

# Estructuras de Datos y de la Información II

## Examen final – 16 de junio 2008

1730

**Observación:** Las preguntas se deberán entregar en tres bloques:

- Bloque 1: preguntas 1 y 2.
- Bloque 2: preguntas 3 y 4.
- Bloque 3: preguntas 5 y 6.

### Bloque 1

2 puntos

- 1) Escribir el pseudocódigo de un algoritmo de complejidad mínima basado en DFS que eliminando el mínimo n° posible de arcos, suprima los ciclos de un grafo dirigido (suponer que el grafo está representado por una lista de adyacencia). Dar una justificación teórica del correcto funcionamiento algoritmo, basada en la clasificación de arcos y el Teorema del camino blanco.

1.5 puntos

- 2) Se dispone de un fichero con 8.000.000 registros de 100 bytes cada uno, y un espacio en RAM en el que caben 10.000 registros. Supóngase que el orden de los registros es aleatorio. Se desea ordenar los registros del fichero mediante un procedimiento de selección con reemplazo y *merge* en dos pasos. Supóngase que el procedimiento de selección con reemplazo utiliza un buffer I/O de 2.500 registros situado en la memoria RAM disponible.
- Estimar el número accesos (lectura y escritura) de la fase de ordenación, y el tamaño promedio de los trozos ordenados resultantes, así como su número. Razona todos los resultados (0.5 puntos).
  - Estimar el número de accesos de lectura en la fase *merge* en dos pasos, suponiendo que para el primer *merge* se hacen agrupaciones de 28 ficheros auxiliares (1 punto).

### Bloque 2

1.5 puntos

- 3) Una base de datos de coches privados contiene para cada coche los datos siguientes: marca, modelo, número de identificación (único), matrícula, DNI del propietario, año de primera matriculación, color.
- ¿Cuáles de estos campos se pueden utilizar como clave primaria? Justificar la respuesta.
  - Dada la base de datos siguiente:

pos.	marca	modelo	núm. ident.	matrícula	DNI prop.	año	color
100	FIAT	punto	5781 —	4973 ES	113678	2004	blanco
247	renault	C4	3347 <	9011 BA	980252	2006	rojo
312	FIAT	uno	9001 —	3999 CD	115673	1995	blanco
379	FIAT	uno	2854 —	0083 FG	889265	1997	blanco
422	renault	C4	3333 · <	8212 BA	187445	2004	azul
501	dodge —	stratus	1035	3999 BA	980252	2007	negro
568	FIAT	punto	0555 · —	9011 ES	490742	2004	blanco

crear un índice primario utilizando uno de los campos posibles como clave primaria y un índice secundario (sin listas invertidas) por marca.



1.5 puntos

- 4) El *diámetro* de un grafo dirigido no ponderado es la máxima distancia (en número de arcos) que se puede dar entre dos vértices del grafo. De un grafo  $G = (V, E)$  se sabe que tiene un diámetro  $D < |V|$ . Utilizando esta información adicional, modificar el algoritmo de Bellman-Ford no optimizado de manera tal que se reduzca su tiempo de ejecución. (El tiempo de ejecución del algoritmo es  $O(VE)$ ).

↑ nos dan el diámetro!

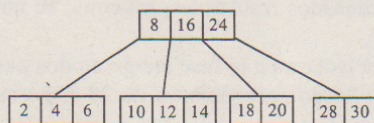
### Bloque 3

1.5 puntos

- 5) En la base de datos del problema 3:
- Suponiendo que la gama de colores incluye: *blanco*, *negro*, *gris*, *rojo*, *azul*, *verde*, *amarillo*, *naranja*, *marrón*, razonar cuál de los siguientes métodos de compresión: notación compacta, run-length, método de Huffman, sería más adecuado para comprimir el campo *color*. Suponiendo que el campo sin comprimir se almacena con longitud fija, ¿cuánto espacio en disco se ahorraría por registro con la compresión (suponiendo 1 byte por carácter)? (0.5 puntos).
  - Mostrar los códigos que se obtienen para cada valor de la cadena de caracteres "abyabcczabaabac" con el método de Huffman. (1 punto).

2 puntos

- 6) Dado el siguiente árbol B de 3 claves por página:



- Insertar la clave 3 (0.5 puntos).
- Partiendo del resultado del apartado anterior, eliminar las claves 3 y 6 (0.5 puntos).
- Partiendo del resultado del apartado anterior, eliminar la clave 2 (0.5 puntos).
- Volviendo al estado inicial, repetir el apartado a) suponiendo que se trata de un árbol  $B^*$  (0.5 puntos).

Nota: Poner solo el resultado final de la operación en cada apartado.