

LABORATORIO No. 6: Mecanismos de Colisión

Estudiante 1: Sofia Torres Ramírez.

Código: 202014872

Correo: s.torres21@uniandes.edu.co

Estudiante 2: Ana Margarita Florez Ruiz

Código: 201922242

Correo: a.florezr@uniandes.edu.co

Paso 1: Responder preguntas preparatorias

En esta parte consideraremos el uso apropiado de los TAD Map respondiendo las siguientes preguntas:

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

Req 1: Se crearía un índice llamado “yearartist” el cual será un map, las llaves serán los años y los valores serán listas con los "ConstituentID" de los artistas que nacieron en dicho año para luego buscarlos en el catálogo de artists.

Req 2: Se crearía un índice llamado “yearartworks” el cual será un map, las llaves serán los años y los valores serán otro map las llaves de ese map serán cada mes en ese año, el valor de ese map será otro map, las llaves de ese map serán los días y los valores serán los "ObjectID" de cada obra que fueron hechas en ese día de ese mes de ese año para luego buscarlas en el catálogo de artworks.

Req 3: Se crearía un índice llamado “constituent ID” El cual será un map, las llaves serán los ID de los artistas y el valor será otro map donde las llaves serán las técnicas y cada valor de dichas técnicas sería una lista con los "ObjectID" que usaron esta técnica para luego buscarlas en el catálogo de artworks.

Req 4: Se crearía un índice llamado “Nationality” que será un map donde las llaves serán las nacionalidades de los artistas y el valor de este sería una lista con los "ObjectID" de cada una de las obras con esta nacionalidad para luego buscarlas en el catálogo de artworks.

Req 5: Se crearía un índice llamado “Department” donde cada llave será cada departamento del museo y cada valor será una lista con el "ObjectID" de las obras de dicho departamento para luego buscarlas en el catálogo.

Req 6:

Criterio 1: Se crearía un índice: “obrasartista” en donde cada llave será el artista y el valor serán las obras.

Criterio 2: Se crearía un índice: “mediosartista” en donde cada llave será el artista y el valor serán los medios.

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

Req 1: “Linear Probing” debido a que se requiere de varias consultas en el índice “yearartist” y se necesita velocidad en este caso

Req 2: “Linear probing” al índice de años ya que este requiere más consultas, los demás maps (de meses y días) serán “separate chaining” debido a que estos no requieren tantas consultas.

Req 3: El índice de constituent ID sería “separate chaining” debido a que este se utiliza pocas veces, en cambio el índice de técnicas sería “Linear probing” debido a que, que se requiere retornar varios datos de las listas de obras.

Req 4: “Separate chaining” debido a que el índice de nacionalidades no requiere tantas consultas.

Req 5: “Separate chaining” porque con un solo llamado al índice ya se pueden tomar las obras de un departamento y hacer las consultas necesarias.

Req 6: “Separate chaining” en ambos índices debido a que solo se necesita el total de obras y de medios

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

Req 1: “yearartist”: $\frac{237}{239} = 0,99$

Req 2: “yearartworks”: $\frac{425}{431} = 0,98$

Map de meses: $\frac{13}{12} = 1,083$

Map de días: $\frac{31}{31} = 1$

Req 3: “constituent ID”: $\frac{15173}{15166} = 1,004$

Map de técnicas: $\frac{21191}{21191} = 1$

Req 4: “Nationality”: $\frac{1227}{119} = 1,067$

Req 5: “Department”: $\frac{11}{9} = 1,22$

Req 6: “obrasartista”: 0,75
“mediosartista”: 0,75

Paso 3: Separate Chaining o Linear Probing

Pruebas de tiempo de ejecución

	Máquina
Procesadores	Intel(R) Core(TM) i5-10210U CPU @ 1.60GHz 2.11 GHz
Memoria RAM (GB)	8,00 GB
Sistema Operativo	Sistema operativo de 64 bits, procesador basado en x64

Modificaciones de la estructura de datos y factor de carga	Tiempos	
	Índice [“Medium”]	Índice [“Nationality”]
“CHAINING” – 4.00	123.574	5.022
“PROBING” – 0,50	75.677	2.398

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

Se puede notar una pequeña diferencia entre las dos modificaciones ejecutadas, llegando a la conclusión de que las funciones corren más rápido al utilizar la estructura de datos “PROBING” con un factor de carga de 0,5.

Paso 4: Modificar el factor de carga

Pruebas de tiempo de ejecución

Modificaciones de la estructura de datos y factor de carga	Tiempos	
	Índice [“Medium”]	Índice [“Nationality”]
“CHAINING” – 2,00	67.613	5.324
“CHAINING” – 8,00	26.327	2.585
“PROBING” – 0,20	58.983	5.09
“PROBING” – 0,80	59.028	2.505

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

Para el índice de medios escogería el mecanismo de colisión “Separate Chaining” con un factor de carga de 8 y un número inicial de elementos de 21191.

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

Para el índice de nacionalidades se podría escoger tanto el mecanismo de colisión “Separate Chaining” con un factor de carga de 8 como el mecanismo de colisión “Linear probing” con un factor de carga de 0,8 y ambos con un número inicial de elementos de 119.