Министерство образования Республики Беларусь

Учреждение Образования

«Белорусский государственный университет информатики и радиоэлектроники»

Факультет компьютерных систем и сетей

Кафедра электронных вычислительных машин

Лабораторная работа №5

по дисциплине «Программирование на языке ассемблера»

на тему «Работа с файлами»

вариант №21

|  |  |  |
| --- | --- | --- |
| Выполнил  студент гр. 250501  Лукьянов Е.О. |  | Проверил  Туровец Н.О. |

Минск 2023

**Цель работы:** Ознакомиться с основными операциями обработки файлов, получить понятие о работе с параметрами командной строк

**Теоретические сведения**

Для выполнения работы требуется рассмотреть следующие элементы языка ассемблера и операционной системы:

1. Работа с файлами.

Для работы с файлами в данной лабораторной работе лучше всего использовать функции DOS, которые обращаются к файлу через 16-битный идентификатор (дескриптор) файла. Такой подход более прост, чем использование более старых описателей файла (37-байтного блока управления файлом FCB) или функций низкого уровня доступа к диску (прерывание BIOS 13h).

Первые пять значений идентификаторов такого формата инициализируются системой следующим образом:

-- 0 – STDIN – стандартное устройство ввода (клавиатура),

-- 1 – STDOUT – стандартное устройство вывода (экран),

-- 2 – STDERR – устройство вывода сообщений об ошибках (всегда экран),

-- 3 – AUX – последовательный порт (СОМ1),

-- 4 – PRN – параллельный порт (LPT1).

Работа с файлами выполняется через функции DOS в стандартном порядке:

1) создание или открытие существующего файла;

2) выполнение файловых операций чтения или записи данных;

3) закрытие файла.

Дополнительно доступны операции: удаление, поиск и управление.

Для создания или открытия существующего файла рекомендуется использовать функции:

-- Функция DOS 3Ch (INT 21h) – создать файл:

Ввод: AH = 3Ch,

СХ = атрибут файла (биты):

-- 7 – файл можно открывать разным процессам,

-- 6 – не используется,

-- 5 – архивный бит (1, если файл не сохранялся),

-- 4 – каталог (должен быть 0 для функции 3Ch),

-- 3 – метка тома (игнорируется функцией 3Ch),

-- 2 – системный файл,

-- 1 – скрытый файл,

-- 0 – файл только для чтения.

DS:DX = адрес ASCIZ-строки с полным именем файла.

Вывод: Если CF = 1, то ошибка операции (в AX код ошибки):

-- 03h – путь не найден,

-- 04h – слишком много открытых файлов,

-- 05h – доступ запрещен.

Если CF = 0, то операция выполнена успешно:

АХ = идентификатор файла.

Особенности: Если файл уже существует, функция все равно открывает его, присваивая ему нулевую длину.

-- Функция DOS 5Bh (INT 21h) – создать и открыть файл:

Ввод: AH = 5Bh,

СХ = атрибут файла,

DS:DX = адрес ASCIZ-строки с полным именем файла.

Вывод: Если CF = 1, то ошибка операции (в AX код ошибки):

-- 03h – путь не найден,

-- 04h – слишком много открытых файлов,

-- 05h – доступ запрещен,

-- 50h – файл уже существует.

Если CF = 0, то операция выполнена успешно

АХ = идентификатор файла, открытого для чтения/записи в режиме совместимости.

-- Функция DOS 5Ah (INT 21h) – создать и открыть временный файл:

Ввод: AH = 5Ah,

СХ = атрибут файла,

DS:DX = адрес ASCIZ-строки с путем, оканчивающимся символом «\», и тринадцатью нулевыми байтами в конце.

Вывод: Если CF = 1, то ошибка операции (в AX код ошибки):

-- 03h – путь не найден,

-- 04h – слишком много открытых файлов,

-- 05h – доступ запрещен.

Если CF = 0, то операция выполнена успешно:

АХ = идентификатор файла, открытого для чтения/записи в режиме совместимости (при этом в строку по адресу DS:DX дописывается имя созданного файла).

Особенности: Функция создает файл с уникальным именем, который не является на самом деле временным, его следует специально удалять, для чего его имя и записывается в строку в DS:DX.

-- Функция DOS 3Dh (INT 21h) – открыть существующий файл:

Ввод: AH = 3Dh,

AL = режим доступа (биты):

-- 0 – открыть для чтения,

-- 1 – открыть для записи,

-- 3 – 2 – 00 (резерв),

-- 6 – 4 – режим доступа для других процессов:

000 – режим совместимости,

001 – все операции запрещены,

010 – запись запрещена

011 – чтение запрещено

100 - запрещений нет.

-- 7 – файл не наследуется порождаемыми процессами.

DS:DX = адрес ASCIZ-строки с полным именем файла.

CL = маска атрибутов файла.

Вывод: Если CF = 1, то ошибка операции (в AX код ошибки):

-- 02h – файл не найден,

-- 03h – путь не найден,

-- 04h – слишком много открытых файлов,

-- 05h – доступ запрещен,

-- 0Ch – неверный режим доступа.

Если CF = 0, то операция выполнена успешно:

АХ = идентификатор файла.

Во всех случаях имя файла (если диск или путь отсутствуют в описании, то системой используются их текущие значения) описывается ASCIZ-строкой (строкой ASCII-символов, оканчивающейся нулем), которая, например, имеет следующий вид:

filename db 'с:\data\filename.txt',0

Для выполнения файловых операций чтения или записи данных рекомендуется использовать функции:

-- Функция DOS 3Fh (INT 21h) – чтение из файла или устройства:

Ввод: АН = 3Fh,

ВХ = идентификатор файла,

СХ = число байт для чтения,

DS:DX = адрес буфера для приема данных.

Вывод: Если CF = 1, то ошибка операции (в AX код ошибки):

-- 05h – доступ запрещен,

-- 06h – неверный идентификатор.

Если CF = 0, то операция выполнена успешно:

АХ = число считанных байт.

Особенности: Каждая следующая операция чтения, так же как и записи, начинается не с начала файла, а с того байта, на котором остановилась предыдущая операция чтения/записи. Если при чтении из файла число фактически считанных байт в АХ меньше, чем заказанное число в СХ, то при чтении был достигнут конец файла.

-- Функция DOS 42h (INT 21h) – переместить указатель чтения/записи: Ввод: АН = 42h,

ВХ = идентификатор файла,

CX:DX = расстояние, на которое надо переместить указатель (знаковое число),

AL = перемещение относительно:

0 – начала файла,

1 – текущей позиции,

2 – конца файла.

Вывод: Если CF = 1, то ошибка операции (в AX код ошибки):

-- 06h – неверный идентификатор.

Если CF = 0, то операция выполнена успешно:

CX:DX = новое значение указателя (в байтах от начала файла). Особенности: Указатель можно установить за реальными пределами файла:

-- если указатель устанавливается в отрицательное число, то следующая операция чтения/записи вызовет ошибку;

-- если указатель устанавливается в положительное число, большее длины файла, следующая операция записи увеличит размер файла.

Эта функция также часто используется для определения длины файла: достаточно вызвать ее с СХ = 0, DX = 0, AL = 2, и в CX:DX будет возвращена длина файла в байтах.

-- Функция DOS 40h (INT 21h) – запись в файл или устройство:

Ввод: АН = 40h,

ВХ = идентификатор файла,

СХ = число байт для записи,

DS:DX = адрес буфера с данными.

Вывод: Если CF = 1, то ошибка операции (в AX код ошибки):

-- 05h – доступ запрещен,

-- 06h – неверный идентификатор.

Если CF = 0, то операция выполнена успешно:

АХ = число записанных байт.

Особенности: Если при записи в файл указать СХ = 0, то файл будет обрезан по текущему значению указателя.

При записи в файл на самом деле происходит запись в буфер DOS, данные из которого сбрасываются на диск при закрытии файла или если их количество превышает размер сектора диска.

-- Функция DOS 68h (INT 21h) – сброс файловых буферов DOS на диск:

Ввод: АН = 68h,

ВХ = идентификатор файла.

Вывод: Если CF = 1, то ошибка операции (в AX код ошибки).

Если CF = 0, то операция выполнена успешно.

-- Функция DOS 0Dh (INT 21h) – сброс всех файловых буферов на диск: Ввод: АН = 0Dh

Для закрытия файла рекомендуется использовать следующую функцию:

-- Функция DOS 3Eh (INT 21h) – закрыть файл:

Ввод: АН = 3Eh,

ВХ = идентификатор файла.

Вывод: Если CF = 1, то ошибка операции (в AX код ошибки):

-- 06h – неверный идентификатор.

Если CF = 0, то операция выполнена успешно.

Особенности: Если файл был открыт для записи, то все файловые буферы сбрасываются на диск, устанавливается время модификации файла записывается его новая длина.

Дополнительные операции с файловой системой:

-- Функция DOS 41h (INT 21h) – удалить файл:

Ввод: АН = 41h,

DS:DX = адрес ASCIZ-строки с полным именем файла.

Вывод: Если CF = 1, то ошибка операции (в AX код ошибки):

-- 02h – файл не найден,

-- 03h – путь не найден,

-- 05h – доступ запрещен.

Если CF = 0, то операция выполнена успешно.

Особенности: Удалить файл можно только после того, как он будет закрыт, так как DOS будет продолжать выполнять запись в несуществующий файл, что может привести к разрушению файловой системы. Данная функция не позволяет использовать маску имени файла (символы «\*» и «?») для удаления сразу нескольких файлов (только с DOS 7.0).

-- Функция DOS 4Eh (INT 21h) – найти первый файл:

Ввод: АН = 4Eh,

AL используется при обращении к функции APPEND,

СХ = атрибуты файла (биты 0 и 5 игнорируются, если бит 3 установлен, то все остальные биты игнорируются),

DS:DX = адрес ASCIZ-строки с именем файла, которое может включать путь и маску для поиска (символы \* и ?).

Вывод: Если CF = 1, то ошибка операции (в AX код ошибки):

-- 02h – файл не найден,

-- 03h – путь не найден,

-- 12h – неверный режим доступа.

Если CF = 0, то операция выполнена успешно и область DTA заполняется данными о найденном файле.

Особенности: Вызов этой функции заполняет данными область памяти DTA (область передачи данных), которая начинается по умолчанию со смещения 0080h от начала блока данных PSP (при запуске com и exe программ сегменты DS и ES содержат сегментный адрес начала PSP), но ее можно переопределить с помощью функции DOS 1Ah.

Функции поиска файлов заполняют DTA следующим образом:

+00h (байт) – биты 0 – 6: ASCII-код буквы диска, бит 7: сетевой диск; +01h (11 байт) – маска поиска (без пути);

+0СН (байт) – атрибуты для поиска;

+0Dh (слово) – порядковый номер файла в каталоге;

+0Fh (слово) – номер кластера начала внешнего каталога;

+11h (4 байта) – зарезервировано;

+15h (байт) – атрибуты найденного файла;

+16h (слово) – время создания файла в формате DOS: биты 15 – 11: час (0 — 23), биты 10 – 05: минута, биты 04 – 00: номер секунды, деленный на 2 (0 – 30).

+18h (слово) – дата создания файла в формате DOS: биты 15 – 09: год, начиная с 1980, биты 08 – 05: месяц, биты 04 – 00: день.

+1Ah (4 байта) – размер файла;

+1Eh (13 байт) - ASCIZ-имя найденного файла (с расширением).

-- Функция DOS 1Ah (INT 21h) – установить область DTA:

Ввод: АН = 1Ah,

DS:DX = адрес начала DTA (128-байтный буфер).

-- Функция DOS 4Fh (INT 21h) – найти следующий файл:

Ввод: АН = 4Fh,

DTA должна содержать данные от предыдущего вызова функции 4Е или 4F.

Вывод: Если CF = 1, то ошибка операции (в AX код ошибки).

Если CF = 0, то операция выполнена успешно и область DTA заполняется данными о следующем найденном файле.

Особенности: Для продолжения поиска следует вызывать функцию 4Fh, пока не будет возвращена ошибка.

-- Функция DOS 39h (INT 21h) – создать каталог:

Ввод: АН = 39h,

DS:DX = адрес ASCIZ-строки с путем, в котором все каталоги, кроме последнего, существуют (для версии DOS 3.3 и более ранних длина всей строки не должна превышать 64 байта).

Вывод: Если CF = 1, то ошибка операции (в AX код ошибки):

-- 03h – путь не найден,

-- 05h – доступ запрещен.

Если CF = 0, то операция выполнена успешно.

-- Функция DOS 3Ah (INT 21h) – удалить каталог:

Ввод: АН = 3Ah,

DS:DX = адрес ASCIZ-строки с путем, последний каталог в котором будет удален (только если он пустой, не является текущим, не занят командой SUBST).

Вывод: Если CF = 1, то ошибка операции (в AX код ошибки):

-- 03h – путь не найден,

-- 05h – доступ запрещен,

-- 10h – текущий каталог нельзя удалить.

Если CF = 0, то операция выполнена успешно.

-- Функция DOS 47h (INT 21h) – определить текущий каталог:

Ввод: АН = 47h,

DL = номер диска (00h — текущий, 01h — А: и т.д.),

DS:SI = 64-байтный буфер для текущего пути (ASCIZ-строка без имени диска, первого и последнего символа «\»).

Вывод: Если CF = 1, то ошибка операции (в AX код ошибки):

-- 0Fh – указан несуществующий диск.

Если CF = 0 и АХ = 0100h, то операция выполнена успешно.

-- Функция DOS 3Bh (INT 21h) – сменить каталог:

Ввод: АН = 3Bh,

DS:DX = адрес 64-байтного ASCIZ-буфера с путем, который станет текущим каталогом.

Вывод: Если CF = 1, то ошибка операции (в AX код ошибки):

-- 03h – путь не найден.

Если CF = 0, то операция выполнена успешно.

1. Работа с командной строкой.

Передача параметров программы через командную строку при запуске программы – решение, которое позволяет построить гибкую по входным данным программу без дополнительных диалогов с пользователем.

При запуске программы DOS помещает всю командную строку (включая последний символ 0Dh) в блок PSP запущенной программы по смещению 81h и ее длину в байт 80h. Длина командной строки, хранящейся в PSP, не может быть больше 126 символов (командная строка большей длины доступна начиная с DOS 4.0 в переменной среды CMDLINE).

При загрузке программы, в начале отводимого для нее блока памяти, создается структура данных PSP (префикс программного сегмента) размером 256 байт (100h). Затем DOS создает копию текущего окружения для загружаемой программы, помещает полный путь и имя программы в конец окружения, заполняет поля PSP, сама программа записывается в память, начиная с адреса PSP:0100h.

При запуске com-программы регистры устанавливаются следующим образом:

1) AL = FFh, если первый параметр командной строки содержит неправильное имя диска (например, z:/something), иначе – AL = 00h.

2) АН = FFh, если второй параметр содержит неправильное имя диска, иначе АН = 00h.

3) CS = DS = ES = SS = сегментный адрес PSP.

4) SP = адрес последнего слова в сегменте (обычно FFFEh или меньше, если не хватает памяти).

При запуске exe-программы регистры SS:SP устанавливаются в соответствии с сегментом стека, определенным в программе.

Затем в стек помещается слово 0000h и выполняется переход на начало программы (PSP:0100h для com, собственная точка входа для exe).

**Код программы (.exe)**

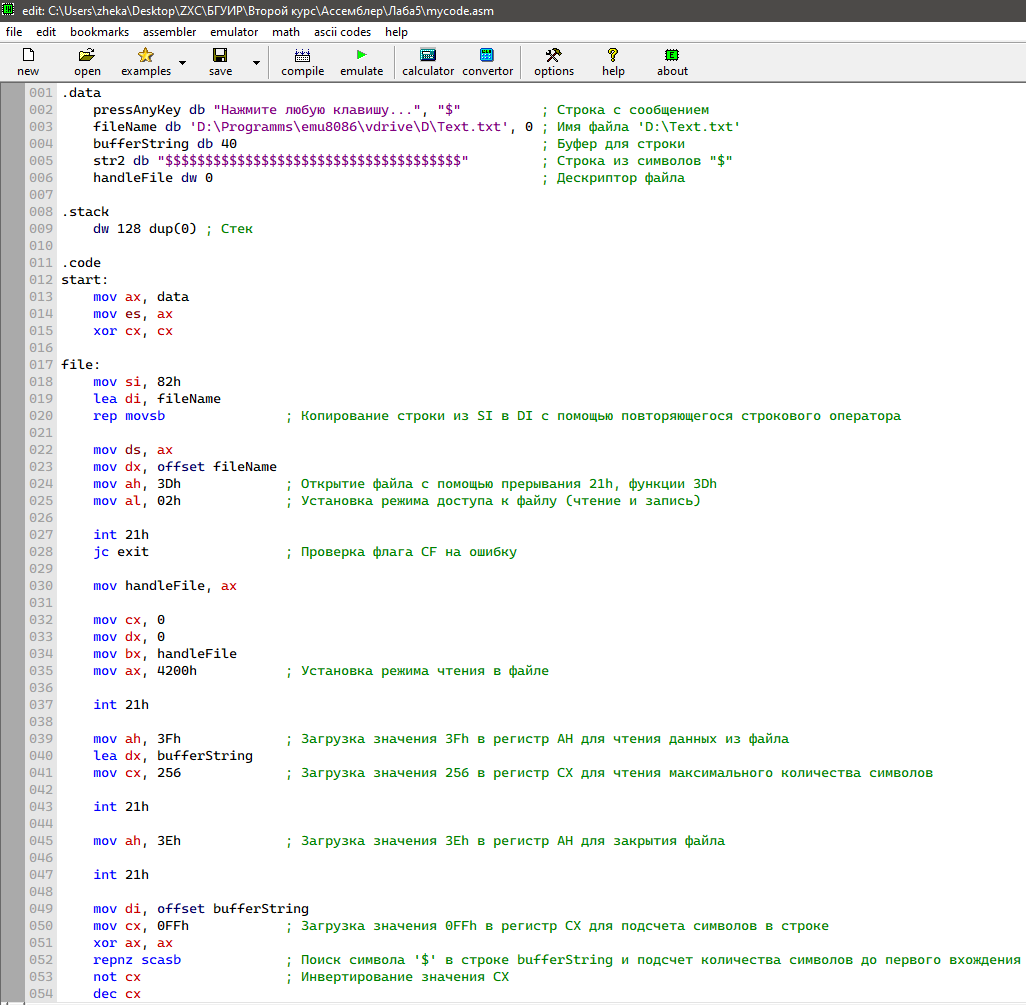
****

Рисунок 1 – Код программы

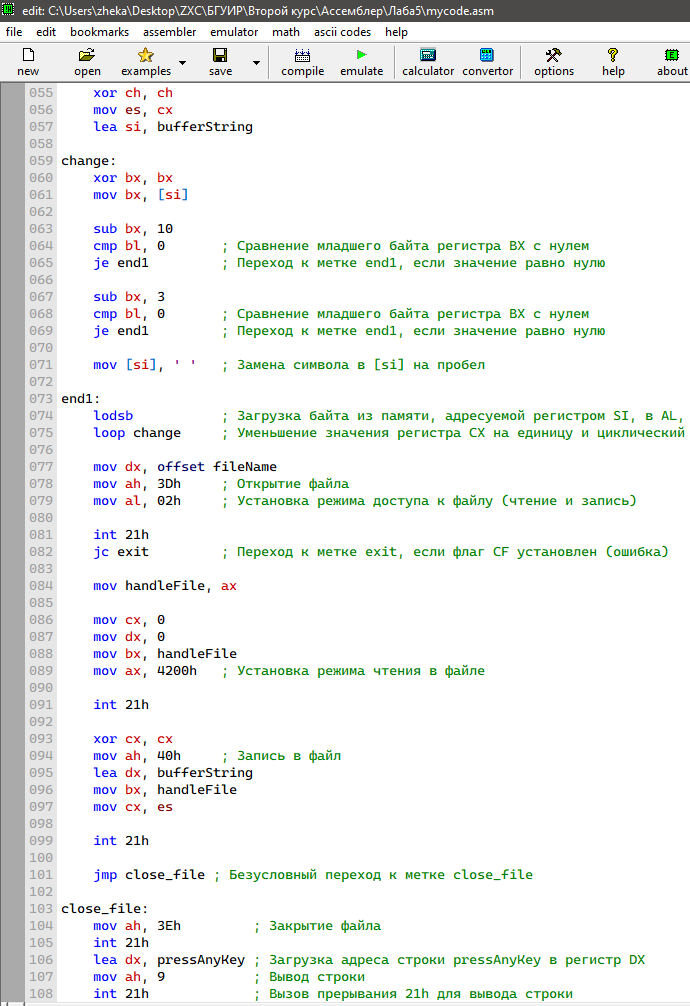


Рисунок 2 – Код программы

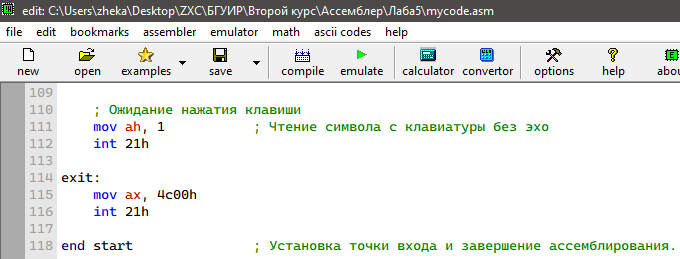
****

Рисунок 3 – Код программы

**Вывод программы**

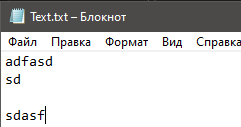


Рисунок 4 – Файл Text.txt до выполнения программы

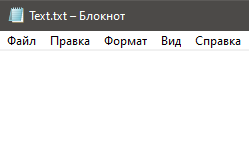


Рисунок 5 – Файл Text.txt после выполнения программы