

2018

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

UN METODO COMPUTACIONAL COMPLEMENTARIO A LOS
MANEJADORES DE BASES DE DATOS: SU APLICACION
EN RECUPERACION BIBLIOGRAFICA

T E S I S

PARA OBTENER EL TITULO DE:

M A T E M A T I C O

P R E S E N T A :

LAURA FREIDBERG GOJMAN

DIRECTOR DE TESIS:

MAT. MA. ISABEL GARCIA HIDALGO

MEXICO D.F. SEPTIEMBRE

1989

TESIS CON
FOLIO DE ORIGEN



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

I N D I C E

Introducción	1
I. Estructuración de la información para recuperación bibliográfica	4
II. Técnicas de Indexamiento	12
III. Un esquema estructural para la información bibliográfica	19
IV. El sistema de recuperación bibliográfica	31
Conclusiones	52
Anexos	
A. Programas y algunas de las tablas que utilizan	55
B. Ejemplos de productos	90
C. Adaptación del formato MARC para El Colegio de México	96
Bibliografía	132

INTRODUCCION

En esta tesis se presenta un sistema computacional para manejar información bibliográfica. Esta información se caracteriza por su alto grado de variabilidad en lo que se refiere a categorías y tamaños.

Las características fundamentales del sistema que se presenta son:

1) Trabaja sobre información codificada en un formato que no impone limitaciones de tamaño o número de ocurrencias de los campos de una unidad de información. Para manejar adecuadamente este tipo de información utiliza algoritmos rápidos de reconocimiento, identificación y selección de los campos de una unidad de información, así como algoritmos de transformación de los contenidos de campos seleccionados.

2) Da la posibilidad de que la transformación de contenidos para campos seleccionados tenga como producto un subconjunto, en formato fijo, del total de la información.

El sistema permite que los usuarios definan, de acuerdo a sus necesidades o etapa de sus investigaciones, qué campos de información habrán de incluir en cada subconjunto -en formato fijo- que van a crear.

Los subconjuntos se obtienen de la información completa por medio del sistema computacional desarrollado para esta tesis, y pueden ser posteriormente trabajados con un manejador de bases de datos.

Usando las herramientas de selección y agrupamiento que proporcionan los manejadores de bases de datos se pueden construir nuevos subconjuntos tomando en cuenta ya no los campos sino los contenidos de los campos seleccionados en la etapa de definición de la base de datos.

El sistema permite que se pueda recuperar la información completa de aquellas unidades seleccionadas en el manejador de bases de datos. Estas unidades seleccionadas pueden ser procesadas nuevamente por el sistema y así obtener, para algunas aplicaciones, nuevos productos que no se podían obtener con el manejador.

En el caso que nos ocupa, el material bibliográfico con el que se trabaja corresponde a obras sobre Asia y Africa catalogadas y clasificadas, utilizando el formato MARC, por la Biblioteca de El Colegio de México que es una biblioteca especializada en Ciencias Sociales y Humanidades. Además de los usuarios normales de una biblioteca, el material bibliográfico sobre Asia y Africa es utilizado por investigadores sumamente especializados, para quienes la recuperación que puede hacerse bajo los criterios de catalogación y clasificación de la propia biblioteca resulta poco precisa. Con el sistema propuesto se da a estos investigadores la posibilidad de utilizar criterios más finos sin invalidar los criterios más generales que dan buenos resultados de recuperación a usuarios menos especializados. Esto se puede hacer agregando descriptores especializados al conjunto de información original, así como resúmenes escritos en lengua

natural en los cuales estén marcadas algunas palabras claves.

El proceso de agregar información especializada no afecta a los usuarios no especializados ya que sus productos pueden ser obtenidos sin ninguna modificación. Paralelamente, los usuarios especializados pueden incluir directamente en sus bases de datos estos nuevos descriptores o las palabras claves marcadas en los resúmenes, y conservar o descartar los descriptores más generales.

I. Estructuración de la Información para recuperación bibliográfica

El término recuperación de información se usa, en la literatura sobre el tema, en dos sentidos distintos. A continuación se discute la diferencia entre la acepción genérica del término "recuperación de información" y la acepción específica que tiene cuando se considera en relación al término "recuperación bibliográfica".

Un científico realizando alguna investigación está envuelto, casi sin excepción, en dos tipos de búsquedas que, aunque están conectadas en su quehacer, conviene mantener separadas en el análisis si se quiere evitar confusión.

Puede querer obtener respuestas muy detalladas a ciertas preguntas específicas que surgen de su investigación ó necesitar las referencias de toda la literatura que contiene material relevante a su problema. Llamaremos a lo primero "recuperación de información" y a lo segundo "recuperación bibliográfica".

La respuesta al primer tipo de búsqueda sería una o más declaraciones afirmativas, mientras que para el segundo sería una lista de referencias bibliográficas, con o sin notas, resúmenes o copias de la literatura relevante.

Para que el conocimiento sobre un tema pueda ser utilizado en un sistema de recuperación de información, debe ser estructurado como un conjunto de proposiciones que conserven la información sobre hechos relativos al tema, o la que es

considerada de interés y valor para el sistema. Debe hacerse hincapié en que no cualquier documento puede ser usado en un recuperador de información debido a que generalmente contienen proposiciones metalingüísticas sobre los resultados obtenidos en la literatura original, en vez de un sumario de los resultados en sí mismos.

En cambio, para que una referencia bibliográfica pueda ser seleccionada mediante un sistema de recuperación bibliográfica es necesario que se use un conjunto de símbolos cuya función es dar una pista al usuario de qué tan útil puede serle el documento. Estos símbolos pueden ser:

- a) asociados a la referencia bibliográfica. En este caso se caracterizan por ser palabras sueltas, frases, encabezamientos de materia, descriptores, etc., que se pueden tomar directamente del documento original, o que pertenecen a algún lenguaje controlado (thesaurus, etc.).
- b) obtenidos a través de un análisis ya sea de la información contenida en las referencias bibliográficas, o de la contenida en las notas o resúmenes asociados a dichas referencias. Estos símbolos se pueden obtener, por ejemplo, calculando la frecuencia de las palabras que aparecen (sin tomar en cuenta las palabras gramaticales como artículos, preposiciones etc.) y marcando o eligiendo las más frecuentes.

Usaremos el término "descriptor" como nombre genérico de los

símbolos asociados o los elementos obtenidos mediante el análisis.

Estos descriptores señalan de alguna manera, el tipo de información contenida en el documento, pero no incluyen como podría pensarse, la información en sí misma. Es un error suponer que una palabra o una frase contiene información en el mismo sentido que la tiene una proposición; la información contenida en una proposición no es la suma o combinación de la información contenida en las frases que la constituyen.

Podemos dar una explicación a la confusión que ha existido entre recuperación de información y recuperación bibliográfica con base en los siguientes hechos: para llegar a responder una consulta a un sistema de recuperación de información a menudo se necesita hacer primero una recuperación bibliográfica con el propósito de seleccionar los documentos que contienen información relevante, y luego obtener la respuesta a la pregunta solicitada buscando en el lugar preciso de los documentos seleccionados. Además, algunas veces se hace tan solo una recuperación bibliográfica porque no se tiene un buen sistema de recuperación de información directo. Bajo estas circunstancias la recuperación bibliográfica es un sustituto muy pobre de un sistema de recuperación de información.

Sin embargo, ningún sistema de recuperación de información (si existe) podrá reemplazar totalmente la lectura de la literatura original que contiene la información, ya que el valor de ésta consiste principalmente en la estimulación que da a

través de su línea general de argumentación y método, más que su riqueza en hechos específicos. Por esto los sistemas de recuperación bibliográfica no deben verse solamente como una primera fase de los sistemas de recuperación de información sino que tienen una importancia propia.

Una vez aclarada la diferencia entre estas dos formas de recuperación (la bibliográfica y la de información), debemos dejar claro también que los métodos para resolver cada uno de los problemas no son necesariamente los mismos, aunque los métodos desarrollados para alguno de estos sistemas puedan aplicarse, con o sin adaptaciones, al otro; y las consideraciones teóricas sobre uno puedan ser relevantes al otro.

Esta tesis se centra en el problema de recuperación bibliográfica.

Discutamos a continuación dentro del marco de la recuperación bibliográfica, la necesidad de estructurar la información -material textual-, y algunos de los problemas que se presentan.

Supongamos por ejemplo que tenemos una obra titulada "Escritos sobre la Historia del México Antiguo", cuyo autor corporativo es la Coordinación de Historia de la Facultad de Filosofía y Letras de la Universidad Nacional Autónoma de México; la obra ha sido editada por Editorial México, en México, en 1980. Supongamos también que elegimos un encabezamiento de materia o un descriptor temático como "México - Historia - 1492-1521".

Es claro que cada ocurrencia de la palabra "México" tiene un

valor diferente: en el título mencionado la palabra México refiere al tema sobre el cual se escribe, en la segunda ocurrencia es parte del nombre que precisa cuál Institución es autora de la obra, en la tercera es parte del nombre de la editorial que publicó el libro, en la cuarta hace una referencia geográfica al lugar donde se publicó la obra y en la última es el tema que ha sido aprobado por el catalogador siguiendo los estándares de una Institución dada. Difícilmente nos encontraremos con un problema de recuperación para el cual no sea necesario discriminar los diferentes valores de estas ocurrencias.

El valor que tiene cada ocurrencia de un dato específico, por ejemplo la palabra "México", es a su vez información que debe captarse para realizar una adecuada recuperación en la computadora.

Podríamos decir que la información Título cualifica a la información "Escritos sobre la Historia del México Antiguo", la información Autor cualifica a "Universidad Nacional Autónoma de México. Facultad de Filosofía y Letras. Coordinación de Historia", la información Editorial cualifica a "Editorial México", "México" tiene la cualidad de ser el Lugar de Publicación y "1980" la de ser el Año de Publicación.

Cuando hemos podido establecer reglas precisas y consistentes para discriminar los campos de autor, de título, etc., en una obra, decimos que esa información está "bien estructurada".

Una referencia bibliográfica es entonces el conjunto de datos discriminados que identifican a la obra. Llamaremos unidad de información de un sistema de recuperación bibliográfica a las referencias bibliográficas junto con la información asociada a la que hemos llamado genéricamente descriptores.

Muchas veces no es posible discriminar o predefinir partes específicas en la información. En esta situación diremos que la información está "mal estructurada". Debe quedar claro que lo bien o mal estructurado de la información depende del propósito de la recuperación. Así, un resumen estará bien estructurado si se ha discriminado del resto de la información, y se desea usar como una componente de la referencia bibliográfica. Sin embargo, un resumen estará mal estructurado, en el momento en que nuestro propósito sea analizarlo para discriminar la información que contiene.

No debemos olvidar que el valor de estos resúmenes es que proporcionan una gran cantidad de información al lector, ya que están escritos en lengua natural y conservan la línea general de argumentación o método además de la descripción de los hechos específicos fundamentales; información que no puede ser proporcionada al lector mediante un análisis por computadora debido a que carecen de una estructura adecuada para ello.

La aplicación de los criterios de discriminación -usualmente llamada análisis e indexamiento- nos permite imponer una estructura a la información que será la que se aproveche durante la recuperación. La exactitud del proceso de recuperación

dependerá directamente del nivel de detalle que se haya alcanzado en la discriminación de las categorías de Información.

Desde hace mucho tiempo los científicos de la computación se han esforzado por proponer métodos automáticos para efectuar análisis del material textual con los que se pueda dar estructura a información mal estructurada. Entre los métodos propuestos podemos enlistar: el uso de las frecuencias de los términos en un texto, la aplicación del análisis contextual y el sintáctico, los métodos semánticos, ... y la lista es larga y mal estructurada.

Con estas técnicas se han obtenido listas, producidas por computadora, de términos asociados al material textual. El hecho de que estas listas de descriptores sean tan diferentes entre sí, en función de las técnicas aplicadas, el orden de aplicación, el material analizado y el programa utilizado, muestra que aún no es eficiente la estructuración automática de la Información. No se puede cambiar la Información mal estructurada a bien estructurada con técnicas computacionales por sí solas aún teniendo un propósito definido; seguirá siendo una tarea humana en la que la computadora deberá usarse como una herramienta de ayuda.

NI el hombre ni las máquinas por sí solos pueden realizar con éxito las tareas de análisis e Indexamiento necesarias para una recuperación de Información eficiente. Se requiere una gran cantidad de Juicio humano y de capacidad intelectual, que no han podido ser incorporados a procesos computacionales. Pero también se requiere el uso de las computadoras por la rapidez de sus operaciones y las facilidades que ofrecen para manejar grandes

volúmenes de información, lo que esta tesis analiza las posibilidades prácticas del hombre.

Para que una solución sea eficiente habrá que encontrar un justo medio en la apropiación de las tareas que pueden ser hechas y las que se ejecutan mediante las computadoras. La que el "justo medio" depende de cada problema específico. Una solución formada de herramientas que puedan ser adaptadas de manera simple para los diversos problemas será eficiente. Uno de los propósitos de esta tesis es presentar herramientas adaptables.

En la mayoría de las áreas en las que se intenta el uso de sistemas de información se tiene una mezcla intrínseca de información bien estructurada y mal estructurada. Esta mezcla es el elemento esencial para que la información sea útil, ya que, si se separa, perdería gran parte de su significado para los usuarios.

El uso que el hombre haga de los sistemas computacionales para apoyo a sus tareas de análisis e indexamiento de la información, debe permitir el manejo conjunto de los dos tipos de información, la mal estructurada y la bien estructurada. En algunos casos el hombre se podrá servir de la información bien estructurada para organizar la información mal estructurada posibilitando un avance en la estructuración.

II.- Técnicas de Indexamiento

En la sección anterior hemos analizado la importancia y el significado de estructurar la información para un sistema de recuperación bibliográfica y hemos concluido que debe permitirse el manejo conjunto de la información bien estructurada y la mal estructurada.

Ya que el almacenamiento de información está íntimamente relacionado con los propósitos de la recuperación queremos distinguir entre dos modalidades de búsqueda: aquéllas en las que el usuario ha hecho la selección de la obra antes de utilizar el sistema de recuperación bibliográfica, y aquéllas en las que el usuario hará la selección mediante el uso del sistema.

En el primer caso se conocen los elementos necesarios para identificar unívocamente a la obra seleccionada, -generalmente autor y título- y se desea recuperar la unidad de información completa que dará al usuario otros datos como la signatura topográfica o el resumen de la obra, que le permitirán respectivamente localizar la obra o juzgar si su selección previa fue adecuada.

En el segundo paso es necesario manejar el conjunto completo de información para de él seleccionar las obras que se refieren a un tema determinado. Para efectuar esta selección se usan -ya sea aislados o combinados mediante relaciones lógicas- los descriptores que, como habíamos mencionado, tienen la función de dar una pista al usuario de qué tan útil puede serle el

documento.

Para asignar dichos descriptores al material textual existen principalmente dos formas. Una de estas formas, que se ha utilizado tanto en métodos manuales como automáticos, es elegir como descriptores los términos tal como se encuentran en el material textual que se está describiendo. A esta forma se le llama "Indexamiento libre". La otra forma es construir los descriptores y organizarlos para después asociarlos al material textual. Esta forma se conoce como "Indexamiento de tesoro". Cada una tiene algunas ventajas así como ciertas desventajas. A continuación discutiremos los dos métodos con el propósito de hacer ver, más adelante, que el sistema desarrollado en esta tesis ofrece herramientas para aprovechar complementariamente ambos métodos y por lo tanto cubrir las necesidades de usuarios de diversos niveles.

En el Indexamiento libre, el Indexador, teóricamente, puede elegir cualesquiera expresiones, ya sean palabras sueltas o frases, que aparezcan en el material textual que está describiendo -expresiones textuales-, o bien expresiones lógicamente derivadas de dichas expresiones textuales, y en general cualquier término que considere adecuado para representar el tema o materia del que trata el documento.

Podría pensarse que en estas condiciones se tiene una libertad absoluta de elegir los descriptores, sin embargo, debe tenerse presente que existen limitantes como son la tradición, el sentido común, restricciones del lenguaje ordinario,

entrenamiento psicológico e Institucional, y acondicionamiento cultural, las cuales restringen en diversa medida la libertad de Indexamiento.

En la práctica el Indexamiento libre es aquél en el que los términos o expresiones que pueden ser escogidos como descriptores deben ser seleccionados únicamente entre los términos o expresiones que aparecen en el material textual. Es decir, los Indexadores examinan el material textual y subrayan o indican aquéllos términos que consideran que son importantes o indicativos de la materia o asunto al que se refiere el documento.

Los resultados de un Indexamiento libre pueden llegar a presentar severas inconsistencias debidas a las distintas apreciaciones de un Indexador con respecto a otros así como las de un Indexador en distintos momentos o circunstancias. Estas inconsistencias se minimizan cuando el Indexador y el buscador son especialistas en una área del conocimiento, ya que ambos perciben de manera similar la organización de los conceptos de su área de especialidad y además utilizan de manera natural un vocabulario estandarizado tomado de la terminología de su materia.

Por tanto, el Indexamiento libre es ventajoso siempre y cuando los Indexadores y quienes consultan sean especialistas en la materia a la que se refieren los documentos.

En el Indexamiento de tesaurus, los descriptores se eligen de listas que contienen términos previamente seleccionados asociados

a cierta información sobre su uso; estas listas se organizan o estructuran mediante relaciones entre los términos. La mayoría de los tesauros son de propósitos especializados en el sentido de que sus términos se toman del vocabulario de una disciplina particular. Junto con cada término listado se pueden incluir algunas referencias a otros términos que tienen el mismo significado, significados más específicos o más genéricos, o términos cuyos significados están relacionados.

El propósito de los tesauros es organizar los conceptos de una área específica del conocimiento en campos semánticos o conceptuales. La organización se establece mediante dos tipos de relaciones: las relaciones jerárquicas y las relaciones de afinidad. La organización jerárquica incorpora los grados relativos de especificidad o generalidad entre los distintos conceptos de un campo semántico. Esta organización no corresponde a un ordenamiento de lo más importante o más usado hacia lo menos importante o menos usado, sino que se basa en conexiones que van de lo general hacia lo específico como son las relaciones entre un conjunto y sus elementos, el todo y sus partes, el género y sus especies, etc. La relación de afinidad se emplea para establecer un nexo entre dos conceptos que están estrechamente relacionados pero entre los que no existe una verdadera relación jerárquica como serían, por ejemplo, los diversos aspectos de un proceso -el aspecto fisiológico y el aspecto anatómico de un proceso biológico-.

Un tesoro permite, por tanto, ampliar las asociaciones

entre los conceptos, facilitar la búsqueda de conceptos periféricos a partir de un tema central y pasar de un campo conceptual a otros relacionados con él.

La organización conceptual de un tesoro se posibilita gracias al establecimiento de un vocabulario estándar controlado. Este vocabulario debe seleccionarse de la terminología del área específica del conocimiento a la que se refiere el tesoro, de manera que constituya un vocabulario mínimo al cual pueda reducirse el universo que se pretende estructurar mediante el uso de sinónimos; en dicho vocabulario deben, además, quedar representadas todas las relaciones conceptuales jerárquicas y de afinidad.

Las ventajas del Indexamiento mediante tesauros son: 1) La existencia de una especie de mapa que indica cómo los conceptos están relacionados entre sí, lo cual ayuda a los Indexadores y a los buscadores a entender la estructura del campo. 2) El uso de un vocabulario estándar con referencias cruzadas entre sus términos -jerárquicas, de preferencia, de afinidad- lo cual asegura que se pueda elegir el término que apropiadamente represente a un concepto dado dentro de la estructura clasificadora del tesoro, y que se use únicamente el término preferente o autorizado de entre el conjunto de sinónimos para Indexar un concepto.

Las referencias cruzadas garantizan que al hacer una búsqueda se elija el término correcto, ya que si un Indexador usara más de un sinónimo para representar un mismo concepto, los

documentos se encontrarían dispersos bajo todos estos sinónimos, y un buscador que eligiera uno de los sinónimos y encontrara documentos indexados bajo éste supondría que ha encontrado el término correcto y detendría su búsqueda sin saber que hay otros documentos útiles indexados bajo los sinónimos. El uso de las referencias jerárquicas y de afinidad permiten a un buscador ser más general o más particular sistemáticamente, si su primera elección de los términos de búsqueda produce ya sea muy pocas o demasiadas referencias al material almacenado.

Uno de los problemas clave en la construcción de un tesoro es la determinación inicial de los candidatos a descriptores. El mejor procedimiento para esto es que especialistas de una disciplina o campo dado hagan la selección original. De aquí se desprende una propuesta de uso complementario del indexamiento libre y el de tesoro. La propuesta incorpora como primer requerimiento que se incluyan en las unidades de información resúmenes o sinopsis del documento original que, como hablamos dicho, proporcionan una gran cantidad de información al lector.

El tipo de resumen que proponemos que se incluya es aquél que, en algunas ocasiones, producen los propios autores de los libros o artículos científicos, es decir, resúmenes hechos por especialistas. Simultáneamente a la producción de estos resúmenes los especialistas efectuarían un indexamiento libre sobre ellos produciendo así una sinopsis integrada que es al mismo tiempo un resumen en lengua natural con todas las ventajas que ya han sido mencionadas, y una serie de palabras clave que

permitirán la recuperación de la unidad de información.

Con estas palabras clave se obtendrá la lista de candidatos a descriptores para la formación del tesoro. De esta manera, la lista de candidatos incorpora el juicio de los especialistas, es decir, los elementos de la organización conceptual y la estandarización terminológica de su área de especialidad, que ellos manejan en forma natural.

Los especialistas envueltos en el proceso deben intentar indexar varios documentos para establecer el vocabulario mínimo que constituirá el tesoro, definir las relaciones jerárquicas y de afinidad entre los términos, y determinar en dónde son necesarias explicaciones que hagan más claros a los descriptores y dónde son relevantes, necesarias o deseables reglas interpretativas y aplicativas.

Por último, el indexamiento con el tesoro resultante puede ser auxiliado por procesos computarizados gracias al indexamiento libre que se tenía.

El agregado automático de los términos autorizados en el tesoro, correspondientes a determinadas palabras claves marcadas mediante el previo indexamiento libre, es un ejemplo de tales procesos computarizados.

De la discusión anterior queda clara la necesidad de que la información que se almacena debe poseer una estructura tal que posibilite recuperación basada en indexamiento tanto libre como con tesoro para poder satisfacer las necesidades de cualquier tipo de usuario.

III. Un esquema estructural para la Información bibliográfica

En el capítulo anterior se discutió fundamentalmente la relación entre las características de la información y los posibles resultados de los procesos de recuperación. En este capítulo analizaremos la estructura específica que se ha decidido imponer a la información bibliográfica del Catálogo de Obras de Asia y Africa (COAA), así como la forma de almacenamiento en la computadora de dicha información. Dejaremos para el siguiente capítulo la descripción de los métodos computacionales mediante los cuales se efectúan los procesos de recuperación y transformación de la información.

Como se mostró en los capítulos anteriores, al imponer una estructura a la información se establecen las posibilidades de identificación de sus partes. Durante la definición de un esquema estructural para información bibliográfica, es necesario considerar el máximo nivel de detalle en la discriminación de las partes debido a que los procesos computacionales de discriminación no siempre son posibles mientras que los de agrupación son relativamente simples. Parfraseando en lenguaje aritmético diríamos que la definición del esquema estructural corresponde a la definición del conjunto de números primos; la imposición de este esquema a la información bibliográfica corresponde a encontrar -discriminar- los factores primos de cada unidad de información; y los procesos de agrupación corresponden

al cálculo de submúltiplos a través de la operación con los factores primos.

Al tratar de definir el esquema estructural que debería imponerse a la Información bibliográfica del COAA se consideró que por lo menos debería incluir discriminadamente la información de las fichas catalográficas que tradicionalmente habían sido producidas por la Biblioteca de El Colegio de México y que eran las que constitulan dicho catálogo. Además, debería poderse agregar a cada unidad de Información los descriptores especializados que los Investigadores del Centro de Estudios de Asia y Africa (CEAA) requerirían durante sus procesos de manipulación bibliográfica, así como los resúmenes de las obras correspondientes.

En 1982 la Biblioteca Daniel Cosío Villegas (BDCV) de El Colegio de México había decidido adoptar el formato MARC (Machine Readable Cataloging) para la catalogación de sus obras, debido a que iniciaba la automatización de sus procesos técnicos.

A pesar de que el proyecto de automatización del COAA había nacido como un proyecto independiente del de automatización de la BDCV, se consideró importante conservar el mismo esquema estructural que usara la BDCV con el propósito de que ambos proyectos utilizaran el mismo sistema computacional de manejo y recuperación de Información catalográfica cuyo desarrollo se iniciaba en aquellos momentos. En realidad, en aquel entonces el proyecto de automatización del COAA se veía como un proyecto piloto de lo que sería el gran proyecto de automatización de la

BDCV. Una de las ideas fundamentales era que al utilizar un estándar como MARC se posibilitaba la catalogación compartida entre la Biblioteca como entidad central y los Centros de Estudios como entidades especializadas. La catalogación compartida ofrecía ventajas en dos modalidades: en una, tanto la Biblioteca como los Centros podrían catalogar obras e intercambiar sus productos evitando con esto la duplicación del trabajo; la otra modalidad sería que la Biblioteca catalogara la información básica y de uso general de cada obra y los Centros agregarán la información propia de sus especialidades como serían descriptores y resúmenes, enriqueciendo así la información. El uso de MARC no sólo permitiría obtener las ventajas de la catalogación compartida al interior de El Colegio de México sino que también permitiría el intercambio de información con otras bibliotecas tanto a nivel nacional como internacional ya que la comunidad bibliotecaria lo ha adoptado no sólo como su esquema estructural estándar sino también como un formato de intercambio del material catalográfico. Particularmente, en México, el formato MARC es utilizado por la Biblioteca Nacional así como por el Sistema de Bibliotecas de la UNAM entre otras.

Damos a continuación un brevíssimo resumen del surgimiento y alcance del formato MARC con el propósito de asentar la importancia de su adopción y poder exponer su composición y su forma de uso dentro de El Colegio de México.

Las investigaciones para establecer la posibilidad de usar técnicas automatizadas en la Library of Congress de los Estados

Unidos (LC) se iniciaron a fines de la década de los años 50. En 1965, con el apoyo del Council on Library Resources y del Committee on Automation of the Association of Research Libraries se llevó a cabo una conferencia en la que participaron representantes de universidades, de agencias de investigación, de agencias del gobierno y de la industria privada. En dicha conferencia se llegó a la conclusión de que la existencia de registros catalográficos estandarizados legibles por computadora facilitaría la automatización de las bibliotecas participantes. Dichos registros catalográficos deberían incluir todos los datos que aparecen en las tarjetas catalográficas junto con la información adicional necesaria para lograr la implantación de técnicas y procedimientos que permitieran el manejo automático de catálogos, búsquedas, indexamientos, y recuperación bibliográfica. Para facilitar la mencionada estandarización se llegó al acuerdo de que la LC diseñara, produjera y distribuyera los registros. Así, en diciembre de 1965 surgió el proyecto piloto denominado MARC bajo la dirección de Henriette D. Avram.

En abril de 1966 la LC estableció el formato MARC I para monografías y produjo cintas con registros MARC los cuales fueron enviados, para su prueba, a las 16 bibliotecas participantes en el proyecto piloto a partir de octubre del mismo año. Con la experiencia ganada en el proyecto piloto y por el interés mostrado por la Biblioteca Nacional Británica, se diseñó el formato MARC II cuya característica principal fue conformar un formato de comunicación e intercambio de datos bibliográficos a

nivel internacional.

El propósito al crear MARC II fue diseñar un formato que pudiera representar físicamente, en un medio legible por computadora, la información bibliográfica pertinente para todas las formas de material tales como libros, series, mapas, música, artículos en revistas, etc. El formato resultante está formado por los siguientes tres componentes: 1) la estructura, o el "contenedor vacío"; 2) los designadores de contenido (etiquetas, indicadores y códigos de subcampo) usados para identificar explícitamente o caracterizar adicionalmente los elementos de los datos; y 3) el contenido, es decir, los datos en sí mismos (nombre de autor, título, etc.).

La estructura de los registros MARC es lo suficientemente detallada que permite la inclusión o exclusión de los elementos de los datos según las necesidades de los usuarios. Durante el año de 1982 cuando se realizó el estudio para definir el formato de captura de los registros catalográficos de la BDCV, se revisó el MARC original y se seleccionó la lista de las etiquetas utilizables por El Colegio de México. Dicha lista incluyó todos los casos eventualmente posibles, aún aquellos que jamás se habían usado en la BDCV, como fue el caso del ISBN. Se adoptaron los códigos alfabéticos para subcampos del MARC original y se respetaron las etiquetas reservadas por la LC, por ejemplo, en la codificación de encabezamientos de materia. La adaptación del MARC a la catalogación de El Colegio de México (MARCOLMEX), la cual se incluye como anexo en esta tesis, consta de 408 campos

entre etiquetas y códigos de subcampos.

FORMA DE ALMACENAMIENTO DE LA INFORMACION

Una vez definidas las categorías de información, es decir una vez definidos los elementos del esquema estructural que van a diferenciarse, existen básicamente dos formas de identificarlos dentro de la computadora: el llamado formato fijo y el formato variable.

En el formato fijo, la identificación de los diversos elementos se logra mediante su posición en el registro del archivo en computadora, o equivalentemente a través de un nombre asignado a cada una de las posiciones predefinidas en el registro, como se hace en los manejadores de bases de datos. Por ejemplo, diríamos que el Autor ocupará los caracteres 1 a 60 del registro, el Título los caracteres del 61 al 180, etc. Nos referimos a la información misma a través de los nombres o las posiciones de los campos. Empezamos a hablar de los elementos de la estructura a través de campos y contenidos de los campos. Por ejemplo, el campo de Título (o los caracteres del 61 al 180 del registro) podría tener el contenido "Escritos sobre la Historia del México Antiguo".

El uso de un formato fijo requiere que se hayan prestablecido todos los campos posibles en que se deseen subdividir o discriminar las unidades de información; en los casos de campos que puedan repetirse, deberá prestablecerse un máximo de repetición; también deberán imponerse longitudes máximas y en el caso de que un campo particular excediera al

máximo establecido, dicho contenido tendría que abreviarse.

Los procesos de recuperación basados en formato fijo son los más eficientes, ya que se buscan los contenidos que se quieren seleccionar solamente en los campos deseados, y las manipulaciones para recuperación con condiciones lógicas son simples de desarrollar.

Este tipo de formato para captación y almacenamiento de información es satisfactorio cuando todas las unidades de información tienen contenidos para todos los campos preestablecidos en la estructura, y las longitudes de los contenidos para cada campo dado son similares en todas las unidades de información.

Sin embargo, cuando la información que se intenta manejar es muy variable en longitudes, posibilidades de repetición o ausencia de determinados campos, el desperdicio de espacio, así como el tamaño que la base puede llegar a tener produce un costoso manejo de los datos.

En el formato variable la identificación de los diversos elementos que se haya decidido discriminar se hace a través de la asociación de etiquetas a cada campo. Estas etiquetas pueden ser un nombre o una etiqueta numérica. Por ejemplo, la palabra Título o la etiqueta 245 de MARC podría ser la etiqueta asociada a cada ocurrencia de un título en una unidad de información. Hablamos entonces de etiquetas y contenidos para un campo dado. En el caso más simple, la información en formato variable está constituida por ocurrencias de parejas etiqueta-contenido. Tanto

las etiquetas como los contenidos pueden ser configuraciones similares de caracteres y además de longitudes variables, lo que hace necesario el uso de un carácter separador como '>' que permita la diferenciación entre las etiquetas y los contenidos. La información se almacena en la computadora como una cadena de etiquetas y contenidos separadas por el carácter especial.

Los procesos de recuperación de información almacenada en formato variable requieren del análisis de toda la información, de ahí la importancia de implantar algoritmos muy rápidos para dicho análisis.

Una vez elegido MARC como el esquema estructural que se impondrá a la información catalográfica del COAA y tomando en cuenta la característica de variabilidad de dicha información, se decidió que la manera más apropiada de almacenarla sería en formato variable.

Presentamos a continuación algunas unidades de información con el propósito de ejemplificar el grado de variabilidad real de la información catalográfica del COAA:

035>bb \$a 001707>
 FCAP>83/08/11>F1>1969>PAIS>ne>ILUS>a>IDIOMA>eng>
 040>bb \$a CM \$c CM>
 092>bb \$a 221.063 \$b In618/1968>
 111>20 \$a International Organization of Old Testament Scholars \$b Congress \$c Roma \$d 1968>
 245>10 \$a Congress volume \$b Rome, 1968>
 260>0b \$a Leiden, Netherlands \$b E. J. Brill \$c 1969>
 300>bb \$a 244 p. \$b Ilus. \$c 25 cm.>
 440>b0 \$a Supplements to Vetus Testamentum \$v 17>
 690>b0 \$a Libros sagrados>
 690>b0 \$a Sagradas escrituras>
 691>b0 \$a Asia \$x Religión>
 695>0b \$a Biblia \$p A. T. \$x Congresos>
 695>0b \$a Biblia \$p A. T. \$x Discursos, ensayos, conferencias>

901>bb \$a 137905>
FIN>

035>bb \$a 001717>
FCAP>83/08/15>F1>1957>PAIS>ne>IDIOMA>eng>
040>bb \$a CM \$c CM>
041>1b \$a eng>
092>bb \$a 221.4 \$b M294m>
130>00 \$a Manual de disciplina \$l Inglés>
245>14 \$a The manual of discipline \$c Translated and annotated
with an Introduction by P. Wernberg-Moiler>
260>0b \$a Leiden, Netherlands \$b E. J. Brill \$c 1957>
300>bb \$a 57 p. \$c 25 cm.>
440>b0 \$a Studies on the Texts of the Desert of Judah \$v 1>
691>b0 \$a Mar Muerto, Rollos>
690>b0 \$a Manuscritos Judios>
691>b0 \$a Jerusalem \$x Destrucción 70 D.C.>
691>b0 \$a Asia \$x Religión>
695>0b \$a Biblia \$x Arqueología>
901>bb \$a 138480>
FIN>

035>bb \$a 001943>
FCAP>83/08/26>F1>1974>PAIS>us>CONT>b>IDIOMA>eng>
040>bb \$a CM \$c CM>
092>bb \$a 016.95204 \$b W262a>
100>10 \$a Ward, Robert Edward>
245>14 \$a The allied occupation of Japan, 1945-1952 \$b an
annotated bibliography of Western-language materials \$c Compiled
and edited for the Joint Committee on Japanese Studies of the
Social Science Research Council-American Council of Learned
Societies and the Center for Japanese Studies of the University
of Michigan by... and Frank Joseph Shulman. With the assistance
of Masashi Nishihara and Mary Tobin Espey>
260>0b \$a Chicago, Ill. \$b American Library Association \$c 1974>
300>bb \$a xx, 867 p. \$c 25 cm.>
691>b0 \$a Japón \$x Historia \$y Ocupación aliada, 1945-1952 \$x
Bibliografía>
691>b0 \$a Asia \$x Historia \$y 1945-1952 \$x Bibliografía>
700>10 \$a Shulman, Frank Joseph \$d 1943- \$e coed.>
901>bb \$a 120474/121029>
FIN>

035>bb \$a 001957>
FCAP>83/08/26>F1>1917>PAIS>us>CONT>h>IDIOMA>eng>
040>bb \$a CM \$c CM>
092>bb \$a 221.48 \$b S215o>
100>10 \$a Sanders, Henry Arthur \$d 1868->
245>14 \$a The Old Testament manuscripts in the Freer collection
\$c by Henry A. Sanders...>
260>0b \$a New York \$b The Macmillan company \$c 1917>
300>bb \$a 357 p. \$b facsim. \$c 28 cm.>

440>b0 \$a University of Michigan Studies Humanistic series \$v 8>
 505>0b \$a I. The Washington manuscript of Deuteronomy and
 Joshua.- II. The Washington manuscript of the Psalms>
 691>b0 \$a Asia \$x Religión>
 695>0b \$a Biblia \$h Manuscritos griegos \$p A.T. \$k Selecciones>
 690>b0 \$a Manuscritos griegos \$x Facsimiles>
 730>00 \$a Biblia \$p A.T. \$t Salmos \$l Griego \$f 1917>
 710>20 \$a Freer gallery of art \$c Washington, D.C. Mss. (greek)>
 901>bb \$a 016884>
 FIN>

Al producir una unidad de Información los catalogadores seleccionan únicamente los campos pertinentes a la obra en tanto que no consideran las etiquetas no aplicables. Compárense por ejemplo las unidades primera y cuarta: la primera incluye el campo de Ilustración (etiqueta ILUS) que no aparece en la cuarta, mientras que la cuarta incluye una nota de contenido (etiqueta 505) y asientos secundarios (etiquetas 710 y 730) que no aparecen en la primera.

Por otra parte, el nivel de detalle en la discriminación de las partes de la Información al que obliga MARC establece muchas posibilidades de elección dentro de un grupo de etiquetas dado, por ejemplo, para el grupo de etiquetas de asiento principal (etiquetas 100, 110, 111 y 130) el catalogador deberá seleccionar una de ellas de acuerdo al tipo de elemento que fungirá como asiento principal de la unidad catalogada; compárense el autor personal (etiqueta 100) del tercer ejemplo con la conferencia o reunión (etiqueta 111) del primero y con el título uniforme (etiqueta 130) del segundo.

Otro factor involucrado en la variabilidad es que la longitud del contenido de un campo dado varía considerablemente

de una unidad de información a otra: compárense, por ejemplo, los 35 caracteres del título (etiqueta 245) de la primera unidad con los 406 caracteres correspondientes al título de la tercera unidad. Queremos hacer notar que en casi todos los casos, como en los ejemplos aquí mencionados, la variabilidad en la longitud está relacionada con el nivel de detalle en la catalogación que es permitido por MARC. La existencia de subcampos da al catalogador la posibilidad de incorporar más información de manera estructurada en una unidad determinada. Por ejemplo, en la primera unidad únicamente se tiene información sobre el título propiamente dicho (subcampo \$a de la etiqueta 245) y sobre el subtítulo (subcampo \$b de la etiqueta 245), mientras que en la tercera unidad además de la información sobre título y subtítulo se tiene la mención de responsabilidad (subcampo \$c de la etiqueta 245).

Por último, MARC permite la repetición de algunas etiquetas o grupos de etiquetas, en particular, los encabezamientos de materia (etiquetas 690, 691, 692, 693, 694 y 695). Es importante considerar el hecho de que mientras más encabezamientos se asignen a una unidad determinada, habrá más puntos de acceso para la recuperación de dicha unidad. Sin embargo, habrá unidades de información que, por su tipo, no llevan encabezamientos de materia, por ejemplo las unidades correspondientes a obras literarias.

Aunando todos los factores de variabilidad que hemos mencionado en los cuatro párrafos anteriores, podemos concluir

que el grado de variabilidad de la información catalogada bajo el esquema estructural de MARC es muy alto. Si intentáramos abordar el problema de análisis de esta información bajo un esquema en formato fijo, tendríamos que preestablecer máximos de longitud para el contenido de las 408 partes en que se puede subdividir una unidad de información -etiquetas y subcampos del MARC-, así como máximos de repetición para los campos repetibles, lo cual nos daría una longitud del orden de 10,000 caracteres para una unidad de información. Simplemente para almacenar las cuatro unidades de información que nos han servido de ejemplo bajo un esquema en formato fijo, tendríamos que reservar 1316 caracteres para cada una; mientras que almacenadas en formato variable la ficha de longitud mínima ocupa 589 caracteres y la de longitud máxima ocupa 857; por lo tanto la cantidad de caracteres usados en formato fijo comparada con la cantidad de caracteres usados en formato variable por la ficha promedio guardan una razón de 2 a 1. Hacemos notar que las 4 fichas con las que ejemplificamos no forman una muestra en la que estén representadas las posibilidades de variación que permite MARC y que entre más fichas se consideren la razón de uso de espacio en memoria aumentará.

IV. El sistema de recuperación bibliográfica

En este capítulo describiremos los métodos computacionales mediante los cuales efectuamos los procesos de recuperación bibliográfica. Para realizar estos procesos sobre información almacenada en formato variable se requiere localizar rápidamente las ocurrencias de determinadas cadenas de caracteres en un texto para posteriormente realizar las operaciones de selección y transformación de la información.

Se verá que el sistema que efectúa los procesos de recuperación es un autómata finito que reconoce el lenguaje conformado por las etiquetas y códigos de subcampos de MARC. Durante el análisis léxico se efectúa el rastreo y la identificación de las etiquetas y al llegar a un estado terminal, se toma la decisión de cómo manipular el contenido correspondiente a la etiqueta identificada, es decir, de qué operaciones de selección o transformación hay que aplicar al contenido del campo.

A continuación presentamos los principales conceptos matemáticos que nos permitirán dar una clara descripción de los algoritmos, estructuras y métodos computacionales que conforman este trabajo.

Un símbolo es una entidad abstracta que no se define formalmente pero que tiene un significado intuitivo suficientemente claro. Las letras, los dígitos y, en general, todos los caracteres son ejemplos de símbolos usados

frecuentemente.

Una cadena o palabra es una secuencia finita de símbolos yuxtapuestos. La longitud de una cadena w se denota por $|w|$ y es el número de símbolos que componen la cadena. La cadena vacía se denota por ϵ y es la cadena que consiste de cero símbolos.

Un alfabeto es un conjunto finito de símbolos. Un lenguaje formal sobre un alfabeto es un conjunto de cadenas de símbolos de dicho alfabeto.

Una gráfica, denotada $G=(V,E)$, consiste de un conjunto finito V de vértices o nodos y un conjunto E de parejas de vértices llamadas lados.

Una trayectoria en una gráfica es una secuencia de vértices v_1, v_2, \dots, v_k , $k \geq 1$ tal que hay un lado (v_i, v_{i+1}) para todo i con $1 \leq i < k$. La longitud de la trayectoria es $k-1$.

Una gráfica dirigida o una digráfica, que se denota $D=(V,E)$, consiste de un conjunto V de vértices y un conjunto E de parejas ordenadas de vértices llamadas arcos. Si $v, w \in V$ denotamos un arco de v a w por $v \rightarrow w$.

Un recorrido en una digráfica es una secuencia de vértices v_1, v_2, \dots, v_k $k \geq 1$ tal que $v_i \rightarrow v_{i+1}$ es un arco para cada i , con $1 \leq i < k$. Se dice que el recorrido es de v_1 a v_k .

Si $v \rightarrow w$ es un arco se dice que v es un predecesor de w y que w es un sucesor de v .

Un árbol dirigido y ordenado es una digráfica con las siguientes propiedades:

- 1) Hay un vértice, llamado raíz, que no tiene

predecesores y desde el cual hay un recorrido hacia cualquier vértice.

- 2) Cualquier vértice que no sea la raíz tiene exactamente un predecesor.
- 3) Los sucesores de cada vértice están ordenados de izquierda a derecha.

En la literatura especializada se utiliza una terminología especial para los árboles que difiere de la terminología general para gráficas: el sucesor de un vértice se llama hijo y el predecesor es llamado padre; a un vértice que no tiene hijos se le llama hoja y a los otros vértices se les llama vértices interiores.

Un autómata finito es un modelo matemático de un sistema con entradas y salidas discretas. El sistema puede estar en cualquiera de un número finito de configuraciones o "estados". El estado del sistema resume la información concerniente a las entradas pasadas que es necesaria para determinar el funcionamiento del sistema en las subsecuentes entradas. En computación se encuentran muchos ejemplos de sistemas de estados finitos y la teoría de autómatas finitos es una herramienta de diseño muy útil para estos sistemas.

Un autómata finito (AF) consiste de un conjunto finito de estados y un conjunto finito de transiciones de estado a estado que ocurren con símbolos de entrada tomados de un alfabeto. Para cada símbolo de entrada hay exactamente una transición hacia afuera de cada estado (posiblemente hacia sí mismo). Un estado

usualmente denotado q_0 , es el estado inicial, en el cual empieza el autómata. Algunos estados se designan como estados finales.

Se asocia una digráfica, llamada diagrama de transiciones, a cada AF de la siguiente forma: los vértices de la gráfica corresponden a los estados del AF, si hay una transición del estado q al estado p con una entrada a , entonces hay un arco etiquetado con a del vértice o estado q al p en el diagrama de transición. Un AF acepta una cadena x si la secuencia de transiciones correspondientes a los símbolos de x lleva del estado inicial a algún estado final.

Formalmente se denota un autómata finito por un quinteto $(Q, \Sigma, \delta, q_0, F)$ donde Q es un conjunto finito de estados, Σ es un alfabeto finito de entrada, $q_0 \in Q$ es el estado inicial, $F \subseteq Q$ es el conjunto de estados finales, y δ es una función de transición $\delta: Q \times \Sigma \rightarrow Q$. Esto es (q, a) es un estado para cada estado $q \in Q$ y símbolo de entrada $a \in \Sigma$.

Una vez dadas las definiciones matemáticas fundamentales, entraremos de lleno a la explicación formal de los métodos computacionales mediante los que se efectúan los procesos de recuperación bibliográfica.

Recordará el lector que la información catalográfica con la que trabajamos ha sido estructurada con base en el esquema de MARC, y se almacena en formato variable como ocurrencias de parejas etiqueta-contenido. Los procedimientos de recuperación bibliográfica constan de dos partes: la primera, a la que llamamos análisis léxico, es la que efectúa la localización

dentro de la información del COAA de las etiquetas MARC que identifican los campos de la estructura; y la segunda, a la que llamaremos procesamiento semántico, que es la aplicación de operaciones de selección y transformación a la información que constituye el contenido de la etiqueta localizada durante el análisis léxico.

La familia de programas mediante los cuales se efectúan los diversos procesos de recuperación bibliográfica, son implantaciones computacionales eficientes de autómatas finitos que realizan el análisis léxico, aunadas a subrutinas o procedimientos en las que se realiza el procesamiento semántico. El autómata finito que realiza el reconocimiento de la clase restringida de expresiones regulares constituida por el lenguaje de MARC es construido automáticamente a partir de una descripción en la que cada etiqueta o código de subcampo del lenguaje MARC está asociada al operador semántico que se desea aplicar al contenido de esa etiqueta durante el momento del procesamiento de la información.

El análisis léxico está controlado por un autómata finito cuyo diagrama de transiciones es un árbol que representa a las etiquetas de la manera que se describe a continuación. Para que esta descripción sea clara, iremos construyendo el árbol correspondiente a algunas etiquetas MARC sin considerar los códigos de subcampos. Cada etiqueta se representa como una secuencia de caracteres alfanuméricos terminada en el carácter separador. La lista de dichas secuencias se debe ordenar

alfabéticamente para que pueda construirse adecuadamente la tabla que representará al árbol.

020>	098>	260>	503>	692>	810>	EDCONM>	LTIP>
035>	099>	300>	504>	693>	811>	F1>	LNBIB>
040>	100>	400>	505>	694>	830>	F2>	LNDES>
041>	110>	410>	507>	695>	840>	FCAP>	LFOR>
043>	111>	411>	520>	700>	901>	FIN>	PAIS>
049>	130>	440>	533>	710>	902>	FSPUB>	PUBOF>
050>	240>	490>	546>	711>	AP>	IDIOMA>	REPROD>
080>	241>	500>	590>	730>	BIO>	ILUS>	
082>	245>	501>	690>	740>	CONF>	IND>	
092>	250>	502>	691>	800>	CONT>	LEST>	

FIGURA 1. Lista de etiquetas MARC.

La raíz del árbol tiene tantos hijos como caracteres diferentes puede haber en el lugar de más a la izquierda de cada etiqueta MARC. Y los arcos que unen a la raíz con cada uno de sus hijos estarán etiquetados con dichos caracteres. En la figura 2 el arco con la marca '1' del primer nivel corresponde tanto a la etiqueta '100>' como a la '110>', la '111>' y la '130>'. De la misma manera, a los arcos con las marcas '2', '3', etcétera les corresponden sus respectivas listas de etiquetas MARC.

Recursivamente, tomando todas las etiquetas MARC que tuvieron el mismo carácter en el lugar de más a la izquierda, construímos subárboles en cada nodo del primer nivel y obtendremos un segundo nivel en el árbol.

Siguiendo este proceso llegará el momento en que todas las

etiquetas MARC estén representadas en el árbol mediante un recorrido desde la raíz hasta una hoja. El carácter separador '>' se considera como el último carácter de cada etiqueta MARC, de manera que todos los arcos etiquetados con ese carácter llegan a una hoja de donde se hace depender un nodo agregado en donde está codificado un operador semántico.

El árbol y los nodos agregados se implantan como una tabla que es generada automáticamente por el programa Arbol.c incluido en el anexo A. Los estados del autómata, es decir, los nodos del árbol, junto con los nodos agregados se numeran por niveles desde la raíz hacia las hojas de izquierda a derecha, de manera que todos los hijos de un nodo determinado están numerados consecutivamente. Esta numeración se corresponde con el número del registro de la tabla. Existen, por lo tanto, dos tipos de registros: los que corresponden a los estados del autómata y los correspondientes a los nodos agregados; el primer carácter de cada registro marca la diferencia entre un estado del autómata y un nodo agregado. El carácter "*" marca los estados del autómata, mientras que cualquier otro carácter se podrá utilizar como el código del operador semántico.

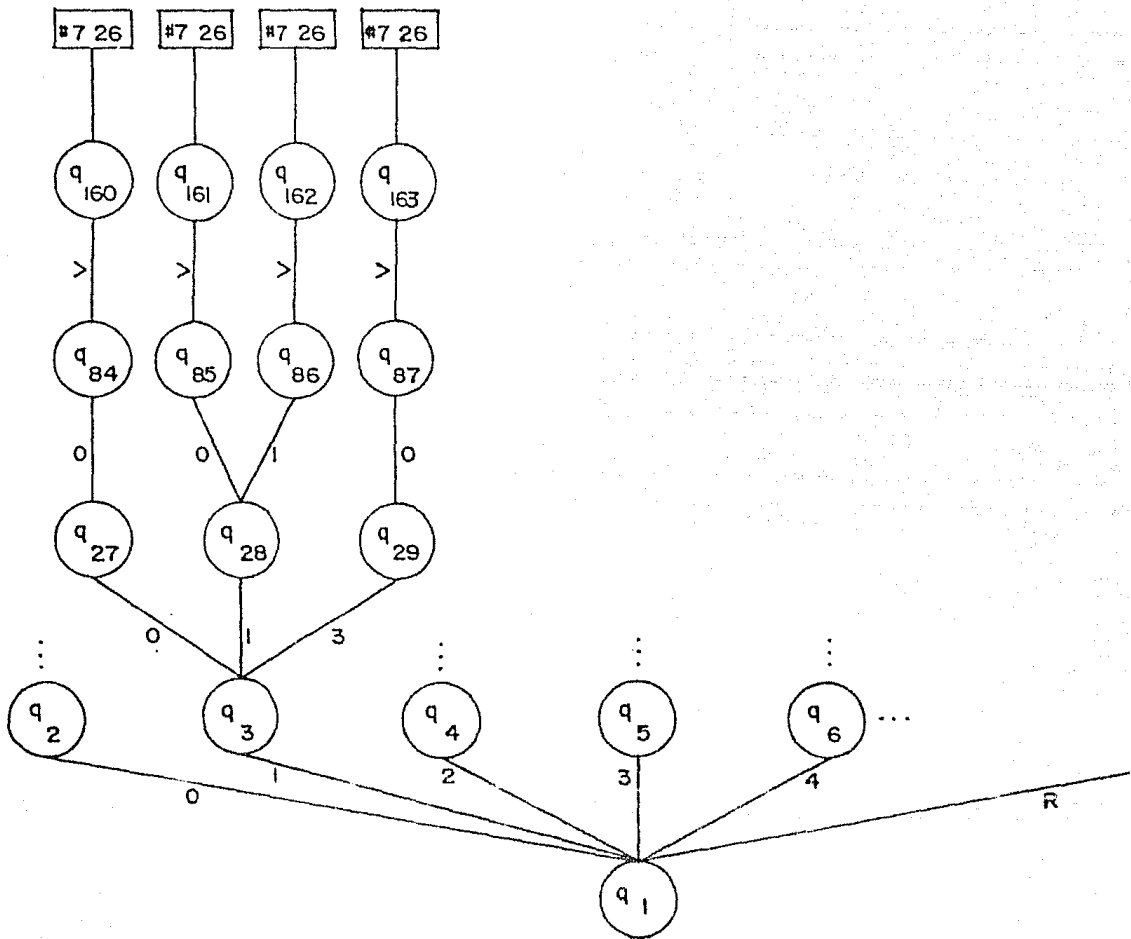


FIGURA 2. Diagrama de transiciones del autómata.

1:	*	<	19	2	51:	*	4	1	125	101:	*	2	1	177
2:	*	0	6	21	52:	*	0	2	126	102:	*	3	1	178
3:	*	1	3	27	53:	*	P	1	128	103:	*	4	1	179
4:	*	2	3	30	54:	*	1	1	129	104:	*	5	1	180
5:	*	3	1	33	55:	*	O	1	130	105:	*	7	1	181
6:	*	4	4	34	56:	*	D	1	131	106:	*	0	1	182
7:	*	5	5	38	57:	*	1	1	132	107:	*	3	1	183
8:	*	6	1	43	58:	*	2	1	133	108:	*	6	1	184
9:	*	7	4	44	59:	*	C	1	134	109:	*	0	1	185
10:	*	8	4	48	60:	*	1	1	135	110:	*	0	1	186
11:	*	9	1	52	61:	*	S	1	136	111:	*	1	1	187
12:	*	A	1	53	62:	*	D	1	137	112:	*	2	1	188
13:	*	B	1	54	63:	*	L	1	138	113:	*	3	1	189
14:	*	C	1	55	64:	*	N	1	139	114:	*	4	1	190
15:	*	E	1	56	65:	*	E	1	140	115:	*	5	1	191
16:	*	F	5	57	66:	*	T	1	141	116:	*	0	1	192
17:	*	I	3	62	67:	*	N	2	142	117:	*	0	1	193
18:	*	L	4	65	68:	*	F	1	144	118:	*	1	1	194
19:	*	P	2	69	69:	*	A	1	145	119:	*	0	1	195
20:	*	R	1	71	70:	*	U	1	146	120:	*	0	1	196
21:	*	2	1	72	71:	*	E	1	147	121:	*	0	1	197
22:	*	3	1	73	72:	*	0	1	148	122:	*	0	1	198
23:	*	4	4	74	73:	*	5	1	149	123:	*	1	1	199
24:	*	5	1	78	74:	*	0	1	150	124:	*	0	1	200
25:	*	8	2	79	75:	*	1	1	151	125:	*	0	1	201
26:	*	9	3	81	76:	*	3	1	152	126:	*	1	1	202
27:	*	0	1	84	77:	*	9	1	153	127:	*	2	1	203
28:	*	1	2	85	78:	*	0	1	154	128:	*	>	1	204
29:	*	3	1	87	79:	*	0	1	155	129:	*	0	1	205
30:	*	4	3	88	80:	*	2	1	156	130:	*	N	2	206
31:	*	5	1	91	81:	*	2	1	157	131:	*	C	1	208
32:	*	6	1	92	82:	*	8	1	158	132:	*	>	1	209
33:	*	0	1	93	83:	*	9	1	159	133:	*	>	1	210
34:	*	0	1	94	84:	*	0	1	160	134:	*	A	1	211
35:	*	1	2	95	85:	*	0	1	161	135:	*	N	1	212
36:	*	4	1	97	86:	*	1	1	162	136:	*	P	1	213
37:	*	9	1	98	87:	*	0	1	163	137:	*	1	1	214
38:	*	0	7	99	88:	*	0	1	164	138:	*	U	1	215
39:	*	2	1	106	89:	*	1	1	165	139:	*	D	1	216
40:	*	3	1	107	90:	*	5	1	166	140:	*	S	1	217
41:	*	4	1	108	91:	*	0	1	167	141:	*	1	1	218
42:	*	9	1	109	92:	*	0	1	168	142:	*	B	1	219
43:	*	9	6	110	93:	*	0	1	169	143:	*	D	1	220
44:	*	0	1	116	94:	*	0	1	170	144:	*	O	1	221
45:	*	1	2	117	95:	*	0	1	171	145:	*	1	1	222
46:	*	3	1	119	96:	*	1	1	172	146:	*	B	1	223
47:	*	4	1	120	97:	*	0	1	173	147:	*	P	1	224
48:	*	0	1	121	98:	*	0	1	174	148:	*	>	1	225
49:	*	1	2	122	99:	*	0	1	175	149:	*	>	1	226
50:	*	3	1	124	100:	*	1	1	176	150:	*	>	1	227

FIGURA 3. Tabla en la que se representa el autómata.

151: * v 1 228	201: * v 1 278	251: 0
152: * v 1 229	202: * v 1 279	252: 0
153: * v 1 230	203: * v 1 280	253: 0
154: * v 1 231	204: 0	254: 0
155: * v 1 232	205: * v 1 281	255: 0
156: * v 1 233	206: * F 1 282	256: 0
157: * v 1 234	207: * T 1 283	257: 0
158: * v 1 235	208: * O 1 284	258: 0
159: * v 1 236	209: 0	259: 0
160: * v 1 237	210: 0	260: 0
161: * v 1 238	211: * P 1 285	261: 0
162: * v 1 239	212: * v 1 286	262: 0
163: * v 1 240	213: * U 1 287	263: 0
164: * v 1 241	214: * O 1 288	264: 0
165: * v 1 242	215: * S 1 289	265: 0
166: * v 1 243	216: * v 1 290	266: 0
167: * v 1 244	217: * T 1 291	267: 0
168: * v 1 245	218: * P 1 292	268: 0
169: * v 1 246	219: * I 1 293	269: 0
170: * v 1 247	220: * E 1 294	270: 0
171: * v 1 248	221: * R 1 295	271: 0
172: * v 1 249	222: * S 1 296	272: 0
173: * v 1 250	223: * O 1 297	273: 0
174: * v 1 251	224: * R 1 298	274: 0
175: * v 1 252	225: 0	275: 0
176: * v 1 253	226: # 0 6	276: 0
177: * v 1 254	227: 0	277: 0
178: * v 1 255	228: 0	278: 0
179: * v 1 256	229: 0	279: 0
180: * v 1 257	230: 0	280: 0
181: * v 1 258	231: 0	281: 0
182: * v 1 259	232: 0	282: * v 1 299
183: * v 1 260	233: 0	283: * v 1 300
184: * v 1 261	234: 0	284: * N 1 301
185: * v 1 262	235: 0	285: * v 1 302
186: * v 1 263	236: 0	286: \$
187: * v 1 264	237: # 7 26	287: * B 1 303
188: * v 1 265	238: # 7 26	288: * M 1 304
189: * v 1 266	239: # 7 26	289: * v 1 305
190: * v 1 267	240: # 7 26	290: 0
191: * v 1 268	241: # 34 29	291: * v 1 306
192: * v 1 269	242: # 34 29	292: * v 1 307
193: * v 1 270	243: # 34 29	293: * B 1 308
194: * v 1 271	244: 0	294: * S 1 309
195: * v 1 272	245: 0	295: * v 1 310
196: * v 1 273	246: 0	296: * v 1 311
197: * v 1 274	247: 0	297: * F 1 312
198: * v 1 275	248: 0	298: * O 1 313
199: * v 1 276	249: 0	299: 0
200: * v 1 277	250: 0	300: 0

Continuación de la FIGURA 3.

301: * M 1 314	309: * > 1 318	317: 0
302: 0	310: 0	318: 0
303: * > 1 315	311: 0	319: 0
304: * A 1 316	312: * > 1 319	320: * > 1 323
305: 0	313: * D 1 320	321: 0
306: 0	314: * > 1 321	322: 0
307: 0	315: 0	323: 0
308: * > 1 317	316: * > 1 322	

Continuación de la FIGURA 3.

En los registros correspondientes a los estados del autómata está codificada la información necesaria para obtener la función de transición del autómata; cada uno de estos registros está formado por 4 campos: en el primero, al que denotaremos como "tipo_nodo", está el "*"; en el segundo, denotado "etiqueta_arco", se almacena el carácter que debe ser comparado con el carácter actual de la cadena de entrada para decidir si se realiza o no la transición al estado correspondiente a este registro; en el tercero, "hijos", se guarda el número de hijos que tiene el nodo correspondiente al registro; y en el cuarto, "primer_hijo", está el apuntador al registro a partir del cual están secuencialmente almacenados los hijos del nodo en cuestión. En cambio, cuando el registro corresponde a un nodo agregado en él estará codificada la información necesaria para una adecuada aplicación de la acción semántica. La figura 3 es la tabla que representa al autómata para el reconocimiento de las etiquetas MARC sin códigos de subcampos. En ella se han codificado, a modo

de ejemplo, tres operadores semánticos con sus parámetros ("0", "*" y "\$").

En la figura 4 se muestra el algoritmo que reproduce la función de transición del autómata a partir de la información almacenada en la tabla. El estado en el que se encuentra el autómata se representa mediante un apuntador "estado_actual" a los registros de la tabla. El primer registro de la tabla corresponde al estado inicial del autómata que es donde se inicia el análisis para cada etiqueta MARC. Para decidir a qué estado debe hacerse la transición para el siguiente carácter de la cadena de entrada, es necesario revisar, de acuerdo al orden secuencial preestablecido, las etiquetas de los arcos que llevan del estado actual a cada uno de sus hijos. Estas etiquetas están almacenadas precisamente en los registros correspondientes a dichos hijos, por lo que la variable estado_actual está siempre un nivel hacia adelante del estado actual real del autómata. En el momento en que el carácter de la cadena de entrada coincide con una de las etiquetas de los hijos del estado actual, se realiza la transición a ese hijo, la cual consiste en la decodificación del número de hijos del estado actual en la variable hijo, la cual controlará el número de revisiones de las etiquetas. Este es el único momento en el que la variable estado_actual apunta exactamente al estado actual real del autómata. Inmediatamente después de esta transición se vuelve a "ver hacia adelante" con la variable estado_actual. Ocasionalmente, después de haber revisado todas las etiquetas de

los arcos que llevan de un estado a sus hijos, es decir, la variable hijo ha sido decrementada hasta cero, no habrá coincidido el carácter de entrada con ninguna de ellas, lo cual significará haber detectado un error en la información.

```

error <-- FALSO
estado_actual <-- 1
hijo <-- tabla[estado_actual].hijos
c <-- obtener carácter de entrada
estado_actual <-- tabla[estado_actual].primer_hijo
MIENTRAS tabla[estado_actual].tipo_nodo=estado del autómata
    Y NO error() EFECTUA
(
    SI c = tabla[estado_actual].etiqueta_arco ENTONCES
    (
        hijo <-- tabla[estado_actual].hijos
        c <-- obtener carácter de entrada
        estado_actual <-- tabla[estado_actual].primer_hijo
    )
    SINO
    (
        estado_actual <-- estado_actual + 1
        hijo <-- hijo - 1
        SI hijo = 0 ENTONCES error <-- CIERTO
    )
)

```

FIGURA 4. Algoritmo que reproduce la función de transición del autómata.

El hecho de que la variable estado_actual esté siempre "viendo hacia adelante" permite que se identifique que el autómata está en un estado final porque el tipo de nodo al que apunta dicha variable tiene el código de un operador semántico. Cuando el autómata llega a un estado final, es decir, acepta una etiqueta MARC, se aplica el operador semántico -cuyo código está en el nodo agregado a ese estado final- a la información de entrada que es en realidad el contenido de la etiqueta identificada.

En la tabla de la figura 3 en la que se representa el autómata y sus nodos agregados fueron codificados los operadores semánticos necesarios para construir un subconjunto de la información catalográfica en formato fijo, que incluya el número de control Colmex (etiqueta 035), el autor (etiquetas 100, 110, 111 y 130) y el título (etiquetas 240, 241 y 245).

Para dejar claro cómo se usa la tabla para el reconocimiento e identificación de etiquetas MARC, así como para tomar la decisión de cuál operador semántico utilizar, listaremos a continuación los registros de la tabla que se van seleccionando durante el procesamiento del siguiente ejemplo:

....245>The allied occupation of Japan>...

<u>estado actual</u>	<u>c</u>	<u>Resto de la cadena de entrada</u>
1: * < 19 2	245>The allied occupation of Japan>...
2: * 0 6 21	245>The allied occupation of Japan>...
3: * 1 3 27	245>The allied occupation of Japan>...
4: * 2 3 30	245>The allied occupation of Japan>...
30: * 4 3 88	45>The allied occupation of Japan>...
88: * 0 1 164	5>The allied occupation of Japan>...
89: * 1 1 165	5>The allied occupation of Japan>...
90: * 5 1 166	5>The allied occupation of Japan>...
166: * > 1 243	>The allied occupation of Japan>...
243: # 34 29	The allied occupation of Japan>...

Al alcanzar el nodo agregado 243 se inicia la aplicación del operador semántico '#' de escritura en formato fijo, y el resultado de esto es poner la información del título "The allied occupation of Japan" en los caracteres del 34 al 62 del registro de salida.

El reconocimiento del carácter separador '>' después del

título produce que se reinicie la operación del autómata.

Por el orden en que la información entra al proceso de análisis, ya debió haberse hecho, para la unidad de información que nos sirve de ejemplo, el reconocimiento y la transformación de los campos de número de control Colmex y de autor, lo cual ha producido que en el registro de salida ya esté escrita adecuadamente dicha información.

A continuación se presenta el resultado del procesamiento dirigido por la tabla de la figura 3 para los cuatro ejemplos de unidades de información del capítulo anterior:

001707	International organization	Congress volume
001717	Manual de disciplina	The manual of discipline
001943	Ward, Robert Edward	The allied occupation of Japan
001957	Sanders, Henry Arthur	The Old Testament Manuscripts

Los operadores semánticos son pequeños procedimientos, subrutinas, o secciones del programa de análisis mediante las que se efectúa la transformación o selección de la información. Estos operadores deben ser programados de acuerdo al objetivo particular del sistema.

En el ejemplo presentado se codificaron tres tipos de operadores semánticos:

- 1) El operador semántico de escritura en formato fijo '*' con sus dos parámetros 'carácter inicial' y 'longitud máxima' permite seleccionar los campos que se incluirán en el subconjunto, y, al mismo tiempo, copiar el

contenido del campo identificado al registro de salida desde la posición indicada por el primer parámetro; en caso de que la longitud del contenido exceda al valor del segundo parámetro el resto del contenido será omitido.

- 2) El operador semántico de omisión 'O' permite la exclusión de los campos que no se desea incluir en el subconjunto.
- 3) El operador de fin de ficha '\$' hace las reinicializaciones apropiadas.

Una vez construida la tabla que representa al autómata que reconoce el lenguaje MARC, es muy simple codificar, en los nodos agregados, los operadores semánticos seleccionados, para las etiquetas que interese. La posibilidad de cambiar los operadores seleccionados, o asignarlos a otras etiquetas, produce, de hecho, que un mismo programa se comporte de formas variadas.

Como se dijo en su momento, la tabla de la figura 3 fue construida sin incluir los códigos de subcampo de MARC, con el propósito de dar claridad a la explicación sobre la construcción del autómata y el análisis léxico. Cuando se quiere analizar la información catalogada bajo el esquema estructural de MARC considerando en forma discriminada los subcampos de cada etiqueta, se construye para cada etiqueta -adicionalmente al autómata de reconocimiento de las etiquetas MARC descrito anteriormente- un autómata de reconocimiento de los códigos de subcampos que corresponden a dicha etiqueta. Cada autómata

adicional se construye con la metodología descrita para el autómata principal. Su tabla correspondiente se incluye en la tabla principal ligada a través de un registro de conexión que está en lugar del nodo agregado al nodo final de reconocimiento de la etiqueta dada. En cada uno de estos registros se codifica el apuntador que lleva al estado inicial del autómata correspondiente. En los registros de conexión se pueden codificar además los parámetros que sean comunes a todo el conjunto de operadores semánticos asociados a los códigos de subcampo de la etiqueta dada. El análisis de los subcampos se efectúa usando el autómata adicional correspondiente -con el mismo algoritmo de la figura 4- hasta encontrar el separador que marca el fin del contenido de la etiqueta, con lo que se regresa al autómata principal para seguir el análisis.

El autómata principal y los autómatas adicionales se implantan con estructuras similares y se utilizan con el mismo algoritmo, de aquí que sea natural representarlos en una misma tabla utilizando los registros de conexión. Usaremos el término "autómatas conectados" para referirnos a los autómatas que se usan en conjunto con la modalidad de representación descrita.

Con los operadores semánticos de exclusión de campos y de transformación a formato fijo que acabamos de explicar puede crearse la estructura de archivos con la información catalográfica que proporcione los mismos resultados que dan los catálogos por autor y por título usados normalmente en una biblioteca. Describiremos brevemente a continuación, también a

manera de ejemplo, la forma en que se puede crear dicha organización de información. Usando los operadores semánticos exactamente como lo describimos en el ejemplo sobre el análisis léxico, se construye un archivo en formato fijo que contiene en cada registro la información del campo de autor, la del campo de título, y además el número de control Colmex para cada unidad de información. Este archivo se carga en cualquier manejador de bases de datos. Por otro lado, se crea un archivo cuyos registros guardan el número de control Colmex para cada unidad de información y los apuntadores al inicio y fin de dicha unidad de información completa en formato variable. Este archivo es creado por supuesto con un programa que utiliza la misma metodología de análisis léxico. Mediante las operaciones del manejador de bases de datos se realizan las selecciones de los registros que cumplan las condiciones impuestas por el usuario a la información de autor y de título. También, mediante operaciones del manejador, se obtienen los apuntadores al inicio y fin de los registros seleccionados, y dichos apuntadores son descargados de la base de datos. Posteriormente un programa muy sencillo toma los apuntadores de registros seleccionados y recupera, del archivo en formato variable, la información completa de las unidades seleccionadas.

En el anexo "A" se incluyen: el programa que construye la tabla en la que se representan autómatas (Arbol.c), el programa que codifica el contenido de los registros de conexión y de los nodos agregados (Semantic.c), el programa que transforma a

formato fijo subconjuntos de la información catalográfica (Indices.c), el programa que crea los apuntadores para la recuperación de las fichas en formato variable (Creapun.c), el programa que recupera dichas fichas (Recupera.c), así como algunas de las tablas que utilizan.

A continuación queremos discutir las similitudes que existen entre lo que hemos denominado autómatas conectados y operadores semánticos, y el modelo de redes de transición aumentadas (ATN). Para esto usaremos algunos de los conceptos y argumentos que presenta W. A. Woods en su artículo "Transition Network Grammars for Natural Language Analysis" [Woods, 1970].

"Una red de transición recursiva es una gráfica dirigida con estados y arcos etiquetados...Esencialmente se parece a un diagrama de transiciones de un autómata finito no determinístico excepto porque las etiquetas de los arcos pueden ser nombres de estados además de símbolos terminales". Podríamos decir que los registros de conexión de nuestro formalismo corresponden a arcos etiquetados con el nombre de un estado. "La interpretación de un arco etiquetado con un nombre de estado es que el estado al que llega el arco será metido a una pila y el control se transferirá (sin avanzar la cadena de entrada) al estado que etiqueta al arco". Efectivamente el apuntador al registro de conexión es conservado en una variable especial con el propósito de regresar a ese "estado" cada vez que se reinicia el reconocimiento de un subcampo para una etiqueta MARC dada. "Los nombres de estados que pueden aparecer en los arcos en este modelo son esencialmente

nombres de construcciones que se pueden encontrar como "frases" de la cadena de entrada. El efecto de un arco etiquetado como estado es que la transición que representa puede llevarse a cabo si una construcción del tipo indicado se encuentra como "frase" en el punto apropiado de la cadena de entrada". En nuestro caso las frases o constituyentes que debemos encontrar son los subcampos de cada etiqueta MARC.

Para el problema que discutimos en esta tesis, es decir, para el reconocimiento y análisis de las fichas catalográficas en lenguaje MARC no es necesaria la aplicación recursiva de los autómatas principal o agregados debido a lo cual la pila queda reducida a una variable y la representación de la red de transición puede ser simplemente la tabla que hemos descrito.

Woods resalta algunas de las características de las ATN: "Las principales características que una gramática transformacional agrega a las gramáticas libres de contexto son la capacidad de mover, copiar o borrar fragmentos de la estructura de la oración, y efectuar estas acciones sobre constituyentes que generalmente dependen del contexto en el que ocurren. Nosotros podemos agregar facilidades equivalentes al modelo de redes de transición agregando a cada arco de la red ... un conjunto de acciones para construir estructura que se ejecutarán si se sigue el arco. Llamamos a esta versión del modelo redes de transición aumentadas." "Las redes de transición aumentadas van construyendo una descripción estructural parcial de la oración conforme van de estado a estado através de la red.

Las piezas de esta descripción parcial son almacenadas en registros ... Las acciones para construir estructuras especificadas en los arcos cambian el contenido de estos registros en función de contenidos previos, de contenidos de otros registros, el símbolo de entrada, y/o los resultados de operaciones de nivel inferior ... las subestructuras eventualmente se incorporarán a estructuras mayores..." .

Podríamos intentar establecer el siguiente paralelismo: la cadena de entrada a nuestro proceso corresponde a la estructura superficial de una ficha MARC, y el producto que resultará del análisis y transformación de una estructura superficial mediante nuestros operadores semánticos corresponde a una especie de estructura profunda. Aquí los operadores semánticos corresponden al conjunto de acciones para construir estructura que se efectuarán durante el análisis cuando se realizan las transferencias entre nodos de la ATN.

En otro intento de establecer un paralelismo la ficha MARC sería una conjunción de estructuras superficial y profunda, donde los contenidos de los campos corresponden a la superficial y el conjunto ordenado de etiquetas y códigos de subcampos MARC que aparecen en una ficha dada son la estructura profunda. Aquí deberemos entender el procesamiento descrito en la tesis como un proceso de traducción. En él se usa la tabla -que puede verse como ATN- para interpretar el lenguaje fuente, y nuestros operadores semánticos se utilizan para caracterizar al lenguaje objeto.

CONCLUSIONES

En esta tesis se presenta una solución al problema de recuperación bibliográfica desde un punto de vista descriptivo, debido a lo cual muchas de las afirmaciones están basadas en ejemplos particulares. Sin embargo, la solución descrita tiene intrínseca una metodología computacional para el manejo de información documental que incluye: la acomodación de un esquema estructural a los datos por analizar, el uso de un formato variable con identificadores y separadores de campos para almacenar la información, la aplicación del análisis léxico dirigido por una tabla en la que se representan autómatas conectados, y el diseño y programación de operadores semánticos.

El método de programación descrito permite obtener productos muy especializados de manera rápida, debido a que ya han sido programados tanto el algoritmo para construir la tabla que representa a los autómatas -a partir de simples descripciones de las etiquetas del esquema estructural-, como el algoritmo de utilización de esa tabla para efectuar el análisis léxico. Al abordar cada nueva aplicación de este método de programación, es necesario solamente programar los operadores semánticos. En la tesis se han ejemplificado algunos de los operadores semánticos más simples, sin embargo, con operadores más complejos, este mismo método ha sido utilizado para elaborar algunos de los productos requeridos por la BDCV como son: los juegos de tarjetas catalográficas con alta calidad de impresión junto con su

correspondiente juego de etiquetas para el marcaje del libro; la lista de obras catalogadas que está formada por los datos de la referencia bibliográfica y en la cual los descriptores (particularmente encabezamientos de materia) se utilizan para conformar un índice a las referencias; bibliografías especializadas; y registros MARC propiamente dichos los cuales se han importado al paquete MICRO CDS/ISIS.

La interrelación entre el sistema descrito y algún manejador de bases de datos permite trabajar con un alto grado de eficiencia. Por un lado, se aprovechan las ventajas de la recuperación selectiva que ofrece el manejador de bases de datos. Y, por otro lado, una vez hecha la selección de un subconjunto de unidades de información, es posible aplicarle alguno de los programas para emisión de productos especializados como los mencionados en el párrafo anterior.

Las herramientas aquí descritas, tanto la metodología como la implementación misma del sistema, poseen un alto grado de adaptabilidad. Cuando un investigador especializado obtiene nueva información que complementa a la ya existente pero que no había sido prevista en su esquema estructural, basta que defina el esquema estructural aumentado, y que, utilizando los criterios metodológicos descritos, agregue a las unidades de información pertinentes sus nuevos datos. A nivel del trabajo computacional definir el esquema estructural aumentado equivale a agregar las etiquetas faltantes a la descripción del esquema original, y aplicar los programas de construcción de la tabla en la que se

representan los autómatas y de codificación de los operadores semánticos. Los programas ya existentes pueden ser usados sin modificación.

La posibilidad de aumentar información, sobre todo resúmenes en lengua natural -probablemente mal estructurados dentro de sí mismos, aunque bien estructurados como componentes de las unidades de información completas- y manipularla en conjunto mediante las herramientas computacionales descritas en esta tesis, permite al investigador encontrar características que él representará mediante descriptores con objeto de dar a su información mejores formas de organización.

En el procedimiento de recuperación bibliográfica descrito en el último capítulo, los programas con los que se ejemplifica la metodología que ha sido objeto de esta tesis están efectuando la función de una operación generalizada de proyección de un manejador de bases de datos relacional. Es por esto que nos propusimos llamar a esta tesis "Un método computacional complementario a los manejadores de bases de datos: su aplicación en recuperación bibliográfica".

ANEXO A

PROGRAMAS Y ALGUNAS DE LAS TABLAS QUE UTILIZAN

Arbol.c

Lista de descripciones para reconocimiento de Marc
Tabla en la que se representa a los autómatas conectados

Semantic.c

Indices.c

Creapun.c

Recupera.c

```
/*XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*/
```

Programa: Arbol.c

Objetivo: Construir tablas que representan automatas simples o conectados.

Datos-de-entrada: Una lista ordenada de descripciones formadas por las etiquetas que reconocera el o los automatas que se van a representar asociadas a codigos de operadores semanticos con sus parametros. En el caso de automatas simples, cada descripcion esta formada por la etiqueta, una coma, y el operador con sus parametros; por ejemplo "FIN,5". En el caso de automatas conectados habra tantas descripciones para cada etiqueta como posibles cadenas se deseen reconocer mediante el automata conectado al automata principal. Los caracteres "<" que van a continuacion de la etiqueta se utilizan para crear el registro de conexion entre los automatas. Por ejemplo:

```
260><<a,#03000
260><<b,#03000
260><<c,#03000
```

Datos-de-salida: La tabla que representa el automata simple o los automatas conectados.

Informacion-practica: El programa se ejecuta de la siguiente manera:

```
arbol.exe <arch-des> <arch-sal>
```

```
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*/
```

```
#include <stdio.h>
#define LONDES 20 /* longitud maxima de una descripcion */
#define NUMDES 450 /* maximo numero de descripciones */
#define MAXNOD 1500 /* maximo numero de nodos en el arbol */
#define LOPTER 7 /* maxima longitud de los operadores terminales */
```

```
FILE *fopen(), *fpent, *fpal;
```

```
main (argc,argv) /* construye arboles */
```

```
int argc;
```

```
char *argv[];
```

```
{
```

```
char des[NUMDES][LONDES]; /* tabla con descripciones */
char arbol[MAXNOD][LOPTER+1]; /* tabla de operadores y caracteres */
int hijos[MAXNOD]; /* tabla de contadores de hijos */
int apun[MAXNOD]; /* tabla con apuntadores */
int par[NUMDES+1]; /* particiones que controlan */
int paraux[NUMDES+1]; /* particiones que se van creando */
int numpar; /* numero de particiones */
int c, l, n; /* del caracter en las descripciones;
```

```

                                contenido, horizontal y vertical */
int i, j, l;
int cd;                                /* indice horizontal en las descripciones */
int cn;                                /* apuntador al nodo actual */
int niv;                                /* indice vertical en las descripciones */
int p;                                  /* indice en las particiones */
int cnb;                                /* apuntador secuencial a hijo */
int apunp;                              /* temporal para crear apuntadores */
int fin;                                /* 0 cuando en la particion hay ceros */
char operc;                             /* operador de % = nodo de base */

/* LECTURA E IMPRESION DE LAS DESCRIPCIONES */

fpent=fopen(argv[1],"r");
fpsal=fopen(argv[2],"w");
for (i=0; i<NUMDES; ++i)
for (j=0; j<LONDES; ++j)
desl[i][j]='\0';

l=0;
n=0;
while (c=getc(fpent)) != EOF)
    if (c != '\n') { desl[l][n]=c; ++n; }
    else { ++n; l=0; }

/* CONSTRUCCION DE LA ARBURESCENCIA */

for (i=0; i<NUMDES; ++i) { par[i]=0; paraux[i]=0; }
for (i=0; i<MAXNOD; ++i)
{
    for (j=0; j<LOPTEN; ++j) arbol[i][j]='\0';
    hijos[i]=0; apun[i]=0;
}
i;
fin=1;
numpar=1;
par[numpar]=0;
niv=0;
n=0;
p=1;
cd=0;
for (j=1; j<=NUMDES; ++j) paraux[j]=1;
cnb=1;
arbol[i][1][1]='%';
arbol[i][1][2]='<<';
cn=2;
while (fin == 1)
{
    for (j=1; j<=numpar; ++j)
    {
        if (par[j] == 0) paraux[p]=0;
        else
        {
            while (n < par[j]-1)
            {
                if (des[niv][cd] == des[niv+1][cd]) ++paraux[p];
                else

```

```

    {
        if (des[niv][ed] == '<,>')
        {
            for (k=1; k<=LOFTER; ++k) arbol[en][k]=des[niv][k+ed];
            paraux[p]=0;
        }
        else
        {
            arbol[en][1]='*';
            arbol[en][2]=des[niv][ed];
            i;
            ++hijos[en];
            ++en;
            ++p;
        }
        i;
        ++niv;
        ++n;
    }
    if (des[niv][ed] == '<,>')
    {
        for (k=1; k<=LOFTER; ++k) arbol[en][k]=des[niv][k+ed];
        paraux[p]=0;
    }
    else
    {
        arbol[en][1]='*';
        arbol[en][2]=des[niv][ed];
    }
    ++hijos[en];
    ++en;
    ++knh;
    while (arbol[en][1] != '*' && en<en) ++en;
    i;
    ++p;
    ++niv;
    n=0;
}
niv=0;
++ed;
numpar = p-1;
fin=0;
for (k=1; k<=numpar; ++k)
if (paraux[k] != 0) fin=1;
for (j=1; j<=NUMDES; ++j) par[j]=paraux[j];
for (j=1; j<=NUMDES; ++j) paraux[j]=1;
p=1;
};

/* CALCULO DE LOS APUNTADES */

apun[1]=2;
aptmp=apun[1]+hijos[1];
for (j=2; j<=en; ++j)
if (arbol[j][1] == '*')
{
    apun[j] = aptmp;
}

```

```
arbol[c]
```

```
apunt[apunt[j]][hijos[j]];  
};  
  
/* IMPRESION DEL ARBOL */  
  
oporec1="Z";  
for (j=1; j<n; ++j)  
{  
    switch(arbol[j][1])  
    {  
        case '8':  
            if (arbol[j][2]!="1")  
                fprintf(fpсал, "%c%c%c%c%c", oporec1,  
                    apunt[j]/1000, (apunt[j]%1000)/100,  
                    (apunt[j]%100)/10, apunt[j]%10);  
            else  
                fprintf(fpсал, "%c%c%c%c%c%c%c", arbol[j][1], arbol[j][2], hijos[j]/10,  
                    hijos[j]%10, apunt[j]/1000, (apunt[j]%1000)/100,  
                    (apunt[j]%100)/10, apunt[j]%10);  
            break;  
        default:  
            k=1;  
            while (arbol[j][k]!="\0" && k<=LONGITER)  
                fprintf(fpсал, "%c", arbol[j][k++]);  
            break;  
    }  
    fprintf(fpсал, "\n");  
};  
};
```

020> <\$a,0	130> <\$a,0	400> <\$t,0	546> <\$a,0
035> <\$a,0	130> <\$d,0	400> <\$v,0	590> <\$a,0
040> <\$a,0	130> <\$f,0	400> <\$w,0	690> <\$a,0
040> <\$b,0	130> <\$g,0	400> <\$x,0	690> <\$b,0
040> <\$c,0	130> <\$h,0	410> <\$a,0	690> <\$c,0
040> <\$d,0	130> <\$k,0	410> <\$b,0	690> <\$d,0
041> <\$a,0	130> <\$l,0	410> <\$e,0	690> <\$f,0
043> <\$a,0	130> <\$n,0	410> <\$d,0	691> <\$a,0
042> <\$a,0	130> <\$p,0	410> <\$e,0	691> <\$b,0
050> <\$a,0	130> <\$s,0	410> <\$f,0	691> <\$k,0
050> <\$b,0	130> <\$t,0	410> <\$g,0	691> <\$y,0
080> <\$a,0	130> <\$w,0	410> <\$k,0	691> <\$z,0
082> <\$a,0	240> <\$a,0	410> <\$l,0	692> <\$a,0
092> <\$a,0	240> <\$d,0	410> <\$n,0	692> <\$b,0
092> <\$b,0	240> <\$f,0	410> <\$p,0	692> <\$c,0
098> <\$a,0	240> <\$g,0	410> <\$t,0	692> <\$d,0
099> <\$a,0	240> <\$h,0	410> <\$v,0	692> <\$e,0
099> <\$b,0	240> <\$k,0	410> <\$w,0	692> <\$f,0
100> <\$a,0	240> <\$l,0	410> <\$x,0	692> <\$g,0
100> <\$b,0	240> <\$m,0	411> <\$a,0	692> <\$h,0
100> <\$c,0	240> <\$n,0	411> <\$b,0	692> <\$k,0
100> <\$d,0	240> <\$p,0	411> <\$c,0	692> <\$l,0
100> <\$e,0	240> <\$r,0	411> <\$d,0	692> <\$m,0
100> <\$f,0	240> <\$s,0	411> <\$e,0	692> <\$n,0
100> <\$t,0	240> <\$w,0	411> <\$f,0	692> <\$o,0
100> <\$l,0	241> <\$a,0	411> <\$k,0	692> <\$p,0
100> <\$n,0	241> <\$b,0	411> <\$l,0	692> <\$q,0
100> <\$p,0	241> <\$c,0	411> <\$n,0	692> <\$r,0
100> <\$q,0	245> <\$a,0	411> <\$p,0	692> <\$s,0
100> <\$t,0	245> <\$b,0	411> <\$q,0	692> <\$t,0
100> <\$w,0	245> <\$c,0	411> <\$t,0	692> <\$w,0
110> <\$a,0	245> <\$h,0	411> <\$v,0	692> <\$x,0
110> <\$b,0	245> <\$n,0	411> <\$w,0	692> <\$y,0
110> <\$c,0	245> <\$p,0	411> <\$x,0	692> <\$z,0
110> <\$d,0	250> <\$a,0	440> <\$a,0	693> <\$a,0
110> <\$e,0	250> <\$b,0	440> <\$n,0	693> <\$b,0
110> <\$f,0	260> <\$a,0	440> <\$p,0	693> <\$c,0
110> <\$g,0	260> <\$b,0	440> <\$v,0	693> <\$d,0
110> <\$k,0	260> <\$c,0	440> <\$x,0	693> <\$e,0
110> <\$l,0	300> <\$a,0	490> <\$a,0	693> <\$f,0
110> <\$n,0	300> <\$b,0	490> <\$l,0	693> <\$g,0
110> <\$p,0	300> <\$c,0	490> <\$n,0	693> <\$h,0
110> <\$t,0	300> <\$e,0	490> <\$v,0	693> <\$k,0
110> <\$w,0	300> <\$f,0	490> <\$x,0	693> <\$l,0
111> <\$a,0	300> <\$g,0	500> <\$a,0	693> <\$m,0
111> <\$b,0	400> <\$a,0	501> <\$a,0	693> <\$n,0
111> <\$c,0	400> <\$b,0	502> <\$a,0	693> <\$o,0
111> <\$d,0	400> <\$c,0	502> <\$b,0	693> <\$p,0
111> <\$e,0	400> <\$d,0	503> <\$a,0	693> <\$r,0
111> <\$f,0	400> <\$e,0	504> <\$a,0	693> <\$s,0
111> <\$t,0	400> <\$f,0	505> <\$a,0	693> <\$t,0
111> <\$l,0	400> <\$k,0	507> <\$a,0	693> <\$w,0
111> <\$n,0	400> <\$l,0	520> <\$a,0	693> <\$x,0
111> <\$p,0	400> <\$n,0	533> <\$a,0	693> <\$y,0
111> <\$q,0	400> <\$p,0	533> <\$b,0	693> <\$z,0
111> <\$w,0	400> <\$q,0	533> <\$c,0	694> <\$a,0

694>1<#b,0	710>1<#i,0	810>1<#a,0	F1>,0
694>1<#c,0	710>1<#s,0	810>1<#b,0	F2>,0
694>1<#d,0	710>1<#t,0	810>1<#c,0	FCAP>,0
694>1<#e,0	710>1<#u,0	810>1<#d,0	FIN>,\$
694>1<#f,0	711>1<#a,0	810>1<#e,0	ESTUB>,0
694>1<#g,0	711>1<#b,0	810>1<#f,0	IDIOMA>,0
694>1<#h,0	711>1<#c,0	810>1<#g,0	ILUS>,0
694>1<#i,0	711>1<#d,0	810>1<#h,0	IND>,0
694>1<#j,0	711>1<#e,0	810>1<#i,0	LEST>,0
694>1<#k,0	711>1<#f,0	810>1<#l,0	LTIP>,0
694>1<#l,0	711>1<#g,0	810>1<#m,0	LNBIB>,0
694>1<#m,0	711>1<#h,0	810>1<#n,0	LNDES>,0
694>1<#n,0	711>1<#i,0	810>1<#p,0	LFOR>,0
694>1<#o,0	711>1<#j,0	810>1<#r,0	PAIS>,0
694>1<#p,0	711>1<#k,0	810>1<#s,0	PUBOP>,0
694>1<#q,0	711>1<#l,0	810>1<#t,0	REPROD>,0
694>1<#r,0	711>1<#m,0	810>1<#v,0	
694>1<#s,0	711>1<#n,0	811>1<#a,0	
694>1<#t,0	711>1<#o,0	811>1<#b,0	
694>1<#u,0	730>1<#a,0	811>1<#c,0	
694>1<#v,0	730>1<#d,0	811>1<#d,0	
694>1<#w,0	730>1<#f,0	811>1<#e,0	
694>1<#x,0	730>1<#g,0	811>1<#f,0	
694>1<#y,0	730>1<#h,0	811>1<#g,0	
694>1<#z,0	730>1<#i,0	811>1<#h,0	
700>1<#a,0	730>1<#k,0	811>1<#i,0	
700>1<#b,0	730>1<#l,0	811>1<#k,0	
700>1<#c,0	730>1<#m,0	811>1<#l,0	
700>1<#d,0	730>1<#n,0	811>1<#n,0	
700>1<#e,0	730>1<#o,0	811>1<#p,0	
700>1<#f,0	730>1<#p,0	811>1<#q,0	
700>1<#g,0	730>1<#r,0	811>1<#t,0	
700>1<#h,0	730>1<#s,0	811>1<#v,0	
700>1<#i,0	730>1<#t,0	830>1<#a,0	
700>1<#j,0	730>1<#u,0	830>1<#d,0	
700>1<#k,0	740>1<#a,0	830>1<#f,0	
700>1<#l,0	740>1<#b,0	830>1<#g,0	
700>1<#m,0	740>1<#c,0	830>1<#i,0	
700>1<#n,0	740>1<#h,0	830>1<#k,0	
700>1<#o,0	740>1<#n,0	830>1<#l,0	
700>1<#p,0	740>1<#p,0	830>1<#m,0	
700>1<#q,0	800>1<#a,0	830>1<#n,0	
700>1<#r,0	800>1<#b,0	830>1<#o,0	
700>1<#s,0	800>1<#c,0	830>1<#p,0	
700>1<#t,0	800>1<#d,0	830>1<#r,0	
700>1<#u,0	800>1<#e,0	830>1<#s,0	
700>1<#v,0	800>1<#f,0	830>1<#t,0	
700>1<#w,0	800>1<#g,0	830>1<#v,0	
700>1<#x,0	800>1<#h,0	840>1<#a,0	
700>1<#y,0	800>1<#i,0	840>1<#v,0	
700>1<#z,0	800>1<#j,0	901>1<#a,0	
710>1<#a,0	800>1<#k,0	902>1<#a,0	
710>1<#b,0	800>1<#l,0	AF>,0	
710>1<#c,0	800>1<#m,0	B10>,0	
710>1<#d,0	800>1<#n,0	CONF>,0	
710>1<#e,0	800>1<#o,0	CONT>,0	
710>1<#f,0	800>1<#p,0	ECONM>,0	
710>1<#g,0	800>1<#q,0		
710>1<#h,0	800>1<#r,0		
710>1<#i,0	800>1<#s,0		
710>1<#j,0	800>1<#t,0		
710>1<#k,0	800>1<#u,0		
710>1<#l,0			
710>1<#m,0			
710>1<#n,0			
710>1<#o,0			
710>1<#p,0			

1:	*K	19	2	52:	*O	2	126	103:	*4	1	179	154:	*>	1	231
2:	*O	6	21	53:	*P	1	128	104:	*5	1	180	155:	*>	1	232
3:	*1	3	27	54:	*1	1	129	105:	*7	1	181	156:	*>	1	233
4:	*P	3	30	55:	*O	1	130	106:	*O	1	182	157:	*>	1	234
5:	*3	1	33	56:	*D	1	131	107:	*3	1	183	158:	*>	1	235
6:	*4	4	34	57:	*1	1	132	108:	*6	1	184	159:	*>	1	236
7:	*5	5	38	58:	*2	1	133	109:	*O	1	185	160:	*>	1	237
8:	*6	1	43	59:	*O	1	134	110:	*O	1	186	161:	*>	1	238
9:	*7	4	44	60:	*1	1	135	111:	*1	1	187	162:	*>	1	239
10:	*8	4	48	61:	*S	1	136	112:	*2	1	188	163:	*>	1	240
11:	*9	1	52	62:	*D	1	137	113:	*3	1	189	164:	*>	1	241
12:	*A	1	53	63:	*L	1	138	114:	*4	1	190	165:	*>	1	242
13:	*B	1	54	64:	*N	1	139	115:	*5	1	191	166:	*>	1	243
14:	*C	1	55	65:	*I	1	140	116:	*O	1	192	167:	*>	1	244
15:	*E	1	56	66:	*T	1	141	117:	*O	1	193	168:	*>	1	245
16:	*F	5	57	67:	*N	2	142	118:	*1	1	194	169:	*>	1	246
17:	*1	3	62	68:	*F	1	144	119:	*O	1	195	170:	*>	1	247
18:	*L	4	65	69:	*A	1	145	120:	*O	1	196	171:	*>	1	248
19:	*P	2	69	70:	*U	1	146	121:	*O	1	197	172:	*>	1	249
20:	*R	1	71	71:	*E	1	147	122:	*O	1	198	173:	*>	1	250
21:	*2	1	72	72:	*O	1	148	123:	*1	1	199	174:	*>	1	251
22:	*3	1	73	73:	*5	1	149	124:	*O	1	200	175:	*>	1	252
23:	*4	4	74	74:	*O	1	150	125:	*O	1	201	176:	*>	1	253
24:	*5	1	76	75:	*1	1	151	126:	*1	1	202	177:	*>	1	254
25:	*8	2	79	76:	*3	1	152	127:	*2	1	203	178:	*>	1	255
26:	*9	3	81	77:	*9	1	153	128:	*>	1	204	179:	*>	1	256
27:	*O	1	84	78:	*O	1	154	129:	*D	1	205	180:	*>	1	257
28:	*1	2	85	79:	*O	1	155	130:	*N	2	206	181:	*>	1	258
29:	*3	1	87	80:	*2	1	156	131:	*C	1	208	182:	*>	1	259
30:	*4	3	88	81:	*2	1	157	132:	*>	1	209	183:	*>	1	260
31:	*5	1	91	82:	*8	1	158	133:	*>	1	210	184:	*>	1	261
32:	*6	1	92	83:	*9	1	159	134:	*A	1	211	185:	*>	1	262
33:	*O	1	93	84:	*O	1	160	135:	*N	1	212	186:	*>	1	263
34:	*O	1	94	85:	*O	1	161	136:	*P	1	213	187:	*>	1	264
35:	*1	2	95	86:	*1	1	162	137:	*1	1	214	188:	*>	1	265
36:	*4	1	97	87:	*O	1	163	138:	*U	1	215	189:	*>	1	266
37:	*9	1	98	88:	*O	1	164	139:	*D	1	216	190:	*>	1	267
38:	*O	7	99	89:	*1	1	165	140:	*S	1	217	191:	*>	1	268
39:	*2	1	106	90:	*5	1	166	141:	*1	1	218	192:	*>	1	269
40:	*3	1	107	91:	*O	1	167	142:	*B	1	219	193:	*>	1	270
41:	*4	1	108	92:	*O	1	168	143:	*D	1	220	194:	*>	1	271
42:	*9	1	109	93:	*O	1	169	144:	*O	1	221	195:	*>	1	272
43:	*9	6	110	94:	*O	1	170	145:	*1	1	222	196:	*>	1	273
44:	*O	1	116	95:	*O	1	171	146:	*B	1	223	197:	*>	1	274
45:	*1	2	117	96:	*1	1	172	147:	*P	1	224	198:	*>	1	275
46:	*3	1	119	97:	*O	1	173	148:	*>	1	225	199:	*>	1	276
47:	*4	1	120	98:	*O	1	174	149:	*>	1	226	200:	*>	1	277
48:	*O	1	121	99:	*O	1	175	150:	*>	1	227	201:	*>	1	278
49:	*1	2	122	100:	*1	1	176	151:	*>	1	228	202:	*>	1	279
50:	*3	1	124	101:	*2	1	177	152:	*>	1	229	203:	*>	1	280
51:	*4	1	125	102:	*3	1	178	153:	*>	1	230	204:	0		

205:	*	1	281	256:	%	330	307:	**	1	378	358:	0			
206:	*I	1	282	257:	%	331	308:	**	1	379	359:	*	1	427	
207:	*I	1	283	258:	%	332	309:	**	1	380	360:	*A	1	428	
208:	*D	1	284	259:	%	333	310:	**	1	381	361:	0			
209:	0			260:	%	334	311:	**	1	382	362:	0			
210:	0			261:	%	335	312:	**	1	383	363:	0			
211:	*P	1	285	262:	%	336	313:	**	1	384	364:	*	1	429	
212:	*D	1	286	263:	%	337	314:	**	1	385	365:	*	1	430	
213:	*H	1	287	264:	%	338	315:	**	1	386	366:	0			
214:	*D	1	288	265:	%	339	316:	**	1	387	367:	0			
215:	*S	1	289	266:	%	340	317:	**	1	388	368:	*	1	431	
216:	*C	1	290	267:	%	341	318:	**	1	389	369:	*D	1	432	
217:	*I	1	291	268:	%	342	319:	**	1	390	370:	**	1	433	
218:	*P	1	292	269:	%	343	320:	**	1	391	371:	**	1	434	
219:	*I	1	293	270:	%	344	321:	**	1	392	372:	**	4	435	
220:	*E	1	294	271:	%	345	322:	**	1	393	373:	**	1	439	
221:	*R	1	295	272:	%	346	323:	**	1	394	374:	**	1	440	
222:	*S	1	296	273:	%	347	324:	**	1	395	375:	**	1	441	
223:	*D	1	297	274:	%	348	325:	**	1	396	376:	**	2	442	
224:	*R	1	298	275:	%	349	326:	**	1	397	377:	**	1	444	
225:	%		299	276:	%	350	327:	**	1	398	378:	**	1	445	
226:	%		300	277:	%	351	328:	**	1	399	379:	**	2	446	
227:	%		301	278:	%	352	329:	**	1	400	380:	**	1	448	
228:	%		302	279:	%	353	330:	**	1	401	381:	**	2	449	
229:	%		303	280:	%	354	331:	**	1	402	382:	**	13	451	
230:	%		304	281:	0		332:	**	1	403	383:	**	13	454	
231:	%		305	282:	*	1	355	333:	**	1	404	384:	**	12	477
232:	%		306	283:	*	1	356	334:	**	1	405	385:	**	12	489
233:	%		307	284:	*N	1	357	335:	**	1	406	386:	**	13	501
234:	%		308	285:	*	1	358	336:	**	1	407	387:	**	3	514
235:	%		309	286:	*			337:	**	1	408	388:	**	6	517
236:	%		310	287:	*B	1	359	338:	**	1	409	389:	**	2	523
237:	%		311	288:	*M	1	360	339:	**	1	410	390:	**	3	525
238:	%		312	289:	*	1	361	340:	**	1	411	391:	**	6	528
239:	%		313	290:	0			341:	**	1	412	392:	**	15	534
240:	%		314	291:	*	1	362	342:	**	1	413	393:	**	15	549
241:	%		315	292:	*	1	363	343:	**	1	414	394:	**	15	564
242:	%		316	293:	*B	1	364	344:	**	1	415	395:	**	5	579
243:	%		317	294:	*C	1	365	345:	**	1	416	396:	**	5	584
244:	%		318	295:	*	1	366	346:	**	1	417	397:	**	1	599
245:	%		319	296:	*	1	367	347:	**	1	418	398:	**	1	599
246:	%		320	297:	*I	1	368	348:	**	1	419	399:	**	2	591
247:	%		321	298:	*D	1	369	349:	**	1	420	400:	**	1	593
248:	%		322	299:	**	1	370	350:	**	1	421	401:	**	1	594
249:	%		323	300:	**	1	371	351:	**	1	422	402:	**	1	595
250:	%		324	301:	**	1	372	352:	**	1	423	403:	**	1	596
251:	%		325	302:	**	1	373	353:	**	1	424	404:	**	1	597
252:	%		326	303:	**	1	374	354:	**	1	425	405:	**	3	598
253:	%		327	304:	**	1	375	355:	0		406:	**	1	601	
254:	%		328	305:	**	1	376	356:	0		407:	**	1	602	
255:	%		329	306:	**	1	377	357:	*M	1	426	408:	**	5	603

409:	af	5	603	460:	ap	1	850	511:	ar	1	901	563:	av	1	952
410:	ah	27	611	461:	aq	1	851	512:	as	1	902	564:	aw	1	953
411:	ai	21	625	462:	at	1	852	513:	aw	1	903	564:	ax	1	954
412:	aj	13	656	463:	aw	1	853	514:	ax	1	904	565:	ay	1	955
413:	ak	14	669	464:	ax	1	854	515:	ay	1	905	566:	az	1	956
414:	al	16	683	465:	ay	1	855	516:	az	1	906	567:	ba	1	957
415:	am	18	699	466:	az	1	856	517:	ba	1	907	568:	bb	1	958
416:	an	15	717	467:	ba	1	857	518:	bb	1	908	569:	bc	1	959
417:	ao	15	732	468:	bb	1	858	519:	bc	1	909	570:	bd	1	960
418:	ap	6	747	469:	bc	1	859	520:	bd	1	910	571:	be	1	961
419:	aq	16	753	470:	bd	1	860	521:	be	1	911	572:	bf	1	962
420:	ar	17	769	471:	be	1	861	522:	bf	1	912	573:	bg	1	963
421:	as	15	786	472:	bf	1	862	523:	bg	1	913	574:	bh	1	964
422:	at	15	801	473:	bg	1	863	524:	bh	1	914	575:	bi	1	965
423:	au	2	816	474:	bh	1	864	525:	bi	1	915	576:	bj	1	966
424:	av	1	818	475:	bi	1	865	526:	bj	1	916	577:	bk	1	967
425:	aw	1	819	476:	bj	1	866	527:	bk	1	917	578:	bl	1	968
426:	ax	1	820	477:	bk	1	867	528:	bl	1	918	579:	bm	1	969
427:	0			478:	bm	1	868	529:	bn	1	919	580:	bn	1	970
428:	ay	1	821	479:	bn	1	869	530:	bo	1	920	581:	bp	1	971
429:	0			480:	bo	1	870	531:	bp	1	921	582:	bq	1	972
430:	0			481:	br	1	871	532:	br	1	922	583:	br	1	973
431:	0			482:	bs	1	872	533:	bs	1	923	584:	bs	1	974
432:	az	1	822	483:	bt	1	873	534:	bt	1	924	585:	bt	1	975
433:	ba	1	823	484:	bu	1	874	535:	bu	1	925	586:	bu	1	976
434:	ba	1	824	485:	bv	1	875	536:	bv	1	926	587:	bv	1	977
435:	ba	1	825	486:	bw	1	876	537:	bw	1	927	588:	bw	1	978
436:	bb	1	826	487:	bx	1	877	538:	bx	1	928	589:	bx	1	979
437:	bc	1	827	488:	by	1	878	539:	by	1	929	590:	by	1	980
438:	bd	1	828	489:	bz	1	879	540:	bz	1	930	591:	ca	1	981
439:	ba	1	829	490:	ca	1	880	541:	ca	1	931	592:	cb	1	982
440:	ba	1	830	491:	cb	1	881	542:	cb	1	932	593:	ca	1	983
441:	ba	1	831	492:	cc	1	882	543:	cc	1	933	594:	ca	1	984
442:	ca	1	832	493:	cd	1	883	544:	cc	1	934	595:	ca	1	985
443:	cb	1	833	494:	ce	1	884	545:	cd	1	935	596:	ca	1	986
444:	ca	1	834	495:	cf	1	885	546:	cd	1	936	597:	ca	1	987
445:	ca	1	835	496:	cg	1	886	547:	ce	1	937	598:	ca	1	988
446:	ca	1	836	497:	ch	1	887	548:	ce	1	938	599:	cb	1	989
447:	cb	1	837	498:	ci	1	888	549:	ca	1	939	600:	cc	1	990
448:	ca	1	838	499:	cl	1	889	550:	cb	1	940	601:	ca	1	991
449:	ca	1	839	500:	cm	1	890	551:	cc	1	941	602:	ca	1	992
450:	cb	1	840	501:	cn	1	891	552:	cd	1	942	603:	ca	1	993
451:	ca	1	841	502:	co	1	892	553:	ce	1	943	604:	cb	1	994
452:	cb	1	842	503:	cp	1	893	554:	cf	1	944	605:	ca	1	995
453:	cc	1	843	504:	cq	1	894	555:	cg	1	945	606:	cy	1	996
454:	cd	1	844	505:	cr	1	895	556:	ch	1	946	607:	ca	1	997
455:	ce	1	845	506:	cs	1	896	557:	ci	1	947	608:	ca	1	998
456:	cf	1	846	507:	ct	1	897	558:	cn	1	948	609:	cb	1	999
457:	ck	1	847	508:	cu	1	898	559:	cp	1	949	610:	ca	1	1000
458:	cl	1	848	509:	cv	1	899	560:	ct	1	950	611:	cy	1	1001
459:	cn	1	849	510:	cp	1	900	561:	cv	1	951	612:	ca	1	1002

613:	sa	1	1003	664:	st	1	1054	715:	xt	1	1105	766:	xa	1	1156
614:	sb	1	1004	665:	sw	1	1055	716:	xw	1	1106	767:	xt	1	1157
615:	sc	1	1005	666:	sx	1	1056	717:	xa	1	1107	768:	xy	1	1158
616:	sd	1	1006	667:	sy	1	1057	718:	xb	1	1108	769:	za	1	1159
617:	se	1	1007	668:	sz	1	1058	719:	xc	1	1109	770:	zb	1	1160
618:	sf	1	1008	669:	sa	1	1059	720:	xd	1	1110	771:	zc	1	1161
619:	sg	1	1009	670:	sd	1	1060	721:	xe	1	1111	772:	zd	1	1162
620:	sh	1	1010	671:	se	1	1061	722:	xf	1	1112	773:	ze	1	1163
621:	sk	1	1011	672:	sh	1	1062	723:	xh	1	1113	774:	zf	1	1164
622:	sl	1	1012	673:	sl	1	1063	724:	xi	1	1114	775:	zg	1	1165
623:	sm	1	1013	674:	sl	1	1064	725:	xl	1	1115	776:	zh	1	1166
624:	sn	1	1014	675:	sm	1	1065	726:	xn	1	1116	777:	zi	1	1167
625:	so	1	1015	676:	sp	1	1066	727:	xp	1	1117	778:	xj	1	1168
626:	sp	1	1016	677:	sq	1	1067	728:	xq	1	1118	779:	xm	1	1169
627:	sq	1	1017	678:	sl	1	1068	729:	xs	1	1119	780:	xn	1	1170
628:	sr	1	1018	679:	sw	1	1069	730:	xt	1	1120	781:	xp	1	1171
629:	ss	1	1019	680:	sa	1	1070	731:	xw	1	1121	782:	xr	1	1172
630:	st	1	1020	681:	sy	1	1071	732:	xa	1	1122	783:	xs	1	1173
631:	sw	1	1021	682:	sz	1	1072	733:	xd	1	1123	784:	xt	1	1174
632:	sa	1	1022	683:	sa	1	1073	734:	xi	1	1124	785:	xy	1	1175
633:	sy	1	1023	684:	sb	1	1074	735:	xg	1	1125	786:	xa	1	1176
634:	sz	1	1024	685:	sc	1	1075	736:	xh	1	1126	787:	xb	1	1177
635:	sa	1	1025	686:	sd	1	1076	737:	xi	1	1127	788:	xc	1	1178
636:	sb	1	1026	687:	se	1	1077	738:	xl	1	1128	789:	xd	1	1179
637:	sc	1	1027	688:	sf	1	1078	739:	xm	1	1129	790:	xe	1	1180
638:	sd	1	1028	689:	sh	1	1079	740:	xn	1	1130	791:	xf	1	1181
639:	se	1	1029	690:	sl	1	1080	741:	so	1	1131	792:	xg	1	1182
640:	sl	1	1030	691:	sl	1	1081	742:	sp	1	1132	793:	xh	1	1183
641:	sg	1	1031	692:	sm	1	1082	743:	sr	1	1133	794:	xi	1	1184
642:	sh	1	1032	693:	sn	1	1083	744:	sa	1	1134	795:	xj	1	1185
643:	sk	1	1033	694:	so	1	1084	745:	st	1	1135	796:	xk	1	1186
644:	sl	1	1034	695:	sp	1	1085	746:	sw	1	1136	797:	xl	1	1187
645:	sm	1	1035	696:	sq	1	1086	747:	sa	1	1137	798:	xm	1	1188
646:	sn	1	1036	697:	st	1	1087	748:	sb	1	1138	799:	xt	1	1189
647:	so	1	1037	698:	sw	1	1088	749:	sc	1	1139	800:	xy	1	1190
648:	sp	1	1038	699:	sa	1	1089	750:	sh	1	1140	801:	za	1	1191
649:	sr	1	1039	700:	sb	1	1090	751:	sn	1	1141	802:	zd	1	1192
650:	ss	1	1040	701:	sc	1	1091	752:	sp	1	1142	803:	ze	1	1193
651:	st	1	1041	702:	sd	1	1092	753:	sa	1	1143	804:	zf	1	1194
652:	sw	1	1042	703:	se	1	1093	754:	sb	1	1144	805:	zh	1	1195
653:	sa	1	1043	704:	sf	1	1094	755:	sc	1	1145	806:	zi	1	1196
654:	sy	1	1044	705:	sg	1	1095	756:	sd	1	1146	807:	xj	1	1197
655:	sz	1	1045	706:	sh	1	1096	757:	se	1	1147	808:	xm	1	1198
656:	sa	1	1046	707:	sl	1	1097	758:	sf	1	1148	809:	xn	1	1199
657:	sb	1	1047	708:	sl	1	1098	759:	sh	1	1149	810:	zo	1	1200
658:	sc	1	1048	709:	sm	1	1099	760:	sk	1	1150	811:	zp	1	1201
659:	sd	1	1049	710:	sn	1	1100	761:	sl	1	1151	812:	zr	1	1202
660:	se	1	1050	711:	so	1	1101	762:	sm	1	1152	813:	zs	1	1203
661:	sg	1	1051	712:	sp	1	1102	763:	sn	1	1153	814:	zt	1	1204
662:	sh	1	1052	713:	sr	1	1103	764:	sp	1	1154	815:	zv	1	1205
663:	sq	1	1053	714:	ss	1	1104	765:	sq	1	1155	816:	za	1	1206

817: 0	868: 0	919: 0	970: 0
818: 0	869: 0	920: 0	971: 0
819: 0	870: 0	921: 0	972: 0
820: 0	871: 0	922: 0	973: 0
821: 0	872: 0	923: 0	974: 0
822: 0	873: 0	924: 0	975: 0
823: 0	874: 0	925: 0	976: 0
824: 0	875: 0	926: 0	977: 0
825: 0	876: 0	927: 0	978: 0
826: 0	877: 0	928: 0	979: 0
827: 0	878: 0	929: 0	980: 0
828: 0	879: 0	930: 0	981: 0
829: 0	880: 0	931: 0	982: 0
830: 0	881: 0	932: 0	983: 0
831: 0	882: 0	933: 0	984: 0
832: 0	883: 0	934: 0	985: 0
833: 0	884: 0	935: 0	986: 0
834: 0	885: 0	936: 0	987: 0
835: 0	886: 0	937: 0	988: 0
836: 0	887: 0	938: 0	989: 0
837: 0	888: 0	939: 0	990: 0
838: 0	889: 0	940: 0	991: 0
839: 0	890: 0	941: 0	992: 0
840: 0	891: 0	942: 0	993: 0
841: 0	892: 0	943: 0	994: 0
842: 0	893: 0	944: 0	995: 0
843: 0	894: 0	945: 0	996: 0
844: 0	895: 0	946: 0	997: 0
845: 0	896: 0	947: 0	998: 0
846: 0	897: 0	948: 0	999: 0
847: 0	898: 0	949: 0	1000: 0
848: 0	899: 0	950: 0	1001: 0
849: 0	900: 0	951: 0	1002: 0
850: 0	901: 0	952: 0	1003: 0
851: 0	902: 0	953: 0	1004: 0
852: 0	903: 0	954: 0	1005: 0
853: 0	904: 0	955: 0	1006: 0
854: 0	905: 0	956: 0	1007: 0
855: 0	906: 0	957: 0	1008: 0
856: 0	907: 0	958: 0	1009: 0
857: 0	908: 0	959: 0	1010: 0
858: 0	909: 0	960: 0	1011: 0
859: 0	910: 0	961: 0	1012: 0
860: 0	911: 0	962: 0	1013: 0
861: 0	912: 0	963: 0	1014: 0
862: 0	913: 0	964: 0	1015: 0
863: 0	914: 0	965: 0	1016: 0
864: 0	915: 0	966: 0	1017: 0
865: 0	916: 0	967: 0	1018: 0
866: 0	917: 0	968: 0	1019: 0
867: 0	918: 0	969: 0	1020: 0

1021: 0	1072: 0	1123: 0	1174: 0
1022: 0	1073: 0	1124: 0	1175: 0
1023: 0	1074: 0	1125: 0	1176: 0
1024: 0	1075: 0	1126: 0	1177: 0
1025: 0	1076: 0	1127: 0	1178: 0
1026: 0	1077: 0	1128: 0	1179: 0
1027: 0	1078: 0	1129: 0	1180: 0
1028: 0	1079: 0	1130: 0	1181: 0
1029: 0	1080: 0	1131: 0	1182: 0
1030: 0	1081: 0	1132: 0	1183: 0
1031: 0	1082: 0	1133: 0	1184: 0
1032: 0	1083: 0	1134: 0	1185: 0
1033: 0	1084: 0	1135: 0	1186: 0
1034: 0	1085: 0	1136: 0	1187: 0
1035: 0	1086: 0	1137: 0	1188: 0
1036: 0	1087: 0	1138: 0	1189: 0
1037: 0	1088: 0	1139: 0	1190: 0
1038: 0	1089: 0	1140: 0	1191: 0
1039: 0	1090: 0	1141: 0	1192: 0
1040: 0	1091: 0	1142: 0	1193: 0
1041: 0	1092: 0	1143: 0	1194: 0
1042: 0	1093: 0	1144: 0	1195: 0
1043: 0	1094: 0	1145: 0	1196: 0
1044: 0	1095: 0	1146: 0	1197: 0
1045: 0	1096: 0	1147: 0	1198: 0
1046: 0	1097: 0	1148: 0	1199: 0
1047: 0	1098: 0	1149: 0	1200: 0
1048: 0	1099: 0	1150: 0	1201: 0
1049: 0	1100: 0	1151: 0	1202: 0
1050: 0	1101: 0	1152: 0	1203: 0
1051: 0	1102: 0	1153: 0	1204: 0
1052: 0	1103: 0	1154: 0	1205: 0
1053: 0	1104: 0	1155: 0	1206: 0
1054: 0	1105: 0	1156: 0	1207: 0
1055: 0	1106: 0	1157: 0	1208: 0
1056: 0	1107: 0	1158: 0	1209: 0
1057: 0	1108: 0	1159: 0	
1058: 0	1109: 0	1160: 0	
1059: 0	1110: 0	1161: 0	
1060: 0	1111: 0	1162: 0	
1061: 0	1112: 0	1163: 0	
1062: 0	1113: 0	1164: 0	
1063: 0	1114: 0	1165: 0	
1064: 0	1115: 0	1166: 0	
1065: 0	1116: 0	1167: 0	
1066: 0	1117: 0	1168: 0	
1067: 0	1118: 0	1169: 0	
1068: 0	1119: 0	1170: 0	
1069: 0	1120: 0	1171: 0	
1070: 0	1121: 0	1172: 0	
1071: 0	1122: 0	1173: 0	


```
#include <stdio.h>
#define MAXNOD 1300
```

```
FILE *fopen(), *fp, *fpd, *fpas;
```

```
struct nodo l
    char arbol[2]; /* operadores y caracteres por comparar */
    int hijos; /* numero de hijos */
    int apun; /* apuntador a primer hijo */
    char desc[30]; /* para guardar lo que sobra de la desc. */
    int ar[MAXNOD];
```

```
char d[135]; /* descripcion que leo */
int cn; /* nodo actual */
int hijo; /* numero de hijos actual */
int base; /* base del subarbol */
int c;
int ap; /* apuntador a descripcion en uso */
int fin, sigue;
int i;
```

```
main(argc, argv)
```

```
int argc;
```

```
char *argv[];
```

```
{
    if (argc!=3)
    {
        printf("faltan o sobran nombres de archivos\n");
        printf("dar el nombre de la descripcion y del arbol a generar\n");
    }
    else
```

```
{
```

```
    if((fpd=fopen(++argv, "r"))==NULL)
        printf("no existe ese archivo para la descripcion\n");
```

```
    else
```

```
    {
        fpas=fopen(++argv, "w");
        fp=fopen("arbolpuro", "r");
```

```
        for(i=0; i<MAXNOD; ++i)
```

```
            ar[i].arbol[0]=' ';
```

```
/* ===== */
/* leemos el arbol puro */
/* ===== */
```

```
cn=1;
```

```
while (fscanf(fp, "%1s", &c)!=EOF)
```

```
{
```

```
    switch (c)
```

```
    {
```

```
        case '*':
```



```

        ar[ten].arbol[0] = e;
        fscanf(fp, "%c%2d%2d",
                &ar[ten].arbol[1], &ar[ten].hijos, &ar[ten].apun);
        break;
    case '4':
    case '0':
        ar[ten].arbol[0] = e;
        break;
    case '%':
        ar[ten].arbol[0] = e;
        fscanf(fp, "%d", &ar[ten].apun);
        break;
} /* del switch */

+ten;
} /* de while */

/*
/*
/*      leemos la descripción y corregimos el arbol
/*
/*

fin=1;
ca='<<';
ca=1;
hijo=ar[1].hijos;
while((fscanf(fp, "%s", &d1) != EOF) && fin)
{
    ap=(-1);
    sigue=1;
    if (d1[0] == '(')
    {
        ca=1;
        ca='<<';
        hijo=ar[1].hijos;
    }
    else
        while (sigue) /* sigue mientras estamos en etiqueta (disrun) */
            switch (ar[ten].arbol[0])
            {
                case '*': disyun();
                    break;
                case '%': subarbol();
                    break;
                case '0': copia();
                    break;
            }; /* del switch */
} /* del while */

/*
/*

```

```

/*          escribimos el arbol listo
*/
    en-1;
    while ((c=ar[en].arbol[0])!= '\0')
        i
        switch (c)
        {
            case '^':
                fprintf(fpas, "%c%c%02d%04d\n",
                    ar[en].arbol[0], ar[en].arbol[1], ar[en].hijos, ar[en].apun);
                break;
            case '$':
            case '0':
                fprintf(fpas, "%c\n", ar[en].arbol[0]);
                break;
            case '%':
                fprintf(fpas, "%c%04d%z\n", ar[en].arbol[0], ar[en].apun,
                    ar[en].desc);
                break;
            default:
                fprintf(fpas, "%c\n", ar[en].desc);
        } /* del switch */

        ++en;
    } /* de while */
/* for (i=1; i<MAXNOD; ++i)
    printf("=%z=\n", ar[i].desc);
*/

    close(fpd);
    close(fp);
    close(fpas);
}

}

/*-----
/*          operadores
/*-----

/*-----*/
/*          DISYUNCIÓN
/*-----*/
disyun()
{
    while (hijo>0)
        i
        if ((c=='\n') || (c==EOF))
            i
            printf("error en la descripción término mal \n");
}

```

```

        sigue=0;
        fin=0;
        return;
    }
    else
    {
        if (c==ar[en].arbol[1]) /* caracter actual igual a caracter del nodo */
        {
            hijo=ar[en].hijos;
            en=ar[en].apun;
            c=d1[d1[ap]]; /* obtengo el sig. car. de la desc. */
            return;
        }
        else
        {
            --hijo;
            if (hijo==0)
            {
                printf("error en desc. termino mal c=%c en=%d\n",c,en);
                sigue=0;
                fin=0;
                return;
            }
        }
    }
} /* del while */
} /* de disyun() */ /* ojo nunca sale por aqui */

/*******/
/*      subarbol      */
/*******/
subarbol()
{
    sigue=0;
    i=0;
    while((ar[en].desc[i++]=d1[ap+i])!='\0')
        ; /* copiamos el resto de la linea */
    base=ar[en].apun;
    c='<<';
    en=ar[en].apun;
    hijo=ar[en].hijos;
    return;
}

/*******/
/*      copia      */
/*******/
copia()
{
    sigue=0;
    i=0;
    ar[en].arbol[0]=d1[ap];
    while((ar[en].desc[i++]=d1[ap+i])!='\0')

```

semantic.c

```

; /* copiamos el resto de la linea */
ca <';
cu=base;
hijo=arfcn1.hijos;
return;
}

```



```

#include <stdio.h>
#define MAXNOD 1000
#define TAMBUF 1000

FILE *fopen(), *fpar, *fpd, *fps;

struct nodo {
    char arbol[2]; /* La Tabla */
    int hijos; /* numero de hijos */
    int apun; /* apuntador a primer hijo */
    int par1, par2; /* parametros para otros operadores */
    int rep; /* numero de veces que ha aparecido la etq.*/
};

int en; /* nodo actual */
int hijo; /* numero de hijos actual */
int base; /* base del subarbol */
int e;
int nofin;
int inietq, lametq; /* inicio y tamaño de etiqueta */
char buff[TAMBUF]; /* buffer de salida */
int i;

extern int finsube();
extern int finetq();

main(argc, argv)
int argc;
char *argv[];
{
    if (argc != 4)
    {
        printf("faltan o sobran nombres de archivos\n");
        printf("dar el nombre del archivo de datos, el arbol y el indice \n");
    }
    else
    {
        if ((fpd=fopen(++argv, "r")) != NULL)
            printf("no existe ese archivo de datos\n");
        else
        {
            if ((fpar=fopen(++argv, "r")) != NULL)
                printf("no existe ese arbol\n");
            else
            {
                fps=fopen(++argv, "w");

                for(i=0; i<MAXNOD; ++i)
                    arbol[i].arbol[0] = '\0';

                /*
                /*
                leemos el arbol
                */
            }
        }
    }
}

```

```

/*
  en1;
  while (fscanf(fpar, "%1s", &e) != EOF)
  {
    switch (e)
    {
      case '*':
        ar[en1].arbol[0] = e;
        fscanf(fpar, "%Z3dZ4d",
              &ar[en1].arbol[1], &ar[en1].hijos, &ar[en1].apun);
        break;
      case '$':
        ar[en1].arbol[0] = e;
        break;
      case '0':
        ar[en1].arbol[0] = e;
        break;
      case '%': /* operador, nueva base, inicio de etq., tamaño de etq. */
        ar[en1].arbol[0] = e;
        fscanf(fpar, "%4dZ3dZ3dZ2d\n", &ar[en1].apun, &ar[en1].par1,
              &ar[en1].par2, &ar[en1].hijos);
        break;
      case '#': /* operador, inicio dentro de etq., máximo */
        ar[en1].arbol[0] = e;
        fscanf(fpar, "%3dZ3d", &ar[en1].par1, &ar[en1].par2);
        break;
    } /* del switch */

    ++en1;
  } /* de while */

/*
  escribimos el arbol
  */
  en1;
  while (en1 < MAXNOD)
  {
    switch (ar[en1].arbol[0])
    {
      case '*':
        printf("%d: %c %Z3dZ4d\n", en1, ar[en1].arbol[0],
              ar[en1].arbol[1], ar[en1].hijos, ar[en1].apun);
        break;
      case '$':
      case '0':
        printf("%d: %c\n", en1, ar[en1].arbol[0]);
        break;
      case '%': /* operador, nueva base, inicio de etq., tamaño de etq. */
        printf("%d: %c %4dZ3dZ3dZ2d\n", en1, ar[en1].arbol[0],
              ar[en1].apun, ar[en1].par1,

```

```

        ar[en].par2, ar[en].hijos);
        break;
    case 'H':
        /* operador, inicio dentro de etq., maximo */
        /*
            printf("%d: %cZ3dZ3d\n", en, ar[en].arbol[0],
                ar[en].par1, ar[en].par2);
            break;
        */
    }
    /* del switch */
    /*
        ++en;
    */
}
/* de while */

/*
    rastreamos la informacion y generamos el indice
*/
for (i=0; i<TAMBUO; ++i) buf[i] = ' ';
for (i=0; i<MAXNOD; ++i) ar[en].rep=0;
nofin=1;

finetq();
while(nofin)
{
    switch (ar[en].arbol[0])
    {
        case '%':
            disyun();
            break;
        case 'Z':
            subarbol();
            break;
        case 'O':
            omile();
            break;
        case 'H':
            rastrea();
            break;
        case '$':
            finfi();
            break;
    }
    /* del switch */
}
/* del while */

close(fpd);
close(fpar);

```



```

        close(fpod);
    } /* 3 er. else */
    i /* 2 in. else */

    } /* 1 gr. else */
printf("fin termino bien\n");
} /* fin de main */

/*-----*/
/*----- operadores -----*/
/*-----*/

/*-----*/
/* DISYUNCIÓN */
/*-----*/
disyun()
{
    while (hijo>0)
    {
        if (c== '\n') c=getc(fpd);
        if (c==EOF)
        {
            printf("error en los datos termino mal, fin de arch. inesperada\n");
            nofin=0;
            return;
        }
        else
        {
            if (c==fenl.arbol[1]) /* caracter actual igual a caracter del nodo */
            {
                hijo=arbol.hijos;
                c=arbol.apun;
                c=getc(fpd); /* obtengo el sig. car. de los datos */
                return;
            }
            else
            {
                --hijos;
                c=cn;
                if (hijos==0)
                {
                    printf("error en datos termino mal c=%c cn=%d\n",c,cn);
                    nofin=0;
                    return;
                }
            }
        }
        i /* del else */
    } /* del while */
} /* de disyun() */ /* ojo nunca sale por aqui */

/*-----*/
/* subarbol */
/*-----*/

```

```

/* ..... */
subarbol()
{
    if (arbol.rep==arbol.hijos) /* mas repeticiones de las permitidas */
    {
        printf("mas rep de permitidas\n");
        while ((c=getc(fpd))!=>)&&(c!=EOF); /* saltamos toda la etq.*/
            c=getc(fpd);
            finetq();
    }
    else
    {
        base=arbol.apun;
        getc(fpd); getc(fpd); /* salto indicadores */

        inietq=arbol.par1;
        tmetq=arbol.par2;

        inietq+=(arbol.rep*tmetq);
        arbol.rep+=1;
        c=base;
        c=<<1;
        hijos=arbol.hijos;
    }
    return;
}

/* ..... */
/*      omite      */
/* ..... */
omite()
{
    int f;

    f=1;
    while(f)
        switch(c)
        {
            case '>': f=finetq();
                    break;
            case '<': f=finsubc();
                    break;
            case EOF: f=0;
                    nofin=0;
                    printf("Error fin de archivo inesperado\n");
                    break;
            default: c=getc(fpd);
                    break;
        }
};

return;
}

```

```

/* ***** */
/*      rastrea      */
/* ***** */
rastrea()
{
    int col,max,f;

    /* inicio relativo dentro de etiqueta */
    col=inieltq+parfen1.par1;
    max=col+parfen1.par2;
    f=1;

    while(f)
    {

        switch(c)
        {

            case '>': f=fineltq();

                    break;

            case '$': f=finsubc();

                    break;

            case EOF:printf("error fin de archivo inesperado\n");

                    f=0;
                    nofin=0;
                    break;

            default:

                    if (col>max)
                        while((cl='>')&&(cl!='$')&&(cl!=EOF)) c=getc(fpd);
                    else
                    {
                        buf[col++]=c;
                        c=getc(fpd);
                    }
                    break;

                } /* de switch */
        } /* del while */

        }

/* ***** */
/*      fin de ficha      */
/* ***** */
finfi() /* escribimos el registro */
{
    for (i=0;i<TAMBUF;++i) fprintf(fpas,"%c",buf[i]);

```

ESTA TESIS NO DEBE
SALIR DE LA BIBLIOTECA

```
fprintf(fpas, "\n");
for (i=0; i<TAMBUF; ++i) buf[i]=' ';
for (i=1; i<MAXNOD; ++i) arbol.rep=0;
finetq();
return;
```

```
/* ..... */
/* Rutinas auxiliares */
/* ..... */
```

```
int finetq()
{
    enl;
    e444;
    hijosarfenl.hijos;
    return(0);
}
```

```
int finsube()
{
    en=arbol.sel.apun; /* ojo yo estamos en el nodo de $ */
    hijosarfenl.hijos;
    return(0);
}
```

```

/*XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*/
%%
%%      Programa: creapun.c                                %%
%%
%%      Programa: Laura Freidberg Bojman                  %%
%%
%%      Objetivo: Este programa genera un archivo con los %%
%%                control colmez de las fichas y sus res- %%
%%                pectores al principio y al final de la %%
%%                ficha.                                     %%
%%
%%      Datos de Entrada: El archivo que se quiere indexar. %%
%%
%%      Datos de salida: El archivo con los numeros de %%
%%                ficha y los apuntadores.                %%
%%
%%      Metodo: Se utiliza una tabla en la que se repre- %%
%%                sentan automa-                            %%
%%                las conectados para el analisis lexico %%
%%                de la informa-                               %%
%%                cion. La tabla tiene codificados unica- %%
%%                mente dos operadores semanticos: en el %%
%%                nodo agregado al estado que reconoce %%
%%                la etiqueta FIN> el operador $ que %%
%%                escribe el numero de control colmez y %%
%%                los apuntadores en el archivo de %%
%%                salida y reinicializa los apuntadores, %%
%%                y en el registro agregado al estado %%
%%                que reconoce el subcampo #a de la %%
%%                etiqueta O35> (numero de control %%
%%                colmez) el operador # que copia en el %%
%%                reg. de salida la informacion %%
%%                correspondiente. Todos los demas reg. %%
%%                agregados tienen codificado el %%
%%                operador que emite 0.
%%
%%      Informacion Practica: el programa se ejecuta %%
%%                de la siguiente manera:                %%
%%                creapun.ex <archent> <archsal>          %%
%%
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*/

```

```

#include <stdio.h>
#define MAXNOD 1300
#define LAMBDF 7

```

```
FILE *fopen(), *fpar, *fpent, *fpaal;
```

```

struct node {
    char arbol[LAMBDF]; /* operadora es y caracteres por comparacion */
    int hijos;          /* numero de hijos */
    int apun;           /* apuntador a primer hijo */
    int ar[MAXNOD];
}

```

```

int en;                /* nodo actual */
int hijo;              /* numero de hijos actual */
int base;              /* base del subarbol */
int c;
int nofin;
int i;
long int inif, contcar;
char bufsal[LAMBDF];

```

```

/XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX/
%%
%% Programa: creapun.c %%
%%
%% Programa: Laura Fridberg Gofman %%
%%
%% Objetivo: Este programa genera un archivo con los numeros de %%
%% control colmez de las fichas y sus respectivos apun %%
%% tadores al principio y al final de la ficha. %%
%%
%% Datos de Entrada: El archivo que se quiere indexar. %%
%%
%% Datos de salida: El archivo con los numeros de ficha y los %%
%% apuntadores. %%
%%
%% Metodo: Se utiliza una tabla en la que se representan automa- %%
%% tas conectadas para el analisis lexico de la informa- %%
%% cion. La tabla tiene codificados unicamente dos oper- %%
%% radores semanticos: en el nodo agregado al estado que %%
%% reconoce la etiqueta FIN> el operador $ que escribe %%
%% el numero de control colmez y los apuntadores en el %%
%% archivo de salida y reinicializa los apuntadores, y %%
%% en el registro agregado al estado que reconoce el %%
%% subcampo %a de la etiqueta O35> (numero de control %%
%% colmez) el operador # que copia en el reg. de salida %%
%% la informacion correspondiente. Todos los demas reg. %%
%% agregados tienen codificado el operador que omite 0. %%
%%
%% Informacion Practica: el programa se ejecuta de la siguiente %%
%% manera: %%
%% creapun.ex <archent> <archsalsal> %%
%%
/XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX/

```

```

#include <stdio.h>
#define MAXNOD 1300
#define TAMBUEF 7

```

```

FILE *fopen(), *fpar, *fpent, *fpsal;

```

```

struct nodo {
    char arbol[E2]; /* La Tabla */
    int hijos; /* operadores y caracteres por comparar */
    int inif; /* numero de hijos */
    int apun; /* apuntador a primer hijo */
    int ar[MAXNOD];
};

int cn; /* nodo actual */
int hijo; /* numero de hijos actual */
int base; /* base del subarbol */
int c;
int nofin;
int i;
long int inif, contcar;
char bufsal[TAMBUEF];

```

```

char arch[11];

extern int finsube();
extern int fineta();
extern copcad();

main(argc, argv)
int argc;
char *argv[];
{
    if (argc != 3)
    {
        printf("faltan o sobran nombres de archivos\n");
        printf("dar el nombre del archivo de entrada y el de salida \n");
    }
    else
    {
        if ((fpent=fopen(argv[1], "r")) == NULL)
            printf("no existe ese archivo de entrada\n");
        else
        {
            fpсал=fopen(argv[2], "w");
            fpar=fopen("arbolnum", "r");
            copcad(arch, argv[1]);

            for(i=0; i<MAXNOD; ++i)
                ar[i].arbol[0]=' ';
/* ===== */
/*          leemos el arbol          */
/* ===== */
            cn=1;
            while (fscanf(fpar, "%1s", &c) != EOF)
            {
                switch (c)
                {
                    case '*':
                        ar[cn].arbol[0]=c;
                        fscanf(fpar, "%c%2d%4d",
                            &ar[cn].arbol[1], &ar[cn].hijos, &ar[cn].apun);
                        break;
                    case '#':
                        ar[cn].arbol[0]=c;
                        break;
                    case '0':
                        ar[cn].arbol[0]=c;
                        break;
                    case 'Z': /* operador, nueva base, numero de hijos */
                        ar[cn].arbol[0]=c;
                        fscanf(fpar, "%4d%2d\n", &ar[cn].apun,
                            &ar[cn].hijos);
                        break;
                    case 'H': /* operador */
                        ar[cn].arbol[0]=c;
                        break;
                }
                cn++;
            } /* del switch */

```

```

    }
    if /* de while */

/* ===== */
/*
/*     escribimos el arbol
/*
/*     en:1;
/*     while (en<MAXNOD)
/*     {
/*         switch (ar[en].arbol[0])
/*         {
/*             case '*':
/*                 printf("%d:%c%2d%4d\n",en,ar[en].arbol[0],
/*                     ar[en].arbol[1],ar[en].hijos,ar[en].apun);
/*                 break;
/*             case '#':
/*             case '0':
/*             case '#':
/*                 printf("%d:%c\n",en,ar[en].arbol[0]);
/*                 break;
/*             case '%':
/*                 /* operador,nueva base,inicio de etq.,tamano de etq.*/
/*
/*                 printf("%d:%c%4d%2d\n",en,ar[en].arbol[0],
/*                     ar[en].apun,
/*                     ar[en].hijos);
/*                 break;
/*             case '#':
/*                 /* operador,inicio dentro de etq., maximo */
/*
/*                 printf("%d:%c\n",en,ar[en].arbol[0]);
/*                 break;
/*         }
/*         /* del switch */
/*         ++en;
/*     }
/*     /* de while */
/*
/* ===== */
/*
/*     rastreamos la informacion
/*
/* ===== */
noinf=1;
contear=0;
inif=0;
for(i=0; i<TAMBUF;++i) buf[sol[i]]=' ';

```



```

finetq();
while(!nofin)
{
    switch (ordenl, arbol101)
    {
        case 'S':
            disyun();
            break;
        case 'Z':
            subarbol();
            break;
        case 'O':
            omite();
            break;
        case 'R':
            rastrea();
            break;
        case 'F':
            finfi();
            break;
    }; /* del switch */
} /* del while */

/* ..... */

    close(fpent);
    close(fpar);
    close(fpsal);
} /* 2 o. else */

} /* 1 er. else */
printf("fin termina bien\n");
} /* fin de main */

/* ..... */
/* ..... operadores ..... */
/* ..... */

/* ..... */
/* ..... DISYUNCION ..... */
/* ..... */
disyun()
{
    while (hijo>0)
    {
        if (c=="\n") {c=getc(fpent); ++contcar; }
        if (c==EOF)
        {
            nofin=0;
            return;
        }
    }
    else

```



```

        nofin=0;
        printf("Error fin de archivo inesperado\n");
        break;
    default:
        c=getc(fpent); ++contcar;
        break;
    }
}

return;
}

/* ***** */
/*      rastrea      */
/* ***** */
rastrea()
{
    int f,ar=0;

    f=1;
    while(f)
    {
        switch(c)
        {
            case '>':
                f=finetq();
                break;

            case '$':
                f=finsube();
                break;

            case EOF:printf("error fin de archivo inesperado\n");
                    f=0;
                    nofin=0;
                    break;

            default:
                bufsallap++;c;
                c=getc(fpent);
                ++contcar;
                break;
        }
        /* de switch */
    }
    /* del while */
}

/* ***** */
/*      fin de ficha      */
/* ***** */
finfi() /* escribimos el registro */
{
    for(i=1; i<TAMBUF;++i) fprintf(fpsal,"%c",bufsall[i]);
    /* ojo escribo desde 1 porque no quiero el blanco */
    fprintf(fpsal,"%9ld %9ld %s\n",ini1,contcar-1,arch);
}

```

```
inif=conten;  
for(i=0; i<TAMBUF; ++i) bufz[i]=0;  
finetsq();  
return;  
}
```

```
/*-----*/  
/* Rutinas auxiliares */  
/*-----*/
```

```
int finetsq()
```

```
{  
    en--;  
    es<<2;  
    hijo=ar[en].hijos;  
    return(0);  
}
```

```
int finsube()
```

```
{  
    en=ar[base].apun; /* ojo ya estamos en el nodo de * */  
    hijo=ar[en].hijos;  
    return(0);  
}
```

```
copiada(s,t) /* copia la cadena t a la cadena s */  
char *s, *t;
```

```
{  
    while (*s++ = *t++);  
}
```

```

/*XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*/
ZZ
ZZ      Programa: recupera.c
ZZ
ZZ      Programa: Laura Freidberg Bojman
ZZ
ZZ      Objetivo: Con este programa se recupera la informacion completa
ZZ                  ta de las fichas seleccionadas.
ZZ
ZZ      Datos de entrada: Un archivo con los apuntadores al inicio y al
ZZ                  fin de las fichas que se desean recuperar, y el arch
ZZ                  chivo original con los datos completos.
ZZ
ZZ      Datos de salida: El archivo con la informacion completa.
ZZ
ZZ
ZZXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*/

```

```

#include <stdio.h>
main(argc,argv)
int argc;
char *argv[];
{
    int fp1;
    FILE *fopen(), *fp2;
    int pri,i,n,m;
    long int posi,posf;
    char c,erro;
    char buff[1536];

    fp1=open(argv[1],0); /* archivo de datos */
    fp2=fopen(argv[2],"w"); /* archivo de salida con las buenas */

    pri=0;
    while ((erro=getchar()) != EOF)
    {
        if (erro=='1' && pri==0)
        {
            close(fp2);
            fp2=fopen("malas","w");
            pri=1;
        }
        /* putchar(erro); */
        posi=posf=0;
        for (i=0; i<=1535; ++i) buff[i]=' ';
        for (i=1; i<=60; ++i) { c=getchar(); /* putchar(c); */ }
        while ((c=getchar())!='\n') posi=posi*10+(c-'0');
        while ((c=getchar())!='\n') posf=posf*10+(c-'0');
        /* printf(" %ld %ld ",posi,posf); */
        lseek(fp1,posi,0);
        m=posf-posi+1;
        /* printf("m: %5d ",m); */
        n=read(fp1,buff,m);
        if (n!=m)

```

```
    printf("MAYOR AL LEER CARACTERES\n");
    printf("n: %5d , %1d\n", n, (posf-posi+1));
    for (i=0; i<n; ++i) printf("%c",buff[i]);
    printf("\n\n");
}
for (i=0; i<n; ++i) puts(buff[i],fp2);
}
```

ANEXO B


Ejemplos de algunos productos elaborados
con programas que siguen la
metodología de programación descrita en la tesis
e incluyen operadores semánticos complejos

JUEGO DE TARJETAS CATALOGRAFICAS

320.98 W9276 no. 54	PUB
<p>Dunkerley, James The crisis in Bolivia [by]...and Rolando Morales. Notre Dame, Ind., Helen Kellogg Institute for International Studies, University of Notre Dame, 1985. 55 p. 28 cm. (Working paper, 54)</p>	
<p>I. Bolivia-Politica y gobierno-1982- . I. Morales, Rolando, coaut. II. Titulo. III. Serie.</p>	
JVS/jva	COLMEX/031850

320.98 W9276 no. 54	The crisis in Bolivia
<p>Dunkerley, James The crisis in Bolivia [by]...and Rolando Morales. Notre Dame, Ind., Helen Kellogg Institute for International Studies, University of Notre Dame, 1985. 55 p. 28 cm. (Working paper, 54)</p>	
<p>I. Bolivia-Politica y gobierno-1982- . I. Morales, Rolando, coaut. II. Titulo. III. Serie.</p>	
JVS/jva	COLMEX/031850

320.98 W9276 no. 54	BOLIVIA-POLITICA Y GOBIERNO- 1982-
<p>Dunkerley, James The crisis in Bolivia [by]...and Rolando Morales. Notre Dame, Ind., Helen Kellogg Institute for International Studies, University of Notre Dame, 1985. 55 p. 28 cm. (Working paper, 54)</p>	
<p>I. Bolivia-Politica y gobierno-1982- . I. Morales, Rolando, coaut. II. Titulo. III. Serie.</p>	
JVS/jva	COLMEX/031850

1 330 38 W9271 no. 54	Morales, Rolando, coaut.
	Dunckerley, James. The crisis in Bolivia [by]... and Rolando Morales. Notre Dame, Ind., Helen Kellogg Institute for International Studies, University of Notre Dame, 1985. 55 p. 29 cm. (Working paper, 54)
	I. Bolivia-Politica y gobierno-1982- Rolando, coaut. II. Titulo. III. Serie.
JVS/iva	 COLMEX/031850

ETIQUETAS PARA EL MARCAJE DEL LIBRO

296644

346.1/M5113n/
 Baja California
 México
 Nacimientos...

296644

346.1/M5113n/
 Baja California
 México
 Nacimientos...

346.1
 M5113n
 Baja Califo

- (069) 051
M277
V. 229
1979
no. 3
sobr
Klare, Michel T., 1942-
South Africa's U.S. weapons connections.
(Washington) Institute for Policy Studies
(1979)
141 p. illus. 29 cm.
Sobretiro de: The Nation, July 28-August 4,
1979
A la cabeza de la portada: Rne corporate
sunrunners.
COLMEX/050793
- (060) f
070.48
F446
El peronismo rural en Africa. Por Paul Ansa (y
otros) Paris, Organización de las
Naciones Unidas para la Educación, la
Ciencia y la Cultura, 1981.
31 p. 27 cm.
COLMEX/009184
- (061) f
070.48
R948
Rural journalism in Africa. By Paul Ansa (et
al.) (Paris) Unesco [1981]
35 p. 27 cm. (Reports and papers on mass
communication, 88)
COLMEX/053122
- (062) 079.6
H313in
Harris, Phil.
La información sobre Africa austral; cómo
informan desde el sur de Africa las agencias
occidentales de noticias [traducción de Homero
Aisnal] (Barcelona) (Serbal) [1984]
188 p. 19 cm. (Colección de temas
africanos, 19)
Coed. con UNESCO.
COLMEX/052995
- (063) 079.6
R974p
Ruth Sloan Associates, Washington, D.C.
The press in Africa. Edited by Helen Kitchen
Washington, 1956.
95 p. 29 cm.
COLMEX/050893
- (064) f
079.6511
L627n
Leselbaum, Charles.
Notes sur "El Correo Espanol" de Orán.
Paris, Hispaniques, 1975.
13 p. 24 cm.
Sobretiro de: Mélanges offerts à Charles
Vincent Aubrun.
COLMEX/052318
- (065) 082.1
8846
1958
v. 86
Wilson, John Albert, 1899-
La cultura egipcia. Tr. de Florentino M.
Torner. México, Fondo de Cultura
Económica [1958]
483 p. 17.5 cm. (Breviarios del Fondo de
Cultura Económica, 86)
COLMEX/052717
- (066) 082.1
8846
v. 185
Paulse, Denise.
Las esculturas del Africa negra (Tr. de
Francisco González Aránguiz) México, Fondo
de Cultura Económica [1962]
171 p. illus. 17.5 cm. (Breviarios del
Fondo de Cultura Económica, 165)
COLMEX/052719
- (067) 082.1
8846
v. 86
Wilson, John Albert, 1899-
La cultura egipcia (Tr. de Florentino M.
Torner) México, Fondo de Cultura Económica
[1958]
483 p. illus. 17.5 cm. (Breviarios del
Fondo de Cultura Económica, 86)
COLMEX/052716
- (068) 082.1
03
v. 1604
Déjeux, Jean.
La littérature Algérienne contemporaine.
Paris, Universitaires de France [1975]
126 p. 18 cm. (Que sais-je?, 1604)
COLMEX/050214
- (069) 082.1
03
v. 241
Labouret, Henri, 1878-
Histoire des noirs d'Afrique. Paris, Presses
Universitaires de France, 1950.
126 p. mapas. 17.5 cm. (Que sais-je, 241)
COLMEX/052721
- (070) 082.1
03
v. 241a
Labouret, Henri, 1878-
L'Afrique précoloniale. Paris, Presses
Universitaires de France, 1959.
126 p. mapas. 17.5 cm. ("Que Sais-Je?"
Le point des connaissances actuelles, No. 241)
Bibliografía: p. [125]-126.
COLMEX/052722
- (071) 082.1
03
v. 316
Klein, Jacques.
La Tunisie. Paris, Presses Universitaires de
France, 1949.
126 p. mapas, diagrs. 17.5 cm. ("Que
sais-Je?". Le point des connaissances actuelles,
316)
COLMEX/052723
- (072) 082.1
03
v. 4
Julien, Charles-André, 1891-
Histoire de l'Afrique. Ed. refondu et mise à
jour. Paris, Presses Universitaires de France,
1955.
136 p. mapas. 17.5 cm. ("Que sais-Je?"
Le point des connaissances actuelles, 4)
Bibliografía: p. 134-135.
COLMEX/052720

- Teherán. Conferencia, 1949) 2762
 Teledifusión - África) 1818
 Temblores - África Oriental - Etiopía) 1935
 Teridofilas - África Ecuatorial Francesa) 3239
 Tesis y disertaciones académicas - Canadá -
 Bibliografía) 0031
 Tesis y disertaciones académicas - EE.UU. -
 Bibliografía) 0031
 Textiles - Industria y telas - India) 1926
 Texto, Libros de - África del Sur) 1766
 Thiers, Adolphe, 1757-1877) 0824
 Tronja (Tribu) 3390
 Tibbu (Pueblo africano) 3282
 Tierra - Tenencia, Derecho Vivienda - África del Sur -
 Botswana) 1247
 Tierra - Tenencia, Derecho de - África Central -
 Camerún) 1251
 Tierra - Tenencia, Derecho de - África Occidental -
 Ghana) 1400
 Tierra - Tenencia, Derecho de - África Occidental -
 Nigeria) 1241
 Tierra - Tenencia, Derecho de - África Oriental -
 Etiopía) 1250, 1252
 Tierra - Tenencia, Derecho de - África Oriental) 3308
 Tierra - Tenencia, Derecho de - África del Sur) 3349
 Tierra - Tenencia, Derecho de - África, Sur del Sahara) 1253
 Tierra - Tenencia, Derecho de - África) 1254, 1256
 Tierra - Tenencia, Derecho de - Babilonia) 2754
 Tierra - Tenencia, Derecho de - Tanganyika) 1705
 Tierra - Tenencia, Derecho de - África) 1255
 Tierra - Tenencia, derecho de - África Oriental -
 Etiopía) 1246
 Tierra - Tenencia, Derecho de - África Oriental -
 Etiopía) 1249
 Tierra, Tenencia - África del Norte - Egipto) 1264
 Tierra, Tenencia - Asia - Iraq) 1264
 Tierra, Tenencia - Asia - Siria) 1264
 Tierra, Tenencia - Derecho de - Cercano Oriente) 1264
 Tierra, Tenencia - Oriente (Cercano Oriente)) 1264
 Tierra, Tenencia - Países Árabes) 1264
 Tierra, Tenencia, Derecho de - África del Norte -
 Marruecos) 1365
 Tierras Áridas - África Occidental) 1261
 Tiabuku - Historia) 3157
 Tiabuku - Vida social y costumbres) 2544
 Tiabuku, Mall - Historia) 3154
 Tiabuku, Sudán - Historia) 3163
 Tiagad, Argelia - Descripción) 2075
 Timur, el Grande, 1336-1405) 2280
 Togo - Civilización) 2566
 Togo - Lenguas) 3241
 Togo - Religión) 1982
 Togoland Francés - Historia) 3209
 Tolomeos, reyes de Egipto) 2719
 Tong King - Historia) 3397
 Tolemis) 2252
 Touré, Sékou, Pres. Guinea, 1922-) 0950
 Touré, Sékou, pres. Guinea - 1922-2) 2539
 Touré, Sékou, Pres. Guinea, 1922-) 2766
 Trabajadores agrícolas) 1137
 Trabajadores agrícola - África del Norte - Liberia) 1369
 Trabajadores agrícolas - África Central) 1140
 Trabajadores agrícolas - África Occidental - historia) 0001, 0002
 Trabajadores agrícolas - África Occidental) 1148
 Trabajadores agrícolas - África Oriental - Tanzania) 1089
 Trabajadores agrícolas - África del Norte - Argelia) 1147
 Trabajadores agrícolas - África del Norte - Liberia) 1368
 Trabajadores agrícolas - África del Norte - Sudán) 1356
 Trabajadores agrícolas - África del Norte - Túnez) 1147
 Trabajadores agrícolas - África, Sur del Sahara) 0711, 0712
 Trabajadores extranjeros - África Occidental Francesa -
 Costa de Marfil) 1155, 2529, 2536
 Trabajadores extranjeros - Europa - Francia) 1136
 Trabajadores extranjeros - Europa - Gran Bretaña) 1156
 Trabajadores extranjeros (Europa) 1154
 Trabajadores extranjeros africanos - Hispanoamérica -
 Jamaica) 1233
 Trabajadores extranjeros - E.U.) 1157
 Trabajadores y trabajadores - África Oriental - Tanzania) 1089
 Trabajadores, Representación ante la gerencia - África del Norte - Argelia) 1016
 Trabajo - Leyes y Legislación - África) 1166
 Trabajo - Movilidad - África del Norte - Argelia) 1112
 Trabajo - Oferta - África Central - Angola - Historia) 0187
 Trabajo - Oferta - África Central - Camerún) 1120
 Trabajo - Oferta - África Oriental - Kenia) 1135
 Trabajo - Oferta - África Oriental - Tanzania) 0735
 Trabajo - Oferta - África del Norte - Argelia) 1136
 Trabajo - Oferta - África del Norte - Egipto) 1134, 1200
 Trabajo - Oferta - África del Norte - Sudán) 0177
 Trabajo - Oferta - África del Sur - Zambia) 1598
 Trabajo - Oferta - África) 1121, 1598
 Trabajo y trabajadores - África - Accra) 1130
 Trabajo y trabajadores - África - Alto Volta) 0402
 Trabajo y trabajadores - África - Congresos) 0994
 Trabajo y trabajadores - África - Discursos, ensayos, conferencias) 1160
 Trabajo y trabajadores - África Central - Camerún) 1120, 1126
 Trabajo y trabajadores - África Central - Rodesia del Norte) 1129

0091245008900113200003002013000015002844400085003998900107007897100070994419917
01230520>840230 1974 us a eng >bb sa 011101>bb sa CM sc CM>bb
sa 301.36 \$b H281p>10 sa Harik, Iliya F->14 sa The political mobilization of peasants
sb a study of an Egyptian community>6b sa Bloomington, Ind. so Indiana University Press sc [c1974]>bb sa x, 309 p. sb illus. sc 24 cm.>bb sa International Development Research Center \$p Studies in Development sv >bb sa
Aldeas de Africa del Norte sz Egipto>bb sa Participacion politica sc Africa del Norte sz Egipto>bb sa Indiana sb University>bb sa 125910>x
0622000002200157000450000000410000003500130004104000150005109200220006910000300
0912450056001242600109001803000023002394400059003126910098000371901001300451>840
330 1961 sp >bb sa 011102>bb sa CM sc CM>bb sa 301.36 \$
> S676p>10 sa Soares, Amadeu de Castilho>10 sa Politica de bem-estar rural em Angola sb Ensaio>0b sa Lisboa \$b Junta de Investigacoes de Ultramar. Centro de Estudos Politicos e Sociais sc 1961>bb sa 278 p. sc 25 cm.>bb sa Estudos de Ciencias Politicas e Sociais sv No. 49>bb sa Angola \$x Condicoes rurales>bb sa 093308>x
0644000002200169000450000000410000003500130004104000150005104900120006909200220
0811000036001032450060001392600078001593000023002775460023003006900138003239010
01300451>840830 1975 ua ara >bb sa 011103>bb sa CM sc CM>bb
sa ARABE>bb sa 301.36 \$b A122a>10 sa [Al-'Abbae'di', 'Abd Alla'>10 sa Annae' wa-nama'di'ey al-mudun al-kubra fi-l-Suda'n>0b sa El Cairo sb Instituto de Investigaciones y Estudios Arabes sc 1975]>bb sa 216 p. sc 24 cm.>bb sa Texto en arab>bb sa Ciudades y pueblos \$x Proyectos y urbanizacio@'n sz Africa del Norte sz Suda'n>bb sa Urbanizacio@'n sz Africa del Norte sz Suda'n>bb sa 15970>x
0756000002200169000450000800410000003500130004104000150005409200210006910000300
0902450045001202600133001653000023002956900655003216018157003767100040005339010
1300573>840830 1958 uk eng >bb sa 011104>bb sa CM sc CM>bb
sa 301.36 \$b E64p>10 sa Epstein, Arnold Leonard>10 sa Politics in an urban African community>0b sa [Manchester] sb Published on the Rhodeslivingstone Institute Northern Rhodesia by Manchester University Press sc [1958]>bb sa 254 p. c 22 cm.>bb sa Ciudades y pueblos sz Rodesia del Norte \$x Casos>bb sa Lusshya, Rodesia del Norte >bb sa Rodesia del Norte sz Razas nativas>bb sa Africa Central \$x Razes nativas>bb sa Ciudades y pueblos sz Africa Central>20 sa Victoria University of Manchester>bb sa 025732>x
0534000002200157000450000800410000003500130004104000150005409200220006910000300
091245009300121280004900214300002800263690004200291710003000323901001300363>840
30 1970 us eng >bb sa 011105>bb sa CM sc CM>bb sa 301.36 \$
L778w>10 sa Little, Kenneth Lindsay>10 sa West African urbanization sb a study of voluntary associations in social change>0b sa [New York] sb Cambridge University sc 1970>bb sa vii, 179 p. sc 21 cm.>bb sa Urbanizacio@'n sz Africa Occidental>10 sa Cambridge sb University>bb sa 185036>x
0814000002200181000450000800410000003500130004104000150005409200220006910000240
0912450125001152600099002403000031003394400018003796910112003887000023005007100
7900533901002000012>340830 1972 au a eng >bb sa 011106>bb sa CM sc CM>bb sa 301.36 \$b S525r>10 sa Shah, Mahendra M->10 sa Rural-urban population projections for Kenya and implications for development sc [by]... [ed] Frans Willekens>0b sa [Laxenburg, Austria] sb International Institute for Applied Systems Analysis sc 1978>bb sa 77 p. sb illus. sc 29 cm.>bb sa RM / 78-55>bb sa Kenia \$x Poblacio@'n rural>bb sa Kenia \$x Poblacio@'n urbana>bb sa Africa Occidental \$x Poblacio@'n rural>10 sa Willekens, Frans se coaut.>20 sa International Institute for Applied Systems Analysis, Laxenburg, Australia>bb sa 1128/174541>x

ANEXO C

ADAPTACION DEL FORMATO MARC

PARA

EL COLEGIO DE MEXICO

L I D E R Y D I R E C T O R I O

<u>POSICION DENTRO DEL LIDER O DIRECTORIO</u>	<u>CODIGO DEL CONTENIDO</u>	<u>DESCRIPCION DEL CAMPO Y DESCRIPCION DEL CONTENIDO</u>
LIDER		
00-04		Longitud lógica del registro
05		Estatus del registro
	c	Registro revisado o corregido
	d	Registro dado de baja
	n	Registro nuevo
06		Tipo de registro
	a	Lenguaje impreso
	b	Lenguaje manuscrito
	c	Música impresa
	d	Música manuscrita
	e	Mapa impreso
	f	Mapa manuscrito
	g	Audiovisuales (excepto i, j, n, o)
	i	Grabaciones de sonido no musical
	j	Grabaciones de sonido musical
	n	Materiales audiovisuales de enseñanza
	o	Juegos audiovisuales (Kits)
07		Nivel bibliográfico
	a	Análítica
	c	Colección
	m	Monografía
	s	Publicación seriada (seriais)

<u>POSICION DENTRO DEL LIDER</u>	<u>CODIGO DEL CONTENIDO</u>	<u>DESCRIPCION DEL CAMPO Y DEL CODIGO DE CONTENIDO</u>
08-09	b	Biancos
10	2	Conteo de los Indicadores de campo
11	2	Conteo de los Indicadores de subcampo
12-16		Base de la dirección de cada registro
17		Nivel de descripción bibliográfica del registro
	b	Registro completo (inspección física)
	1	Registro parcial (tomado de fichas)
18		Forma de la catalogación descriptiva
	b	Antes de AACR1 revisadas
	a	AACR2
	l	AACR1 revisadas (ISBD completo)
	p	AACR1 revisadas (ISBD parcial)
19	b	Bianco
20	4	Número de caracteres de la longitud del campo
21	5	Número de caracteres de la posición del primer carácter en cada campo
22-23	00	

DIRECTORIO

00-02		Número de etiqueta
03-06		Longitud del campo
07-11		Posición del primer carácter del campo

CAMPOS FIJOS

<u>ETIQUETA</u> <u>COLMEX</u>	<u>POSICION</u> <u>DENTRO DEL</u> <u>CAMPO</u>	<u>CODIGO</u>	<u>DESCRIPCION DEL CAMPO</u> <u>Y DESCRIPCION DEL CONTENIDO</u>
FCAP	00-05		Fecha de entrada al archivo
FSPUB	06		Código de tipo de fecha de publicación
		c	Dos fechas: real y de copyright
		d	Fecha detallada con año, mes y día
		m	Dos fechas: de inicio y de terminación
		n	Fechas de publicación desconocidas
		q	Fecha dudosa
		r	Dos fechas: original y de reimpresión
		s	Fecha única (o probable)
F1	07-10		Fecha 1
F2	11-14		Fecha 2
PAIS	15-17		Código de país de publicación
ILUS	18-21		Código de ilustraciones
		b	Sin ilustraciones
		a	Ilustraciones
		b	Mapas
		h	Fascículos
		o	Fotografías
	22		<u>Código intelectual</u>
		b	No aplica

<u>ETIQUETA</u> <u>COLMEX</u>	<u>POSICION</u> <u>DENTRO DEL</u> <u>CAMPO</u>	<u>CODIGO</u>	<u>DESCRIPCION DEL CAMPO</u> <u>Y DESCRIPCION DEL CODIGO DEL</u> <u>CONTENIDO</u>
REPROD	23		Forma de la reproducción
		b	No es reproducción
		a	Microfilm
		b	Microficha
CONT	24-27		Código de contenido
		b	No especificado
		t	Reportes técnicos
		b	Bibliografías
		c	Catálogos
		l	Índices
		l	Legislación
		a	Resúmenes
		d	Diccionarios
		e	Enciclopedias
		r	Directorios
		y	Anuarios
		s	Estadísticas
		f	Manuales
o	Reseñas		
p	Textos programados		
PUBOF	28		Código de publicación oficial
		b	No oficial
		f	Federal/Nacional

<u>ETIQUETA</u> <u>COLMEX</u>	<u>POSICION</u> <u>DENTRO DEL</u> <u>CAMPO</u>	<u>CODIGO</u>	<u>DESCRIPCION DEL CAMPO</u> <u>Y DESCRIPCION DEL CODIGO DEL</u> <u>CONTENIDO</u>
		I	Organismos Internacionales
		I	Municipio o ciudades
		O	Nivel no determinado
		S	Estatal
		U	De origen dudoso
		Z	Otros (oficiales)
CONF	29		Indicador de conferencias
		0	No es una publicación de conferencias
		1	SI es una publicación de conferencias
EDCONM	30		Indicador de Edición conmemorativa (festschraft)
		0	No
		1	SI
IND	31		Indicador de indice
		0	No tiene indice
		1	SI tiene indice
AP	32		Indicador de asiento principal
		0	No es parte del registro
		1	SI es parte del registro
	33	W	Indicador de ficción

<u>ETIQUETA</u> <u>COLMEX</u>	<u>POSICION</u> <u>DENTRO DEL</u> <u>CAMPO</u>	<u>CODIGO</u>	<u>DESCRIPCION DEL CAMPO</u> <u>Y DESCRIPCION DEL CODIGO DEL</u> <u>CONTENIDO</u>
BIO	34		Código biográfico
		b	No es material biográfico
		a	Autobiografía
		b	Biografía individual
		c	Biografía colectiva
		d	Contiene información biográfica
IDIOMA	35-37		Código de idioma
008	38		Código de registros modificados
		b	Registro idéntico a la tarjeta
		r	El registro está romanizado y la tarjeta está parcialmente romanizada
		o	El registro está romanizado y la tarjeta también
		z	El registro no contiene los diacríticos
008	39		Fuente de catalogación
		d	No es LC

CAMPOS DE LONGITUD VARIABLE

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
020	bb	\$a	Número Internacional Normalizado del Libro (ISBN)
035	bb	\$a	No. Control Col. Mex.
040	bb		Fuentes de la Catalogación
		\$a	Fuente original de catalogación
		\$b	Idioma de catalogación
		\$c	Agencia que introduce el registro
		\$d	Agencia modificadora
041	b		Idiomas
	0		Varios Idiomas
	1		Traducción
		\$a	Idioma del texto o de su traducción
043	bb	\$a	Código del área geográfica
049	bb	\$a	Registro de colecciones
050	bb		Clasificación de la Biblioteca del Congreso de EUA (LCC)
		\$a	Número clasificador asignado por la Bib. del Congreso EUA
		\$b	Número de autor
080	bb	\$a	Clasificación Decimal Universal
082	bb		Clasificación Decimal Dewey (DDC) asignada por Bib. Congreso EUA
		\$a	Número clasificador asignado

<u>ETIQUETA MARC</u>	<u>VALOR DE INDICADORES</u>	<u>CODIGO DE SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL INDICADOR O DEL CODIGO DE SUBCAMPO</u>
092	b b		Clasificación Decimal Dewey (DDC) Colmex
		\$a	Número clasificador
		\$b	Número de autor
098	b b		Otros sistemas de clasificación
		\$a	Sistema de Naciones Unidas
099	b b		Clasificación de texto libre
		\$a	Primer elemento
		\$b	Segundo elemento
100			Asiento principal-Autor personal
	0		Nombre en orden directo
	1		Apellido simple
	2		Apellido compuesto
	3		Nombre de familia
	0		El asiento principal no es asiento de materia
	1		El asiento principal si es asiento de materia
		\$a	Nombre del autor
		\$b	Numeración usada como elemento asociado al nombre
		\$c	Títulos u otras designaciones asociadas al nombre
	\$d	Fechas de nac. y/o muerte	
	\$e	Términos que describen la relación de autoría	

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
		\$f	Fecha de la obra
		\$k	Sub asiento de forma
		\$l	Idioma de la obra
		\$n	Número de la parte o sección de la obra
		\$p	Nombre de la parte o sección de la obra
		\$q	Forma completa del nombre (como adicción)
		\$t	Título usado como subasiento del nombre
		\$w	Forma del asiento verificado según AACR2
110			Asiento principal-Autor corporativo
	0		APELLIDO [(por Inversión) requerido sólo para asientos pre-AACR2]
	1		Lugar (jurisdicción)
	2		Nombre en orden directo
	0		El asiento principal no es asiento de materia
	1		El asiento principal sí es asiento de materia
		\$a	Nombre del autor corporativo
		\$b	Nombre de la entidad subordinada
		\$c	Lugar (Conferencia)

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
		\$d	Fecha (Conferencia)
		\$e	Terminos que describen la relación de autoría
		\$f	Fecha de la obra
		\$g	Información adicional
		\$k	Subasiento de forma
		\$l	Idioma de la obra
		\$n	Número de la parte ó sección
		\$p	Nombre de la parte ó sección de la obra
		\$t	Título de la obra
		\$w	Forma del asiento verificada según AACR2
111			Asiento principal-Conferencia o Reunión
	0		Apellido [(por inversión) requerido solo para asientos pre-AACR2]
	1		Lugar o lugar más nombre (requerido solo para asientos pre-AACR2)
	2		Nombre en orden directo
	0		El asiento principal no es asiento de materia
	1		El asiento principal si es asiento de materia
		\$a	Nombre de la conferencia, reunión, etc. (o, elemento geográfico)

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
		\$b	Número de la conferencia ó reunión (utilizado exclusivamente para asientos pre-AACR2)
		\$c	Lugar de realización de la conferencia ó reunión
		\$d	Fecha de la conferencia ó reunión
		\$e	Nombre de la entidad subordinada a la conferencia ó reunión
		\$f	Fecha de la obra
		\$k	Subasiento de forma
		\$l	Idioma
		\$n	Número de la conferencia ó reunión; parte o sección
		\$p	Nombre de la parte o sección
		\$q	Nombre de la conferencia ó reunión a continuación del elemento geográfico
		\$w	Forma del asiento verificada según AACR2
130			Asiento principal-Título uniforme
	0-9		Número de caracteres que no se tomarán en cuenta para la ordenación alfabética del asiento
	0		El asiento principal no es asiento de materia
	1		El asiento principal es asiento de materia

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
		\$a	Título o elemento básico del asiento de título
		\$d	Fecha de celebración de tratado
		\$f	Fecha de la obra
		\$g	Información adicional
		\$h	Designación general de material
		\$k	Subasiento de forma
		\$l	Idioma
		\$n	Número de la parte o sección de la obra
		\$p	Nombre de la parte o sección de la obra
		\$s	Versión
		\$t	Título
		\$w	Forma del asiento verificada según AACR2
240			Título uniforme
	0-9		Número de caracteres que no se tomarán en cuenta para la ordenación alfabética del asiento
		\$a	Título uniforme
		\$d	Fecha de celebración del tratado
		\$f	Fecha de edición de la obra
		\$g	Información adicional
		\$h	Designación general de material
		\$k	Subencabezamiento de forma

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
		\$l	Idioma
		\$m	Medio de ejecución
		\$n	Número de la parte o sección
		\$p	Nombre de la parte o sección
		\$r	Clave
		\$s	Versión
		\$w	Forma de asiento verificado según AACR2
241			Título romanizado (CoIMex)
	1		Si se requiere asiento secundario de título
	0-9		Número de caracteres que no se tomarán en cuenta para la ordenación alfabética del asiento
		\$a	Título
		\$b	Subtítulo
		\$c	Mención de responsabilidad
245			Título proplamente dicho
	0		No se requiere asiento secundario de título
	1		Si se requiere asiento secundario de título
	0-9		Número de caracteres que no se tomarán en cuenta para la ordenación alfabética del asiento
		\$a	Título proplamente dicho

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
		\$b	Subtítulo; otra información acerca del título
		\$c	Mención de responsabilidad
		\$h	Designación general de material (DGM)
		\$n	Número de la parte o sección
		\$p	Nombre de la parte o sección
250	0 0		Mención de edición
		\$a	Mención de edición propiamente dicha
		\$b	Mención de responsabilidad relacionada con la edición
260	0		Ple de Imprenta
	0		El editor no es el asiento principal
	1		El editor es el asiento principal
		\$a	Lugar de publicación
		\$b	Nombre del editor
		\$c	Fecha de publicación
300	0 0		Descripción física
		\$a	Número de páginas o volúmenes
		\$b	Mención de ilustración
		\$c	Tamaño
		\$e	Material acompañante
		\$f	Tipo de unidad
		\$g	Tamaño de unidad

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
400			Mención de serie-Autor personal/Título. (solo pre-AACR2)
	0		Nombre en orden directo
	1		Apellido simple
	2		Apellido compuesto
	3		Nombre de familia
	0		El elemento principal del asiento de la serie no está representado por un pronombre
	1		El elemento principal del asiento de la serie si está representado por un pronombre ("su")
		\$a	Nombre del autor
		\$b	Numeración usada como elemento asociado al nombre
		\$c	Títulos u otras designaciones asociadas al nombre
		\$d	Fechas de nac. y/o muerte
		\$e	Términos que describen la relación de autoría
		\$f	Fecha de la obra
		\$k	Subasiento de forma
		\$l	Idioma de la obra
		\$n	Número de la parte o sección
		\$p	Nombre de la parte o sección
		\$q	Forma completa del nombre

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
		\$t	Título de la serie usado como subasiento del nombre
		\$v	Volumen o número de la serie
		\$x	ISSN
		\$w	Forma del asiento verificado según AACR2
410			Mención de serie-Autor corporativo/Título (pre-AACR2)
	0		Apellido [(por Inversión) requerido sólo para asientos pre-AACR2]
	1		Lugar (Jurisdicción)
	2		Nombre en orden directo
	0		El elemento principal del asiento de la serie no está representado por un pronombre ("su")
	1		El elemento principal del asiento de la serie sí está representado por un pronombre ("su")
		\$a	Nombre del autor corporativo
		\$b	Nombre de la entidad subordinada
		\$c	Lugar (conferencia)
		\$d	Fecha (conferencia)
		\$e	Términos que describen la relación de autoría
		\$f	Fecha de la obra
		\$g	Información adicional
		\$k	Subasiento de forma

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
		\$l	Idioma de la obra
		\$n	Número de la parte o sección
		\$p	Nombre de la parte o sección
		\$t	Título de la serie usado como subasiento del nombre
		\$v	Volumen o número de la serie
		\$w	Forma del asiento verificado según AACR2
		\$x	ISSN
411			Mención de serie-Conferencia o Reunión/Título
	0		Apellido [(por inversión) requerido sólo para asientos pre-AACR2]
	1		Lugar o lugar más nombre (requerido sólo para asientos pre-AACR2)
	2		Nombre en orden directo
	0		El elemento principal del asiento de la serie no está representado por un pronombre ("su")
	1		El elemento principal del asiento de la serie sí está representado por un pronombre ("su")
		\$a	Nombre de la conferencia, reunión, etc. (ó, elemento geográfico)

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
		\$b	Número de la conferencia ó reunión (utilizado exclusivamente para asientos pre-AACR2)
		\$c	Lugar de realización de la conferencia ó reunión
		\$d	Fecha de la conferencia ó reunión
		\$e	Nombre de la entidad subordinada a la conferencia o reunión
		\$f	Fecha de la obra
		\$k	Subasiento de forma
		\$l	Idioma de la obra
		\$n	Número de la conferencia ó reunión; parte ó sección
		\$p	Nombre de de la conferencia ó reunión a continuación del elemento geográfico
		\$q	Nombre de la conferencia ó reunión a continuación del elemento geográfico
		\$t	Título de la serie usado como subasiento del nombre de la conferencia ó reunión
		\$v	Volumen o número de la serie
		\$x	ISSN
		\$w	Forma del asiento verificado según AACR2
			Mención de serie-Título distintivo

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
	0-9		Número de caracteres que no se tomarán en cuenta, para la ordenación alfabética del asiento
		\$a	Título
		\$n	Número de la parte o sección
		\$p	Nombre de la subserie
		\$v	Volumen o número de la serie
		\$x	ISSN
490	b		Mención de serie-Serie no consignada en el registro, ó, serie consignada de manera diversa en el registro.
	0		Serie no consignada en el registro
	1		Serie consignada de manera diversa en el registro
		\$a	Mención de serie (incluye mención de responsabilidad)
		\$i	Número de la Biblioteca del Congreso de E.U.
		\$n	Número de la Biblioteca Nacional de Canada
		\$x	ISSN
		\$v	Volumen o número de la serie (La información de este campo no se utiliza para generar asientos secundarios de serie sin embargo cada vez que el valor del primer indicador sea "1" habrá una mención de serie correspondiente en 800-840, que dará la serie en la forma que será consignada en el registro)

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
500	b b	\$a	Nota general
501	b b	\$a	Nota "con"
502	b b		Nota de tesis académica
		\$a	Mención de tesis y grado académico
		\$b	Institución
503	b b	\$a	Nota de Historia bibliográfica
504	b b	\$a	Nota de bibliografías
505	b	\$a	Nota de contenido
	0		Contenido completo
	1		Contenido incompleto
	2		Contenido parcial
507	b b	\$a	Nota de escala
520	b b	\$a	Nota de resumen, alcance, etc.
533	b b		Nota de fotoreproducción
		\$a	Tipo de reproducción
		\$b	Lugar de reproducción
		\$c	Agencia responsable
546	b b	\$a	Nota de idioma
590	b b	\$a	Nota local
690	b		Encabezamientos de materia - Tópicos
	0		No validados
	1		Kidder (mecanografiado)
	2		DPT (Kidder manuscrito)

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
	3		BNM (Escamilla)
	4		UP (Roviera)
	5		Consulta especialistas
	6		Otras fuentes
		\$a	Encabezamiento de materia
		\$b	Elemento geográfico adicional al nombre del lugar
		\$x	Subdivisión general
		\$y	Subdivisión cronológica
		\$z	Subdivisión geográfica
691	0		Encabezamiento de materia - Geográfico
	0		No validados
	1		Kidder (mecanografiado)
	2		DPT (Kidder manuscrito)
	3		BNM (Escamilla)
	4		UP (Roviera)
	5		Consulta especialistas
	6		Otras fuentes
		\$a	Encabezamiento de materia
		\$b	Elemento geográfico adicional al nombre
		\$x	Subdivisión general
		\$y	Subdivisión cronológica
		\$z	Subdivisión geográfica

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
692	b		Encabezamiento de materia - Autor personal
	0		Nombre en orden directo
	1		Apellido simple
	2		Apellido múltiple
	3		Nombre de familia
		\$a	Nombre del autor personal
		\$b	Numeración usada como elemento del nombre
		\$c	Títulos u otras designaciones asociadas al nombre
		\$d	Fecha de nacimiento y/o muerte
		\$e	Términos que describen la relación de autoría
		\$f	Fecha de la obra
		\$g	Información adicional
		\$h	Designación general de material
		\$k	Subasiento de forma
		\$l	Idioma de la obra
		\$m	Medio de ejecución
		\$n	Número de la parte o sección de la obra
		\$o	Mención de arreglo
		\$p	Nombre de la parte o sección
		\$q	Forma completa del nombre (como adición)
		\$r	Clave

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
		\$s	Versión
		\$t	Título usado como subasiento del nombre
		\$x	Subdivisión general
		\$y	Subdivisión cronológica
		\$z	Subdivisión geográfica
		\$w	Forma del asiento verificado según AACR-2
693	0		Encabezamiento de materia - Auto corporativo
	0		Apellido [(por Inversión) requerido sólo para asientos pre-AACR2]
	1		Lugar (Jurisdicción) ó lugar más nombre (requerido sólo para asientos pre-AACR2)
	2		Nombre en orden directo
		\$a	Nombre del autor corporativo
		\$b	Unidad subordinada
		\$c	Lugar (conferencia)
		\$d	Fecha de la conferencia o año de tratado
		\$e	Términos que describen la relación de autoría
		\$f	Fecha de la obra
		\$g	Información adicional
		\$h	Designación general de material
		\$k	Subasiento de forma
		\$l	Idioma

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
		\$m	Medio de ejecución
		\$n	Número de la parte o sección, o, número de conferencia
		\$o	Mención sobre arreglo
		\$p	Nombre de la parte o sección
		\$r	Clave
		\$s	Versión
		\$t	Título usado como subasiento del nombre
		\$x	Subdivisión general
		\$y	Subdivisión cronológica
		\$z	Subdivisión geográfica
		\$w	Forma del asiento verificado
694	b		Encabezamiento de materia - Conferencia, reunión, etc.
	0		Apellido (Por Inversión) requerido sólo para asientos pre-AACR2
	1		Lugar ó lugar más nombre (requerido sólo para asientos pre-AACR2)
	2		Nombre en orden directo
		\$a	Nombre de la conferencia, reunión, etc.
		\$b	Número de la conferencia o reunión (utilizado exclusivamente para asientos pre-AACR2)
		\$c	Lugar de realización de la conferencia o reunión

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
		\$d	Fecha de realización de la conferencia o reunión
		\$e	Nombre de la unidad subordinada
		\$g	Información adicional
		\$h	Designación general de material
		\$q	Nombre de la reunión a continuación de lugar
		\$t	Título de la obra usado como subsiento
		\$x	Subdivisión general
		\$y	Subdivisión cronológica
		\$z	Subdivisión geográfica
		\$w	Forma del asiento verificada según AACR2
695	W		Encabezamiento de materia - Título uniforme
	0-9		Número de caracteres que no se tomarán en cuenta para la ordenación alfabética del asiento
		\$a	Título o elemento básico del asiento
		\$d	Fecha de tratado
		\$s	Versión
		\$f	Fecha de la obra
		\$h	Designación general de material
		\$k	Subsiento de forma
		\$l	Idioma de la obra

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
		\$n	Número de la parte o sección
		\$p	Nombre de la parte o sección
		\$t	Título de la obra
		\$x	Subdivisión general
		\$y	Subdivisión cronológica
		\$z	Subdivisión geográfica
		\$w	Forma del asiento verificada según AACR2
700			Asiento secundario - Autor personal
	0		Nombre en orden directo
	1		Apellido simple
	2		Apellido compuesto
	3		Nombre de familia
	0		Asiento secundario (Autor)
	1		Asiento secundario (Autor Título)
	2		Asiento analítico
		\$a	Nombre del autor
		\$b	Numeración usada como elemento asociado al nombre
		\$c	Títulos u otras designaciones asociadas al nombre
		\$d	Fechas de nacimiento y/o muerte
		\$e	Términos que describen la relación de autoría
		\$f	Fecha de la obra

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
		\$h	Designación general de material
		\$k	Subasiento de forma
		\$l	Idioma de la obra
		\$m	Medio de ejecución
		\$n	Número de la parte o sección
		\$o	Mención sobre arreglo
		\$p	Nombre de la parte o sección
		\$q	Forma completa del nombre (como adición)
		\$t	Título usado como subasiento del nombre
		\$w	Forma de asiento verificado según AACR2
710			Asiento secundario - Autor corporativo
	0		Apellido [(por inversión) requerido sólo para asientos pre-AACR2]
	1		Lugar (jurisdicción) ó lugar más nombre (requerido sólo para asientos pre-AACR2)
	2		Nombre en orden directo
	0		Asiento secundario (Autor)
	1		Asiento secundario (Autor Título)
	2		Asiento analítico
		\$a	Nombre del autor corporativo
		\$b	Unidad subordinada

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
		\$c	Lugar (conferencia)
		\$d	Fecha de conferencia, ó, año de tratado
		\$e	Términos que describen la relación de autoría
		\$f	Fecha de la obra
		\$g	Información adicional
		\$h	Designación general de material
		\$k	Subasiento de forma
		\$l	Idioma de la obra
		\$m	Medio de ejecución
		\$n	Número de la parte o sección
		\$p	Nombre de la parte o sección
		\$o	Mención sobre arreglo
		\$r	Clave
		\$s	Versión
		\$t	Título de la obra
		\$w	Forma de asiento verificado según AACR2
711			Asiento secundario - Conferencia, reunión, etc.
	0		Apellido [(Por Inversión) requerido sólo para asientos pre-AACR2]
	1		Lugar (jurisdicción) ó lugar más nombre (requerido sólo para asientos pre-AACR2)
	2		Nombre en orden directo

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
	0		Asiento secundario (Autor)
	1		Asiento secundario (Autor- Titulo)
	2		Asiento analítico
		\$a	Nombre de la conferencia, reunión
		\$b	Número de la conferencia o reunión
		\$c	Lugar de realización de la conferencia o reunión
		\$d	Fecha de realización de la conferencia o reunión
		\$e	Nombre de la unidad subordinada
		\$f	Fecha de la obra
		\$h	Designación general de material
		\$k	Subasiento de forma
		\$l	Idioma
		\$n	Número de la conferencia o reunión; parte o sección
		\$p	Nombre de la parte o sección
		\$q	Nombre de la reunión a continuación de lugar (requerido sólo para asientos pre-AACR2)
		\$s	Versión
		\$t	Título de la obra usado como subasiento
		\$w	Forma del asiento verificado según AACR2

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
730			Asiento secundario - Título uniforme
	0-9		Número de caracteres que no se tomarán en cuenta en la ordenación alfabética del asiento
	0		Asiento secundario (Autor)
	1		Asiento secundario (Autor-Título)
	2		Asiento analítico
		\$a	Título o elemento básico del asiento
		\$d	Fecha de tratado
		\$f	Fecha de la obra
		\$g	Información adicional
		\$h	Designación general de material
		\$k	Subasiento de forma
		\$l	Idioma de la obra
		\$m	Medio de ejecución
		\$n	Número de la parte o sección
		\$o	Mención sobre arregio
		\$p	Nombre de la parte o sección
		\$r	Clave
		\$s	Versión
		\$t	Título de la obra
		\$w	Forma del asiento verificado según AACR

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
740			Asiento secundario - Variante del título
	0-9		Número de caracteres que no se tomarán en cuenta en la ordenación alfabética del asiento
	1		Asiento secundario propiamente dicho
	2		Asiento analítico
		\$a	Variante del título
		\$b	Subtítulo; otra información acerca del título
		\$c	Mención de responsabilidad
		\$h	Designación general de material
		\$n	Número de la parte o sección
		\$p	Nombre de la parte o sección
800	\$		Asiento secundario de serie - Autor personal/Título
	0		Nombre en orden directo
	1		Apellido simple
	2		Apellido compuesto
	3		Nombre de familia
		\$a	Nombre del autor
		\$b	Numeración usada como elemento asociado al nombre
		\$c	Titulos u otras designaciones asociadas al nombre
		\$d	Fechas de nac. y/o muerte

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
		\$e	Términos que describen la relación de autoría
		\$f	Fecha de la obra
		\$h	Designación general del material
		\$k	Subasiento de forma
		\$l	Idioma de la obra
		\$m	Medio de ejecución
		\$n	Número de la parte o sección de la obra
		\$p	Nombre de la parte o sección de la obra
		\$q	Forma completa del nombre
		\$s	Versión
		\$t	Título de la serie usado como subasiento del nombre
		\$v	Volumen ó número de la serie
810	b		Asiento secundario de serie - Autor corporativo/Título
	0		Apellido [(por inversión) requerido sólo para asientos pre-AACR2]
	1		Lugar (Jurisdicción)
	2		Nombre en orden directo
		\$a	Nombre del autor corporativo
		\$b	Nombre de la unidad subordinada
		\$c	Lugar (Conferencia)
		\$d	Fecha

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
		\$e	Términos que describen la relación de autoría
		\$f	Fecha de la obra
		\$g	Información adicional
		\$h	Designación gen. de material
		\$k	Subasiento de forma
		\$l	Idioma de la obra
		\$m	Medio de ejecución
		\$n	Número de la parte o sección de la obra
		\$p	Nombre de la parte o sección de la obra
		\$r	Clave
		\$s	Version
		\$t	Título de la serie usado como subasiento del nombre
		\$v	Volumen o número de la serie
811	b		Asiento secundario de serie - Conferencia o Reunión/Título
	0		Apellido [(por inversión) requerido sólo para asientos pre-AACR2]
	1		Lugar (Jurisdicción)
	2		Nombre en orden directo
		\$a	Nombre de la conferencia, reunión, etc. (o elemento geográfico)

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
		\$b	Número de la conferencia, reunión (utilizado exclusivamente para asentos pre-AACR2)
		\$c	Lugar de realización de la conferencia ó reunión
		\$d	Fecha de la conferencia ó reunión
		\$e	Nombre de la entidad subordinada a la conferencia ó reunión
		\$f	Fecha de la obra
		\$g	Información adicional
		\$h	Designación gral. de material
		\$k	Subasiento de forma
		\$l	Idioma de la obra
		\$n	Número de la parte ó sección de la obra
		\$p	Nombre de la parte ó sección de la obra
		\$q	Nombre de la reunión a continuación del elemento geográfico
		\$t	Título de la serie usado como subasiento del nombre
		\$v	Volumen ó número de la serie
830	b		Asiento secundario de serie - Título uniforme
	0-9		Número de caracteres que no se tomarán en cuenta para la ordenación alfabética del asiento

<u>ETIQUETA</u> <u>MARC</u>	<u>VALOR DE</u> <u>INDICADORES</u>	<u>CODIGO DE</u> <u>SUBCAMPO</u>	<u>DESCRIPCION DEL CAMPO, DEL</u> <u>INDICADOR O DEL CODIGO DE</u> <u>SUBCAMPO</u>
		\$a	Título ó elemento básico del asiento de título
		\$d	Fecha de celebración de tratado
		\$f	Fecha de la obra
		\$g	Información adicional
		\$h	Designación gral. de material
		\$k	Subasiento de forma
		\$l	Idioma de la obra
		\$m	Medio de ejecución
		\$n	Número de la parte ó sección de la obra
		\$o	Arreglo
		\$p	Nombre de la parte ó sección de la obra
		\$r	Clave
		\$s	Versión
		\$t	Título
		\$v	Volumen ó número de la serie
840	Ø		Asiento secundario serie - Título (Pre-AACR2)
	0-9		Número de caracteres que no se tomarán en cuenta para la ordenación alfabética del asiento
		\$a	Título de la serie
		\$v	Volumen ó número de la serie
901	Ø Ø	\$a	Número de adquisición
902	Ø Ø	\$a	Claves adicionales

BIBLIOGRAFIA

BIBLIOGRAFIA

- Aho, A. V. y Corasick, M. J. "Efficient string matching: An aid to bibliographic search" Comm. ACM 18:6 p. 333-340, 1975.
- Aho, A. V. y Ullman, J. D. Principles of compiler design, Addison-Wesley, Reading Mass., 1977.
- Aho, A. V. y Ullman, J. D. The theory of parsing, translation, and compiling, Prentice-Hall, Englewood Cliffs, N.J., 1972.
- Avram, Henriette D. MARC, its history and implications, Washington, D.C.: Library of Congress, 1975.
- Bar-Hillel, Yehoshua. "A logician's reaction to theorizing on information search systems" p. 313-329. En Language and Information. Selected essays on their theory and application, Addison-Wesley, Reading Mass., 1964.
- Bar-Hillel, Yehoshua. "Theoretical aspects of the mechanization of literature searching" p. 330-364. En Language and Information. Selected essays on their theory and application, Addison-Wesley, Reading Mass., 1964.
- Encyclopedia of Library and Information Science, Marcel Dekker, Inc. N.Y. Vol. 30 p. 416-463, 1980.
- García Barbabosa, Ageo. Campos MARC para la Biblioteca Daniel Cosío Villegas, El Colegio de México, [Documento de circulación interna].
- García Hidalgo, Ma. Isabel. "Metodología" En Manejo automatizado de acervos documentales Unidad de Computo de El Colegio de México, accésit CREI DE PLATA 1985.
- García Hidalgo, Ma. Isabel y Freidberg, Laura. "Procedimiento computacional para compatibilizar el sistema de la biblioteca de El Colegio de México con microslis", Segundo Congreso Nacional de Micro CDS/ISIS, 1989.
- García Hidalgo, Ma. I., Tejeda, C. y Quijano, A. "El proyecto de automatización de la Biblioteca Daniel Cosío Villegas: Adquisiciones y Catalogación y Clasificación", XV Jornadas Mexicanas de Biblioteconomía, Tlaxcala, México, 1984.

- Hammer, D. The Information Age; its development, its impact, The Scarecrow Press, Inc. Metuchen, N.J. Scarecrow, 1976.
- Heaps, H. S. Information retrieval; computational and theoretical aspects Academic Press, San Francisco Calif. 1978.
- Hopcroft, J. E. y Ullman, J. D. Introduction to automata theory, languages, and computation, Addison-Wesley, Reading, Mass. 1979.
- Kidder, Ione Marion, comp. Encabezamientos de materia, castellano-Inglés, Inglés-castellano, México: Biblioteca Benjamin Franklin, 1947.
- Mallincoln, S. M. y Fasana, P. J. The future of the catalog: the library's choices, Professional Librarian series; N.Y.: Knowledge Industry Publications, 1979.
- Wessel, A. Computer-aided information retrieval, Melville, Los Angeles Calif. 1975.
- Woods, W. A. "Transition Network Grammars for Natural Language Analysis" Comm. ACM 13:10 p. 591-606, 1970.