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 14

Выполнила:

Рудникова Виктория Олеговна

K3125

Проверила:

Артамонова В.Е.

Санкт-Петербург 2022 г.

Содержание отчета

Содержание отчета	2
Задачи по варианту	3
Задача №1. Название задачи [N баллов]	3
Дополнительные задачи	4
Задача №1. Название задачи [N баллов]	4
Вывод	5

Задачи по варианту

Задача №1. Сортировка слиянием

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

- Формат входного файла (input.txt). В первой строке входного файла содержится число п ($1 \le n \le 2 \cdot 10^{**}4$) число элементов в массиве. Во второй строке находятся п различных целых чисел, по модулю не превосходящих $10^{**}9$.
- Формат выходного файла (output.txt). Одна строка выходного файла с отсортированным массивом. Между любыми двумя числами должен стоять ровно один пробел.
- Ограничение по времени. 2сек.
- Ограничение по памяти. 256 мб.

```
from time import process_time
from tracemalloc import start, get_traced_memory
```

```
def merge_sort(arr):
    if len(arr) > 1:
        mid = len(arr) // 2
        r = arr[:mid]
        l = arr[mid:]
        merge_sort(l)
        merge_sort(r)
        merge(arr, l, r)
    return arr
```

```
def merge(arr, 1, r):
    i = j = k = 0
    while i < len(l) and j < len(r):
        if l[i] < r[j]:
            arr[k] = l[i]
            i += 1
    else:</pre>
```

```
if __name__ == '__main__':
    start()

with open('input.txt', 'r') as f:
    n = int(f.readline())
    a = [int(i) for i in f.readline().split()]
    a = merge_sort(a)

with open('output.txt', 'w+') as g:
    g.write(' '.join(str(i) for i in a))
    print('Time:', str(process_time()), 'sec')
    print('Memory usage:', str(get_traced_memory()[1] / 1024),
'KB')
```

В функции merge_sort реализуется сначала алгоритм разделения списка на более маленькие списки, а затем вызывается функция merge, которая определяет, в каком порядке соединять маленькие списки в более большие.

Результат работы кода на примерах из текста задачи:

```
      Дополнительные задачи\input.txt ×

        ☐ Задачи по варианту\input.txt ×

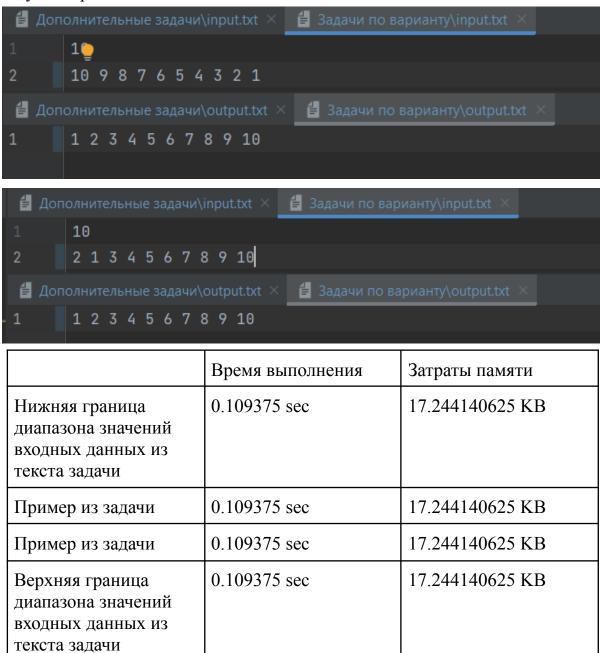
      1
      10

      2
      15 15 15 11 15 234 15 556 15 2

      ☐ Дополнительные задачи\output.txt ×
      ☐ Задачи по варианту\output.txt ×

      1
      2 11 15 15 15 15 15 15 234 556
```

Результат работы кода на максимальных и минимальных значениях:



Вывод по задаче: в задаче я использовала алгоритм сортировки слиянием.

Задача №4. Бинарный поиск

В этой задаче вы реализуете алгоритм бинарного поиска, который позволяет очень эффективно искать (даже в огромных) списках при условии, что список отсортирован. Цель - реализация алгоритма двоичного (бинарного) поиска.

- Формат входного файла (input.txt). В первой строке входного файла содержится число п ($1 \le n \le 10^{**}5$) число элементов в массиве, и последовательность a0 < a1 < ... < a(n-1) из п различных положительных целых чисел в порядке возрастания, $1 \le ai \le 10^{**}9$ для всех $0 \le i < n$. Следующая строка содержит число k, $1 \le k \le 10^{**}5$ и k положительных целых чисел b0, ...b(k-1), $1 \le bj \le 10^{**}9$ для всех $0 \le j < k$.
- Формат выходного файла (output.txt). Для всех і от 0 до k-1 вывести индекс $0 \le j \le n-1$, такой что ai = bj или -1, если такого числа в массиве нет.
- Ограничение по времени. 2сек.
- Ограничение по памяти. 256 мб.

```
from task1 import merge_sort
from time import process_time
from tracemalloc import start, get_traced_memory

def binary_search(arr, l, r, n):
    if l <= r:
        mid = (r - l) // 2 + l
        if arr[mid] == n:
            return mid
        if arr[mid] > n:
            return binary_search(arr, l, mid - l, n)
        return binary_search(arr, mid + l, r, n)
    return -1
```

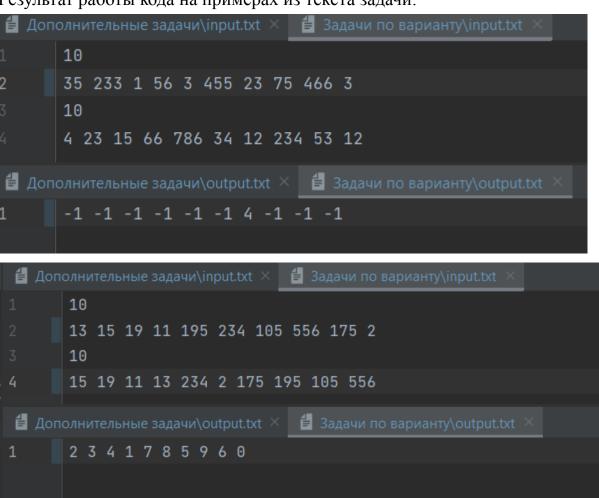
```
if __name__ == '__main__':
    start()

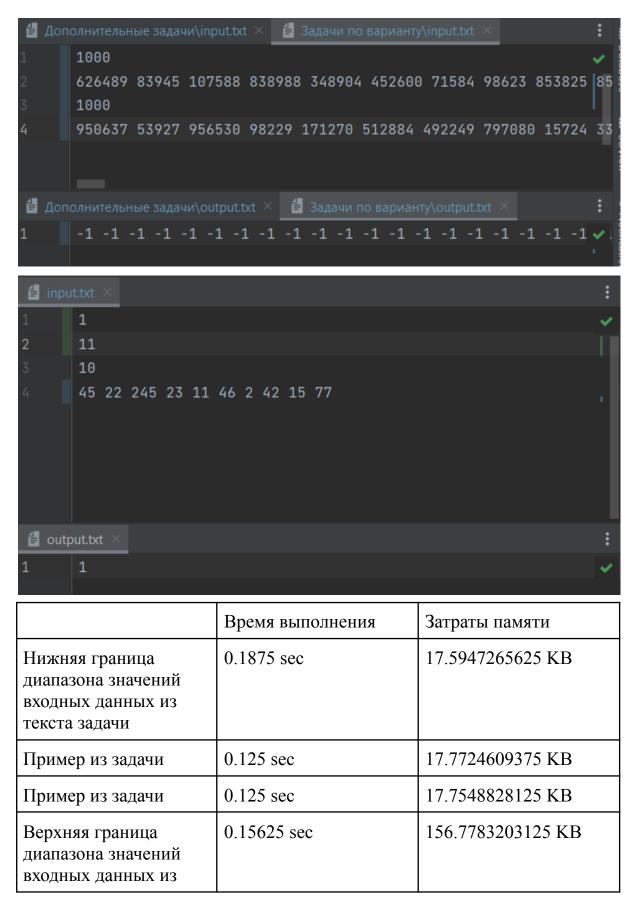
with open('input.txt', 'r') as f:
    n1 = int(f.readline())
    a = [int(i) for i in f.readline().split()]
    n2 = int(f.readline())
    b = [int(i) for i in f.readline().split()]
    b = merge_sort(b)
    idxs = [binary_search(b, 0, len(b) - 1, i) for i in a]
```

```
with open('output.txt', 'w+') as g:
    g.write(' '.join([str(i) for i in idxs]))
    print('Time:', str(process_time()), 'sec')
    print('Memory usage:', str(get_traced_memory()[1] / 1024),
'KB')
```

Массив сортируется с помощью сортировки слиянием, импортированной из первой задачи. В функции binary_search реализуется алгоритм бинарного поиска: искомое число сравнивается с серединой отсортированного списка, и если оно больше, то дальше поиск проводится только в левой половине массива, а если меньше - только в правой. Таким образом, когда в списке остаётся один элемент, выводится его индекс (а если не остаётся, значит искомого элемента нет).

Результат работы кода на примерах из текста задачи:





текста задачи	

Вывод по задаче: в этой задаче я реализовала алгоритм бинарного поиска.

Задача №5. Представитель большинства

Правило большинства - это когда выбирается элемент, имеющий больше половины голосов. Допустим, есть последовательность А элементов a1, a2, ...an, и нужно проверить, содержит ли она элемент, который появляется больше, чем n/2 раз. Ваша цель - использовать метод "Разделяй и властвуй" для разработки алгоритма проверки, содержится ли во входной последовательности элемент, который встречается больше половины раз, за время O(n log n).

- 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 \$ положительных целых чисел, по модулю не превосходящих $10^{**}9, \ 0 \le ai \le 10^{**}9.$
- Формат выходного файла (output.txt). Выведите 1, если во входной последовательности есть элемент, который встречается строго больше половины раз; в противном случае 0.
- Ограничение по времени. 2сек.
- Ограничение по памяти. 256 мб.

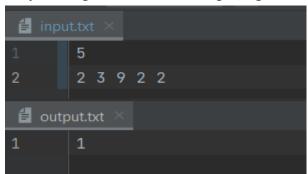
```
rom task1 import merge sort
rom time import process time
rom tracemalloc import start, get traced memory
def left binary search(arr, l, r, n, idx):
  if 1 <= r:
      mid = (r - 1)
           return left binary search(arr, 1, mid - 1, n, mid)
      elif arr[mid] > n:
          return left binary search (arr, 1, mid - 1,
      return left binary search (arr, mid + 1, r, n, idx)
  return idx
   right binary search(arr, 1, r, n, idx):
  if 1 <= r:
      mid = (r - 1) // 2 + 1
      if arr[mid] == n:
          return right binary search(arr, mid + 1, r,
```

right binary search (arr, mid

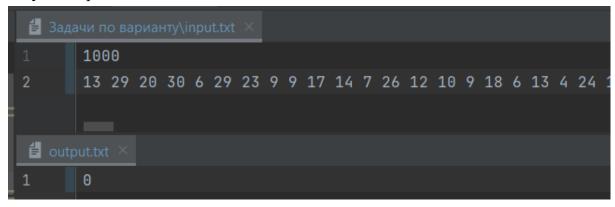
```
def majority(arr, n):
   if left binary search (arr, 0, len (arr) - 1, n, -1) != -1
right binary search (arr,
                                len(arr)
      return 1
  return 0
  start()
  with open('input.txt', 'r') as f:
      a = [int(i) for i in f.readline().split()]
  a = merge sort(a)
      if majority(a, i) == 1:
           flag = True
  with open('output.txt', 'w+') as g:
       if flag:
          g.write('1')
  print('Time:', str(process time()), 'sec')
   print('Memory usage:', str(get traced memory()[1] /
'KB')
```

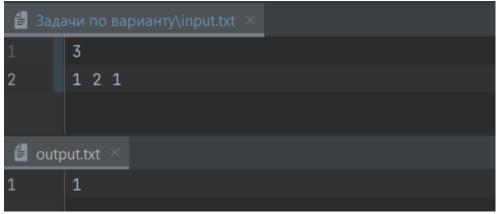
Реализованы функции левого и правого бинарного поиска, которые ищут самый левый и самый правый подходящий элемент соответственно. Выполняется проверка, больше ли половины длины списка разность индекса самого правого и самого левого искомого элемента.

Результат работы кода на примерах из текста задачи:









	Время выполнения	Затраты памяти
Нижняя граница	0.203125 sec	17.5087890625 KB

диапазона значений входных данных из текста задачи		
Пример из задачи	0.234375 sec	17.5166015625 KB
Пример из задачи	0.21875 sec	17.5126953125 KB
Верхняя граница диапазона значений входных данных из текста задачи	0.265625 sec	59.958984375 KB

Вывод по задаче: в задаче я использовала алгоритмы левого и правого бинарного поиска.

Дополнительные задачи

Задача №3. Число инверсий

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.

Подсказка: чтобы сделать это быстрее, можно воспользоваться модификацией сортировки слиянием.

- 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). В выходной файл надо вывести число инверсий в массиве.
- Ограничение по времени. 2сек.
- Ограничение по памяти. 256 мб.

```
def inversions(arr):
    if len(arr) < 2:
        return 0
    mid = len(arr) // 2
    l = arr[:mid]
    r = arr[mid:]
    return inversions(l) + inversions(r) + merge(arr, l, r)</pre>
```

from tracemalloc import start, get traced memory

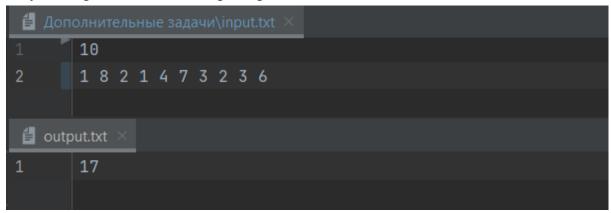
```
def merge(arr, l, r):
```

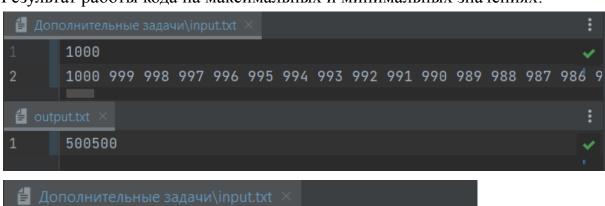
```
i = j = k = 0
while i < len(l) or j < len(r):
    if i == len(l):
        arr[i + j] = r[j]
        j += 1
    elif j == len(r):
        arr[i + j] = l[i]
        i += 1
    elif l[i] <= r[j]:
        arr[i + j] = l[i]
        i += 1
    else:
        arr[i + j] = r[j]
        k += len(l) - i
        j += 1
    return k</pre>
```

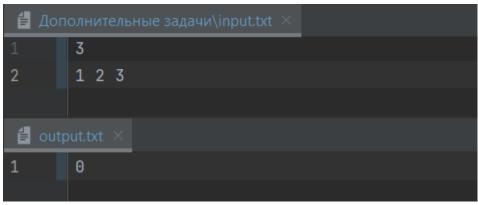
```
if __name__ == '__main__':
    start()
    with open('input.txt', 'r') as f:
        n = int(f.readline())
        a = [int(i) for i in f.readline().split()]
    inv = inversions(a)
    with open('output.txt', 'w+') as g:
        g.write(str(inv))
    print('Time:', str(process_time()), 'sec')
    print('Memory usage:', str(get_traced_memory()[1] / 1024),
'KB')
```

В функции merge считаем инверсии, проверяя соседние элементы соединяемых списков. В функции inversions список делится на маленькие списки, для каждого из которых считается количество инверсий.

Результат работы кода на примерах из текста задачи:







	Время выполнения	Затраты памяти
Нижняя граница диапазона значений входных данных из текста задачи		
Пример из задачи	0.140625 sec	17.537109375 KB
Верхняя граница диапазона значений входных данных из	0.109375 sec	96.921875 KB

текста задачи

Вывод по задаче: в задаче я немного изменила алгоритм сортировки слиянием, чтобы посчитать количество инверсий.

Задача №7. Поиск максимального подмассива за линейное время

Можно найти максимальный подмассив линейное 3a время, воспользовавшись следующими идеями. Начните с левого конца массива и вправо, отслеживая найденный данному максимальный подмассив. Зная максимальный подмассив массива А[1...i], поиск максимального распространите ответ на подмассива, индексом і + 1, воспользовавшись следующим заканчивающегося 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 A[i..i + 1] за константное время, зная максимальный подмассив, заканчивающийся индексом і.

В этом случае у вас возможны 2 варианта тестирования: первый предполагает создание рандомного массива чисел, аналогично задаче №1 (в этом случае формат входного и выходного файла смотрите там). Второй вариант - взять любые данные по акциям какой-либо компании, аналогично задаче №6.

```
from time import process_time
from tracemalloc import start, get traced memory
```

```
def maxSubArray(arr):
    idx = 0
    idx_len = 0
    summ = 0
    max_sum = 0
    for i in range(len(arr)):
        if summ < 0:
            summ = 0
        if summ == 0:
            idx = i
        summ > max_sum:
```

```
max_sum = summ
    idx_len = i

if max_sum == 0:
    idx = -1

    idx_len = -1

return arr[idx: idx_len + 1]

if __name__ == '__main__':
    start()

with open('input.txt', 'r') as f:
    n = f.readline()
    a = [int(i) for i in f.readline().split()]

res = maxSubArray(a)

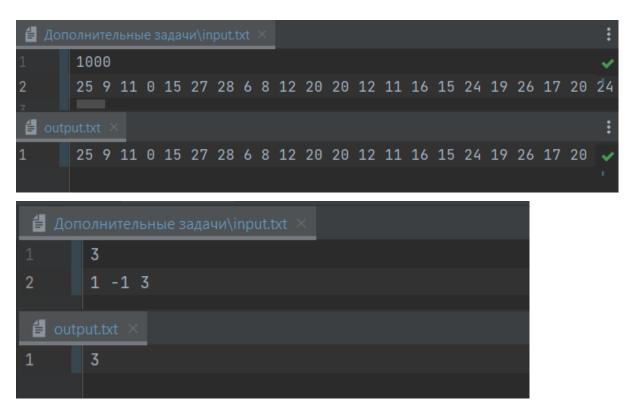
with open('output.txt', 'w+') as g:
    g.write(' '.join([str(i) for i in res]))
    print('Time:', str(process_time()), 'sec')

print('Memory_usage:', str(get_traced_memory()[1] / 1024),
'KB')
```

Проходимся по всем элементам списка один раз и смотрим, не стала ли сумма элементов, включая следующий, меньше, чем была (если стала, прекращаем поиск и запоминаем этот подмассив).

Результат работы кода на примерах из текста задачи:

```
    Дополнительные задачи\input.txt ×
    1 9
    2 2 1 -3 4 -1 2 1 -5 4
    ₫ output.txt ×
    1 4 -1 2 1
```



Проверка задачи на (openedu, астр и тд при наличии в задаче). (скрин)

	Время выполнения	Затраты памяти
Нижняя граница диапазона значений входных данных из текста задачи	0.140625 sec	17.560546875 KB
Пример из задачи	0.125 sec	17.244140625 KB
Верхняя граница диапазона значений входных данных из текста задачи	0.109375 sec	92.58984375 KB

Вывод по задаче: в этой задаче я использовала алгоритм нахождения максимального подмассива за линейное время.

Вывод:

В работе я вспомнила метод "разделяй и властвуй" и некоторые алгоритмы, работающие за логарифмическое время.