МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение

высшего образования

**«Челябинский государственный университет»**

**(ФГБОУ ВО «ЧелГУ»)**

Институт информационных технологий

Кафедра информационных технологий и экономической информатики

ОТЧЕТ

по лабораторной работе №2

Выполнил:

Студент группы ПрИ-202 Приходько Даниил Александрович

Студент группы ПрИ-202 Саламатин Алексей Юрьевич

Студент группы ПрИ-202 Скоробогатов Максим Дмитриевич

Принял:

Преподаватель ИИТ Николаев Иван Евгеньевич

Отчет защищен \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

дата оценка

Челябинск 2024 г.

Содержание

Введение..................................................................................................................3

Задание 1………......................................................................................................4

Задание 2................................................................................................................17

Задание 3................................................................................................................29

Введение

Цель работы: понять и доступно объяснить Шестикласснику, как работают различные сортировки данных. Написать приложение с визуальной частью, благодаря которой будет видно весь процесс сортировок.

Задание 1: Реализовать демонстрацию работы 4 алгоритмов внутренней сортировки. А конкретно, нами были выбраны: Select Sort, Shell Sort, Quick Sort, Merge Sort.

Задание 2: Дана таблица, состоящая из нескольких полей. Таблица находится в файле. Реализовать программу, которая одним из методов (прямым, естественным или многопутевым слиянием) сортирует записи и записывает их в результирующий файл. Метод внешней сортировки и ключевой атрибут, по которому осуществляется сортировка выбирается пользователем. Количество атрибутов в таблица может быть произвольным.

Задание 3: Пусть имеется текст, состоящий из слов. Необходимо разбить текст на отдельные слова и провести их сортировку в лексикографическом порядке. Причем одинаковые слова будут в отсортированной последовательности идти друг за другом. Надо выбрать два алгоритма сортировки, один из которых базовый или усовершенствованный, и другой алгоритм ABC- сортировка или redix сортировка. После получения отсортированного массива необходимо пройтись по нему и для каждого слова подсчитать, сколько раз оно встречается. Результаты подсчетов вывести на экран. Далее необходимо провести эксперименты с использованием обоих алгоритмов по сортировке текстов различной длины (100, 500, 1000, 2000, 5000 слов или более). Причем для каждого эксперимента произвести замеры времени сортировки. Получившиеся

данные оформить в таблицу.

Задание 1

В рамках этой части были реализованы алгоритмы внутренней сортировки Select Sort, Shell Sort, Quick Sort, Merge Sort.

Сортировка выбором – это алгоритм, суть которого заключается в постоянном сравнении элементов неотсортированой части. Наименьший элемент в неотсортированой части меняется местами с первым ее элементом и сразу становится отсортированой частью.

Чтобы доступно объяснить процесс сортировки, код был подвержен постоянному логированию. Так же внутри него есть множество вставок кода который отвечает за визуализацию сортировки.

Код сортировки представлен на рисунке 1.1.1 – 1.1.3.

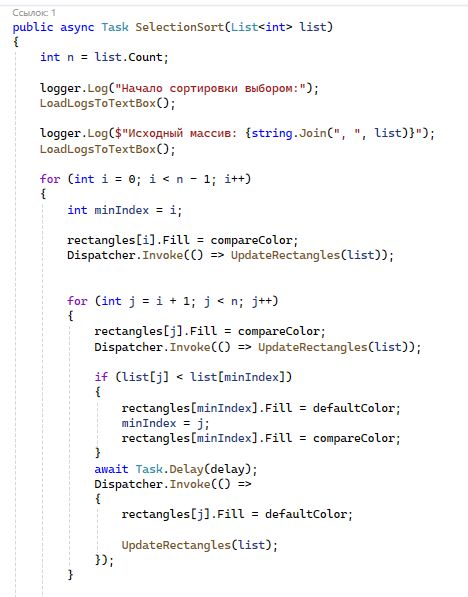


Рис.1.1.1 Код сортировки выбором

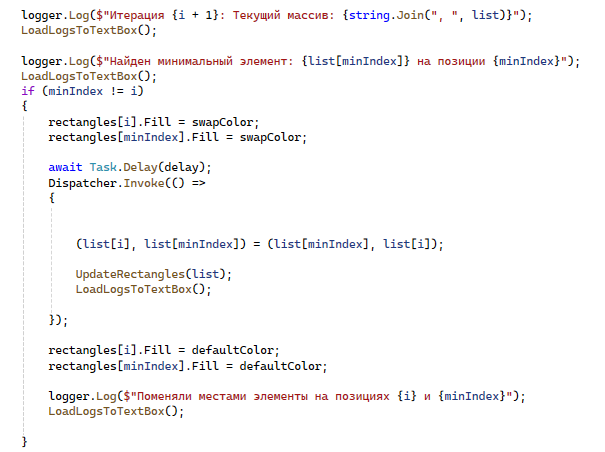


Рис.1.1.2 Код сортировки выбором

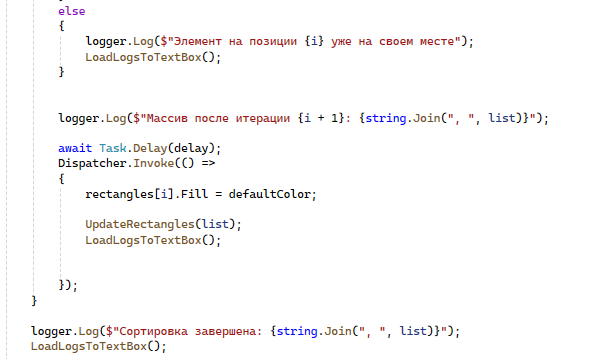


Рис.1.1.3 Код сортировки выбором

Сортировка Шелла – это модифицированая сортировка вставками. Ее суть в том, что массив данных разбивается на так называемые Gap’ы (зазоры). Каждая итерация сортировки – это сравнение элементов массива на расстоянии текущего зазора друг от друга, а затем упорядочивание этих элементов. С каждым проходом по массиву зазор уменьшается вдвое и эти действия повторяются до окончания сортировки.

Чтобы доступно объяснить процесс сортировки, код был подвержен постоянному логированию. Так же внутри него есть множество вставок кода который отвечает за визуализацию сортировки.

Код сортировки представлен на рисунке 1.2.1 – 1.2.2.

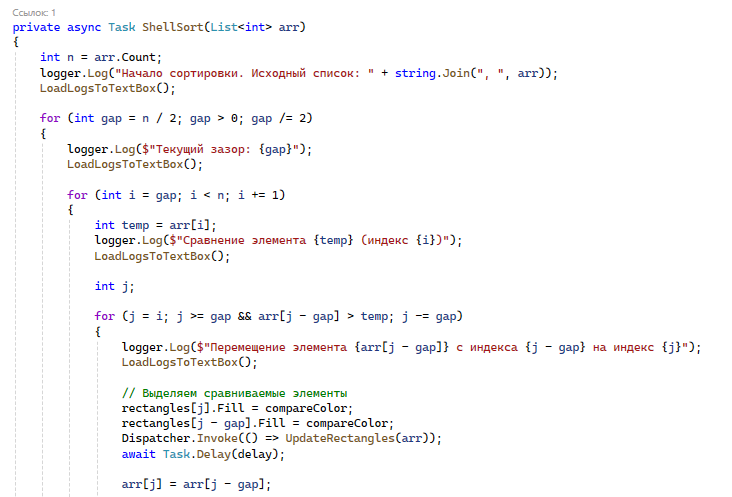


Рис.1.2.1 Код сортировки Шелла

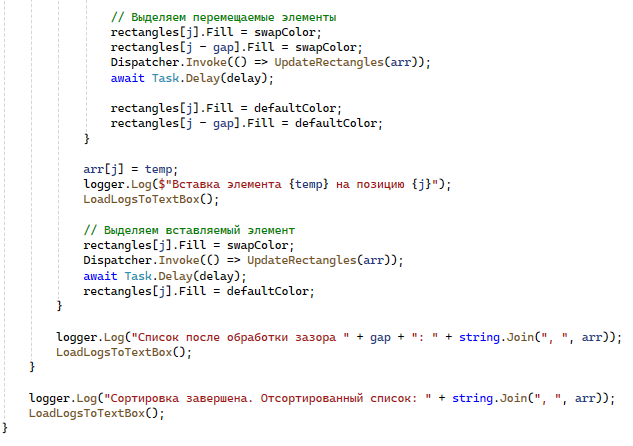


Рис.1.2.2 Код сортировки Шелла

Быстрая сортировка - это алгоритм "разделяй и властвуй", который рекурсивно сортирует массив, разделяя его на части относительно опорного элемента и объединяя отсортированные части.

Чтобы доступно объяснить процесс сортировки, код был подвержен постоянному логированию. Так же внутри него есть множество вставок кода который отвечает за визуализацию сортировки.

Код сортировки представлен на рисунке 1.3.1 – 1.3.3.

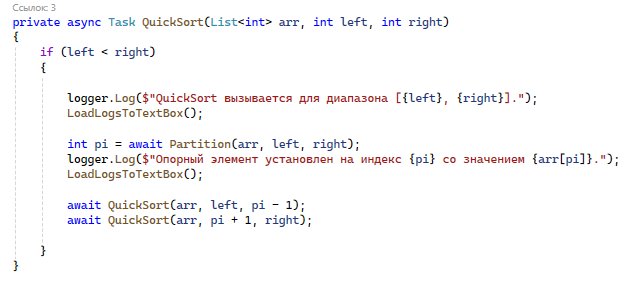


Рис.1.3.1 Код быстрой сортировки

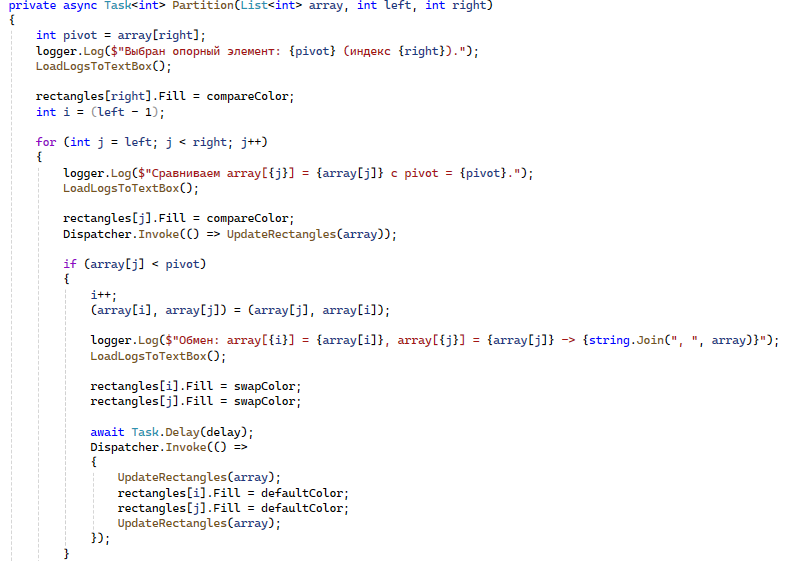


Рис.1.3.2 Код быстрой сортировки

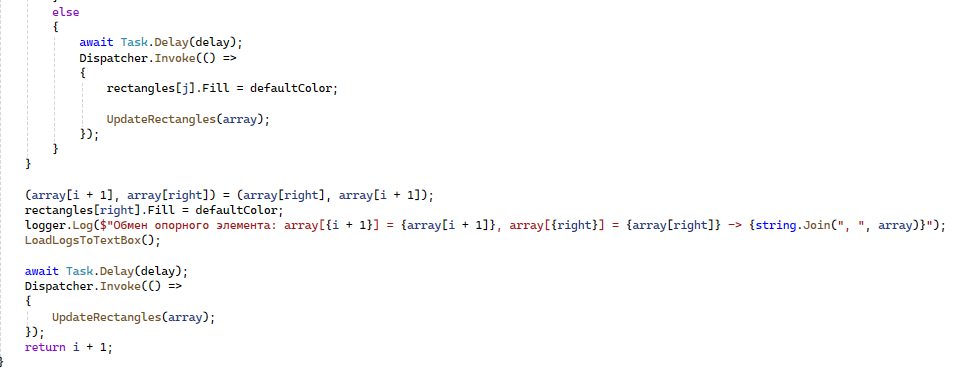


Рис.1.3.3 Код быстрой сортировки

Сортировка слиянием – это алгоритм который разделяет массив данных на две равные части, потом вызывает сам себя для каждой из этих частей и так до тех пор, пока массивы не станут единичными. Дальше эти подмассивы сравниваются со своей «половинкой» и в упорядоченом виде объединяются в новый отсортированый массив.

Данный алгоритм является внутренним, несмотря на то что есть его аналоги в виде внешних сортировок.

Чтобы доступно объяснить процесс сортировки, код был подвержен постоянному логированию. Так же внутри него есть множество вставок кода который отвечает за визуализацию сортировки.

Код сортировки представлен на рисунке 1.4.1 – 1.4.4.

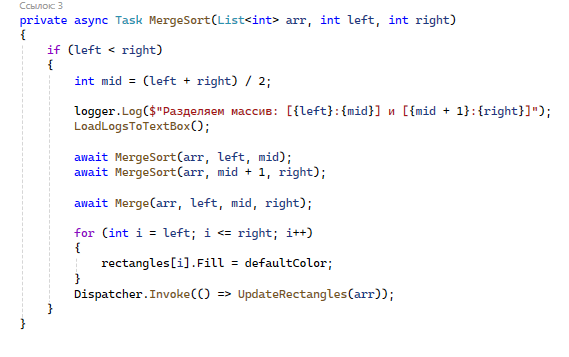


Рис.1.4.1 Код сортировки слиянием

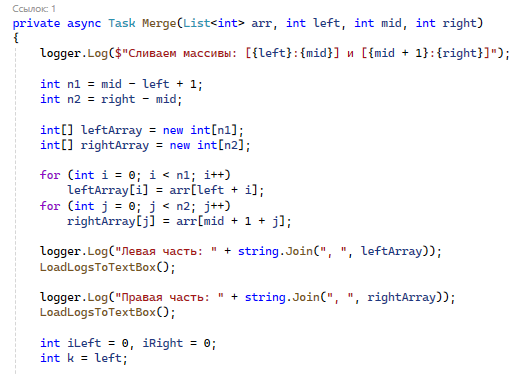


Рис.1.4.2 Код сортировки слиянием

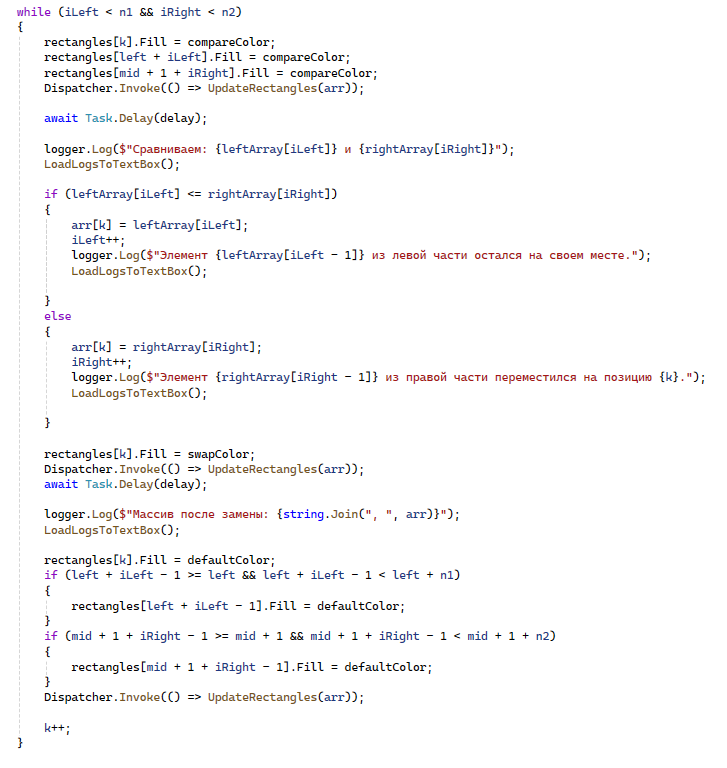


Рис.1.4.3 Код сортировки слиянием

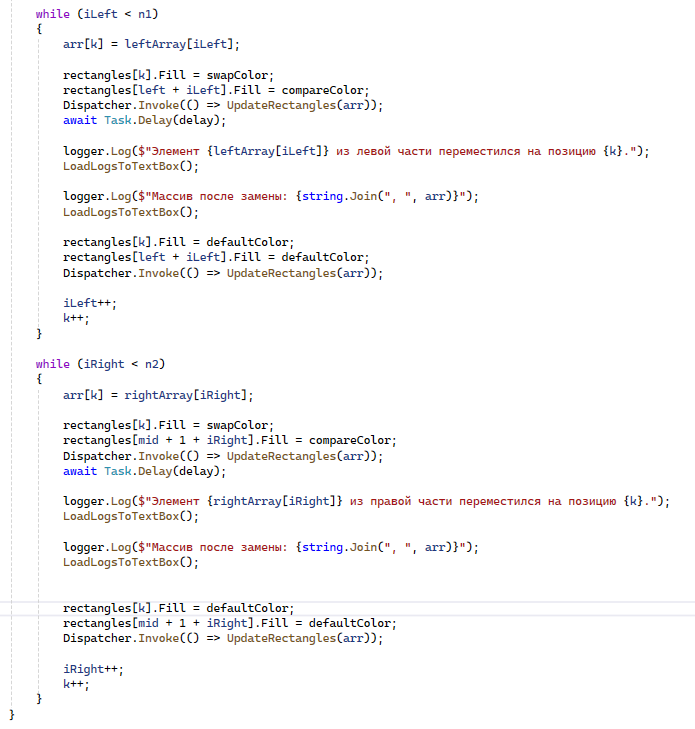


Рис.1.4.4 Код сортировки слиянием

Чтобы наглядно посмотреть на процесс сортировки был написан GUI для данной части (рис.1.5), где видно как проходит сортировка элементов, которые представлены в виде столбцов с соответствующей величиной. Так же есть окно, куда с определенной задержкой, которую выбирает пользователь, выводятся сопутствующие объяснения. Сам массив пользователь может ввести в верхнем окне ввода, через запятую или пробел. Выбрать желаемую сортировку можно в панели управления, в правой части окна.

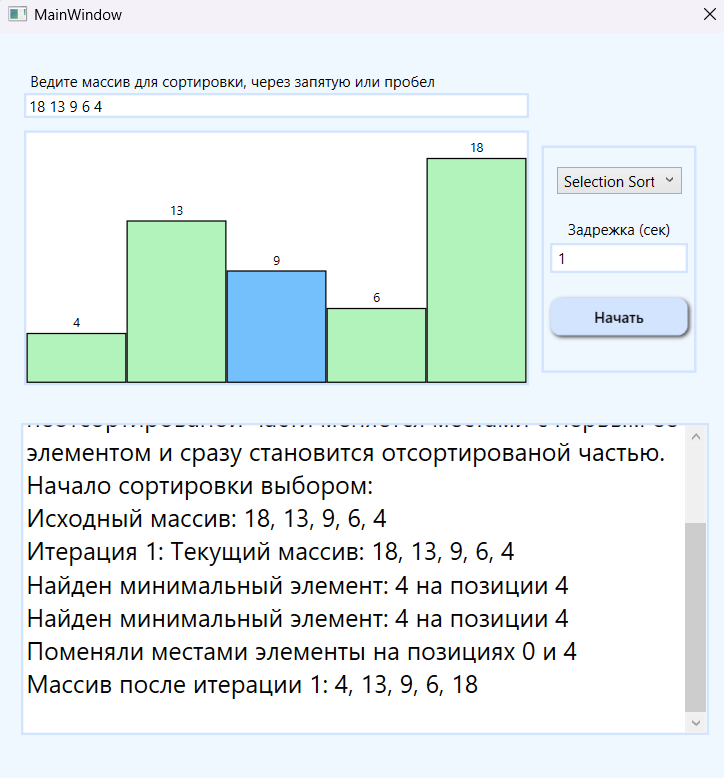


Рис.1.5 визуализация первого задания

Задание 2

Эта часть работы заключается в том, чтобы показать наглядно, и объяснить, как работают внешние сортировки, такие как прямое, естественное и многопутевое слияния.

Прямая сортировка слиянием:

Принцип прямой сортировки, реализованный в данном коде, заключается в следующем: данные разбиваются на несколько отсортированных частей (чанков), которые затем сливаются в один отсортированный список. Процесс состоит из двух основных этапов: сортировка внутри каждого чанка и слияние отсортированных чанков.

На первом этапе каждый чанк сортируется отдельно с использованием алгоритма сортировки вставками. Этот алгоритм работает следующим образом: для каждого элемента в чанке, начиная со второго, элемент сравнивается с предыдущими элементами и вставляется на правильное место, чтобы поддерживать порядок. Этот процесс повторяется для всех элементов в чанке, что приводит к его полной сортировке.

На втором этапе отсортированные чанки сливаются в один отсортированный список с использованием очереди приоритетов. Очередь приоритетов хранит минимальные элементы из каждого чанка и позволяет эффективно извлекать минимальный элемент из всех чанков. Процесс слияния начинается с добавления первых элементов каждого чанка в очередь приоритетов. Затем, пока очередь не пуста, извлекается минимальный элемент, который добавляется в итоговый список. Если в чанке, из которого был извлечен элемент, есть следующий элемент, он добавляется в очередь приоритетов. Этот процесс продолжается до тех пор, пока все элементы из всех чанков не будут добавлены в итоговый список.

Таким образом, прямая сортировка в данном коде сочетает в себе сортировку вставками для отдельных чанков и слияние с использованием очереди приоритетов для объединения этих чанков в один отсортированный список. Этот подход позволяет эффективно обрабатывать большие объемы данных, разбивая их на управляемые части и сливая их по мере необходимости.

Чтобы доступно объяснить процесс сортировки, код был подвержен постоянному логированию. Так же внутри него есть множество вставок кода который отвечает за визуализацию сортировки.

Код сортировки представлен на рисунке 2.1.1 – 2.1.3.

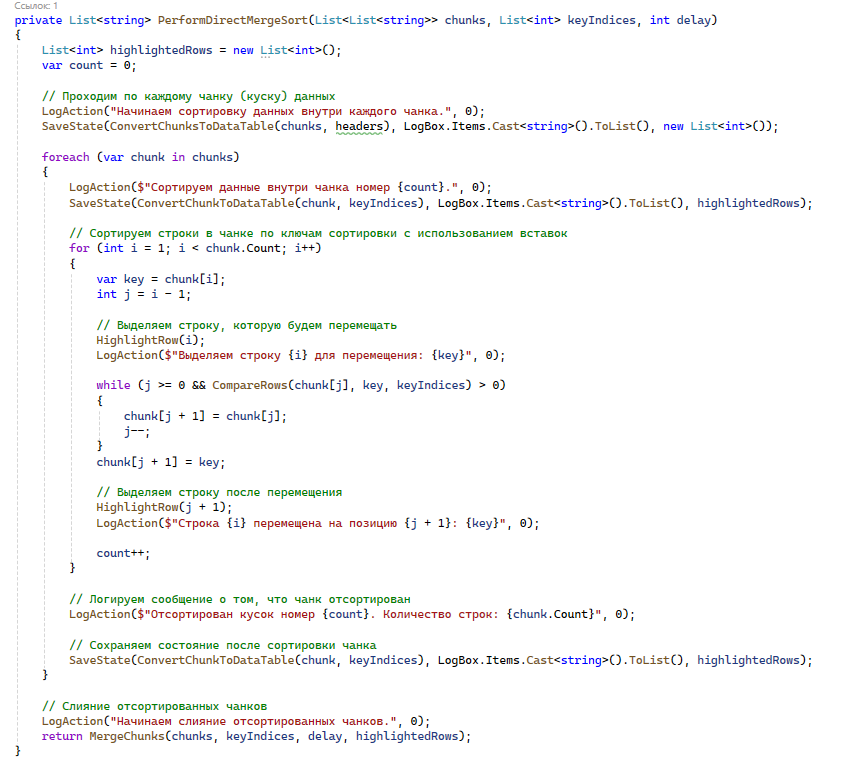


Рис.2.1.1 Код сортировки прямым слиянием



Рис.2.1.2 Код сортировки прямым слиянием



Рис.2.1.3 Код сортировки прямым слиянием

Естественное слияние:  
  
Принцип естественной сортировки слиянием, реализованный в данном коде, заключается в разбиении данных на естественные подпоследовательности (или "ранны"), которые затем сливаются в один отсортированный список. Этот метод сортировки называется естественной сортировкой слиянием, так как он использует естественные подпоследовательности, которые уже частично отсортированы, что позволяет уменьшить количество операций сравнения и перемещения.

Процесс начинается с разбиения исходных данных на естественные подпоследовательности. Это делается путем последовательного сравнения элементов данных. Если текущий элемент меньше или равен предыдущему, он добавляется в текущую подпоследовательность. Если текущий элемент больше предыдущего, текущая подпоследовательность завершается, и начинается новая подпоследовательность с текущего элемента. Этот процесс продолжается до тех пор, пока все данные не будут разбиты на подпоследовательности.

После разбиения данных на естественные подпоследовательности, начинается процесс их слияния. Слияние осуществляется с использованием очереди приоритетов, которая хранит минимальные элементы из каждой подпоследовательности. Процесс слияния начинается с добавления первых элементов каждой подпоследовательности в очередь приоритетов. Затем, пока очередь не пуста, извлекается минимальный элемент, который добавляется в итоговый список. Если в подпоследовательности, из которой был извлечен элемент, есть следующий элемент, он добавляется в очередь приоритетов. Этот процесс продолжается до тех пор, пока все элементы из всех подпоследовательностей не будут добавлены в итоговый список.

Таким образом, естественная сортировка слиянием сочетает в себе разбиение данных на естественные подпоследовательности и их последующее слияние с использованием очереди приоритетов. Этот подход позволяет эффективно обрабатывать данные, используя их естественные свойства, что уменьшает количество операций и улучшает производительность сортировки.

Чтобы доступно объяснить процесс сортировки, код был подвержен постоянному логированию. Так же внутри него есть множество вставок кода который отвечает за визуализацию сортировки.

Код сортировки представлен на рисунке 2.2.1 – 2.2.4.

  
Рис.2.2.1 Код сортировки естественным слиянием

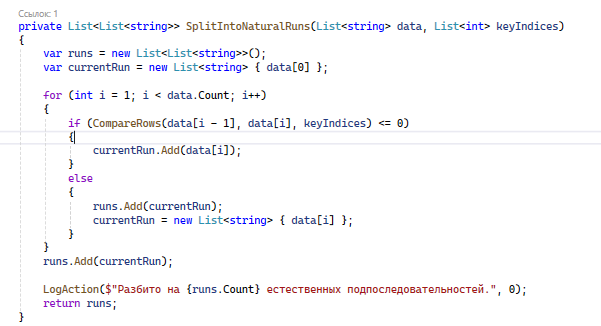


Рис.2.2.2 Код сортировки естественным слиянием



Рис.2.2.3 Код сортировки естественным слиянием



Рис.2.2.4 Код сортировки естественным слиянием

Многопутевое слияние:

Принцип многопутевой сортировки слиянием, реализованный в данном коде, заключается в параллельной сортировке и слиянии данных, разбитых на несколько частей (чанков). Этот метод позволяет эффективно обрабатывать большие объемы данных, используя параллелизм для ускорения процесса сортировки.

Процесс начинается с параллельной сортировки каждого чанка. Для этого используется алгоритм сортировки вставками, который применяется к каждому чанку независимо. Сортировка вставками работает следующим образом: для каждого элемента в чанке, начиная со второго, элемент сравнивается с предыдущими элементами и вставляется на правильное место, чтобы поддерживать порядок. Этот процесс повторяется для всех элементов в чанке, что приводит к его полной сортировке. Параллельная обработка позволяет сортировать все чанки одновременно, что значительно ускоряет процесс.

После завершения сортировки всех чанков начинается процесс их параллельного слияния. Для этого используется очередь приоритетов, которая хранит минимальные элементы из каждого чанка и позволяет эффективно извлекать минимальный элемент из всех чанков. Процесс слияния начинается с добавления первых элементов каждого чанка в очередь приоритетов. Затем, пока очередь не пуста, извлекается минимальный элемент, который добавляется в итоговый список. Если в чанке, из которого был извлечен элемент, есть следующий элемент, он добавляется в очередь приоритетов. Этот процесс продолжается до тех пор, пока все элементы из всех чанков не будут добавлены в итоговый список.

Параллельная обработка очереди приоритетов позволяет одновременно извлекать и добавлять элементы из разных чанков, что ускоряет процесс слияния. В результате, многопутевая сортировка слиянием эффективно объединяет отсортированные чанки в один отсортированный список, используя параллелизм для ускорения как сортировки, так и слияния данных.

Чтобы доступно объяснить процесс сортировки, код был подвержен постоянному логированию. Так же внутри него есть множество вставок кода который отвечает за визуализацию сортировки.

Код сортировки представлен на рисунке 2.3.1 – 2.3.3.



Рис.2.3.1 Код сортировки многопутевым слиянием



Рис.2.3.2 Код сортировки многопутевым слиянием

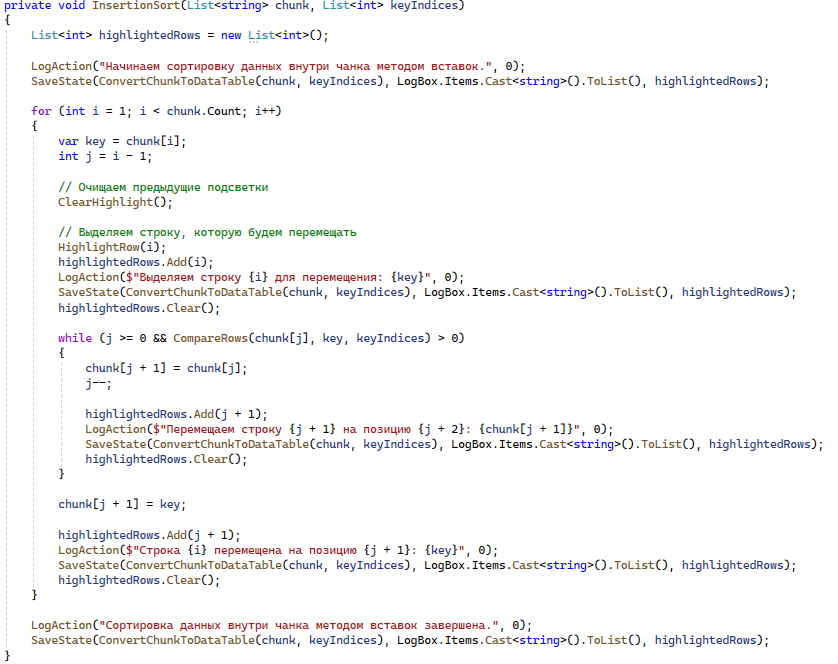


Рис.2.3.3 Код сортировки вставками в сортировке многопутевым слиянием

Чтобы наглядно посмотреть на процесс сортировки был написан GUI для данной части (рис.2.4), где видно как проходит сортировка элементов таблицы. Пользователю представлен выбор сортировки, ключа. Так же он может пошагово отслеживать изменения во время процесса, благодаря кнопкам «Вперед» и «Назад», или включить автопроигрывание.

В отдельных окнах выводится визуализация сортировки в виде состояния таблицы в текущий момент времени, текстовая документация процесса и финальный результат.

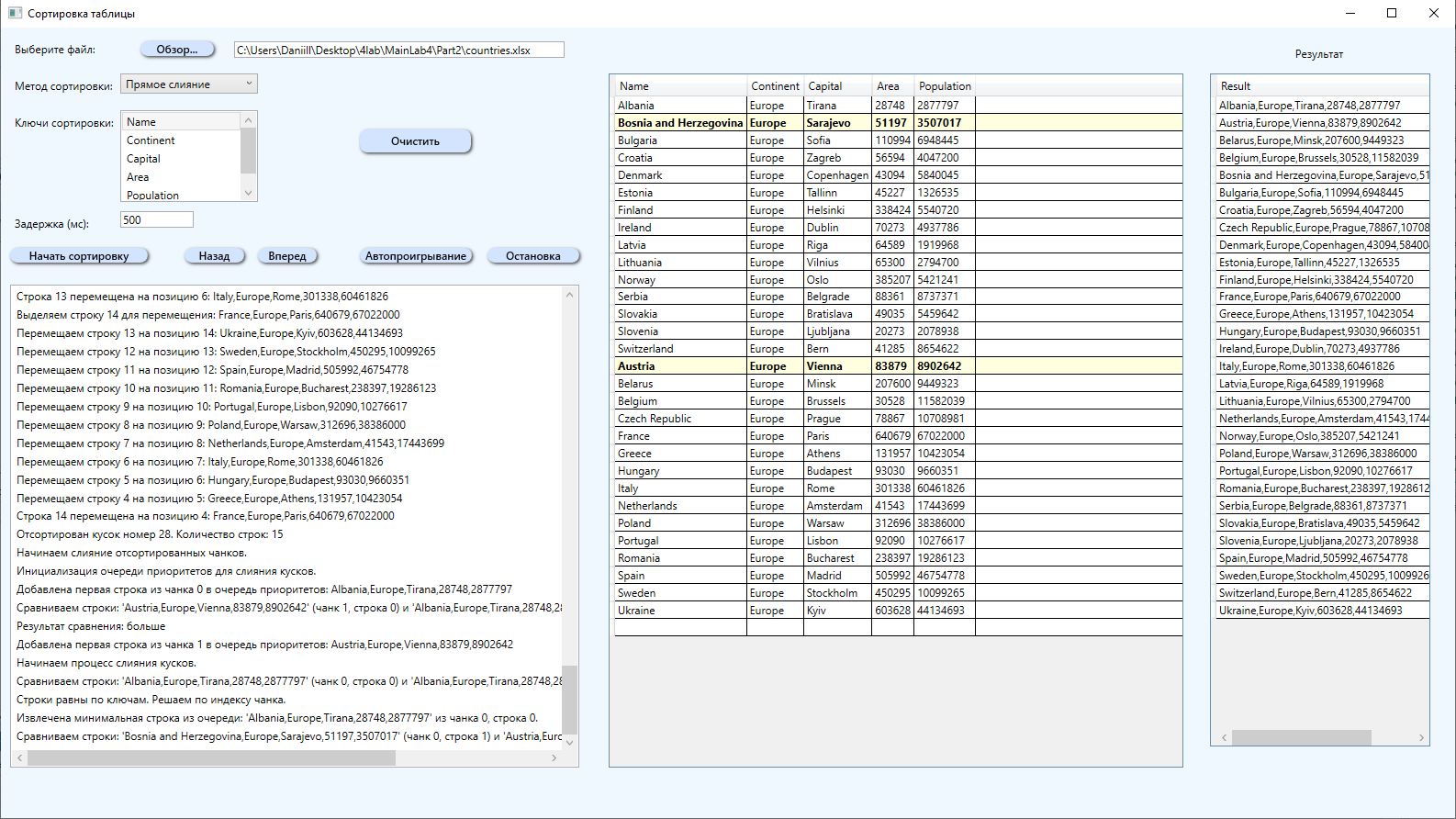


Рис.2.4 Часть 2 GUI

Задание 3

В рамках этой части были реализованы алгоритмы сортировки Modificated Buble Sort и Radix Sort.

Modificated Bubble Sort — это оптимизированная версия стандартного алгоритма сортировки пузырьком, которая позволяет завершить сортировку раньше, если массив уже отсортирован. Этот алгоритм прост в реализации и может быть эффективен для небольших массивов или частично отсортированных данных.

Чтобы доступно объяснить процесс сортировки, код был подвержен постоянному логированию. Так же внутри него есть множество вставок кода который отвечает за визуализацию сортировки.

Код сортировки представлен на рисунке 3.1.1 – 3.1.2.



Рис.3.1.1 Код сортировки пузырьком.

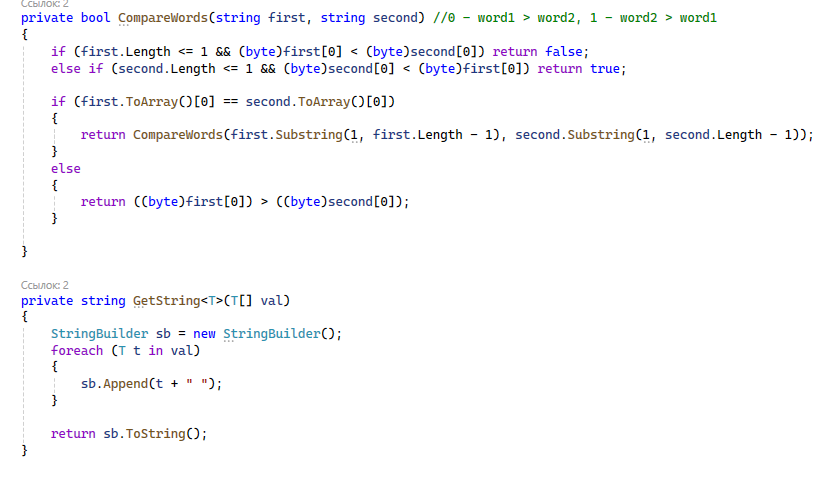


Рис.3.1.2 Код сортировки пузырьком.

Radix Sort - выполняет поразрядную сортировку, начиная с последнего символа и двигаясь к первому. Все строки выравниваются до одинаковой длины, заполняясь символом '0' (или пробелом) для корректного сравнения. На каждом этапе используется стабильная сортировка, например, CountingSort, для сортировки символов текущей позиции.

Чтобы доступно объяснить процесс сортировки, код был подвержен постоянному логированию. Так же внутри него есть множество вставок кода который отвечает за визуализацию сортировки.

Код сортировки представлен на рисунке 3.2.1 – 3.2.2.

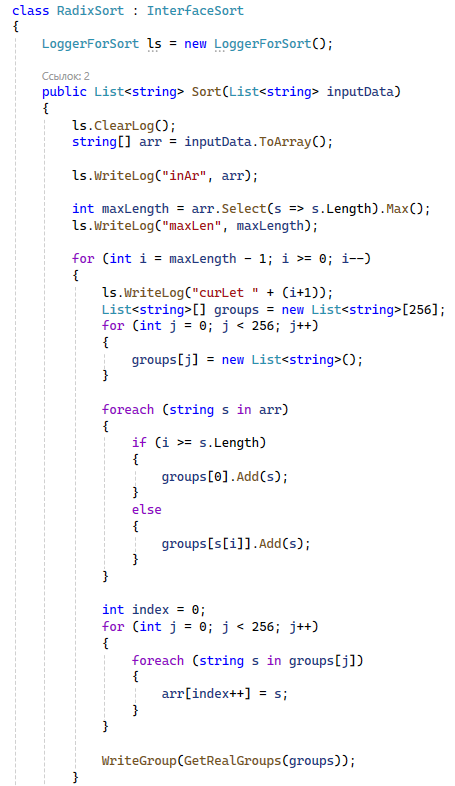


Рис.3.2.1 Код Radix Sort

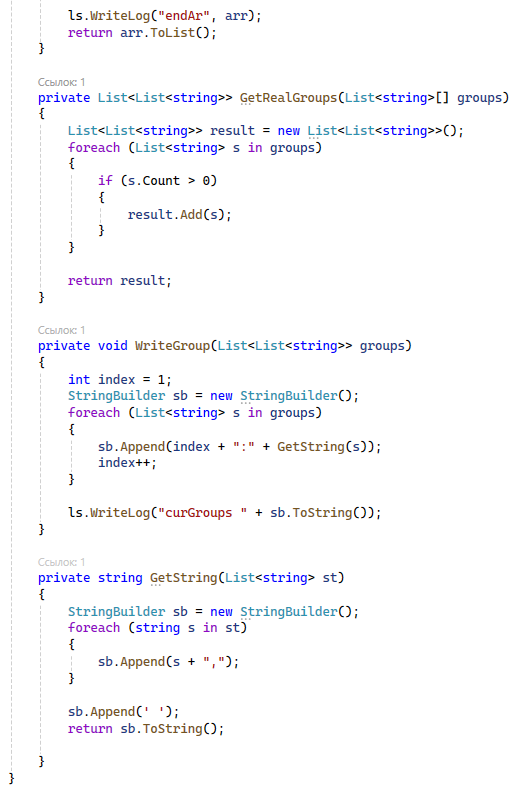


Рис.3.2.2 Код Radix Sort

Чтобы наглядно показать работу этих сортировок было создано два окна под Buble Sort (рис.3.3.1) и Radix Sort (рис.3.3.2), соответсвтенно. В каждом окне есть возможность выгрузить файл с массивом данных с компьютера. Так же пользователь может отслеживать состояния массива прямо во время сортировки, переключая шаги

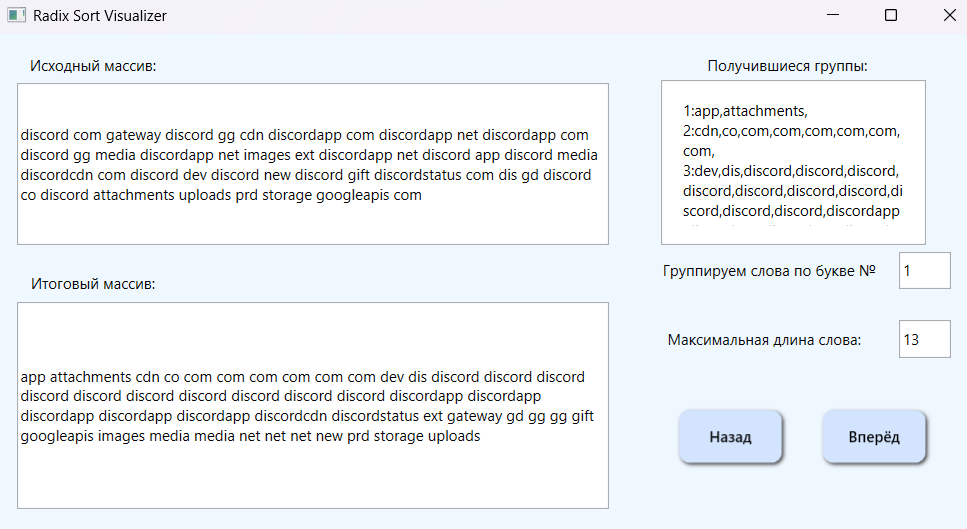


Рис.3.3.2 Визуализация Radix Sort