

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
FACULTAD DE CIENCIAS E INGENIERÍA

ALGORITMIA
Laboratorio 4
(Primer semestre 2016)

Indicaciones generales:

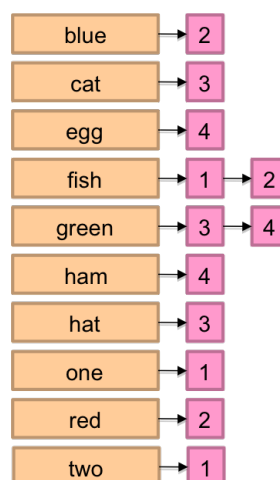
- Duración: 2h 55min.
- Materiales o equipos a utilizar: Apuntes de clase personales escritos a mano.
- Al inicio de cada programa, el alumno deberá incluir, a modo de comentario, la estrategia que utilizará para resolver el problema. De no incluirse dicho comentario o si la implementación es significativamente diferente a la estrategia indicada, el alumno no obtendrá el puntaje completo en dicha pregunta.
- Un programa que no muestre resultados coherentes y/o útiles será corregido sobre el 60 % del puntaje asignado a dicha pregunta.
- **Debe utilizar comentarios para explicar la lógica seguida en el programa elaborado.**
- El orden será considerado dentro de la evaluación.
- Su trabajo deberá ser subido a PAIDEIA en el espacio indicado por los Jefes de Práctica.
- **La presentación, la ortografía y la gramática de los trabajos influirá en la calificación.**

Puntaje total: 20 puntos

Cuestionario:

En Recuperación de Información, el **índice invertido** de una colección de documentos es básicamente una estructura de datos que vincula cada término único en la colección con la lista de todos los documentos que contienen el término. Por ejemplo:

- **Doc1:** one fish two fish
- **Doc2:** red fish blue fish
- **Doc3:** cat in the green hat
- **Doc4:** green eggs and ham



Luego, se pueden hacer consultas a la estructura del tipo:

- *blue AND fish*: se obtienen todos los documentos que contienen la palabra *blue* y la palabra *fish*.
- *red OR green*: se obtienen todos los documentos que contienen la palabra *red* o la palabra *green*.

Pregunta 1 (3 puntos)

Declare las estructuras y tipos de datos necesarios para implementar el TAD Índice Invertido.

NOTA: Si la declaración de estructuras y tipos de datos es incorrecta y/o no corresponde con lo mostrado, solo podrá obtener como máximo la mitad del puntaje en las preguntas siguientes.

Pregunta 2 (10 puntos)

Agregue a su TAD Índice Invertido la función ***CrearIndice*** que reciba como parámetro (entre otros que considere necesarios) un arreglo que contiene las rutas de diferentes documentos. Utilizando estos documentos, su función debe crear un Índice Invertido. Para facilitar el proceso, asuma que todos estos documentos contienen una palabra por cada línea.

Luego de implementar su función, cree un pequeño programa para probar su funcionamiento. Para ello, deberá invocar a la función ***CrearIndice*** y luego imprimir el índice creado. La función que imprime el índice debe ser parte del TAD.

IMPORTANTE: Observe que las palabras en el índice están ordenadas, y que los documentos dentro de cada lista también están ordenados.

Pregunta 3 (3.5 puntos)

Agregue a su TAD Índice Invertido la función ***ProcesarAND*** que reciba como parámetros (entre otros que considere necesarios) dos palabras. Para ello siga los siguientes pasos:

- Ubicar la palabra 1 en el índice y obtener su lista de documentos.
- Ubicar la palabra 2 en el índice y obtener su lista de documentos.
- Crear una tercera lista que es la “intersección” de las dos listas anteriores. Este proceso debe ser $O(x + y)$, donde x e y son las longitudes de las listas.
- A partir de la lista “intersección”, imprimir las **rutas** de los documentos.

Pruebe que su implementación funciona correctamente desde el programa principal.

Pregunta 4 (3.5 puntos)

Agregue a su TAD Índice Invertido la función ***ProcesarOR*** que reciba como parámetros (entre otros que considere necesarios) dos palabras. El proceso es similar a ***ProcesarAND***, solo que en este caso debe obtener la “unión” y no la “intersección”. Pruebe que su implementación funciona correctamente desde el programa principal.

Profesores del curso: Fernando Alva
Iván Sipirán

Pando, 24 de mayo de 2016