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 №4

Авторы отчета С.А. Костюк ПрИ-202

подпись инициалы, фамилия группа

Д. Дударов ПрИ-202

подпись инициалы, фамилия группа

Е. Хотенов ПрИ-202

подпись инициалы, фамилия группа

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

дата оценка

Челябинск 2023 г.

**Цель работы:** понять и реализовать алгоритмы внутренней и внешней сортировки, а также реализовать демонстрацию алгоритмов внутренней сортировки, и научиться их применять на практике.

**Задание 1.**

Реализовать демонстрацию работы 2 алгоритмов внутренней сортировки.

По два из каждой группы. (могут повторяться)

* **Квадратичные и субквадратичные алгоритмы**
  + Сортировка выбором (SelectSort)
  + Сортировка пузырьком (BubbleSort) и ее улучшения
  + Сортировка простыми вставками (InsertSort)
  + Cортировка методом Шелла (ShellSort)
  + …
* **Усовершенствованные алгоритмы**
  + Бинарная пирамидальная сортировка (HeapSort)
  + Быстрая сортировка Хоара (QuickSort)
  + …

Программа должна выводить на экран последовательные шаги сортировки с небольшой задержкой, которую устанавливает пользователь. Пояснения для 6-классника. Графический интерфейс.

Программа должна вести лог всех совершенных сравнений и перестановок и выводить их на экран.

**Задание 2. Алгоритмы внешней сортировки**

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.

Примеры таблиц и заданий

1. Дан полный перечень всех стран, который включает в себя: название, континент, столицу, площадь, численность населения. Указать сведения о государствах заданного континента в порядке возрастания численности населения.
2. Даны сведения о химических веществах, которые включает в себя: класс вещества, название вещества, молекулярная масса вещества. Упорядочить по возрастанию молекулярных масс все вещества указанного класса.
3. В файле хранится последовательность русских слов. Упорядочить ее в алфавитном порядке. Использовать внешнюю сортировку.

Программа должна выводить на экран последовательные шаги сортировки с небольшой задержкой, которую устанавливает пользователь. Пояснения для 6-классника. Графический интерфейс.

Программа должна вести лог всех совершенных сравнений и перестановок и выводить их на экран.

**Внешняя сортировка прямым слиянием.**

Алгоритм сортировки простым слияния является простейшим алгоритмом внешней сортировки, основанный на процедуре слияния серией.

В данном алгоритме длина серий фиксируется на каждом шаге. В исходном файле все серии имеют длину 1, после первого шага она равна 2, после второго – 4, после третьего – 8, после k -го шага – 2k.

*Алгоритм сортировки простым слиянием*

Шаг 1. Исходный файл f разбивается на два вспомогательных файла f1 и f2.

Шаг 2. Вспомогательные файлы f1 и f2 сливаются в файл f, при этом одиночные элементы образуют упорядоченные пары.

Шаг 3. Полученный файл f вновь обрабатывается, как указано в шагах 1 и 2. При этом упорядоченные пары переходят в упорядоченные четверки.

Шаг 4. Повторяя шаги, сливаем четверки в восьмерки и т.д., каждый раз удваивая длину слитых последовательностей до тех пор, пока не будет упорядочен целиком весь файл ( рис. 43.1).

После выполнения i проходов получаем два файла, состоящих из серий длины 2i. Окончание процесса происходит при выполнении условия 2i>=n. Следовательно, процесс сортировки простым слиянием требует порядка O(log n) проходов по данным.

Признаками конца сортировки простым слиянием являются следующие условия:

длина серии не меньше количества элементов в файле (определяется после фазы слияния);

количество серий равно 1 (определяется на фазе слияния).

при однофазной сортировке второй по счету вспомогательный файл после распределения серий остался пустым. (Реализация алгоритма представлена на рисунках 2.1 и 2.2)

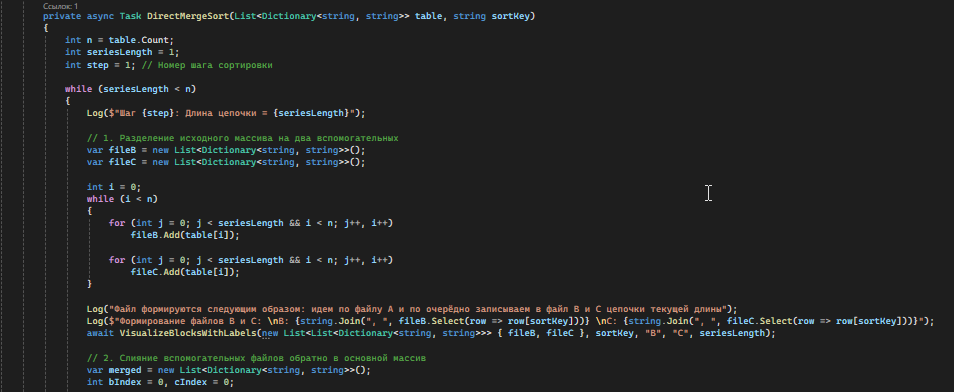


Рис.2.1(Код сортировки прямым слиянием)

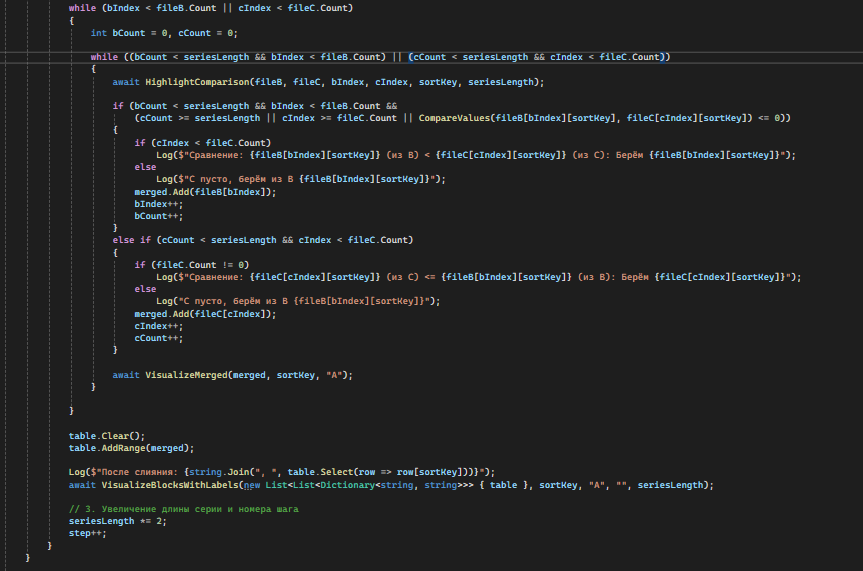


Рис.2.2(Код сортировки прямым слиянием, продолжение)

Пример работы программы представлен на рисунках 2.3 и 2.4

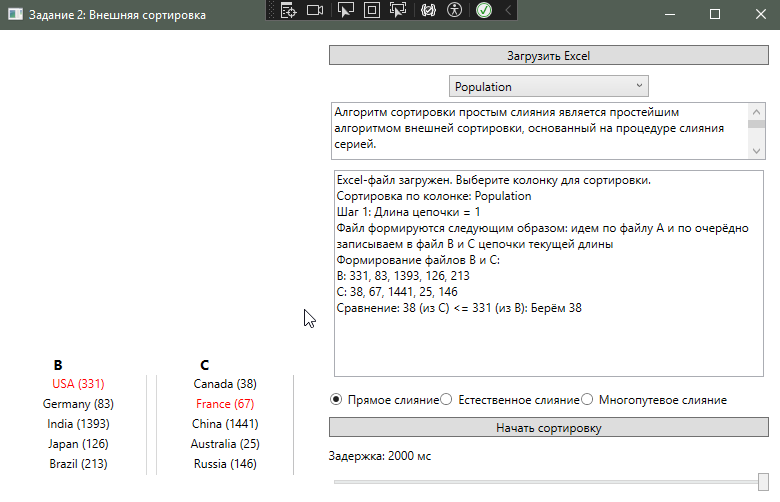


Рис.2.3(Пример прямой сортировки)

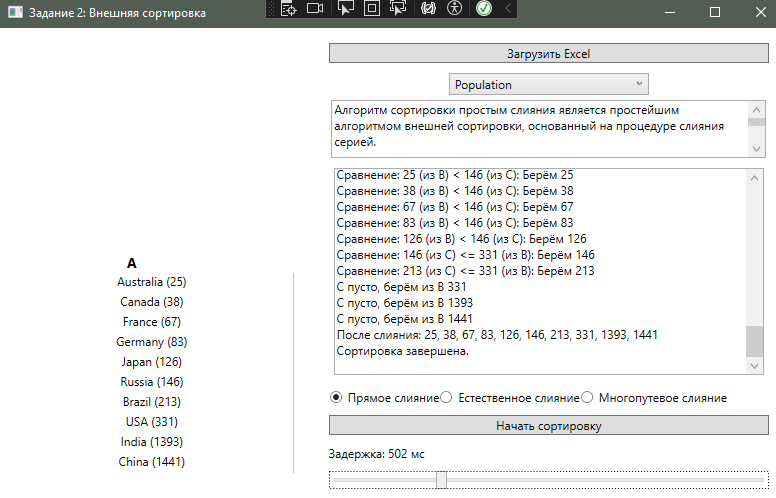


Рис.2.4(Пример прямой сортировки)

**Внешняя сортировка естественным слиянием.**

Сортировка, при которой всегда сливаются две самые длинные из возможных последовательностей, является естественным слиянием. В данной сортировке объединяются серии максимальной длины.

*Алгоритм сортировки естественным слиянием*

Шаг 1. Исходный файл f разбивается на два вспомогательных файла f1 и f2. Распределение происходит следующим образом: поочередно считываются записи ai исходной последовательности (неупорядоченной) таким образом, что если значения ключей соседних записей удовлетворяют условию f(ai)<=f(ai+1), то они записываются в первый вспомогательный файл f1. Как только встречаются f(ai)>f(ai+1), то записи ai+1 копируются во второй вспомогательный файл f2. Процедура повторяется до тех пор, пока все записи исходной последовательности не будут распределены по файлам.

Шаг 2. Вспомогательные файлы f1 и f2 сливаются в файл f, при этом серии образуют упорядоченные последовательности.

Шаг 3. Полученный файл f вновь обрабатывается, как указано в шагах 1 и 2.

Шаг 4. Повторяя шаги, сливаем упорядоченные серии до тех пор, пока не будет упорядочен целиком весь файл.

Признаками конца сортировки естественным слиянием являются следующие условия:

количество серий равно 1 (определяется на фазе слияния).

при однофазной сортировке второй по счету вспомогательный файл после распределения серий остался пустым.

Естественное слияние, у которого после фазы распределения количество серий во вспомогательных файлах отличается друг от друга не более чем на единицу, называется сбалансированным слиянием, в противном случае – несбалансированное слияние. Код сортировки естественным слиянием представлен на рисункк 2.5

**Внешняя сортировка многопутевым слиянием слиянием.**

Процесс многопутевого слияния почти как две капли воды схож с процессом прямого слияния. За одним лишь тем исключением, что мы будем использовать больше двух подфайлов.

**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.

P.S. Рекомендуется использовать тексты на английском языке, чтобы уменьшить количество различных словоформ.

\* Отчет оформляется в электронном (не печатном) виде.

\* К отчету приложить грамотно оформленный код с комментариями.