

«АЛГОРИТМЫ И СТРУКТУРЫ ДАННЫХ»

ЛАБОРАТОРНАЯ РАБОТА №1

Сортировки. Определение сложности алгоритма

Автор: Гуков Сергей Юрьевич

Версия: 1.0 (01.09.2024)

*Для допуска к экзамену необходимо **посетить минимум 50% лекций и сдать любые пять из восьми лабораторных работ до конца зачетной недели.** Для получения за экзамен автоматом оценки «Удовлетворительно» или «Хорошо» необходимо **посетить минимум 50% лекций и сдать все лабораторные работы в установленный для них срок.** Оценка зависит от качества выполнения и защиты лабораторных работ. Однако, на экзамене есть возможность получить оценку на балл выше. Порядок сдачи лабораторных работ не важен. Если хотя бы одна лабораторная защищена и загружена позже установленного срока, то студент отправляется на экзамен без возможности получить оценку автоматом. Оценка «Отлично» можно получить только на экзамене, при условии нужного количества посещений и сданных в срок всех лабораторных.*

Цель

Изучить работу алгоритмов сортировок. Научиться определять сложность алгоритмов.

Задание

Программа может быть написана на любом языке программирования.

Написать программу, для сортировки и анализа текстовых данных. В соответствии с параметрами варианта задания нужно реализовать сортировку слов в тексте и провести его простейший анализ. Варианты заданий приведены на рисунке 3.

Необходимо реализовать программу, которая будет считывать текст из файла (например, «original.txt») и с минимальными погрешностями разбивать его на массив отдельных слов. Далее, согласно параметрам выбранного варианта, необходимо отсортировать этот массив слов. Путь либо название считываемого файла должно вводиться через консоль.

Примечание. Символы и знаки препинания игнорируются и не включаются в слово/длину слова.

Отсортированный массив необходимо вывести в выходной файл, например, «result.txt». В консоль результат выводить не нужно.

Также необходимо вывести информацию с простейшим анализом введенного текста (эту информацию выводим в консоль и в файл, например, «analysis.txt»):

- исходный текст
- параметры выбранного варианта
- количество слов в исходном тексте
- время выполнения сортировки
- количество слов на каждую букву алфавита (для варианта «сортировка по алфавиту»)
- количество слов каждой длины (для варианта «сортировка по количеству символов в слове»)

Необходимо провести 10 тестов, то есть повторить такие действия с разными по длине (1 тыс., 5 тыс., 15 тыс., 30 тыс., 50 тыс., 70 тыс., 85 тыс., 100 тыс., 115 тыс. и 130 тыс. символов) текстами. Соответственно, должно получиться 30 файлов – 10 файлов с входными текстами, 10 с отсортированными массивами и 10 с анализированием текстов. В отчете нужно наглядно изобразить (график/гистограмма/диаграмма) зависимость времени выполнения сортировки от количества слов в тексте и сделать выводы. Также необходимо оценить и обосновать сложность получившегося алгоритма в О-нотации.

Предположим, что параметры выбранного варианта такие: *кириллица, по алфавиту, по возрастанию, игнорировать числа, быстрая сортировка.*

Исходный текст следующий:

За годы в метро человек не накопил сил, чтобы с триумфом подняться вверх по ступеням сияющего эскалатора, везущего его к былой мощи и великолепию. Напротив, он измельчал, привык к темноте и тесноте.

В таком случае файл «result.txt» может выглядеть примерно следующим образом (рисунок 1):

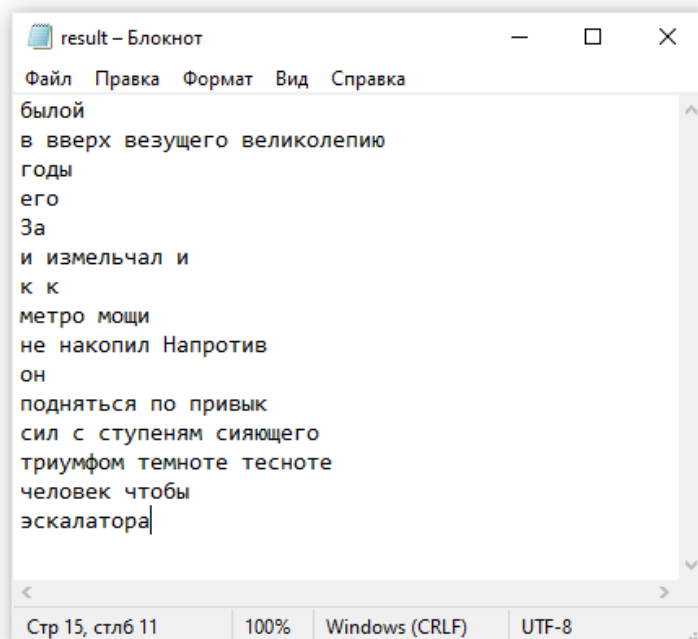


Рисунок 1 – Примерный вид файла «result.txt»

А файл «analysis.txt» может выглядеть, например, так (рисунок 2):

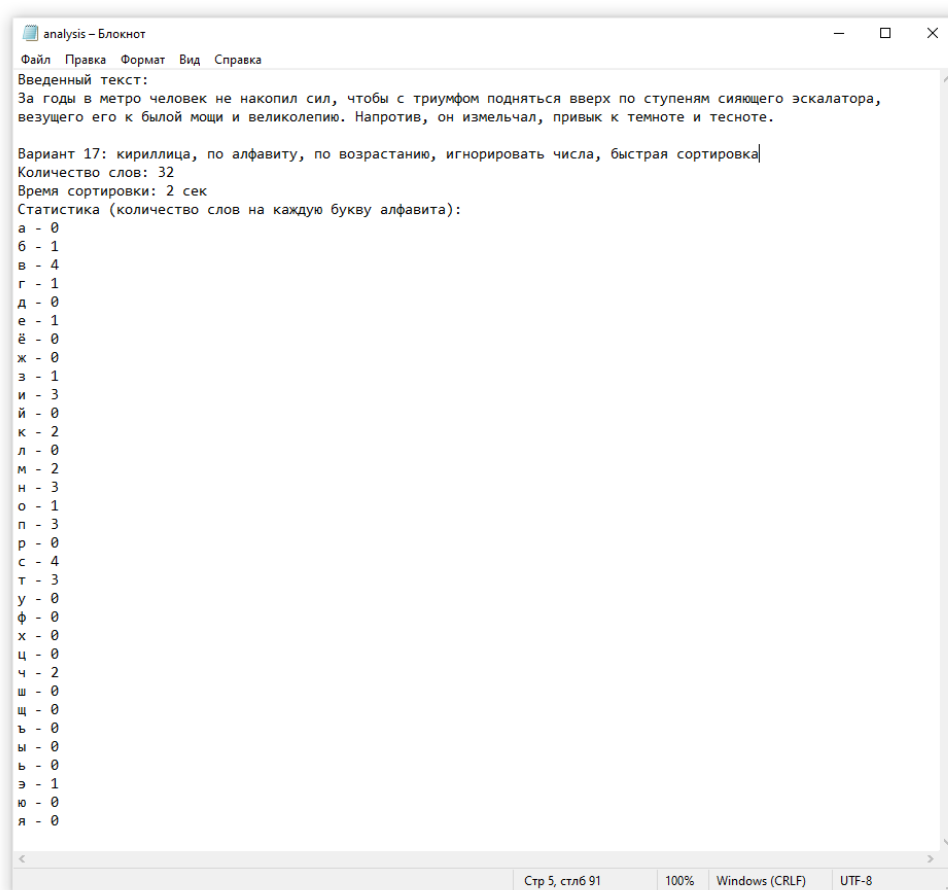


Рисунок 2 – Примерный вид файла «analysis.txt»

Варианты заданий

Далее в таблице на рисунке 3 представлены варианты заданий. Варианты старше 24-го начинаются опять с 1-го.

Вариант / Задание	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Кириллица	v		v		v			v		v		v	v		v		v			v		v		v
Латиница		v		v		v	v		v		v			v		v		v	v		v		v	
По алфавиту	v		v		v			v		v		v	v		v		v			v		v		v
По кол-ву символов в слове		v		v		v	v		v		v			v		v		v	v		v		v	
По возрастанию	v		v		v			v		v		v	v		v		v			v		v		v
По убыванию		v		v		v	v		v		v			v		v		v	v		v		v	
Учитывать числа	v		v		v		v		v		v			v		v		v		v		v		v
Игнорировать числа		v		v		v		v		v		v	v		v		v		v		v		v	
Сортировка Пузырьком	v						v						v						v					
Сортировка Расческой		v						v						v						v				
Сортировка Вставками			v						v						v						v			
Сортировка Шелла				v						v						v						v		
Быстрая сортировка					v						v						v						v	
Сортировка Слиянием						v						v						v						v

Рисунок 3 – Варианты заданий

Пункт «Кириллица / Латиница» означает, с символами какого алфавита придется работать. Кириллица – «абвгдеёжзийклмнопрстуфхцчшщъыьэюя1234567890». Латиница – «abcdefghijklmnopqrstuvwxyz1234567890».

Пункт «По алфавиту / По количеству символов в слове» означает, по какому признаку необходимо сортировать слова в тексте. По алфавиту – проверяем первую букву в каждом слове и сортируем относительно алфавита (цифры идут после букв). По количеству символов в слове – вычисляем количество букв в каждом слове и сортируем относительно этого количества.

Пункт «По возрастанию / По убыванию» означает, в каком направлении сортировать массив слов.

Пункт «Учитывать числа / Игнорировать числа» означает, необходимо ли идентифицировать отдельные числа/цифры в тексте как отдельные слова. Если числа учитываются, то они считаются словом и к ним применяются правила из пункта «По алфавиту / По количеству символов в слове». Если игнорируются, то эти числа/цифры просто опускаются при сортировке.

Пункт «Сортировка ***» означает, какой метод сортировки необходимо использовать.

Требования к структуре проекта

- ✓ Весь программный код должен быть разбит на отдельные функции/методы, писать весь код внутри функции «void main()» не допускается
- ✓ Имена функций и переменных должны отражать их назначение
- ✓ Основные части кода должны иметь комментарии

Отчет должен содержать

1. Цель работы
2. Задание
3. Краткое описание хода разработки
4. Блок-схема (при необходимости)
5. Исходный код программы (с комментариями)

6. Результаты работы программы с примерами

7. Выводы

При оценке выполнения работы будут учитываться грамотность оформления исходного кода, работоспособность программы и соответствие отчета правилам оформления текстовых документов по ГОСТ 7.32-2017. Титульные листы лабораторных работ представлены на сайте ГУАП (<https://guap.ru/standart/doc>).