

CURSO 2003/2004

Asignatura	Estructura de Datos y de la Información
Titulación	Ingeniero Técnico en Informática de Gestión Ingeniero Técnico en Informática de Sistemas
Carácter	Troncal
Código	151062010/151052010
Curso	Segundo
Cuatrimestre	Anual

TEMARIO

Teoría:

Parte 1. Nivel físico de un sistema gestor de bases de datos: Almacenamiento de datos y estructuras avanzadas de acceso.

Tema 1: Tipos Abstractos de Datos.

- 1.1 Introducción.
- 1.2 Especificación Abstracta. Encapsulamiento.
- 1.3 Interfaz e implementación.
- 1.4 Operaciones.
- 1.5 Contenedores asociativos: conjuntos y tablas.

Tema 2: Ficheros.

- 2.1 Introducción.
- 2.2 Ficheros físicos y lógicos.
- 2.3 Operaciones elementales sobre ficheros.
- 2.4 Dispositivos de almacenamiento secundario.

Tema 3: Estructura de Ficheros.

- 3.1 Introducción: Un fichero como secuencia de bytes.
- 3.2 Estructuras de campos.
- 3.3 Estructuras de registros.
- 3.4 Extracción de registros por claves: Formas canónicas para las claves.
- 3.5 Búsqueda secuencial.
- 3.6 Acceso directo.
- 3.7 Registros de encabezado.
- 3.8 Acceso y organización de ficheros.

Tema 4: Mantenimiento de Ficheros.

- 4.1 Introducción.
- 4.2 Compactación.
- 4.3 Eliminación de registros de longitud fija.
- 4.4 Eliminación de registros de longitud variable.
- 4.5 Fragmentación del almacenamiento.
- 4.6 Estrategias de colocación.

Tema 5: Indexación.

- 5.1 ¿Qué es un índice?.
- 5.2 Ordenación.
- 5.3 Índice simple.
 - 5.3.1 Operaciones básicas en un fichero secuencial indexado con índice en RAM.
 - 5.3.2 Índices demasiado grandes para caber en memoria.
- 5.4 Indexación para proporcionar acceso mediante varias claves.
- 5.5 Extracción de información mediante combinaciones de claves secundarias.
- 5.6 Listas invertidas.
- 5.7 Índices selectivos.

5.8 Enlace (binding).

Tema 6: Árboles 2-3, 2-3-4, Rojo-Negro Y Multicamino (Multiway Tree).

- 6.1 Introducción.
- 6.2 Notación preliminar y conceptos básicos de árboles.
- 6.3 Definición de árbol 2-3 y 2-3-4.
- 6.4 Inserción en un árbol 2-3-4.
- 6.5 Transformación de un árbol 2-3-4 a un árbol rojo-negro.
- 6.6 Definición de árbol de búsqueda multicamino.
- 6.7 Búsqueda en un árbol multicamino.
- 6.8 Recorrido en un árbol multicamino.
- 6.9 Inserción en un árbol multicamino.
- 6.10 Borrado en un árbol multicamino.

Tema 7: Árboles-B (B-Tree).

- 7.1 Introducción y definiciones.
- 7.2 Inserción de un árbol-B.
- 7.3 Algoritmo de inserción en un árbol-B.
- 7.4 Análisis del algoritmo de inserción.
- 7.5 Inserción en un árbol-B (Rotación-I/1).
- 7.6 Inserción en un árbol-B (Rotación-I/N).
- 7.7 Algoritmo de inserción en un árbol-B (Rotación-I/N).
- 7.8 Extracción en un árbol B.
- 7.9 Algoritmo de extracción en un árbol-B (Rotación-E/1, Recombinación-2/1).
- 7.10 Análisis del algoritmo de extracción.
- 7.11 Extracción en un árbol-B (Rotación-E/N).
- 7.12 Extracción en un árbol-B (Recombinación-3/2).
- 7.13 Algoritmo de extracción en un árbol-B (Rotación-E/N, Recombinación-3/2).

Tema 8: Arboles-B*.

- 8.1 Definición de árbol-B*.
- 8.2 Inserción de un árbol-B* (Partición-2/3).
- 8.3 Algoritmo de inserción en un árbol-B* (Partición-2/3).

Tema 9: Arboles-B⁺.

- 9.1 Introducción y definiciones.
- 9.2 Algoritmo de búsqueda en un árbol-B⁺.
- 9.3 Algoritmo de recuperación secuencial en un árbol-B⁺.
- 9.4 Algoritmo de recuperación secuencial a partir de un valor de clave en un árbol-B⁺.
- 9.5 Recuperación secuencial en rango en un árbol-B⁺.
- 9.6 Inserción en un árbol-B⁺.
- 9.7 Inserción en un árbol-B⁺ (Rotación-I/N+).
- 9.8 Inserción en un árbol-B⁺ (Partición-2/3+).
- 9.9 Algoritmo de inserción en un árbol-B⁺.
- 9.10 Extracción en un árbol-B⁺ (Rotación-E/N+).
- 9.11 Extracción en un árbol-B⁺ (Recombinación-2/1+).
- 9.12 Extracción en un árbol-B⁺ (Recombinación-3/2+).
- 9.13 Algoritmo de extracción en un árbol-B⁺.

Tema 10: Trie.

- 10.1 Introducción.
- 10.2 Búsqueda en trie.
- 10.3 Inserción en trie.
- 10.4 Extracción en trie.
- 10.5 Búsqueda e inserción en trie con factor de salto.
- 10.6 Extracción en trie con factor de salto.

Parte 2. Bases de datos relacionales.

Tema 11: Introducción a las bases de datos.

- 11.1 El procesamiento tradicional de datos (sistemas de ficheros).
- 11.2 Sistemas de bases de datos. Ventajas e inconvenientes.
- 11.3 Abstracción de la información e independencia de datos.
- 11.4 Modelos de datos.
- 11.5 Lenguajes de bases de datos.
- 11.6 Usuarios de un sistema de bases de datos.
- 11.7 Estructura general de un sistema gestor de bases de datos.
- 11.8 Arquitectura de aplicaciones.

Tema 12: El modelo relacional.

- 12.1 Introducción.
- 12.2 Estructura de las bases de datos relacionales. Definiciones y conceptos básicos. Notación.
- 12.3 Componentes del modelo relacional de datos.

Tema 13: Álgebra relacional.

- 13.1 Definición.
- 13.2 Operadores fundamentales (o esenciales o primitivos).
 - 13.2.1 Selección.
 - 13.2.2 Proyección
 - 13.2.3 Producto cartesiano.
 - 13.2.4 Unión
 - 13.2.5 Diferencia
- 13.3 Operadores derivados (no esenciales).
 - 13.3.1 Intersección.
 - 13.3.2 Yunción.
 - 13.3.3 Yunción natural.
 - 13.3.4 Cociente.
- 13.4 Modificación de la base de datos.
- 13.5 Operación de complemento.
- 13.6 Potencia expresiva del álgebra.
- 13.7 Propiedades de los operadores.
- 13.8 Valores nulos.
- 13.9 Vistas.
 - 13.9.1 Usos.
 - 13.9.2 Inconvenientes.
- 13.10 Integridad de claves primarias.
- 13.11 Integridad referencial.

Tema 14: Cálculo relacional.

- 14.1 Introducción.
- 14.2 Nociones básicas sobre lógica de predicados. Operadores y cuantificadores lógicos.
- 14.3 Definición informal al cálculo relacional de t-uplas.
- 14.4 Cálculo restringido de t-uplas.
- 14.5 Potencia expresiva del cálculo restringido de t-uplas.
- 14.6 Definición formal de cálculo relacional de t-uplas.
- 14.7 Cálculo relacional de dominios.

Tema 15: SQL interactivo.

- 15.1 Introducción.
- 15.2 Recuperación de datos.
 - 15.2.1 Cláusulas SELECT y FROM.
 - 15.2.2 Cláusula WHERE.
 - 15.2.3 Operaciones sobre conjuntos: Unión, intersección y diferencia.
 - 15.2.4 Funciones de agregación.
 - 15.2.5 Cláusulas GROUP BY y HAVING.
 - 15.2.6 Subconsultas (o consultas anidadas).
 - 15.2.6.1 Subconsultas que retornan un único valor.
 - 15.2.6.2 Subconsultas que retornan múltiples valores.

- 15.2.6.3 Subconsultas que retornan más de una columna.
- 15.2.6.4 Múltiples subconsultas.
- 15.2.6.5 Vistas temporales.
- 15.2.7 Cláusula ORDER BY.
- 15.3 Instrucciones para actualización de datos.
 - 15.3.1 Inserción.
 - 15.3.2 Modificación.
 - 15.3.3 Borrado.
- 15.4 Instrucciones de definición de datos.
 - 15.4.1 Tipos de dominios en SQL.
 - 15.4.2 Sentencias relacionadas con tablas.
 - 15.4.3 Sentencias relacionadas con vistas.
 - 15.4.4 Sentencias relacionadas con índices.
 - 15.4.5 Otras sentencias.

Tema 16: Control de integridad en SQL.

- 16.1 Restricciones de los dominios.
- 16.2 Declaración de claves.
- 16.3 Integridad de claves primarias.
- 16.4 Integridad referencial.
- 16.5 Aserciones.
- 16.6 Disparadores (triggers).

Tema 17: Seguridad y autorización en SQL.

- 17.1 Violaciones de la seguridad.
- 17.2 Control de acceso a la base de datos.
- 17.3 Tipos de autorización.
- 17.4 Autorizaciones y vistas.
- 17.5 Concesión de privilegios.
- 17.6 Eliminación de privilegios.
- 17.7 El concepto de rol o papel.
- 17.8 Limitaciones de la autorización SQL.

Tema 18: Control de transacciones en SQL.

- 18.1 Concepto de transacción.
- 18.2 Propiedades ACID.
- 18.3 Estados de una transacción.
- 18.4 Definición de transacciones en SQL.
- 18.5 Problemas del acceso concurrente.
- 18.6 Bloqueo de datos en SQL.
- 18.7 La situación de abrazo mortal.

Tema 19: SQL inmerso (o embebido) en C.

- 19.1 Introducción.
- 19.2 Sentencias ejecutables y no ejecutables.
- 19.3 Delimitadores.
- 19.4 Área de comunicación del SQL.
- 19.5 Variables huésped.
- 19.6 Representación de valores nulos.
- 19.7 Manipulación de datos sin usar cursores.
- 19.8 Programación con cursores.
- 19.9 SQL estático y dinámico.
- 19.10 SQL dinámico sin recuperación de datos.
- 19.11 SQL dinámico para consultas de lista fija.
- 19.12 SQL dinámico para consultas de lista variable.

Tema 20: Introducción al diseño de bases de datos.

- 20.1 Introducción.
- 20.2 Diseños equivalentes.
- 20.3 Descomposición reversible por yunción.

- 20.4 Dependencias funcionales.
 - 20.4.1 Sistema de inferencia para dependencias funcionales.
 - 20.4.2 Tipos de dependencias funcionales.
 - 20.4.3 Cierre de un conjunto de dependencias funcionales.
 - 20.4.4 Cierre de un conjunto de atributos.
 - 20.4.5 Recubrimiento canónico de un conjunto de dependencias funcionales.
 - 20.4.6 Propagación de las dependencias funcionales al descomponer.
 - 20.4.7 Anomalías de actualización asociadas a dependencias funcionales.
- 20.5 Forma normal de Boyce-Codd (FNBC).
- 20.6 Tercera forma normal (3FN).
- 20.7 Segunda forma normal (2FN).
- 20.8 Primera forma normal (1FN).
- 20.9 Dependencias plurales.
 - 20.9.1 Sistema de inferencia para dependencias plurales.
 - 20.9.2 Propagación de las dependencias plurales al descomponer.
 - 20.9.3 Anomalías de actualización asociadas a dependencias plurales.
- 20.10 Cuarta forma normal (4FN).
- 20.11 Otras formas normales.
- 20.12 Proceso general del diseño de bases de datos.

Tema 21: ORACLE: un sistema gestor de bases de datos relacionales.

- 21.1 Antecedentes históricos.
- 21.2 Estructura de la base de datos ORACLE.
- 21.3 Organización del diccionario de datos.
- 21.4 El procesador de consultas interactivo: SQLPLUS.
- 21.5 El precompilador de SQL inmerso en C: PRO*C.
- 21.6 Herramientas de Oracle.

Prácticas:

Parte 1.

- Práctica Nº 1. Entorno de programación en lenguaje C++.
- Práctica Nº 2. Tipos Abstractos de datos.
- Práctica Nº 3. Operaciones básicas con ficheros.
- Práctica Nº 4. Inserción de registros de longitud fija.
- Práctica Nº 5. Inserción de registros de longitud variable.
- Práctica Nº 6. Eliminación de registros.
- Práctica Nº 7. Índice primario
- Práctica Nº 8. Índice secundario.
- Práctica Nº 9. Árboles multicamino.
- Práctica Nº 10. Árboles-B.
- Práctica Nº 11. Árboles-B*.
- Práctica Nº 12. Árboles-B⁺.
- Práctica Nº 13. Trie.

Parte 2.

- Práctica Nº 14. Toma de contacto con el sistema gestor: SQLPLUS, alta de usuarios y creación de la base de datos.
- Práctica Nº 15. Consultas elementales: condiciones lógicas en cláusula WHERE.
- Práctica Nº 16. Consultas con yunción natural y subconsultas.
- Práctica Nº 17. Consultas avanzadas: agrupamientos y funciones de grupo.
- Práctica Nº 18. Control de transacciones. Causas frecuentes de fallo.
- Práctica Nº 19. Definición de datos: modificación del esquema de la base de datos.
- Práctica Nº 20. Autorizaciones y bloqueos explícitos de tablas
- Práctica Nº 21. Estructura y uso del diccionario de la base de datos.
- Práctica Nº 22. Control de integridad: disparadores.
- Práctica Nº 23. SQL inmerso en C/C++.
- Práctica Nº 24. Desarrollo de aplicaciones utilizando SQL estático.
- Práctica Nº 25. Desarrollo de aplicaciones utilizando SQL dinámico.

BIBLIOGRAFÍA

- [1] M.J. FOLK. Y B. ZOELICK, "*Estructuras de Archivos*", Addison–Wesley Iberoamericana (1992).
- [2] R.H. AUSTING Y L.N. CASSEL, "*Gestión de Ficheros*", Anaya Multimedia (1988).
- [3] D.E. KNUTH, "*El Arte de Programar Ordenadores. Clasificación y Búsqueda*", Vol. III, Ed. Reverté (1987).
- [4] Apuntes de programación de la E.U. de Informática de la U.L.P.G.C.
- [5] M.A. WEISS, "*Estructuras de Datos y Algoritmos*", Addison–Wesley Iberoamericana (1995).
- [6] R. SEDGEWICK., "*Algorithms in C++*", Addison–Wesley (1992).
- [7] A.V. AHO, J.E. HOPCROFT, Y J.D. ULLMAN, "*The Design and Analysis of Computer Algorithms*", Addison-Wesley (1995).
- [8] Artículos en revistas especializadas.
- [9] J. MARTIN, "*Organización de las Bases de Datos*", Prentice Hall International.
- [10] G. WIEDERHOLD, "*Diseño de Bases de Datos*", McGraw-Hill.
- [11] ELMASRI Y NAVATHE, "*Fundamentals of Database Systems*", The Benjamin/Cummings Pub. Co. Inc., 2ª ed. (1994).
- [12] J. BENAVIDES, J. OLAIZOLA Y E. RIVERO CORNELIO, E., "*SQL Para Usuarios y Programadores*", Paraninfo, (1991).
- [13] SILBERSCHATZ, H. KORTH Y S. SUDARSHAN, "*Fundamentos de Bases de Datos*", McGraw-Hill, Cuarta edición, (2002).
- [14] E. RIVERO CORNELIO, "*Bases de Datos Relacionales*", Paraninfo, (1988).

BREVE DESCRIPCIÓN DE LA ASIGNATURA

En la primera parte de esta asignatura se estudian las técnicas de almacenamiento y recuperación de la información organizados en estructuras de datos que residen en memoria secundaria. Estas técnicas constituyen el nivel físico de un sistema gestor de bases de datos.

Se plantea al alumno el problema del almacenamiento de grandes cantidades de información y la conveniencia de encontrar estructuras de ficheros eficientes para el mantenimiento y consulta de los datos.

Se estudian distintas organizaciones de archivos mediante registros de longitud fija y longitud variable. Se evalúan las técnicas de reutilización del espacio liberado al eliminar registros. Se plantea la indexación como técnica para acelerar la búsqueda de registros y se estudian las estructuras de datos arbóreas para su implementación.

La segunda parte de la asignatura presenta inicialmente los conceptos fundamentales del modelo relacional de datos, tratando la relevancia del álgebra relacional y cálculo relacional para, a continuación, realizar un estudio en detalle del lenguaje SQL así como del desarrollo de programas de aplicaciones en SQL inmerso.

Por último, se realiza una introducción a la normalización de bases de datos relacionales basada en la teoría de las dependencias funcionales y plurales, con énfasis en la motivación y el significado intuitivo de cada forma normal.

OBJETIVOS DOCENTES

- Conocer y evaluar las características de las estructuras de datos.
- Conocer la organización física de los sistemas de almacenamiento secundario.
- Analizar los requerimientos de espacio y tiempo de acceso para almacenar un esquema de datos.

- Comprensión de las características del modelo relacional (Tablas, t-uplas, condiciones de integridad, ...).
- Aprendizaje y uso de los principales lenguajes de consulta (teóricos y comerciales) para el modelo relacional (Álgebra relacional, Cálculo relacional de t-uplas y de dominios, SQL-3)
- Programación de aplicaciones con SQL inmerso en C.
- Diseño de una base de datos relacional.

MÉTODOS DE EVALUACIÓN

La evaluación se realizará a partir de las siguientes pruebas individuales:

1. Examen de teoría.
2. Prácticas de laboratorio.

La Laguna, 16 de Junio de 2003

Los Profesores

Director del Departamento

Fdo.: Jesús Alberto González Martínez

Fdo.: Dr. Carlos González Alcón

Fdo: Dr. Félix César García López

Fdo: Luz Marina Moreno de Antonio

Fdo: Casiano Rodríguez León

Fdo: Jesús M. Jorge Santiso