamiento de líneas en el cual los datos históricos tienen indudable valor de experiencias anteriores reales.

El Esquema de la figura 43, no pretende ser exhaustivo pero sí dar una imagen suficiente de cómo debemos ir progresando en esta etapa de desarrollo del proyecto de Base de Datos, etapa que es realizada en paralelo con el planteamiento de las relaciones lógicas entre registros, del que es elemento inseparable que vamos a examinar en el siguiente capítulo.

3.4.2. DISEÑO DE RELACIONES LOGICAS

El planteamiento de las relaciones lógicas define ni más ni menos que los criterios de enlace de registros en la Base de Datos y es el elemento que dará flexibilidad y potencia de proceso a nuestro proyecto. Una vez más, hemos de indi-

car que no procede seguir un camino normalizado para el establecimiento de las relaciones lógicas y únicamente habrá que tener presentes determinados criterios directores de nuestra actividad creadora:

- Un criterio consistirá en *contemplar todas las posibles relaciones lógicas que entre los registros que estamos perfilando podemos establecer, para quedarnos después con las que dan más consistencia al modelo o con las que son más útiles* a la vista de los procesos informáticos de aplicación de los que en forma somera tenemos ya una orientación a nivel de objetivos.

Es decir, primero plantearemos diferentes alternativas de relación y después depuraremos el conjunto de posibles relaciones con un criterio selectivo que estará presidido por la razón y por nuestro “conocimiento analógico” del problema que a esta altura debemos ya de poseer.

- *Evitar duplicidad de relación.* La estructura de la Base de Datos ha de ser ciertamente conexa y todos los registros habrán de estar enlazados al conjunto para garantizar los caminos de acceso. Sin embargo, el tener en la Base de Datos diferentes alternativas de acceso, implica el tener relaciones lógicas duplicadas. Esta duplicidad, no tiene por qué ser perjudicial, pero siempre es un elemento que encarece el proyecto e incluso muchas veces puede restar flexibilidad a la estrategia de almacenamiento de registros. Por ello, sólo se habrán de establecer relaciones lógicas redundantes, cuando con ello mejoremos tiempos de acceso que estemos de importancia básica, lo que nos llevará a aplicar como criterio, el no disponer de redundancias nada más que en la medida estrictamente necesaria.
- *Enjuiciar la cantidad de relaciones.* Nuevamente, esto está en relación con el equilibrio entre las estructuras de datos internas a los registros y las estructuras externas de las que nos estamos ocupando. Posiblemente una complejidad de relaciones lógicas, sea diagnóstico de una definición inadecuada de registros y como viceversa, una red lógica demasiado elemental puede ser reflejo de que estamos construyendo un modelo pobre por haber definido registros demasiado compactos. No existe *a priori* una vía segura para determinar cuál ha de ser la proporción, pero sí es un factor que permanentemente debemos tener en cuenta, pues guarda una estrecha relación con la viabilidad práctica del rodaje operativo de la Base de Datos.
- *Los registros muy relevantes no deben ser excesivamente dependientes.* Quiere decirse que en cuanto más relevante es un fichero para el conjunto de nuestro modelo, más altura tiene que tener la jerarquía de relación por razones obvias de seguridad orgánica del sistema.
- *Los registros más estructurales han de ser registros raíz o próximos a las raíces.* Obsérvese que nuestro modelo de la figura 43, los registros

raíces, son registros de estructura geográfica, de estructura temporal o de Centros de Actividad estables.

Vamos a examinar algunos de los 25 Sets definidos en la figura 43 para mostrar las estructuras lógicas que en esta etapa del proyecto determinamos.

En el Set S20 asociamos los Centros de Actividad X, Y y Z a la ciudad en la que están radicados o de la cual administrativamente dependen, subordinando de esta forma, los registros R23, R24 y R25 al registro R17. En la Base de Datos existirán tantos anillos lógicos como ciudades correspondientes al fichero R17 que tenemos almacenados y de esta organización se refleja la orgánica de la red territorial de ventas, la cual es complementada con el Set S22 que agrupa en anillos lógicos las ciudades que pertenecen geográficamente a un mismo país. Los Sets S23, S24 y S25, prolongan la información que recogemos en la Base de cada uno de los Centros X, Y o Z constituyendo anillos con dichos Centros como cabeceras, y proporcionando la evolución año a año de la actividad de cada uno de los referidos Centros. Análogamente, el Set S9, proporciona un conjunto de anillos encabezados por cada uno de los países del fichero R22, proporcionando información de la evolución año a año registrada en cada país.

El Set S21 agrupa a los registros de tipos R1 y R2 en anillos dependientes de registros del fichero R17. Los registros R1 y R2 corresponden respectivamente a los diferentes servicios de transporte que pueden recibir los pasajeros y a los diferentes tramos operativos que pueden realizar los aviones; de esta forma, para cada una de las ciudades tenemos enlazadas en un anillo lógico, todas las rutas de aviones que salen o llegan a esa determinada ciudad y todas las posibles rutas de pasajeros que en esa ciudad pueden embarcar o que a esa ciudad pueden llegar. Este Set S21 refuerza considerablemente la estructura de operación de tráfico de pasajeros y aviones y sirve de marco sólido a la captación de información derivada de los sucesos.

Los Sets S6 y S7, establecen una estructura temporal para la organización de nuestra información a diferentes niveles de síntesis o análisis con dependencia de la actualidad de los datos. El Set S7 produce una distribución de un determinado mercado de acción en sus diferentes temporadas o períodos de tiempo y el Set S6 engloba para cada temporada los registros que se refieren a una semana concreta; de esta forma, el fichero R11 permitirá que para cada una de las semanas se aglutinen en torno a un registro la información de tráfico de pasajeros (por medio del Set S5), y el conjunto de registros R11, una vez completada toda una temporada de actuación, proporcionará datos de síntesis que podemos asociar al fichero R16 a través del Set S6. De esta forma, cuando por imperativos de tipo económico, una vez lejana en el pasado la operación de una temporada, procedamos a cancelar de la Base registros del tipo R6, R13 o R18, podremos ir acumulando información resumida de la actuación pasada en esta estructura temporal a que aludimos formada por los Sets S6 y S7, pudiendo mantener de esta manera la potencia de correlación que a la Base de Datos como modelo de actividad debe de exigírsele. El Set S8, constituye la definición de cobertura de cada mercado agrupando para cada registro del tipo R21, los servicios de transporte definidos por el fichero R1 que de

acuerdo con el criterio comercial que impongamos, pertenecerán al sector de mercado en cuestión. El Set S8, puede ser un ejemplo de Set con registros miembros enlazados *manualmente* por medio de un determinado programa de proceso que maneje los criterios comerciales que en cada ocasión queramos aplicar.

El Set S2, permite ordenar los registros del fichero R6, que corresponden a los segmentos programados, agrupando estos anillos encabezados por los pares de ciudades de salida y llegada al cual pertenecen por jerarquía natural. Se sugiere inmediatamente que la ordenación de registros miembros esté realizada por una clasificación de fechas y horas de salida correspondientes a cada registro de forma que, recorriendo cada uno de los anillos, tendremos ordenada la secuencia en tiempo de la oferta de la Compañía. Análogamente, el Set S3, engloba en anillos las operaciones físicas de los aviones entre cada par de ciudades definidas por el fichero R2 y se sugiere también inmediatamente que la ordenación de los registros miembros sea realizada en base a las fechas y horas de operación, lo que proporciona una secuencia en tiempos de las salidas y llegadas de aviones a cada ciudad, estructura de interés para cualquier proceso de aplicación que necesite contemplar estas características.

El Set S4 es una familia de anillos cuyas cabeceras son los registros del fichero R12 que corresponden a las líneas programadas. Los registros miembros del Set, son por un lado los diferentes tramos operativos de los aviones, en las diferentes fechas, que están programados dentro de esa línea, y por otro lado, las ofertas a pasajeros que en las diferentes fechas, están igualmente programadas dentro de cada línea. La ordenación de registros miembros en base a la secuencia de tiempo, parece igualmente conveniente. El Set S12, define la estrategia que vamos a seguir en la operación de los aviones puesto que enlaza todos los tramos operativos correspondientes al fichero R7 que van a ser servidos sucesivamente por una misma unidad de avión, unidad de avión desde el punto de vista lógico. La asignación de un avión físico del fichero R4 a una rotación lógica del fichero R9, la realizamos por medio de registros del fichero R8, enlazando estos registros por medio del propio Set 12 y del Set S16, siendo estas asignaciones dinámicas y suponiendo la aparición de nuevos registros del tipo R8, nuevas decisiones de asignación que anulan a las anteriores. Obsérvese que la secuencia de los registros R7 en el Set S12, ha de ser en base a fechas de salida y llegada de los aviones y deberá de existir un programa de aplicación que garantice permanentemente el que dicha secuencia tenga una lógica y sea viable de operar, puesto que los sucesivos servicios definidos por estos registros en el Set S12, han de ser realizados en secuencia por una misma unidad de avión y no puede haber solape entre la llegada de un tramo y la salida del tramo siguiente, entre los cuales tendrá incluso que existir las oportunas latencias de tiempo debidas a los tiempos de escala en Aeropuertos. Naturalmente en esta secuencia de registros R7 dentro de los anillos S12, el punto de llegada definido en un registro R7, deberá coincidir con el punto de salida del registro que sigue en la secuencia. También deberá haber lagunas de tiempos muertos que corresponderán a períodos de revisión de los aviones, en cuya problemática no vamos a entrar.

El Set S5, al que antes nos hemos referido, permite la agrupación de los datos de la actividad de tráfico por los períodos semanales a los que corresponden.

Del análisis del ejemplo, puede comprenderse que necesariamente, la actividad de diseño de las relaciones lógicas ha de ser paralela a la tarea de definición de los registros, puesto que ambas definiciones son conceptos duales fuertemente imbricados.

3.4.3. DIMENSIONADO PREVIO DEL ALMACENAMIENTO DE DATOS

Una vez diseñado el Esquema, y aun cuando éste siempre va a estar sujeto a revisión a lo largo del proyecto, conviene hacer una evaluación del volumen de almacenamiento de datos que va a ser necesario disponer para la operación de la Base, pues antes de proseguir adelante con nuestro trabajo, hemos de contrastar si desde el punto de vista económico estamos en la línea que en un principio habíamos previsto o por el contrario, el tema está desenfocado. Ello puede diagnosticarse con bastante aproximación a través del análisis del volumen de datos.

En la figura 44, recogemos un cuadro en el que se presenta la forma en que esta evaluación puede realizarse. En la primera columna, recogemos el número de registros que de cada tipo se prevé habrá en la Base cuando se realice la carga inicial de datos, cifra que no es preciso que sea de gran aproximación, pues en algunos casos podrá ser incluso absolutamente desconocida; simplemente se trata de seleccionar el orden de magnitud que con la mayor aproximación posible podamos enjuiciar.

La segunda columna, expresa la longitud de cada uno de los registros tipo, que obviamente será una longitud media en el caso de que los registros fueran de longitud variable y naturalmente, la longitud fija que el registro tipo tenga, si es que tiene esta condición como normalmente sucederá. En este grado de aproximación, y aun cuando el Esquema lo hayamos definido, no conoceremos todavía la longitud de los registros, pero sí dispondremos de una idea aproximada de los campos de datos que los registros van a contener, por lo que estamos en condiciones de hacer un tanteo de su longitud total. En la longitud de los registros, debemos de incluir los campos que serán necesarios para contener las asociaciones lógicas previstas, cuestión fácil de evaluar atendiendo a las relaciones lógicas que hemos determinado en el Esquema.

El producto de las columnas 1 y 2, dará para cada fichero el volumen de bytes que ocupará en la carga inicial de la Base de Datos. A partir de dicha carga inicial, conviene evaluar la tasa de crecimiento que cada registro va a experimentar en lo que a volumen de almacenamiento de datos se refiere, puesto que nuestro proyecto no se ciñe a la carga inicial, sino que tendrá que tener una pers-

	REGISTROS	NUMERO REGISTROS	LONGITUD REGISTRO	TASA ANUAL CRECIMIENTO	VOLUMEN DE BYTES ESTIMABLES	PORCENTAJE
1	SEGMENTOS TIPO	3.000	20	10	159.720	0,24
2	TRAMOS TIPO	1.600	26	8	52.403	0,08
3	AVION/TRAMO TIPO	7.400	45	10	443.223	0,68
4	AVION	120	250	5	34.728	0,05
5	FLOTA	30	208	—	6.240	0,01
6	SEGMENTOS PROGRAMADOS	680.000	40	12	38.213.280	58,32
7	TRAMOS PROGRAMADOS	360.000	46	10	22.041.360	33,64
8	ROTACIONES ASIGNADAS	100	20	5	2.315	0,01
9	ROTACIONES DE AVIONES	100	20	5	2.315	0,01
10	HISTORIA DE FLOTA	30	500	—	15.000	0,02
11	MERCADO/TEMPORADA/SEMANA	12.500	66	8	1.039.252	1,59
12	LINEAS	1.000	82	—	82.000	0,12
13	VUELOS CHARTER	100	55	40	15.092	0,02
14	TRIPULACIONES TECNICAS	125	85	5	12.300	0,02
15	TRIPULANTES TECNICOS	370	370	5	158.475	0,24
16	MERCADOS/TEMPORADA	250	34	8	10.707	0,01
17	CIUDADES	340	54	7	22.491	0,03
18	VUELOS DE POSICION	70	60	—	4.200	0,01
19	TRIPULACIONES AUXILIARES	125	62	5	8.971	0,01
20	TRIPULANTES AUXILIARES	633	200	4	142.412	0,22
21	MERCADO	25	42	8	1.323	0,01
22	PAISES	50	40	6	2.382	0,01
23	OFICINA VENTA PROPIA (X)	170	140	3	26.006	0,03
24	AGENCIA DE VIAJES (Y)	8.800	120	4	1.187.894	1,82
25	OFICINA OTRA COMPANIA (Z)	163	65	—	12.545	0,02
26	PAIS/AÑO	1.000	175	6	208.425	0,32
27	HISTORIA DE LINEA	1.100	320	3	384.630	0,59
28	OFICINA DE VENTAS/AÑO	850	55	1	48.167	0,07
29	AGENCIA/AÑO	42.000	27	1	1.168.360	1,78
30	OTRA COMPANIA/AÑO	770	18	—	13.860	0,02
	TOTAL.....	1.122.721			65.520.076	100,00

Fig. 44

pectiva de evolución que debe ser prevista de antemano para disponer de los medios técnicos necesarios para el almacenamiento físico de los datos. Esta tasa de crecimiento, puede medirse por períodos anuales en forma de porcentaje y en el ejemplo de la figura 44, está recogido en la columna 3.

En la columna 4, vamos a recoger el volumen total de bytes para los que tenemos que prever medios técnicos de almacenamiento físico desde un primer momento, evaluación en la que conviene aplicar un margen de seguridad holgado, y ello por dos motivos: porque las ampliaciones de medios técnicos requieren unas latencias de tiempo holgadas para evitar posibles saturaciones, y en segundo lugar porque dado que estamos en una etapa previa de dimensionado, hay una tendencia natural a que el producto en desarrollo vaya aumentando de volumen a medida que vamos avanzando hacia las etapas finales, porque con el progreso del proyecto, profundizando en detalles, se suscitan continuamente nuevas sugerencias. También hemos de cubrirnos para evitar sorpresas, de todo el contenido de incertidumbre que en este momento tienen las cifras que hemos tanteado.

Por ello, proponemos como una dimensión a contemplar, el volumen que la Base de Datos tendrá a tres años vista, de acuerdo con las cifras de las tres primeras columnas, es decir, de acuerdo con el número de registros de cada fichero, con su longitud y con su tasa de crecimiento anual, y suponiendo que al cabo de esos tres años, el factor de carga de los discos, sea de un 50 por 100 de su capacidad total. Es decir, las cifras de la columna 4, serán los productos de las cifras de las columnas 1 y 2, aplicándolos en tres veces sucesivas el porcentaje de la columna 3 en forma acumulativa y posteriormente, doblando dichas cifras. La suma de la columna 4, nos dará el total de volumen de almacenamiento que necesitamos, en nuestro caso del ejemplo, en torno a los 65 millones de bytes.

La columna 5 viene a reflejar el porcentaje con que cada fichero participa en el almacenamiento previsto.

Si del examen de las cifras obtenidas, comprobáramos que el almacenamiento necesario es excesivo para el nivel económico del que disponemos, o por el contrario comprobáramos que todavía podemos ampliarlo considerablemente, enriqueciendo de esta forma nuestro modelo, podremos tantear sobre el cuadro reseñado en la figura 44, sobre qué ficheros hemos de rectificar. Por ejemplo, si en nuestro caso estimáramos excesiva la cifra de 65 millones de bytes, cosa que en realidad no lo es, de un simple vistazo a la columna 5, vemos que son los ficheros R6 y R7 los que tienen una influencia decisiva en el volumen total y sobre los que habríamos de aplicar una eventual restricción, lo que lograríamos reduciendo el horizonte temporal que dichos registros cubren; esto es, cargando los registros R6 y R7 en la Base con menor antelación a las fechas de los vuelos, y a su vez, cancelándolos más prontamente una vez que las fechas son ya pretéritas. Si por el contrario, dispusiéramos de un mayor margen económico, y los objetivos nuestros se viesan enriquecidos cubriendo un mayor horizonte temporal, este podríamos ampliarlo aumentando en consecuencia las cifras de volumen de datos.

Con este dimensionado previo, obtenemos ya un Esquema con sus registros y relaciones lógicas definidas, y con un ajuste de dimensionado que tiene ya visos de realidad.

3.5. ESTUDIO DE LA RELACION DATOS/ACTIVIDAD

En la secuencia de desarrollo del proyecto, nos encontramos en un momento en que hemos estudiado suficientemente la actividad productiva de la Organización y disponemos de un conocimiento analógico de los flujos de la Red de Producción. Por otra parte, hemos “aislado” los Sucesos relevantes para los objetivos de la Base de Datos, e incluso hemos estudiado la interrelación que entre estos Sucesos tiene lugar, ya que la Matriz de Sucesos de la que anteriormente nos ocupamos, establece el conocimiento de cómo unos Sucesos inducen a otros, estudio que como dijimos, puede completarse fijando incluso términos de probabilidad para cada relación de inducción, encontrándonos incluso en condiciones de simular la dinámica de Sucesos en la secuencia del tiempo.

Por otra parte, hemos perfilado ya un Esquema y tenemos en consecuencia definida la estructura de datos que prevemos para la organización lógica de la Base. Ahora nuestra tarea ha de consistir en enfrentar el Esquema estructural que hemos propuesto con la dinámica de la actividad empresarial que el paquete de Sucesos productivos nos determina.

Habremos de analizar en qué medida los Sucesos relevantes impactan a la Base de Datos, atendiendo no sólo a su frecuencia de producción, sino a las actividades de creación, actualización y cancelación de registros que producen y a las modificaciones de sus relaciones lógicas.

Habremos también de analizar cuál es la dinámica que cada fichero va a experimentar al producirse las cadenas de Sucesos definidos en la Matriz de Sucesos.

También hay que comenzar a plantear las especificaciones que vamos a establecer para el acceso de los registros en la Base de Datos. En función de todo ello y persiguiendo siempre la armonía del conjunto que como modelo de actividad real debe contener, habremos de reconsiderar, en la medida en que proceda, el diseño del Esquema.

3.5.1. EXPRESION GRAFICA DE LA RELACION DATOS/ACTIVIDAD

En la figura 45, se recoge un estudio parcial de la relación Datos/Actividad correspondiente a nuestro ejemplo, la cual ha sido realizada con una presentación análoga a la que ya se recogía en la figura 6.

REGISTROS ↓		SUCESOS → (FIGURA 41)		18	19	20	21	22	23	24	25	27	28	37	44
		FRECUCENCIA →		SEMESTRAL	QUINCEN.	QUINC.	QUINC.	QUINC.	DIARIA	DIARIA	DIARIA	SEHEST.	DIARIA	DIARIA	DIARIA
		DATOS RELEVANTES ↓	SUCESOS INDUCIDOS →	19/22	20/21/23	24	20/25	19/20/21	24/25 37/44		28	18/28	19/20/21/22 23/24/25	28	
R1 SEGMENTOS TIPO	PUNTO SALIDA PUNTO LLEGADA			R6→S2	R6→S2			R6→S2	R6→S2		A/R6→S2				
R2 TRAMOS TIPO	PUNTO SALIDA PUNTO LLEGADA DISTANCIA			R7→S3	R7→S3			R7→S3	R7→S3		A/R7→S3				
R4 AVION	MATRICULA FLOTA											R8→S16	R8→S16	R8→S16	
R6 SEGMENTOS PROGRAMADOS	LINEA /PID. SALIDA Y LLEGADA FECHA /HORA SALIDA Y LLEG. PASAJ. RESERVADOS Y VOL.			C/R6	C/R6	X/R6	A/R6	C/R6 A/R6 X/R6	C/R6	X/R6	A/R6				A/R6
R7 TRAMOS PROGRAMADOS	LINEA /S. Y LLEG. /TIP. AVION FECHA /HORA S. Y LLEG. PASAJEROS RES. Y VOL. INGRESO /YIELD /L.F.			C/R7	C/R7	X/R7	A/R7	C/R7 A/R7 X/R7	C/R7	X/R7	A/R7			A/R7	A/R7
R8 ROTACIONES ASIGNADAS	Nº DE ROTACION AVION ASIGNADO											C/R8	C/R8	C/R8	
R9 ROTACIONES DE AVIONES	Nº DE ROTACION FLOTA				R7→S12			R7→S12 A/R7→S12	R7→S12		A/R7→S12	C/R9 C/S12, S17 R7→S12 R8→S12	A/R7→S12 R8→S12	A/R7→S12 R8→S12	
R11 MERCADO/TEMP/ SEMANA	PAIS DE ORIGEN Y DESTINO SECTOR TEMPORADA / SEMANA			C/R11 R6→S5	R6→S5			R6→S5	R6→S5		A/R6→S5				
R12 LINEAS	Nº LINEA /Nº TRAMOS FRECUCENCIA PERIODO DE VALIDEZ.			C/R12 C/S4 R6→S4 R7→S4	C/R12 R6→S4 R7→S4	X/R12 X/S4	A/R12	C/R12 A/R12 X/R12 R6, R7→S4	R6→S4 R7→S4						

Fig. 45

En la primera columna y en la primera fila, están relacionados, respectivamente, los ficheros y los Sucesos de nuestro ejemplo que de una forma parcial y a título de ejemplo, estamos considerando. Es decir, incluimos en el cuadro de la figura, los ficheros R1, R2, R4, R6, R7, R8, R9, R11 y R12 y por otro lado los Sucesos 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 37 y 44, cuya numeración corresponde a la Matriz de Sucesos.

En la segunda fila, expresamos una indicación de la periodicidad media con la que los Sucesos se producen, lo que no presupone su volumen. En la tercera fila, se recogen los Sucesos inducidos contemplados en el cuadro, datos que concuerdan con los de la Matriz de Sucesos de la figura 41.

En la segunda columna, indicamos los datos más relevantes que los ficheros considerados contienen.

En el resto del cuadro, que consiste en una tabla de doble entrada, se recoge la actividad que cada suceso produce en cada fichero. Así, por ejemplo, el Suceso 37 produce en relación con el fichero R9, la actualización de la inserción de un registro del fichero R7 en el Set S12 del cual un registro R9 es cabecera. También produce la inserción de un registro del tipo R8 en el Set S12. Análogamente, el Suceso 27 produce, respecto del fichero R8, la creación de un nuevo registro y de la misma forma, por citar otro ejemplo, el Suceso 20 respecto del fichero R12, genera una actividad de cancelación de un registro y de un Set del tipo S4 del cual un registro R12 es cabecera. No desarrollamos más el contenido de la figura, por haber sido convenientemente explicado al tratar la figura 6 que presenta un perfil análogo.

Si de la columna 45, contemplamos la fila que a cada fichero corresponde, estamos observando la dinámica de cambio que tiene en la secuencia de Sucesos. Si contemplamos las columnas de la figura, estaremos atendiendo la incidencia cualitativa que cada Suceso produce en la Base de Datos, El conjunto de los dos aspectos, completan una expresión de la relación que existe entre los datos que hemos definido por medio del Esquema y entre los Sucesos que hemos “aislado” de la Red de Producción y estudiado en la Matriz de Sucesos.

3.5.2. ESTUDIO DE LA DINAMICA DE REGISTROS

Atendiendo a la expresión cualitativa de la relación Datos/Actividad, analizaremos para cada uno de los ficheros, su evolución a través de los Sucesos que tienen lugar, para lo que reflexionaremos contemplando la fila de la figura 45 correspondiente a cada fichero y por otro lado, las frecuencias y volúmenes con que los Sucesos se producen. Los datos correspondientes a las frecuencias de los Sucesos, están recogidos en forma aproximada en la propia figura 45, mientras que los volúmenes de actividad de los Sucesos, habremos de tenerlos en mente y suficientemente conocidos en orden de magnitud como consecuencia de nuestro estudio de la Red de Producción y de la reflexión que hemos realizado al hacer el tanteo previo de dimensionado del almacenamiento.

Adicionalmente, habremos de tener en cuenta para cada fichero, las posibles creaciones o cancelaciones de registros que tendrán lugar en base a los procesos de síntesis de horizontes de tiempo que tienen lugar por motivos de economía, que más o menos tenemos perfilados, pues están en relación con el dimensionado de almacenamiento, y que obviamente no están asociados a Sucesos directos porque obedecen a otras motivaciones de proyecto independientes de la Red de Producción.

El conjunto de evolución en el tiempo que para cada fichero en cuestión queda de esta manera reflejado, habrá de ser juzgado a la luz de la lógica, considerando:

- Las tasas de creación de nuevos registros de fichero y su periodicidad.
- Las tasas de cancelación de registros.
- Las tasas de actualización de registros.
- Las tasas de recuperación de registros. Estas realmente, no estamos en condiciones de evaluarlas, puesto que dependerán de las aplicaciones de explotación de la Base de Datos, pero podemos empezar a contemplarlas en lo que se refiere exclusivamente a las necesidades de leer registros en la realización de los procesos de captación de datos que los Sucesos generan, y por lo menos puede observarse a través de los caminos lógicos que hemos de seguir, teniendo el Esquema a la vista, más bien para detectar fallos de inconsistencia y errores de bulto que para medir realmente la actividad.
- Los cambios de enlaces lógicos.
- Los aspectos relevantes que se detecten y sean dignos de tener en cuenta.

Este es también el momento de comenzar a perfilar la estrategia de almacenamiento de los registros, pues estamos ya midiendo en la Base de Datos, la estabilidad de cada fichero y en qué medida es estructural o en qué medida contiene registros cambiantes o rotaciones en su población.

Un subproducto interesante del estudio de la dinámica de registros, es el ir elaborando un programa de control estadístico, fijando desde este momento algunas especificaciones que deberán cubrir los programas de proceso que accedan a la Base de Datos. A título de ejemplo, pudiéramos citar como sugerencia en el caso al que de forma continua nos estamos refiriendo, una posible conveniencia de medir en forma automática, las frecuencias de cancelación, actualización y creación de registros del tipo R7 en nuestra Base de Datos. Para ello, los programas de software que acceden a estos registros, deberán de generar "información contable" utilizando al efecto los oportunos contadores o incluso si esta actividad fuera muy importante, disponiendo en la Base de Datos de nuevos registros que recojan este tipo de información. En nuestro caso, esta motivación pudiera radicar en que los registros R7, suponen la tercera parte de la Base de Datos (véase fig. 44) y su dinámica de evolución nos interesa tenerla muy medida para garantizar la adecua-

ción de medios técnicos en todo momento. Ello parte de que en este caso concreto nos puede proporcionar una información de diagnóstico, ya que el significado de la dinámica de este fichero, es una correlación entre la actividad de operación de los aviones y la antelación con que dicha actividad es programada e incluso con la latencia de tiempo con que mantenemos los datos de los sucesos en la Base de Datos una vez que éstos han ocurrido.

3.5.3. INFLUENCIA DE LOS SUCESOS Y CAPTACION DE LOS DATOS

El análisis del impacto que cada Suceso produce en la Base de Datos, puede ser realizado contemplando las columnas de la figura 45. Del examen de esta información que de esta forma hemos estructurado, tendremos una imagen precisa de la relevancia de los sucesos en función de la actividad que generan en la Base de Datos, lo que habremos de complementar en sus aspectos cuantitativos teniendo en cuenta las frecuencias de producción de los sucesos y los volúmenes de dicha producción. Ello nos calificará a cada uno de los sucesos.

Adicionalmente, podemos seguir la secuencia de los sucesos definida en la Matriz de Sucesos, y chequear la armonía que existe entre las frecuencias que hemos previsto para los sucesos y que recogemos en la figura, y entre las inducciones de unos sucesos con otros, en lo que debe existir ciertamente compatibilidad, examen que nos revelará posibles inconsistencias lógicas.

También es un momento adecuado para analizar las oportunidades de captación de datos en función de la relevancia e impacto de cada uno de los sucesos generadores de los datos. Recordemos a estos efectos, los criterios que manejábamos en los primeros capítulos:

- *Captación de datos como subproducto de la actividad.* En nuestro ejemplo, pudiéramos plantear la actualización de los registros R6 y R7 que tiene lugar cuando se produce la salida de un avión (Suceso 44), como subproducto del propio procedimiento de despacho del vuelo, obligando a que el desarrollo de dicho procedimiento de trabajo, incluya la emisión de un mensaje que vaya en directo al sistema de Base de Datos con la información pertinente del vuelo. El condicionar de esta forma la actividad productiva, puede estar motivada por lo relevante que es el tener esa información disponible con rapidez, ya que nos puede ofrecer un diagnóstico inmediato de cómo se está desarrollando el tráfico de pasajeros y qué ingresos económicos está produciendo, lo que constituye el pulso comercial, de indudable interés para la gestión diaria.
- *Captación de datos como actividad independiente.* El Suceso 18, sería un ejemplo en nuestro caso, pues en la creación de registros que genera, no tenemos el factor tiempo como elemento condicionante.
- *Captación como actividad condicionante.* En nuestro ejemplo, el modificar la rotación de un avión, es decir la creación de registros del tipo

R8, es una circunstancia importante que puede requerir cambiar todo el planteamiento de trabajo que da lugar a la decisión, pues la demora en la captación de datos y los correspondientes procesos de control, pueden conducir a inversiones de tiempo indeseables.

- *Condicionantes técnicos, geográficos y de rentabilidad.* En nuestro ejemplo anterior de actualización de los registros R6 y R7, para los que hemos considerado la conveniencia de envío de mensajes por teleproceso, habremos de contemplar las circunstancias geográficas inherentes, lo que trae consigo el tener presente los medios técnicos disponibles o que habría que disponer, en este caso de líneas de comunicación, y su dimensión económica a efectos de rentabilidad. Conviene señalar que la captación de datos, no tiene por qué ser uniforme, y que para los registros de un mismo fichero, pudieran coexistir diferentes especificaciones de captación de datos.

3.5.4. ESPECIFICACIONES DE ACCESO A LOS DATOS

Se trata de estudiar las limitaciones que hay que imponer a las operaciones de acceso a la Base de Datos en función de la dinámica lógica que del estudio de la relación Datos/Actividad se deriva. Continuando con nuestro caso, citaremos a título de ejemplo:

- Mientras se actualiza la inserción de un registro del fichero R7 en el Set S12, cosa que producen los Sucesos 22, 25, 28 y 37, las áreas de datos que contienen al fichero R7, debieran de estar protegidas de otros accesos concurrentes, o al menos las páginas de datos involucradas.
- La importancia del fichero R12, nos invita a especificar actualización exclusiva y protegida para sus registros.
- Las cancelaciones que por síntesis históricas se produzcan de los registros del tipo R11, habrán de incluir como especificación el chequeo previo de los registros miembros de los Set tipo S5.

Es decir, para cada operación de las que se establecen en la tabla de doble entrada de la figura 45, se habrán de considerar las condiciones técnicas y limitaciones de acceso que los diferentes programas de software de aplicación deberán respetar, construyendo de esta forma todo un conjunto de especificaciones para la programación.

También hay que definir los tipos de acceso que convenga determinar para el almacenamiento de los datos, esto es, en qué casos de los registros de nuestro Esquema habremos de disponer de situación en acceso directo, en qué casos deberemos definir procedimientos específicos para calcular la localización de registros y en qué casos la estrategia de búsqueda y almacenamiento quedará simplemente condicionada por la participación de registro como miembros de Sets. A estos efectos, habremos de tener en cuenta criterios elementales tales como:

- Los registros que son raíz de estructura, deberán disponer de unas condiciones de acceso directo o de acceso calculado. El criterio de separación será una consecuencia lógica del volumen de registros dependientes que desciendan de dicha raíz, para homogeneizar y compactar en la medida de lo posible, la estructura física en las áreas lógicas de datos y por consiguiente en las unidades de almacenamiento.
- Los registros que sean objeto de altas tasas de recuperación y sean miembros de Sets muy voluminosos, es decir, formen anillos lógicos muy densos, habrán de tener sus registros principales en especificaciones de acceso rápido. Por ejemplo en nuestro caso, los Sets tipo S4, son voluminosos y quizás los registros del tipo R12, convenga disponerlos en acceso directo.
- Las relaciones estructurales de registros que sirven de marco a Sucesos se consultan mucho en los procesos informáticos, y por ello hay que disponer organizaciones físicas muy compactas para evitar accesos sucesivos innecesarios que dilatan los tiempos de respuesta de la Base de Datos. Por ejemplo, en nuestro caso, los registros tipo R22 quizá convenga disponerlos con un acceso calculado, con un procedimiento de cálculo de direcciones de áreas y páginas que nos proporcionen un espaciado suficiente entre los registros que constituyen el fichero R22, para permitir una distribución racional de los elementos de los Sets S22, S21, S1 y S8, ya que estos Sets engloban registros estructurales que se prevén serán muy consultables, es decir, tendrán altas tasas de recuperación, y conviene que estén suficientemente compactados.

Sin embargo, quizás en los Sets S2 y S3 no nos importe en exceso la separación física de los miembros, ya que quizás no sean procesados estos anillos con demasiada frecuencia, al menos en conjunto.

Con este tipo de razonamientos vamos configurando las especificaciones de acceso de los datos, lo que en definitiva vienen a determinar en cierto modo la estrategia previa de almacenamiento que la unidad de administración de la Base de Datos tendrá que disponer.

3.5.5. RECONSIDERACION DEL DISEÑO DEL ESQUEMA

El estudio de la relación entre los datos y la actividad nos ha permitido profundizar considerablemente en los detalles de proyecto de la Base de Datos, y debemos en este momento de haber diagnosticado con una cierta seguridad los puntos débiles, las faltas de armonía o los conflictos de intereses que existan en nuestro planteamiento.

En función de todos estos elementos, habremos de dar marcha atrás para rectificar o perfeccionar las definiciones de los registros, las relaciones lógicas entre dichos registros e incluso la tipología de los Sucesos, remontándonos si es preciso al análisis de la Red de Producción. La repetición de todo el proceso seguido hasta el

momento es elemental puesto que simplemente necesitamos actualizar algunos de los elementos, siguiendo esta actitud permanente de aproximaciones sucesivas con que los proyectos informáticos se desarrollan. Es preciso por lo tanto una actitud de ir consolidando y retocando si es preciso lo ya construido a medida que vamos profundizando en el desarrollo del producto.

3.6. ESPECIFICACIONES DE LOS DATOS

Las anteriores etapas de desarrollo de proyecto habrán sido normalmente seguidas en secuencia varias veces, aplicando el concepto de aproximaciones sucesivas, lo que habrá conducido a tener diseñado el Esquema de la Base de Datos y estar ya suficientemente contrastado en sus aspectos más relevantes.

A partir de esta posición se realizará un análisis a mayor nivel de detalle para llegar a establecer el conjunto de especificaciones que definirán la Base de Datos, lo que será punto de partida para aplicar el correspondiente lenguaje de definición de datos dependiendo del sistema de software a utilizar.

La definición de los datos es a partir de este momento una actividad análoga a la que se desarrolla para la definición de ficheros en cualquier sistema informático convencional, pues el esfuerzo de diseño integrado de los datos, está ya en su mayor parte realizado y disponemos, en consecuencia, de unos patrones bastante consolidados de cómo van a ser los registros y de un marco estable de las relaciones lógicas que entre ellos van a existir.

La definición detallada de los registros tipo o ficheros constituirá una labor más o menos convencional, aunque siempre habrá que tener en cuenta los aspectos propios de su pertenencia a una Base de Datos, como puede ser la previsión de pointers para las relaciones lógicas, las zonas de áreas lógicas designadas para almacenamiento o las especificaciones de protección o de acceso.

La definición de los Sets, es decir, de las relaciones lógicas utilizadas, también debe de desarrollarse al mayor nivel de detalle, sin precisar esfuerzo creativo adicional, puesto que las relaciones están ya establecidas en el Esquema e incluso ya tenemos un juicio de cuáles han de ser los criterios de ordenación de los elementos miembros en cada Set. Sí tiene cierta importancia el revisar cuidadosamente cuáles son los Sets en los cuales la gestión de enlace de los miembros puede ser confiada a un procedimiento standard o en qué casos tendremos que diseñar procedimientos ad-hoc para la gestión de enlace de los miembros, lo que dependerá de si el enlace de registros está en función del significado de claves contenidas en el propio registro o adicionalmente existe alguna otra condición a chequear independientemente de los valores de las claves.

Los criterios de ordenación de los anillos lógicos nos proporcionan la forma de seleccionar un registro dentro de un anillo ya seleccionado. Sin embargo, para cada

Set tipo habrá que definir cuál es el criterio que el software de utilización habrá que aplicar para seleccionar un determinado anillo dentro de la familia que el Set tipo en cuestión define.

Completarán la definición de los datos el diseño de las áreas lógicas, que constituirán el marco de almacenamiento contemplado por los analistas de Sistemas de aplicación, para los que como ya hemos repetido, la organización física real será transparente. El diseño de las áreas lógicas de datos tiene una componente muy importante de normativa propia de cada Centro de Cálculo y quizás no convenga pronunciarnos demasiado acerca del número de áreas que debe de haber, el número de páginas que cada área debe de poseer, si estas páginas conviene que tengan una mayor o menor uniformidad entre las diferentes áreas, etc. Esta problemática no tiene demasiado interés para el diseño lógico, y más bien está asociada al dimensionado de la configuración para adaptar las áreas lógicas a las necesidades físicas de almacenamiento, y por otra parte a los "buffers" con los que podamos contar en la memoria central para el manejo de páginas en el acceso a los datos de la Base.

Sin embargo sí tiene mucha importancia, la asignación de registros a áreas y páginas lógicas de almacenamiento, porque aquí radica gran parte de la estrategia de la administración de datos, función a la que ya nos hemos referido y que se va perfeccionando fundamentalmente por información de tipo estadístico derivado de la propia información de la Base.

3.6.1. DEFINICION DE REGISTROS

La definición de Registros tipo equivale a la definición de ficheros individuales, aunque por tratarse de una Base de Datos ya conozcamos que éstos van a estar interrelacionados. Vamos a hacer un repaso de todos los aspectos que la definición de Registros tipo debe de contemplar, muchos de ellos, los esenciales, ya concebidos en todo el proceso racional de diseño del Esquema, y otros varios más convencionales, que habrá que establecer, y cuya tipología no merece mayores comentarios.

No se propone un formato determinado para el establecimiento de estos datos por considerar que ello puede ser realizado en diferentes maneras de acuerdo con las costumbres del Centro de Cálculo, pero sin embargo sí interesa puntualizar que la definición de registros debe de estar bajo formato standard en una u otra forma, con una perfecta descripción y explicación de detalles, que permita que el manejo de cada fichero pueda ser comprendido y realizado con independencia de las personas que lo diseñaron siguiendo una actitud objetiva al respecto. La descripción de los ficheros ha de servir para diferentes analistas de aplicaciones en base a la multifunción de los datos que es característica esencial de las Bases de Datos. Ello nos obliga en una medida superior con respecto a las aplicaciones convencionales a ejercer con rigor la labor de documentación.

Vamos por lo tanto a ver cuáles son los conceptos a definir para cada fichero que realmente constituirán los conceptos que también habrán de ser documentados:

— *Identificación del registro tipo.* El fichero tendrá a partir de este momento un nombre codificado o simbólico y por lo tanto procesable, y una descripción del contenido.

— *Definición de relaciones lógicas.* Para cada fichero convendrá indicar para debida información de los programas que lo manejen, los Sets en los que participa como registro principal. Igualmente deberá recogerse la denominación de los Sets en los que el fichero participa, aportando registros miembro, diferenciando si son automáticos enlazados por un procedimiento standard o si se trata de miembros manuales que disponen de un procedimiento **ad-hoc**.

El número de pointers que son dedicados en el registro a las relaciones lógicas también habrá de ser detallado y concordará con la oportuna reserva de campos de datos en el registro para alojar dichos pointers.

— *Descripción de los campos de datos.* Se trata de la determinación de toda la estructura interna del registro, lo que viene a suponer una denominación análoga a la que en cualquier definición de fichero COBOL se realiza.

Para cada campo de datos definido, se indicará naturalmente la longitud y la tipología de los caracteres, numéricos, alfanuméricos, etc., indicando también los nombres simbólicos asociados a los campos de información.

Entre los campos del registro, habrá que definir también los pointers necesarios para las relaciones lógicas, aunque existen dos tipos de ellos: los que serán manejados por un procedimiento standard del sistema o los que serán manejados por software **ad-hoc** de aplicación. Estos últimos quizá requieran de una definición a este nivel, mientras que los primeros son por lo general, en los paquetes de software de base de mercado, ya tenidos en cuenta y anexionados a los registros en tiempo de compilación de la definición global de los datos, en base a las relaciones lógicas automáticas que hayamos declarado.

— *Zonas para almacenamiento del registro.* Para el fichero habremos de establecer cuáles son las áreas lógicas de datos y en qué páginas dentro de ellas vamos a disponer el almacenamiento de los registros. Esta definición de áreas y páginas deberán estar de acuerdo con la descripción general de las áreas lógicas.

Aquí está explicitada realmente la estrategia de almacenamiento que se ha ido “destilando” a lo largo del proceso de diseño del Esquema, en base principalmente al espaciado que hemos de realizar entre registros raíces para que los registros dependientes tengan sitio en las zonas de almacenamiento para agruparse en torno a ellos, buscando la mayor compactación posible de almacenamiento.

El acceso a los datos se realiza siempre recuperando páginas enteras de información, por lo que los registros relacionados en una misma página

serán procesados en memoria y, por lo tanto, con mucha mayor rapidez que si se encuentran en diferentes páginas, lo que obliga a operaciones consecutivas de acceso que obviamente son más lentas. Esta es la motivación para perseguir un almacenamiento compactado, agrupando y almacenando físicamente lo más conexos posibles, los elementos de estructura lógica que vayan a ser procesados en secuencia.

- *Dinámica del fichero.* Deberá recogerse el número de registros habidos en la carga inicial, tanto los previstos en un principio como a posteriori, los que realmente entraron, además del volumen de bytes que ello representa.

También deberá indicarse y consiguientemente prever, las tasas de recuperación de registros del fichero por un determinado período e incluso de perfil de su distribución. También la tasa de operaciones de actualización, la tasa de creación de nuevos registros y la tasa de cancelación de registros, todo ello medido por períodos iguales, quizás mensuales.

Como dato de interés que determinará en la explotación real la dinámica de cambio, está la definición de los elementos estadísticos de control que a este respecto tenemos que especificar para que después sea tenido en cuenta tanto en el software de aplicación como en las rutinas de administración de la Base de Datos. De acuerdo con esta especificación, la evolución de la dinámica de cambio será conocida en la explotación real, y estos datos se acompañarán a la documentación del fichero porque de este modo el fichero, como elemento vivo que evoluciona con el tiempo, será conocido en toda su trayectoria, y por consiguiente se tendrá por un lado una medida de su proyección hacia el futuro y por otra parte irán saltando a la vista las inconsistencias que la unidad de administración deberá ir perfeccionando.

- *Especificaciones de protección de Datos.* En la definición de fichero deberá explicitarse los requerimientos de copiado de información de seguridad en los procesos de actualización, los bloqueos de páginas que haya que disponer en las operaciones de acceso y las exclusividades que los registros hayan de imponer a los programas por posible conflicto de simultaneidad.
- *Especificaciones de acceso.* Se determinará la forma de localización o acceso de posiciones de almacenamiento que se determinan para los registros del fichero. Se determinará si el acceso es directo, y en su caso, se establecerá el criterio de asignación de posiciones, o bien se establecerá si el acceso es calculado, y a tal efecto se explicitarán los parámetros de entrada al procedimiento de cálculo y la definición del propio procedimiento, o bien se establecerá cuál es el camino lógico que en el Esquema se va a utilizar para acceder a los registros, estableciendo las claves de paso que permiten seguir dicho camino lógico.

3.6.2. DEFINICION DE SETS

Los Sets deberán ser definidos uno a uno como entidades individuales. Aunque cada uno de los Sets sea un componente más del Esquema, desde el punto de vista lógico, y desde el punto de vista físico, un subconjunto de la Base de Datos, el Set es una unidad lógica completa que establece la forma de relación de los registros de varios ficheros.

Para cada uno de los Sets que hemos definido en el Esquema habrá que establecer la definición detallada de los siguientes conceptos:

- *Identificación del Set.* El Set como unidad individual tendrá un nombre simbólico procesable que le proporciona la identidad necesaria. Igualmente es obligado una descripción normalizada de las relaciones lógicas que agrupa.
- *Tipo de Set.* Deberá determinarse el tipo de estructura lógica que va a utilizar el Set. Si nos referimos a las posibilidades que ofrece el Sistema DBMS de CODASYL, deberá decidirse si el Set va a ser en estructura de anillo o bien va a estar organizado en base a una matriz de pointers.
- *Ficheros componentes.* Para cada Set se determinará cuál es el fichero principal, lo que equivale a decir cuáles serán los registros principales de la familia de Sets específicos que pertenecen al Set tipo que estamos definiendo.

También se determinará cada uno de los ficheros que van a participar como miembros en el Set, lo que determina el conjunto de registros que pueden formar parte como miembros de los Sets específicos. Para cada uno de los registros tipo miembro debe decidirse si va a ser miembro automático o miembro manual, es decir, si el procedimiento de inserción de un registro en un anillo va a ser decidido y realizado bajo el control de un procedimiento standard, o bien van a existir criterios de decisión específicos de una determinada aplicación que realice dicha función.

- *Criterio de organización de los anillos.* En cada Set tipo los anillos lógicos que constituyen los Sets específicos tienen un determinado criterio de organización lógica que es preciso establecer de forma unívoca. Por ello, para cada Set se definirá cuál es el criterio en base al cual se organiza la secuencia lógica de registros, y así dicho criterio nos determinará el problema doble de cómo localizar un determinado registro dentro de un anillo y en qué posición insertar un nuevo registro dentro del anillo que le corresponde.

El criterio de ordenación lógica de registros puede estar definido en base al orden de almacenamiento de los registros, lo que hace que la ordenación sea independiente del significado de la información.

También puede definirse un orden que esté basado en decisiones toma-

das por una determinada rutina, lo que equivale a decir que el criterio residirá en cada momento en la memoria principal de proceso como una "pieza" modificable. Ello supone que el criterio sea abierto.

Generalmente, el criterio de Organización será una clasificación lógica de registros en base al valor expresado por claves que serán campos contenidos en los propios registros, lo que equivale a los criterios tradicionales de clasificación de registros.

- *Criterio de selección de anillos.* El que esté definido el problema de cómo insertar o recuperar un registro en o de un determinado anillo, nos resuelve la situación después de que dicho anillo lógico haya sido identificado.

Deberá definirse para completar la resolución de los criterios de Organización lógica, la forma en que para un determinado registro se identifica y localiza cuál es el anillo al que dicho registro está asociado de entre la familia de anillos que pertenecen a un mismo Set tipo.

Generalmente, hay dos criterios para establecer esta definición. Uno de ellos es reducir el problema de seleccionar un anillo al problema de seleccionar un registro del fichero principal del Set, ya que hay una correspondencia biunívoca entre los registros que constituyen el Set tipo. Como para el fichero principal hemos establecido el criterio de cómo localizar o dónde almacenar un registro específico, este criterio resuelve también la fórmula de selección de anillos. Hay que hacer destacar no obstante que la forma de selección del registro principal puede estar definida como realizable a través de la participación de los registros del fichero como miembros en algún otro Set, por lo que el proceso de identificación de un registro podrá llevar anejo el aplicar a su vez un nuevo criterio para seleccionar anillos de otros Set para conducirnos a resolver la localización de nuestros registros principales. Dicha selección de anillos estará definida oportunamente en las especificaciones del Set tipo al que correspondan, por lo que el problema queda determinado. Naturalmente, la selección de un Set específico puede por este procedimiento llevar aparejado toda una cadena de selección de anillos en varios Sets lógicamente relacionados, pero siempre la cadena tendrá un fin, pues ya hemos indicado que los registros raíz de la estructura lógica tienen que poseer necesariamente una forma de localización independiente de la estructura lógica.

El otro criterio de selección de anillos está basado en disponer en cada programa de aplicación que vaya a utilizar como estructura lógica el Set tipo en cuestión, un contador que contenga la dirección de un registro que esté formando parte del Set en el almacenamiento. Dicho registro nos define lógicamente el anillo específico al cual está enlazado, y por lo tanto la información contenida en el referido contador es definidora de un anillo específico que de esta forma se selecciona. Como

el contenido de ese contador es manejable por la propia aplicación, ello quiere decir que aplicamos un criterio abierto.

Determinando todas estas especificaciones para cada uno de los Sets tipo, hemos completado la definición de la estructura lógica del Esquema de la Base de Datos. Sin embargo, la Base de Datos puede disponer de estructuras lógicas adicionales tales como Cadenas, que para determinadas aplicaciones o para determinadas formas de almacenamiento de registros es preciso disponer. Toda otra estructura que se prevea la Base de Datos va a utilizar deberá ser igualmente definida.

3.6.3. DEFINICION DE AREAS LOGICAS DE DATOS

El concepto de Area es en realidad el elemento intermedio que hace posible la transferencia de las posiciones físicas de almacenamiento de los datos a los programas de utilización. Estas áreas habrán de definirse de forma individual cada una de ellas, con lo que se completa el marco lógico al cual hará referencia las aplicaciones informáticas que operen con la Base de Datos. El problema posterior de asignación de áreas lógicas a zonas físicas de almacenamiento queda como problema técnico de operación en un Centro de Cálculo.

Para cada Area lógica habrán de definirse las siguientes entidades:

- *Identificación del Area.* Cada Area es una unidad independiente de las demás que estará identificada en la Base de Datos con un nombre simbólico procesable.
- *Páginas componentes.* Para cada área se definirá el número de páginas que va a poseer y cuál es el tamaño en bytes de la página, tamaño que es un standard dentro del área. El tamaño de la página viene un poco condicionado por la estrategia de disposición de las zonas de la memoria principal, ya que la página es la unidad de transferencia a la memoria y es la unidad, por lo tanto, en base a la cual se realiza el intercambio de información entre la Base de Datos y la memoria principal.

Sin embargo, el tamaño de la página también deberá estar decidido por otro lado en base a la dimensión de las estructuras lógicas que dentro de las páginas del área vayan a almacenarse, e incluso a la posible conveniencia de minimizar tiempos de proceso de dichas estructuras. Debe de tenerse en cuenta que el seguimiento de una estructura es ultrarrápido si la estructura está contenida en una página ya que dicho proceso sería realizado en la memoria principal. Por el contrario, de estar dividida la estructura en varias páginas, el proceso se ve afectado por tantas operaciones de entrada-salida como páginas sea preciso recuperar. El número de accesos necesario para procesar una determinada estructura puede ser especificado como un límite máximo deseable, lo que nos daría en su caso un factor a tener en cuenta en la determinación del tamaño de la página.

También ha de definirse las páginas destinadas a “overflow” su cuantía y su situación dentro del resto de las páginas.

- *Densidad de carga de las Areas.* La densidad de carga debe estar en consonancia con las necesidades que tengamos de que los datos estén más o menos compactados en torno a las estructuras lógicas a las que pertenecen. En cuanto mayor sea esta necesidad más baja habrá de ser la densidad y viceversa. Efectivamente, para que el eventual almacenamiento de un registro pueda ser realizado con proximidad a otro determinado elemento con el que está relacionado debe de ser posible encontrar cerca de dicho elemento una posición libre de almacenamiento, lo que será más probable en cuanto más baja sea la densidad de ocupación. Entendemos como proximidad de almacenaje de dos elementos el que éstos estén situados en una misma página o, en su defecto, el que el ejercicio sucesivo de almacenaje de elementos afines nos lleve a una distribución en un número máximo determinado de páginas del Area.

Para la fijación de esta densidad aplicaremos por un lado la definición del número de páginas del Area que, en unión de la definición del tamaño de la página, nos da una potencia de almacenamiento que referida al volumen neto que efectivamente vamos a necesitar, nos determina realmente un factor de densidad. Establecida esta densidad por la definición del Area, es preciso adicionalmente determinar cuál va a ser el criterio que controla la carga inicial de registros en la Base de Datos, porque para que los objetivos de compactación de datos puedan ser posteriormente viables en la operación normal de la Base de Datos es preciso que esta carga inicial se efectúe disponiendo una distribución uniforme de registros en las páginas del Area.

Esta distribución uniforme puede conseguirse fijando como objetivo de la carga inicial el que la densidad sea un standard máximo de las páginas. Es decir, fijamos un porcentaje que establece la densidad máxima que vamos a permitir para cada página en la carga inicial de registros. Ello quiere decir que la distribución de registros en las páginas se va realizando de manera que se considerará una página ya completa cuando el porcentaje máximo de densidad establecido ha sido ya alcanzado, momento a partir del cual la página queda cerrada a la admisión de nuevos registros durante dicho proceso de carga inicial. Esta condición solamente se levantará en el caso de que todas las páginas hayan alcanzado ya dicho factor y todavía existan registros a almacenar.

La determinación para cada Area para el factor de densidad de carga inicial es, como vemos, un elemento que decide en gran medida la racionalidad de la estrategia de almacenamiento y condiciona decisivamente la eficiencia de la explotación.

- *Especificaciones de protección.* Para las Areas y sus páginas habrá de definirse los criterios de seguridad que vamos a utilizar, naturalmente en función de la importancia de la información que en el Area vayamos a almacenar. Es importante la condición del “copiado” preventivo

de páginas antes de que un programa de aplicación haya efectuado modificación de registros dentro de ellas, siendo ésta un elemento fundamental que permitirá reconstruir el estado de la Base de Datos ante cualquier incidencia en la explotación de una aplicación. Los elementos de protección referidos a las Areas, a las que ya nos hemos referido, deberán ser en esta etapa establecidos.

4. CRITERIOS GENERALES DE DISEÑO DE LA BASE DE DATOS

La Base de Datos tiene dos significados informáticos de naturaleza independiente. En primer lugar, como soporte de datos de un conjunto de aplicaciones, tiene el valor de constituir la Organización de la información que este conjunto de aplicaciones van a manejar, lo que sitúa a la Base de Datos dentro de la utilidad que un conjunto de ficheros más o menos sofisticados proporciona. Por otro lado, la Base de Datos marca una concepción nueva que orienta en una nueva forma la explotación de aplicaciones informáticas, y ello en base a que ese conjunto de ficheros sofisticados que nos va a servir de soporte de las aplicaciones de informática constituye un Sistema. Este Sistema está muy estructurado y confiere a la explotación de aplicaciones unos atributos y propiedades muy superiores a la suma de propiedades que los ficheros aportarían con una contribución individual.

Esta significación de tipo general nos obliga a que el diseño de la Base de Datos tenga unas exigencias propias que son independientes de las necesidades informáticas que los objetivos de las diferentes aplicaciones nos impondrán. Estas exigencias o especificaciones generales de diseño estarán por lo tanto asociadas a esta nueva concepción de explotación informática, o, lo que es lo mismo, estarán asociadas al elemento fundamental que hemos introducido que es la condición que nos hemos impuesto de que la Base de Datos constituya un Modelo de la Actividad de la Organización o Empresa a la que estamos aplicando estas técnicas.

Vamos a examinar por lo tanto cuáles son estos criterios generales de diseño que siempre han de anteponerse a las necesidades concretas de las aplicaciones, de las que en gran medida son independientes.

Así, vamos a examinar el concepto que hemos denominado de “tiempo presente”, que nos permitirá que el contenido de la Base de Datos no sólo sea un modelo de la actividad de la Organización, sino una proyección de su evolución en el tiempo y que todo ello esté en disposición de ser procesado informáticamente.

También vamos a examinar la necesidad de que los Datos sean únicos por corresponder a hechos únicos, lo que quiere decir que las duplicidades de información han de ser mínimas y estar bien controladas.

También nos vamos a referir a la orientación morfológica que el Modelo deberá poseer como organismo o Sistema que es y que habrá de ser objeto de formalización por etapas a lo largo de la vida de la Base de Datos.

El conocimiento de la actividad como elemento indispensable previo a todo diseño se resalta también como condición imprescindible que decidirá la utilidad del

conjunto así como la debida dimensión de futuro que todo el planteamiento deberá de contener.

La rentabilidad es obviamente también un criterio general a contemplar.

4.1. EL CONCEPTO DE TIEMPO PRESENTE

Dentro de la característica esencial de Modelo que pretendemos conseguir nos obligamos a mantener la Base de Datos con un contenido informático que sea procesable y que refleje todos los datos definidores de la actividad en su diferente estado de síntesis o análisis, correspondiendo a diferentes sucesos, actuales, pasados o futuros, que en cada momento o tiempo presente de posible explotación de la Base de Datos puedan ser de relevancia para los Sistemas de información de gestión.

Esto nos obliga a mantener un contenido de la evolución histórica de los datos, es decir, que la información que dispongamos en la Base deberá permitirnos reconstruir la frecuencia de Sucesos que en la Red de Producción han tenido lugar a lo largo del tiempo. Sin embargo, el contenido de información de la Base de Datos no habrá de ser uniforme respecto del tiempo al cual están asociados los Sucesos, porque las necesidades de información que un Sistema informático tenga respecto de la Base de Datos en un momento determinado, son diferentes en función de la coordenada tiempo a la que los datos corresponden. Así, generalmente, los datos más actuales con respecto al momento en que se procesan será necesario que contengan un grado de detalle o de análisis más completo mientras que los datos que pertenecen a tiempos más pretéritos respecto del momento del proceso será utilizados en forma más sintetizada o globalizada. E igualmente puede ser necesario que dispongamos en la Base de Datos una cierta prospección de los Sucesos futuros, lo que naturalmente estará en relación con las posibilidades de prospección que el modelo que diseñamos realmente contenga en términos estadísticos, aunque en todo caso, la característica de modelo estructurado de la actividad nos proporcionará un marco hacia el futuro que habrá de servir de soporte informático a los Sucesos que en el futuro se produzcan y en consecuencia de un camino determinado de cómo los diferentes nuevos registros correspondientes a dichos Sucesos van a pasar a incorporarse a la Base de Datos.

La Base de Datos, y todos sus procedimientos de explotación, habrá de realizar, con independencia de las aplicaciones informáticas, una especie de "destilación" de los datos en función de la coordenada de tiempo a la cual pertenecen. Quiere decir que cuando un dato entra a formar parte de la Base éste estará recogido en uno o varios registros que tendrán una vida estable durante un cierto tiempo. Pasado este tiempo, el dato ya no será necesario que permanezca en la Base como entidad individual en el grado de detalle en el que estaba cuando su actualidad pertenecía a tiempos próximos al tiempo de captación. Pasado este cierto tiempo que de hecho está definido a través de la estructura del Esquema de la Base de Datos, el dato es sintetizado, englobándose con otros datos afines que en conjunto desapa-

recen como elementos individuales y son cancelados los correspondientes registros que los albergan, pero que sin embargo vienen a crear como contrapartida otros registros de síntesis con referencia a períodos acumulados por conceptos de tiempo generalmente o bien por conceptos relativos a otras entidades de orden geográfico u orgánico.

Este proceso de destilación de los datos hace que el volumen de la Base de Datos no crezca proporcionalmente con el tiempo, lo que supondría por un lado una penalidad de la economía del Sistema y por otro lado supondría el realizar procesos informáticos ad-hoc para obtener información de síntesis, procesos que penalizarían la explotación y que la Base de Datos debe evitar anticipándose en la definición del modelo a muchas de las necesidades de síntesis que las aplicaciones vayan a tener.

Este concepto que venimos a denominar de Tiempo Presente nos configura la Base de Datos con el contenido óptimo de información que en cada momento presente de la explotación va a ser preciso tanto para procesar la información más actual como para reconstruir la historia de la actividad o para poder proyectar la información hacia el futuro.

Hemos dicho, por lo tanto, que cada dato se integrará en la Base con su propia trayectoria de evolución histórica y que a medida que pase el tiempo, la trayectoria del dato que recoge la Base de Datos va a ser modificada manteniendo siempre el dato a nivel atomizado en los períodos de tiempo más presentes y efectuando síntesis o resúmenes de la trayectoria del dato en tiempos más pretéritos, síntesis cuya elaboración se realizará con los oportunos criterios de tipo estadístico, recogiendo las medias aritméticas o medias ponderadas o desviaciones típicas que se contemplen como relevantes.

El contenido histórico de la trayectoria de los datos nos aporta la reconstrucción de la actividad de la Organización y todo ello debe estar alojado en la Base de Datos en términos procesables, es decir, con las debidas codificaciones, porque de esta forma podremos mediante cualquier aplicación informática establecer las correlaciones estadísticas que en nuestros Sistemas de información de gestión coyunturalmente puedan surgir. Esta potencia de correlación es una característica importante que viene a ser una medida de la potencia que la Base de Datos tiene como modelo de la actividad, porque determina la capacidad analógica de nuestro modelo. Tengamos muy presente que un modelo es un instrumento de analogía y que nuestro modelo es un instrumento a su vez para los Sistemas de información de gestión que vamos a procesar.

El hecho de que los datos más pretéritos estén más sintetizados viene a resolver un compromiso entre la rentabilidad del volumen de almacenamiento y la racionalización de las necesidades de explotación, pero en ningún caso debe de constituir una limitación que reste potencia de análisis a la Base de Datos por el hecho de que cancelamos de ella los datos analíticos correspondientes a períodos más antiguos. Tengamos en cuenta que un dato deberá estar contenido en la Base solamente el tiempo en que dicho dato atomizado vaya a ser necesario en aplicaciones frecuentes o en su defecto muy relevantes. Ello porque el tener almacenado un dato supone no sólo el coste del espacio que ocupa, sino el correspondiente a

que dicho espacio ocupado esté “online” en la configuración en disposición de ser accedido, coste que evidentemente es de orden superior.

Sin embargo, los datos que son cancelados de la Base de Datos por pertenecer a épocas más antiguas y ser su utilidad en momentos presentes reducida, podrán ser dispuestos en soportes informáticos (cintas o discos) que situaremos en nuestra biblioteca de datos “off-line” que permitirá seguir conservándolo en términos rentables. Cualquier aplicación que coyunturalmente necesitara de un análisis individualizado de estos datos podrá procesar estos soportes informáticos como complemento de la Base de Datos porque en realidad la biblioteca “off-line” de datos es una extensión de la Base de Datos.

Es decir, el proceso de destilación de los datos viene a conducir a que pasados determinados períodos de tiempo en la trayectoria de los datos se produce un fenómeno de síntesis de información que queda recogido en la Base de Datos en base a criterio de tiempo presente, mientras que los datos atomizados que han servido para realizar estas síntesis abandonan la Base de Datos “online” y pasan a soportes informáticos de la biblioteca “off-line”.

Esta estructura de tiempos que recogemos en el Esquema constituye el marco director del contenido de tiempo presente de nuestro modelo, y será el elemento que nos permitirá “vivir al día” disponiendo en forma estructurada y procesable en cualquier momento los datos actuales, teniendo accesible el contenido de historia del modelo y disponiendo asimismo de un marco procesable abierto a la incorporación de datos futuros.

4.2. UNICIDAD DE LOS DATOS

La unicidad de los datos es un objetivo a perseguir, porque el dato corresponde a un hecho pragmático que tuvo lugar en la Organización y por lo tanto atendiendo a su realidad es una entidad única.

La captación del dato también debe de estar orientada de forma unívoca, porque captar un dato viene a ser medir un aspecto de la realidad con un determinado instrumento que es el procedimiento de captación. Es obvio que diferentes mediciones de un mismo hecho pueden llevarnos a la obtención de datos diferentes aun cuando el procedimiento de captación fuera el mismo, porque éste puede no ser perfecto. Las diferencias que se obtengan, aun cuando sean elementales, siempre introducirán confusión e indeterminación.

El hecho de que la captación de datos haya de ser única, no está en contraposición a que en el proceso de captación dispongamos de las necesarias redundancias para poder establecer “cuadros” en la información que nos detecten los posibles errores de procedimiento, en lo que no entramos porque ya está expuesto en otro capítulo y pertenece a otro orden de ideas.

El dato es único y debiera ser de una forma única captado, y su almacenamiento en la Base de Datos también deberá ser único, y vamos a analizar cuáles son las razones para ello.

Desde el punto de vista de actualización del dato, solamente el que éste sea único puede garantizar de forma unívoca que la actualización será debidamente realizada, porque de lo contrario el procedimiento de actualización no estaría a salvo de los problemas derivados de la necesaria característica de simultaneidad de proceso por diferentes aplicaciones.

El mantenimiento del dato desde el punto de vista de Administración de la Base de Datos sólo puede igualmente garantizarse si los datos están identificados de forma unívoca porque todas las operaciones de revisión de las estructuras lógicas, y más relevantemente el proceso de síntesis o “destilación” del dato, exige que no existan incertidumbres ni haya que ejercer criterios de decisión de cuál versión del dato tiene mayor protagonismo.

En lo que se refiere a las especificaciones de protección, el concepto de control obliga a la unicidad, pues es complicado a nivel operativo mantener procedimientos paralelos del mismo nivel de seguridad para proteger diferentes versiones de una misma entidad. Aun cuando tal procedimiento fuera viable, su aplicación reiterada podría conducir a que la degeneración que de un dato pudiera llegar a producirse, fuera de distinto grado para las diferentes versiones del dato, lo que siempre tiene un contenido de incertidumbre indeseable. Ello no quiere decir que dentro del mismo criterio de protección del dato se establezca como acción precautoria el disponer de “duplicados” lo que se realiza a nivel general para toda la Base de Datos pero que pertenece a otro nivel de conceptos.

Si el dato es único atendiendo a su procedencia, y si de él existieran diferentes versiones dentro de la Base de Datos, las relaciones lógicas que dichas versiones del dato tengan dentro de la Base, habrán de ser exactamente iguales, pues lo contrario sería una vulneración del criterio de lógica que hemos aplicado en el proceso creativo del Esquema. Si estas relaciones lógicas han de ser iguales por extensión nos llevaría a disponer de estructuras lógicas “paralelas”, lo que por extrapolación, nos conduce al concepto de “elementos duplicados” que es de un orden diferente al concepto de “redundancias”. Lo que viene a suponer que el propio concepto de relaciones lógicas que contiene el Esquema nos lleva en gran parte a excluir por principio la existencia de datos duplicados.

La rentabilidad es otro factor que nos impide que mantengamos información duplicada porque una Base de Datos es un elemento “online” en el Centro de Cálculo y el espacio de almacenamiento tiene un coste asociado elevado.

Desde el punto de vista filosófico, que de hecho la Base de Datos contiene como nuevo concepto de explotación de Sistemas, la propia característica de modelo impide que existan redundancias de los elementos constituyentes, ya que van contra la lógica general de los Sistemas, en los que sólo se establecen elementos duplicados por motivos estrictos de seguridad.

La propia filosofía de control de la Base de Datos queda favorecida en cuanto a su realización práctica por el criterio de unicidad de la información, porque que-

dan separados en un Sistema de Base de Datos los conceptos de control y protección de la información en los procesos de utilización de la misma por un lado, y por el otro el de seguridad técnica. El primero de ellos, de protección y control, se ve favorecido como decimos por la unicidad, mientras que la seguridad técnica dispone de los elementos suficientes de copiado preventivo de Areas y páginas para que el hecho de que la información sea única, no provoque el que la Base de Datos sea más vulnerable, asegurando por otra vía la compatibilidad entre la unicidad de datos y su integridad.

4.3. LA ORIENTACION MORFOLOGICA

La Base de Datos va a ser un Sistema con toda la potencia lógica que la Teoría General de Sistemas nos aporta. La vida operativa de la Base de Datos, que comprende todas las evoluciones de la Base desde que ésta fue diseñada, va constituyendo un proceso de artificialización de la actividad de la Organización, proceso que irá siendo más racional y más perfecto a medida que la Base de Datos está más integrada en la vida de la Organización y a su vez los diseñadores y administradores de la Base de Datos, van enriqueciendo su conocimiento de la actividad con nuevos matices. Es un proceso que analógicamente tiene un cierto paralelismo con el fenómeno general de artificialización que el hombre, en todos sus órdenes, va generando en su paso por el mundo.

La Base de Datos como Sistema tiene que ir incorporando por lo tanto nuevas formas, nuevas estructuras, que estén en condiciones de integrarse en un proceso continuo al diseño inicial, y el diseño inicial ha de permitir lógicamente esta proyección natural del Sistema.

La modelización de la actividad ha de ser objetiva y la Base de Datos tiene que recoger en su línea operativa todos los fenómenos informáticos relevantes con su significado objetivo y de una forma abstracta, es decir, independientemente de la coyuntura y de las excepciones que en la vida de la Organización pueden llegar a producirse. Dicho de otra manera, la analogía entre la Base de Datos como modelo y la actividad de la Organización que estamos modelizando ha de ser una analogía sólida con contenido propio y ser independiente de la coyuntura.

La forma de lograr esta identidad analógica entre la Base de Datos y la Organización, consiste en que la Base de Datos tenga un cierto contenido de "órgano vivo" y que éste por sí mismo se desarrolle en evolución paralela a la Organización, para que de esta forma, pueda ir incorporando como evolución natural, aquellos nuevos atributos que estarán en relación con nuevas circunstancias que en la realidad de la Organización se están produciendo.

Se sugiere como idea conceptual que aporta ilustración a este respecto, el concepto de evolución que manejamos en las especies vivas del mundo que nos rodea, lo que nos aporta la idea de potencia evolutiva que nuestro Sistema debe de contemplar.

El establecer una Base de Datos, excesivamente “dedicada” a alguna aplicación en particular, anteponiendo esa vocación al plano de generalidad que por sí mismo debe de poseer, equivale al concepto de “especialización” que en la evolución de las especies es observable. La especialización puede ser incluso necesaria, pero equivale a cerrar puertas a futuras evoluciones posibles. En la medida en que una Base de Datos está dedicada a una aplicación, está en cierto modo imposibilitada para utilidades futuras de tipo más general. El modelo que ofrece una Base de Datos especializada, es un modelo deformado de la realidad. En caso de que esta cierta polarización del modelo sea necesaria atendiendo a las propias necesidades de los Sistemas de información de gestión, hay que velar porque la “polarización” sea mínima, lo que debe constituir un objetivo.

Para perseguir este objetivo, habremos de actuar con una actitud de evitar todas las rigideces en el Esquema que sean innecesarias o cuyo valor puede ser sustituido por otros atributos, que pudieran limitar desde el principio, las posibilidades de evolución del modelo. La Base de Datos tendrá un porvenir futuro mayor en la medida en que tenga vida propia y pueda irse adaptando a nuevas necesidades que en la realidad siempre surgen en una Organización.

El replanteamiento de una Base de Datos que habría de tener lugar porque el Sistema no tenga posibilidad de evolución, supone una condición de mutación o ruptura que es generadora de crisis en la Organización, pues a medida que los Sistemas informáticos son más complejos e integrados, el Activo de información que se maneja en la Empresa, tiene mayor importancia, y el tener que cambiar todos los supuestos funcionales bajo los que ese Activo se maneja, es un factor indudable de crisis.

Conviene distinguir este tipo de ruptura funcional, de los cambios bruscos que en el orden tecnológico pueden tener lugar por el progreso de la técnica, pues en este caso, es decir en los cambios de hardware y software que periódicamente siempre se producen, la crisis se reduce exclusivamente al Centro de Cálculo y afecta a cuestiones internas de procedimiento sin que necesariamente los problemas derivados, hayan de inducir problemas en la Organización. Sin embargo, un cambio brusco de las bases funcionales del Sistema de Base de Datos, tiene una interrelación manifiesta con el desarrollo de la actividad empresarial y si dicha ruptura llega a producirse, debiera de ser únicamente una consecuencia de una ruptura análoga que en la actividad real se ha producido, nunca por la condición viceversa, pues en la analogía dual Empresa-Modelo, la Empresa es el elemento principal y el Modelo el elemento subordinado.

Por lo tanto, proyectar una Base de Datos, es crear un Modelo que tiene que tener en sí mismo una potencia permanente de evolución pues la Base de Datos no se proyecta completa desde un principio y su morfología ha de ser completada con el transcurso del tiempo.

Así, a la Base de Datos le irán creciendo nuevas estructuras, y la expansión de nuevos registros y de sus relaciones lógicas que ello supone, ha de ser en primer lugar posible y en segundo lugar compatible. Para ello, las posibilidades de crecimiento, han de estar contempladas desde un principio y el Esquema debe de preverlas, pues el Sistema de Base de Datos, debe de contener en sí mismo algu-

na forma de conocimiento de cómo la evolución puede irse orientando. En este caso, al hablar de Sistema, quizá convenga referirse a un entorno más general que el propiamente técnico, englobando a los Analistas y Administradores en simbiosis con la Base de Datos.

El diseñador de la Base de Datos, tiene que tener muy presente que está configurando un Sistema complejo, con un contenido dinámico importante, y tiene que emplearse a fondo aplicando su creatividad al máximo.

Es realmente difícil de establecer guías para comprobar si nuestro proyecto está siendo consecuente con estos principios y quizá aquí juegue un papel muy importante la intuición y el sentido analógico que proporciona el ejercicio prolongado de la profesión.

Adicionalmente y aunque pudiera parecer chocante, conviene muchas veces echar mano de criterios asociados a las condiciones de sensibilidad humana tal como pudiera ser la Estética. El creador de un Esquema, debe de estar identificado con él como producto de sus potencias creativas y si tal figura se produce, debe de sentir un cierto placer hedonista al contemplar su obra, y si algún aspecto de su obra llegara a repugnarle bien por intuición o por un mecanismo incluso inconsciente, deberá repasar los aspectos afectados y considerarlos muy a fondo, porque ello puede corresponder a una cierta deficiencia del Modelo. La sensibilidad del creador de la Base de Datos, para estar en fase con la obra que está produciendo, es una guía útil que no debemos descartar.

Sin embargo, para poder ejercer la creatividad hasta estos grados, es preciso que el diseñador opere con una ingenuidad infantil, desvinculándose de sus propias obras y manteniendo una actitud abierta a toda crítica y a considerar toda nueva posibilidad por poco probable o por indeseable que parezca. La potencia de evolución del Modelo, no puede comenzar a ser frenada por la impotencia de evolución del proyectista.

4.4. CONOCIMIENTO DE LA ACTIVIDAD

El diseño de la Base de Datos, debe de incluir la experiencia de las personas que participan en el desarrollo de la actividad empresarial que se está modernizando e incluso, debe de incluir también la experiencia que ciertamente existe de los gremios que engloban a dichas personas porque muchas veces “la ciencia” pertenece al gremio y no a los elementos que lo componen.

Esta experiencia de personas y asociaciones de personas, debe de “fertilizar” el diseño de la Base de Datos para que la germinación sea fecunda, pero sin embargo, debe a su vez de ser evitado que la propia experiencia sea un condicionante, porque contemplada la Organización desde un plano general, las experiencias de las partes integrantes siempre tienen algunos factores de deformación inherentes a la especialización de su actividad.

Por ello, el equipo diseñador de la Base de Datos, tiene que tener un conocimiento profundo de la actividad empresarial sobre la que está construyendo su modelo, porque un simple conocimiento superficial es inhábil para contemplar los efectos de segundo orden que siempre se producen de la condición de complejidad de los Sistemas en que se opera, y hemos de tener muy presente la ley universal de que pequeñas causas pueden provocar grandes efectos. En consecuencia, los posibles efectos de segundo orden han de estar contemplados de antemano y ser correctamente evaluados para que las consecuencias estén debidamente acotadas.

El estudio de la Red de Producción ha de ser realizado muy al detalle y para ello no debe de olvidarse la participación de las personas o grupos de personas que pueden aportar experiencias abundantes, para lo que se recomienda cuidar el factor de convivencia entre los grupos humanos involucrados.

El diseñador debe de adquirir conocimiento a través de esta convivencia de los aspectos relevantes de la actividad, ejerciendo la cualidad que debe de poseer en alto grado de saber penetrar en los detalles relevantes sin caer en los problemas de orden menor que los detalles siempre traen consigo. Es decir, debe de ser capaz de destilar los aspectos relevantes de la actividad, despreciando los aspectos "deformantes" innecesarios que toda actividad ciertamente lleva consigo.

El criterio de conocimiento de la actividad, es quizá el principal de todo el planteamiento, pues es en definitiva el que de una forma más sólida, viene a garantizar las propiedades de uso general de la Base de Datos y el que dará vida propia a nuestro modelo.

4.5. LA DIMENSION DE FUTURO

Ya hemos expuesto anteriormente el contenido de futuro que la Base de Datos debe de poseer desde el punto de vista de la evolución en el tiempo, es decir, la potencia de transformación que debe de existir.

Simplemente, vamos ahora a reseñar que desde el punto de vista técnico, han de ser contemplados la expansión de la Base de Datos a nuevas zonas de acción en la Organización, además de que como ya hemos dicho, para una zona determinada, el modelo se vaya enriqueciendo.

Esta condición requiere que vislumbremos objetivos futuros de la informática en la Empresa, lo que supone una previsión de la expansión de los Sistemas de información de gestión y la previsión de los medios técnicos que van a ser necesarios.

Los medios técnicos, son factores secundarios, pero pueden llegar a ser altamente condicionantes, por lo que a la hora de elegir el hardware y software de aplicación, tenemos que tener presente que debe estar garantizada la continuidad futura del Sistema. Es decir, la disposición de nuevas unidades de almacenamiento,

nuevos sistemas operativos, ampliación de memoria y en general, la configuración de la instalación, deberá estar contemplada con un panorama de tiempo suficiente y las posibles evoluciones de los medios técnicos de proceso de datos, deberán disponerse con la anticipación oportuna para que las condiciones funcionales de la explotación, no se vean sensiblemente afectadas. La ruptura técnica no debe nunca provocar la ruptura funcional.

Conviene puntualizar estas ideas porque generalmente hay que proyectar la Base de Datos desde el punto de vista informático con un horizonte de tiempo muy superior al que contemplan las planificaciones más generales de la Empresa, y en cierto modo, la Base de Datos deberá prever los aspectos relevantes de estos períodos no contemplados a nivel general.

En la medida en que por necesidades de diseño de los Sistemas de información, nos estamos anticipando a la planificación general de la Empresa, estamos en alguna forma condicionando el futuro de la Organización y esta consecuencia, es muy seria y nos obliga a que la actividad de diseño sea realizada con gran objetividad y con sentido de prospección, ya que obviamente no es lícito condicionar aunque sea en pequeño grado, objetivos generales desde una parte, y ello deberá tener consecuencias mínimas. Resaltemos, que se puede condicionar el futuro tanto por exceso como por defecto, es decir, tanto por introducir orientaciones equivocadas en la Base de Datos, como por omitir el recoger estructuras con aspectos que en el futuro pueden llegar a ser relevantes.

4.6. LA RENTABILIDAD

La rentabilidad es obviamente un factor que siempre ha de estar presente en nuestro proyecto. Como se trata de un proyecto, que como tal tiene unos objetivos que lo motivaron, en su momento tuvo la oportuna asignación de recursos para su desarrollo e igualmente, se realizaron unas previsiones económicas.

Como hay que suponer que la actividad previa de previsión económica fue realizada de una forma consciente, ello establece un nivel en la dotación de recursos para la explotación de la Base de Datos. El proyecto tendrá naturalmente que ajustarse a estos niveles porque una separación exagerada de ellos, puede deteriorar el factor de confianza que es en definitiva el que proporciona el apoyo necesario para que la Base de Datos se diseñe en colaboración con las unidades orgánicas involucradas.

Si la separación con respecto a estos niveles, se detectara conveniente porque un enriquecimiento del Modelo aporte aspectos cualitativos que abren posibilidades interesantes, ello deberá estar siempre muy bien documentado y las propuestas oportunas a los niveles de decisión, habrán de ser colocadas con la debida antelación y siempre bajo la idea de ofrecer alternativas nuevas, nunca bajo unos argumentos que condicionen la ejecución del proyecto, porque de lo contrario, se deterioraría igualmente el factor **confianza**.

La rentabilidad es un criterio importante en el diseño del contenido de “tiempo presente” que la Base de Datos va a poseer, pues combinando adecuadamente ambos criterios, se pueden conseguir soluciones válidas para diferentes niveles económicos, es decir, para diferentes posibilidades.

La economicidad del Sistema, nos limita el volumen del almacenamiento “on-line” y también la dimensión de la memoria principal y en consecuencia, la agilidad de los tiempos de proceso, lo que no debe constituir un condicionante esencial para nuestro proyecto, pues siempre las soluciones que como proyectistas ofrezcamos, estarán gobernadas por una limitación de recursos, lo que constituye el planteamiento que en la práctica siempre se produce.

Como la orientación del proyecto se realiza por aproximaciones sucesivas, las desviaciones importantes en lo económico, habrán de ser reajustadas con una nueva aproximación.

5. EL SISTEMA DBMS DE CODASYL

Como hasta este capítulo hemos visto el enfoque funcional para el diseño de una Base de Datos, o lo que es lo mismo, los aspectos más de aplicación y problemas a resolver, consideramos que ahora procede dar un repaso del instrumento de software de Base que podemos utilizar para la programación y desarrollo técnico de la Base de Datos proyectada. Todos los paquetes de software de Bases de Datos, utilizan estructuras lógicas del estilo de las que hemos visto en el primer capítulo o en todo caso, combinaciones de ellas y, además, los diferentes Sistemas tratan de perseguir objetivos análogos, de independencia de datos y programas, de unicidad de la información, ... etc., cosa que consiguen en mayor o menor grado.

Nosotros vamos a analizar un poco en detalle, uno de estos Sistemas, ya que los demás podrían ser igualmente utilizados por analogía, ya que en la presente publicación no procede analizar diferentes paquetes de software, porque el propósito es simplemente ofrecer una visión de cuál puede ser el instrumento básico a aplicar para llevar a la práctica nuestro proyecto funcional de la Base de Datos. Hemos seleccionado el Sistema DBMS de la Conferencia CODASYL, por entender que es en estos momentos el que tiene un mayor grado de generalidad, y el que está difundido en mayor medida y puede, por lo tanto, considerarse como más normalizado.

La Conferencia de Lenguajes de Sistemas de Datos (CODASYL), agrupa a usuarios y proveedores del sector de la informática, y desde tiempo atrás ha sido un instrumento útil para el estudio de Sistemas comunes que permitan el intercambio de procedimientos y la compatibilidad de software en la comunidad de utilizadores de servicios informáticos en todo el mundo. Fruto de este intento, fue el lenguaje COBOL, hoy en día de uso generalizado en la industria para programar los Sistemas de Información de Gestión.

Desde los años 60, en la Conferencia CODASYL, se ha planteado como objetivo, el estudio permanente de los Sistemas de Gestión de Bases de Datos, y a tal efecto, existe un grupo de trabajo denominado DBTG (Data Base Task Group) que es el órgano de CODASYL para estos propósitos. El DBTG ha producido varios informes, pero fundamentalmente nos vamos a referir al denominado "April 71 Report" que contiene un conjunto de especificaciones básicas para el desarrollo de software básico para la gestión de Bases de Datos, informe que ha sido desarrollado por un conjunto de expertos en el tema, procedentes de diversas Empresas de los Estados Unidos, tanto del campo de los proveedores como del área de los usuarios. Entre aquéllos, participó personal de Honeywell Information Systems, NCR, IBM, UNIVAC, RCA, Control Data, International Computer Limited y Computer Sciences Corporation, y entre éstos, La Bell Telephone Laboratories, Montgomery World,

General Motors, United States Navy, Equitable Life Assurance Society y Goodrich Chemical Company.

El Sistema DBMS es un Sistema host-language y el lenguaje de utilización es una extensión del COBOL, estando realizada la definición de la Base de Datos por un lenguaje ad-hoc que tiene una estructura parecida al COBOL, pero que no es un lenguaje de procedimiento y contiene exclusivamente instrucciones declarativas.

Por su indudable interés, vamos a recoger los elementos más relevantes de los lenguajes de definición y de utilización de datos tal y como han sido especificados en el "April 71 Report" del DBTG, y a dicho informe podrá acudir el lector para examinar en profundidad todo el conjunto de especificaciones. Dicho informe, consideramos es de obligado estudio para todo profesional que haya de abordar un Sistema de Base de Datos o para cualquier estudio sobre el tema.

Como presentación, vamos a reproducir en la figura 46, el prólogo de dicho informe de Abril de 1971, cosa que por otra parte es de obligada referencia, pues el referido informe es la fuente original de donde hemos tomado las especificaciones que en el presente capítulo estamos comentando.

5.1. OBJETIVOS DEL SISTEMA DBMS

A continuación reseñamos el conjunto de características que el DBTG se impuso como objetivo a la hora de desarrollar las especificaciones del DBMS.

- Estructurar los datos de la forma más conveniente a cada aplicación, sin atender al hecho de que muchos datos son usados por otras aplicaciones, y alcanzar tal flexibilidad sin que exista redundancia de datos.
- Permitir que varias unidades de tratamiento, accedan en forma concurrente a la Base de Datos para operaciones de lectura o de actualización.
- Proporcionar y permitir el uso de diferentes estrategias de búsqueda.
- Protección de la Base de Datos ante accesos no autorizados y ante posible conflicto por interacción de varios programas.
- Facilidades para controlar centralizadamente, el almacenamiento físico de los datos.
- Independencia de datos y programas.
- Variedad de estructuras lógicas de datos, incluso las estructuras de Red.
- Establecer descripciones separadas de los datos que se encuentran en la Base de Datos y de los datos que un programa ha de manejar.
- Describir la Base de Datos de forma universal, no asociada de forma decisiva a ningún lenguaje en particular.
- Disponer una arquitectura que permita que la descripción de la Base de Datos y la propia Base de Datos, puede ser compatible con diferentes lenguajes de proceso.

CODASYL DATA BASE TASK GROUP REPORT / ABRIL 1971

PROLOGO

La Conferencia de Lenguajes de Sistemas de Datos (CODASYL) tiene el placer de ofrecer para su distribución general un segundo informe del "Data Base Task Group" (DBTG). Este informe refleja muchos de los cambios, mejoras y extensiones sugeridas como resultado de la publicación del primer informe. También refleja la petición del Comité Ejecutivo de Codasyl de que las especificaciones se separen en dos categorías, a saber: 1) aquellos elementos que serían necesarios para un lenguaje independiente de descripción de los datos, y 2) aquellos elementos que permitirían al Cobol interrelacionarse adecuadamente con los datos definidos en este lenguaje independiente de descripción de los datos.

Con este informe como base, Codasyl ha establecido ahora un nuevo Comité, el Comité de Lenguajes de Descripción de Datos, independiente, pero análogo, a los Comités de Planificación, de Sistemas y de Lenguajes de Programación. Se prevé que desde la Base ya establecida por el informe del DBTG, este nuevo Comité rematará las Especificaciones de un lenguaje común de descripción de datos, independiente de, pero común a, muchos otros lenguajes de programación de alto nivel. Simultáneamente, el Comité de Lenguajes de Programación, desarrollará extensiones del Cobol basados en la Sección 4 (Sub-esquema Cobol) y Sección 5 (Cobol DML). Se espera que las Organizaciones responsables de otros lenguajes de programación, desarrollarán sub-esquemas y lenguajes de manejo de datos apropiados a las características de su lenguaje.

Aunque este informe no está clasificado como unas especificaciones definitivas de Codasyl para un lenguaje común, representa muchos años de trabajo y el mejor pensamiento de muchos expertos reconocidos. Por lo tanto, creemos que es práctico y apropiado para hacer implementaciones basadas en estas especificaciones, de tal forma que pueda establecerse una Base de experiencia para un desarrollo y perfeccionamiento futuro de las especificaciones de un lenguaje común para la descripción de datos.

Todo comentario será bienvenido, y les estimulamos a participar en este nuevo y excitante empeño.

Firmado: J. L. JONES

Presidente del Comité Ejecutivo de Codasyl.

Como puede verse, este paquete de objetivos, son condiciones universales que un Sistema eficiente de Base de Datos tiene que poseer, y está en línea con el tipo de instrumento que a lo largo de los capítulos anteriores, se sugiere como necesario.

5.2. DESCRIPCION GENERAL DEL SISTEMA DBMS

En base a la característica de sistema host-language del Sistema DBMS, por un lado se establece la definición de la estructura lógica y física de los datos mediante un lenguaje de definición de datos denominado DDL (Data Definition Language). Este procedimiento para definición de datos, permite que la Base de Datos tenga una vida con entidad propia, independiente de los programas que la utilizan.

Mediante el lenguaje DDL, se obtiene la definición en lenguaje fuente de todo el conjunto de ficheros y relaciones lógicas que estarán albergados en las unidades de almacenamiento de datos. La definición obtenida con este lenguaje declarativo, se denomina “esquema fuente”, el cual será compilado por el procesador DDL y de esta forma se obtendrá el “esquema objeto” que constituye la definición de la Base de Datos que habrá de ser residente en la memoria principal, sirviendo de interfase y soporte en la explotación del Sistema.

El Esquema así obtenido, constituye la más completa definición de toda la Base de Datos, pero también se contempla el que puedan existir Sub-esquemas o definiciones parciales que corresponden a sub-conjuntos de la Base de Datos. Los Sub-esquemas serán invocados por los programas de utilización, de manera que existirá una correspondencia unívoca de cada programa de utilización con un determinado Esquema que complementará al programa de utilización para que pueda operar en la Base de Datos. Obviamente, un mismo Sub-esquema podrá dar servicio a varios programas.

La motivación del concepto de Sub-esquema, viene a tener dos principales bases:

- La simplicidad de uso, pues el diseñador de una aplicación, sólo tendrá que relacionarse y utilizar un sub-conjunto de la Base.
- Las garantías de protección que ofrece el poder delimitar a nivel de sub-esquema el campo de actividad de un programa.

Por supuesto, las definiciones que los sub-esquemas establecen, han de ser compatibles entre sí y no estar en contradicción lógica con el Esquema.

La utilización de la Base de Datos, se realiza por medio del lenguaje COBOL, complementado por un paquete de macroinstrucciones que en conjunto se denomina DML (Data Manipulation Language). Un programa de aplicación que utiliza la Base de Datos será por lo tanto escrito en COBOL más DML, utilizando las macro-instrucciones del DML para las operaciones de acceso a la Base de Datos y el resto de instrucciones COBOL, para determinar los procesos que la aplicación re-

quiere, e incluso los accesos convencionales a otros ficheros ajenos a la Base de Datos.

Los programas de utilización, podrán ser procesados de forma simultánea, residiendo varios de ellos a un tiempo en la memoria principal, lo que obligará a que asimismo cohabiten en la memoria principal, los respectivos sub-esquemas que dichos programas utilizan.

En la figura 47, tomada del informe del DBTG, se representan de forma ideográfica, los diferentes elementos conceptuales que participan en la explotación de la Base de Datos, cuya secuencia de uso vamos a exponer para mayor ilustración, siguiendo la numeración asignada a las flechas indicadoras:

- 1) El programa de utilización número 1, solicita del paquete de control DBMS, un determinado dato del almacenamiento, naturalmente por medio de una instrucción de DML.
- 2) El DBMS, analiza la petición y complementa los argumentos de búsqueda con la información contenida en el Sub-esquema 1 invocado por el programa y en el Esquema.
- 3) El DBMS, requiere al Sistema operativo de la instalación, para la oportuna operación de Entrada/Salida.
- 4) El Sistema operativo interactúa en el almacenamiento.
- 5) Se produce una transferencia de datos entre el almacenamiento y el buffer general del Sistema.
- 6) El DBMS provoca que los datos que ahora están en el buffer del Sistema, sean depositados en el área de trabajo ad-hoc del programa de utilización 1 que solicitó dichos datos.
- 7) El DBMS proporciona información de estado, para control del programa, de sus relaciones con la Base de Datos.
- 8) El dato que ahora se encuentra en el área de trabajo, puede ser manipulado en memoria de acuerdo con los procedimientos que el programa contiene.
- 9) El DBMS, administra el buffer del Sistema, el cual es compartido por todos los programas que utilizan en forma simultánea la Base de Datos.

5.3. DDL O LENGUAJE DE DEFINICION DE DATOS

El lenguaje DDL para redacción de un Esquema, o definición de una estructura de datos, consta de 4 tipos de "Entrada" las cuales se utilizan para:

- Identificar el Esquema.
- Definir las Areas lógicas de almacenamiento.
- Definir los Registros tipo o Ficheros.
- Definir los Sets o relaciones lógicas del Esquema.

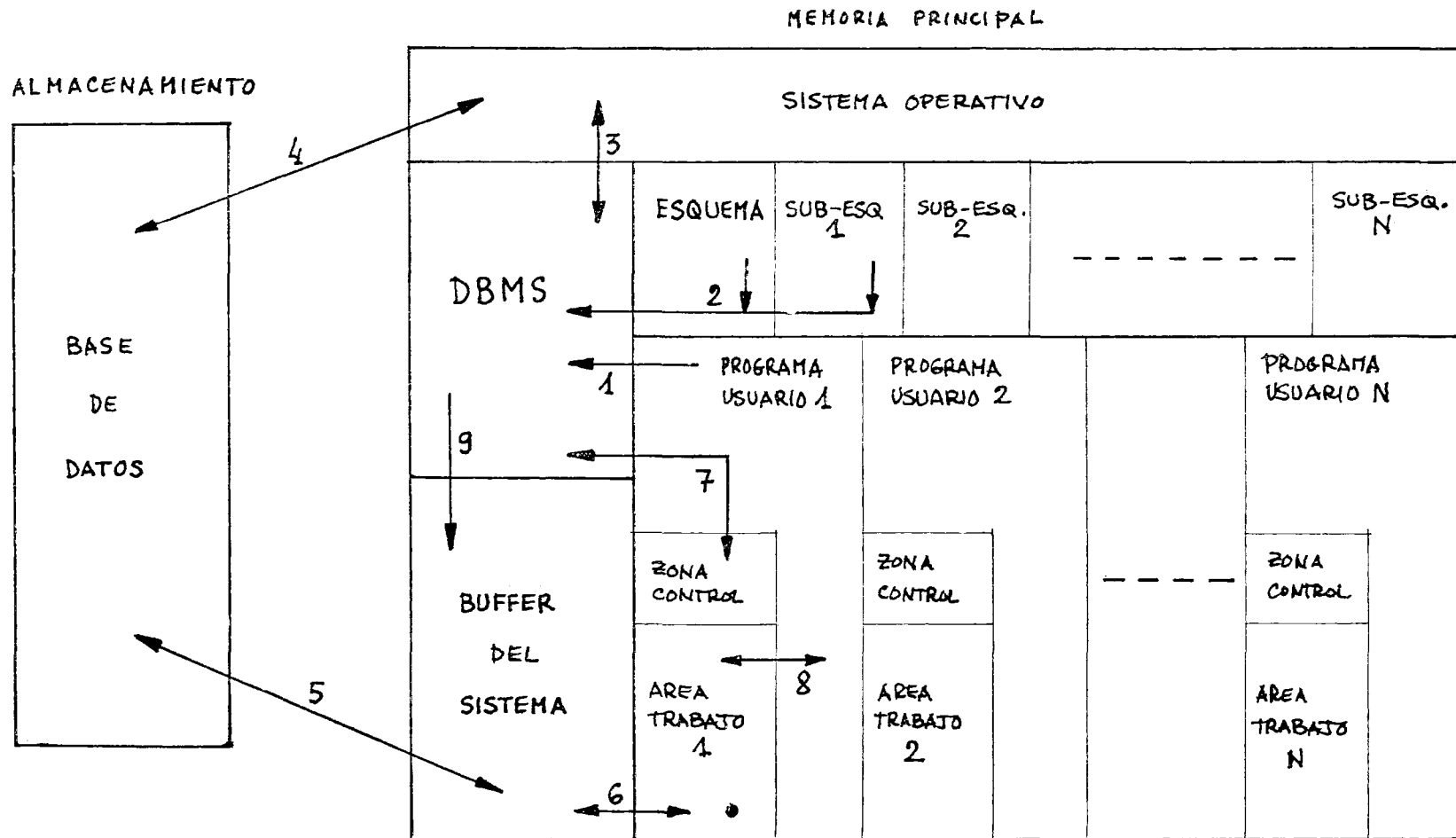


FIG. 47. Sistema DBMS

La Entrada que define el Esquema, deberá ser única, mientras que las Entradas que definen Areas, Registros y Sets, deberán ser tantas en número como elementos de dichos 3 tipos existan para definir.

Una Entrada consiste en una o más Cláusulas, que en el caso de descripción de Registros o de Sets se agrupan también formando Sub-entradas, pudiendo estas Sub-entradas repetirse varias veces dentro de una misma Entrada. En la descripción de las instrucciones, se terminarán las Entradas y las Sub-entradas por un punto, y dentro de las Entradas y Sub-entradas se describirán en secuencia las Cláusulas, las cuales pueden contener a su vez otras Cláusulas subordinadas.

En las figuras que a continuación expresarán las Entradas, Sub-entradas y Cláusulas del DDL, se ha utilizado un metalenguaje de presentación en el que se incluye en letras mayúsculas el texto fijo de las instrucciones, y en letras minúsculas los elementos que deben ser definidos por el Analista o Programador. Los elementos en "negrita", constituyen texto obligado aun cuando puede ser opcional, mientras que el texto en letras mayúsculas normales, siempre es de utilización opcional (*).

Los corchetes rectos utilizados, expresan módulos opcionales y que por lo tanto, pueden ser no utilizados. Si dentro de este corchete recto hubiera más de una opción, hay que tener en cuenta que solamente puede ser escogida una de ellas.

Las rectas dobles paralelas, ofrecen varias alternativas de especificación de las cuales obligadamente ha de seleccionarse una, pero opcionalmente pueden especificarse varias o incluso todas. Los corchetes arqueados, incluyen opciones de las que debe de seleccionarse una y solamente una.

No vamos a entrar en los tipos de caracteres y símbolos que se ofrecen en el lenguaje DDL por considerar que son detalles irrelevantes.

5.3.1. ENTRADA DE ESQUEMAS

Esta Entrada cuyo formato se expone en la figura 48, es única para cada Esquema, identifica dicho Esquema en el Sistema, para lo cual el Administrador o Analista le asigna un nombre simbólico que servirá para su identificación de forma unívoca. Tiene un paralelismo esta Entrada con la Identification Division del COBOL.

Opcionalmente, puede establecerse una Cláusula PRIVACY, cuya función es el establecer protecciones sobre dicho Esquema, ante las siguientes operaciones que pueden realizarse y que pertenecen a funciones de Administración de la Base de Datos:

- ALTER. Esta operación permite, a nivel administración, la alteración de todo el Esquema a excepción de sus Cláusulas PRIVACY.
- COPY. Cuando está especificada en lenguaje DDL para definir un Sub-

* Nota.—En el "Abril 1971 Report" el texto de uso obligado figura subrayado en lugar de figurar en "negrita".

esquema, permite que en dicho Sub-esquema se incluyan Entradas que corresponden al Esquema.

- **DISPLAY.** Permite la lectura del Esquema con excepción de sus claves de privacidad.
- **LOCKS.** Permite leer, crear o cambiar las Cláusulas del Esquema que se refieren a Protecciones.

Mediante la Cláusula **PRIVACY**, se pueden disponer protecciones para cada una de las operaciones mencionadas, y dichas protecciones pueden establecerse por un literal, una clave paramétrica o en base a un procedimiento.

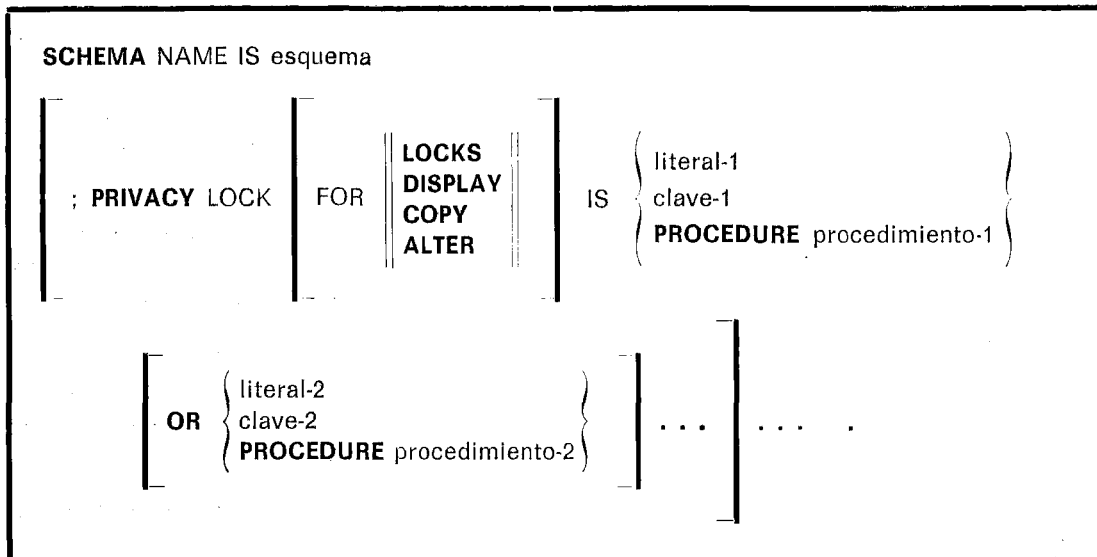


Fig. 48. Entrada SCHEMA

Si las claves de protección son Literales, su valor podrá ser modificado por una operación **LOCKS** de Administración, y en caso contrario por lo general, permanecerá constante. Si utilizamos una Clave simbólica, su valor siempre puede ser inicializado por un programa, y si utilizamos un Procedimiento, podremos establecer un criterio de protección tan complejo como sea necesario.

Cuando un programa pretende realizar las operaciones indicadas que afectan al Esquema, la clave de Protección de que dispone el programa, será chequeado con las claves definidas para el Esquema, y en caso de que no coincidan, la operación será invalidada.

5.3.2. ENTRADA DE AREA

En la figura 49, se presenta el formato propuesto por CODASYL para la definición de las Areas lógicas de la Base de Datos, formato que como los demás, ha sido tomado del "April 71 Report" del DBTG.

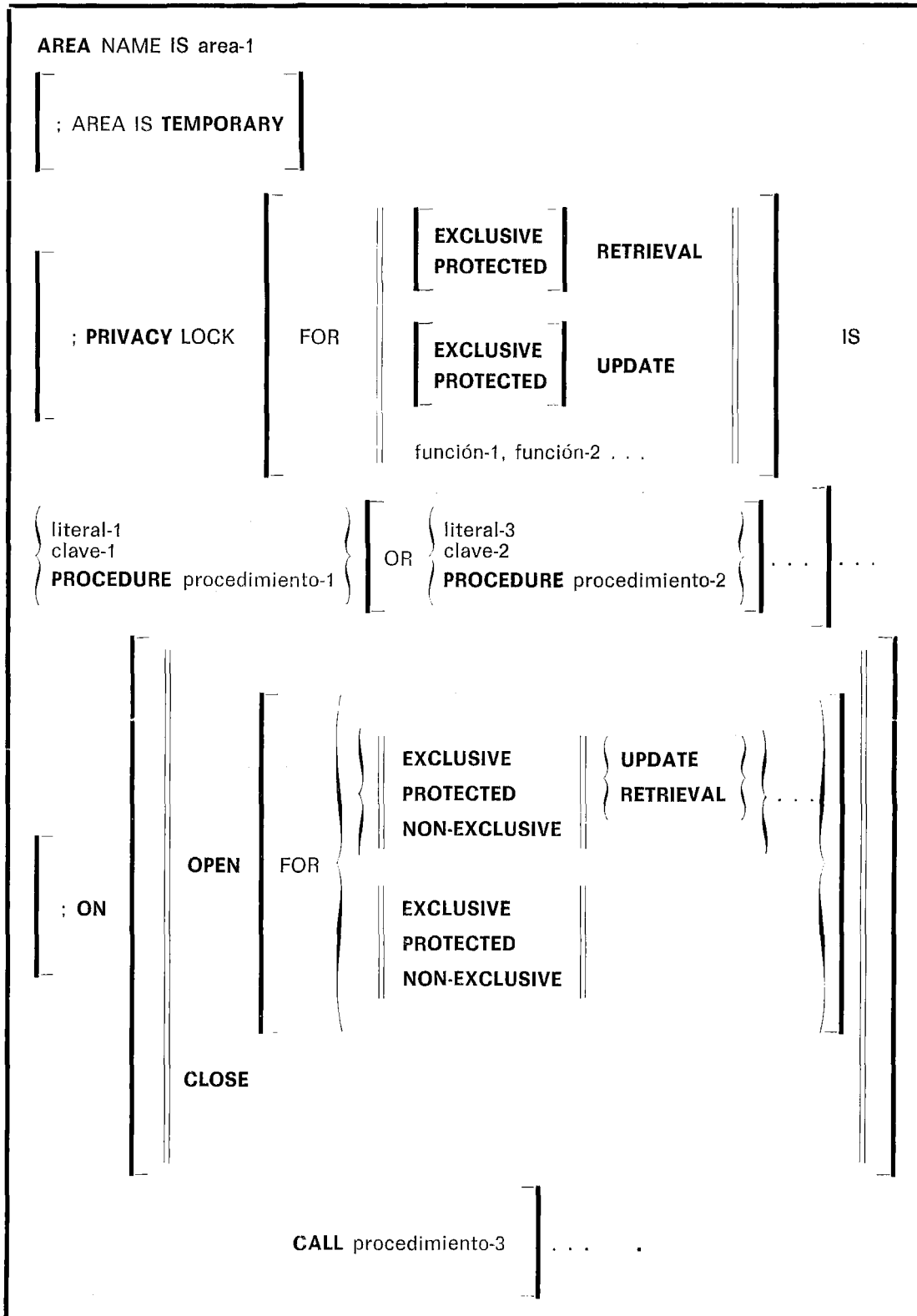


FIG. 49. *Entrada AREA*

Deberá incluirse una Entrada de este tipo para cada una de las Areas lógicas de almacenamiento que existan en la Base de Datos, y cada una de ellas, deberá ser identificada por un nombre "área-1" que el Administrador de la Base de Datos debe de asignar, nombre que será utilizado por los programas para acceder a dicha Area.

La Cláusula TEMPORARY, califica a un Area, obligándola a que no pueda ser compartida por unidades de tratamiento concurrentes, es decir, sólo podrá ser accedida a un tiempo por una sola unidad de tratamiento. Los registros que están en un Area de este tipo, no pueden participar ni como miembros ni como registros principales de Sets que contengan registros que no están en Areas de este tipo.

La Cláusula PRIVACY tiene una vocación análoga a la ya indicada en la Entrada de Esquema y su funcionamiento operativo es del mismo tipo, aunque en este caso se referirá al uso de este Area en particular. Las figuras que son objeto de protección, son diferentes utilizaciones que un programa de proceso puede realizar en el Area:

- *Exclusive Retrieval.* Función que permite la recuperación de registros de ese Area en forma exclusiva.
- *Protected Retrieval.* Función que permite garantizar que los registros recuperados de ese Area durante un proceso, no van a ser modificados por el acceso concurrente de otros programas.
- *Exclusive Update.* Función que permite que un programa actualice datos en un Area con exclusividad.
- *Protected Update.* Se trata de la actualización de datos en el Area por una unidad de tratamiento en forma protegida de otras actualizaciones que pudieran ser concurrentes.
- Otras funciones de soporte tales como carga inicial de datos, copiado de Area, preparación de errores detectados en mantenimiento, ... etc., que corresponden a operaciones de Administración.

Para que un programa pueda abrir un Area para utilización bajo una de estas condiciones, es preciso que el programa presente la clave adecuada que será contrastada con la definición establecida en la Cláusula PRIVACY para autorizar o desautorizar la operación.

La Cláusula ON, permite determinar opcionalmente el que sea ejecutado un procedimiento de programación denominado en la figura "procedimiento-3", cada vez que un programa intenta realizar las operaciones de abrir o cerrar el Area en la forma en que la propia Cláusula ON establezca. Si no se incluyen las opciones OPEN o CLOSE en la Cláusula ON, el procedimiento sería ejecutado en cualquiera de los casos.

El procedimiento es ejecutado inmediatamente antes de que el control de la secuencia de instrucciones vuelva a ser ejercido por la unidad de tratamiento. Son parámetros de entrada en el procedimiento, el código de comando que se ha tratado de ejecutar y el nombre del Area involucrada, tomándose también como argu-

mento de entrada, el código de error de la zona de control de la unidad de tratamiento, código que ofrecerá un diagnóstico de alguna posible circunstancia que se haya producido al efectuar la operación en el Area.

El procedimiento definido por la Cláusula ON, nos permite establecer criterios específicos de decisión ante excepciones ocurridas en las operaciones de apertura y cierre de Areas.

5.3.3. ENTRADA DE REGISTRO

La Entrada de Registro, se divide en dos tipos de Sub-entrada, la Sub-entrada de Registro y la Sub-entrada de Dato. Para cada tipo de Registro, es decir, para cada Fichero contemplado en la Base de Datos, deberá incluirse una Sub-entrada de Registro, la cual define la forma de localización de dicho Registro en la Base de Datos, las zonas autorizadas para almacenar ese Registro, y unas especificaciones adicionales de protección y de control de errores.

Para cada Sub-entrada de Registro, se dispondrán tantas Sub-entradas de Datos como campos deban ser definidos del Registro, en una forma análoga de estructura de niveles a la utilizada en las especificaciones de la Data Division del COBOL.

La Sub-entrada de Registro, identifica a cada Fichero con un nombre simbólico "registro-1", nombre que es el utilizado por los programas al referirse al Fichero.

La forma de localización y almacenamiento del Registro en la Base de Datos puede ser establecida por una de las tres siguientes fórmulas:

- DIRECT. Bajo esta forma de localización, la dirección lógica de almacenamiento de Registros, es la definida por el nombre simbólico contenido en la Cláusula, el cual debe ser previamente inicializado con una clave de Base de Datos, es decir, con la dirección lógica específica que se quiere acceder o donde se quiere almacenar. El campo simbólico puede estar definido por un identificador dispuesto por el Administrador por un dato contenido en el propio segmento. Se denomina como acceso directo porque los programas de proceso manejan direcciones lógicas en directo. Hay que puntualizar, que el Sistema DBMS, mantendrá las oportunas correspondencias biunívocas de estas claves lógicas con posiciones físicas en los discos, lo cual es transparente para las aplicaciones. Es decir, no es un acceso directo pero desde el punto de vista lógico, se utiliza como tal.
- CALC. El cálculo de direcciones para acceso o almacenamiento de Registros, es realizado por un procedimiento de cálculo establecido por el Administrador o Analista, que utiliza como argumentos de entrada, determinados parámetros que habrán de ser inicializados previamente por las instrucciones de la unidad de tratamiento, y que pueden ser valores de campos que el propio registro objeto contiene.

El procedimiento de cálculo, ofrece como valor de salida, la dirección del Registro objeto. El Administrador puede por lo tanto diseñar procedimientos para cálculo de direcciones en la forma más conveniente para

la compactación de datos en la Base, porque dicho procedimiento puede contemplar criterios de densidad, y por lo tanto de reparto de Registros en las Areas y Páginas.

- VIA SET. La forma de localización del Registro, se realiza por un camino lógico derivado de la participación de dicho Registro en el Set "set-1". De acuerdo con la definición del Registro objeto como miembro del referido Set, definición que más adelante estableceremos al describir la Entrada de Set, se seleccionará un determinado anillo específico y un punto lógico de inserción del Registro objeto dentro del anillo en cuestión, y si la operación a realizar es de almacenamiento, éste será realizado en la posición libre disponible más próxima a dicho punto lógico de inserción, naturalmente en compatibilidad con las posiciones autorizadas para el almacenaje de ese tipo de Registros. Si la operación a realizar es de recuperación, el camino lógico que la definición del Set establece permitirá seleccionar la dirección del Registro buscado.

La Cláusula WITHIN, permite definir las Areas en las que los Registros del fichero pueden ser almacenados.

La Cláusula opcional ON, provoca la ejecución de un procedimiento de cálculo definido por el Analista o por el Administrador cuando un Registro del fichero es accedido o almacenado por un programa, pudiendo cualificar el tipo de instrucción o comando que provocará la ejecución de dicho procedimiento. El procedimiento tiene una vocación análoga al ya expresado para la Entrada de Area.

La Cláusula PRIVACY, establece protecciones ante operaciones de acceso o almacenamiento por parte de los programas, pudiendo cualificarse según se ve en la figura 50, el tipo o tipos de operación que van a ser objeto de protección. El mecanismo operativo de las claves de protección, es el mismo que ya hemos indicado anteriormente.

Las figuras 51 y 52, establecen el formato de definición de la Sub-entrada de Datos que va a servirnos para determinar los campos de Registros. Cada dato deberá poseer un número de nivel que es opcional y un nombre simbólico, en forma análoga a la definición de datos en COBOL, por lo que no nos vamos a extender en la exposición. Igualmente, podrán existir unas Cláusulas PICTURE, TYPE y OCCURS, que describen respectivamente los caracteres que el campo de datos va a contener, el tipo de campo de que se trata, y el número de veces que dicho campo se repite en el Registro.

Las Cláusulas RESULT y SOURCE, pueden utilizarse para controlar la actualidad del Dato y vigilar sus actualizaciones en base a un procedimiento o al valor dado por un determinado identificador.

La Cláusula CHECK, permite controlar mediante un procedimiento, la validez de un dato después de cualquier operación que con él se haya realizado.

La Cláusula ON, permite opcionalmente, la ejecución de un procedimiento de decisión inmediatamente antes de que después de una operación STORE, GET o

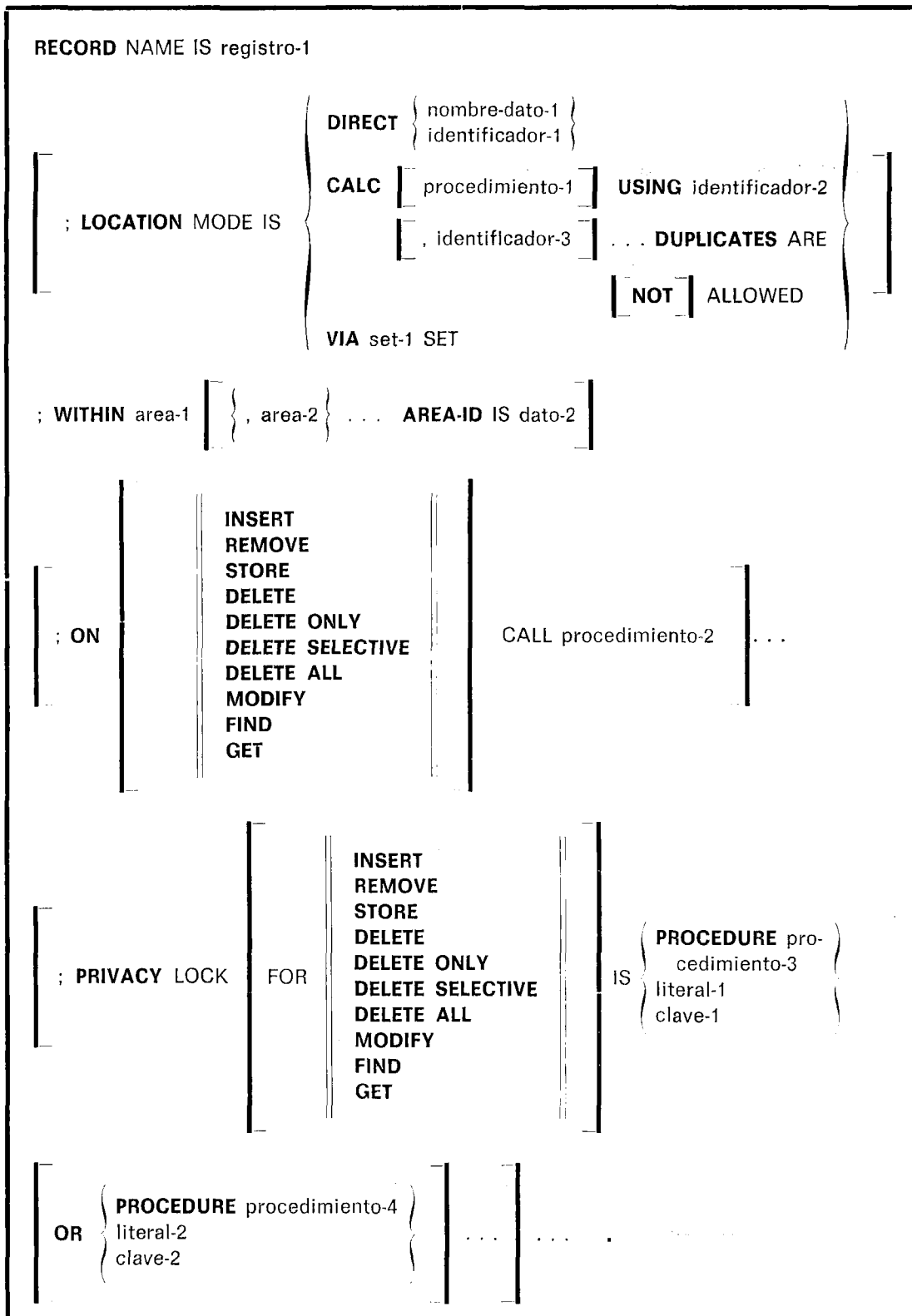


Fig. 50. Sub-entrada RECORD

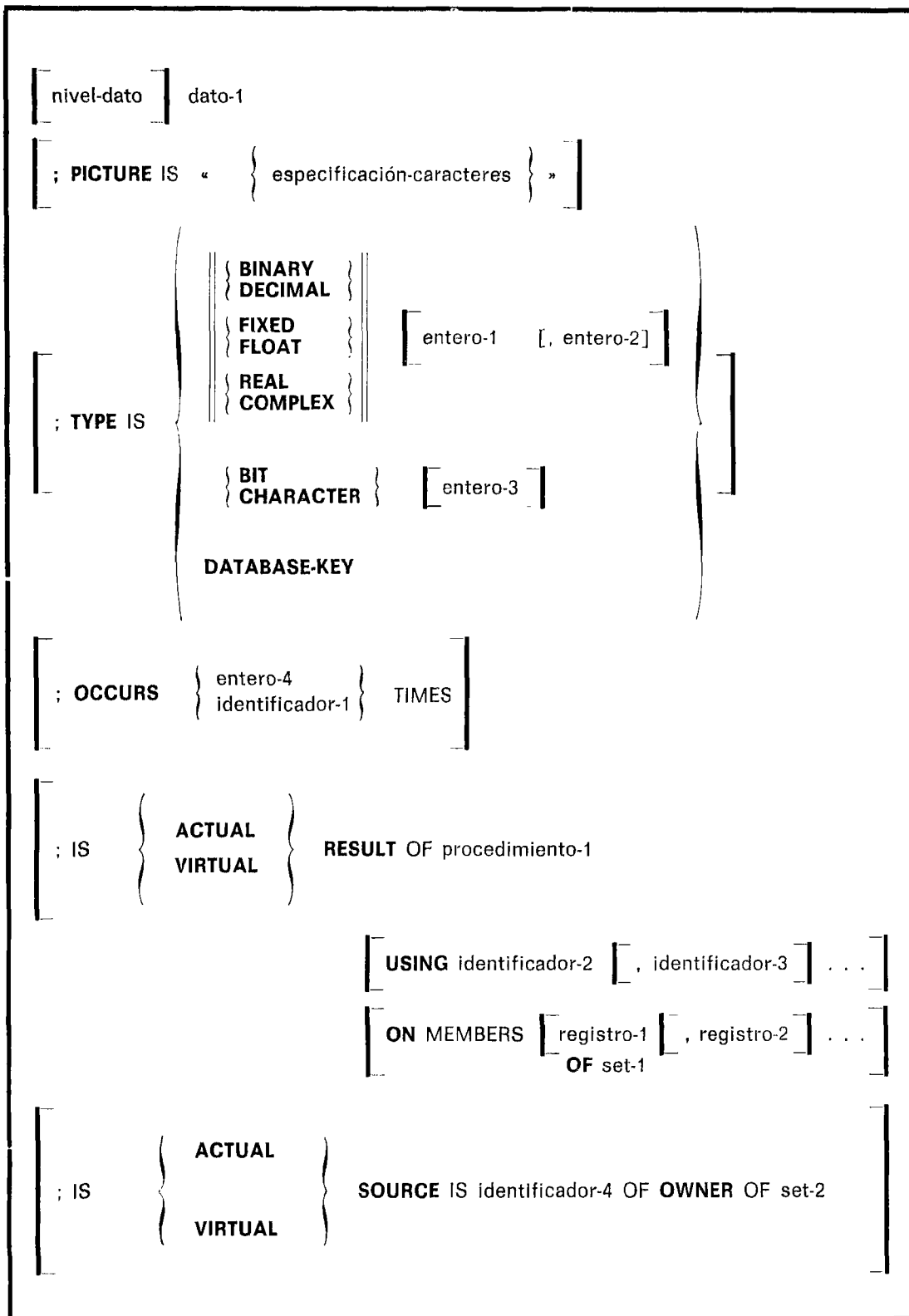


FIG. 51. Sub-entrada DATO

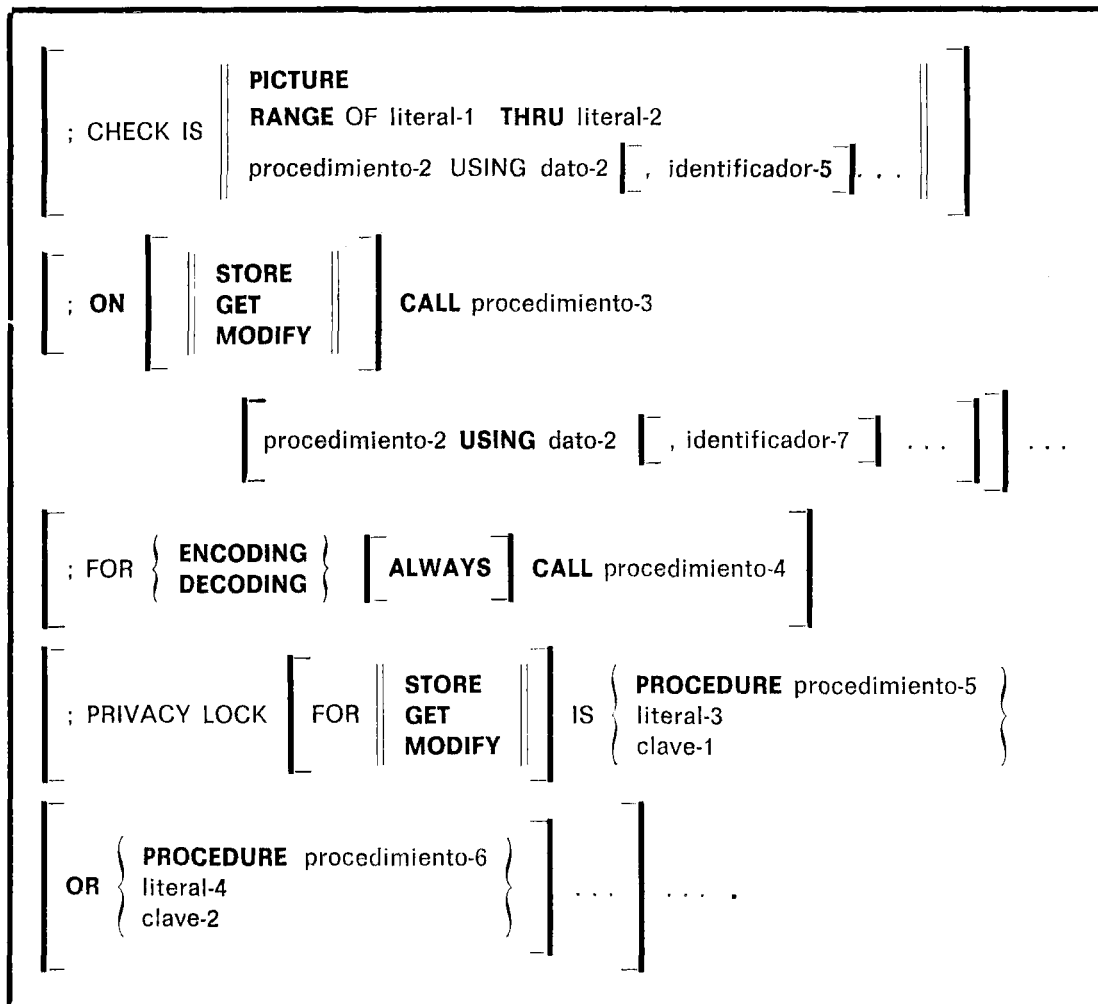


Fig. 52. Sub-entrada DATO (continuación)

MODIFY, el control de la secuencia de instrucciones, sea devuelto a la unidad de tratamiento.

También pueden establecerse claves de Protección, mediante la Cláusula PRIVACY, para cada uno de los Datos del Registro.

5.3.4. ENTRADA DE SET

La Entrada de Set, está compuesta por dos tipos de Sub-entrada. La Sub-entrada de Set, identificará y definirá cada uno de los Sets definidos en el Esquema de la Base de Datos, y para cada Sub-entrada de Set, se incluirán tantas Sub-entradas de Miembros como número de ficheros miembros se declaren como posibles participantes en dicho Set.

En la figura 53, se presenta el formato que CODASYL propone para determinar

SET NAME IS set-1

; MODE IS { CHAIN [LINKED TO PRIOR]
 { POINTER - ARRAY [DYNAMIC] }

Format 1

; ORDER IS [ALWAYS] { FIRST
 { LAST
 { NEXT
 { PRIOR }

Format 2

; ORDER IS SORTED [INDEXED [NAME IS indice-1]]

WITHIN RECORD-NAME
BY DATABASE-KEY

DUPLICATES ARE

[FIRST
 LAST
 NOT] ALLOWED

[; ON [ORDER
 INSERT
 REMOVE] CALL procedimiento-1 ...

[; PRIVACY LOCK FOR [ORDER
 INSERT
 REMOVE
 FIND] IS { PROCEDURE procedimiento-2
 { literal-1
 { clave-1 }

[OR { PROCEDURE procedimiento-3
 { literal-2
 { clave-2 }] ...

; OWNER IS { registro-1
 { SYSTEM }

Fig. 53. Sub-entrada SET

un Set. El nombre simbólico "set-1" identificará de forma unívoca a la unidad lógica que estamos describiendo y será un nombre utilizado por las unidades de tratamiento.

La Cláusula **MODE**, establece de los dos tipos de Set que el Sistema contempla, cuál es al que el Set definido pertenece. Los Sets pueden ser en forma de cadena, o mejor dicho, de anillo, especificando opcionalmente enlaces dobles. También puede estar organizado el Set en base a la secuencia establecida en una matriz de pointers enlazada al Registro principal del Set.

La Cláusula **ORDER** puede, como vemos, tener dos formatos diferentes:

- El formato 1, permite establecer dónde un nuevo Registro va a ser situado dentro de un Set específico, admitiéndose la posibilidad de que sea situado en primer lugar, en último lugar, o en la situación siguiente o anterior a un determinado Registro de referencia, cuya dirección puede ser controlada por un procedimiento de programa.
- El formato 2, permite disponer un orden clasificado en base a claves, por lo que un Registro irá a incorporarse en la secuencia del anillo, en el punto donde su propia clave determine.

La Cláusula **ON**, permite la ejecución de un procedimiento diseñado al efecto, que será ejecutado antes de completarse un acceso de un programa al Set que nos ocupa, pudiendo disponerse selectivamente que ello ocurra, cuando las operaciones que el programa está realizando, sean **ORDER**, **INSERT** o **REMOVE**. El procedimiento se ejecuta y decide antes de que el control de instrucciones sea devuelto al programa de proceso.

La Cláusula **PRIVACY**, dispone de procedimientos de protección análogos a los que ya hemos establecido para otras Entradas.

La Cláusula **OWNER**, de uso obligatorio, indica cuál es el fichero principal del Set.

La Sub-entrada **Miembro**, cuyo formato se ofrece en la figura 54, define a cada uno de los Ficheros que son autorizados para que sus Registros participen como miembros en el Set, pudiendo cualificar si un Registro siempre ha de formar parte de forma obligada del Set o dicha circunstancia es opcional a decidir por un procedimiento de programa.

El miembro, puede definirse como Automático o Manual. Un miembro Automático es aquel que utilizará un procedimiento standard para insertar el Registro en el Set al cual pertenece, mientras que un miembro Manual, deberá ser insertado una vez almacenado el Registro por un procedimiento ad-hoc aportado por la unidad de tratamiento interesada en la operación.

Para cada Miembro, puede opcionalmente disponerse, el que exista un pointer de enlace con el Registro principal, figura que puede resultar útil para procesos de anillos que requieran sucesivas "vueltas" al Registro principal. También se puede establecer si se permite la duplicidad de Registros Miembros en un anillo, en-

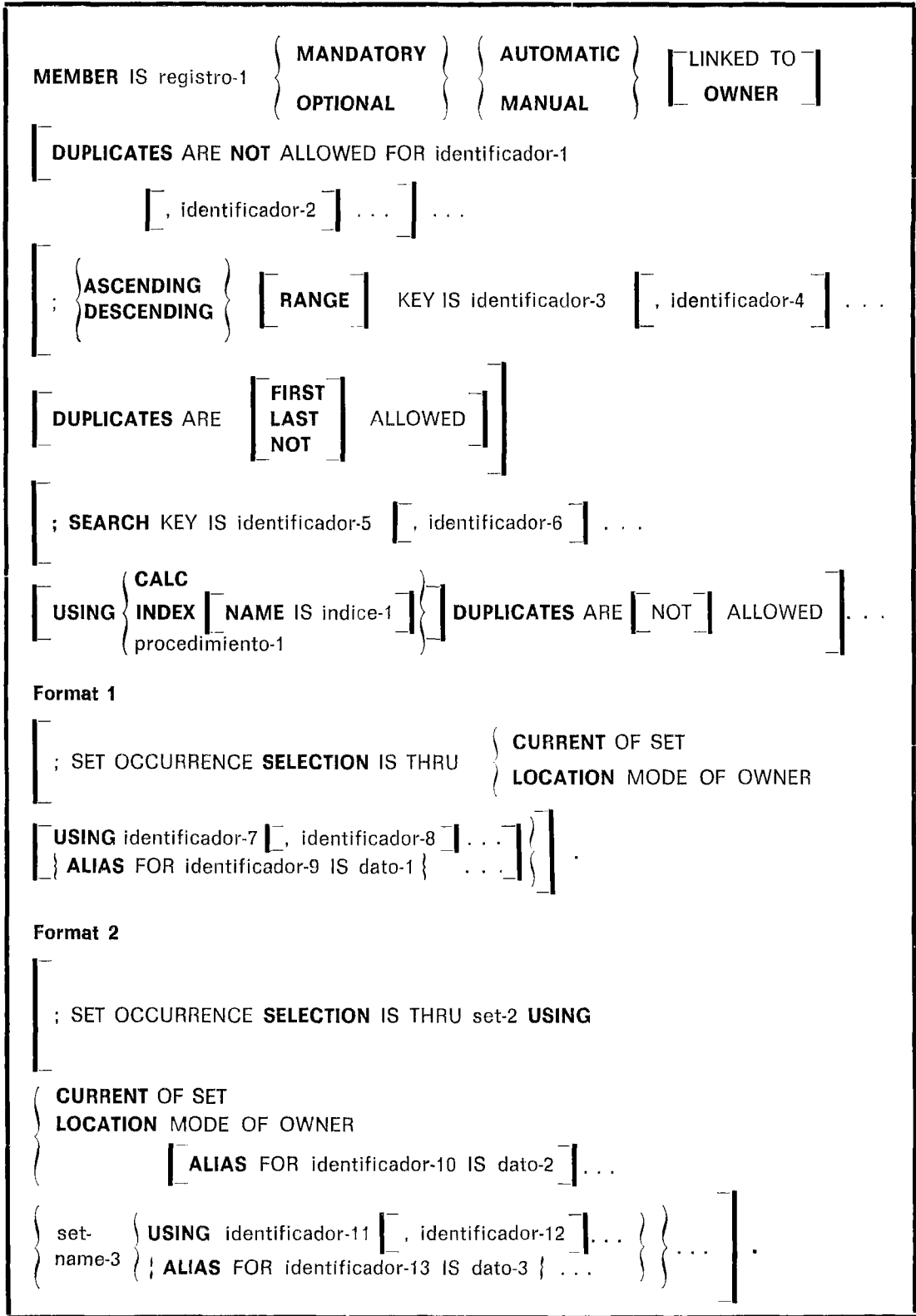


FIG. 54. Sub-entrada MIEMBRO

tendiendo por duplicidad la coincidencia de claves en base a las cuales los Registros son clasificados.

La Cláusula KEY, establece las claves para clasificación de los Registros Miembros en los anillos, lo cual define por completo el criterio de ordenación de cada fichero miembro dentro de los anillos, y por extensión define asimismo el orden total de los anillos del Set.

La Cláusula SELECTION, que tiene dos formatos, viene a definir el procedimiento para seleccionar un anillo específico dentro de todos los que pertenecen a la misma familia definida por el Set tipo.

El primer formato, establece que la forma de selección de un anillo, puede realizarse en base a la forma de localización que para el fichero principal del Set se ha establecido, en base a la correspondencia biunívoca que existe entre los Set específicos y los registros del fichero principal del Set.

También puede seleccionarse la opción CURRENT, que determina que la selección de anillos sea realizada en base a la referencia de dirección de uno de los registros contenido en la familia de Sets, dirección que por corresponder a un Registro único y determinado, de hecho implica ya la selección del anillo al cual pertenece. Este procedimiento de selección, permite ser manejado a nivel de procedimiento.

El formato 2, viene a ser análogo al formato 1, pero con la diferencia de que permite establecer todo un camino lógico a través de las relaciones de Sets, para definir la selección de anillos. Tengamos en cuenta que utilizando el formato 1, y utilizando la opción LOCATION, la forma de localización del Registro principal puede estar definida en su correspondiente Entrada de Registro, como VIA SET, lo que implica la nueva aplicación de selección de anillos utilizando la nueva fórmula que hayamos establecido en la Entrada de Miembro correspondiente a ese Registro principal cuando está definido en su condición de miembro del nuevo Set.

5.4. EL DML O LENGUAJE DE UTILIZACION DE LA BASE DE DATOS

El lenguaje DML, es una extensión del lenguaje COBOL que de esta forma se utiliza como "host-language", y constituye el procedimiento para que las aplicaciones informáticas accedan a la Base de Datos.

Vamos a referirnos a las modificaciones que el DML introduce en la Identification Division, en la Data Division y en la Procedure Division que el programa de aplicación va a tener como programa estructurado en lenguaje COBOL.

Vamos a señalar como concepto de interés a tener en cuenta, la utilización de datos específicos para contener direcciones lógicas de la Base de Datos, datos que

los denominamos Claves de Base de Datos. La utilización de este tipo de Claves, nos permite manejar direcciones lógicas de almacenamiento.

También vamos a manejar en los programas de proceso, determinados conceptos que referidos a la Base de Datos, tienen una personalidad definida, como pueden ser nombres simbólicos de Areas, nombres de Registros, códigos de error, ... etcétera, cuyo significado trasciende incluso del propio programa de aplicación, pues tiene que ser compatible con el orden establecido por el órgano Administrador de la Base de Datos.

A continuación presentamos las instrucciones de DML que el DBGT propone en su informe de Abril de 1971.

5.4.1. INSTRUCCIONES DML EN LA IDENTIFICATION DIVISION

Además de las instrucciones propias de identificación del programa COBOL, se puede incluir opcionalmente, las claves de privacidad que recogemos en las figuras 55 y 56.

La **PRIVACY KEY FOR COMPILE**, se emplea para invocar en la Data Division, un Sub-esquema especificado que tiene una protección declarada para él, "literal-1", es la clave de protección que el programa presenta y que será contrastada en el Sub-esquema correspondiente.

A continuación puede expresarse opcionalmente, claves de protección o procedimientos de protección para el acceso a Areas por parte del programa, pudiendo selectivamente establecer Areas determinadas o incluir operaciones completas determinadas sobre dichas Areas. Como puede advertirse en el formato de la figura 55, el chequeo de protección puede realizarse en base a claves o incluso en base a procedimientos de decisión. La mecánica de funcionamiento de las claves de Protección, es análoga a la esbozada al tratar de la definición del Esquema, pues constituyen con dichas definiciones anteriores, aspectos duales.

También pueden establecerse protecciones para las operaciones concretas de **INSERT**, **REMOVE**, **STORE**, **DELETE**, ... etc., que corresponden a instrucciones específicas de acceso de Registros de la Base de Datos que más adelante se determinan. Estas protecciones serán aplicadas cuando se produce una de estas operaciones con los Registros de la Base, o selectivamente con alguno o algunos de ellos, pues dicha diferenciación permite establecer las instrucciones que en la figura 55 se exponen.

Igualmente se pueden establecer protecciones a nivel de datos internos a los Registros, siendo de aplicación el control de protección cuando otros datos van a ser objeto de alteración, siendo en este caso las condiciones de funcionamiento de las claves, análoga al caso de Registros.

También pueden disponerse **PRIVACY KEY** referidas a Sets para ser aplicadas como órgano de protección para aquellas operaciones que suponen alteraciones de orden lógico en los Sets, tales como **ORDER**, **REMOVE** o **INSERT**, pudiendo en

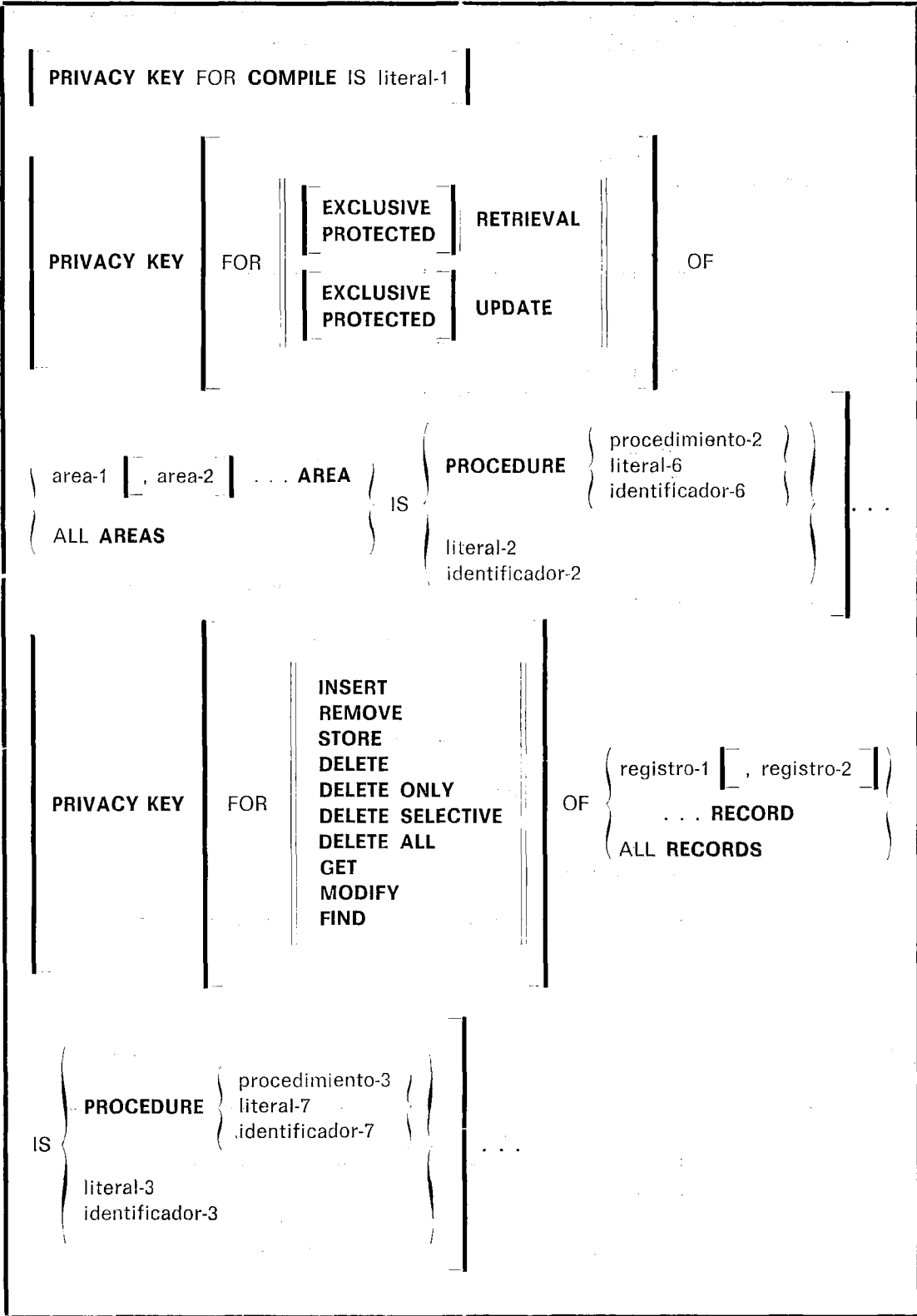


FIG. 55

forma análoga aplicarse de forma general a todos los Sets contenidos en el Esquema o selectivamente a aquéllos en particular que convenga.

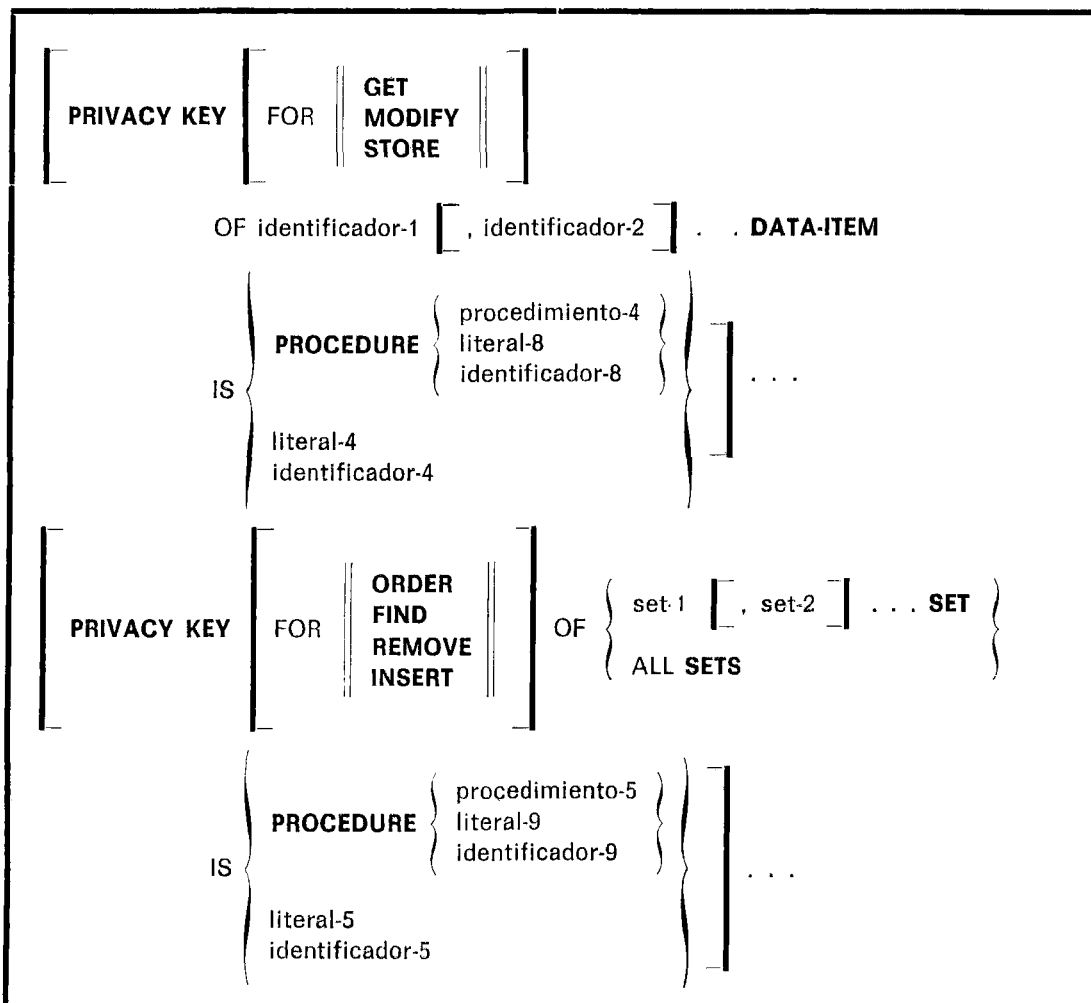


FIG. 56 (Continuación)

5.4.2. INSTRUCCIONES DML EN LA DATA DIVISION

La Data Division del programa redactado en COBOL incluye una nueva sección que recibe el nombre de **SCHEMA SECTION**, que hace referencia a un Sub-esquema previamente definido en lenguaje DDL y que produce que el Sistema efectúe las oportunas reservas de nombres y direcciones y zonas de área de trabajo, complementando al programa de aplicación en todas aquellas partes que están definidas en el Sub-esquema invocado y que de esta forma se incorporan al programa de proceso.

La Cláusula principal del Schema Section es la Cláusula **INVOKE**, que presenta el siguiente formato:

INVOKE SUB-SCHEMA nombre de Sub-esquema
OF SCHEMA nombre de esquema

por el cual se determina cuál es el Sub-esquema que va a regular el acceso del programa a la Base de Datos.

Mediante secciones COBOL standard adicionales se determinará la definición de todos aquellos datos que el programa vaya a manejar y que deberán estar debidamente especificados, en cuyos formatos no vamos a entrar, ya que no aportan ningún nuevo concepto en especial.

5.4.3. INSTRUCCIONES DML EN LA PROCEDURE DIVISION

La División de Procedimientos es una secuencia de instrucciones imperativas Cobol, que son de hecho las que componen el conjunto de criterios que constituyen el programa de proceso. Dentro de todo este flujo de instrucciones se insertan aquellas específicas para manejo de la Base de Datos, instrucciones que vamos a examinar una a una viendo su función y formato de utilización.

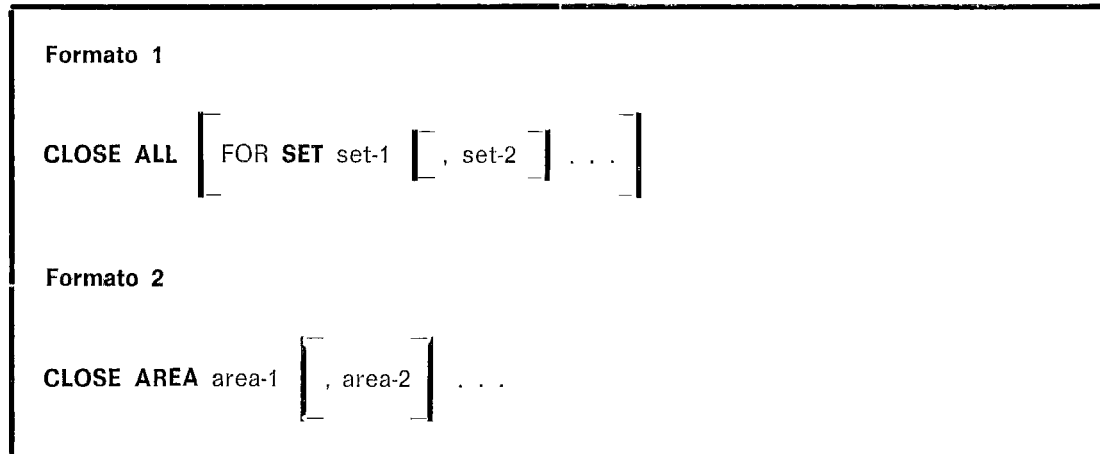


FIG. 57. Instrucción CLOSE

La instrucción CLOSE se presenta en la figura 57 y como podemos contemplar, tiene dos diferentes formatos. Si en el primer formato solamente se reseña la instrucción CLOSE ALL, el ejercicio de esta función produce un cierre de comunicación entre el programa y la Base de Datos, pues supone el cerrar todas las Areas al uso del programa. Cualquier intento de acceso a partir de la ejecución de dicha instrucción produce una condición de error.

Si se utiliza la opción de incluir determinados Sets cualificando la instrucción, solamente se produce el cierre de las Areas que contengan registros que participen en los Sets declarados bien como registros principales o bien como miembros, quedando el resto de las Areas abiertas al acceso de programas.

El segundo formato permite definir selectivamente cuáles son las Areas que se desean cerrar.

La función DELETE produce la cancelación de un Registro objeto de la Base de Datos a partir de cuyo momento no volverá a estar disponible para el acceso. Adicionalmente, el registro es separado de todos los Sets específicos en los que participa como miembro.

Asimismo se produce la cancelación de todos los Registros específicos que son miembros obligatorios de los Sets específicos, de los cuales el Registro objeto que es cancelado era el Registro principal. Cuando los miembros en vez de ser obligatorios estaban definidos en el Esquema como opcionales, simplemente se produ-

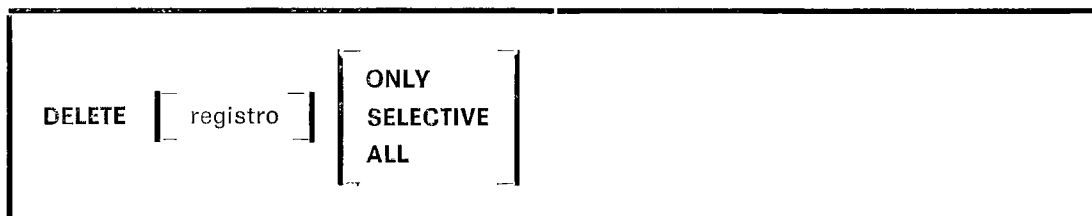


Fig. 58. Instrucción DELETE

ce una cancelación de la relación lógica, pudiendo el Registro continuar siendo accedido, aunque, naturalmente, ya no podrá recuperarse a través de su participación en ese Set específico que como entidad ha desaparecido. Estas funciones se realizan con la opción ONLY.

Si se utiliza la opción ALL, al cancelarse el Registro objeto se produce la cancelación incondicional de los miembros participantes en los Sets que dicho Registro objeto encabeza.

La opción SELECTIVE se diferencia de la opción ONLY en que los Registros miembros opcionales solamente son suprimidos si no participan como miembros en otros Sets. El formato de la instrucción puede verse en la figura 58.

Añadamos que el Registro objeto de la operación de cancelación ha sido previamente establecido por otros Comandos de selección de Registros que pueden utilizarse en el programa, como por ejemplo el Comando FIND.

El Comando FIND, cuyo formato se presenta en la figura 59, sirve para realizar la localización de la dirección de un Registro de la Base de Datos, lo cual se lleva a cabo por medio de la expresión "rse", en la que se incluyen los criterios de selección que vengan al caso.

Una vez realizada la selección de Registros, la dirección del Registro seleccionado se incluirá en unos contadores de estado que cada programa que utiliza la Base de Datos dispone al efecto y que denominaremos "Contadores de CURRENCY" para la debida identificación con las figuras que transcribimos del informe de DBTG.

Cada fichero tiene un Contador de Currency al igual que cada Area y cada Set. Así, cuando un Registro es seleccionado, se constituye en Registro CURRENT del fichero al que pertenece, del Area al que pertenece, y de cada uno de los Sets en los cuales participa, bien como Registro principal o como Registro miembro. Estos

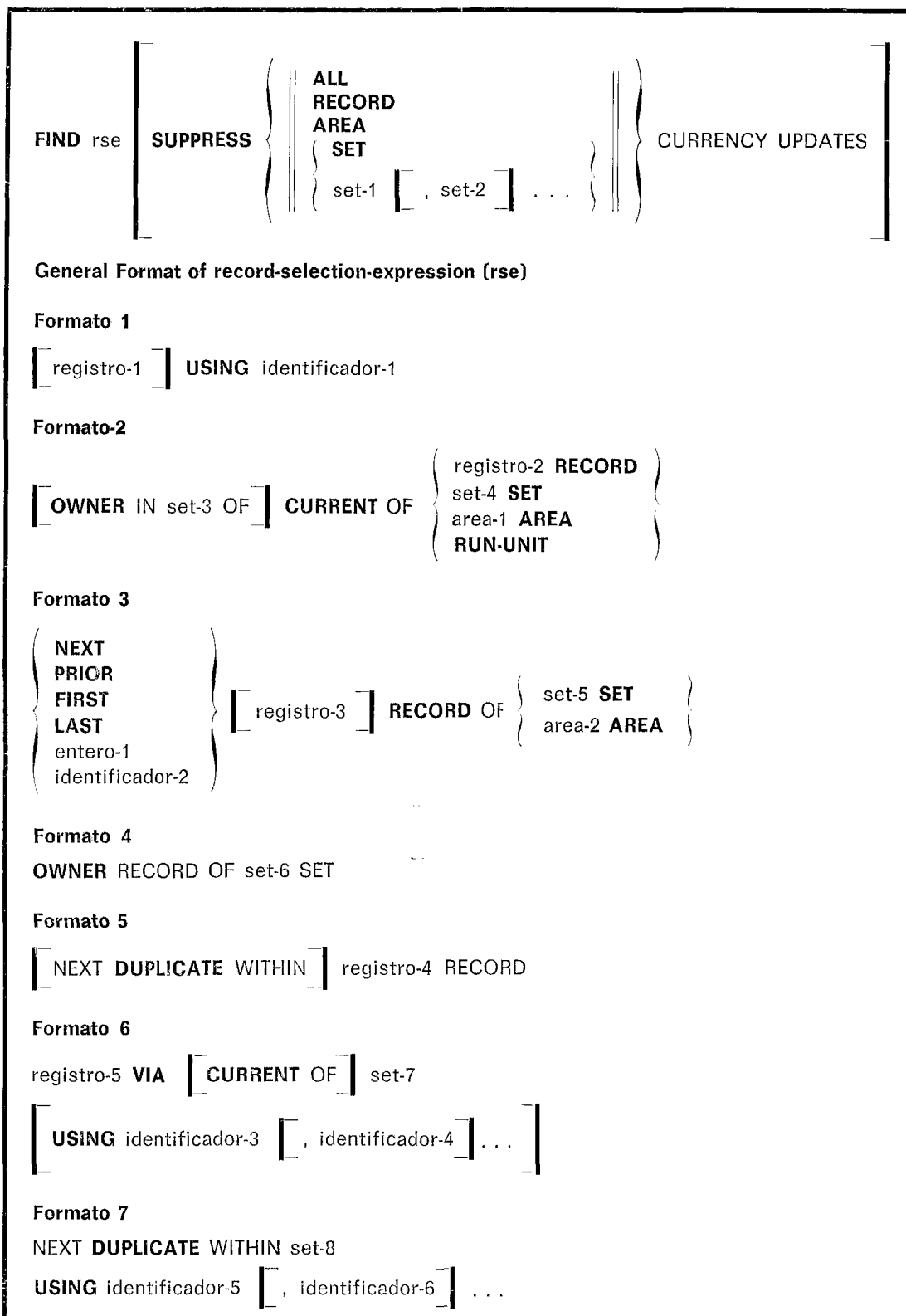


FIG. 59. Instrucción FIND

contadores de Currency son actualizados después de la ejecución de un comando FIND, a no ser que se hayan utilizado las opciones que la Cláusula SUPPRESS ofrece, según puede verse en la figura 59.

La instrucción FIND ofrece siete diferentes formatos para la expresión de selección de Registros que corresponden a las siguientes posibilidades:

- Acceso directo de un Registro, indicando el nombre del Fichero y la Clave de Base de Datos.
- Recuperación en base al Registro objeto definido por los contadores de CURRENCY.
- En el formato 3 se puede especificar la selección de Registros dentro de un Area o bien de un determinado Set. Se puede además detallar el nombre del Fichero al que el Registro ha de pertenecer. Ello define en base a los contadores de Currency un Registro objeto, y por medio de las opciones NEXT, PRIOR, FIRST, LAST, pueden definirse posiciones lógicas relativas en la secuencia de Registros del Area o del Anillo específico de que se trate, relativas, repetimos, al Registro objeto.
- El formato 4 prevé la recuperación del Registro principal de un Anillo, definido por el Registro CURRENT de "set-6".

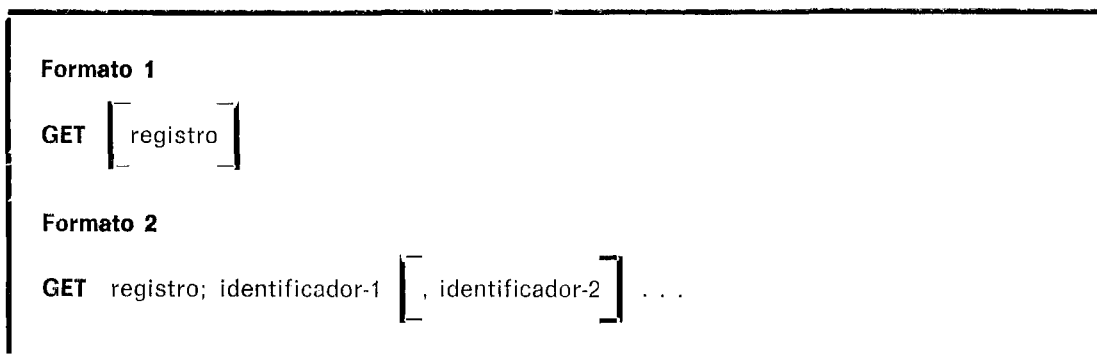


FIG. 60

- El formato 5 será utilizado para explorar la existencia de un Registro duplicado cuando se está procesando un fichero, entendiéndose por duplicados aquellos cuyas claves de clasificación son iguales.
- El formato 6 permite la localización de Registros explorando por claves de clasificación y determinando el Anillo en el cual se debe de explorar, en base al Registro objeto definido por el contador de Currency del Set "set-7".
- El formato 7 se utiliza análogamente al 5, para explorar Registros duplicados en el proceso de un Anillo.

La instrucción GET provoca la transferencia a la memoria principal de un Registro previamente localizado por una instrucción FIND, quedando depositado dicho Registro en el Area de trabajo del programa.

El Registro objeto para la ejecución de la instrucción GET es obviamente el CURRENT RECORD del fichero. Se presenta el formato de la instrucción en la figura 60.

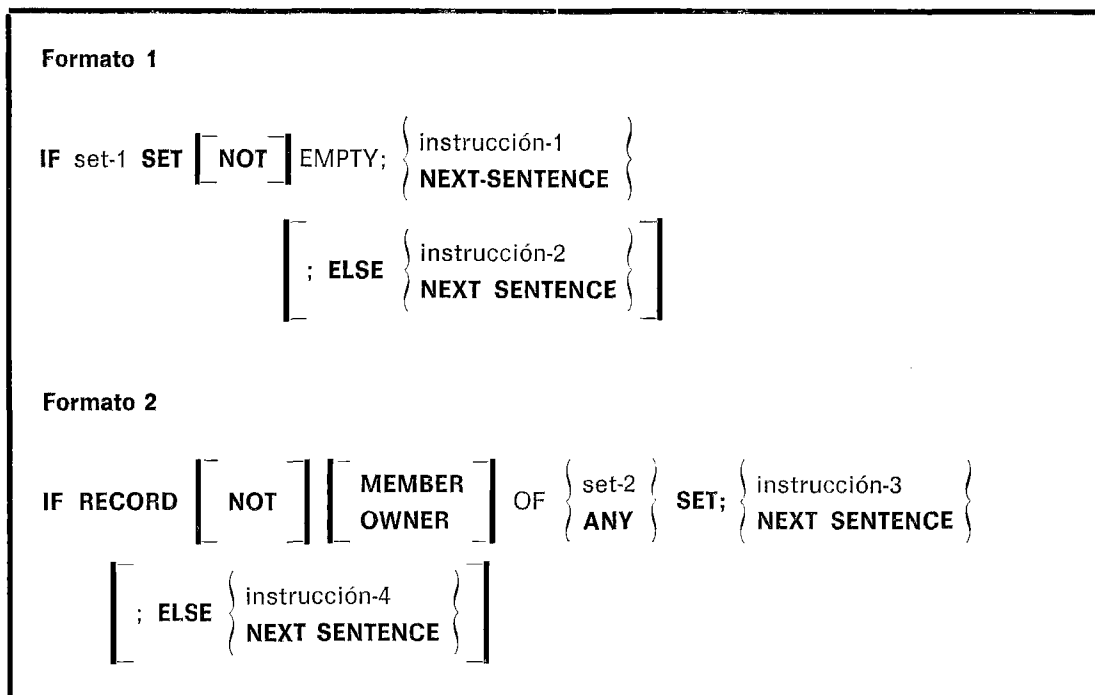


Fig. 61. IF

La instrucción IF cuyo formato se recoge en la figura 61 produce el establecimiento de una condición que ha de ser evaluada, dependiendo la continuación de la secuencia de instrucciones del resultado de la evaluación de la condición, es decir, de si ésta es verdadera o de si es falsa. La instrucción IF puede tener dos formatos.

En el primer formato se chequea el Set específico u ocurrencia que corresponde a la familia definida por "set-1", estando determinado el Set por el CURRENT RECORD de dicha entidad. El chequeo que se efectúa es, si el anillo indicado está vacío, es decir, no contiene miembros, o, por el contrario, tiene miembros dependientes. Según sea el resultado de la condición, se bifurcará a la instrucción indicada por el programador o bien se continuará con la instrucción siguiente en la secuencia.

En el formato 2 se hace una comprobación sobre el CURRENT RECORD general de la unidad de tratamiento, comprobándose si es o no miembro o registro principal de un determinado Set o de cualquier Set, bifurcándose a una instrucción predeterminada o a la instrucción siguiente, según sea el resultado de la comprobación.

```

Formato 1

INSERT [registro] INTO set-1 [, set-2] ...

Formato 2

INSERT [registro] INTO ALL SETS

```

FIG. 62. INSERT

La instrucción INSERT, cuyo formato se recoge en la figura 62, puede ser utilizada para insertar un Registro o, mejor dicho, el CURRENT RECORD de un Fichero, en uno o varios Sets especificados. Naturalmente, para que pueda producirse la operación de inserción, el registro que estamos considerando ha de pertenecer a un fichero que haya sido declarado en el Sub-esquema invocado por el programa, como posible miembro de los Sets especificados.

El Comando MODIFY, cuyos dos formatos se indican en la figura 63, puede ser utilizado para actualizar un registro, o campos de un registro, por los correspondientes valores que en la secuencia anterior de programación han sido depositados en una zona ad-hoc del área de trabajo del programa.

Este mecanismo también puede ser utilizado para alterar el conexionado lógico de miembros de Sets, pues si la inserción lógica de un registro miembro se puede realizar en base al significado de claves de clasificación contenidas como claves dentro del propio registro, al actualizar estas claves se provoca una necesaria ordenación de los Sets específicos afectados.

```

Formato 1

MODIFY [registro] [USING identificador-1 [, identificador-2] ...]

Formato 2

MODIFY registro; identificador-3 [, identificador-4] ...

[USING identificador-5 [, identificador-6] ...]

```

FIG. 63

El registro objeto de la instrucción **MODIFY** es el **CURRENT RECORD** del fichero que se especifica. Los valores que ofrece la Cláusula opcional **USING** sirven para calcular la nueva posición del registro como miembro de **Set**.

El formato 2 de la instrucción **MODIFY** permite definir mediante identificadores, los campos del registro que puedan ser objeto de modificación, estando obligados los restantes campos a permanecer inalterados.

La instrucción **MOVE CURRENCY STATUS** se puede utilizar para recoger el contenido de un contador de **CURRENCY** en un identificador o campo simbólico manejado por el programa, instrucción que puede tener dos formatos que se recogen en la figura 64.

La figura 65 presenta la instrucción **OPEN**, cuya función es la de abrir la comunicación entre una o varias Areas y un Programa de proceso, permitiéndose a partir de ese momento el acceso del referido Programa a los registros de las Areas indicadas en la instrucción.

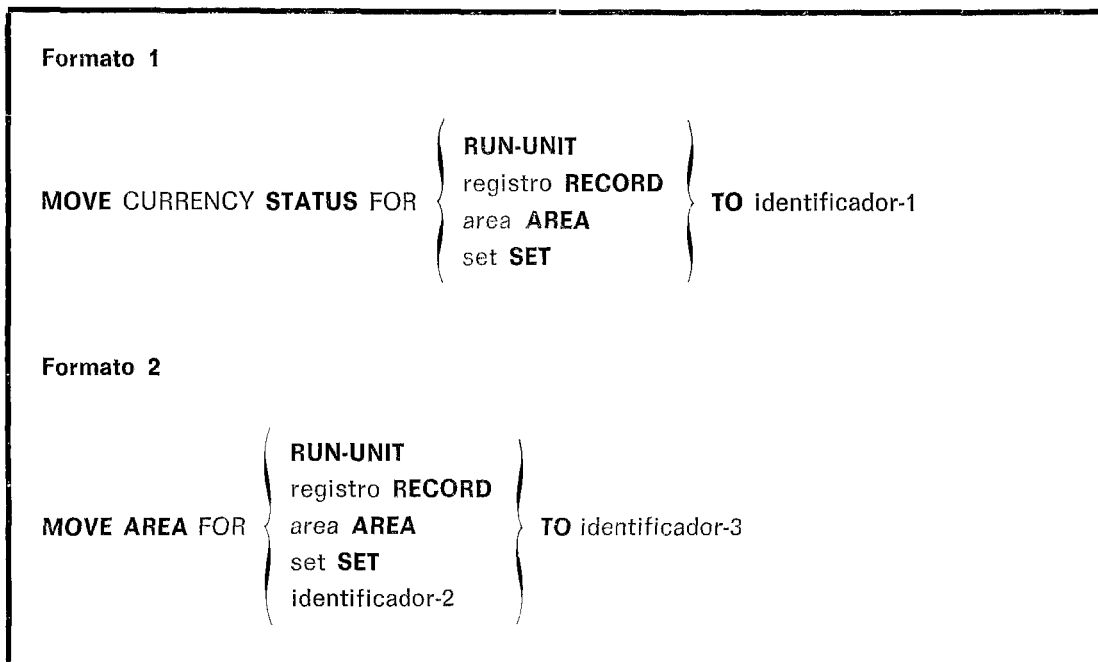


FIG. 64

Si se utiliza el formato **OPEN ALL** se puede limitar además que los registros que vayan a ser objeto de acceso pertenezcan a unas determinadas familias de **Sets** que pueden definirse en la propia instrucción, o también, opcionalmente, puede explicitarse el tipo concreto de utilización que el programa vaya a realizar a partir de ese momento.

Con el formato número 2 se puede explicitar las Areas concretas que van a ser accedidas, estableciéndose igualmente las condiciones de utilización que van a prevalecer.

En la figura 66 se recoge un cuadro tomado igualmente del informe "April 71 de DBTG", que establece las relaciones entre dos programas concurrentes en cuanto a las incompatibilidades que puedan presentarse para acceso de un mismo Area.

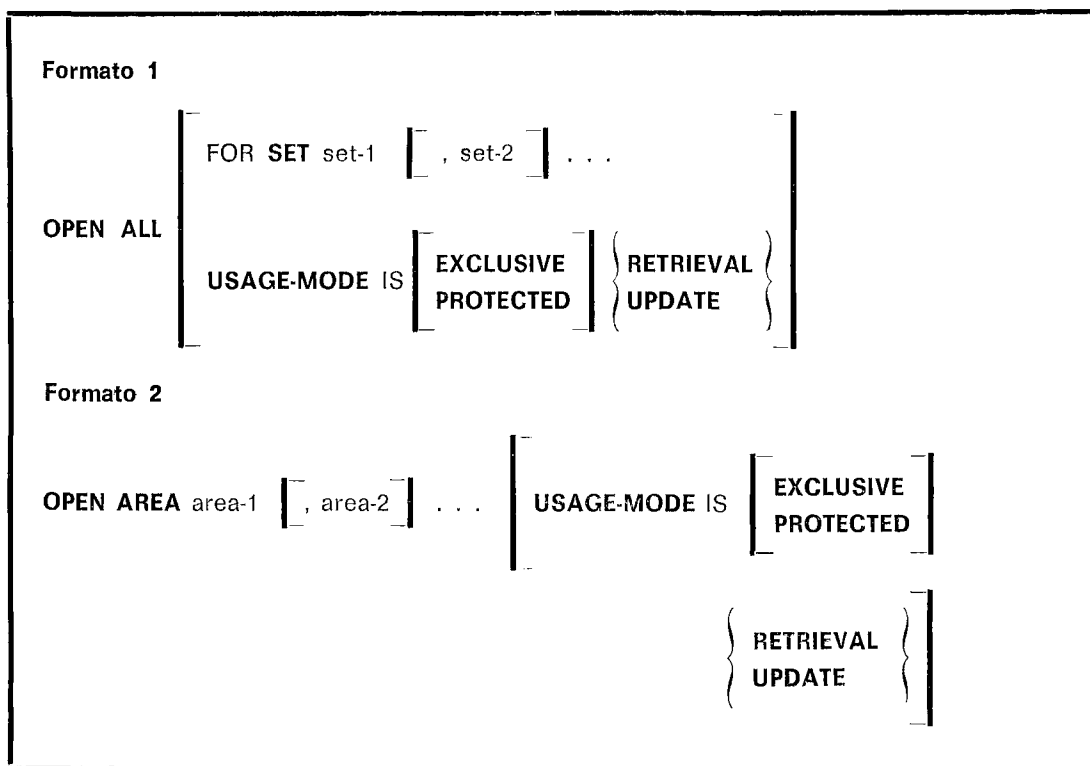


FIG. 65

En la columna del margen izquierdo se establecen los tipos de utilización para los que el primer Programa abrió el Area en cuestión, indicándose en la fila de cabecera asimismo la utilización que del Area pretende realizar el segundo Programa concurrente.

Así, por ejemplo, si el primer Programa abrió el Area para el uso de PROTECTED RETRIEVAL, es decir, para recuperación protegida de Registros, se expresa mediante una N en la tabla de doble Entrada que el segundo Programa no podrá realizar ninguna forma de actualización en el Area porque, obviamente, la necesaria protección para las recuperaciones del primer Programa lo impide. Así,

USAGE-MODE	OPEN-NO USAGE-MODE	RETRIEVAL	UPDATE	PROTECTED RETRIEVAL	PROTECTED UPDATE	EXCLUSIVE RETRIEVAL	EXCLUSIVE UPDATE
OPEN-NO USAGE-MODE	S	S	S	S	S	N	N
RETRIEVAL	S	S	S	S	S	N	N
UPDATE	S	S	S	N	N	N	N
PROTECTED RETRIEVAL	S	S	N	S	N	N	N
PROTECTED UPDATE	S	S	N	N	N	N	N
EXCLUSIVE RETRIEVAL	N	N	N	N	N	N	N
EXCLUSIVE UPDATE	N	N	N	N	N	N	N

FIG. 66

de esta forma, tampoco podrá el segundo Programa realizar recuperaciones en exclusiva, porque existe ya un Programa anterior que está realizando tales recuperaciones, lo que obviamente implica que dicha exclusividad no puede ser concedida. Sin embargo, sí puede abrirse el Area, para acceso del segundo Programa, para recuperación simple o recuperación protegida, o simplemente puede abrirse el Area sin especificar el uso, lo que solamente dará lugar a poderse realizar operaciones compatibles con las condiciones del primer Programa, estableciéndose las correspondientes condiciones de error en caso de tratar de vulnerar las condiciones.

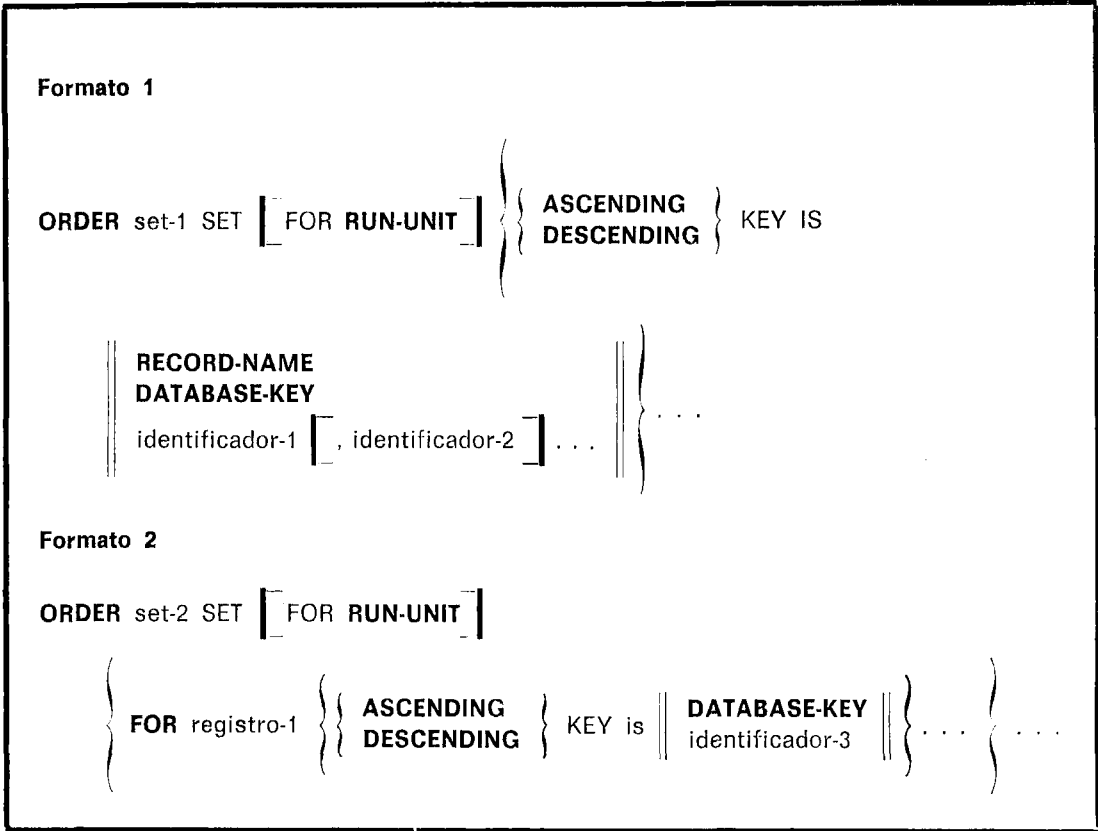


FIG. 67

La Instrucción ORDER recogida en la figura 67 provoca que el Set específico definido por el CURRENT RECORD del Set "set-1" sea reorganizado en lo que a la secuencia lógica del registro se refiere, y ello en función de las claves que me-

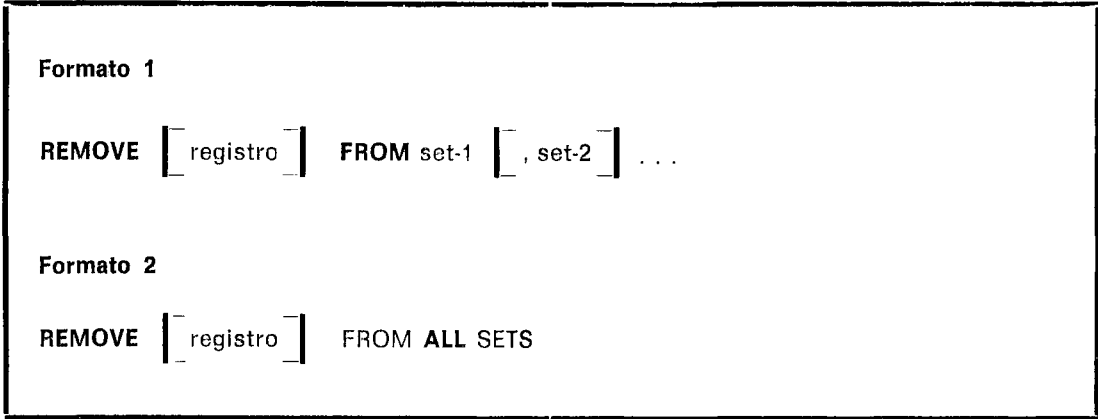


FIG. 68

diante la propia Instrucción determinamos, pudiendo ello ser realizado bajo los dos formatos que en la figura pueden contemplarse de manera prácticamente autoexplicativa.

La Instrucción REMOVE provoca la cancelación de la relación lógica de un Registro como miembro de uno o varios Sets, para los cuales ha de estar definido como miembro opcional.

Con el primer formato se pueden seleccionar los Sets en particular, de los cuales se quiera separar el Registro, mientras que el formato 2 prevé la separación de todos los Sets en los que el Registro participa como miembro opcional. El Registro objeto de esta Instrucción es el CURRENT RECORD del fichero que se determina en la propia Instrucción, cuyo formato está recogido en la figura 68.

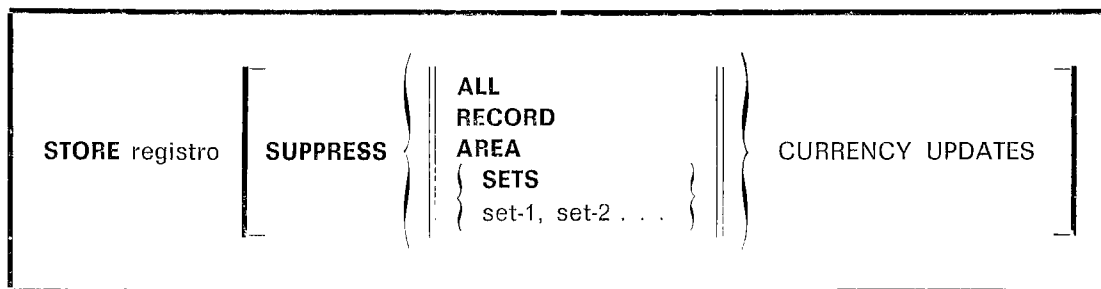


FIG. 69

La figura 69 ofrece el formato de la Instrucción STORE, cuya función determina el almacenamiento de un Registro en la Base de Datos. El Registro objeto que ha de ser almacenado, debe estar depositado en el Area de trabajo del programa, en la posición destinada al fichero en cuestión, fichero que está definido en la propia Cláusula STORE. La posición de almacenamiento será calculada en base a la forma de localización que para dicho fichero se establece en las definiciones de Sub-esquema que el programa invoca. Para ello, naturalmente, el programa habrá inicializado adecuadamente los parámetros oportunos previamente a la ejecución de la Instrucción.

La ejecución de la Instrucción STORE provoca también que el registro, una vez almacenado, sea insertado automáticamente en todos los Sets para los que se encuentra el fichero declarado en el Sub-esquema, como miembro Automático. Análogamente, la ejecución de la Instrucción STORE provoca la creación de un Set específico si el fichero al cual el registro pertenece está declarado como fichero principal de algún Set.

La Instrucción STORE produce una actualización de los contadores de CURRENCY, a menos que opcionalmente se haya declarado una especificación en contra.

6. BIBLIOGRAFIA

- IAN R. PALMER: *Database Systems: A Practical Reference*. C.A.C.I. Inc., London 1975.
- JOHN K. LYON: *An Introduction to Data Base Design*. John Wiley & Sons, Inc., 1971.
- Data Base Management Systems. Proceedings of the Share Working Conference on DBMS*. Donald A. Jardine, Canada 1973
- CHARLES T. MEADOW: *The analysis of Information Systems: An Introduction to Information Retrieval*. John Wiley & Sons., Inc., 1967/1973.
- Y. HARRAND: *Traitement des files et des listes*. Dunod 1967.
- C. JOUFFROY y C. LETANG: *Les fichiers. Pratique et choix de l'organisation des données informatiques*. Dunod 1974.
- M. THOMAS: *Les banques de données dans l'administration publique*. OCDE, Paris 1971.
- J. D. WARNIER: *L'Organisation des Données d'un Système*. Les Editions d'Organisation, Paris 1974.
- Data Base Systems. State of the Art Report*. Infotech, London 1975.
- DATA BASE TASK GROUP: *April 71 Report*, Codasyl 1971.
- SYSTEMS COMMITTEE: *Feature Analysis of Generalized Data Base Management Systems*. Codasyl 1971.
- PROCEEDINGS OF 1971 ACM-SIGFIG ET WORKSHOP: *Data Description Access and Control*, ACM 1971.
- PROCEEDING OF 1972 ACM-SIGFIDET WORKSHOP: *Data Description Access and Control*, ACM 1972.
- SHARE SSD, 246 JUNE 10, 74: *Establishment of a Data Base Administration Function*. Share Inc., New York 1974.
- JOINT SHARE/GUIDE D. B. REQUIREMENTS GROUP: *Data Base Management System Requirements*, Share Inc., Noviembre 1970
- GUIDE, INFORMATION MANAGEMENT GROUP: *Basic Requirements for a Data Base Management System*. Guide, febrero 1973.
- GUIDE, INFORMATION MANAGEMENT GROUP: *The Data Base Administrator*. Guide, abril 1973.
- Informatique et Gestion*, núm. 28, mayo 1971.
- Comparison of DBMS Reports, DBTG y GUIDE-SHARE*. IBM, USA 1971.
- Requirements for the Data Dictionary/Directory within the Guide/Share DBMS concept*. Guide, enero 1974.
- GUIDE, INFORMATION MANAGEMENT GROUP: *Requirements for a User Language*. Guide, febrero 1972.
- SHARE, DATA BASE PROJECT: *Technical Requirements Committee Report*. Share, SSD247, mayo 1974.
- SHARE, DATA BASE PROJET: *Data Base-User's View*. Share, Inc.
- One Man's Data Base*. Share, SSD 234, marzo 1973, Share, Inc.
- Survey of Generalized Data Base Management Systems*. CODASYL, mayo 1969.
- JAMES MARTIN: *Data Base Organization*. McGraw & Hill, 1975.
- E. F. CODD: *A Relational Model of Data for large Shared Data Banks*. Communications of ACM, junio 1970.
- Data Management Software Reports*. Auerbach Publishers Inc.
- Creating the Corporate Data Base*. EDP Analyzer, febrero 1970.
- The Debate on Data Base Management*. EDP Analyzer, marzo 1972.
- Trends in Data Management, Part I*. EDP Analyzer, mayo 1971
- Trends in Data Management, Part 2*. EDP Analyzer, junio 1971.

- The Cautious Path to a Data Base*. EDP Analyzer, junio 1973.
- DBMS: A Critical and Comparative Analysis*. Information Sciences Inc., New York 1973.
- Data Management Software Reports*. Datapro Research Corporation.
- D. E. CUOZZO/J. F. KURTZ: *Building a Base for Data Base: A Management Perspective*. Datamation, octubre 1973.
- Banques de Données: Des responsabilités nouvelles*. Informatique et Gestion, noviembre 1973.
- CLAUDER SAUBRINER: *Les Systems de gestion de bases de données*. 0.1. Informatique, enero/febrero y marzo 1974.
- PARSONS, DALE y YURKANAN: *Data manipulation language requirements for DBMS*. Computer Journal, mayo 1974.
- CHARLES W. BACKNAM: *Bases de données, un nouvel art de la navigation*. 01 Informatique, junio 1974.
- K. RESIDE y T. SEITER: *The evolution of an integrated Data Base*. Datamation, septiembre 1974.
- Data Base*, a quarterly newsletter of SIGBBP. Publicación trimestral de ACM. Association for Computer Machinery.
- The Data Dictionary/Directory function*. EDP Analyzer, noviembre 1974.
- E. S. PAGE & L. B. WILSON: *Information Representation and Manipulation in a Computer*. Cambridge University Press, 1973.
- D. E. KNUTH: *The Art of Computer Programming*. Addison-Wesley, 1968.



SERVICIO DE PUBLICACIONES DEL MINISTERIO DE EDUCACION Y CIENCIA