



Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ Робототехники и комплексной автоматизации

КАФЕДРА Системы автоматизированного проектирования (РК-6)

ОТЧЕТ О ВЫПОЛНЕНИИ ЛАБОРАТОРНОЙ РАБОТЫ

Студент Гусаров Аркадий Андреевич
Группа РК6-53Б
Тип задания Лабораторная работа №1
Тема лабораторной работы Перекодировка текстовых файлов

Студент _____ Гусаров А.А.
подпись, дата фамилия, и.о.

Оценка _____

Москва, 2021 г.

Задание на лабораторную работу

Разработать программу перекодирования текстовых файлов из кодировки KOI-8 в CP1251. Форсированная перезапись. При вызове программы должен быть предусмотрен специальный ключ, который форсирует запись результирующего файла для перекодировки, если он уже существует. Если ключ не указан, то перезапись не происходит, а выводится только диагностическое сообщение.

Теоретическая часть. Описание алгоритма

При запуске программа проверяет входные данные на ошибки, составляется таблица перекодировки из KOI-8 в CP1251, считывается посимвольно и поблочно поток данных в текстовом файле и параллельно перекодируется, после чего формируется выходной файл.

Код программы

```
#include <unistd.h>
// #include <errno.h>
#include <string.h>
#include <stdio.h>
#include <stdlib.h>
// Библиотека для работы с файлами
#include <fcntl.h>

int diag(char *, char *, int);
int inputErrors(int);
unsigned char *recode(int);
static unsigned char buf[BUFSIZ];

int main(int argc, char **argv)
{
    int fds;
    int fdt;
    int num;
    int numOfArgs = inputErrors(argc);
    const char *source = argv[1];

    // Если fds < 0 => файл не существует. O_RDONLY - доступ только на
    чтение
    if ((fds = open(argv[1], O_RDONLY)) < 0)
    {
```

```

        printf("Input file doesn't exist\n");
        exit(11);
    }

    // Если == 0 => F_OK - файл уже существует
    if (access(argv[2], F_OK) == 0)
    {
        if (numOfArgs == 3)
        {
            printf("Output file already exists. Specify [w] key to
rewrite the file\n");
            exit(10);
        }
        else if (numOfArgs == 4)
        {
            char key = *argv[numOfArgs - 1];

            if (key == 'w')
                printf("The output file will be rerwritten\n");
            else
            {
                printf("Wrong key\n");
                exit(101);
            }
        }
    }

    // Доступ на создание выходного файла
    if ((fdt = creat(argv[2], 0644)) < 0)
    {
        printf("Can not create the output file\n");
        exit(12);
    }

    // BUFSIZ - размер буфера
    while ((num = read(fds, buf, BUFSIZ)) > 0)
        write(fdt, recode(num), num);

    close(fds);
    close(fdt);

    return 0;
}

// Функция проверяет стандартный ввод на ошибки
int inputErrors(int argc)
{
    if (argc == 1)
    {

```

```

        printf("Specify the name of the input and output files and
flag[w] (optional)\n");
        exit(126);
    }

    if (argc == 2)
    {
        printf("Specify the name of the output file and flag[w]
(optional)\n");
        exit(127);
    }

    if (argc == 3)
    {
        printf("You can specify the flag[w] (optional)\n");
        return argc;
    }
    if (argc == 4)
        return argc;
    if (argc > 4)
    {
        printf("Too many arguments (max 4)\n");
        exit(128);
    }
}

```

```

unsigned char *recode(int n)
{
    unsigned char c;
    // Индекс - порядковый номер символа в КОИ-8
    // Значение - порядковый номер символа CP1251
    static char tab[64] = {
        //ю а б ц д е ф г х и й к л м н о
        30, 0, 1, 22, 4, 5, 20, 3, 21, 8, 9, 10, 11, 12, 13, 14,
        //п я р с т у ж в ь ы з ш э щ ч ь
        15, 31, 16, 17, 18, 19, 6, 2, 28, 27, 7, 24, 29, 25, 23, 26};

    for (int i = 0; i < n; i++)
    {
        c = buf[i];

        // Не кириллица
        if (c < 192)
            continue;
        // Строчные буквы
        if (c < 224)
            buf[i] = tab[c - 192] + 224;
        // Заглавные буквы
        else
            buf[i] = tab[c - 224] + 192;
    }
}

```

```
    }  
    return buf;  
}
```

Результат работы программы

KOI-8: Гусаров Аркадий

CP1251: Гусаров Аркадий