

Universidade Federal do Rio Grande do Norte Instituto Metrópole Digital

Análise de algoritmos clássicos de ordenação e busca

Bianca Medeiros, Gabriel Carvalho, Marina Medeiros, Vinicius de Lima

Estrutura de Dados Básica II

Trabalho da Primeira Unidade

Sumário

1	Aná	Análise Teórica								
	1.1	Bubble Sort	1							
	1.2	Merge Sort	2							
		1.2.1 Merge Sort Recursivo	3							
		1.2.2 Merge Sort Iterativo	3							
		1.2.3 Função auxiliar Merge	3							
		1.2.4 Análise de complexidade	6							
	1.3	Quick Sort	9							
2	Análise de Algoritmo 10									
	2.1	Função iterativa	LO							
	2.2	Função recursiva	LO							
	2.3	Implementação dos algoritmos de ordenação	LO							
		2.3.1 Bubble Sort	LO							
		2.3.2 Quick Sort	L 1							
		2.3.3 Merge Sort	12							
		2.3.4 Organização do projeto e corretude do algoritmo	L4							
	2.4	Análise de performance dos algoritmos de ordenação	16							

Capítulo 1

Análise Teórica

1.1 Bubble Sort

TODO

1.2 Merge Sort

Desenvolvido por Jon Von Neumann em 1945, o *Merge Sort* é um algoritmo de dividir para conquistar, que subdivide uma lista em singletons e os mescla em sublistas ordenadas até que exista apenas uma sublista, esta sublista é a lista original ordenada. Imagine o seguinte caso:

Você tem um baralho de cartas e gostaria de organizá-lo, seguindo o conceito do *Merge Sort* você trabalha da seguinte forma:

- **Divisão:** Primeiro você divide o baralho em dois baralhos menores. Cada um desses baralhos é dividido novamente até que cada sub-baralho tenha uma carta.
- "Merge" (Mescla): Nesse momento, com cada sub-baralho com uma carta, todos estão ordenados, então você começa a juntar os sub-baralhos comparando duas cartas de cada baralho e colocando-as em ordem crescente. Assim dois grupos de uma carta se tornam um baralho de duas cartas ordenadas. Em seguida, dois baralhos de duas cartas são mesclados para formar um baralho de quatro cartas, e assim por diante, até que todas as cartas estejam combinadas novamente, mas agora ordenadas.

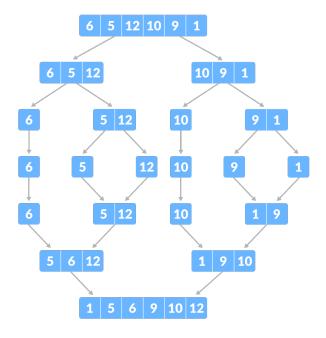


Figure 1.1: Diagrama exemplo de um Merge Sort

1.2.1 Merge Sort Recursivo

Dito isso, agora fica mais fácil estabelecer o pseudocódigo na forma recursiva. O caso base será se a lista tem no máximo elemento, pois já está ordenada. No caso recursivo, pelo teorema da recursão temos acesso ao "caso anterior", ou seja a primeira metade e segunda metade da lista original ordenadas, portanto, podemos mesclá-las.

```
Input: lista = x_0, x_1, ..., x_{N-1}

1: function MergeSort(lista)

2: if tamanho da lista \leq 1 then return

3: end if

4: return Merge(MergeSort(1^a metade da lista), MergeSort(2^a metade da lista)

5: end function
```

1.2.2 Merge Sort Iterativo

TODO

1.2.3 Função auxiliar Merge

Nos resta agora apenas definir a função **Merge**, que vai ser responsável por mesclar duas listas que estão ordenadas. Para tal, a função deve comparar um elemento da primeira lista com um da segunda e anexar o menor elemento entre os dois no final da lista resultado. Por exemplo:

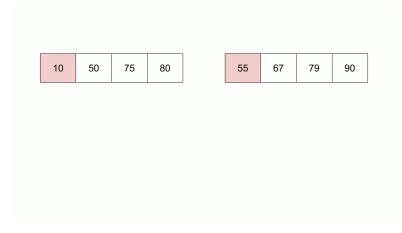


Figure 1.2: Comparando o primeiro elemento da primeira lista com o primeiro da segunda

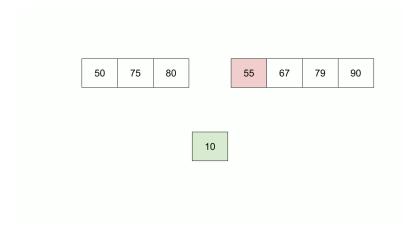


Figure 1.3: 10 é menor que 55, então é anexado no final da lista resultado

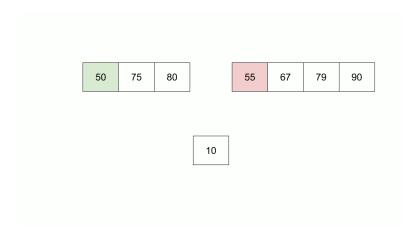


Figure 1.4: Comparando o próximo elemento da primeira lista

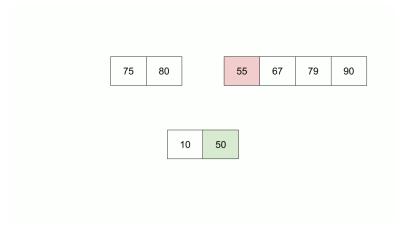


Figure 1.5: 50 é menor que 55, então é anexado no final da lista resultado

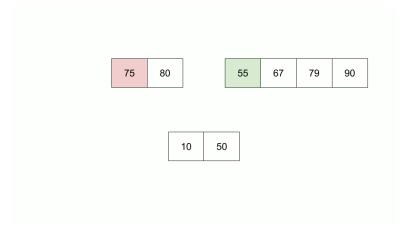


Figure 1.6: Comparando o próximo elemento da primeira lista

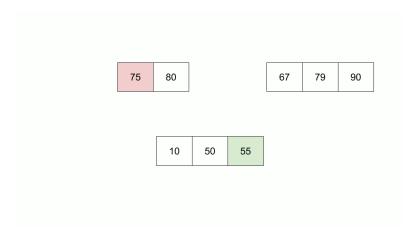


Figure 1.7: 55 é menor, então é anexado no final da lista resultado

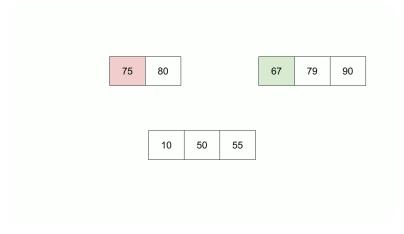


Figure 1.8: Comparando o próximo elemento da segunda lista

E o algoritmo vai continuar até que a lista resultado esteja completa com os elementos da primeira e segunda lista.

Dito isso, construímos o pseudocódigo dessa função da seguinte forma:

```
Input: IistaEsquerda = x_0, x_1, \dots, x_{N-1}, IistaDireita = y_0, y_1, \dots, y_{M-1}
Output: A listaResultado ordenada com os elementos da listaEsquerda e listaDireita
 1: function Merge(listaEsquerda, listaDireita)
                                                  ▷ Esses serão os indexadores de cada lista
       E, D, R = 0
 2:
 3:
       listaResultado = 0, 0, \ldots, 0
       while E < N e D < M do
 4:
           if listaEsquerda(E) < listaDireita(D) then
 5:
              listaResultado(R).push(listaDireita(D))
 6:
               E = E + 1
                                                        ▶ Partimos para o próximo elemento
 7:
           else
 8:
              listaResultado(R).push(listaEsquerda(D))
 9:
               D = D + 1
10:
           end if
11:
           R = R + 1
12:
       end while
13:
14:
       > Como uma das listas vai esgotar primeiro que a outra, copiamos os elementos
       restantes para a listaResultado
15:
       while listaEsquerda ou listaDireita tiver elementos do
16:
           adicione os elementos na listaResultado
17:
       end while
18:
       return listaResultado
19:
20: end function
```

1.2.4 Análise de complexidade

Para analisar qual é a complexidade de tempo da **Merge sort**, nas suas duas versões, vamos estabelecer primeiro qual é a complexidade da função **Merge**. A partir do pseudocódigo da função, é facil de ver que durante sua execução ela sempre vai percorrer o tamanho máximo entre a **listaEsquerda** e **listaDireita**, portanto, sua complexidade na notação de complexidade assintótica é O(n), $\Theta(n)$ e $\Omega(n)$.

Análise da versão iterativa

TO_DO

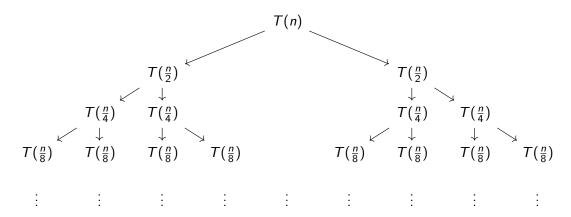
Análise da versão recursiva

Analisando o corpo da função, teremos dois casos: caso o tamanho da lista (n) seja menor ou igual 1 e caso contrário.

No primeiro caso, é facil de ver que a complexidade da função para uma lista de qualquer tamanho é constante, ou seja, O(1), $\Theta(1)$ e $\Omega(1)$. Caso contrário, observa-se que a função chama a si mesmo duas vezes, uma para cada metade da lista. Ademais, a função Merge, com complexidade linear para qualquer entrada, é chamada em cada etapa recursiva. Portanto, estabelecemos a relação de recorrência da Merge sort recursiva como:

$$T(n) = egin{cases} O(1), & ext{se } n \leq 1 \ 2T(rac{n}{2}) + O(n), & ext{caso contrário} \end{cases}$$

Uma vez estabelecida a relação de recorrência, vamos usar primeiro a **Árvore de recorrência** como método. Comecemos estabelecendo seu diagrama:



Uma vez estabelecida a árvore de recorrência, podemos tabular o tamanho da entrada, seu custo por nó e quantidade de nós para cada nível da árvore:

Nível da árvore	Tamanho da entrada	Custo por nó	Quantidade de nós
0	n	n	$1 = 2^0$
1	$\frac{n}{2^1}$	$\frac{n}{2^1}$	$2 = 2^1$
2	n	$\frac{n}{2^2}$	$4 = 2^{2}$ $8 = 2^{3}$
3	$\frac{\overline{2^2}}{\frac{n}{2^3}}$	$\frac{\frac{n}{2^2}}{\frac{n}{2^3}}$	$8 = 2^3$
:	:	:	:
i	$\frac{n}{2^i}$	$\frac{n}{2^i}$	2 ⁱ

Em seguida, para estabelecer o somatório que calcula a complexidade da função, precisamos identificar o valor de i para quando $T(\frac{n}{2^i}) = T(1)$, assim, teremos:

$$\frac{n}{2^{i}} = 1 \Longrightarrow n = 2^{i}$$

$$\Longrightarrow \log_{2} n = i$$

Dessa forma, a complexidade da função para $\Omega,\ \Theta$ e O será dada pelo resultado do somatório:

$$\sum_{i=0}^{\log_2 n} \frac{n}{2^i} \cdot 2^i = \sum_{i=0}^{\log_2 n} n$$

$$= n \sum_{i=0}^{\log_2 n} 1$$

$$= n \cdot \log_2 n$$

Pelo método do **Teorema mestre**, o mesmo resultado ocorre em ainda menos etapas. Pelo teorema, temos que estabelecendo a relação de recorrência nessa forma:

$$T(n) = aT\left(\frac{n}{h}\right) + \Theta\left(n^k\right)$$

Para algum $a \ge 1$, $b \ge 1$, e $k \ge 0$. Vale que:

- 1. Se $a \ge b^k$, então T(n) é $\Theta(n^{\log_b a})$.
- 2. Se $a = b^k$, então $T(n) \in \Theta(n^k \cdot \log_b a)$.
- 3. Se $a < b^k$, então $T(n) \in \Theta(n^k)$.

A partir da relação de recorrência estabelecida, tome a=2, b=2 e k=1. Como $b^k=2=a$, logo, $T(n)=\Theta(n\cdot\log_2 n)$. Como foi estabelecido que o pior caso, o melhor e caso médio iam ter a mesma complexidade, logo, a *Merge sort* também tem as complexidades $O(n\cdot\log_2 n)$ e $\Omega(n\cdot\log_2 n)$.

1.3 Quick Sort

Capítulo 2

Análise de Algoritmo

2.1 Função iterativa

2.2 Função recursiva

2.3 Implementação dos algoritmos de ordenação

Os algoritmos de ordenação Merge Sort, Quick Sort e Bubble Sort foram implementados em suas versões iterativas e recursivas na linguagem de programação Rust. Além disso, com a exceção do Bubble sort, todas as funções esperam um slice mutável de tipo T que implementa uma ordem parcial e pode ser copiado, estes são os traits PartialOrd e Copy, respectivamente. Todas as funções alteram esse slice diretamente e não retornam nada.

2.3.1 Bubble Sort

Iterativo

```
pub fn iterative_bubble_sort<T: PartialOrd>(arr: &mut [T]) {
   for i in 0..arr.len() {
     for j in 1..arr.len() - i {
        if arr[j - 1] > arr[j] {
            arr.swap(j - 1, j);
        }
    }
}
```

Recursivo

Na versão recursiva, foi utilizada a função auxiliar bubble_sort_pass

```
pub fn recursive_bubble_sort<T: PartialOrd>(arr: &mut [T]) {
   if arr.is_empty() {
      return;
   }
   let last_element_position = arr.len();
   bubble_sort_pass(arr, 1, last_element_position);
   recursive_bubble_sort(&mut arr[..last_element_position - 1]);
}
```

Função auxiliar

Essa função move o maior elemento de um slice para a posição last_element_position no arr. O iterator é usado para percorrer o slice do inicio ao fim.

2.3.2 Quick Sort

Iterativo

Como o quick sort naturalmente é um algoritmo recursivo, precisamos de algo que simulasse a pilha de execução do computador, esta simulação se deu a partir do uso da variável stack que armazena as posições das sub-listas que precisam ser ordenadas, permitindo a progressão iterativa da divisão da lógica de divisão e conquista.

```
1 pub fn iterative_quick_sort<T: PartialOrd + Copy>(arr: &mut [T]) {
   if arr.is_empty() {
     return;
   let mut stack = vec![(0, arr.len() - 1)];
    while let Some((low, high)) = stack.pop() {
      if low < high {</pre>
8
        let pivot_index = partition(arr, low, high);
9
        if pivot_index > 0 {
          stack.push((low, pivot_index - 1)); // Left side
10
11
        stack.push((pivot_index + 1, high)); // Right side
12
13
    }
14
15 }
```

Recursivo

Para manter o único parâmetro arr, a recursive_quick_sort só encapsula a versão de fato recursiva.

```
pub fn recursive_quick_sort<T: PartialOrd + Copy>(arr: &mut [T]) {
   if arr.is_empty() {
      return;
   }
   _recursive_quick_sort(arr, 0, arr.len() - 1);
   }
}
```

```
1 fn _recursive_quick_sort <T: PartialOrd + Copy > (
  arr: &mut [T],
    lower_bound: usize,
   upper_bound: usize,
5){
   if lower_bound >= upper_bound {
7
     return;
   }
8
   let pivot_index = partition(arr, lower_bound, upper_bound);
if pivot_index > 0 {
      _recursive_quick_sort(arr, lower_bound, pivot_index - 1);
11
12 }
   _recursive_quick_sort(arr, pivot_index + 1, upper_bound);
13
14 }
```

Listing 2.1: Versão correta

Função auxiliar

Note que ambas as versões do *quick sort* utilizam a função auxiliar partition. Essa função seleciona o último elemento como pivô e rearranja a lista, de modo que todos os elementos menores que o pivó estejam na esquerda e os maiores ou iguais ao pivô na direita. Ao final, ela retorna o índice do pivô após a partição. Os parâmetros lower_bound, upper_bound são os índices do elemento inicial e final do slice a ser particionado.

```
1 fn partition <T: PartialOrd + Copy > (arr: &mut [T], lower_bound: usize,
     upper_bound: usize) -> usize {
    let pivot = arr[upper_bound];
    let mut left_item = lower_bound as isize - 1;
   for right_item in lower_bound..upper_bound {
     if arr[right_item] < pivot {</pre>
        left_item += 1;
6
        arr.swap(left_item as usize, right_item);
7
8
    }
   let new_pivot = (left_item + 1) as usize;
   arr.swap(new_pivot, upper_bound);
    new_pivot
13 }
```

2.3.3 Merge Sort

Iterativo

Novamente, como o *merge sort* naturalmente é um algoritmo recursivo, foi necessário simular a pilha de execução do computador. Primeiro, começamos com pequenos segmentos da lista (de tamanho 1) e iterativamente dobramos o tamanho dos segmentos a cada passo, mesclando-os.

```
pub fn iterative_merge_sort<T: PartialOrd + Copy>(arr: &mut [T]) {
   if arr.len() <= 1 {
      return;
   }
   let mut temp_arr = arr.to_vec();
   let mut segment_size = 1;
   let arr_len = arr.len();
   while segment_size < arr.len() {
    let mut start = 0;</pre>
```

```
while start < arr_len {</pre>
        let mid = (start + segment_size).min(arr_len);
11
        let end = (start + 2 * segment_size).min(arr_len);
12
        merge(&mut temp_arr[start..end], &arr[start..mid], &arr[mid..end]);
13
        start += 2 * segment_size;
14
15
16
      arr.copy_from_slice(&temp_arr);
17
      segment_size *= 2;
    }
18
19 }
```

Recursivo

A principal diferença em relação pseudo código, é que a agora a função edita diretamente a lista, em vez de retorna-la como resultado.

```
pub fn recursive_merge_sort<T: PartialOrd + Copy>(arr: &mut [T]) {
   if arr.len() <= 1 {
      return;
   }
   let mid = arr.len() / 2;
   let mut left_arr = arr[..mid].to_vec();
   let mut right_arr = arr[mid..].to_vec();
   recursive_merge_sort(&mut left_arr);
   recursive_merge_sort(&mut right_arr);
   merge(arr, &left_arr, &right_arr);
}</pre>
```

Função auxiliar

E a *merge* também foi alterada de forma parecida, de forma que receba como argumento a lista resultado (arr).

```
1 fn merge<T: PartialOrd + Copy>(arr: &mut [T], left_arr: &[T], right_arr:
      &[T]) {
    let left_arr_len = left_arr.len();
    let right_arr_len = right_arr.len();
    let (mut i, mut 1, mut r) = (0, 0, 0);
    while 1 < left_arr_len && r < right_arr_len {</pre>
      if left_arr[l] < right_arr[r] {</pre>
6
        arr[i] = left_arr[l];
        1 += 1;
8
      } else {
        arr[i] = right_arr[r];
11
        r += 1;
12
      }
13
      i += 1;
    }
14
    while 1 < left_arr_len {</pre>
15
      arr[i] = left_arr[l];
16
      i += 1;
17
      1 += 1;
18
19
    while r < right_arr_len {</pre>
20
      arr[i] = right_arr[r];
21
22
      i += 1;
      r += 1;
23
    }
24
25 }
```

2.3.4 Organização do projeto e corretude do algoritmo

Para testar os algoritmos o projeto foi subdivido em duas partes, a que testa a performance e a que estipula a corretude, para tal foi usado a separação padrão do cargo entre o módulo principal e o módulo de testes. A organização dos arquivos ocorreu da seguinte forma:

```
/
   out/
   entries.txt
   output.txt
   src/
   algorithms.rs
   lib.rs
   main.rs
   tests/
   test_sorts.rs
   Cargo.lock
   Cargo.toml
   rustfmt.toml
```

No Cargo.toml A unica depêndencia usada foi rand, para poder gerar números aleatórios de 0 a 100000

```
[dependencies]
rand = "0.8.5"
```

Listing 2.2: Trecho do Cargo.toml

Quando executado com cargo test, o cargo executa as funções de teste definidas em test_sorts.rs para estipular a corretude de cada algoritmo. Os testes definidos foram:

```
1 fn reverse_list_test(func: fn(&mut [i32])) {
   let mut arr = [5, 3, 2, 4, 1];
   func(&mut arr);
    assert_eq!(arr, [1, 2, 3, 4, 5], "reverse_list_test failed");
5 }
7 fn duplicates_list_test(func: fn(&mut [i32])) {
   let mut arr = [4, 2, 3, 2, 1, 4];
   func(&mut arr);
   assert_eq!(arr, [1, 2, 2, 3, 4, 4], "duplicates_list_test failed");
11 }
13 fn already_sorted_list_test(func: fn(&mut [i32])) {
14 let mut arr = [1, 2, 3, 4, 5];
  func(&mut arr);
    assert_eq!(arr, [1, 2, 3, 4, 5], "already_sorted_list_test failed");
16
17 }
18
19 fn singleton_list_test(func: fn(&mut [i32])) {
   let mut arr = [42];
    func(&mut arr);
    assert_eq!(arr, [42], "singleton_list_test faild");
23 }
25 fn empty_list_test(func: fn(&mut [i32])) {
  let mut arr: [i32; 0] = [];
  func(&mut arr);
```

```
assert_eq!(arr, [], "empty_list_test failed");

Listing 2.3: Trecho de test_sorts.rs
```

Todos os testes unitários foram definidos como esse:

```
1 #[test]
2 fn test_recursive_quick_sort() {
3    reverse_list_test(recursive_quick_sort);
4    duplicates_list_test(recursive_quick_sort);
5    already_sorted_list_test(recursive_quick_sort);
6    singleton_list_test(recursive_quick_sort);
7    empty_list_test(recursive_quick_sort);
8 }
```

Listing 2.4: Trecho de test_sorts.rs

Executando o comando cargo test, podemos ver que todos os testes foram bem sucedidos:

```
1 running 6 tests
2 test test_iterative_bubble_sort ... ok
3 test test_iterative_merge_sort ... ok
4 test test_recursive_bubble_sort ... ok
5 test test_iterative_quick_sort ... ok
6 test test_recursive_merge_sort ... ok
7 test test_recursive_quick_sort ... ok
9 test result: ok. 6 passed; 0 failed; 0 ignored; 0 measured; 0 filtered out
    ; finished in 0.00s
```

Os testes de performance foram implementados em main.rs, onde criam e escrevem os arquivos entries.txt e output.txt com as listas geradas e os resultados de performance, respectivamente.

```
1 type SortFn<T> = fn(&mut [T]);
3 fn main() -> io::Result <()> {
    let sort_functions = Vec::from([
        ("ITE BUBBLE SORT", iterative_bubble_sort as SortFn < i32 > ),
       ("REC BUBBLE SORT", recursive_bubble_sort),
("REC QUICK SORT", recursive_quick_sort),
("ITE QUICK SORT", iterative_quick_sort),
("REC MERGE SORT", recursive_merge_sort),
("ITE MERGE SORT", iterative_merge_sort),
6
     ]);
11
12
     fs::create_dir_all("out")?;
13
    let mut entries_file = File::create(ENTRIES_FILENAME)?;
14
    let mut output_file = File::create(OUTPUT_FILENAME)?;
15
    let title = "Performance test for sort algorithms";
17
    writeln!(output_file, "{}\n{}\n", title, "=".repeat(title.len()))?;
18
19
    writeln!(
20
       output_file,
       "ITE - stands for iterative\nREC - stands for recursive\n"
21
    )?;
22
23
    for n in 1..=4 {
24
       writeln!(entries_file, "List with {} entries:\n", TEN.pow(n))?;
25
       run_entry(
26
27
         &sort_functions,
28
          TEN.pow(n),
```

Listing 2.5: Trecho de main.rs

2.4 Análise de performance dos algoritmos de ordenação

Uma vez estabelecidas as implementações dos algoritmos de ordenação, em conjunto com testes que estipulavam sua corretude. Executamos o programa principal com todas as otimizações do compilador (cargo run --release) e obtivemos os seguintes resultados:

Algoritmo	10 entradas	100 entradas	1000 entradas	10000 entradas
Bubble Sort (Iterativo)	281 ns	6.723 μs	314.678 μs	31.196169 ms
Bubble Sort (Recursivo)	221 ns	7.374 µs	411.449 µs	43.688736 ms
Quick Sort (Recursivo)	320 ns	2.554 μs	31.329 µs	388.807 μs
Quick Sort (Iterativo)	581 ns	2.775 μs	33.122 µs	403.624 μs
Merge Sort (Recursivo)	882 ns	5.69 µs	62.497 µs	744.751 μs
Merge Sort (Iterativo)	531 ns	2.976 µs	36.809 µs	467.994 μs

Table 2.1: Tabela de resultados

Começando pelo **Bubble Sort**, é possível notar que, apesar de uma performance inicial superior da versão recursiva, a implementação iterativa é mais eficiente conforme o tamanho da lista aumenta. Acredita-se que o overhead causado pelas chamadas recursivas e pelo contexto adicional na pilha de execução (função auxiliar recursiva) resultou em uma performance pior em comparação com a iterativa.

Em seguida, no **Quick Sort**, tivemos um cenário diferente, em que a versão recursiva se saiu como a mais eficiente em todas as entradas em relação à iterativa. Isso pode ter ocorrido pela instanciação da estrutura de dados pilha como um Vec<(usize, usize)>, que fica armazenado na heap, e um overhead de operações push e pop, enquanto as chamadas recursivas não lidam com essas operações e ficam diretamente na pilha do sistema, que é mais rápida.

Por último, no **Merge Sort**, que, apesar de naturalmente ser um algoritmo recursivo, ficou muito atrás de sua versão iterativa, a qual foi até duas vezes mais rápida na maioria dos casos. Entretanto, é fácil entender a origem dessa diferença, considerando que, na implementação recursiva, são instanciados dois Vec<T> em cada chamada recursiva, enquanto na versão iterativa é instanciado um único Vec<T> durante toda a execução.

Referências Bibliográficas

Contribuidores da Wikipedia (2024a), 'Divide-and-conquer algorithm — wikipedia, the free encyclopedia'. [Acessado: 26 de Outubro, 2024].

URL: https://en.wikipedia.org/wiki/Divide-and-conquer_algorithm

Contribuidores da Wikipedia (2024b), 'John von neumann — wikipedia, the free encyclopedia'. [Acessado: 26 de Outubro, 2024].

URL: https://pt.wikipedia.org/wiki/John_von_Neumann

Contribuidores da Wikipedia (2024c), 'Merge sort'. [Acessado: 26 de Outubro, 2024].

URL: https://en.wikipedia.org/wiki/Merge_sort

Contribuidores da Wikipedia (2024*d*), 'Quicksort — wikipedia, the free encyclopedia'. [Acessado: 26 de Outubro, 2024].

URL: https://en.wikipedia.org/wiki/Quicksort

Contribuidores do GeeksforGeeks (2024a), 'Merge sort'. [Acessado: 26 de Outubro, 2024].

URL: https://www.geeksforgeeks.org/merge-sort/

Contribuidores do GeeksforGeeks (2024b), 'Quick sort'. [Acessado: 26 de Outubro, 2024].

URL: https://www.geeksforgeeks.org/quick-sort-algorithm/

Mahmud, S. (2024), 'Merge sort: A detailed explanation'. [Acessado: 26 de Outubro, 2024].

URL: https://blog.shahadmahmud.com/merge-sort/