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

Отчет по лабораторной работе №5 по курсу «Алгоритмы и структуры данных»

Тема: Пирамидальная сортировка Вариант 5

Выполнил:

Артемов И. В.

K3141

Проверил:

Афанасьев А. В.

Санкт-Петербург 2024 г.

# Содержание отчета

Содержание отчета	
Задачи по варианту	3
Задача №1. Куча ли?	3
Задача №6: Очередь с приоритетами	7
Дополнительные задачи	12
Задача №4. Построение пирамиды	12
Задача №7. Снова сортировка	18
Вывол	22

## Задачи по варианту

## Задача №1. Куча ли?

Текст задачи.

## 1 задача. Куча ли?

Структуру данных «куча», или, более конкретно, «неубывающая пирамида», можно реализовать на основе массива.

Для этого должно выполнятся основное свойство неубывающей пирамиды, которое заключается в том, что для каждого 1 < i < n выполняются условия:

- 1. если  $2i \leq n$ , то  $a_i \leq a_{2i}$ ,
- 2. если  $2i + 1 \le n$ , то  $a_i \le a_{2i+1}$ .

Дан массив целых чисел. Определите, является ли он неубывающей пирамидой.

- Формат входного файла (input.txt). Первая строка входного файла содержит целое число n ( $1 \le n \le 10^6$ ). Вторая строка содержит n целых чисел, по модулю не превосходящих  $2 \cdot 10^9$ .
- **Формат выходного файла (output.txt).** Выведите «YES», если массив является неубывающей пирамидой, и «NO» в противном случае.
- Ограничение по времени. 2 сек.
- Ограничение по памяти. 256 мб.
- Примеры:

No	input.txt	output.txt
1	5	NO
	10120	
2	5	YES
	1 3 2 5 4	

#### Листинг кода.

```
# Импортируем библиотеки для отслеживания памяти и времени выполнения программы import tracemalloc import time import os from lab5.utils import open_file, write_file, delete_prev_values, get_output_path, print_output_file

# Запускаем таймер для измерения времени работы программы t_start = time.perf_counter()

# Включаем отслеживание памяти tracemalloc.start()

current_dir = os.path.dirname(os.path.abspath(__file__)) # Текущая директория txtf_dir = os.path.join(os.path.dirname(current_dir), "txtf") # Директория для input.txt и output.txt input_path = os.path.join(txtf_dir, "input.txt")
```

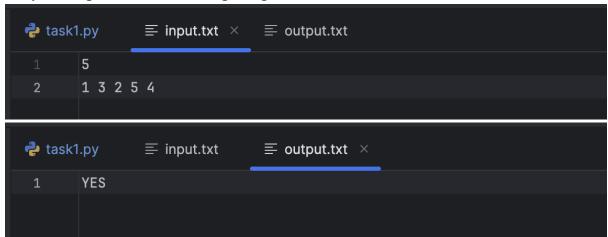
```
for i in range(n):
    if left child index < n and array[i] > array[left child index]:
    if right child index < n and array[i] > array[right child index]:
n str, array str = open file(input path)
n = int(n_str.strip()) # Преобразуем первую строку в число n
array = list(map(int, array str.strip().split())) # Преобразуем
    result = "YES" if is heap(array, n) else "NO"
   output_path = get_output_path(1)
   write_file(result, output_path)
   print output file(1)
print("Затрачено памяти:", tracemalloc.get traced memory()[1], "байт")
tracemalloc.stop()
```

#### Текстовое объяснение решения.

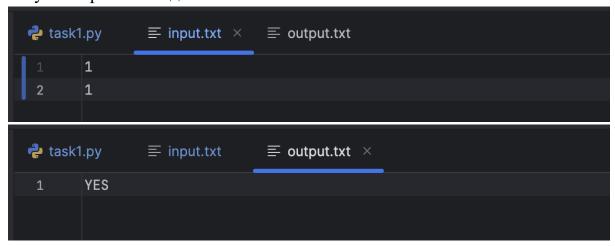
- 1. Импорт библиотек: `tracemalloc`, `time`, `os`, и функций из `lab5.utils`.
- 2. Запуск таймера для измерения времени выполнения.
- 3. Включение отслеживания памяти с помощью 'tracemalloc'.
- 4. Определение текущей директории.
- 5. Определение пути к директории с файлами 'input.txt' и 'output.txt'.

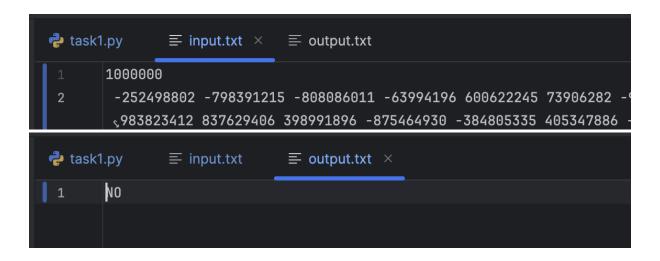
- 6. Определение пути к файлу 'input.txt'.
- 7. Определение функции `is\_heap` для проверки массива на соответствие свойствам пирамиды.
- 8. Чтение данных из файла 'input.txt' через 'open\_file'.
- 9. Преобразование первой строки данных в целое число 'n'.
- 10. Преобразование второй строки данных в массив целых чисел.
- 11. Проверка корректности входных данных: диапазон для `n` и элементов массива.
- 12. Вывод задачи и входных данных, если они корректны.
- 13. Очистка файла результатов для задачи 1 с помощью 'delete prev values'.
- 14. Проверка массива на соответствие свойствам пирамиды с использованием 'is\_heap'.
- 15. Запись результата проверки в файл 'output.txt'.
- 16. Вывод содержимого файла результатов с помощью 'print output file'.
- 17. Вывод сообщения об ошибке, если входные данные некорректны.
- 18. Вывод времени выполнения программы.
- 19. Вывод объема использованной памяти.
- 20. Остановка отслеживания памяти с помощью `tracemalloc.stop()`.

Результат работы кода на примере из задачи:



Результат работы кода на максимальных и минимальных значениях:





	Время выполнения	Затраты памяти
Нижняя граница диапазона значений входных данных из текста задачи	0.0004983750404790044 секунд	15324 байт
Пример из задачи	0.0008398750214837492 секунд	15374 байт
Верхняя граница диапазона значений входных данных из текста задачи	0.4109084579977207 секунд	10765530 байт

Вывод по задаче: мною был изучен алгоритм определения, является ли массив неубывающей пирамидой.

## Задача №6: Очередь с приоритетами

Текст задачи.

## 6 задача. Очередь с приоритетами

Реализуйте очередь с приоритетами. Ваша очередь должна поддерживать следующие операции: добавить элемент, извлечь минимальный элемент, уменьшить элемент, добавленный во время одной из операций.

• Формат входного файла (input.txt). В первой строке входного файла содержится число n ( $1 \le n \le 10^6$ ) - число операций с очередью.

Следующие n строк содержат описание операций с очередью, по одному описанию в строке. Операции могут быть следующими:

- A x требуется добавить элемент x в очередь.
- X требуется удалить из очереди минимальный элемент и вывести его в выходной файл. Если очередь пуста, в выходной файл требуется вывести звездочку «\*».
- $D\,x\,y$ —требуется заменить значение элемента, добавленного в очередь операцией A в строке входного файла номер x+1, на y. Гарантируется, что в строке x+1 действительно находится операция A, что этот элемент не был ранее удален операцией X, и что y меньше, чем предыдущее значение этого элемента.

В очередь помещаются и извлекаются только целые числа, не превышающие по модулю  $10^9$ .

- Формат выходного файла (output.txt). Выведите последовательно результат выполнения всех операций X, по одному в каждой строке выходного файла. Если перед очередной операцией X очередь пуста, выведите вместо числа звездочку «\*».
- Ограничение по времени. 2 сек.
- Ограничение по памяти. 256 мб.
- Пример:

input.txt	output.txt
8	2
A 3	1
A 4	3
A 2	*
X	
D 2 1	
X	
X	
X	

#### Листинг кода.

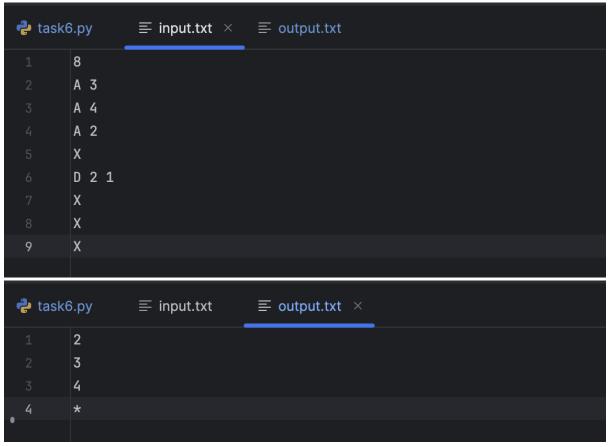
```
Импортируем библиотеки для отслеживания памяти и времени выполнения
import time
import tracemalloc
from lab5.utils import *
t start = time.perf counter()
tracemalloc.start()
current dir = os.path.dirname(os.path.abspath( file )) # Текущая
txtf dir = os.path.join(os.path.dirname(current dir), "txtf") # Путь к
input path = os.path.join(txtf dir, "input.txt")
def process_priority_queue(operations):
    import heapq
    id_map = \{\} # Словарь: идентификатор строки \rightarrow элемент
    for i, operation in enumerate (operations):
        parts = operation.split()
        if parts[0] == "A":
            heapq.heappush(heap, (x, current id))
            element map[current id] = x
            id map[i + 1] = current id # i+1 соответствует строке x + 1
            current id += 1
        elif parts[0] == "X":
            while heap and heap[0][1] in removed:
                heapq.heappop(heap) # Пропускаем удаленные элементы
                _, element id = heapq.heappop(heap)
                result.append(str(element map[element id]))
                result.append("*")
        elif parts[0] == "D":
            y = int(parts[2])
            element id = id map[x] # Получаем идентификатор элемента
            element map[element id] = y
```

## Текстовое объяснение решения.

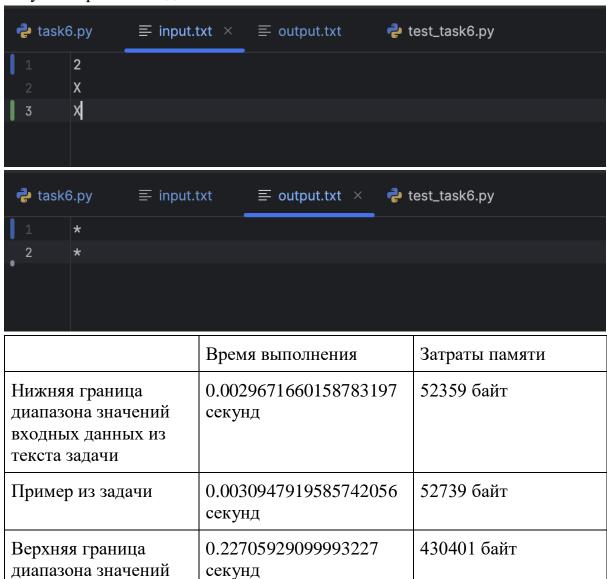
- 1. Импорт библиотек: 'time', 'tracemalloc' и функций из 'lab5.utils'.
- 2. Запуск таймера для измерения времени выполнения программы.
- 3. Включение отслеживания памяти с помощью 'tracemalloc'.
- 4. Определение текущей директории.
- 5. Определение пути к директории с файлами `input.txt` и `output.txt`.
- 6. Определение пути к файлу 'input.txt'.
- 7. Определение функции `process\_priority\_queue` для обработки операций с очередью с приоритетом.
- 8. Инициализация минимальной кучи.
- 9. Инициализация словаря для отслеживания значений элементов.
- 10. Инициализация словаря для соответствия идентификаторов строк и элементов.
- 11. Инициализация множества для удаленных элементов.
- 12. Инициализация списка для хранения результатов операций 'X'.
- 13. Обработка операций 'А' (добавление элемента).
- 14. Обработка операций 'X' (удаление минимального элемента).

- 15. Обработка операций 'D' (изменение значения элемента).
- 16. Возврат результатов операций.
- 17. Чтение данных из файла 'input.txt' через 'open file'.
- 18. Преобразование первой строки в целое число 'n'.
- 19. Формирование списка операций.
- 20. Проверка корректности входных данных.
- 21. Вывод задачи и входных данных, если они корректны.
- 22. Очистка предыдущих результатов задачи 6 с помощью `delete\_prev\_values`.
- 23. Вызов функции 'process priority queue' для обработки операций.
- 24. Запись результатов операций 'X' в файл 'output.txt'.
- 25. Вывод содержимого файла результатов через `print\_output\_file`.
- 26. Вывод сообщения об ошибке, если входные данные некорректны.
- 27. Вывод времени выполнения программы.
- 28. Вывод объема использованной памяти.
- 29. Остановка отслеживания памяти с помощью 'tracemalloc.stop()'.

Результат работы кода на примере из текста задачи:



## Результат работы кода на максимальных и минимальных значениях:



Вывод по задаче: мною был изучен алгоритм реализации очереди с приоритетами.

входных данных из

текста задачи

## Дополнительные задачи

## Задача №4. Построение пирамиды

Текст задачи.

## 4 задача. Построение пирамиды

В этой задаче вы преобразуете массив целых чисел в пирамиду. Это важнейший шаг алгоритма сортировки под названием HeapSort. Гарантированное время работы в худшем случае составляет  $O(n\log n)$ , в отличие от *среднего* времени работы QuickSort, равного  $O(n\log n)$ . QuickSort обычно используется на практике, потому что обычно он быстрее, но HeapSort используется для внешней сортировки, когда вам нужно отсортировать огромные файлы, которые не помещаются в памяти вашего компьютера.

Первым шагом алгоритма HeapSort является создание пирамиды (heap) из массива, который вы хотите отсортировать.

Ваша задача - реализовать этот первый шаг и преобразовать заданный массив целых чисел в пирамиду. Вы сделаете это, применив к массиву определенное количество перестановок (swaps). Перестановка - это операция, как вы помните, при которой элементы  $a_i$  и  $a_j$  массива меняются местами для некоторых i и j. Вам нужно будет преобразовать массив в пирамиду, используя только O(n) перестановок. Обратите внимание, что в этой задаче вам нужно будет использовать min-heap вместо max-heap.

• Формат ввода или входного файла (input.txt). Первая строка содержит целое число n ( $1 \le n \le 10^5$ ), вторая содержит n целых чисел  $a_i$  входного массива, разделенных пробелом ( $0 \le a_i \le 10^9$ , все  $a_i$  - различны.)

• Формат выходного файла (output.txt). Первая строка ответа должна содержать целое число m - количество сделанных свопов. Число m должно удовлетворять условию  $0 \le m \le 4n$ . Следующие m строк должны содержать по 2 числа: индексы i и j сделанной перестановки двух элементов, индексы считаются с 0. После всех перестановок в нужном порядке массив должен стать пирамидой, то есть для каждого i при  $0 \le i \le n-1$  должны выполняться условия:

```
1. если 2i + 1 \le n - 1, то a_i < a_{2i+1},
```

2. если 
$$2i + 2 \le n - 1$$
, то  $a_i < a_{2i+2}$ .

Обратите внимание, что все элементы входного массива различны. Любая последовательность свопов, которая менее 4n и после которой входной массив становится корректной пирамидой, считается верной.

- Ограничение по времени. 3 сек.
- Ограничение по памяти. 512 мб.

#### • Пример 1:

input.txt	output.txt
5	3
5 4 3 2 1	1 4
	0 1
	1 3

После перестановки элементов в позициях 1 и 4 массив становится следующим: 5 1 3 2 4.

Далее, перестановка элементов с индексами 0 и 1: 1 5 3 2 4. И напоследок, переставим 1 и 3: 1 2 3 5 4, и теперь это корректная неубывающая пирамида.

#### • Пример 2:

input.txt	output.txt
5	0
1 2 3 4 5	

#### Листинг кода.

```
# Импортируем библиотеки для отслеживания памяти и времени выполнения программы import tracemalloc import time from lab5.utils import *

# Запускаем таймер для измерения времени работы программы t_start = time.perf_counter()

# Включаем отслеживание памяти tracemalloc.start()

# Определяем пути к файлам current_dir = os.path.dirname(os.path.abspath(__file__)) # Текущая директория task1/src

txtf_dir = os.path.join(os.path.dirname(current_dir), "txtf") # Директория task1/txtf
```

```
input path = os.path.join(txtf dir, "input.txt")
    n = len(data)
        right = 2 * i + 2 \# индекс правого ребенка
            swaps.append((i, min index)) # Запоминаем перестановку
            data[i], data[min index] = data[min index], data[i]
        sift down(i)
   lines = open file(input path)
       swaps = build min heap(m)
       output path = get output path(4)
        output data = [f''(swaps))''] + [f''(i) (j)'' for i, j in swaps]
       write file("\n".join(output data), output path)
       print output file(4)
```

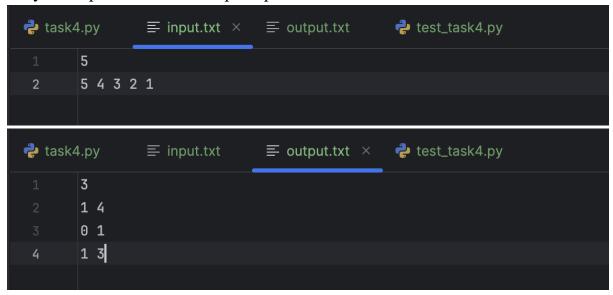
```
# Выводим время работы программы
print("Время работы: %s секунд" % (time.perf_counter() - t_start))
# Выводим количество памяти, затраченной на выполнение программы
print("Затрачено памяти:", tracemalloc.get_traced_memory()[1], "байт")
# Останавливаем отслеживание памяти
tracemalloc.stop()
```

#### Текстовое объяснение решения.

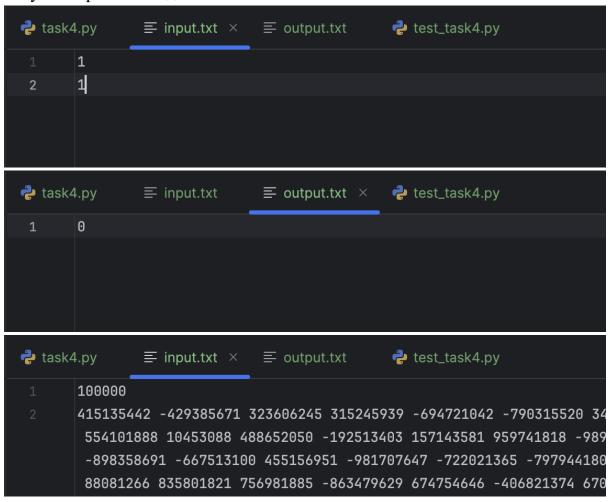
- 1. Импорт библиотек: `tracemalloc`, `time`, и функций из `lab5.utils`.
- 2. Запуск таймера для измерения времени выполнения программы.
- 3. Включение отслеживания памяти с помощью 'tracemalloc'.
- 4. Определение текущей директории.
- 5. Определение пути к директории с файлами `input.txt` и `output.txt`.
- 6. Определение пути к файлу 'input.txt'.
- 7. Определение функции `build\_min\_heap` для преобразования массива в min-heap и сбора перестановок.
- 8. Определение вспомогательной функции 'sift\_down' для просеивания элементов вниз.
- 9. Просеивание узлов массива с последнего родительского элемента до корня.
- 10. Возврат списка перестановок из функции 'build\_min\_heap'.
- 11. Чтение данных из файла 'input.txt' через 'open\_file'.
- 12. Удаление пробелов и символов новой строки из первой строки данных.
- 13. Разделение второй строки данных на элементы.
- 14. Преобразование первой строки в целое число 'n'.
- 15. Преобразование элементов второй строки в массив целых чисел.
- 16. Проверка корректности входных данных: диапазон для `n` и элементов массива.
- 17. Вывод задачи и входных данных, если они корректны.
- 18. Очистка предыдущих результатов задачи 4 с помощью `delete\_prev\_values`.
- 19. Вызов функции `build\_min\_heap` для построения min-heap и получения перестановок.
- 20. Определение пути к файлу результатов через 'get\_output\_path'.
- 21. Формирование содержимого выходного файла с количеством перестановок и их парами.
- 22. Запись содержимого выходного файла через `write\_file`.
- 23. Вывод содержимого файла результатов через 'print\_output\_file'.
- 24. Вывод сообщения об ошибке, если входные данные некорректны.
- 25. Вывод времени выполнения программы.

- 26. Вывод объема использованной памяти.
- 27. Остановка отслеживания памяти с помощью 'tracemalloc.stop()'.

Результат работы кода на примере из текста задачи:



Результат работы кода на максимальных и минимальных значениях:



Время выполнения	Затраты памяти
------------------	----------------

Нижняя граница диапазона значений входных данных из текста задачи	0.0005116250249557197 секунд	15324 байт
Пример из задачи	0.0007557080243714154 секунд	15374 байт
Верхняя граница диапазона значений входных данных из текста задачи	0.1820174169843085 секунд	10481917 байт

Вывод по задаче: мною был изучен алгоритм построения пирамиды.

## Задача №7. Снова сортировка

Текст задачи.

## 7 задача. Снова сортировка

Напишите программу пирамидальной сортировки на Python для последовательности в **убывающем порядке**. Проверьте ее, создав несколько рандомных массивов, подходящих под параметры:

- Формат входного файла (input.txt). В первой строке входного файла содержится число n ( $1 \le n \le 10^5$ ) число элементов в массиве. Во второй строке находятся n различных целых чисел, **по модулю** не превосходящих  $10^9$ .
- Формат выходного файла (output.txt). Одна строка выходного файла с отсортированным по невозрастанию массивом. Между любыми двумя числами должен стоять ровно один пробел.

#### Листинг кода.

```
import tracemalloc
t start = time.perf counter()
current dir = os.path.dirname(os.path.abspath( file )) # Директория
txtf dir = os.path.join(os.path.dirname(current dir), "txtf") # Директория
input path = os.path.join(txtf dir, "input.txt")
    largest = i
    right = 2 * i + 2
    if left < n and arr[left] > arr[largest]:
       largest = left
    if right < n and arr[right] > arr[largest]:
       largest = right
        arr[i], arr[largest] = arr[largest], arr[i]
       heapify(arr, n, largest)
    n = len(arr)
```

```
heapify(arr, n, i)
    heapify(arr, i, 0)
\# Читаем данные из файла input.txt с помощью функции open file
lines = open_file(input_path)
m = list(map(int, m_str.split())) # Преобразуем строки в список чисел
    heapsort(m)
    m.reverse()
    output path = get output path(7)
    # Записываем результат в файл output.txt
write_file(" ".join(map(str, m)), output_path)
    print_output file(7)
print("Затрачено памяти:", tracemalloc.get_traced_memory()[1], "байт")
tracemalloc.stop()
```

#### Текстовое объяснение решения.

#### 1 Импорт модулей:

- tracemalloc для отслеживания памяти.
- time для измерения времени выполнения.
- Функции из lab5.utils для работы с файлами.
- os для работы с файловой системой.

#### 2 Инициализация замеров производительности:

- Запуск таймера с помощью time.perf\_counter().
- Включение отслеживания памяти через tracemalloc.start().

#### 3 Определение директорий:

- Получение текущей директории файла.
- Определение пути к директории txtf, содержащей файлы для ввода/вывода.

• Установка пути к файлу input.txt.

#### 4 Реализация функций:

- heapify(arr, n, i):
  - Просеивает элемент вниз, чтобы преобразовать массив в maxheap.
  - о Сравнивает узел с его левым и правым дочерними элементами и выполняет перестановки для сохранения структуры max-heap.
- heapsort(arr):
  - о Преобразует массив в тах-heap.
  - о Последовательно извлекает наибольший элемент, перемещая его в конец массива, и восстанавливает max-heap на оставшейся части.

### 5 Основной блок программы:

- Чтение данных:
  - о Входные данные читаются из файла input.txt с помощью open\_file.
  - Первую строку преобразуют в число n, а вторую в список чисел m.
- Проверка корректности данных:
  - Условие:  $1 \le n \le 1051 \le n \le 105$ , все элементы массива mm находятся в пределах  $-10^9 \le x \le 10^9 10^9 \le x \le 109$ .
  - о Если данные некорректны, выводится сообщение об ошибке.
- Сортировка данных:
  - о Сортировка массива m с использованием heapsort.
  - о Инвертирование массива для получения убывающего порядка.
- Запись результата:
  - о Результат записывается в файл output.txt через write\_file.
  - о Выводится содержимое выходного файла через print\_output\_file.

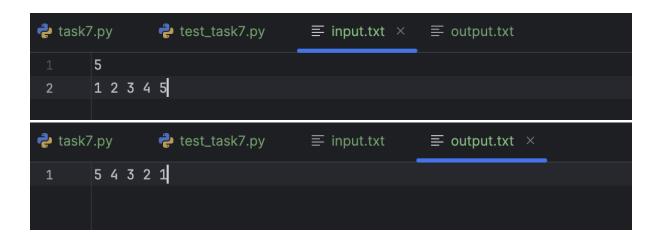
## 6 Вывод результатов производительности:

- Выводится время выполнения программы.
- Выводится максимальный объем памяти, использованной программой.

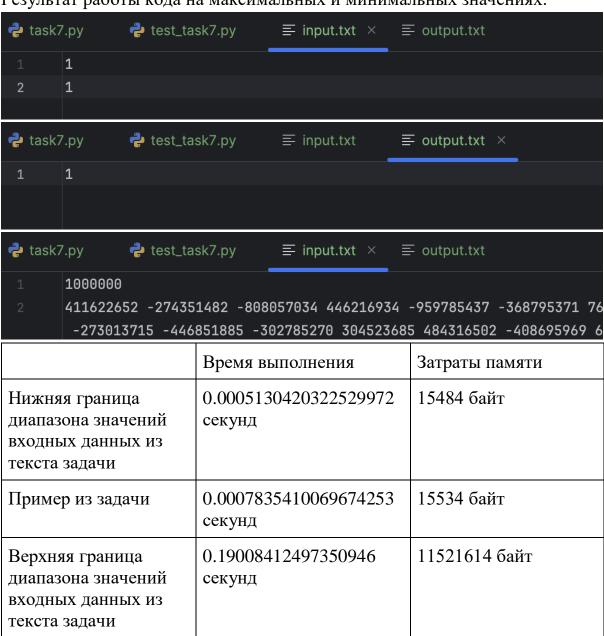
## 7 Завершение отслеживания памяти:

• Выключение отслеживания памяти с помощью tracemalloc.stop().

Результат работы кода на примере из текста задачи:



Результат работы кода на максимальных и минимальных значениях:



Вывод по задаче: мною был изучен алгоритм пирамидальной сортировки.

# Вывод

В ходе лабораторной работы был изучен алгоритм пирамидальной сортировки.