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

Отчет по лабораторной работе №6 по курсу «Алгоритмы и структуры данных» Тема: Хеширование. Хеш-таблицы Вариант 5

Выполнил:

Артемов И. В.

K3141

Проверил:

Афанасьев А. В.

Санкт-Петербург 2024 г.

Содержание отчета

Содержание отчета	2
Задачи по варианту	3
Задача №1. Множество	3
Задача №2: Телефонная книга	7
Задача №4. Прошитый ассоциативный массив	13
Задача №7. Драгоценные камни	21
Вывол	26

Задачи по варианту

Задача №1. Множество

Текст задачи.

1 задача. Множество

Реализуйте множество с операциями «добавление ключа», «удаление ключа», «проверка существования ключа».

- Формат входного файла (input.txt). В первой строке входного файла находится строго положительное целое число операций N, не превышающее $5\cdot 10^5$. В каждой из последующих N строк находится одна из следующих операций:
 - А x добавить элемент x в множество. Если элемент уже есть в множестве, то ничего делать не надо.
 - D x удалить элемент x. Если элемента x нет, то ничего делать не нало.
 - ?x если ключ x есть в множестве, выведите «Y», если нет, то выведите «N».

Аргументы указанных выше операций – **целые числа**, не превышающие по модулю 10^{18} .

- Формат выходного файла (output.txt). Выведите последовательно результат выполнения всех операций «?». Следуйте формату выходного файла из примера.
- Ограничение по времени. 2 сек.
- Ограничение по памяти. 256 мб.
- Пример:

input.txt	output.txt
8	Y
A 2	N
A 5	N
A 3	
? 2	
? 4	
A 2	
D 2	
? 2	

Листинг кода.

```
# Импортируем библиотеки для отслеживания памяти и времени выполнения программы import tracemalloc import time import os from lab6.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()

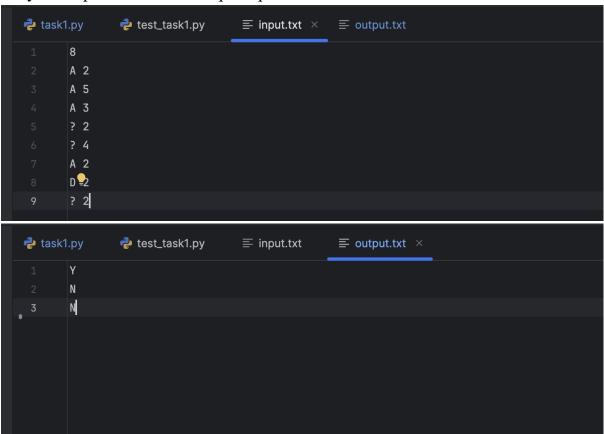
# Включаем отслеживание памяти tracemalloc.start()
```

```
current dir = os.path.dirname(os.path.abspath( file )) # Директория
txtf dir = os.path.join(os.path.dirname(current dir), "txtf") # Директория
    for operation in operations:
       cmd, x = operation.split()
            if x in current set:
               results.append("Y")
                results.append("N")
    return results
   data = open file(input path)
   n = int(data[0]) # Первая строка содержит количество операций
   operations = data[1:] \# Последующие строки — это операции
10**18 for op in operations):
       print(f"\nTask: 1\nInput:\n{n}\n{operations}")
       results = process operations(n, operations)
       output path = get output path(1)
       write file("\n".join(results), output_path)
       print output file(1)
    print("Время работы: %s секунд" % (time.perf counter() - t start))
    print("Затрачено памяти:", tracemalloc.get traced memory()[1], "байт")
    tracemalloc.stop()
```

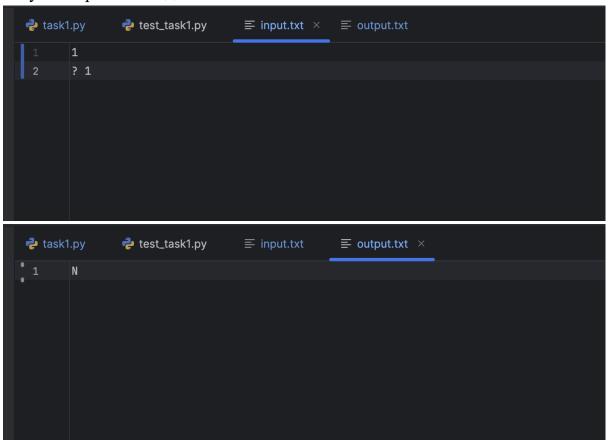
Текстовое объяснение решения.

- 1. Импортируются библиотеки для отслеживания памяти и времени.
- 2. Запускается таймер для измерения времени работы программы.
- 3. Включается отслеживание памяти.
- 4. Устанавливаются пути к файлам.
- 5. Определяется функция `process_operations`, которая выполняет операции над множеством.
- 6. В основном блоке программы:
 - Читаются данные из файла 'input.txt'.
 - В первой строке извлекается количество операций 'n'.
 - В последующих строках извлекаются операции.
- 7. Проверяется корректность входных данных (ограничения на количество операций и их формат).
- 8. Если данные корректны, выводится информация о задаче и входных данных.
- 9. Выполняются операции над множеством с помощью функции `process_operations`.
- 10. Результаты операций записываются в файл 'output.txt'.
- 11. Выводится информация о времени работы программы.
- 12. Выводится информация о затраченной памяти.
- 13. Останавливается отслеживание памяти.

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	0.0007107499995981925 секунд	15368 байт
Пример из задачи	0.0006332500006465125 секунд	15715 байт
Верхняя граница диапазона значений входных данных из текста задачи	0.4109084579977207 секунд	10765530 байт

Вывод по задаче: мною был изучен алгоритм реализации множества с некоторыми операциями.

Задача №2: Телефонная книга

Текст задачи.

2 задача. Телефонная книга

В этой задаче ваша цель - реализовать простой менеджер телефонной книги. Он должен уметь обрабатывать следующие типы пользовательских запросов:

- add number name это команда означает, что пользователь добавляет в телефонную книгу человека с именем name и номером телефона number. Если пользователь с таким номером уже существует, то ваш менеджер должен перезаписать соответствующее имя.
- del number означает, что менеджер должен удалить человека с номером из телефонной книги. Если такого человека нет, то он должен просто игнорировать запрос.
- find number означает, что пользователь ищет человека с номером телефона number. Менеджер должен ответить соответствующим именем или строкой «not found» (без кавычек), если такого человека в книге нет.
- Формат ввода / входного файла (input.txt). В первой строке входного файла содержится число N ($1 \le N \le 10^5$) количество запросов. Далее следуют N строк, каждая из которых содержит один запрос в формате, описанном выше.

Все номера телефонов состоят из десятичных цифр, в них нет нулей в начале номера, и каждый состоит не более чем из 7 цифр. Все имена представляют собой непустые строки из латинских букв, каждая из которых имеет длину не более 15. Гарантируется при проверке, что не будет человека с именем «not found».

- Формат вывода / выходного файла (output.txt). Выведите результат каждого поискового запроса find имя, соответствующее номеру телефона, или «not found» (без кавычек), если в телефонной книге нет человека с таким номером телефона. Выведите по одному результату в каждой строке в том же порядке, как были заданы запросы типа find во входных данных.
- Ограничение по времени. 6 сек.
- Ограничение по памяти. 512 мб.
- Примеры:

	input.txt
	12
	add 911 police
	add 76213 Mom
	add 17239 Bob
	find 76213
	find 910
	find 911
	del 910
	del 911
l:	find 911
	find 76213
	add 76213 daddy
	find 76213
	output.txt
	Mom
	not found
	police
	not found
	Mom
	daddy

input.txt	
8	
find 3839442	
add 123456 me	
add 0 granny	
find 0	
find 123456	
del 0	
del 0	
find 0	
output.txt	
not found	
granny	
me	
not found	

2:

Описание примера 1. 76213 - это номер Мот, 910 - нет в телефонной книге, 911 - это номер police, но затем он был удален из телефонной книги, поэтому второй поиск 911 вернул «not found». Также обратите внимание, что когда daddy был добавлен с тем же номером телефона 76213, что и номер телефона Мот, имя контакта было переписано, и теперь поиск 76213 возвращает «daddy» вместо «Мот».

Листинг кода.

```
import tracemalloc
import time
import os
from lab6.utils import open_file, write_file, get_output_path,
delete_prev_values, print_output_file

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

# Включаем отслеживание памяти
tracemalloc.start()

# Пути к входному и выходному файлам
current_dir = os.path.dirname(os.path.abspath(__file__)) # Текущая
директория
txtf_dir = os.path.join(os.path.dirname(current_dir), "txtf") # Директория
txtf
input_path = os.path.join(txtf_dir, "input.txt")
def is_valid_name(name):
```

```
return name.isalpha() and len(name) <= 15</pre>
def is valid number(number):
    return number.isdigit() and 1 <= len(number) <= 7 and number[0] != '0'
    results = []
            number, name = command[1], command[2]
if is_valid_number(number) and is_valid_name(name):
                 results.append("invalid") # Для отладки можно использовать
            number = command[1]
             if is valid number(number):
                phone_book.pop(number, None)
            number = command[1]
             if is valid number(number):
                 results.append(phone book.get(number, "not found"))
                 results.append("not found") # Некорректный номер считается
    \# Читаем данные из файла input.txt
    lines = open file(input path)
    n = int(lines[0].strip()) # Количество запросов
    queries = lines[1:] # Список запросов
        results = process phone book(queries)
```

```
# Формируем путь к выходному файлу output_path = get_output_path(1)

# Записываем результаты в файл output.txt write_file("\n".join(results), output_path) print_output_file(1)

else:
    print("Введите корректные данные")

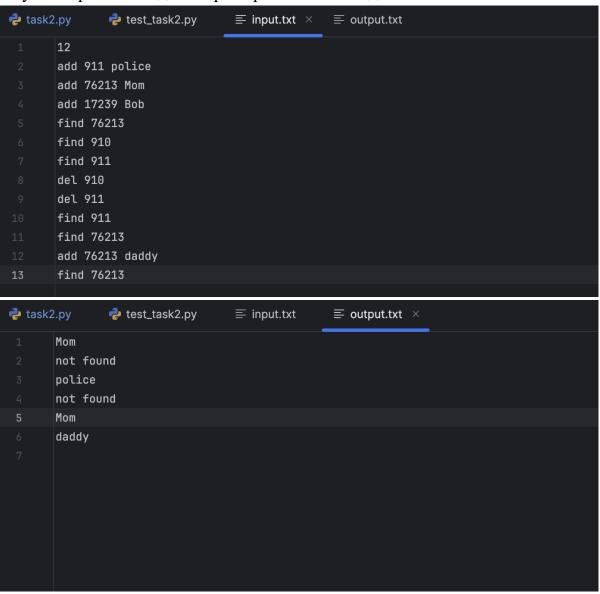
# Выводим время работы программы print("Время работы: %s секунд" % (time.perf_counter() - t_start)) # Выводим количество памяти, затраченной на выполнение программы print("Затрачено памяти:", tracemalloc.get_traced_memory()[1], "байт")

# Останавливаем отслеживание памяти tracemalloc.stop()
```

Текстовое объяснение решения.

- 1. Импортируются библиотеки для отслеживания памяти и времени.
- 2. Запускается таймер для измерения времени работы программы.
- 3. Включается отслеживание памяти.
- 4. Устанавливаются пути к файлам.
- 5. Определяются функции:
 - 'is valid name' для проверки валидности имени.
 - `is_valid_number` для проверки валидности номера телефона.
 - `process_phone_book` для обработки запросов телефонной книги.
- 6. В основном блоке программы:
 - Читаются данные из файла `input.txt`.
 - В первой строке извлекается количество запросов 'n'.
 - В последующих строках извлекаются запросы.
- 7. Проверяется корректность входных данных (ограничения на количество запросов).
- 8. Если данные корректны, выводится информация о задаче и входных данных.
- 9. Обрабатываются запросы с помощью функции `process_phone_book`.
- 10. Формируется путь к выходному файлу.
- 11. Результаты записываются в файл `output.txt`.
- 12. Выводится информация о времени работы программы.
- 13. Выводится информация о затраченной памяти.
- 14. Останавливается отслеживание памяти.

Результат работы кода на примере из текста задачи:



Результат работы кода на максимальных и минимальных значениях:

```
task2.py test_task2.py input.txt × input.t
```



	Время выполнения	Затраты памяти
Нижняя граница диапазона значений входных данных из текста задачи	0.0006144999997559353 секунд	15664 байт
Пример из задачи	0.000861916000758356 секунд	16335 байт
Верхняя граница диапазона значений входных данных из текста задачи	0.4109084579977207 секунд	430401 байт

Вывод по задаче: мною был изучен алгоритм реализации простого менеджера телефонной книги.

Задача №4. Прошитый ассоциативный массив

Текст задачи.

4 задача. Прошитый ассоциативный массив

Реализуйте прошитый ассоциативный массив. Ваш алгоритм должен поддерживать следующие типы операций:

- get x если ключ x есть в множестве, выведите соответствующее ему значение, если нет, то выведите <none>.
- prev x вывести значение, соответствующее ключу, находящемуся в ассоциативном массиве, который был вставлен позже всех, но до x, или <none>,

если такого нет или в массиве нет x.

- next x вывести значение, соответствующее ключу, находящемуся в ассоциативном массиве, который был вставлен раньше всех, но после x, или <none>, если такого нет или в массиве нет x.
- put x y поставить в соответствие ключу x значение y. При этом следует учесть, что
 - если, независимо от предыстории, этого ключа на момент вставки в массиве не было, то он считается только что вставленным и оказывается самым последним среди добавленных элементов – то есть, вызов next с этим же ключом сразу после выполнения текущей операции put должен вернуть <none>;
 - если этот ключ уже есть в массиве, то значение необходимо изменить, и в этом случае ключ не считается вставленным еще раз, то есть, не меняет своего положения в порядке добавленных элементов.
- delete x удалить ключ x. Если ключа в ассоциативном массиве нет, то ничего делать не надо.
- Формат входного файла (input.txt). В первой строке входного файла находится строго положительное целое число операций N, не превышающее $5 \cdot 10^5$. В каждой из последующих N строк находится одна из приведенных выше операций. Ключи и значения операций строки из латинских букв длиной не менее одного и не более 20 символов.
- Формат выходного файла (output.txt). Выведите последовательно результат выполнения всех операций get, prev, next. Следуйте формату выходного файла из примера.
- Ограничение по времени. 4 сек.
- Ограничение по памяти. 256 мб.
- Примеры:

input.txt	output.txt
14	С
put zero a	b
put one b	d
put two c	c
put three d	a
put four e	e
get two	<none></none>
prev two	
next two	
delete one	
delete three	
get two	
prev two	
next two	
next four	

• P.s. Задача на openedu, 8 неделя.

Листинг кода.

```
import tracemalloc
import time
from collections import OrderedDict
from lab6.utils import *

# 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__)) # Директория
task/src
txtf_dir = os.path.join(os.path.dirname(current_dir), "txtf") # Директория
task/txtf
input_path = os.path.join(txtf_dir, "input.txt")

def process_commands(commands):
    """
    Oбрабатывает команды работы с ассоциативным массивом.
    """
    assoc_array = OrderedDict()
    results = []

for line in commands:
    parts = line.strip().split()
    command = parts[0]

    if command == "but":
```

```
x, y = parts[1], parts[2]
            if x in assoc array:
                assoc array[x] = y
                assoc array[x] = y
            x = parts[1]
            results.append(assoc array.get(x, "<none>"))
            x = parts[1]
                keys = list(assoc array.keys())
                    results.append(assoc array[keys[idx - 1]])
                    results.append("<none>")
                results.append("<none>")
            x = parts[1]
            if x in assoc array:
                keys = list(assoc array.keys())
                idx = keys.index(x)
                if idx < len(keys) - 1:</pre>
                    results.append(assoc array[keys[idx + 1]])
                    results.append("<none>")
                results.append("<none>")
            x = parts[1]
            assoc array.pop(x, None)
    return results
   \# Читаем данные из файла input.txt с помощью функции open_file
   lines = open file(input path)
   if 1 \le n \le 5 * 10 ** 5 and all(len(cmd.split()) in (2, 3) for cmd in
commands):
        print(f"\nTask 4\nInput:\n{lines}")
        delete prev values (4)
        results = process commands(commands)
```

```
write_file("\n".join(results), output_path)
    print_output_file(4)

else:
    # Выводим сообщение об ошибке, если данные некорректны
    print("Введите корректные данные")

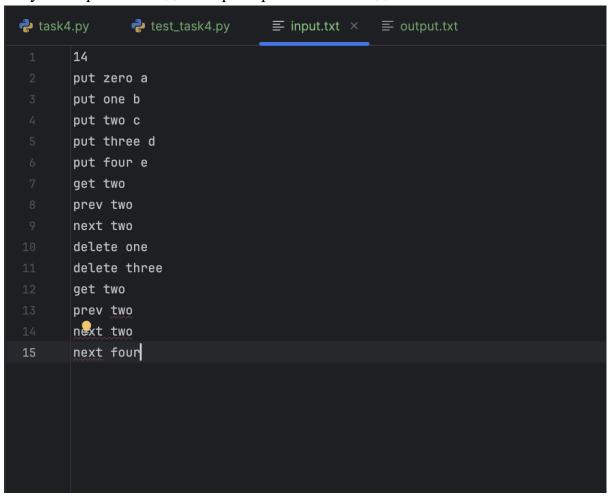
# Выводим время работы программы
print("Время работы: %s секунд" % (time.perf_counter() - t_start))
# Выводим количество памяти, затраченной на выполнение программы
print("Затрачено памяти:", tracemalloc.get_traced_memory()[1], "байт")

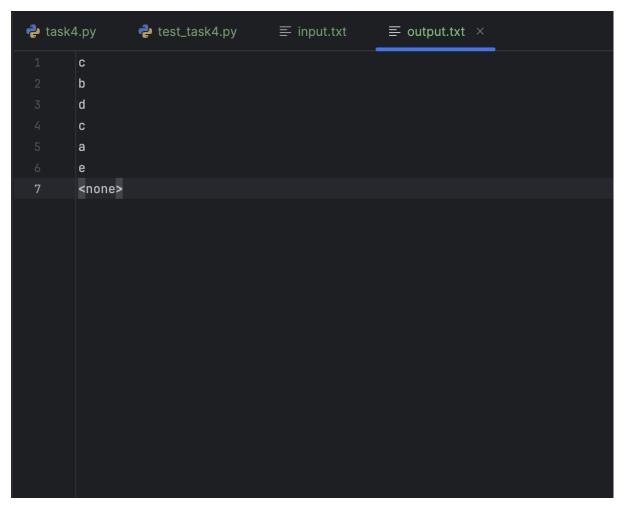
# Останавливаем отслеживание памяти
tracemalloc.stop()
```

Текстовое объяснение решения.

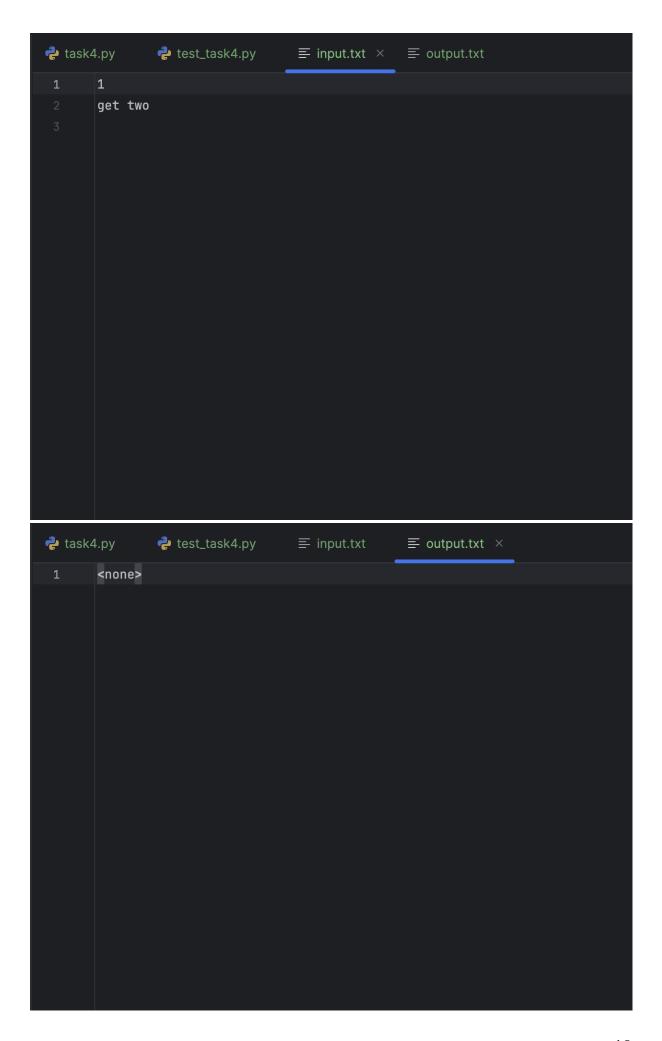
- 1. Импортируются библиотеки для отслеживания памяти и времени, а также `OrderedDict` для ассоциативного массива.
- 2. Запускается таймер для измерения времени работы программы.
- 3. Включается отслеживание памяти.
- 4. Устанавливаются пути к файлам.
- 5. Определяется функция 'process_commands', которая выполняет команды над ассоциативным массивом:
 - `put`: добавление или обновление ключа.
 - 'get': получение значения по ключу.
 - `prev`: поиск предыдущего ключа.
 - `next`: поиск следующего ключа.
 - `delete`: удаление ключа.
- 6. В основном блоке программы:
 - Читаются данные из файла 'input.txt'.
 - В первой строке извлекается количество операций 'n'.
 - В последующих строках извлекаются команды.
- 7. Проверяется корректность входных данных (ограничения на количество операций и формат команд).
- 8. Если данные корректны, выводится информация о задаче и входных данных.
- 9. Обрабатываются команды с помощью функции 'process_commands'.
- 10. Результаты записываются в файл `output.txt`.
- 11. Выводится информация о времени работы программы.
- 12. Выводится информация о затраченной памяти.
- 13. Останавливается отслеживание памяти.

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	0.0004989579992979998 секунд	15341 байт
Пример из задачи	0.0008108750007522758 секунд	16102 байт
Верхняя граница диапазона значений входных данных из текста задачи	0.1820174169843085 секунд	10481917 байт

Вывод по задаче: мною был изучен алгоритм реализации прошитого ассоциативного массива.

Задача №7. Драгоценные камни

Текст задачи.

7 задача. Драгоценные камни

В одной далекой восточной стране до сих пор по пустыням ходят караваны верблюдов, с помощью которых купцы перевозят пряности, драгоценности и дорогие ткани. Разумеется, основная цель купцов состоит в том, чтобы подороже продать имеющийся у них товар. Недавно один из караванов прибыл во дворец одного могущественного шаха.

Купцы хотят продать шаху п драгоценных камней, которые они привезли с собой. Для этого они выкладывают их перед шахом в ряд, после чего шах оценивает эти камни и принимает решение о том, купит он их или нет. Видов драгоценных камней на Востоке известно не очень много всего 26, поэтому мы будем обозначать виды камней с помощью строчных букв латинского алфавита. Шах обычно оценивает камни следующим образом. Он заранее определил несколько упорядоченных пар типов камней: $(a_1,b_1), (a_2,b_2), ..., (a_k,b_k)$. Эти пары он называет красивыми, их множество мы обозначим как P. Теперь представим ряд камней, которые продают купцы, в виде строки S длины n из строчных букв латинского алфавита. Шах считает число таких пар (i,j), что $1 \le i < j \le n$, а камни S_i и S_j образуют красивую пару, то есть существует такое число $1 \le q \le k$, что $S_i = a_q$ и $S_i = b_q$.

Если число таких пар оказывается достаточно большим, то шах покупает все камни. Однако в этот раз купцы привезли настолько много камней, что шах не может посчитать это число. Поэтому он вызвал своего визиря и поручил ему этот подсчет. Напишите программу, которая находит ответ на эту задачу.

- Формат ввода / входного файла (input.txt). Первая строка входного файла содержит целые числа n и k ($1 \le n \le 100000, 1 \le k \le 676$) число камней, которые привезли купцы и число пар, которые шах считает красивыми. Вторая строка входного файла содержит строку S, описывающую типы камней, которые привезли купцы.
 - Далее следуют k строк, каждая из которых содержит две строчных буквы латинского алфавита и описывает одну из красивых пар камней.
- Формат вывода / выходного файла (output.txt). В выходной файл выведите ответ на задачу количество пар, которое должен найти визирь.
- Ограничение по времени. 1 сек.

- Ограничение по памяти. 64 мб.
- Примеры:

№	input.txt	output.txt
1	7 1	6
	abacaba	
	aa	
2	7 3	7
	abacaba	
	ab	
	ac	
	bb	

Листинг кода.

```
Импортируем библиотеки для отслеживания памяти и времени выполнения
t start = time.perf counter()
current dir = os.path.dirname(os.path.abspath( file )) # Текущая
txtf dir = os.path.join(os.path.dirname(current dir), "txtf") # Папка с
input path = os.path.join(txtf dir, "input.txt")
def count_beautiful_pairs(n, k, S, beautiful_pairs):
        for a, b in beautiful pairs:
                 result += count.get(a, 0)
    input data = open file(input path)
    n, k = map(int, input_data[0].split())
S = input_data[1].strip()
```

```
pairs = input_data{2:} # Читаем пары

# Формируем множество красивых пар
beautiful_pairs = set()
for pair in pairs:
    beautiful_pairs.add((pair[0], pair[1]))

print(beautiful_pairs)

# Проверяем корректность входных данных
if (1 <= n <= 100000) and (1 <= k <= 676) and len(S) == n:
    print(f"\nTask 7\nInput:\n{n} {k}\n{S}\n{beautiful_pairs}")
    delete_prev_values(7)

# Считаем количество красивых пар
    result = count_beautiful_pairs(n, k, S, beautiful_pairs)

    output_path = get_output_path(7)
    # Записываем результат в файл output.txt
    write_file(str(result), output_path)
    print_output_file(7)

else:
    # Сообщаем об ошибке, если данные некорректны
    print('Введите корректные данные')

# 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")

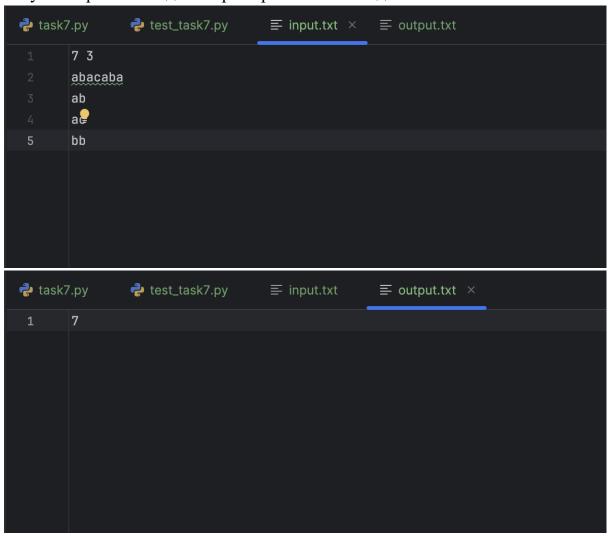
# Останавливаем отслеживание памяти
tracemalloc.stop()</pre>
```

Текстовое объяснение решения.

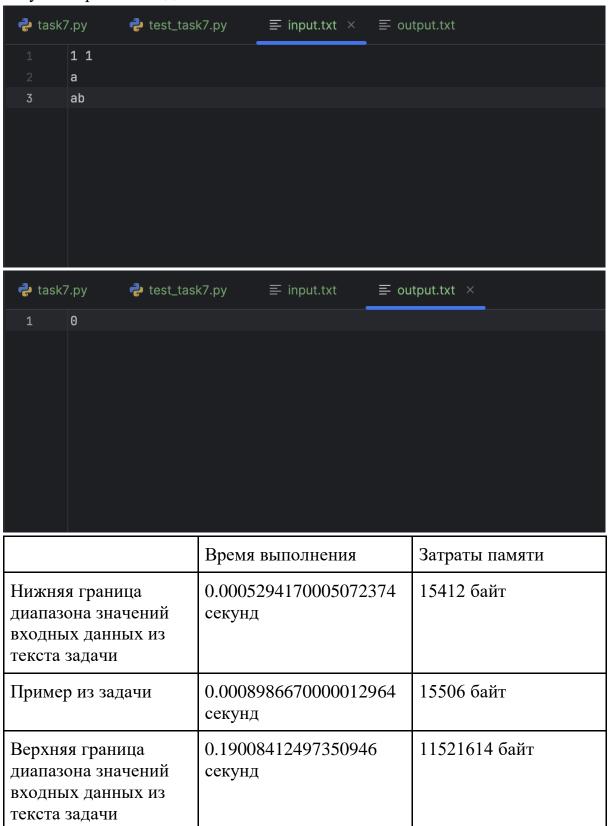
- 1. Импортируются библиотеки для отслеживания памяти и времени, а также функции из `lab6.utils`.
- 2. Запускается таймер для измерения времени работы программы.
- 3. Включается отслеживание памяти.
- 4. Устанавливаются пути к файлам.
- 5. Определяется функция 'count beautiful pairs', которая:
- Принимает количество символов, количество пар, строку символов и множество красивых пар.
- Подсчитывает количество красивых пар в строке, проходя по ней и проверяя, образует ли текущий символ пару с ранее встреченными.
- 6. В основном блоке программы:
 - Читаются данные из файла `input.txt`.
 - Извлекаются значения 'n', 'k', строка 'S' и пары.
- 7. Формируется множество красивых пар.
- 8. Проверяется корректность входных данных (ограничения на количество символов, пар и длину строки).
- 9. Если данные корректны, выводится информация о задаче и входных данных.

- 10. Вычисляется количество красивых пар с помощью функции `count_beautiful_pairs`.
- 11. Результат записывается в файл 'output.txt'.
- 12. Выводится информация о времени работы программы.
- 13. Выводится информация о затраченной памяти.
- 14. Останавливается отслеживание памяти.

Результат работы кода на примере из текста задачи:



Результат работы кода на максимальных и минимальных значениях:



Вывод по задаче: мною был изучен алгоритм поиска "красивых пар".

Вывод

В ходе лабораторной работы я узнал про хеширование и хеш-таблицы.