# САНКТ-ПЕТЕРБУРГСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ ФАКУЛЬТЕТ ИНФОКОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ

Отчет по лабораторной работе №2 по курсу «Алгоритмы и структуры данных» Тема: Сортировка слиянием. Метод декомпозиции. Вариант 12

Выполнила:

Мкртчян.К.Г.

K3141

Проверил:

Афанасьев А.В.

Санкт-Петербург 2024 г.

# Содержание отчета

Содержание отчета	2
Задачи по варианту	3
Задача №1. Сортировка слиянием [N баллов]	3
Задача №3. Число инверсий [N баллов]	6
Задача №8. Умножение многочленов [N баллов]	9
Дополнительные задачи	12
Задача №2. Сортировка слиянием+ [N баллов]	12
Задача №7. Поиск максимального подмассива за линейное время	14
[N баллов]	14
Вывод по всей лабораторной	17

#### Задачи по варианту

#### Задача №1. Сортировка слиянием [N баллов]

Текст задачи: Используя псевдокод процедур Merge и Merge-sort из презентации к Лекции 2 (страницы 6-7), напишите программу сортировки слиянием на Python и проверьте сортировку, создав несколько рандомных массивов, подходящих под параметры.

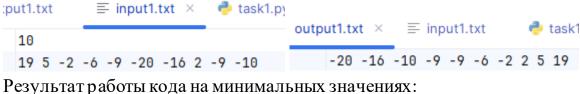
```
def verification(n, array, attempt=1):
    res = 1
    if type(n) is int and 1 <= n <= 10 ** 5:
       if type(array) is list and n == len(array) and all(type(x) is int
and abs(x) \le 10 ** 9 for x in array):
           return merge sort(n, array)
        else:
           res *= 0
    else:
       res *= 0
    if res == 0:
       if attempt == 3:
           return 'Ошибка!'
        else:
            print("Введите данные ещё раз, соблюдая ограничения: ")
                new n = int(input())
                new array = list(map(int, input().split(" ")))
                return verification(new n, new array, attempt + 1)
               return 'Ошибка!'
def merge sort main(*args):
    if len(args) == 1:
       path = args[0]
        file input = open(path, 'r')
       len arr = int(file input.readline().strip())
       array = list(map(int, file input.readline().strip().split(" ")))
        file output = open('output' + path[5:], 'w')
       result = " ".join(map(str, verification(len arr, array)))
       file output.write(result)
    else:
       len arr, array = args
        return verification(len arr, array)
def merge sort(len arr, array):
   middle = len arr // 2
   list a, list b = array[:middle], array[middle:]
   len a, len b = middle, len arr - middle
   if len a > 1:
        list a = merge sort(len a, list a)
   if len b > 1:
        list b = merge sort(len b, list b)
    return merge(len a, len b, list a, list b)
```

```
def merge(len a, len b, array a, array b):
    len c = len a + len b
    array c = [0] * len c
    index a, index b = \overline{0}, 0
    for index c in range(len c):
        if index b >= len b:
            array c[index c] = array a[index a]
            index a += 1
        elif index a >= len a:
            array c[index c] = array b[index b]
            index b += 1
            if array a[index a] <= array b[index b]:</pre>
                array c[index c] = array a[index a]
                index a += 1
                array_c[index_c] = array b[index b]
                index b += 1
    return array c
```

- 1. В первой функции verification, которая принимает на вход количество элементов в массиве, сам массив и число попыток ввода данных, мы проверяем ТИП переменных И соответствие ИХ значений ограничениям. Если все данные удовлетворяют условиям, то переходим к выполнению функции merge sort, если это не так, то просим пользователя ввести данные ещё 2 раза. Если пользователь ввёл неподходящие нам данные уже в третий раз, то программа сообщит об ошибке и завершится. Также мы считываем две строки из терминала, поэтому, если пользователь ввёл все данные в одну строку, то программа выдаст ошибку.
- 2. В функции merge\_sort мы находим середину массива и делим его пополам, если длина массива всё ещё больше единицы, то повторяем эту операцию. К двум получившимся половинкам применяем функцию merge.
- 3. В функции merge мы соединяем два списка следующим образом: создаем массив array\_с и заполняем его нулями (длина списка равна сумме длин списков array\_а и array\_b). Создадим также переменные для хранения индекса текущего элемента из первого и второго списков. Затем в цикле будем записывать в итоговой массив элемент из array\_a и прибавлять единицу к index\_a, если array\_a[index\_a] <= array\_b[index\_b]. Если это не так, то запишем элемент из второго списка и увеличим index\_b. Если мы добавили все элементы из первого (второго) массива, тогда добавим оставшиеся из второго (первого). Функция возвращает массив array\_c результат слияния двух списков.

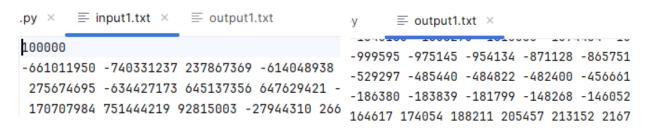
4. Функция merge sort main нужна для того, чтобы считать данные из файла, проверить данные из терминала или файла, запустив функцию verification, которая запустит merge sort, если удовлетворяют условиям задачи.

#### Результат работы кода на примере:



= in	put1.txt ×	= (		
1	1		output1.txt ×	
2	13		13	

# Результат работы кода на максимальных значениях:



	Время выполнения, сек	Затраты памяти, МБ
1 0	0.000011	19.62
1000 -349649722987715729	0.00109	20.02
10000 750431581265752675 - 102766916	0.01324	20.68
100000 -1951942875553582	0.1596	23.98

#### Вывод по задаче:

Мы научились применять сортировку слиянием, которая эффективна по времени работы и используемой памяти. Поняли принцип «разделяй и властвуй», а также повторили ввод и вывод данных из файла. Сортировка слиянием на массиве из 1000 элементов работает в 20 раз быстрее сортировки вставками. На таких данных оба алгоритма используют одинаковый объем памяти – около 20 МБ.

## Задача №3. Число инверсий [N баллов]

Текст задачи: Инверсией в последовательности чисел A называется такая ситуация, когда i < j, а Ai > Aj. Количество инверсий в последовательности в некотором роде определяет, насколько близка данная последовательность к отсортированной. Например, в сортированном массиве число инверсий равно 0, а в массиве, сортированном наоборот - каждые два элемента будут составлять инверсию (всего n(n-1)/2). Дан массив целых чисел. Ваша задача — подсчитать число инверсий в нем.

```
result = 0
def verification(n, array, attempt=1):
   res = 1
   if type(n) is int and 1 \le n \le 10 ** 5:
       if type(array) is list and n == len(array) and all(type(x) is int
and abs(x) \le 10 ** 9 for x in array):
           return merge sort(n, array)
       else:
           res *= 0
    else:
       res *= 0
    if res == 0:
       if attempt == 3:
            return 'Ошибка!'
            print("Введите данные ещё раз, соблюдая ограничения: ")
            try:
                new n = int(input())
                new array = list(map(int, input().split(" ")))
                return verification(new_n, new array, attempt)
                return 'Ошибка!'
def merge sort main(args):
   global result
    result = 0
    if len(args) == 1:
       path = args[0]
        file input = open(path, 'r')
        len arr = int(file input.readline().strip())
        array = list(map(int, file input.readline().strip().split(" ")))
       path = 'output' + path[5:]
        file output = open(path, 'w')
        result = str(number of permutations(len arr, array))
        file output.write(result)
       file output.close()
    else:
       len arr, array = args
        return verification(len arr, array)
```

```
def merge sort(len arr, array):
    middle = len arr // 2
    list a, list b = array[:middle], array[middle:]
    len a, len b = middle, len arr - middle
    if len a > 1:
        list a = merge sort(len a, list a)
    if len b > 1:
        list b = merge sort(len b, list b)
    return merge(len a, len b, list a, list b)
def merge(len a, len b, array a, array b):
    global result
    len c = len a + len b
    array c = [0] * len c
    count = 0
    index a, index b = 0, 0
    for index c in range(len c):
        if index b >= len b:
            array c[index c] = array a[index a]
            step = abs(index c - index a)
            index a += 1
        elif index a >= len a:
            array c[index c] = array b[index b]
            step = abs(index c - len a - index b)
            index b += 1
        else:
            if array a[index_a] <= array_b[index_b]:</pre>
                array_c[index_c] = array a[index a]
                step = abs(index c - index a)
                index a += 1
            else:
                array c[index c] = array b[index b]
                step = abs(index c - len a - index b)
                index b += 1
       count += step
    result += count//2
    return array c
def number of permutations(*args):
    merge sort main([i for i in args])
    return result
```

- 1) Создадим глобальную переменную result, которая будет хранить число инверсий.
- 2) Функции merge\_sort\_main, merge\_sort, verification работают аналогично одноименным функциям в предыдущей задаче, только в функции merge\_sort\_main мы записываем результат функции number of permutations.
- 3) В функцию merge добавим переменную-счётчик count и переменную шаг step. После каждого действия сортировки мы считаем шаг количество перестановок для данной переменной: для левой части массива он будет равен разности нового и старого индексов (index\_c-index\_a), а для правой разности нового и старого индекса минус длина левой части массива (index\_c-len\_a-index\_b).

После каждой итерации прибавляем значение шага к общему числу перестановок (count += step). По завершение сортировки добавляем count // 2 к результату. Делим значение count для того, чтобы не учитывать изменение индексов элементов, которые автоматически попали бы на свое место при, например, сортировке вставками.

4) Для того, чтобы посчитать число перестановок вызовем одноименную функцию – number\_of\_permutations. Она принимает на вход либо имя файла, в котором содержатся данные, либо длину массива и сам массив. Возвращает функция переменную result.

Результат работы кода на примере из задачи:



Результат работы кода на минимальных значениях:



Результат работы кода на максимальных значениях:

	Время выполнения, сек	Затраты памяти, МБ
1 0	0.0000071	19.96
1000 -511304447 -95059240 373123010	0.00138	19.98
10000 679843547 -348593807	0.01723	20.53
100000 -5269212 916664534 -608053615	0.2159	23.85

Вывод по задаче: мы научились считать число перестановок необходимых для сортировки массива и дополнили функцию сортировки слиянием, а также проверили получившиеся значения с помощью сортировки пузырьком.

#### Задача №8. Умножение многочленов [N баллов]

Текст задачи: Задача. Даны 2 многочлена порядка n-1: an-1x n-1+an-2x n-1+...+a1x+a0 и bn-1x n-1+bn-2x n-1+...+b1x+b0. Нужно получить произведение: 7 c2n-2x 2n-2+c2n-3x 2n-3+...+c1x+c0.

```
from math import log2, ceil
def verification(n, A, B, attempt=1):
   res = 1
    if type(n) is int and n > 0 and type(A) is list and type(B) is list:
        if len(A) == len(B) == n:
           pass
        else:
           res *= 0
    else:
       res *= 0
    if res == 0:
       if attempt == 3:
           return 'Ошибка!'
           print("Введите данные ещё раз, соблюдая ограничения: ")
                new n = int(input())
                new A = list(map(int, input().split(" ")))
                new B = list(map(int, input().split(" ")))
               return verification(new n, new A, new B, attempt + 1)
            except:
               return 'Ошибка!'
   else:
      return res
def karatsuba polynomial multiply(*args):
    if len(args) == 1:
       path = args[0]
        file input = open(path, 'r')
        n = int(file input.readline())
       A = list(map(int, file input.readline().strip().split()))
        B = list(map(int, file input.readline().strip().split()))
        if verification(n, A, B):
            file output = open('output' + path[5:], 'w')
            file output.write(reformat(multiply(power of two(A),
power of two(B))))
   else:
        n, A, B = args
        if verification(n, A, B):
           return reformat(multiply(power of two(A), power of two(B)))
def multiply(poly1, poly2):
   n = len(poly1)
   if n == 1:
```

```
return [poly1[0] * poly2[0]]
    mid = n // 2
    a = poly1[:mid]
    b = poly1[mid:]
    c = poly2[:mid]
    d = poly2[mid:]
    ac = multiply(a, c)
    bd = multiply(b, d)
    third product = multiply([a[i] + b[i] for i in range(mid)], [c[i] +
d[i] for i in range(mid)])
   middle = [third product[i] - ac[i] - bd[i] for i in
range(len(third product))]
    result = []
    n = []
    i = 3
    while i <= len(ac):</pre>
       n.append(i)
       i = 2*i + 1
    if len(ac) in n:
        result += union(ac, middle, bd)
        return result
    else:
        return ac + middle + bd
def union(a, b, c):
   n = len(a) // 2
   return a[:-n] + [a[-n + i] + b[i] for i in range(n)] + b[n:-n] + [b[-n
+ i] + c[i] for i in range(n)] + c[n:]
def power of two(m):
    n = len(m)
    return [0] * (2 ** ceil(log2(n)) - n) + m
def reformat(x):
    for i in range(len(x)):
        if x[i] != 0:
            return " ".join(map(str, x[i:]))
```

- 1) karatsuba\_polynomial\_multiply главная функция, с ее помощью мы считываем данные из файла, при необходимости, и вызываем функцию проверки, а затем функцию умножения полиномов.
- 2) Функция verification также, как и предыдущие одноименные функции проверяет полученные данные. В этой задаче нет строгих ограничений, поэтому проверим, чтобы п порядок многочленов являлся натуральным числом, а длины массивов A, B были равны друг другу и равны n.
- 3) Если введенные значения удовлетворяют условиям задачи, то переходим к выполнению функции multiply, которая принимает на

- вход два массива коэффициенты полиномов. Для того, чтобы массивы всегда делились ровно пополам, добавим в начало каждого массива нули, чтобы их длина равнялась двойке в какой-либо степени. Для этого при вызове функции multiply применим функцию power of two к обоим массивам.
- 4) В функции multiply разделим первый полином на а, b, а второй на с, d. Переменные ас, bd содержат произведение этих переменных соответственно. Переменная third\_product будет содержать произведение (a + b)(c + d), а переменная middle содержит разность third\_product и ас, bd. third\_product = (a + b)(c + d) ас bd = ac + ad + bc + bd ac bd = ad + bc это и есть коэффициенты из середины многочлена. Использование трех произведений вместо четырех, помогает быстрее выполнить перемножение многочленов. Такой способ предложил советский математик Анатолий Карацуба. Алгоритм Карацубы имеет сложность  $n^{\log_2 3}$  вместо  $n^2$ .
- 5) Затем мы объединяем результаты трех произведений с помощью функции union.
- 6) С помощью функции reformat мы убираем лишние нули в начале массива и преобразуем массив в строку, где коэффициенты разделены пробелами.

Результат работы кода на примерах из текста задачи:



Результат работы кода на минимальных значениях:

= inpu	ut8.txt	×	≣	
1	1		output8.txt ×	
2	4			
3	5		20	

	Время выполнения, сек	Затраты памяти, МБ
1	0.000012	19.79
5		
100 -744 647 -629 279 365 -545	0.00162	20.00

1000	0.0428	20.37
73334 50189 3616		
-64417 -58002 -95495		

Вывод по задаче: в этой задаче мы научились перемножать многочлены, используя принцип «разделяй и властвуй», а также алгоритм Карацубы.

#### Дополнительные задачи

#### Задача №2. Сортировка слиянием+ [N баллов]

Текст задачи: Дан массив целых чисел. Ваша задача — отсортировать его в порядке неубывания с помощью сортировки слиянием. Чтобы убедиться, что Вы действительно используете сортировку слиянием, мы просим Вас, после каждого осуществленного слияния (то есть, когда соответствующий подмассив уже отсортирован!), выводить индексы граничных элементов и их значения.

```
def verification(n, array, attempt, source):
    res = 1
    if type(n) is int and 1 <= n <= 10 ** 5:
        if type(array) is list and n == len(array) and all(type(x) is int
and abs(x) \le 10 ** 9 for x in array):
            return merge sort(0, n-1, array, source)
        else:
           res *= 0
    else:
       res *= 0
    if res == 0:
       if attempt == 3:
           return 'Ошибка!'
        else:
            print("Введите данные ещё раз, соблюдая ограничения: ")
                new n = int(input())
                new array = list(map(int, input().split(" ")))
                return verification (new n, new array, attempt + 1, '')
            except:
                return 'Ошибка!'
def merge sort main(*args):
   if len(args) == 1:
       path = args[0]
        file input = open(path, 'r')
        len arr = int(file input.readline().strip())
        array = list(map(int, file input.readline().strip().split(" ")))
        path = 'output' + path[5:]
        file output = open(path, 'w')
        file output.close()
        verification(len arr, array, 1, path)
        len arr, array = args
```

```
return verification(len arr, array, 1, '')
def merge sort(start, end, array, source):
   middle = (start + end + 1) // 2
   list a, list b = array[start:middle], array[middle:end+1]
   len a, len b = middle - start, end + 1 - middle
   if len a > 1:
       list a = merge sort(start, middle-1, array, source)
    if len b > 1:
       list b = merge sort(middle, end, array, source)
    return merge(start, end, list a, list b, source)
def merge(start, end, array a, array b, source):
   array c = []
   index a, index b = 0, 0
    for i in range (end - start + 1):
        if index b == len(array b):
            array c.extend(array a[index a:])
           break
        elif index a == len(array a):
            array c.extend(array b[index b:])
            break
        else:
            if array a[index a] <= array b[index b]:</pre>
                array c.append(array a[index a])
                index a += 1
            else:
                array c.append(array b[index b])
                index b += 1
    if len(source) == 0:
       print(start+1, end+1, *array c)
        file = open(source, 'a')
        file.write(f'{start+1} {end+1} {' '.join(map(str, array c))}\n')
        file.close()
    return array c
```

- 1) В функции merge\_sort\_main мы считываем данные из файла, если это необходимо, и проверяем их, вызывая функцию verification. Если данные удовлетворяют заданным ограничениям, то переходим к функции merge\_sort, которая разделяет массив и вызывает функцию merge для половинок массива.
- 2) Мы немного изменили функцию merge из первого задания. Теперь помимо двух массивов она принимает на вход индекс начала и конца первоначального массива (до разделения), а также source источник данных, если это файл, то мы запишем результаты в него, если данные введены из терминала, то source = " и мы выведем данные в консоль.
- 3) После сортировки мы выводим в консоль или записываем в файл индекс начала и конца части массива, а также эту, уже отсортированную часть.

Результат работы кода на примере из задачи:

	input2.txt
	1 1 2 1 8
	2 4514
	3 5 1 2 4
	4 1 5 1 1 2 4 8
	5 6737
≡ input2.txt × ≡ output2.txt	6 9 10 3 6
= inputz.txt \rightarrow = outputz.txt	7 8 10 2 3 6
1 10	8 6 10 2 3 3 6 7
2 1 8 2 1 4 7 3 2 3 6	9 1 10 1 1 2 2 3 3 4 6 7 8

	Время выполнения, сек	Затраты памяти, МБ
1 0	0.000011	19.62
1000 -349649722987715729	0.00109	20.02
10000 750431581265752675 - 102766916	0.01324	20.68
100000 -1951942875553582	0.1596	23.98

# Задача №7. Поиск максимального подмассива за линейное время [N баллов]

Текст задачи: можно найти максимальный подмассив за линейное время, воспользовавшись следующими идеями. Начните с левого конца массива и вправо, отслеживая найденный к двигайтесь данному моменту максимальный подмассив. Зная максимальный подмассив массива А[1...], распространите ответ на поиск максимального подмассива, заканчивающегося индексом і + 1, воспользовавшись следующим наблюдением: максимальный подмассив массива А[1..і + 1] представляет собой либо максимальный подмассив массива А[1..і], либо подмассив А[і..і +1] для некоторого  $1 \le i \le j+1$ . Определите максимальный подмассив вида А[і..і + 1] за константное время, зная максимальный подмассив, заканчивающийся индексом ј.

```
def line_find_max_subarray_main(*args):
    if len(args) == 1:
```

```
path = args[0]
        file_input = open(path, 'r')
        n = int(file input.readline().strip())
        array = list(map(int, file input.readline().strip().split(" ")))
        result = " ".join(map(str, line find max subarray(n, array)))
        file output = open('output' + path[5:], 'w')
        file output.write(result)
        file output.close()
        n, array = args
        line find max_subarray(n, array)
def line find max subarray(n, array):
    if verification(n, array):
       \max sum = 0
        start index = 0
        end index = 0
        sums = 0
        for i in range(n):
            if sums == 0:
               start index = i
            sums += array[i]
            if max sum < sums:</pre>
               max sum = sums
               end index = i
            if sums < 0:
                sums = 0
        return start index, end index, max sum
def verification(n, array, attempt=1):
    res = 1
    if type(n) is int and 1 \le n \le 10 ** 5:
        if type(array) is list and n == len(array) and all(type(x) is int
and abs(x) \le 10 ** 9 for x in array):
            res = 1
        else:
           res *= 0
    else:
       res *= 0
    if res == 0:
        if attempt == 3:
            return 'Ошибка!'
        else:
            print("Введите данные ещё раз, соблюдая ограничения: ")
            try:
                new n = int(input())
                new array = list(map(int, input().split(" ")))
                return verification(new n, new array, attempt + 1)
            except:
               return 'Ошибка!'
    else:
       return res
```

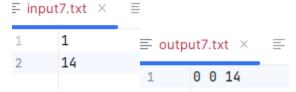
1) В функции line\_find\_max\_subarray\_main мы считаем данные и передадим их в функцию line\_find\_max\_subarray. В ней мы проверим данные с помощью функции verification, если они удовлетворяют условиям задачи, то перейдем к следующему алгоритму.

2) Создадим переменные: max\_sum – максимальная сумма подмассива, start\_index – индекс первого элемента подмассива, end\_index – индекс последнего элемента подмассива, sums – сумма элементов подмассива на данный момент. Изначально все они равны нулю. Затем пройдемся в цикле по массиву, если сумма равно нулю, то обновим индекс начала, затем прибавим к сумме і-ый элемент массива. Если сумма стала больше максимальной, то обновим максимальную сумму и индекс последнего элемента, если сумма стала отрицательной, то обнулим ее. После окончания цикла выведем индексы начала и конца подмассива, а также сумму элементов максимального подмассива.

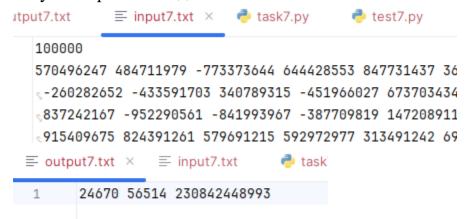
#### Результат работы кода на примере:



#### Результат работы кода на минимальных значениях:



## Результат работы кода на максимальных значениях:



	Время выполнения, сек	Затраты памяти, МБ
1 0	0.0000042	19.98
1000 484787009 - 933307925	0.000145	20.00
100000	0.0131	24.57

-370554242 -561616364	

Вывод по задаче: мы научились искать максимальный подмассив за линейное время при помощи алгоритма Кадане.

# Вывод по всей лабораторной

В ходе лабораторной работы №2 мы научились использовать быстрый алгоритм сортировки — сортировку слиянием, а также алгоритм поиска в отсортированном массиве — бинарный поиск. Также мы познакомились с алгоритмом Карацубы и Кадане и научились применять принцип «разделяй и властвуй» для разных задач.