BBM 101 – Introduction to Programming I

Fall 2013, Lecture 13

Instructors: Aykut Erdem, Erkut Erdem, Fuat Akal
TAs: Yasin Sahin, Ahmet Selman Bozkir, Gultekin Isik
Today

- **File Input and Output**
  - Data Files
  - Data Hierarchy
  - Files and Streams
  - Creating a Sequential Access File
  - Reading Data from a File
  - Reading Data from a Sequential Access File
  - Random-Access Files
  - Creating a Randomly Accessed File
  - Writing Data Randomly to a Randomly Accessed File

- **Strings**
  - The Data Type char
  - Characters and Integers
  - Input and Output of Characters
  - Declaring Strings
  - Initializing Strings
  - Strings and Pointers
  - String Handling Functions (*string.h*)
Today

- **File Input and Output**
  - Data Files
  - Data Hierarchy
  - Files and Streams
  - Creating a Sequential Access File
  - Reading Data from a File
  - Reading Data from a Sequential Access File
  - Random-Access Files
  - Creating a Randomly Accessed File
  - Writing Data Randomly to a Randomly Accesssed File

- **Strings**
  - The Data Type char
  - Characters and Integers
  - Input and Output of Characters
  - Declaring Strings
  - Initializing Strings
  - Strings and Pointers
  - String Handling Functions (*string.h*)
Introduction

Data files

- Can be created, updated, and processed by C programs
- Are used for permanent storage of large amounts of data
  - Storage of data in variables and arrays is only temporary

When you use a file to store data for use by a program, that file usually consists of text (alphanumeric data) and is therefore called a text file.
The Data Hierarchy

Data Hierarchy:

- Bit – smallest data item
  - Value of 0 or 1
- Byte – 8 bits
  - Used to store a character
    - Decimal digits, letters, and special symbols
- Field – group of characters conveying meaning
  - Example: your name
- Record – group of related fields
  - Represented by a **struct** or a **class**
  - Example: In a payroll system, a record for a particular employee that contained his/her identification number, name, address, etc.
The Data Hierarchy

- Data Hierarchy (continued):
  - File – group of related records
    - Example: payroll file
  - Database – group of related files

![Diagram of data hierarchy]

Fig. 11.1 The data hierarchy.
The Data Hierarchy

- **Data files**
  - Record key
    - Identifies a record to facilitate the retrieval of specific records from a file
  - Sequential file
    - Records typically sorted by key
Files and Streams

- C views each file as a sequence of bytes
  - File ends with the \textit{end-of-file marker}
    - Or, file ends at a specified byte

- Stream created when a file is opened
  - Provide communication channel between files and programs
  - Opening a file returns a pointer to a \texttt{FILE} structure
    - Example file pointers:
      - \texttt{stdin} - standard input (keyboard)
      - \texttt{stdout} - standard output (screen)
      - \texttt{stderr} - standard error (screen)
Files and Streams

- **FILE** structure
  - File descriptor
    - Index into operating system array called the open file table
  - File Control Block (FCB)
    - Found in every array element, system uses it to administer the file

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>…</th>
<th>n-1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 11.2  C’s view of a file of n bytes.
Files and Streams

- Read/Write functions in standard library
  - `fgetc`
    - Reads one character from a file
    - Takes a `FILE` pointer as an argument
    - `fgetc(stdin)` equivalent to `getchar()`
  - `fputc`
    - Writes one character to a file
    - Takes a `FILE` pointer and a character to write as an argument
    - `fputc('a', stdout)` equivalent to `putchar('a')`
  - `fgets`
    - Reads a line from a file
  - `fputs`
    - Writes a line to a file
  - `fscanf / fprintf`
    - File processing equivalents of `scanf` and `printf`
/* Create a sequential file */
#include <stdio.h>

int main()
{
    int account;
    char name[30];
    double balance;
    FILE *cfPtr; /* cfPtr = clients.dat file pointer */

    if ( ( cfPtr = fopen( "clients.dat", "w" ) ) == NULL )
        printf( "File could not be opened\n" );
    else
    {
        printf( "Enter the account, name, and balance.\n" );
        printf( "Enter EOF to end input.\n" );
        printf( "? " );
        scanf( "%d %s %.2f", &account, name, &balance );

        while ( !feof( stdin ) ) {
            fprintf( cfPtr, "%d %s %.2f\n", account, name, balance );
            printf( "? " );
            scanf( "%d %s %lf", &account, name, &balance );
        }

        fclose( cfPtr );
    }

    return 0;
}
Program Output

Enter the account, name, and balance.
Enter EOF to end input.
? 100 Jones 24.98
? 200 Doe 345.67
? 300 White 0.00
? 400 Stone -42.16
? 500 Rich 224.62
? ^Z
Creating a Sequential Access File

■ Creating a File

- FILE *myPtr;
  - Creates a FILE pointer called myPtr
- myPtr = fopen(filename, openmode);
  - Function fopen returns a FILE pointer to file specified
  - Takes two arguments – file to open and file open mode
  - If open fails, NULL returned

<table>
<thead>
<tr>
<th>Computer system</th>
<th>Key combination</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNIX systems</td>
<td>&lt;return&gt; &lt;ctrl&gt; d</td>
</tr>
<tr>
<td>IBM PC and compatibles</td>
<td>&lt;ctrl&gt; z</td>
</tr>
<tr>
<td>Macintosh</td>
<td>&lt;ctrl&gt; d</td>
</tr>
</tbody>
</table>
Creating a Sequential Access File

<table>
<thead>
<tr>
<th>Mode</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>Open a file for reading.</td>
</tr>
<tr>
<td>w</td>
<td>Create a file for writing. If the file already exists, discard the current contents.</td>
</tr>
<tr>
<td>a</td>
<td>Append; open or create a file for writing at end of file.</td>
</tr>
<tr>
<td>r+</td>
<td>Open a file for update (reading and writing).</td>
</tr>
<tr>
<td>w+</td>
<td>Create a file for update. If the file already exists, discard the current contents.</td>
</tr>
<tr>
<td>a+</td>
<td>Append; open or create a file for update; writing is done at the end of the file.</td>
</tr>
<tr>
<td>rb</td>
<td>Open a file for reading in binary mode.</td>
</tr>
<tr>
<td>wb</td>
<td>Create a file for writing in binary mode. If the file already exists, discard the current contents.</td>
</tr>
<tr>
<td>ab</td>
<td>Append; open or create a file for writing at end of file in binary mode.</td>
</tr>
<tr>
<td>rb+</td>
<td>Open a file for update (reading and writing) in binary mode.</td>
</tr>
<tr>
<td>wb+</td>
<td>Create a file for update in binary mode. If the file already exists, discard the current contents.</td>
</tr>
<tr>
<td>ab+</td>
<td>Append; open or create a file for update in binary mode; writing is done at the end of the file.</td>
</tr>
</tbody>
</table>
Creating a Sequential Access File

- `fprintf`
  - Used to print to a file
  - Like `printf`, except first argument is a `FILE` pointer (pointer to the file you want to print in)

- `feof( FILE pointer )`
  - Returns true if end-of-file indicator (no more data to process) is set for the specified file

- `fclose( FILE pointer )`
  - Closes specified file
  - Performed automatically when program ends
  - Good practice to close files explicitly

Details

- Programs may process no files, one file, or many files
- Each file must have a unique name and should have its own pointer
Reading Data from a File

- Reading a sequential access file
  - Create a `FILE` pointer, link it to the file to read
    ```c
    myPtr = fopen( "myfile.dat", "r" );
    ```
  - Use `fscanf` to read from the file
    ```c
    fscanf( myPtr, "%d %s %f", &account, name, &balance );
    ```
  - Data read from beginning to end
  - File position pointer
    - Indicates number of next byte to be read / written
    - Not really a pointer, but an integer value (specifies byte location)
    - Also called byte offset
    ```c
    rewind( myPtr )
    ```
    - Repositions file position pointer to beginning of file (byte 0)
/* Reading and printing a sequential file */
#include <stdio.h>

int main()
{
    int account;
    char name[30];
    double balance;
    FILE *cfPtr;  /* cfPtr = clients.dat file pointer */

    if ( ( cfPtr = fopen( "clients.dat", "r" ) ) == NULL )
        printf( "File could not be opened\n" );
    else
    {
        printf( "%-10s%-13s%s
", "Account", "Name", "Balance" );
        fscanf( cfPtr, "%d %s %lf", &account, name, &balance );

        while ( !feof( cfPtr ) )
        {
            printf( "%-10d%-13s%7.2f\n", account, name, balance );
            fscanf( cfPtr, "%d %s %lf", &account, name, &balance );
        }

        fclose( cfPtr );
    }

    return 0;
}
```c
#include <stdio.h>

int main()
{
    FILE *fileA, /* first input file */
        *fileB, /* second input file */
        *fileC; /* output file to be created */
    int num1, /* number to be read from first file */
          num2; /* number to be read from second file */
    int f1, f2;

    /* Open files for processing */
    fileA = fopen("class1.txt","r");
    fileB = fopen("class2.txt","r");
    fileC = fopen("class.txt","w");
```
/* As long as there are numbers in both files, read and compare numbers one by one. Write the smaller number to the output file and read the next number in the file from which the smaller number is read. */

f1 = fscanf(fileA, "%d", &num1);
f2 = fscanf(fileB, "%d", &num2);

while ((f1!=EOF) && (f2!=EOF)){
    if (num1 < num2){
        fprintf(fileC,"%d\n", num1);
        f1 = fscanf(fileA, "%d", &num1);
    }
    else if (num2 < num1) {
        fprintf(fileC,"%d\n", num2);
        f2 = fscanf(fileB, "%d", &num2);
    }
    else {  /* nums are equal:read from both files */
        fprintf(fileC,"%d\n", num1);
        f1 = fscanf(fileA, "%d", &num1);
        f2 = fscanf(fileB, "%d", &num2);
    }
}
while (f1!=EOF){ /* if reached end of second file, read the remaining numbers from first file and write to output file */
    fprintf(fileC,"%d\n", num1);
    f1 = fscanf(fileA, "%d", &num1);
}
while (f2!=EOF){ /* if reached the end of first file, read the remaining numbers from second file and write to output file */
    fprintf(fileC,"%d\n", num2);
    f2 = fscanf(fileB, "%d", &num2);
}

/* close files */
fclose(fileA);
fclose(fileB);
fclose(fileC);
return 0;
} /* end of main */
Reading Data from a Sequential Access File

- Sequential access file
  - Cannot be modified without the risk of destroying other data
  - Fields can vary in size
    - Different representation in files and screen than internal representation
    - 1, 34, -890 are all ints, but have different sizes on disk

300 White 0.00 400 Jones 32.87 (old data in file)
If we want to change White's name to Worthington,

300 Worthington 0.00

300 White 0.00 400 Jones 32.87

300 Worthington 0.00 ones 32.87

Data gets overwritten
Random-Access Files

- Random access files
  - Access individual records without searching through other records
  - Instant access to records in a file
  - Data can be inserted without destroying other data
  - Data previously stored can be updated or deleted without overwriting

- Implemented using fixed length records
  - Sequential files do not have fixed length records

```
0  100  200  300  400  500
```

```
| 100 bytes | 100 bytes | 100 bytes | 100 bytes | 100 bytes | 100 bytes |
```
Creating a Randomly Accessed File

- Data in random access files
  - Unformatted (stored as "raw bytes")
    - All data of the same type (ints, for example) uses the same amount of memory
    - All records of the same type have a fixed length
    - Data not human readable
Creating a Randomly Accessed File

- Unformatted I/O functions
  - fwrite
    - Transfer bytes from a location in memory to a file
  - fread
    - Transfer bytes from a file to a location in memory
  - Example:
    ```c
    fwrite( &number, sizeof( int ), 1, myPtr );
    ```
    - &number – Location to transfer bytes from
    - sizeof( int ) – Number of bytes to transfer
    - 1 – For arrays, number of elements to transfer
      - In this case, "one element" of an array is being transferred
    - myPtr – File to transfer to or from
Creating a Randomly Accessed File

- **Writing structs**
  ```c
  fwrite( &myObject, sizeof (struct myStruct), 1, myPtr );
  ```
  - `sizeof` – returns size in bytes of object in parentheses

- **To write several array elements**
  - Pointer to array as first argument
  - Number of elements to write as third argument
Creating a randomly accessed file sequentially */

#include <stdio.h>

/* clientData structure definition */
struct clientData {
    int acctNum;  /* account number */
    char lastName[15];  /* account last name */
    char firstName[10];  /* account first name */
    double balance;  /* account balance */
};  /* end structure clientData */

int main()
{
    int i;  /* counter */

    /* create clientData with no information */
    struct clientData blankClient = {0, "", "", 0.0};

    FILE *cfPtr;  /* credit.dat file pointer */

    /* fopen opens the file; exits if file cannot be opened */
    if ((cfPtr = fopen("credit.dat", "wb")) == NULL) {
        printf("File could not be opened.\n");
    } /* end if */
else {

    /* output 100 blank records to file */
    for ( i = 1; i <= 100; i++ ) {
        fwrite( &blankClient, sizeof( struct clientData ), 1, cfPtr );
    } /* end for */

    fclose ( cfPtr ); /* fclose closes the file */
} /* end else */

return 0; /* indicates successful termination */
} /* end main */
Writing Data Randomly to a Randomly Accessed File

- **fseek**
  - Sets file position pointer to a specific position
  - `fseek( pointer, offset, symbolic_constant );`
    - `pointer` – pointer to file
    - `offset` – file position pointer (0 is first location)
    - `symbolic_constant` – specifies where in file we are reading from
      - `SEEK_SET` – seek starts at beginning of file
      - `SEEK_CUR` – seek starts at current location in file
      - `SEEK_END` – seek starts at end of file
/**
 * Writing to a random access file */
#include <stdio.h>

/* clientData structure definition */
struct clientData {
    int acctNum;    /* account number */
    char lastName[15];    /* account last name */
    char firstName[10];  /* account first name */
    double balance;      /* account balance */
}; /* end structure clientData */

int main()
{
    FILE *cfPtr;    /* credit.dat file pointer */

    /* create clientData with no information */
    struct clientData client = { 0, "", "", 0.0 }; 

    /* fopen opens the file; exits if file cannot be opened */
    if ( ( cfPtr = fopen("credit.dat", "rb+" ) ) == NULL ) {
        printf( "File could not be opened.\n" );
    } /* end if */
    else {

/* require user to specify account number */
printf( "Enter account number"
        " ( 1 to 100, 0 to end input )\n? " );
scanf( "%d", &client.acctNum );

/* user enters information, which is copied into file */
while ( client.acctNum != 0 ) {
    /* user enters last name, first name and balance */
    printf( "Enter lastname, firstname, balance\n? " );

    /* set record lastName, firstName and balance value */
    fscanf( stdin, "%s%s%lf", client.lastName,
            client.firstName, &client.balance );

    /* seek position in file of user-specified record */
    fseek( cfPtr, ( client.acctNum - 1 ) *
            sizeof( struct clientData ), SEEK_SET );

    /* write user-specified information in file */
    fwrite( &client, sizeof( struct clientData ), 1, cfPtr );

    /* enable user to specify another account number */
    printf( "Enter account number\n? " );
    scanf( "%d", &client.acctNum );
} /* end while */

fclose( cfPtr ); /* fclose closes the file */
} /* end else */

return 0; /* indicates successful termination */

} /* end main */

Enter account number ( 1 to 100, 0 to end input )?
? 37
Enter lastname, firstname, balance
? Barker Doug 0.00
Enter account number
? 29
Enter lastname, firstname, balance
? Brown Nancy -24.54
Enter account number
? 96
Enter lastname, firstname, balance
? Stone Sam 34.98
Enter account number
? 88
Enter lastname, firstname, balance
? Smith Dave 258.34
Enter account number
? 33
Enter lastname, firstname, balance
? Dunn Stacey 314.33
Enter account number
? 0
Writing Data Randomly to a Randomly Accessed File

Fig. 11.14 The file position pointer indicating an offset of 5 bytes from the beginning of the file.
Reading Data Randomly from a Randomly Accessed File

- **fread**
  - Reads a specified number of bytes from a file into memory
    
    ```c
    fread( &client, sizeof (struct clientData), 1, myPtr );
    ```
  - Can read several fixed-size array elements
    - Provide pointer to array
    - Indicate number of elements to read
  - To read multiple elements, specify in third argument
/* Fig. 11.15: fig11_15.c */
Reading a random access file sequentially */
#include <stdio.h>

/* clientData structure definition */
struct clientData {
    int acctNum;    /* account number */
    char lastName[15];  /* account last name */
    char firstName[10];  /* account first name */
    double balance;    /* account balance */
}; /* end structure clientData */

int main()
{
    FILE *cfPtr; /* credit.dat file pointer */

    /* create clientData with no information */
    struct clientData client = { 0, '', '', 0.0 }; 

    /* fopen opens the file; exits if file cannot be opened */
    if ( ( cfPtr = fopen( "credit.dat", "rb" ) ) == NULL ) {
        printf( "File could not be opened.\n" );
    } /* end if */
```c
else {
    printf( "%-6s%-16s%-11s%10s\n", "Acct", "Last Name", "First Name", "Balance" );

    /* read all records from file (until eof) */
    while ( !feof( cfPtr ) ) {
        fread( &client, sizeof( struct clientData ), 1, cfPtr );

        /* display record */
        if ( client.acctNum != 0 ) {
            printf( "%-6d%-16s%-11s%10.2f\n", 
                    client.acctNum, client.lastName, 
                    client.firstName, client.balance );
        } /* end if */
    } /* end while */

    fclose( cfPtr ); /* fclose closes the file */
} /* end else */

return 0;
```

<table>
<thead>
<tr>
<th>Acct</th>
<th>Last Name</th>
<th>First Name</th>
<th>Balance</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>Brown</td>
<td>Nancy</td>
<td>-24.54</td>
</tr>
<tr>
<td>33</td>
<td>Dunn</td>
<td>Stacey</td>
<td>314.33</td>
</tr>
<tr>
<td>37</td>
<td>Barker</td>
<td>Doug</td>
<td>0.00</td>
</tr>
<tr>
<td>88</td>
<td>Smith</td>
<td>Dave</td>
<td>258.34</td>
</tr>
<tr>
<td>96</td>
<td>Stone</td>
<td>Sam</td>
<td>34.98</td>
</tr>
</tbody>
</table>
Today

- **File Input and Output**
  - Data Files
  - Data Hierarchy
  - Files and Streams
  - Creating a Sequential Access File
  - Reading Data from a File
  - Reading Data from a Sequential Access File
  - Random-Access Files
  - Creating a Randomly Accessed File
  - Writing Data Randomly to a Randomly Accessed File

- **Strings**
  - The Data Type char
  - Characters and Integers
  - Input and Output of Characters
  - Declaring Strings
  - Initializing Strings
  - Strings and Pointers
  - String Handling Functions (*string.h*)
The Data Type `char`

- Each character is stored in a machine in one byte (8 bits)
  - 1 byte is capable of storing $2^8$ or 256 distinct values.
- When a character is stored in a byte, the contents of that byte can be thought of as either a character or as an integer.
The Data Type `char`

- A character constant is written between single quotes.
  
  `a`
  `b`

- A declaration for a variable of type `char` is
  
  `char c;`

- Character variables can be initialized
  
  `char c1='A', c2='B', c3='*';`
In C, a character is considered to have the integer value corresponding to its ASCII encoding.

<table>
<thead>
<tr>
<th>lowercase</th>
<th>‘a’</th>
<th>‘b’</th>
<th>‘c’</th>
<th>…</th>
<th>‘z’</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASCII value</td>
<td>97</td>
<td>98</td>
<td>99</td>
<td>…</td>
<td>122</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>uppercase</th>
<th>‘A’</th>
<th>‘B’</th>
<th>‘C’</th>
<th>…</th>
<th>‘Z’</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASCII value</td>
<td>65</td>
<td>66</td>
<td>67</td>
<td>…</td>
<td>90</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>digit</th>
<th>‘0’</th>
<th>‘1’</th>
<th>‘2’</th>
<th>…</th>
<th>‘9’</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASCII value</td>
<td>48</td>
<td>49</td>
<td>50</td>
<td>…</td>
<td>57</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>other</th>
<th>‘&amp;’</th>
<th>‘*’</th>
<th>‘+’</th>
<th>…</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASCII value</td>
<td>38</td>
<td>42</td>
<td>43</td>
<td>…</td>
</tr>
</tbody>
</table>
Characters and Integers

- There is no relationship between the character ‘2’ (which has the ASCII value 50) and the constant number 2.
- ‘2’ is not 2.
- ‘A’ to ‘Z’ 65 to 90
- ‘a’ to ‘z’ 97 to 112

Examples:

- `printf(“%c’, ’a’);`
- `printf(“%c”, 97); have similar output.`
- `printf(“%d’, ’a’);`
- `printf(“%d”, 97); have also similar output.`
The Data Type `char`

- Some nonprinting and hard-to-print characters require an escape sequence.
- For example, the newline character is written as `\n` and it represents a single ASCII character.

<table>
<thead>
<tr>
<th>Name of character</th>
<th>Written in C</th>
<th>Integer Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>alert</td>
<td>\a</td>
<td>7</td>
</tr>
<tr>
<td>backslash</td>
<td>\</td>
<td>92</td>
</tr>
<tr>
<td>double quote</td>
<td>&quot;</td>
<td>34</td>
</tr>
<tr>
<td>horizontal tab</td>
<td>\t</td>
<td>9</td>
</tr>
</tbody>
</table>
Input and Output of Characters

- **getchar( )** reads a character from the keyboard.
  
  ```c
  c = getchar(); /* variable c contains the next character of input */
  ```

- **putchar(c)** prints a character to the screen.
  
  ```c
  putchar(c); /* prints the contents of the variable c as a character */
  ```
/* Illustrating the use of getchar( ) and putchar( ) */

#include <stdio.h>
int main (void)
{
    char c;
    while ((c=getchar()) != EOF) {
        putchar(c);
        putchar(c);
    }
}

EOF : It is control-d in Unix; control-z in DOS.
/* Capitalize lowercase letters and 
 * double space */
#include <stdio.h>

int main(void)
{
    int c;
    while (((c=getchar()) != EOF)){
        if (‘a’ <= c && c <= ‘z’)
            putchar(c+’A’-’a’); /*convert to uppercase*/
        else if (c == ‘\n’){
            putchar (‘\n’);
            putchar (‘\n’);
        }
        else putchar (c);
    }
}

Fundamentals of Strings and Characters

■ Characters
  ▪ Building blocks of programs
    ▪ Every program is a sequence of meaningfully grouped characters
  ▪ Character constant
    ▪ An int value represented as a character in single quotes
    ▪ 'z' represents the integer value of z

■ Strings
  ▪ Series of characters treated as a single unit
    ▪ Can include letters, digits and special characters (*, /, $)
  ▪ String literal (string constant) - written in double quotes
    ▪ "Hello"
  ▪ Strings are arrays of characters in C
    ▪ String is a pointer to first character
    ▪ Value of string is the address of first character
Strings

- A string constant such as “a string” is an array of characters.
- Each element of the array stores a character of the string.
- In its internal representation, the array is terminated with the null character ‘\0’ so that the end of the string can be found easily.
- Thus, the length of the array is defined one more than the number of characters between the double quotes.
Declaring Strings

```c
char myString[10];

myString[0] = 'H';
myString[1] = 'e';
myString[2] = 'l';
myString[3] = 'l';
myString[4] = 'o';
myString[5] = '\0';
```

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>'H'</td>
<td>'e'</td>
<td>'l'</td>
<td>'l'</td>
<td>'o'</td>
<td>'\0'</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>
Initializing Strings

- Character arrays can be initialized when they are declared:

```c
char name[5] = {'E','L','M','A','\0'};
char name[5] = "ELMA"; /* compiler automatically adds '\0' */
char name[] = "ELMA"; /* compiler calculates the size of the array */
```
Strings and Pointers

We can declare and initialize a string as a variable of type

```c
char *
char *color = "blue";
```

But the interpretation is different. “blue” is stored in memory as a string constant. The variable `color` is assigned the address of the constant string in memory.

If we declare it as:

```c
char c[] = "blue";
```
the array `c` contains the individual characters followed by the null character.
## String Handling Functions (*string.h*)

String handling library has functions to

- Manipulate string data
- Search strings
- Tokenize strings
- Determine string length

<table>
<thead>
<tr>
<th>Function prototype</th>
<th>Function description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>char *strcpy( char *s1, char *s2 )</code></td>
<td>Copies string <code>s2</code> into array <code>s1</code>. The value of <code>s1</code> is returned.</td>
</tr>
<tr>
<td><code>char *strncpy( char *s1, char *s2, int n )</code></td>
<td>Copies at most <code>n</code> characters of string <code>s2</code> into array <code>s1</code>. The value of <code>s1</code> is returned.</td>
</tr>
<tr>
<td><code>char *strcat( char *s1, char *s2 )</code></td>
<td>Appends string <code>s2</code> to array <code>s1</code>. The first character of <code>s2</code> overwrites the terminating null character of <code>s1</code>. The value of <code>s1</code> is returned.</td>
</tr>
<tr>
<td><code>char *strncat( char *s1, char *s2, int n )</code></td>
<td>Appends at most <code>n</code> characters of string <code>s2</code> to array <code>s1</code>. The first character of <code>s2</code> overwrites the terminating null character of <code>s1</code>. The value of <code>s1</code> is returned.</td>
</tr>
</tbody>
</table>
String Handling Functions (cont.)

- **unsigned strlen(char *s);**
  - A count of the number of characters before \0 is returned.

- **int strcmp(char *s1, char *s2);**
  - Compares string s1 to s2
  - Returns a negative number if s1 < s2, zero if s1 == s2 or a positive number if s1 > s2

- **int strncmp(char *s1, char *s2, int n );**
  - Compares up to n characters of string s1 to s2
  - Returns values as above
strcpy() and strncpy()  

- We cannot change the contents of a string by an assignment statement.

```c
char str[10];
str = "test";  /*Error! Attempt to change the base address*/
```

- Thus, we need to use string copy functions
  - `strcpy(str, "test");` /*contents of str changed*/
  - `strncpy(str, "testing", 5);`
    ```c
    str[5] = '\0';  /* str contains testi only */
    ```
  - `strcpy(str, "a very long string");` /*overflow of array boundary */
**strcat() and strncat()**

```c
char s[8]="abcd";
strcat(s,"FGH");  /* s keeps abcdFGH */

char t[10]="abcdef";
strcat(t,"GHIJKLM");  /* exceeds string length! */

strncat(t, "GHIJKLM",3);
t[9] = '\0';  /* t keeps abcdefGHI */
```
**strcmp() and strncmp()**

- We can compare characters with <, >, <= etc.
  
  e.g. ‘A’ < ‘B’

- But we cannot compare strings with the relational operators.
  
  e.g. `str1 < str2` will compare the memory addresses pointed by `str1` and `str2`.

- Therefore we need to use string comparison functions.

  ```c
  strcmp("abcd", "abcde")    returns a negative number
  strcmp("xyz", "xyz")      returns zero
  strcmp("xyz", "abc")      positive number
  strncmp("abcde", "abcDEF", 3) zero
  strncmp("abcde", "abcDEF", 4) positive number
  ```
Examples

char s1[] = "beautiful big sky country";
char s2[] = "how now brown cow";

<table>
<thead>
<tr>
<th>Expression</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>strlen(s1)</td>
<td>25</td>
</tr>
<tr>
<td>strlen(s2+8)</td>
<td>9</td>
</tr>
</tbody>
</table>

Statements | What is printed
--- | ---
printf("%s", s1+10); | big sky country
strcpy(s1+10, s2+8) |
strcat(s1, "s!"); |
printf("%s",s1); | beautiful brown cows!
#include <stdio.h>
#include <string.h>
#define LENGTH 20

/* A string is a palindrome if it reads the same backwards and forwards.
   e.g. abba, mum, radar. This programs checks whether a given string is
   palindrome or not. */

int isPalindrom(char s[]);  /* function prototype */
int main()
{
    char str[LENGTH];

    /* read the string */
    printf("Enter a string ");
    scanf("%s", str);

    /* Check if it is a palindrome. */
    if (isPalindrom(str))
        printf("%s is a palindrome.\n", str);
    else
        printf("%s is not a palindrome.\n", str);
}
int isPalindrome(char str[])
{
    int i, j, flag;

    i = 0;          /* index of the first character */
    j = strlen(str) - 1; /* index of the last character */
    flag = 1;       /* assume it is a palindrome */

    while ((i<j) && flag){
        /* compare the ith and jth. characters */

        if (str[i] != str[j])
            flag = 0; /* if not same then string cannot be a palindrome. */

        else {
            i++;
            j--;
        }
        /* advance to next characters */
    }

    return flag;
}