

Daniel Dorado Toro 201821010

Juan Camilo Neira Campos: 201922746.

Laboratorio 06.

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

Implementaremos cinco índices, es decir, uno por cada requerimiento. Un índice por año de nacimiento, año de adquisición, medio, nacionalidad y departamento. Esto se debe a que así agrupamos los datos y de tal forma retornar una pareja (key, value) es  $O(1)$ . En este caso, esta información es una referencia a los datos ya almacenados.

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

En el caso de la fecha de adquisición usaremos Chaining, ya que en caso de ser Probing, el tamaño de la lista tendría que ser el doble del número de fechas de adquisición, lo cual podría utilizar un espacio significativo de memoria principal. Esto mismo aplica para los años de nacimiento, ya que en Chaining

En el caso de nacionalidad vamos a implementar Probing, ya que hay un aproximado de 193 nacionalidad en todo el mundo, y, dado que el factor de carga es de 0.5, el tamaño del maps es de aproximadamente de 386, lo cual no consume un espacio significativo en memoria y sí hace que el requerimiento sea veloz. Este mismo razonamiento aplica para el índice de departamentos, ya que hay un número reducido de departamento y por tanto utilizar Probing no consumiría mucha memoria y sí sería veloz.

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

Para Probing utilizaremos un factor de carga de 0.5 y para Chaining un factor de carga de 2.

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

Como se puede ver en las tablas, con linear probing la carga es más rápida aproximadamente 7 segundos.

**¿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.**

Para el índice de técnicas o medios, usaríamos linear probing con un factor de carga de 0.8 y un tamaño inicial de 1291.

**¿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.**

Para el índice de técnicas o medios usaría separate chaining con un factor de carga de 2.0 y un tamaño inicial de 101, el número primo más cercano a 100 (asumiendo que hay cerca de 200 nacionalidades). Como el conjunto de posibles nacionalidades no es tan grande, esperamos que no ocupe demasiada memoria.

	<b>Separate Chaining</b>
<b>Factor de carga = 2.0</b>	85103.4
<b>Factor de carga = 4.0</b>	96842.8
<b>Factor de carga = 8.0</b>	123788.9

Tabla. Tiempo para cargar datos 50pct.

	<b>Linear Probing</b>
<b>Factor de carga = 0.2</b>	99711.8
<b>Factor de carga = 0.5</b>	89402.1
<b>Factor de carga = 0.8</b>	86864.2

Tabla. Tiempo para cargar datos 50pct.

Para encontrar el tamaño de la tabla M para el respectivo factor de carga se hizo la suposición de que hay un N de 1000 medios y 120 países.