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

Отчет по лабораторной работе №3 по курсу «Алгоритмы и структуры данных» Тема: Быстрая сортировка, сортировки за линейное время Вариант 15

Выполнил:

Левахин Лев Александрович

K3140

Проверил:

--

Санкт-Петербург 2024 г.

Содержание отчета

Содержание отчета	1
Задачи по варианту	2
Задача №1. Быстрые сортировки +	
Задача №2. Анти-быстрая сортировка	
Задача №4. Точки и отрезки	
Дополнительные задачи	3
Задача №3. Сортировка пугалом	
Задача №5. Индекс Хирша	
Задача №6. Сортировка целых чисел	
Вывод	4

Задачи по варианту

Задача №1. Быстрые сортировки +

Листинг кода.

1) Быстрая сортировка +

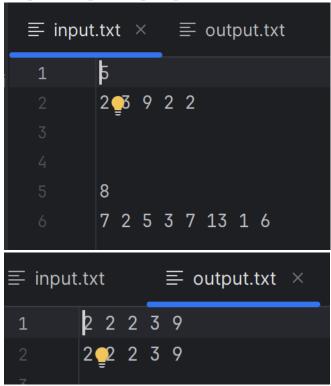
```
from lab3 import utils
    if left < right:</pre>
         quick sort(lst, m + 1, right)
         k = randint(left, right)
lst[left], lst[k] = lst[k], lst[left]
        m = partition(lst, left, right)
    \overline{n}, lst = utils.read file() # Количество элементов и список с
    randomized_quick sort(lst1, 0, n1 - 1)
   utils.write file([lst, lst1])
```

Текстовое объяснение решения.

Делим массив на две части. С помощью функции партишон находим элемент м. Выбираем опорный элемент, и запоминаем начальный, проходимся по остальным элементам и меняем их местами. Возвращаем индекс последней перестановки.

Используем метод декомпозиции и рекурсивно запускаем функцию быстрой сортировки(до м и от м).

Результат работы кода на примерах из текста задачи: (скрины input output файлов)



Тесты к задаче:

Листинг кода:

```
from lab3.task1.src.task1 import quick_sort, randomized_quick_sort
from lab3.task1.src.task2 import randomized_quick_sort_three
from lab3 import utils
import random
import unittest

class TaskTest1(unittest.TestCase):

# Тест функции быстрой сортировки на данных из примера
def test_quick_sort(self):
    """Тест быстрой сортировки на данных из примера"""
# given
    n, lst = utils.read_file()

# when
    utils.test_memory_and_time_lst_n(lst, n, quick_sort, True)
```

```
self.assertEqual(lst, [2, 2, 2, 3, 9])
         n, lst = utils.read file()
         self.assertEqual(lst, [2, 2, 2, 3, 9])
        \overline{n} \ \overline{hud} = 10 * * 5
         self.assertEqual(lst hud, sorted(lst hud))
randomized quick sort, False)
         self.assertEqual(lst hud, sorted(lst hud))
```

```
utils.test memory and time 1st n(1st hud, n hud,
randomized quick sort three, False)
          self.assertEqual(lst hud, sorted(lst hud))
          utils.test memory and time 1st n(res3, n hud,
          self.assertEqual(res1, sorted(res1))
self.assertEqual(res2, sorted(res2))
self.assertEqual(res3, sorted(res3))
          self.assertEqual(lst hud, sorted(lst hud))
    unittest.main()
```

Скрины работы тестов:

2) Быстрая сортировка с разделением на 3 части Листинг кола.

```
m = partition three(lst, left, right)
x = lst[left] # Выбираем "опорный" элемент
j = left
i = left - 1
while j<=p:</pre>
    if lst[j] < x:
        lst[i], lst[j] = lst[j], lst[i]
        lst[j], lst[p] = lst[p], lst[j]
```

```
j += 1

return p

if __name__ == "__main__":
    n, lst = utils.read_file() # Количество элементов и список с
элементами
    randomized_quick_sort_three(lst, 0, n - 1)

res =
[1,2,2,2,1,1,14,5,67,7,5,7,4342,2,34,234,23,423,4,553456,1,1,1,43,324,23,5,23452,1]
    randomized_quick_sort_three(res,0, len(res)-1)

utils.write_file([lst, res])
```

Текстовое объяснение решения.

Аналогичный подход 1ой задачи, но добавляется третий рассматриваемый случай. То есть рассматриваем случаи до ключа, после, и если равен ключу.

Вывод по задаче:

- 1) Быстрая сортировка соответствует названию и работает быстро, однако обычная её версия медленно сортирует одинаковые элементы
- 2) Быструю сортировку можно модернизировать и использовать для сортировки списка с большим количеством одинаковых элементов.
- 3) Случайный выбор элемента иногда может ускорить процесс работы сортировки.

Задача №2. Анти-быстрая сортировка

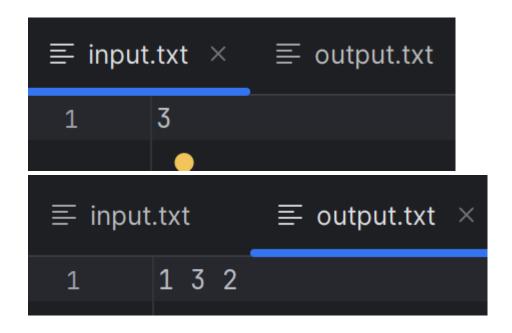
Листинг кода.

```
i = left
\overline{n} = \overline{utils.read} file 1()
```

Текстовое объяснение решения.

Используя алгоритм быстрой сортировки(описанный выше) генирируем из заданной длины списка список, в котором сделается наибольшее количество перестановок. Создаем список, состоящий из п элементов, меняем местами все элементы с шагом і. Возвращаем список.

Результат работы кода на примерах из текста задачи: (скрины input output файлов)



Тесты к задаче:

Листинг кода:

```
from lab3.task2.src.task2 import anti_qsort
from lab3 import utils
import unittest

class TaskTest2(unittest.TestCase):

# Tect функции быстрой сортировки на данных из примера
def test_quick_sort(self):
    """Tect быстрой сортировки на данных из примера"""
    # given
    n = utils.read_file_1()

# when
    utils.test_memory_and_time_lst(n, anti_qsort, True)

# then
    self.assertEqual(anti_qsort(n), [1, 3, 2])

if __name__ == "__main__":
    unittest.main()
```

Скрины работы тестов:

```
Просчитаем время и память работы Сортировки <function anti_qsort at 0x000001BAC72E2F20>
[1, 3, 2]
Итоговое время: 0:00:00
Используемая память: 0.000496 МБ
Память на пике: 0.006181 МБ
```

	Время выполнения	Затраты памяти
Пример из задачи	0:00:00	0.000496 Мб

Вывод по задаче:

1) Чтобы сгенирировать список, на котором быстрая сортировка будет делать максимальное количество перемещений, нужен инверсированный список, в котором произведены дополнительные инверсии.

Задача №4. Точки и отрезки

Листинг кода.

```
file in = open(f"../txtfiles/{file name}.txt")
   cnt segment, cnt dot = map(int, file in.readline().split())
   list with segments, list with dots = [], []
    for segment in range(cnt segment):
            for el in (list(range(a, b + 1))):
                list with segments.append(int(el))
                list with segments.append(int(el))
    return cnt segment, cnt dot, list with segments, list with dots
def lottery_find(list_with_segments:list, list with dots:list) -> list:
```

```
cnt found dot in segments = line search(list with segments,
         res.append(cnt found dot in segments)
    cnt segment, cnt dot, list with segments, list with dots =
make_lottery_data("input")
    res1 = lottery_find(list_with_segments, list_with_dots)
    cnt segment 2, cnt dot 2, list with segments 2, list with dots 2 =
make_lottery_data("input2")
    res2 = lottery_find(list_with_segments_2, list_with_dots_2)
make_lottery_data("input3")
    res3 = lottery find(list with segments 3, list with dots 3)
```

Текстовое объяснение решения.

Создаем функцию, в которой будем обрабатывать входные данные. Принимаем в ней название файла, для которого будем принимать. Возвращаем количество интервалов, количество точек, полный список всех интервалов, список с точками.

Далее, используя линейный поиск, находим количество раз, которое заданное число встречается в обработанных данных.

При запуске файла, выводим результат с помощью функции из utils.

Результат работы кода на примерах из текста задачи: (скрины input output файлов)

```
≡ input.txt ×
                \equiv \equiv input.txt \times \equiv input2.txt \times
         2 3
   1
                           1 3
         0 95
                           -10 10
         7 10
                            -100 100 0
         1 6 11
                  \equiv input2.txt \equiv input3.txt
≡ input.txt
        3 2
 1
        0 €5 √
        7 10
        1 6
     input 1:
     1 0 0
     input 2:
     0 0 1
     input 3:
     2 🗐
```

Тесты к задаче:

Листинг кода:

```
from lab3.task4.src.task4 import *
from lab2 import utils
import datetime
import tracemalloc
import unittest
```

```
cnt segment, cnt dot, list with segments, list with dots =
make lottery data("input")
        cnt segment 2, cnt dot 2, list with segments 2, list with dots 2 =
make lottery data ("input2")
        cnt segment 3, cnt dot 3, list with segments 3, list with dots 3 =
make lottery data("input3")
        print(lottery find(list with segments, list with dots))
        print("Итоговое время:", finish time - start time) # Выводим
        current, peak = tracemalloc.get traced memory() # Присваеваем двум
        res3 = lottery find(list with segments 3, list with dots 3)
        self.assertEqual(res1, [1, 0, 0])
self.assertEqual(res2, [0, 0, 1])
self.assertEqual(res3, [2, 0])
   unittest.main()
```

Скрины работы тестов:

Просчитаем время и память работы алгоритма на данных из примера [1, 0, 0]

Итоговое время: 0:00:00

Используемая память: 0.000328 МБ

Память на пике: 0.004677 МБ

	Время выполнения	Затраты памяти
Пример из задачи	0:00:00	0.000328 Мб

Вывод по задаче:

- 1) Задачу можно решить быстрее используя бинарный поиск
- 2) Большая часть способа решения задачи зависит от метода обработки входных данных.

Дополнительные задачи

Задача №3. Сортировка пугалом

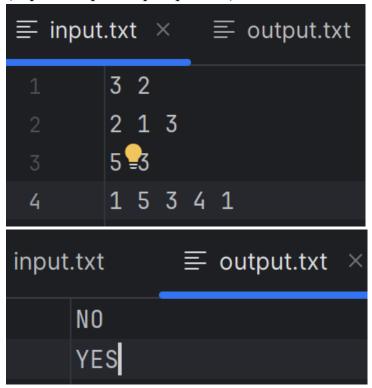
Листинг кода.

```
from lab3 import utils
def scarecrow sort(n:int, k:int, lst:list) -> str:
         swapped = False
                    swapped = True
          if not swapped:
     \overline{\text{file in}} = \overline{\text{open}}("../\overline{\text{txtfiles/input.txt"}})
    res2 = scarecrow sort(n2, k2, 1st2)
```

Текстовое объяснение решения.

Очень необычный вид сортировки, в котором используем перемену элементов местами, чтобы отслеживать, можно ли отсортировать список, сравнивая его с уже отсортированным. Проходимся по элементам на заданном расстоянии, и смотрим, отсортирован ли список при смене элементов местами. Используем переменную-флаг для отслеживания.

Результат работы кода на примерах из текста задачи: (скрины input output файлов)



Тесты к задаче:

Листинг кода:

```
from lab3.task3.src.task3 import scarecrow_sort
from lab3 import utils
import datetime
import tracemalloc
import unittest

class TaskTest3(unittest.TestCase):

# Тест функции быстрой сортировки на данных из примера
def test_scarecrow_sort(self):
    """Тест сортировки Пугалом на данных из примера"""
    # given
    file_in = open("../txtfiles/input.txt")

    n1, k1 = map(int, file_in.readline().split())
    lst1 = list((map(int, file_in.readline().split())))

    n2, k2 = map(int, file_in.readline().split())
    lst2 = list((map(int, file_in.readline().split())))
    file_in.close()

# when
    print("Просчитаем время и память работы алгоритма")
```

```
tracemalloc.start() # Запускаем счётчик памяти start_time = datetime.datetime.now() # Запускаем счётчик времени print(scarecrow_sort(n1,k1, lst1))

finish_time = datetime.datetime.now() # Измеряем время конца paботы print("Итоговое время:", finish_time - start_time) # Выводим итоговое время

current, peak = tracemalloc.get_traced_memory() # Присваеваем двум переменным память, используемую сейчас, и на пике print(
 f"Используемая память: {current / 10 ** 6} МБ\nПамять на пике: {peak / 10 ** 6} МБ\n") # Выводим время работы в мегабайтах

# then self.assertEqual(scarecrow_sort(n1,k1, lst1), "NO") self.assertEqual(scarecrow_sort(n2, k2, lst2), "YES")

if __name__ == "__main__": unittest.main()
```

Скрины работы тестов:

Просчитаем время и память работы алгоритма NO

Итоговое время: 0:00:00

Используемая память: 0.000208 МБ

Память на пике: 0.006021 МБ

	Время выполнения	Затраты памяти (Мб)
Значения из примера	0:00:00	0.000208

Вывод по задаче:

1) Сортировка пугалом легко отображается на реальной жизни.

Задача №5. Индекс Хирша

Листинг кода.

```
###

Ahtu-быстрая copтировка

from lab3 import utils

def index_harsh(citations: list) -> int:
    """

Индекс Хирша
    - Принимает: (список - каждая цифра списка - количество использований

статьи)
    - Возвращает: (число - индекс Хирша учёного)
    """

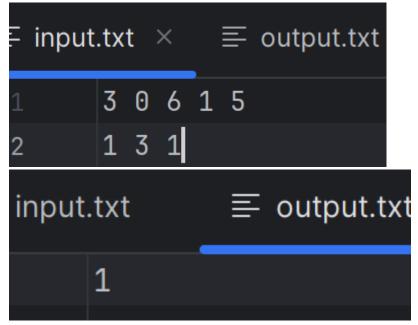
for h in range(len(citations), 0, -1):
    cnt_citation = 0
    for el in citations:
        if el >= h:
            cnt_citation += 1
        if cnt_citation >= h:
            return h

if __name__ == "__main__":
    lst = utils.read_file_1_list()
    res = index_harsh(lst)
    utils.write_file([res])
```

Текстовое объяснение решения.

Перебираем индексы в обратном порядке(от максимума до 0). Смотрим на количество элементов, если оно больше или равно индексу, то прибавляем количество цитирований. Возвращаем индекс Хирша учёного.

Результат работы кода на примерах из текста задачи: (скрины input output файлов)



Тесты к задаче:

Листинг кода:

```
from lab3.task5.src.task5 import index_harsh
from lab3 import utils
import unittest

class TaskTest5(unittest.TestCase):

# Tect функции быстрой сортировки на данных из примера
def test_quick_sort(self):
    """Тест быстрой сортировки на данных из примера"""
    # given
    lst1 = utils.read_file_1_list()
    lst2 = [1, 3, 1]

# when
    utils.test_memory_and_time_lst(lst1, index_harsh, True)

# then
    self.assertEqual(index_harsh(lst1), 3)
    self.assertEqual(index_harsh(lst2), 1)

if __name__ == "__main__":
    unittest.main()
```

```
Скриныработытестов:Просчитаем время и память работы Сортировки <function index_harsh at 0x000001EA73C82DE0>3Итоговое время: 0:00:00Используемая память: 0.000208 МБ
```

	Время выполнения	Затраты памяти (Мб)
Сортировка вставкой средние значения	0:00:00	0.000208

Вывод по задаче:

Память на пике: 0.005893 МБ

1) Используем двойной перебор для поиска индекса Хирша.

Задача №6. Сортировка целых чисел

Листинг кода.

```
n, m = map(int, file in.readline().split())
 lst b = list(map(int, file in.readline().split()))
 n, m, lst_a, lst_b = read_input_file_2_numbers_2_lists()
 lst c = get data c(lst a, lst b)
 res = sort z numbers(lst c)
utils.write file([res])
```

Текстовое объяснение решения.

Создаём функцию, которая обрабатывает входные данные. Также используем дополнительную функцию, которая перемножает одинаковые по индексу элементы из двух списков. Используем быструю сортировку для сортировки массива, возвращаем сумму каждого 10-го элемента в отсортированном списке.

Результат работы кода на примерах из текста задачи: (скрины input output файлов)

Тесты к задаче:

Листинг кода:

```
from lab3.task6.src.task6 import *
from lab3 import utils
import unittest

class TaskTest6(unittest.TestCase):

# Тест функции быстрой сортировки на данных из примера
def test_quick_sort(self):
    """Тест быстрой сортировки на данных из примера"""
    # given
    n, m, lst_a, lst_b = read_input_file_2_numbers_2_lists()
    lst_c = get_data_c(lst_a, lst_b)

# when
    utils.test_memory_and_time_lst(lst_c, sort_z_numbers, True)

# then
    self.assertEqual(sort_z_numbers(lst_c), 51)

if __name__ == "__main__":
    unittest.main()
```

Скрины работы тестов:

```
Просчитаем время и память работы Сортировки <function sort_z_numbers at 0x000001F8483ED260>
51
Итоговое время: 0:00:00
Используемая память: 0.000208 МБ
Память на пике: 0.005893 МБ
```

	Время выполнения	Затраты памяти
Пример из задачи	0:00:00	0.000208 Мб

Вывод по задаче:

- 1) Чтобы решить сложную задачу, используем несколько маленьких разных функций, которые в дальнейшем используем для большей.
- 2) Обрабатывать входные данные можно по-разному.

Вывод

Существует множество видов сортировок, как самые оптимизированные, так и самые неоптимизированные. С ними можно совершать огромное количество действий и решать множество задач, также подстраивая их под свои цели. Они отличаются между собой по времени выполнения и затрачиваемой памяти. Чтобы использовать много видов сортировок и подбирать их под свои нужды, нужно понимать логику работы таких сортировок, о которых идёт речь в данной лабораторной. Я постарался использовать все свои знания, чтобы грамотно написать и применить их. Существует множество задач, которые пересекаются между собой и используют ранее написанные методы или сортировки.