

CSE1301 Computer Programming: Lecture 22 File I/O

1

Topics

- **Sequential** File Handling in C
- Basic File I/O Functions
- Example: Count words in a file

- Readings:
- King: Sections 22.1 to 22.3
- D&D (2/e or 3/e) Sections 11.1 to 11.5
- Also: Brookshear, Sections 3.1 to 3.3, 8.1 to 8.3

2

File Handling in C

- Files need to be *opened* before use.
 - Associate a "**file handler**" to each file
 - Modes: read, write, or append
- File input/output functions use the file handler (*not* the filename).
- Need to *close* the file after use.
- Basic file handling functions: **fopen()**, **fclose()**, **fscanf()**, **fprintf()**.

3

File I/O Example

- Write a program which:
 - inputs a list of names from a file called *names.lst*
 - counts the number of names in the list
 - asks a mark for each name in the file
 - outputs the name and corresponding mark to the file *names_marks.dat*
- Note: the tasks above are not necessarily accomplished in that order.

4

File I/O (Header)

- **Step 0:** Include **stdio.h**.

```
#include <stdio.h>

int main()
{
    ...
    return 0;
}
```

5

File I/O (File Pointer)

- **Step 1:** Declare a **file handler** (a.k.a. **file pointer**) as **FILE *** for each file.

```
int main()
{
    FILE *inputfile = NULL;
    FILE *outputfile = NULL;
    FILE *currentfile = NULL;
    ...
    return 0;
}
```

6

File I/O (Open)

- Step 2: Open file using `fopen()`.

```

int main()
{
    FILE *inputfile = NULL;
    FILE *outputfile = NULL;
    FILE *currentfile = NULL;

    inputfile = fopen("Names.txt", "r");
    outputfile = fopen("marks.dat", "w");
    currentfile = fopen("logfile.txt", "a");

    ...
    return 0;
}
    
```

File I/O (Open)

- Step 2: Open file using `fopen()`.

```

int main()
{
    FILE *inputfile = NULL;
    FILE *outputfile = NULL;
    FILE *currentfile = NULL;

    inputfile = fopen("Names.txt", "r");
    outputfile = fopen("marks.dat", "w");
    currentfile = fopen("logfile.txt", "a");

    ...
    return 0;
}
    
```

File name

File I/O (Open)

- Step 2: Open file using `fopen()`

```

int main()
{
    FILE *inputfile = NULL;
    FILE *outputfile = NULL;
    FILE *currentfile = NULL;

    inputfile = fopen("Names.txt", "r");
    outputfile = fopen("marks.dat", "w");
    currentfile = fopen("logfile.txt", "a");

    ...
    return 0;
}
    
```

Mode
r : read
w : write
a : append

Warning: The "w" mode overwrites the file, if it exists.

File I/O (Open)

- Step 2: Open file using `fopen()`.

```

int main()
{
    FILE *inputfile = NULL;
    FILE *outputfile = NULL;
    FILE *currentfile = NULL;

    inputfile = fopen("Names.txt", "r");
    outputfile = fopen("marks.dat", "w");
    currentfile = fopen("logfile.txt", "a");

    ...
    return 0;
}
    
```

Associate a file handler for every file to be used.

File I/O (Error Check)

- Step 3: Check if file is opened successfully.

```

int main()
{
    FILE *inputfile;

    inputfile = fopen("Names.txt", "r");

    if (inputfile == NULL)
    {
        printf("Unable to open input file.\n");
        return 1;
    }

    ...
    return 0;
}
    
```

File I/O (Error Check)

- Step 3: Check if file is opened successfully.

```

int main()
{
    FILE *inputfile;

    inputfile = fopen("Names.txt", "r");

    if (inputfile == NULL)
    {
        printf("Unable to open input file.\n");
        return 1;
    }

    ...
    return 0;
}
    
```

File handler becomes NULL when an fopen() error occurs.

File I/O (Error Check)

- Step 3: Check if file is opened successfully.

```
int main()
{
    FILE *inputfile;

    inputfile = fopen("Names.txt", "r");

    if (inputfile == NULL)
    {
        printf("Unable to open input file.\n");
        return 1;
    }

    ...
    return 0;
}
```

Ends program if inside main () function.

13

File I/O (Input)

- Step 4a: Use `fscanf()` for input.

```
#include <stdio.h>
#define MAXLEN 100

int main()
{
    FILE *inputfile = NULL;
    char name[MAXLEN];
    int count;
}
```

listnames.c

14

File I/O (Input)

- Step 4a: Use `fscanf()` for input.

```
#include <stdio.h>
#define MAXLEN 100

int main()
{
    FILE *inputfile = NULL;
    char name[MAXLEN];
    int count;
}
```

Recall: Macro definition

15

File I/O (Input)

- Step 4a: Use `fscanf()` for input.

```
/* Assuming "names.lst" contains a
list of names, open this file for
reading. */

inputfile = fopen("names.lst", "r");

if (inputfile == NULL)
{
    printf("Error opening names file.\n");
    return 1;
}
```

listnames.c

16

File I/O (Input)

- Step 4a: Use `fscanf()` for input.

```
/* Read in each name, and keep count how
many names there are in the file. */

count = 0;
while ( fscanf(inputfile, "%s", name) == 1 )
{
    count++;
    printf("%d. %s\n", count, name);
}

printf("\nNumber of names read: %d\n", count);
return 0;
}
```

listnames.c

17

File I/O (Input)

- Step 4a: Use `fscanf()` for input.

```
/* Read in each name, and keep count how
many names there are in the file. */

count = 0;
while ( fscanf(inputfile, "%s", name) == 1 )
{
    count++;
    printf("%d. %s\n", count, name);
}

printf("\nNumber of names
return 0;
}
```

listnames.c

Requires the file handler ("stream"), not the file name.

18

File I/O (Input)

- **Step 4a:** Use `fscanf()` for input.

listnames.c

```

/* Read in each name, and keep count how
many names there are in the file. */

count = 0;
while ( fscanf(inputfile, "%s", name) == 1 )
{
    count++;
    printf("%d. %s\n", count, name);
}

printf("\nNumber of names: %d\n", count);
return 0;
    
```

Other parameters:
like ordinary `scanf()`.

File I/O (Input)

- **Step 4a:** Use `fscanf()` for input.

listnames.c

```

/* Read in each name, and keep count how
many names there are in the file. */

count = 0;
while ( fscanf(inputfile, "%s", name) == 1 )
{
    count++;
    printf("%d. %s\n", count, name);
}

printf("\nNumber of names: %d\n", count);
return 0;
    
```

fscanf() returns the number of
input items converted and
assigned successfully.

File I/O (Input)

- **Step 4a:** Use `fscanf()` for input.

listnames.c

```

/* Read in each name, and keep count how
many names there are in the file. */

count = 0;
while ( fscanf(inputfile, "%s", name) == 1 )
{
    count++;
    printf("%d. %s\n", count, name);
}

printf("\nNumber of names: %d\n", count);
return 0;
    
```

Used to check if a read or
assignment error occurred, or end of
input file has been reached.

File I/O (Output)

- **Step 4b:** Use `fprintf()` for output.

listnames2.c

```

#include <stdio.h>
#define MAXLEN 100

int main()
{
    FILE *inputfile = NULL;
    FILE *outfile = NULL;
    char name[MAXLEN];
    int count;
    float mark;

    /* Assuming "names.lst" contains a list of names,
    open this file for reading. */

    inputfile = fopen("names.lst", "r");
    if (inputfile == NULL)
    {
        printf("Error opening names file.\n");
        return 1;
    }
}
    
```

File I/O (Output)

- **Step 4b:** Use `fprintf()` for output.

listnames2.c

```

/* The output file "names_marks.dat" will
contain the list of names and
corresponding marks. */

outfile = fopen("names_marks.dat", "w");

if (outfile == NULL)
{
    printf("Error opening output file.\n");
    return 1;
}
    
```

File I/O (Output)

- **Step 4b:** Use `fprintf()` for output.

listnames2.c

```

/* Read in each name, ask for the mark, and write name and mark to
output file. Also keep count how many names there are in the
file. */

count = 0;
while ( fscanf(inputfile, "%s", name) == 1 )
{
    count++;
    printf("Enter mark for %s: ", name);
    scanf("%f", &mark);

    if ( fprintf(outfile, "%s %f\n", name, mark) <= 0 )
    {
        printf("Error writing to output file.\n");
        return 1;
    }
}
/**/ etc /**/
    
```

File I/O (Output)

- **Step 4b:** Use `fprintf()` for output.

```

/* Read in each name, ask for the mark, and write name and mark to
output file. Also keep count how many names there are in the
file. */

count = 0;
while ( fscanf(inputfile, "%s", name) == 1 )
{
    count++;

    printf("Enter mark for %s: ", name);
    scanf("%f", &mark);

    if ( fprintf(outfile, "%s %f\n", name, mark) <= 0 )
    {
        printf("Error writing to output file.\n");
        return 1;
    }
}
/** etc **/
    
```

listnames2.c

25

File handler, not the file name.

File I/O (Output)

- **Step 4b:** Use `fprintf()` for output.

```

/* Read in each name, ask for the mark, and write name and mark to
output file. Also keep count how many names there are in the
file. */

count = 0;
while ( fscanf(inputfile, "%s", name) == 1 )
{
    count++;

    printf("Enter mark for %s: ", name);
    scanf("%f", &mark);

    if ( fprintf(outfile, "%s %f\n", name, mark) <= 0 )
    {
        printf("Error writing to output file.\n");
        return 1;
    }
}
/** etc **/
    
```

listnames2.c

26

Other parameters: like ordinary printf().

File I/O (Output)

- **Step 4b:** Use `fprintf()` for output.

```

/* Read in each name, ask for the mark, and write name and mark to
output file. Also keep count how many names there are in the
file. */

count = 0;
while ( fscanf(inputfile, "%s", name) == 1 )
{
    count++;

    printf("Enter mark for %s: ", name);
    scanf("%f", &mark);

    if ( fprintf(outfile, "%s %f\n", name, mark) <= 0 )
    {
        printf("Error writing to output file.\n");
        return 1;
    }
}
/** etc **/
    
```

listnames2.c

fprintf() returns the number of characters written out successfully, or negative if an error occurs.

File I/O (Close)

- **Step 5:** Close file using `fclose()`

```

int main()
{
    /** etc **/

    printf("\n");
    printf("Number of names read: %d\n", count);

    fclose(inputfile);
    fclose(outfile);

    return 0;
}
    
```

28

File I/O (Close)

- **Step 5:** Close file using `fclose()`

```

int main()
{
    /** etc **/

    printf("\n");
    printf("Number of names read: %d\n", count);

    fclose(inputfile);
    fclose(outfile);

    return 0;
}
    
```

29

File handler, not the file name.

File I/O (Close)

- **Step 5:** Close file using `fclose()`

```

int main()
{
    /** etc **/

    printf("\n");
    printf("Number of names read: %d\n", count);

    fclose(inputfile);
    fclose(outfile);

    return 0;
}
    
```

- Clears input buffer.
- Flushes output buffer.
- `fclose()` fails when the file was not opened successfully.

Notes on Filenames

- Unless a directory path is specified, the program will look for the file in the current directory.
- Directory paths in filenames: DOS/Windows

```
sysFile = fopen("C:\\win\\system.ini", "r");
```
- Directory paths in filenames: Unix

```
passFile = fopen("/usr/etc/passwd", "r");
```

31

Notes on Filenames

- Variable filenames:

```
FILE *outFile = NULL;
char someName[MAX_NAME_LEN];

printf("Please enter output filename: ");
scanf("%s", someName);

outFile = fopen(someName, "w");
```

32

File I/O and Streams

- *Recall*: A **stream** serves as a channel to convey characters between I/O devices and programs.
- Standard streams: **stdin**, **stdout**, **stderr**
- A file handler/pointer serves as a **stream** to/from a file.
- Once an item is read from an input stream, the **file position** moves automatically, and the next item (if any) becomes available for the next read ("**sequential access**").

33

Notes on Strings and fscanf()

- Reading in a string:
 - fscanf(stream, "%s", string)**
 - Reads only a "word" at a time.
 - Words are separated by a *white-space*: (space, tab, newline, or any combination of these)
 - Moves to the next word in the stream automatically after each read.
- **scanf("%s", string)**
 - behaves similarly, except input stream is **stdin**.
 - eg: `eg := fscanf(stdin, "%s", string)`

34

Checking for EOF

- Both **scanf()** and **fscanf()** return:
 - the **number of input items** converted and assigned successfully
 - or the constant value **EOF** when an error or end-of-file occurs, *but...*

Not recommended!

```
count = 0;
while ( fscanf(inputfile, "%s", name) != EOF )
{
    count++;
    printf("%d. %s\n", count, name);
}
```

35

Checking for EOF

- Warning! Use **EOF** with caution!

Can cause bad problems if the conversion specifiers do not match the file's contents.

```
while ( fscanf(inpfile, "%s %f", name, &mark) != EOF )
{
    printf("%s\t%f\n", name, mark);
}
```

listmarks.c

Example: `infile: Jake absent`

36

Checking for EOF

- To check for **end-of-file** (or any other **input error**), check that the **number of items** converted and assigned successfully is **equal** to the **expected** number of items.

```
while ( fscanf(inpF, "%s %f", name, &mark) == 2 )
{
    printf("%s\t%f\n", name, mark);
}
```

listmarks.c

37

Checking for EOF

- To check for **end-of-file** (or any other **input error**), check that the **number of items** converted and assigned successfully is **equal** to the **expected** number of items.

Ditto for scanf().

```
if ( scanf("%d %d %d", &page, &row, &col) != 3 )
{
    printf( "I cannot go on without 3 integers :-( \n" );
    exit(1);
}
```

testscanf1.c

38

Checking for EOF

- To check for end-of-file (or any other input error), check that the number of items converted and assigned successfully is equal to the expected number of items.

The `exit()` function causes the program to terminate immediately; Requires `#include <stdlib.h>`

```
if ( scanf("%d %d %d", &page, &row, &col) != 3 )
{
    printf( "I cannot go on without 3 integers :-( \n" );
    exit(1);
}
```

39

Example: Count Words

- Write a program which counts the number of "words" in a file.
 - Note that as far as `scanf()` and `fscanf()` are concerned, any sequence of non-whitespace characters is a "word."

40

Count Words: Algorithm

ask the user for the name of the file

open the file

check if file is opened successfully

count the number of words in the file

print out the count

close the file

Duh?

41

Count Words: Algorithm

To count the number of words in a file:

```
set count to 0
loop
{
    read a word from the file
    if attempt to read a word failed
    then { exit loop }
    add 1 to count
}
```

Gotta write a function for this!

42

Function: countWords()

- Function prototype:
`int countWords (FILE *inpf);`
- Description:
 - This function returns the number of "words" in the input stream `inpf`.

43

Function: countWords()

- PRE-Conditions:
 - It assumes that `inpf` is a pointer to a file which has been opened successfully. There is no check for that within the function.
 - It also assumes that the file position is at the start of the input file/stream.
 - Note that a "word" here means any string of characters, separated from other words by any whitespace (ie. space, tab, newline, or combination).
 - It assumes that no "word" in the file has more than **(MAXLEN - 1)** characters.

44

Function: countWords()

- PRE-Conditions:
 - It assumes that `inpf` is a pointer to a file which has been opened successfully. There is no check for that within the function.
 - It also assumes that the file position is at the start of the input file/stream.
 - Note that a "word" here means any string of characters, separated from other words by any whitespace (ie. space, tab, newline, or combination).
 - It assumes that no "word" in the file has more than **(MAXLEN - 1)** characters.

We will use a #define to set the value of MAXLEN.
(See `countwords.h`)

45

Function: countWords()

- POST-Conditions:
 - At the end of the function, the file position will be at the end of file.
 - The function returns an integer value which is the number of "words" in the file.

46

Function: countWords()

```

int
countWords ( FILE *inpf )
{
    char word[MAXLEN];
    int count = 0;

    while ( fscanf(inpf, "%s", word) == 1 )
    {
        count++;
    }

    return count;
}
    
```

`countwords.c` 47

Count Words: Algorithm

ask the user for the name of the file
 open the file
 check if file is opened successfully

count the number of words in the file

print the count
 close the file

I can also write a reusable function for these!

48

Function: openInput ()

- Function prototype:
`FILE* openInput (void);`
- Description:
 - This function keeps asking the user for a filename, until it is able to open the file successfully for input.

49

Function: openInput ()

- PRE-Condition:
 - It assumes that the filename fits in a string of size MAXLEN (including the '\0').

50

Function: openInput ()

- POST-Conditions:
 - It can cause the program to terminate if the user chooses to abort the operation.
 - It returns the file handler/pointer for the specified file.
 - It assumes that the calling function has the corresponding variable to catch the return value.
 - It also assumes that the calling function takes care of closing the file.

51

```

FILE* openInput ( void )
{
    FILE *handle = NULL;
    char theFile[MAXLEN];
    int option;

    while (1)
    {
        printf("Enter file name: ");
        scanf("%s", theFile);

        if ((handle = fopen(theFile, "r")) == NULL )
        {
            /* Insert code to handle open error. */
        }
        else
        {
            break;
        }
    }
    return handle;
}
    
```

52

Code to handle open error:

```

printf("Error opening file.\n");

option = 0; /* Set default to abort. */

printf("\nEnter 1 to try again, ");
printf("or any number to abort: ");

scanf("%d", &option);
printf("\n");

if ( option != 1 )
{
    printf("Alright then. ");
    printf("Program terminated.\n");
    exit(1);
}
    
```

53

Main Algorithm

```

set file to be the result of openInput()

set count to the result of countWords(file)

print out the count

close the file
    
```

54

Test Program #1

```
#include <stdio.h>
#include <stdlib.h>

#include "countwords.h"
#include "countwords.c"

int main()
{
    FILE *inputFile = NULL;
    int count;

    inputFile = openInput();

    count = countWords(inputFile);
    printf("\nThere are %d words in the file.\n", count);

    fclose(inputFile);

    return 0;
}
```

testcount1.c 55

Test Program #2

```
#include <stdio.h>
#include <stdlib.h>
#include "countwords.h"
#include "countwords.c"

int main()
{
    FILE *inputFile = NULL;
    int count;

    inputFile = openInput();

    count = countWords(inputFile);
    printf("\nThere are %d words in the file.\n", count);

    count = countWords(inputFile);
    printf("\nThere are %d words in the file.\n", count);

    fclose(inputFile);
    return 0;
}
```

What is the result if we call the countWords() function a second time over the same file?

testcount2.c

Summary

- Sequential File handling
- EOF/Error Checking
- Additions to your "C vocabulary":
 - FILE *
 - fopen(), fscanf(), fprintf(), fclose()
 - EOF
 - exit()

57