

Juan Sebastián Sánchez Delgado

Preguntas previas

a) Teniendo en cuenta cada uno de los requerimientos ¿Cuántos índices implementaría en el Reto? y ¿Por qué?

R/: En principio utilizaría 2 índices, uno con los medios de la obras y otro con las nacionalidad de los artistas. Esto para facilitar la implementación de los requisitos 3 y 4, ya que así se evitaría hacer dobles recorridos u ordenar los arreglos para posteriormente realizar una búsqueda binaria. En otras palabras, el utilizar estos índices hace que la complejidad temporal disminuya considerablemente.

b) Según los índices propuestos ¿en qué caso usaría Linear Probing o Separate Chaining en estos índices? y ¿Por qué?

R/: Creo que lo mas conveniente sería utilizar Separate Chaining para ambos casos, puesto que su factor de carga recomendado para un buen desempeño es de 1, en lugar del 0.5 como máximo sugerido para Linear Probing. Puede parecer un detalle sin importancia, pero el tener un factor de carga de 0.5 implica tener una tabla con el doble de parejas a almacenar, lo que en ultimas implica un mayor uso de memoria.

Por otra parte, Separate Chaining puede tener una tabla con una cantidad ligeramente menor tamaño, de tal forma que la búsqueda de elementos sigue siendo eficiente y ocupando menos espacio en memoria que su contraparte.

c) Dado el número de elementos de los archivos MoMA, ¿Cuál sería el factor de carga para estos índices según su mecanismo de colisión?

R/: Como se menciona en el numeral anterior, si el mecanismo es Separate Chaining, idealmente el factor de carga debe ser 1 o por lo menos un numero cercano a este, para así mantener un equilibrio entre la memoria a usar y la complejidad temporal de búsqueda. Por otro lado, para el caso de Linear Probing se recomienda que su factor de carga sea menor a 0.5, debido a que de esta manera se evita un crecimiento desmedido de la complejidad temporal para almacenar nuevas parejas.

Paso 3

d) ¿Qué diferencias en el tiempo de ejecución notan al ejecutar la cargar los datos al cambiar la configuración de Linear Probing a Separate Chaining?

R/: Si se utilizan los archivos de extensión “large”, la implementación de la función con “Medium” con Probing se demora 6 segundos menos que la hecha con Chaining, también existe una diferencia de aproximadamente un segundo para el caso de la función de “Nationality”. En conclusión, la implementación que en resulto ser mas eficiente en todos los casos fue la de Probing.

Paso 4

e) ¿Qué configuración de ADT Map escogería para el índice de técnicas o medios?, especifique el mecanismo de colisión, el factor de carga y el numero inicial de elementos.

R/: La variante que mejores resultados arrojo fue la implementación con Probing y factor de carga de 0.2 (11.36seg), sin embargo, si se quisiera mejorar su rendimiento será necesario tener una tabla de hash de 5 veces mas grande que la cantidad de parejas, por lo que se ocuparía demasiada memoria. Personalmente, escogería un map con Chaining, un factor de carga de 1 y un tamaño de 60 o 70 elementos.

f) ¿Qué configuración de ADT Map escogería para el índice de nacionalidades?, especifique el mecanismo de colisión, el factor de carga y el numero inicial de elementos

R/: Por los motivos explicados anteriormente, a pesar de que en las pruebas el que reporto un mejor resultado fue una implementación con probing, la configuración que usarria seria la misma del caso anterior, es decir, escogería un map con Chaining, un factor de carga de 1 y un tamaño de 80 elementos.