FACULTAD DE INGENIERIA
Universidad de Buenos Aires

# Planificaciones

## 7506 - Organización de Datos

Docente responsable: SERVETTO ARTURO CARLOS

## OBJETIVOS

Que el alumno:
Interprete los conceptos fundamentales que sustentan la estructuración de la Información para su almacenamiento/recuperación.
Comprenda las fortalezas y debilidades de cada técnica.
Sea capaz de evaluar la adecuación de una determinada técnica con aplicación a casos concretos.
Sea capaz de inferir por sí mismo nuevas soluciones.

## CONTENIDOS MÍNIMOS

-

## PROGRAMA SINTÉTICO

Sistemas y organización de archivos.
Recuperación de textos.
Compresión y archivos multimedia.
Criptografía.

## PROGRAMA ANALÍTICO

Sistemas y organización de archivos: directorios y archivos; organización lógica de discos; seguridad y auditoría de accesos; buffering;estudio de casos (FAT32,NTFS, EXT4, HFS); organizaciones básicas de archivo y primitivas (archivos secuenciales -desordenados y ordenados-, ordenamiento externo de archivos secuenciales, archivos B, B*, B+ y B#, archivos directos estáticos y dinámicos; organizaciones generales (secuencial indexada e indexada); tipos de índice y resolución de consultas con múltiples índices.

Recuperación de textos: índices invertidos booleanos y vectoriales, construcción de índices; compresión de léxico y de listas invertidas; firmas e índices de porciones de firmas; índices especiales (de n-gramas, de léxico rotado, para frases y términos próximos, de espacios métricos para términos parecidos).

Compresión y archivos multimedia: introducción a la teoría de la información; compresión estadística (Huffman estático y dinámico; manejo eficiente de Huffman dinámico; Half Coding; compresión aritmética; aritmética de contextos; PPMC); compresión no estadística (Run Length, LZ77, LZ78, LZHUFF, LZP); localidad en archivos (BS + MTF + SHANNON / Estructurado); métodos de compresión para datos multimedia (algoritmos JPEG para imágenes y MPEG para video).

Criptografía: introducción; objetivos; criptografía clásica; criptgrafia de clave privada; introducción a la criptografía asimétrica; criptografia de clave pública; firmal digital; certificados digitales; Pretty Good Privacy (PGP).

## BIBLIOGRAFÍA

[Ian Witten, Alistair Moffat, Timothy Bell] Managing Gigabytes 2nd edition
Website: http://www.cs.mu.oz.au/mg/
1999; Morgan Kaufmann Publishers; ISBN: 1558605703 ; 576 paginas;
Temas: Compresion de datos (Algunos metodos), Huffword, Indices, Indices invertidos, almacenamiento y construccion de indices invertidos, front-coding, hashing perfecto, resolucion de consultas, metodo del coseno, compresion de imagenes.

[Simon St Laurent] XML a primer3rd edition
2001;Hungry Minds, Inc; ISBN: 0764547771 ; 500 paginas;
Temas: XML

[Timothy Bell, Ian Witten, John Clearly] Text compression
1990;Prentice Hall; ISBN: 0139119914 ; 318 paginas;
Temas: Compresion de datos, compresion estadistica, Huffman, metodos dinamicos, compresion aritmetica, PPM, modelos de PPM de orden superior. LZ77, LZ78, otros metodos de compresion de datos.

[Michael Folk, Bill Zoellick, Greg Ricciardi] File structures: An object oriented approach with C++
Website: http://cseng.aw.com/book/0,3828,0201874016,00.html
1998; Addison-Wesley; ISBN: 0-201-87401-6; 724 paginas;
Temas: Organizacion de archivos, operaciones entre archivos, sort, arboles B, B*, B+ paginado, hashing, hashing extensible, encapsulamiento de distintas organizaciones.

[Bruce Schneier] Applied Cryptography 2nd edition

Website: http://www.counterpane.com/applied.html
1995;John Wi

## RÉGIMEN DE CURSADA

### Metodología de enseñanza

Clases teórico-prácticas: exposición teórica de conceptos fundamentales, con resolución metódica de problemas tipo y ensayos sobre objetivos.
Clases prácticas: resolución por parte de los alumnos y controlada por los docentes auxiliares de problemas correspondientes a las unidades temáticas del programa, ya sea por escrito o por máquina (programas). En general se tratará de problemas abiertos, que generen dudas y motiven la consulta a los docentes y la profundización del conocimiento a través de la bibliografía.

### Modalidad de Evaluación Parcial

Evaluación de manejo de conceptos, aplicación de conocimientos y dominio de técnicas, mediante la respuesta a preguntas y la resolución de problemas por escrito en evaluaciones parciales e integradoras, y el desarrollo controlado de trabajos prácticos en computadora. Las evaluaciones parciales e integradoras son por unidades o subunidades temáticas. La evaluación de los trabajos por computadora es por presentación en tiempo y forma (plazos y formato establecido), método de desarrollo (aplicación de método de desarrollo de programas visto en el curso) y corrección del resultado (cumplimiento de objetivos del programa).

## CALENDARIO DE CLASES

| Semana | Temas de teoría | Resolución de problemas | Laboratorio | Otro tipo | Fecha entrega Informe TP | Bibliografía básica |
|---|---|---|---|---|---|---|
| <1> 11/03 al 16/03 | Sistemas de archivo | | | casos de estudio | | |
| <2> 18/03 al 23/03 | Archivos secuenciales ordenados y desordenados | Ordenamiento externo de archivos | | | | |
| <3> 25/03 al 30/03 | Archivos B, B*, B+ y B# | Altas y bajas en archivos B, B*, B+ y B# | | | | |
| <4> 01/04 al 06/04 | Archivos directos estáticos y dinámicos | Operaciones con dispersión lineal y doble y con áreas de desborde independientes | | | | |
| <5> 08/04 al 13/04 | Índices y organizaciones secuencial indexada e indexada | Operaciones con dispersión extensible | | | | |
| <6> 15/04 al 20/04 | Índices invertidos booleanos y vectoriales | Construcción de índices | | | | |
| <7> 22/04 al 27/04 | Compresión del léxico y de listas invertidas | Front coding parcial y compresión gamma, delta y golomb | | | | |
| <8> 29/04 al 04/05 | Firmas, índices de porciones de firmas e índices especiales | Construcción de índices para frases y términos próximos, de n-gramas, de léxico rotado y de porciones de firmas | | | | |
| <9> 06/05 al 11/05 | Introducción a la teoría de la información y compresión estadística | Huffman estático y dinámico; manejo eficiente de Huffman dinámico | | | Primera entrega de trabajo práctico grupal | |
| <10> 13/05 al 18/05 | Compresión estadística | Half Coding; compresión aritmética; aritmética de contextos; PPMC | | | | |
| <11> 20/05 al 25/05 | Compresión no estadística | Run Length, LZ77, LZ78, LZHUFF, LZP | | | | |
| <12> 27/05 al 01/06 | Localidad en archivos | BS + MTF + SHANNON / Estructurado | | | | |
| <13> 03/06 al 08/06 | Compresión multimedia | Algoritmos JPEG para imágenes y MPEG para video | | | | |
| <14> 10/06 al 15/06 | Criptografía clásica; | criptgrafia de clave privada | | | | |

| Semana | Temas de teoría | Resolución de problemas | Laboratorio | Otro tipo | Fecha entrega Informe TP | Bibliografía básica |
|---|---|---|---|---|---|---|
| | criptgrafia de clave privada | | | | | |
| <15> 17/06 al 22/06 | ntroducción a la criptografía asimétrica; criptografia de clave pública | criptografia de clave pública | | | | |
| <16> 24/06 al 29/06 | firmal digital; certificados digitales; Pretty Good Privacy | | | | Segunda entrega de trabajo práctico grupal | |

# CALENDARIO DE EVALUACIONES

## Evaluación Parcial

| Oportunidad | Semana | Fecha | Hora | Aula |
|---|---|---|---|---|
| 1º | 6 | 19/04 | 19:00 | |
| 2º | 9 | 10/05 | 19:00 | |
| 3º | 14 | 14/06 | 19:00 | |
| 4º | | 05/07 | 19:00 | |

| Observaciones sobre el Temario de la Evaluación Parcial |
|---|
| Se toma una evaluación parcial obligatoria de cada unidad del programa durante el cursado (ver calendario), y en fechas de evaluación integradora se puede rendir cada unidad en 2ª y 3ª oportunidad. |
| Otras observaciones |
| La evaluación de Criptografía es en la semana 17. |