

El algoritmo de ordenación por mezcla

¿En qué medida las técnicas de concurrencia optimizan el
algoritmo de ordenación *Merge Sort*?

MONOGRAFIA DE INFORMÁTICA NS

Cómputo de palabras: xxxx

Código del alumno: lqv708

Índice

1. Introducción	2
1.1. Merge Sort	2
1.2. Ejecución serial y paralela	3
1.3. Iteración y recursividad	3
1.4. Complejidad	4
2. Implementaciones	6
2.1. Merge Sort Recursivo Serial (MSRS)	6
2.1.1. Complejidad	7
2.2. Merge Sort Iterativo Serial (MSIS)	9
2.2.1. Complejidad temporal	9
2.3. Merge Sort Recursivo Paralelo (MSRP)	11
2.4. Merge Sort Iterativo Paralelo (MSIP)	15
3. Experimentación	18
3.1. Procedimiento	18
4. Discusión de resultados	19
4.1. Comparación	21
4.2. Conclusión	21
5. Bibliografía	21
Referencias	21
6. Anexo	22

1. Introducción

En el día a día usamos estructuras de datos ordenadas, como la lista de contactos del teléfono, en los almacenes para la gestión de inventario, en los resultados de una búsqueda en internet, etc. El proceso de colocar los datos en un cierto orden se llama ordenación.¹ La ordenación es una operación común en los sistemas informáticos y ha sido ampliamente estudiada.² Aun así, no existe un algoritmo de ordenación perfecto. Actualmente, se siguen desarrollando, y además estos relacionan una gran variedad de conceptos de informática, como la concurrencia, el paralelismo, la ejecución serial, la iteración o la recursión.

La finalidad de esta investigación evaluar holísticamente cuatro implementaciones del algoritmo de ordenación Merge Sort en términos de complejidad temporal y espacial, del peor caso; tiempos de ejecución y memoria usados; además de la naturaleza específica de cada implementación. Concretamente, se comparan el Merge Sort Iterativo Serial (MSIS), el Merge Sort Recursivo Serial (MSRS), el Merge Sort Recursivo Paralelo (MSRP), y el Merge Sort Iterativo Paralelo (MSIP).

1.1. Merge Sort

El ordenamiento por mezcla es un algoritmo de ordenación inspirado en la técnica divide y vencerás. Es capaz de ordenar un conjunto de datos a través de, primero, dividir la colección en dos mitades, dividir las sub-colecciones en más mitades hasta que contengan cero o un elemento, ordenar cada sub-colección, unir (*merge*) todas las sub-colecciones de forma ordenada, y, finalmente la colección queda ordenada.³

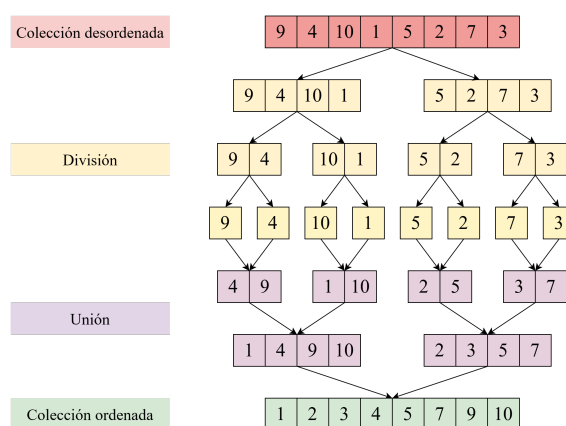


Figura 1: Funcionamiento del *Merge Sort*

¹ (Knuth, 1997)

² (McMillan, 2007)

³ (Skiena, 2008)

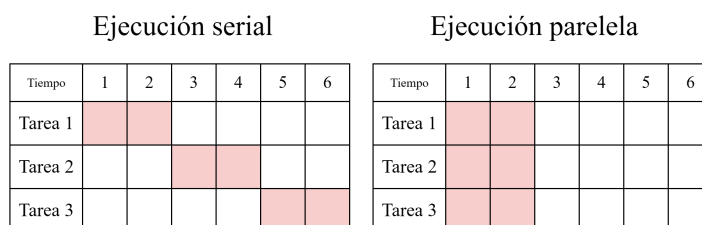


Figura 2: Comparación de los diagramas de Gantt (//Nota: Probablemente cambiar por dos árboles mostrando la secuencia de instrucciones)

1.2. Ejecución serial y paralela

Un **programa informático** es un conjunto de instrucciones que un sistema informático ejecuta. A su vez, un programa se divide en partes más pequeñas e independientes que llamamos tareas. Cuando las tareas se ejecutan una tras otra durante períodos de tiempo no superpuestos, hablamos de **ejecución serial**. A veces, la ejecución de una tarea depende del resultado de la tarea anterior. En dicho caso, la tarea bloquea a la siguiente. Esto es la **computación secuencial**. En contraposición, la **ejecución paralela** consiste en ejecutar varias tareas simultáneamente. Sin embargo, el paralelismo real solo es posible si el sistema consta de más de una unidad de procesamiento y si las tareas del algoritmo son independientes.⁴ El estudio pretende aplicar el paralelismo a la ordenación por mezcla para aumentar la eficiencia del algoritmo.

1.3. Iteración y recursividad

El paradigma de la **programación estructurada** considera que todo programa informático está formado por las estructuras de control de Secuencia, Selección y Repetición.⁵ La recursión es una estructura de repetición.⁶ Si un programa incorpora una estructura recursiva, quiere decir que hay una función que se llama a sí misma.⁷ Toda función recursiva contiene caso recursivo, estructura condicional que llama a la propia función, y un caso base, que retorna un valor constante o finaliza la función.

[

Nota: añadir gráfico para mostrar la recursión??]

⁴ (Bobrov, 2023)

⁵ ([ELI] y [NOVA], s.f.)

⁶ (College, 2000)

⁷ (Bhargava, 2016)

1.4. Complejidad

Esta investigación evalúa la aplicación de todos estos conceptos al algoritmo de ordenación por mezcla. La eficiencia de un algoritmo de ordenación usualmente se cuantifica en términos de complejidad temporal y espacial. La complejidad temporal es una medida de la variación del tiempo de ejecución de un algoritmo a medida que el tamaño de la entrada crece. El tiempo de ejecución se ve afectado por variables como el soporte físico y lógico (CPU⁸, RAM⁹, lenguaje de programación, SO...) ¹⁰ Es por ello, que se estudia el comportamiento asintótico de la complejidad, es decir, cuando el tamaño de la entrada n tiende al infinito. Lo hace mediante las funciones del: peor caso posible $O(n)$, el caso promedio $\theta(n)$, y el mejor caso posible $\Omega(n)$. ⁸ De ahí la notación *Big-O*, que permite comparar algoritmos a través de distintas plataformas y predecir el tiempo de ejecución de un algoritmo.

La **complejidad temporal** se mide normalmente mediante la función del peor caso posible por varios motivos: es más sencillo de analizar al no ser necesario conocer la distribución de datos de entrada, garantiza el buen funcionamiento del programa en el que se incorpore en caso de situaciones desfavorables... (NOTA: Más ejemplos!!) Caracterizar un algoritmo mediante su complejidad $O(n)$ es una abstracción, ya que se fundamenta en el modelo de Máquina de Acceso Aleatorio, que considera que las *operaciones básicas*, como las aritméticas, lógicas o los accesos a la memoria, duran una sola unidad de tiempo. Asimismo, se considera que cada línea de código es una *instrucción básica* que ejecuta un número constante de operaciones básicas. Los bucles (`for`, `for`, `while`...) y llamadas a funciones se evalúan sumando sus instrucciones básicas.

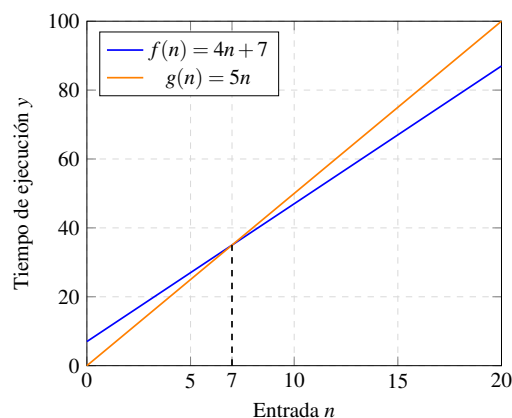


Figura 3: Ejemplo de *Big-O*

⁸ Central Processing Unit

⁹ Random Access Memory

¹⁰ (Heineman, Pollice, y Selkow, 2008)

La notación *Big-O* proporciona un límite superior para el cual una función nunca lo sobrepasará. Esta notación estudia el orden de crecimiento de una función. Formalmente, se define como: dadas dos funciones $f(n)$, $g(n)$, entonces $f(n) = O(g(n))$ siempre que existan las constantes c y n_0 tal que $f(n) \leq c \cdot g(n)$, para todo $n \geq n_0$. Esto significa que para una función $f(n)$ solo existirá un Big-O $O(g(n))$ si todos los valores de su entrada n son inferiores al producto entre una constante c y una función $g(n)$, que es el límite superior. De forma que, $f(n)$ nunca crecerá más que $g(n)$.

Por ejemplo, se da un algoritmo que toma un tiempo de ejecución $f(n) = 4n + 7$ y queremos saber si se comporta de forma lineal. Esto es $g(n) = n$ Entonces para que $f(n) = O(g(n))$ hay que encontrar un valor de c para el que se cumpla que $4n + 7 \leq c \cdot n$. Con $c = 5$ la inecuación se cumple, por tanto: $f(n) = O(g(n)) = O(n)$; y esto solo se cumple para $n \geq n_0$, en este caso $n \geq 7$ porque: $4n + 7 \leq 5n \rightarrow 7 \leq 5n - 4n \rightarrow 7 \leq n$. De forma que el *Big-O* de $4n + 7$ es $O(n)$ para cualquier entrada mayor que 7.

2. Implementaciones

2.1. Merge Sort Recursivo Serial (MSRS)

El algoritmo de ordenación por mezcla clásico se basa en el paradigma «divide y vencerás». Es decir, primero se divide un problema en otros subproblemas, se solucionan los subproblemas, y, se combinan para llegar a la solución final.¹¹

```
1 public static void sort(int[] arr, int[] aux, int left, int right) {  
2     if (left >= right) return;  
3  
4     int mid = left + (right - left) / 2;  
5  
6     sort(arr, aux, left, mid);  
7     sort(arr, aux, mid+1, right);  
8  
9     merge(arr, aux, left, mid, right);  
10 }
```

Figura 4: Función `sort()` del Merge Sort Recursivo Serial

El algoritmo a estudiar consta de una función `sort()` que toma una colección `arr`, un índice inicial `left` y un índice final `right`. (Véase la Figura 4) Después calcula el pivote `mid` desde donde dividir la colección original `arr[]` y se ejecuta una llamada recursiva a `sort()` para cada mitad `arr[left]... arr[mid]` y `arr[mid+1]... arr[right]`.

```
1 private static void merge(int[] arr, int[] aux, int left, int mid, int right) {  
2     for (int i = left; i <= right; i++) aux[i] = arr[i];  
3  
4     int i = left, j = mid + 1, k = left;  
5  
6     while (i <= mid && j <= right) arr[k++] = (aux[i] <= aux[j])? aux[i++] : aux[j++];  
7  
8     while (i <= mid) arr[k++] = aux[i++];  
9 }
```

Figura 5: Función `merge()` del Merge Sort Recursivo Serial

Finalmente, se unen las mitades mediante la función `merge()` de la Figura 5 que toma los mismos argumentos que `sort()`, además del parámetro `mid`. Para controlar el recorrido de las tres colecciones se emplean tres índices: `i` para la mitad izquierda `arr[left]... arr[mid]`, `j` para la mitad derecha `arr[mid+1]... arr[right]` y `k` para el arreglo original `arr[left]... arr[right]`. Un primer bucle copia en un arreglo auxiliar `aux[]` todos los elementos de `arr[]`. Después, un segundo bucle recorre simultáneamente las dos

¹¹ (Sedgewick, 2003)

mitades y la colección auxiliar; mientras coloca en $arr[k]$ el elemento más pequeño entre $aux[i]$ y $aux[j]$. Existen dos posibilidades: primera, que el bucle se detenga porque $i > mid$, que implica que todos los elementos de la primera mitad $aux[left] \dots arr[mid]$ han sido copiados en $arr[]$; segunda, que el bucle se detenga porque $j > right$, que implica que todos los elementos de la segunda mitad $aux[mid+1] \dots arr[right]$ han sido copiado en $arr[]$. Sin embargo, si i no alcanza mid quedan elementos sin copiar en $arr[]$: entonces un tercer bucle copia los elementos restantes de $aux[i] \dots aux[mid]$ en $arr[]$.

Todas las implementaciones (MSRS, MSIS, MSRP, MSIP) emplean el mismo método `merge()` para unir las sucesivas mitades.

2.1.1. Complejidad

La complejidad temporal del algoritmo será la suma de las complejidades de cada línea. Según el modelo MAN¹², la comprobación del caso base (línea 2, Figura 4) y el cálculo del pivote mid toman tiempo constante $O(1)$ al ser operaciones básicas. La inicialización de L y R toma tiempo $O(n)$ porque en Java la creación de un arreglo implica la asignación de espacio en la memoria proporcional al tamaño de este, en este caso $length = n$. La copia de las sub-colecciones (bucles de las líneas 8-9) toman $O(n)$ al su tiempo de ejecución depender de n . Por último, `merge` (Figura 5) toma $O(n)$ porque en el peor caso se realizaran $left + right = length$ comparaciones $L[i] \leq R[j]$, que es lo mismo que n . Por tanto, el tiempo de ejecución respecto la entrada por ahora es $f(n) = 2O(1) + 2O(n) + 2O(n) + O(n)$, que es igual a $O(n)$ ya que según la notación *Big-O* se ignoran los términos de menor orden y los factores constantes, puesto que estos no afectan significativamente al crecimiento cuando n tiende a un número grande.

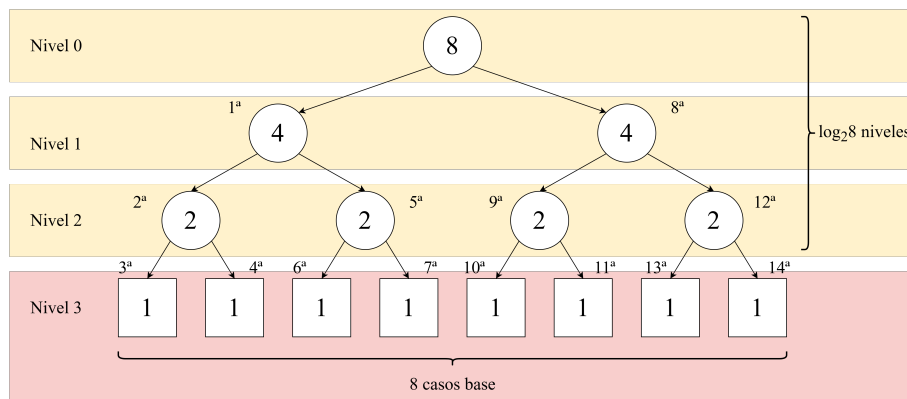


Figura 6: Árbol binario para $n = 8$

Lo anterior corresponde a una llamada a `sort()`, pero esta función es recursiva, por tanto, hay que

¹² Máquina de Acceso Aleatorio

determinar cuantas veces se llama a `sort()` en función del tamaño de la entrada n . Esto se facilita si rastreamos las llamadas a `sort()`, por ejemplo mediante un árbol binario como el de la Figura 6 donde cada nodo representa la longitud de la entrada `length`. Se observa que para 8 llamadas hay tres niveles en el árbol, la relación entre 8 y 3 es que $\log_2 8 = 3$ ya que $8 = 2^3$. Por extensión, para un n tamaño de entrada hay $\log_2 n$ niveles. Esto se cumple siempre que n sea múltiplo de 2, lo que da lugar a un árbol balanceado como el de Figura 8, en caso contrario está desbalanceado y hay llamadas extras. El trabajo realizado en cada nivel i es $2^i \cdot n \cdot \frac{n}{2^i} = n$ como se observa en la Figura 7 donde n es el tamaño de entrada original. Finalmente, el trabajo $f(n)$ de una llamada a `sort()` es la suma del trabajo en cada nivel, menos el del último nivel porque los casos bases requieren $O(1)$ al realizarse un `return`. Esto es:

$$f(n) = \sum_{i=0}^{\log_2 n - 1} n = n \sum_{i=0}^{\log_2 n - 1} 1 = n \cdot \log_2 n$$

A continuación se aplica la definición de la notación *Big-0*, donde $f(n) = n \log_2 n$ y $g(n) = n \log n$. Existe un $f(n) = O(n \log n)$ siempre que:

$$n \log_2 n \leq c \cdot n \log n$$

Aislamos la constante c :

$$n \log_2 n \leq c \cdot n \log n \longrightarrow c \geq \frac{\log_2 n}{\log_{10} n} \longrightarrow c \geq \frac{\frac{\log_{10} n}{\log_{10} 2}}{\log_{10} n} \longrightarrow c \geq \frac{1}{\log 2}$$

Esto significa que la complejidad del MSRS se aleja de $O(n \log_n)$ en un factor aproximado de 3,32% siempre que la entrada sea mayor que 1 y sea múltiplo de 2. Aun así, la literatura considera que el *Merge Sort* tiene complejidad $O(n \log n)$ por razones prácticas.¹³

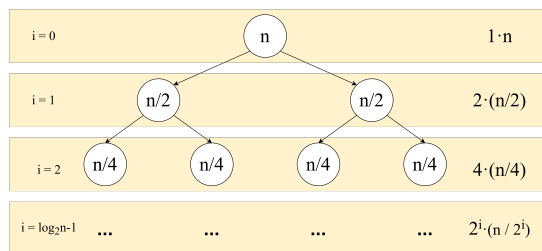


Figura 7: Tamaño de la entrada a lo largo de las llamadas a `sort()`

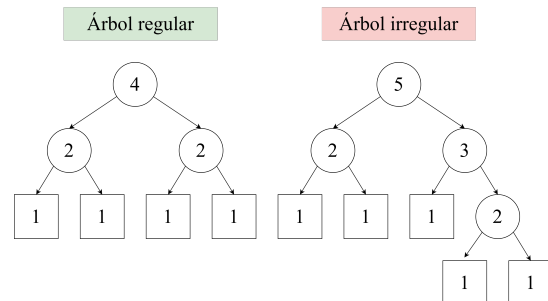


Figura 8: Árbol balanceado versus árbol desbalanceado del MSRS

¹³ (Sedgewick, 2003)

2.2. Merge Sort Iterativo Serial (MSIS)

El MSIS es la versión análoga al clásico MSRS, en esta implementación (Figura 9) se divide la colección en partes más pequeñas, después las partes adyacentes se unen, y se aumenta el tamaño de las partes. Estos tres pasos se repiten hasta que la parte tenga el tamaño de la colección original. El primer bucle determina el tamaño `length` de las partes: `size = 2, 4, 8...` El segundo bucle determina el índice `left` desde el cual comienza la parte `arr[left]...arr[left+size-1]` que será unida a su adyacente `arr[mid]...arr[right]` mediante `merge()`. Este índice toma valores en intervalos de $2 \cdot \text{size}$. En la Figura 10 se presenta el proceso de ordenación de una colección: las casillas coloreadas representan las sucesivas iteraciones de `left` del segundo bucle; las casillas azules corresponden a la respectiva parte izquierda de una iteración; y las verdes la parte derecha.

```
1 public static void sort(int[] arr, int[] aux) {
2     int n = arr.length;
3
4     for (int size = 1; size < n; size *= 2) {
5         for (int left = 0; left < n - size; left += 2 * size) {
6             int mid = left + size - 1;
7             int right = Math.min(left + 2 * size - 1, n - 1);
8             merge(arr, aux, left, mid, right);
9         }
10    }
11 }
```

Figura 9: Función `sort()` del Merge Sort Iterativo Serial

2.2.1. Complejidad temporal

El MSIS recorre el árbol (véase la Figura 11) desde la base de la recursión hasta la parte superior, ya que se realizan uniones entre partes de longitud 1-1, 2-2- 4-4... Como en el MSRS, el árbol será balanceado solo si el tamaño de entrada es múltiplo de 2. Siguiendo el mismo método y tomando las mismas suposiciones que en el cálculo de la complejidad temporal del MSRS, la complejidad del MSRS es $O(n \log n)$ porque hay $\log_2 n - 1$ llamadas a `sort()` en las cuales se realiza un trabajo de $O(n)$ en el peor caso.

2.3. Merge Sort Recursivo Paralelo (MSRP)

Un proceso es la ejecución de las instrucciones de un programa; después de que estas instrucciones se hayan movido desde la memoria secundaria (SSD¹⁴ por ejemplo) hasta la primaria (RAM¹⁵). El sistema operativo crea los procesos y guarda en la memoria su información asociada: el un identificador de proceso único (*PID*); un espacio de direcciones de memoria, ya que un proceso lee y modifica datos guardados en la memoria; el estado del proceso; y, una lista de los archivos que el programa usa. En cambio, un hilo de ejecución es una secuencia de instrucciones que el planificador del sistema operativo puede manejar independientemente.¹⁶ Hasta el momento se han mostrado programas que al ejecutarse toman la forma de un proceso con un solo hilo de ejecución (MSIS y MSRS). La idea es mejorar el rendimiento del *Merge Sort* haciendo uso de más de un hilo de ejecución.

La herramienta más apropiada que proporciona Java para este problema es la clase `ForkJoinPool`.¹⁷ Una piscina de hilos (*thread pool*), es un espacio en el que se mantienen un conjunto fijo de hilos de ejecución reutilizables que esperan a que se les pase un conjunto de instrucciones a ejecutar. Una piscina evita la necesidad de crear y destruir hilos constantemente, hecho que conlleva un costo computacional elevado.¹⁸

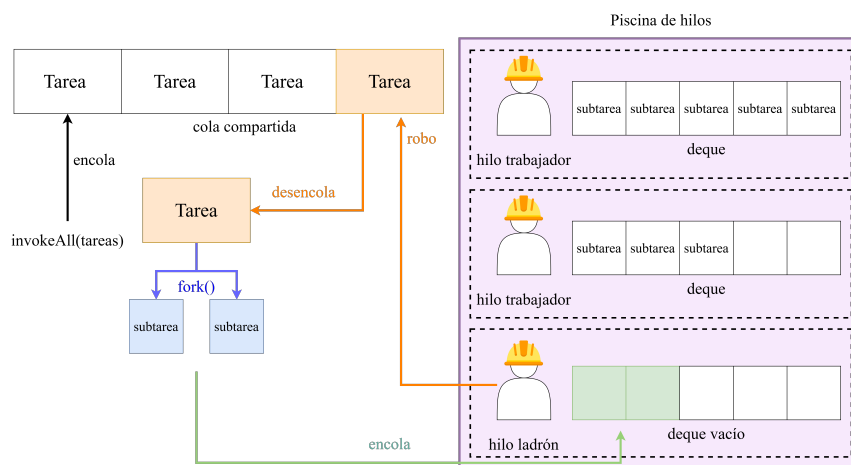


Figura 12: Funcionamiento del `ForkJoinPool`

La `ForkJoinPool` permite crear piscinas basadas en un algoritmo de robo de trabajo (*work-stealing algorithm*).¹⁹ Esto significa que las tareas se acumulan en una cola compartida entre todos los hilos, pero, además, cada hilo consta de su propia cola doblemente terminada (decola). Los hilos extraen de la cola las

¹⁴ Solid-state drive ¹⁵ Random access memory ¹⁶ (Bobrov, 2023) ¹⁷ (*ForkJoinPool (java SE 10 & JDK 10)*, s.f.) ¹⁸ (Engle, 2022) ¹⁹ (Ramgir y Samoylov, 2017)

tareas y las ejecutan, si la tarea produce subtareas estas se guardan en su decola. Puede ocurrir que la decola de un hilo se vacíe, en ese caso, el hilo desencola una tarea de la cola compartida.²⁰ La `ForkJoinPool` es útil si el algoritmo genera muchas subtareas, ya que el uso de decolas propias reduce significativamente la cantidad de accesos a la cola compartida.¹⁷ Adicionalmente, las tareas de la cola compartida serán siempre de mayor tamaño que las subtareas que generen tareas ubicadas en las decolas. Por ende, `ForkJoinPool` es apropiada para la paralelización del *Merge Sort*, en tanto que se generan subtareas que después se volverán a unir.

```

1 ForkJoinPool pool = new ForkJoinPool(parallelismLevel);
2 MSrecursivoParalelo task = new MSrecursivoParalelo(arr, aux, left, right);
3 pool.invoke(task);

```

Figura 13: Inicialización de la `ForkJoinPool`

Para crear una `ForkJoinPool` se llama a su constructor y se pasa el número de hilos trabajadores que tendrá la piscina (`parallelismLevel`).¹⁷ Después se debe pasar una tarea a la piscina que esta procesará cuando se llama al método `.invoke()`. Existen dos tipos tareas: aquellas que no retornan ningún valor (`RecursiveAction`) y aquellas que sí retornan (`RecursiveTask`). Ambas subclases heredan de `ForkJoinTask`.¹⁷ En este caso, se pasa una tarea del tipo `RecursiveAction`. (Véase la Figura 13)

Concretamente, la lógica del MSRP se inscribe en una clase `MSrecursivoParalelo` que hereda de la clase abstracta `RecursiveAction` como se observa en la Figura 14.

```

1 public class MSrecursivoParaleloSinUmbral extends RecursiveAction {
2     private final int[] arr, aux;
3     private final int right, left;
4
5     public MSrecursivoParaleloSinUmbral(int[] arr, int[] aux, int left, int right) {
6         this.arr = arr;
7         this.aux = aux;
8
9         this.left = left;
10        this.right = right;
11    }
12
13    @Override
14    protected void compute() { }
15
16    private static void merge(int[] arr, int[] aux, int left, int mid, int right) { }
17 }

```

Figura 14: Esquema de la clase `MergeSortRecursivoParalelo`

²⁰ (Kumar, 2024)

En dicha clase el algoritmo de *Merge Sort* se queda encapsulado en el método abstracto `compute()`, como se observa en la Figura 15. Primero se comprueba el caso base. Después se crean dos subtareas, por ahora inactivas, para cada lado de la colección. A continuación, se pasan las subtareas a `invokeAll()`, que las encola en la cola compartida de piscina y espera a que `Left` y `Right` finalicen. Finalmente, se unen las dos partes con el mismo algoritmo común `merge()`.

```
1  @Override
2  protected void compute() {
3      if (left >= right) return;
4
5      int mid = left + (right - left) / 2;
6
7      final MSrecursivoParalelo Left = new MSrecursivoParalelo(arr, aux, left, mid);
8      final MSrecursivoParalelo Right = new MSrecursivoParalelo(arr, aux, mid + 1, right);
9
10     invokeAll(Left, Right);
11
12     merge(arr, aux, left, mid, right);
13 }
```

Figura 15: Método `compute()` del MSRP

Normalmente llega un momento en que generar más subtareas torna ineficiente dado que el costo de gestión de hilos es mayor que el costo que conllevaría ejecutar las tareas de forma serial. En este caso, colecciones con baja longitud ralentizarían al MSRP. Por ende es común establecer un umbral (*threshold*) a partir del cual se ejecutan las subtareas serialmente.

Se ha conducido un test piloto para determinar el umbral adecuado para esta implementación. Se han tomado 50 muestras del tiempo de ejecución para 20 longitudes de arreglo para el Merge Sort Recursivo Serial y para el Merge Sort Recursivo Paralelo, pero sin un umbral definido (como el de la Figura 15). En la Tabla 1 se comparan los promedios de las muestras obtenidas. Los tamaños de las colecciones son potencias de dos, en tanto que se intenta conseguir un árbol binario balanceado, que aprovecha mejor los recursos. Obsérvese que hasta el tamaño 2^{15} el algoritmo más rápido es el MSRS. Entonces se modifica el método `compute()` para que, en caso de que la longitud del arreglo de entrada sea inferior al umbral de 2^{15} , se ejecute la versión serial. (Figura 17)

Tamaño	Tiempo MSRS (ms)	Tiempo MSRP sin umbral (ms)
2	0,002	0,271
4	0,004	0,325
8	0,003	0,375
16	0,003	0,402
32	0,005	0,519
64	0,004	0,751
128	0,005	0,806
256	0,013	0,924
512	0,027	0,987
1024	0,057	1,109
2048	0,120	1,227
4096	0,263	1,424
8192	0,529	1,468
16384	1,099	1,689
32768	2,224	2,186
65536	4,603	2,529
131072	9,332	3,770
262144	19,450	6,248
524288	39,380	11,012
1048576	81,661	25,592

Tabla 1: Datos de prueba para MSRS y MSRP sin umbral

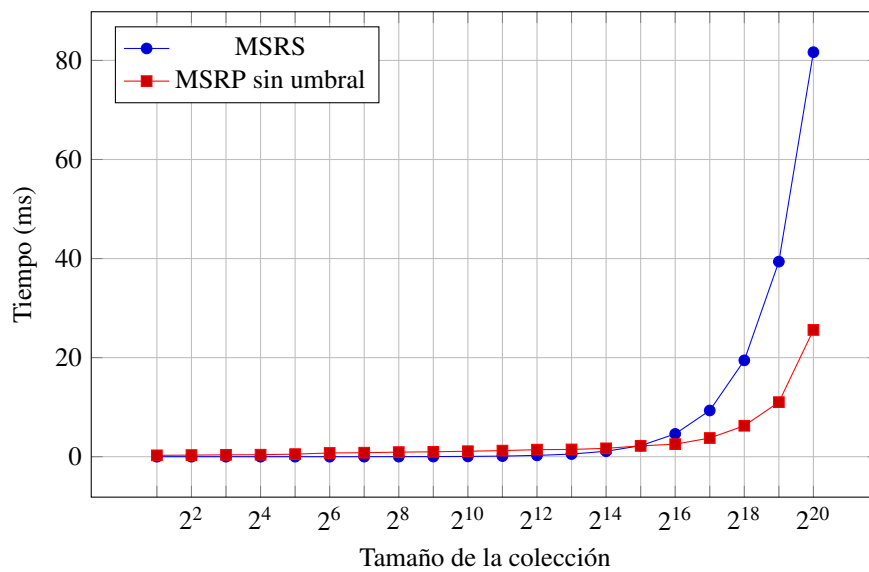


Figura 16: Datos de prueba para MSRS y MSRP sin umbral

```

1 protected void compute() {
2     int length = (right + 1 - left);
3     if (length <= THRESHOLD) {
4         MSrecursivoSerial.sort(arr, aux, left, right);
5     } else {
6         // Ejecucion paralela
7     }
8 }

```

Figura 17: MSRP con umbral

2.4. Merge Sort Iterativo Paralelo (MSIP)

Para la paralelización del algoritmo iterativo se ha hecho uso de la interfaz `ExecutorService` que proporciona Java y que proporciona un marco para gestionar y controlar la ejecución de tareas asincrónicas. A diferencia de las piscinas `ForkJoin`, el `ExecutorService` utiliza un algoritmo de trabajo compartido (*work-sharing algorithm*). Esto implica que solo hay una cola compartida entre todos los hilos: una vez termina un hilo de ejecutar una subtarea, extrae otra de la cola. Este flujo de ejecución es apropiado para tareas independientes entre ellas.²¹

```

1 public static void sort(int[] array, int aux[], ExecutorService executor) {
2     int length = arr.length;
3     List<Future<?>> futures = new ArrayList<>();
4
5     for (int size = 1; size < length; size *= 2) {
6
7         for (int left = 0; left < length - size; left += 2 * size) {
8             int mid = left + size - 1;
9             int right = Math.min(left + 2 * size - 1, length - 1);
10            int finalLeft = left; //El resto son efectivamente finales
11            futures.add(executor.submit(() -> merge(arr, aux, finalLeft, mid, right)));
12        }
13        for (Future<?> future : futures) {
14            try {
15                future.get();
16            } catch (Exception e) {
17                e.printStackTrace();
18            }
19        }
20        futures.clear();
21
22    }
23    executor.shutdown();
24 }

```

Figura 18: Método `sort()` del MSIP

En este caso, la implementación iterativa paralela se hace en una función `sort()` como la de la Figura 18. Para crear una instancia de `ExecutorService` se usan los métodos que proporciona la clase `Executors`

²¹ (Oracle, 2025)

de Java. Particularmente, `.newFixedThreadPool(parallelismLevel)` crea una piscina con un número de hilos fijo determinado.

```
1 ExecutorService executorService = Executors.newFixedThreadPool(parallelismLevel);  
2 MSIterativoParalelo.sort(arr, aux, executor);
```

Figura 19: Inicialización del `ExecutorService`

A continuación, para cada iteración se encola internamente en `executor` una tarea `merge()` mediante una expresión lambda que recibe `executor.submit()` como argumento. Entonces un número determinado de hilos trabajadores toman las tareas de la cola interna y las ejecutan. Si todos los hilos están ocupados, la tarea esperará en la cola hasta que un hilo esté disponible. Una vez que un hilo esté disponible, tomará la tarea de la cola y la ejecutará. Cada llamada a `merge()` implica que `.submit()` retorne un objeto de la clase `Future` que representa el resultado de una operación asíncrona. En este caso, se añaden los `Future` a una lista. Finalmente se recorre la lista y para cada `Future` se llama a `future.get()` que obliga al hilo principal a esperar a que acaben de procesarse todas las tareas encoladas. De lo contrario, podríamos pasar al siguiente `size` (del bucle inicial) sin asegurarse de que todas las partes están ordenadas. Una vez completado todos los bucles se llama a `executor.shutdown()`, que espera a que, una vez terminen de ejecutarse todas las tareas encoladas anteriormente, cierra la piscina `executor` y libera los hilos.

Al igual que en el MSRP se ha tratado de establecer un umbral, para en este caso particular cambiar a la versión iterativa serial (MSIS). Los resultados del test preliminar (Figura) (mismo procedimiento que el anterior) arrojan que no existe tamaño de colección alguno para el cual esta implementación (MSIP) sea más rápida.

Tamaño	Tiempo MSIS (ms)	Tiempo MSIP (ms)
2	0,004	0,240
4	0,005	0,432
8	0,004	0,649
16	0,005	0,645
32	0,006	0,504
64	0,004	0,504
128	0,006	0,605
256	0,011	0,685
512	0,024	0,663
1024	0,051	0,624
2048	0,111	0,931
4096	0,239	1,513
8192	0,509	2,686
16384	1,062	5,074
32768	2,197	9,538
65536	4,448	18,727
131072	8,980	45,305
262144	18,553	78,727
524288	37,924	173,462
1048576	78,274	320,937

Tabla 2: Datos de prueba para MSIS y MSIP sin umbral

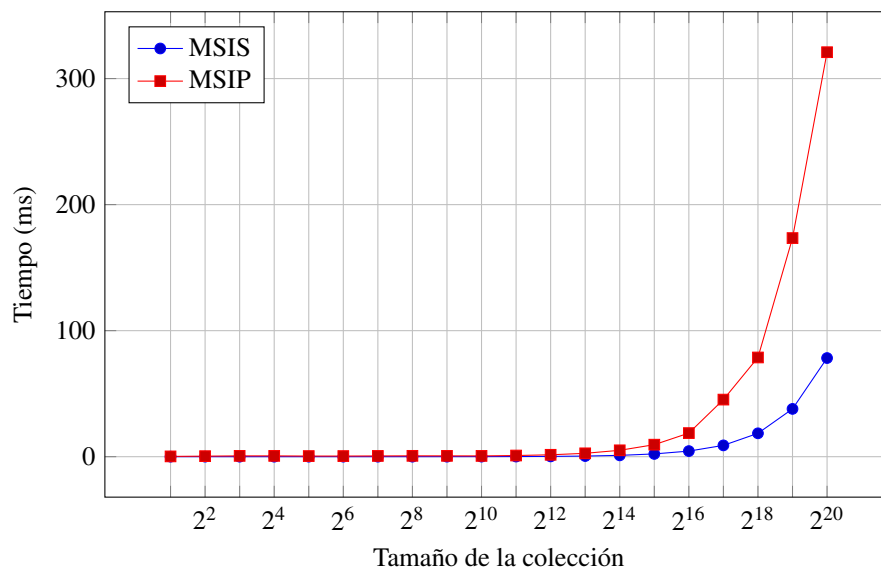


Figura 20: Datos de prueba para MSIS y MSIP sin umbral

3. Experimentación

Para la realización de una comparación empírica de los cuatro algoritmos (MSRS, MSIS, MSRP, MSIP) se toman 50 muestras del tiempo de ejecución para n tamaños de colección de entrada para cada algoritmo. Los algoritmos son ejecutados en un computador con procesador Intel(R) Core(TM) i5-11400F @ 2,60GHz y 16GB de RAM DDR4 3200MHz. En el SO de Windows 10 Pro 22H2 (x64). El IDE de ejecución es IntelliJ Idea Community Edition 2023.3.3.

3.1. Procedimiento

1. Se generan las longitudes crecientes de las colecciones de entrada siguiendo la secuencia $n = 10, 45, 100, 450, 1000, \dots, 10^8$. El máximo tamaño es 10^8 ya que a partir de este tamaño la memoria del computador empleado es insuficiente.
2. Para cada tamaño n se genera una colección mediante la clase `SplittableRandom` de Java, que en este caso genera colecciones de tipo `int` con valores pseudoaleatorios extraídos de una distribución uniforme.²² Particularmente, el generador produce números del 0 al 1000 y siempre emplea la misma semilla (6180339887), por tanto, en todas las muestras la secuencia de elementos a ordenar es la misma. Esto se hace para que la variación del tiempo de ejecución a lo largo del tamaño de entrada creciente se deba a la naturaleza misma del algoritmo y no influya el estado inicial del arreglo ya que son «idénticos».
3. Para cada toma de muestra se instancia un nuevo arreglo y se copia en este el arreglo generado anteriormente; de lo contrario en la siguiente muestra se estaría ordenando un arreglo ya ordenado. Para la colección auxiliar se sigue el mismo procedimiento.
4. El tiempo de ejecución se mide mediante la función `System.nanoTime` que retorna el tiempo actual más preciso disponible en el sistema. El valor devuelto son los nanosegundos desde un tiempo arbitrario y provee de precisión nanosegundaria, pero no necesariamente exactitud nanosegundaria. Cada ejecución se realizan entre una variable `long startTime` y `long endTime`. El tiempo de ejecución es la diferencia entre estas.²³
5. Después se elimina el peor y mejor tiempo de entre las 50 ejecuciones, quedando así 48.

²² (*SplittableRandom (Java Platform SE 8)*, 2024)

²³ (*System (Java Platform SE 8)*, 2024)

6. En el caso de los algoritmos paralelos se establece un mismo número de hilos para cada piscina para que la comparación sea justa. Concretamente, `parallelism=10` ya que el computador empleado consta de 12 hilos y se reservan 2 en caso de que se ejecute algún proceso en segundo plano que los necesite.
7. El código de las implementaciones, el código del *benchmark* y los resultados del *benchmark* quedan recogidos en los Apéndices A, B y C respectivamente.

4. Discusión de resultados

En la Tabla se presentan los resultados obtenidos en los tests realizados.

Tamaño	Tiempo MSRS (ms)	Tiempo MSIS (ms)	Tiempo MSRP (ms)	Tiempo MSIP (ms)
10	0,005	0,009	0,217	0,842
45	0,008	0,013	0,227	0,654
100	0,006	0,009	0,230	0,629
450	0,024	0,023	0,265	0,704
1.000	0,054	0,049	0,330	0,893
4.500	0,275	0,247	0,530	1,762
10.000	0,636	0,592	0,936	3,204
45.000	3,067	2,774	2,137	13,043
100.000	6,855	6,301	2,765	33,804
450.000	32,441	30,049	7,527	134,608
1.000.000	74,761	68,275	14,576	322,416
4.500.000	354,405	332,042	62,903	1.498,327
10.000.000	804,113	749,785	142,320	3.200,823
45.000.000	3.825,093	3.570,820	678,702	15.648,251
100.000.000	8.534,543	7.889,013	1.535,203	–

Tabla 3: Media de los tiempos de ejecución en ms

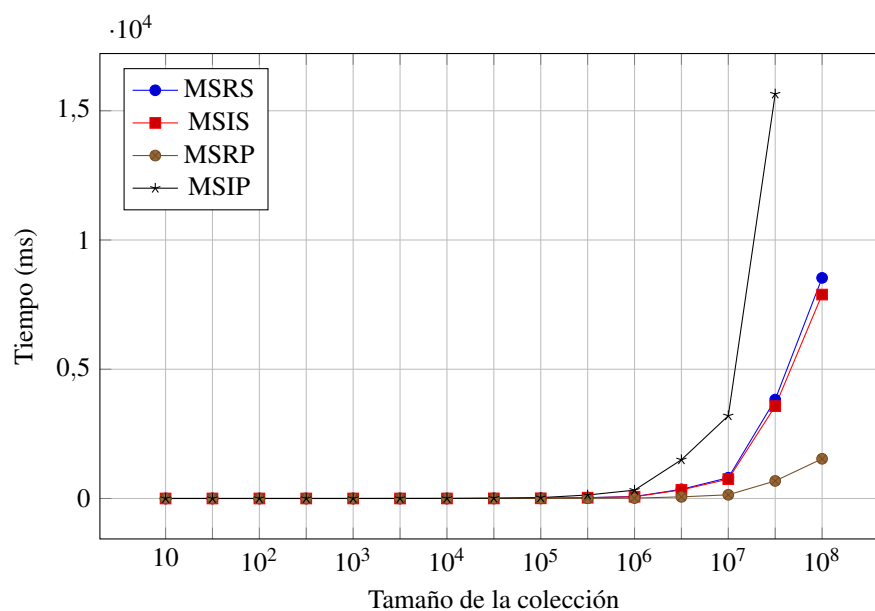


Figura 21: Media de los tiempos de ejecución en ms

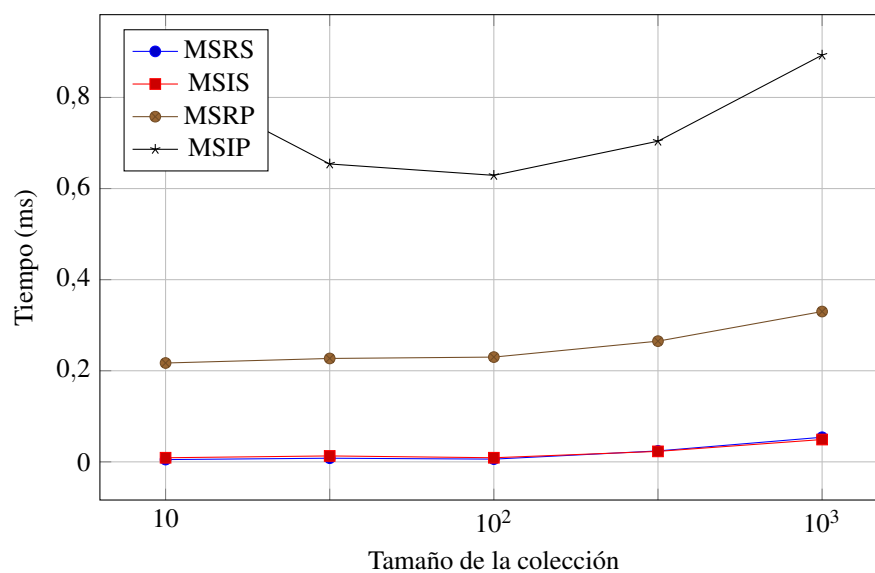


Figura 22: Media de los tiempos de ejecución en ms para tamaños pequeños

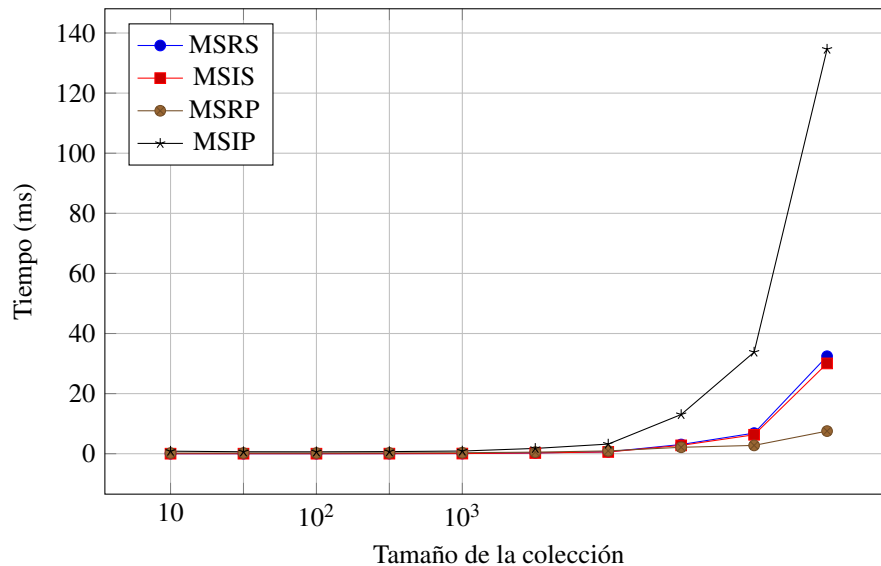


Figura 23: Media de los tiempos de ejecución en ms para tamaños medianos

4.1. Comparación

4.2. Conclusión

5. Bibliografía

Referencias

Bhargava, A. Y. (2016). *Grokking algorithms*. Manning Publications.

Bobrov, K. (2023). *Grokking concurrency*. Manning. Descargado de <https://www.manning.com/books/grokking-concurrency>

College, W. (2000). *Forms of iteration*. Descargado de https://cs111.wellesley.edu/archive/cs111_spring00/public_html/lectures/iteration.html

[ELI], E. L. I., y [NOVA], N. V. C. C. (s.f.). *Reading: Structured Programming*. Descargado de <https://courses.lumenlearning.com/sanjacinto-computerapps/chapter/reading-structured-programming/>

Engle, S. (2022). *Thread pools and work queues*. Descargado de <https://usf-cs272-spring2022.github.io/files/Thread%20Pools%20and%20Work%20Queues.pdf> (CS 272 Software Development, Department of Computer Science, University of San Francisco, Contact: sjen-gle@cs.usfca.edu)

- ForkJoinPool* (java SE 10 & JDK 10). (s.f.). <https://docs.oracle.com/javase/10/docs/api/java/util/concurrent/ForkJoinPool.html>. (Accessed: 2024-12-15)
- Heineman, G. T., Pollice, G., y Selkow, S. (2008). *Algorithms in a nutshell*. Sebastopol, CA, Estados Unidos de América: O'Reilly Media.
- Knuth, D. E. (1997). *The Art of Computer Programming: Seminumerical algorithms*. Addison-Wesley Professional.
- Kumar, R. (2024). *A deep dive into java's forkjoinpool mechanics*. Descargado de <https://medium.com/@reetesh043/a-deep-dive-into-javas-forkjoinpool-mechanics-556f82d160fb> (Medium, Contact: reetesh043@medium.com)
- McMillan, M. (2007). Basic sorting algorithms. En *Data structures and algorithms using c#* (p. 42–43). Cambridge University Press.
- Oracle. (2025). *Executorservice* (java platform se 8). Descargado de <https://docs.oracle.com/javase/8/docs/api/java/util/concurrent/ExecutorService.html> (Accedido: 19-ene-2025)
- Ramgir, M., y Samoylov, N. (2017). *Java 9 high performance*. Birmingham, Inglaterra: Packt Publishing.
- Sedgewick, R. (2003). *Algorithms in java* (3.^a ed.). Boston, MA, Estados Unidos de América: Addison-Wesley Educational.
- Skiena, S. S. (2008). *The Algorithm Design Manual*. Springer London. Descargado de <https://doi.org/10.1007/978-1-84800-070-4> doi: 10.1007/978-1-84800-070-4
- SplitTableRandom* (Java Platform SE 8). (2024, 9). Descargado de <https://docs.oracle.com/javase/8/docs/api/java/util/SplittableRandom.html>
- System* (Java Platform SE 8). (2024, 9). Descargado de <https://docs.oracle.com/javase/8/docs/api/java/lang/System.html>

6. Anexo