

## Observaciones lab - 6

Daniel R Alonso A - 201419873 - [dr.alonso10@uniandes.edu.co](mailto:dr.alonso10@uniandes.edu.co)

Nicolás Díaz M - 202021006 - [n.diaz9@uniandes.edu.co](mailto:n.diaz9@uniandes.edu.co)

- a) Teniendo en cuenta cada uno de los requerimientos ¿Cuántos índices implementaría en el Reto? y ¿Por qué?
- b) Según los índices propuestos ¿en qué caso usaría Linear Probing o Separate Chaining en estos índices? y ¿Por qué?
- c) Dado el número de elementos de los archivos MoMA, ¿Cuál sería el factor de carga para estos índices según su mecanismo de colisión?
- d) ¿Qué diferencias en el tiempo de ejecución notan al ejecutar la carga de los datos al cambiar la configuración de Linear Probing a Separate Chaining?
- 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 número inicial de elementos.
- f) ¿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.

## Respuestas:

- a) Viendo los requerimientos solo se necesitarían 5 índices: uno para nacionalidades; uno para medios; uno para departamentos; uno para fechas de las obras y uno para el ID de los artistas. Debido a que se volvería más eficiente buscar las obras o artistas que está pidiendo el programa si estos ya están guardados en un diccionario. Entonces por ejemplo si quisiera buscar todas las obras que contengan la nacionalidad "Colombian" entonces el programa solo debería encontrar la llave "Colombian" y adentro se encontraría todas las obras con esa nacionalidad.
- b) Pensaría que los datos que contienen más elementos como los son los artistas o las fechas de las obras serían se utilizaría linear probing por lo que se demoraría menos la carga de datos, mientras que para los demás índices se utilizaría linear chaining para el resto de índices para ahorrar espacio.
- c) Factores de carga:
  - IDs de Artistas: 0.5
  - Fechas de Obras: 0.5
  - Departamentos: lo más cercano a 1
  - Medios: lo más cercano a 1
  - Nacionalidades: lo más cercano a 1
- d) Se pudo notar que ejecutar la carga de datos por medio de Linear Probing era un poco más rápida que por Separate Chaining, siendo un promedio de 470.74 milisegundos más rápida en los archivos "large".

CHAINING	
Factor de carga	Carga de Datos
4	10309,09
2	9679,866
8	10255,85
PROBING	
Factor de carga	Carga de Datos
0,5	9541,14
0,2	9782,3
0,8	9509,16

- e) Para la carga del índice de técnicas o medios escogeríamos la configuración de Separate Chaining, con un factor de carga de 2.00 y un numero inicial de elementos de 10000.
- f) Para la carga del índice de nacionalidades escogeríamos la configuración de Linear Probing, con un factor de carga de 0.8 y un numero inicial de elementos de 1000.