

UNIVERSIDAD DE JAÉN
ESCUELA POLITÉCNICA SUPERIOR DE JAÉN



Práctica 7:
Tabla Hash Cerrado

Curso 2017-2018
Grado de Ingeniería Informática

Funciones de exploración consideradas:

Exploración cuadrática

Exploración doble

Realizado por:

Nombre: David Díaz Jiménez Correo electrónico: ddj00003@red.ujaen.es

Nombre: Manuel García López Correo electrónico: mgl00050@red.ujaen.es

Grupo 1 Prácticas: Martes 8:30-10:30

Resultados a entregar:

Ficha técnica de los resultados de las dos funciones de dispersión (cuadrática y doble):

Los resultados los hemos dividido en dos tablas, una con un límite de intentos, establecidos a 10, ya que así lo requería la práctica, y otra sin límite de intentos. Esta última tabla es la que con el tamaño más grande probado sin que el factor de carga decaiga por debajo del 0,6.

Tabla con límite de intentos = 10.

Intentos de inserción = 10	Tamaño 1: 104491	Tamaño 2: 111121	Tamaño 3: 132367
Dispersión Cuadrática	Máximo Colisiones: 9 Media Colisiones:0,890951 Factor Carga: 0,752199 Número de palabras insertadas: 78598	Máximo Colisiones: 9 Media Colisiones:0,802272 Factor Carga: 0,71055 Número de palabras insertadas:78957	Máximo Colisiones: 9 Media Colisiones:0,582553 Factor Carga: 0,599281 Número de palabras insertadas: 79325
Dispersión Doble	Máximo Colisiones:9 Media Colisiones:0,948547 Factor Carga: 0,754228 Número de palabras insertadas:78810	Máximo Colisiones:9 Media Colisiones:0,854825 Factor Carga: 0,711567 Número de palabras insertadas:79070	Máximo Colisiones:9 Media Colisiones:0,645418 Factor Carga: 0,599553 Número de palabras insertadas: 79361

En esta tabla se puede observar que la dispersión cuadrática tiene una media de colisiones inferior a la dispersión doble, sin embargo, es la dispersión doble la que inserta un mayor número de palabras, utilizando mejor los recursos disponibles. Además, se ha observado experimentalmente, que el tiempo de ejecución (carga de palabras), crece conforme se reduce el tamaño de la tabla hash con respecto al volumen de datos.

Tabla con límite de intentos = ilimitados. (Se ha limitado a 50 para no entrar en bucle infinito)

Intentos de inserción = ilimitados	Tamaño 1: 104491	Tamaño 2: 111121	Tamaño 3: 132367
Dispersión Cuadrática	Máximo Colisiones: 30 Media Colisiones:1,02302 Factor Carga: 0,760085 Número de palabras insertadas:79422	Máximo Colisiones: 32 Media Colisiones:0,872718 Factor Carga: 0,714734 Número de palabras insertadas:79422	Máximo Colisiones: 16 Media Colisiones:0,595704 Factor Carga: 0,600014 Número de palabras insertadas: 79422
Dispersión Doble	Máximo Colisiones:32 Media Colisiones:1,04495 Factor Carga: 0,760085 Número de palabras insertadas:79422	Máximo Colisiones:29 Media Colisiones:0,907658 Factor Carga: 0,714734 Número de palabras insertadas:79422	Máximo Colisiones:18 Media Colisiones:0,654088 Factor Carga: 0,600014 Número de palabras insertadas: 79422

Como se puede observar en esta tabla, y comparándola con la anterior, las ventajas de no limitar las inserciones, reside en que finalmente, las funciones de dispersión, consiguen insertar todos los datos. Además se observa, que a mayor tamaño de tabla, menor media de colisiones, y el máximo de estas también baja considerablemente. Todas insertan el total de las palabras del diccionario.

Para la realización de la práctica elegiremos la función de dispersión doble con número de intentos ilimitado y tamaño 132367. Esto nos da la inserción total de las palabras, un factor de carga idóneo de 0,600014, y una media de colisiones de 0,654088. El único dato negativo son las hasta 61 palabras que requieren 10 o más intentos para ser insertadas.

Comparación de tiempos de las prácticas

(Bajo ejecución en Visual C++ y Visual Studio 2017)

Tiempo de ejecución de la práctica 5 AVL : 6.6909 segundos

Tiempo de ejecución de la práctica 6 Mapas : 7,10172 segundos

Tiempo de ejecución de la práctica 7 (Tabla Hash Cerrada con dispersión doble, intentos ilimitados y tamaño 132367) : 4,13275 segundos.