CIS 330:

Unix and C/C++

Lecture 6:
File I/O, Redirection, Pipes

AND

Enums, Structs, and Unions

April 17th, 2015  Hank Childs, University of Oregon
New topic:
SCRAMBLE FOR 4A
Comments on 4A

• p1.c: only 2 errors, not 3
  – And only 1 error until 1pm today!
    • (redownload)

• What to turn in:
  – sentence / partial sentence
    • what kind of memory error
    • where it occurred
pointers to pointers

• int **p = malloc(2*sizeof(int *));
• int *p2 = malloc(2*sizeof(int));
There have been various extensions to ASCII ... 
now more than 128 characters
Many special characters are handled outside this convention
Unix and Windows difference

• Unix:
  – “\n”: goes to next line, and sets cursor to far left

• Windows:
  – “\n”: goes to next line (cursor does not go to left)
  – “\m”: sets cursor to far left

• Text files written in Windows often don’t run well on Unix, and vice-versa
  – There are more differences than just newlines

vi: “set ff=unix” solves this
signed vs unsigned chars

• signed char ("char"):  
  – valid values: -128 to 127  
  – size: 1 byte  
  – used to represent characters with ASCII  
  • values -128 to -1 are not valid

• unsigned char:  
  – valid values: 0 to 255  
  – size: 1 byte  
  – used to represent data
character strings

• A character “string” is:
  – an array of type “char”
  – that is terminated by the NULL character

• Example:
  char str[12] = “hello world”;
  – str[11] = ‘\0’ (the compiler did this automatically)

• The C library has multiple functions for handling strings
Character strings example

```c
#include <stdio.h>

int main()
{
    char str[12] = "hello world";
    char *str2 = str+6;

    printf("str is \"%s\" and str2 is \"%s\"\n", str, str2);

    str[5] = '\0';

    printf("Now str is \"%s\" and str2 is \"%s\"\n", str, str2);
}
```

```
128-223-223-72-wireless:330 hank$ cat string.c
128-223-223-72-wireless:330 hank$ gcc string.c
128-223-223-72-wireless:330 hank$ ./a.out
str is "hello world" and str2 is "world"
Now str is "hello" and str2 is "world"
```
Useful C library string functions

• `strcpy`: string copy
• `strncpy`: string copy, but just first N characters
• `strlen`: length of a string

```c
#include <string.h>
#include <stdio.h>

int main()
{
    char str[12] = "hello world";
    char str2[6], str3[7];
    strcpy(str2, str+strlen("hello "));
    strncpy(str3, str, strlen("hello "));
    printf("%s,%s\n", str2, str3);
}
```
Useful C library string functions

- `strcpy`: string copy
- `strncpy`: string copy, but just first N characters
- `strlen`: length of a string

```c
#include <string.h>
#include <stdio.h>

int main()
{
    char str[12] = "hello world";
    char str2[7], str3[6];
    strcpy(str2, str+strlen("hello "));
    strncpy(str3, str, strlen("hello "));
    printf("%s,%s\n", str2, str3);
}
```

What happened here?

```
128-223-223-72-wireless:330 hank$ cat strcpy.c
#include <string.h>
#include <stdio.h>

int main()
{
    char str[12] = "hello world";
    char str2[7], str3[6];
    strcpy(str2, str+strlen("hello "));
    strncpy(str3, str, strlen("hello "));
    printf("%s,%s\n", str2, str3);
}
128-223-223-72-wireless:330 hank$ gcc strcpy.c
128-223-223-72-wireless:330 hank$ ./a.out
world,hello world
```
More useful C library string functions

<table>
<thead>
<tr>
<th>Functions</th>
<th>Searching:</th>
<th>Other:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copying:</td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>memcpy</code></td>
<td>Locate character in block of memory (function)</td>
<td><code>memset</code> Fill block of memory (function)</td>
</tr>
<tr>
<td><code>memmove</code></td>
<td>Locate first occurrence of character in string (function)</td>
<td><code>strerror</code> Get pointer to error message string (function)</td>
</tr>
<tr>
<td><code>strcpy</code></td>
<td>Get span until character in string (function)</td>
<td><code>strerror</code> Get pointer to error message string (function)</td>
</tr>
<tr>
<td><code>strncpy</code></td>
<td>Locate characters in string (function)</td>
<td><code>stab</code> Get span of character set in string (function)</td>
</tr>
<tr>
<td>Concatenation:</td>
<td></td>
<td><code>strlen</code> Get string length (function)</td>
</tr>
<tr>
<td><code>strcat</code></td>
<td>Locate substring (function)</td>
<td></td>
</tr>
<tr>
<td><code>strncat</code></td>
<td>Split string into tokens (function)</td>
<td></td>
</tr>
<tr>
<td>Comparison:</td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>memcmp</code></td>
<td>Compare two blocks of memory (function)</td>
<td></td>
</tr>
<tr>
<td><code>strcmp</code></td>
<td>Compare two strings (function)</td>
<td></td>
</tr>
<tr>
<td><code>strcoll</code></td>
<td>Compare two strings using locale (function)</td>
<td></td>
</tr>
<tr>
<td><code>strnccmp</code></td>
<td>Compare characters of two strings (function)</td>
<td></td>
</tr>
<tr>
<td><code>strxfrm</code></td>
<td>Transform string using locale (function)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Macros</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><code>NULL</code></td>
<td>Null pointer (macro)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Types</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><code>size_t</code></td>
<td>Unsigned integral type (type)</td>
</tr>
</tbody>
</table>
memcpy

memcpy(3) BSD Library Functions Manual MEMCPY(3)

NAME
memcpy -- copy memory area

LIBRARY
Standard C Library (libc, -lc)

SYNOPSIS
#include <string.h>

void *
memcpy(void *restrict dst, const void *restrict src, size_t n);

DESCRIPTION
The memcpy() function copies n bytes from memory area src to memory area
dst. If dst and src overlap, behavior is undefined. Applications in
which dst and src might overlap should use memmove(3) instead.

RETURN VALUES
The memcpy() function returns the original value of dst.

I mostly use C++, and I still use memcpy all the time
sscanf

• like printf, but it parses from a string

\[
\text{sscanf(str, } \"%s\n%d \n%d\n\n\", \text{ magicNum, } \\
&width, &height, &maxval)\);
\]

on:

\[
\text{str=“P6\n1000 1000\n255\n”};
\]
gives:

\[
\text{magicNum = “P6”, width = 1000, } \\
\text{height = 1000, maxval = 255}
\]
int val = (X < 2 ? X : 2);

\[
\begin{align*}
\text{if } (X < 2) & \quad \{ \\
\text{val} & = X;
\} \\
\text{else} & \quad \{ \\
\text{val} & = 2;
\}
\end{align*}
\]
Outline

• File I/O
• Project 2E
• Redirection
• Pipes
• Enums
• Structs
• Unions
• Bonus Slides
Outline

- File I/O
- Project 2E
- Redirection
- Pipes
- Enums
- Structs
- Unions
- Bonus Slides
File I/O: streams and file descriptors

• Two ways to access files:
  – File descriptors:
    • Lower level interface to files and devices
      – Provides controls to specific devices
    • Type: small integers (typically 20 total)
  – Streams:
    • Higher level interface to files and devices
      – Provides uniform interface; easy to deal with, but less powerful
    • Type: FILE *

Streams are more portable, and more accessible to beginning programmers. (I teach streams here.)
File I/O

• Process for reading or writing
  – Open a file
    • Tells Unix you intend to do file I/O
    • Function returns a “FILE *
      – Used to identify the file from this point forward
    • Checks to see if permissions are valid
  – Read from the file / write to the file
  – Close the file
Opening a file

• FILE *handle = fopen(filename, mode);

The argument mode points to a string beginning with one of the following sequences (Additional characters may follow these sequences.):

```
``r''  Open text file for reading. The stream is positioned at the beginning of the file.
``r+''  Open for reading and writing. The stream is positioned at the beginning of the file.
``a+''  Open for reading and writing. The file is created if it does not exist, otherwise it is truncated. The stream is positioned at the beginning of the file.
``a''  Open for writing. The file is created and opened for writing. Subsequent writes to the file will always end up at the then current end of file, irrespective of any intervening fseek(3) or similar.
``w''  Create or overwrite the file. The file is opened for writing. All previous contents are lost.
``w+''  Create or overwrite the file. The file is opened for writing and reading. All previous contents are lost.
``x''  Create and open the file for writing. If the file already exists, fopen() fails. (Note: on some systems, if the file does not exist, `x' is also treated as `w'.)
``x+''  Create and open the file for reading and writing.
``b''  The file is treated as binary.
``b+''  The file is treated as binary and opened for reading and writing.
``"('\''r''|\''a''|\''w''|\''x''|\''b''|\''\n'')"  Open for reading, writing, or both. The mode character can be any of \''r'' (read), \''w'' (write), \''a'' (append), or \''x'' (create and append). A trailing `\n' specifies text mode.
``"(\''r'','\''w''|\''a''|\''x''|\''\n'')"  Open for reading, writing, or both. The first mode character determines the mode; the second character can be one of `\''r''', `\''w''', `\''a''', or `\''x'''. A trailing `\n' specifies text mode.
``b\"  The file is treated as binary.
``b\"+''  The file is treated as binary and opened for reading and writing.
``"(\''r'','\''w''|\''a''|\''x''|\''\n'')b\"  Open for reading, writing, or both. The first character determines the mode; the second character can be one of `\''r''', `\''w''', `\''a''', or `\''x'''. A trailing `\n' specifies text mode.
``b\"\"  The file is treated as binary.
``b\"\"+''  The file is treated as binary and opened for reading and writing.
``"(\''r'','\''w''|\''a''|\''x''|\''\n'')b\"\"  Open for reading, writing, or both. The first character determines the mode; the second character can be one of `\''r''', `\''w''', `\''a''', or `\''x'''. A trailing `\n' specifies text mode.
``"r\"  Open text file for reading. The stream is positioned at the beginning of the file.
``r\"+''  Open for reading and writing. The stream is positioned at the beginning of the file.
``a\"  Open for writing. The file is created and opened for writing. All previous contents are lost.
``a\"+''  Create and open the file for reading and writing.
``w\"  Create or overwrite the file. The file is opened for writing. All previous contents are lost.
``w\"+''  Create or overwrite the file. The file is opened for reading and writing. All previous contents are lost.
``x\"  Create and open the file for writing. If the file already exists, fopen() fails. (Note: on some systems, if the file does not exist, `x' is also treated as `w'.)
``x\"+''  Create and open the file for reading and writing.
``"(\''r'','\''w''|\''a''|\''x''|\''\n'')b\"  Open for reading, writing, or both. The first character determines the mode; the second character can be one of `\''r''', `\''w''', `\''a''', or `\''x'''. A trailing `\n' specifies text mode.
``b\"\"  The file is treated as binary.
``b\"\"+''  The file is treated as binary and opened for reading and writing.
``"(\''r'','\''w''|\''a''|\''x''|\''\n'')b\"\"  Open for reading, writing, or both. The first character determines the mode; the second character can be one of `\''r''', `\''w''', `\''a''', or `\''x'''. A trailing `\n' specifies text mode.
``Note: #include <stdio.h>
FREAD(3)

NAME
fread, fwrite — binary stream input/output

LIBRARY
Standard C Library (libc, -lc)

SYNOPSIS
#include <stdio.h>

size_t
fread(void *restrict ptr, size_t size, size_t nitems, FILE *restrict stream);

size_t
fwrite(const void *restrict ptr, size_t size, size_t nitems,
      FILE *restrict stream);

DESCRIPTION
The function fread() reads nitems objects, each size bytes long, from the stream
pointed to by stream, storing them at the location given by ptr.

The function fwrite() writes nitems objects, each size bytes long, to the stream
pointed to by stream, obtaining them from the location given by ptr.

RETURN VALUES
The functions fread() and fwrite() advance the file position indicator for the
stream by the number of bytes read or written. They return the number of objects
read or written. If an error occurs, or the end-of-file is reached, the return
value is a short object count (or zero).
Example

C02LN00GFD58:330 hank$ cat rw.c
#include <stdio.h>
#include <string.h>

int main(int argc, char *argv[])
{
    char *hello = "hello world: file edition\n";
    FILE *f = fopen("330", "w");
    fwrite(hello, sizeof(char), strlen(hello), f);
    fclose(f);
}
C02LN00GFD58:330 hank$ gcc rw.c
C02LN00GFD58:330 hank$ .a.out
C02LN00GFD58:330 hank$ cat 330
hello world: file edition
fseek

int fseek(FILE *stream, long offset, int whence);

The `fseek()` function sets the file position indicator for the stream pointed to by `stream`. The new position, measured in bytes, is obtained by adding `offset` bytes to the position specified by `whence`. If `whence` is set to SEEK_SET, SEEK_CUR, or SEEK_END, the offset is relative to the start of the file, the current position indicator, or end-of-file, respectively. A successful call to the `fseek()` function clears the end-of-file indicator for the stream and undoes any effects of the `ungetc(3)` and `ungetwc(3)` functions on the same stream.
The `ftell()` function obtains the current value of the file position indicator for the stream pointed to by `stream`. 
We have everything we need to make a copy command...

• fopen
• fread
• fwrite
• fseek
• ftell

Can we do this together as a class?
argc & argv

• two arguments to every C program
• argc: how many command line arguments
• argv: an array containing each of the arguments
• “./a.out hank childs”
• → argc == 3
```c
#include <stdio.h>
#include <string.h>
#include <stdlib.h>

int main(int argc, char *argv[])
{
    FILE *f_in, *f_out;
    int buff_size;
    char *buffer;

    if (argc != 3)
    {
        printf("Usage: %s <file1> <file2>\n", argv[0]);
        exit(EXIT_FAILURE);
    }

    f_in = fopen(argv[1], "r");
    fseek(f_in, 0, SEEK_END);
    buff_size = ftell(f_in);
    fseek(f_in, 0, SEEK_SET);

    buffer = malloc(buff_size);
    fread(buffer, sizeof(char), buff_size, f_in);

    printf("Copying %d bytes from %s to %s\n", buff_size, argv[1], argv[2]);

    f_out = fopen(argv[2], "w");
    fwrite(buffer, sizeof(char), buff_size, f_out);

    fclose(f_in);
    fclose(f_out);

    return 0;
}
```
Return values in shells

C02LN00GFD58:330 hank$ ./a.out copy.c copy2.c
Copying 697 bytes from copy.c to copy2.c
C02LN00GFD58:330 hank$ echo $?
0
C02LN00GFD58:330 hank$ ./a.out copy.c
Usage: ./a.out <file1> <file2>
C02LN00GFD58:330 hank$ echo $?
1

$? is the return value of the last executed command
Printing to terminal and reading from terminal

• In Unix, printing to terminal and reading from terminal is done with file I/O

• Keyboard and screen are files in the file system!
  – (at least they were ...)

Standard Streams

• Wikipedia: “preconnected input and output channels between a computer program and its environment (typically a text terminal) when it begins execution”

• Three standard streams:
  – stdin (standard input)
  – stdout (standard output)
  – stderr (standard error)

What mechanisms in C allow you to access standard streams?
printf

• Print to stdout
  – printf(“hello world\n”);
  – printf(“Integers are like this %d\n”, 6);
  – printf(“Two floats: %f, %f”, 3.5, 7.0);
fprintf

• Just like printf, but to streams
  • `fprintf(stdout, "helloworld\n");`
    – → same as printf
  • `fprintf(stderr, "helloworld\n");`
    – prints to “standard error”
  • `fprintf(f_out, "helloworld\n");`
    – prints to the file pointed to by FILE *f_out.
Outline

• File I/O
• Project 2E
• Redirection
• Pipes
• Enums
• Structs
• Unions
• Bonus Slides
CIS 330: Project #2E
Assigned: April 17th, 2015
Due April 23rd, 2015
(which means submitted by 6am on April 24th, 2015)
Worth 4% of your grade

Assignment: Write a program that reads the file “2E_binary_file”. This file contains a
two-dimensional array of integers, that is 10x10. You are to read in the 5x5 bottom
left corner of the array. That is, the values 0-4, 10-14, 20-24, 30-34, and 40-44. You
may only read 25 integers total. Do not read all 100 and throw some out. You will
then write out the new 5x5 array. Please write this as strings, one integer per line
(25 lines total). You should be able to “cat” the file afterwards and see the values.

Use Unix file streams for this project (i.e., fopen, fread, fseek, fprintf). Your program
will be checked for good programming practices. (Close your file streams, use
memory correctly, etc. I am not referring to style, variable initialization, etc.)

Also, add support for command line arguments (argc and argv).

Your program should run as:
./<prog_name> <input_name> <output_name>

(The input_name will be 2E_binary_file, unless you change it.)

Finally, note that I am handing you a binary file. I think we are all little endian, and
so it will be fine. But, if it is big endian, then we will have a problem. You can check
if it is little endian by printing the first two values of the file. They should be “0” and
“1”.

Please submit a tarball with (1) a Makefile (should be simple), (2) your source code,
and (3) the output ASCII file from running your program, with the name
“ASCII_output”.

Outline

• File I/O
• Project 2E
• Redirection
• Pipes
• Enums
• Structs
• Unions
• Bonus Slides
Unix shells allows you to manipulate standard streams.

- “>” redirect output of program to a file
- Example:
  - `ls > output`
  - `echo “this is a file” > output2`
  - `cat file1 file2 > file3`
Unix shells allows you to manipulate standard streams.

• “<” redirect file to input of program

• Example:
  – python < myscript.py
    • Note: python quits when it reads a special character called EOF (End of File)
    • You can type this character by typing Ctrl-D
    • This is why Python quits when you type Ctrl-D
      – (many other programs too)
Unix shells allows you to manipulate standard streams.

- “>>” concatenate output of program to end of existing file
  - (or create file if it doesn’t exist)
- Example:
  - echo “I am starting the file” > file1
  - echo “I am adding to the file” >> file1
  - cat file1

    I am starting the file
    I am adding to the file
What’s happening here?

```bash
C02LN00GFD58:330 hank$ mkdir tmp
C02LN00GFD58:330 hank$ cd tmp
C02LN00GFD58:tmp hank$ touch f1
C02LN00GFD58:tmp hank$ ls f1 f2 > out
ls: f2: No such file or directory
C02LN00GFD58:tmp hank$ cat out
f1
```

ls is outputting its error messages to stderr
Redirecting stderr in a shell

C02LN00GFD58:Documents hank$ cd ~/330
C02LN00GFD58:330 hank$ mkdir tmp
C02LN00GFD58:330 hank$ cd tmp
C02LN00GFD58:tmp hank$ touch f1
C02LN00GFD58:tmp hank$ ls f1 f2 > out
ls: f2: No such file or directory
C02LN00GFD58:tmp hank$ cat out
f1
C02LN00GFD58:tmp hank$ ls f1 f2 > out 2>out_error
C02LN00GFD58:tmp hank$ cat out_error
ls: f2: No such file or directory
Redirecting stderr to stdout

```bash
C02LN00GFD58:330 hank$ mkdir tmp
C02LN00GFD58:330 hank$ cd tmp
C02LN00GFD58:tmp hank$ touch f1
C02LN00GFD58:tmp hank$ ls f1 f2 > out
ls: f2: No such file or directory
C02LN00GFD58:tmp hank$ cat out f1
C02LN00GFD58:tmp hank$ ls f1 f2 > out 2>out_error
C02LN00GFD58:tmp hank$ cat out_error
ls: f2: No such file or directory
C02LN00GFD58:tmp hank$ ls f1 f2 > out 2>&1
C02LN00GFD58:tmp hank$ cat out
ls: f2: No such file or directory
f1
```

Convenient when you want both to go to the same stream
Outline

- File I/O
- Project 2E
- Redirection
- Pipes
- Enums
- Structs
- Unions
- Bonus Slides
c functions: fork and pipe

• fork: duplicates current program into a separate instance
  – Two running programs!
  – Only differentiated by return value of fork (which is original and which is new)

• pipe: mechanism for connecting file descriptors between two forked programs

Through fork and pipe, you can connect two running programs. One writes to a file descriptor, and the other reads the output from its file descriptor

Only used on special occasions. (And one of those occasions is with the shell.)
pipes in Unix shells

- represented with “|”
- output of one program becomes input to another program
Very useful programs

- grep: keep lines that match pattern, discard lines that don’t match pattern

```bash
C02LN00GFD58:Documents hank$ ls -l | grep ppt
-rw-r--r--@ 1 hank staff 3278589 Apr 5 11:40 CIS330_Lec2.pptx
-rw-r--r--@ 1 hank staff 2220104 Apr 8 20:57 CIS330_Lec3.pptx
-rw-r--r-- 1 hank staff 3899863 Jan 21 09:26 CIS610_lec2.pptx
-rw-r--r-- 1 hank staff 4629257 Jan 30 10:24 CIS610_lec3.pptx
-rw-r--r-- 1 hank staff 21382185 Mar 25 12:40 CIS_colloquium2013.pptx
-rw-r--r-- 1 hank staff 21382185 Jan 7 12:21 CIS_colloquium_2013.pptx
-rw-r--r--@ 1 hank staff 2172179 Dec 20 15:24 ICS_results.pptx
-rw-r--r--@ 1 hank staff 4841050 Nov 13 10:10 MBTI.pptx
-rw-r--r--@ 1 hank staff 2031749 Apr 5 16:20 SC14_flow.pptx
-rw-r--r-- 1 hank staff 17972476 Mar 25 12:43 VMV_2013.pptx
-rw-r--r--@ 1 hank staff 98149068 Apr 1 10:25 aachen.pptx
-rw-r--r-- 1 hank staff 9815146 Feb 24 07:00 childs_poster_SDAV_AHM_2014.pptx
-rw-r--r--@ 1 hank staff 592243 Feb 26 04:09 childs_sdav_slides.pptx
-rw-r--r--@ 1 hank staff 15765504 Feb 13 14:57 cig_exascale.ppt
-rw-r--r--@ 1 hank staff 16699392 Jan 7 12:14 cis610_Lec1.ppt
-rw-r--r-- 1 hank staff 3159872 Jan 7 11:15 egpgv_cgf.pptx
-rw-r--r--@ 1 hank staff 15767552 Mar 23 02:48 eu_regional_school.ppt
-rw-r--r--@ 1 hank staff 35099136 Mar 25 09:42 eu_regional_school_part1.ppt
-rw-r--r--@ 1 hank staff 10775552 Mar 25 04:49 eu_regional_school_part1B.ppt
-rw-r--r--@ 1 hank staff 72966144 Mar 26 08:43 eu_regional_school_part2.ppt
-rw-r--r-- 1 hank staff 7571317 Mar 25 12:53 ilm_booth_talk.pptx
```
Very useful programs

• **sed**: replace pattern 1 with pattern 2
  – `sed s/pattern1/pattern2/g`
    • `s` means substitute
    • `g` means “global” ... every instance on the line

*sed is also available in “vi”*
: `%s/pattern1/pattern2/g` (% means all lines)
: `103,133s/p1/p2/g` (lines 103-133)
Outline

• File I/O
• Project 2E
• Redirection
• Pipes
• Enums
• Structs
• Unions
• Bonus Slides
Enums

• Enums make your own type
  – Type is “list of key words”

• Enums are useful for code clarity
  – Always possible to do the same thing with integers

• Be careful with enums
  – ... you can “contaminate” a bunch of useful words
enum example

C keyword "enum" – means enum definition is coming

```c
enum StudentType {
    HighSchool,
    Freshman,
    Sophomore,
    Junior,
    Senior,
    GradStudent
};
```

This enum contains 6 different student types

semi-colon!!!
enum example

```c
int AverageAge(enum StudentType st) {
    if (st == HighSchool)
        return 16;
    if (st == Freshman)
        return 18;
    if (st == Sophomore)
        return 19;
    if (st == Junior)
        return 21;
    if (st == Senior)
        return 23;
    if (st == GradStudent)
        return 26;

    return -1;
}
```
enums translate to integers ... and you can set their range

```
#include <stdio.h>

enum StudentType
{
    HighSchool = 105,
    Freshman,
    Sophomore,
    Junior,
    Senior,
    GradStudent
};

int main()
{
    printf("HighSchool = %d, GradStudent = %d\n", HighSchool, GradStudent);
}
```

```
128-223-223-72-wireless:330 hank$ cat enum2.c
128-223-223-72-wireless:330 hank$ gcc enum2.c
128-223-223-72-wireless:330 hank$ ./a.out
HighSchool = 105, GradStudent = 110
```
But enums can be easier to maintain than integers

```c
enum StudentType {
    HighSchool, Freshman, Sophomore, Junior, Senior, PostBacc, GradStudent
};

int AverageAge(enum StudentType st) {
    if (st == HighSchool) return 16;
    if (st == Freshman) return 18;
    if (st == Sophomore) return 19;
    if (st == Junior) return 21;
    if (st == Senior) return 23;
    if (st == PostBacc) return 24;
    if (st == GradStudent) return 26;

    return -1;
}
```

If you had used integers, then this is a bigger change and likely to lead to bugs.
Outline

• File I/O
• Project 2E
• Redirection
• Pipes
• Enums
• Structs
• Unions
• Bonus Slides
Data types

- float
- double
- int
- char
- unsigned char

All of these are simple data types
Structs: a complex data type

• Construct that defines a group of variables
  – Variables must be grouped together in contiguous memory
• Also makes accessing variables easier ... they are all part of the same grouping (the struct)
C keyword “struct” – means struct definition is coming

struct syntax

```
struct Ray
{
    double origin[3];
    double direction[3];
};

int main()
{
    struct Ray r;
    r.origin[0] = 0;
    r.origin[1] = 0;
    r.origin[2] = 0;
    r.direction[0] = 1;
    r.direction[1] = 0;
    r.direction[2] = 0;
}
```

This struct contains 6 doubles, meaning it is 48 bytes

Declaring an instance

“.” accesses data members for a struct
Nested structs

```c
struct Origin
{
    double originX;
    double originY;
    double originZ;
};

struct Direction
{
    double directionX;
    double directionY;
    double directionZ;
};

struct Ray
{
    struct Origin ori;
    struct Direction dir;
};

int main()
{
    struct Ray r;
    r.ori.originX = 0;
    r.ori.originY = 0;
    r.ori.originZ = 0;
    r.dir.directionX = 0;
    r.dir.directionY = 0;
    r.dir.directionZ = 0;
}
```

accesses dir part of Ray
accesses directionZ part of Direction (part of Ray)
typedef

• typedef: tell compiler you want to define a new type

```c
struct Ray
{
  double origin[3];
  double direction[3];
};

int main()
{
  struct Ray r:
  r.origin[0] = 0;
  r.origin[1] = 0;
  r.origin[2] = 0;
  r.direction[0] = 1;
  r.direction[1] = 0;
  r.direction[2] = 0;
}
```

```c
typedef struct
{
  double origin[3];
  double direction[3];
} Ray;

int main()
{
  Ray r;
  r.origin[0] = 0;
  r.origin[1] = 0;
  r.origin[2] = 0;
  r.direction[0] = 1;
  r.direction[1] = 0;
  r.direction[2] = 0;
}
```

saves you from having to type "struct" every time you declare a struct.
Other uses for typedef

• Declare a new type for code clarity
  – typedef int MilesPerHour;
    • Makes a new type called MilesPerHour.
    • MilesPerHour works exactly like an int.

• Also used for enums & unions
  – same trick as for structs ... typedef saves you a word
Outline

• File I/O
• Project 2E
• Redirection
• Pipes
• Enums
• Structs
• Unions
• Bonus Slides
Unions

• Union: special data type
  – store many different memory types in one memory location

    typedef union
    {
      float x;
      int y;
      char z[4];
    } cis330_union;

When dealing with this union, you can treat it as a float, as an int, or as 4 characters.

This data structure has 4 bytes
Why are unions useful?

```c
#include <stdio.h>

typedef union
{
    float x;
    int   y;
    char  z[4];
} cis330_union;

int main()
{
    cis330_union u;
    u.x = 3.5;  /* u.x is 3.5,  u.y and u.z are not meaningful */
    u.y = 3;    /* u.y is 3,   now u.x and u.z are not meaningful */
    printf("As u.x = %f, as u.y = %d\n", u.x, u.y);
}
```

As u.x = 0.000000, as u.y = 3
Unions Example

typedef struct
{
    int firstNum;
    char letters[3];
    int endNums[3];
} CA_LICENSE_PLATE;

typedef struct
{
    char letters[3];
    int nums[3];
} OR_LICENSE_PLATE;

typedef struct
{
    int nums[6];
} WY_LICENSE_PLATE;

typedef union
{
    CA_LICENSE_PLATE ca;
    OR_LICENSE_PLATE or;
    WY_LICENSE_PLATE wy;
} LicensePlate;
typedef enum
{
    CA,
    OR,
    WY
} US_State;

typedef struct
{
    char *carMake;
    char *carModel;
    US_State state;
    LicensePlate lp;
} CarInfo;

int main()
{
    CarInfo c;
    c.carMake = "Chevrolet";
    c.carModel = "Camaro";
    c.state = OR;
    c.lp.or.letters[0] = 'X';
    c.lp.or.letters[1] = 'S';
    c.lp.or.letters[2] = 'Z';
    c.lp.or.nums[0] = 0;
    c.lp.or.nums[1] = 7;
    c.lp.or.nums[2] = 5;
}

typedef struct
{
    int firstNum;
    char letters[3];
    int endNums[3];
} CA_LICENSE_PLATE;

typedef struct
{
    char letters[3];
    int nums[3];
} OR_LICENSE_PLATE;

typedef struct
{
    int nums[6];
} WY_LICENSE_PLATE;

typedef union
{
    CA_LICENSE_PLATE ca;
    OR_LICENSE_PLATE or;
    WY_LICENSE_PLATE wy;
} LicensePlate;
Why are Unions useful?

• Allows you to represent multiple data types simultaneously
  – But only if you know you want exactly one of them

• Benefit is space efficiency, which leads to performance efficiency

Unions are also useful for abstracting type. We will re-visit this when we talk about C++’s templates.
Outline

• File I/O
• Project 2E
• Redirection
• Pipes
• Enums
• Structs
• Unions
• Bonus Slides
Problem with C...

C02LN00GFD58:330 hank$ cat doubler.c
float doubler(float f) { return 2*f; }
C02LN00GFD58:330 hank$ gcc -c doubler.c
C02LN00GFD58:330 hank$ cat doubler_example.c
#include <stdio.h>

int doubler(int);

int main()
{
    printf("Doubler of 10 is %d\n", doubler(10));
}
C02LN00GFD58:330 hank$ gcc -c doubler_example.c
C02LN00GFD58:330 hank$ gcc -o doubler_example doubler.o doubler_example.o
C02LN00GFD58:330 hank$ ./doubler_example
Doubler of 10 is 2
Problem with C...

C02LN00GFD58:330  hank$ nm doubler.o
00000000000000048  s  EH_frame0
000000000000000000  T  _doubler
00000000000000060  S  _doubler.eh
C02LN00GFD58:330  hank$ nm doubler
  doubler.c          doubler_example  doubler_example.o
doubler.o           doubler_example.c  doubler_user.o
C02LN00GFD58:330  hank$ nm doubler_example.o
00000000000000068  s  EH_frame0
00000000000000032  s  L_.str
U  _doubler
000000000000000000  T  _main
00000000000000080  S  _main.eh
U  _printf

No checking of type...
Problem is fixed with C++...

C02LN00GFD58:330 hank$ cat doubler.c
float doubler(float f) { return 2*f; }
C02LN00GFD58:330 hank$ g++ -c doubler.c
clang: warning: treating 'c' input as 'c++' when in C++ mode, this behavior is deprecated
C02LN00GFD58:330 hank$ cat doubler_example.c
#include <stdio.h>

int doubler(int);

int main()
{
    printf("Doubler of 10 is %d\n", doubler(10));
}
C02LN00GFD58:330 hank$ g++ -c doubler_example.c
clang: warning: treating 'c' input as 'c++' when in C++ mode, this behavior is deprecated
C02LN00GFD58:330 hank$ g++ -o doubler_example doubler_example.o doubler.o
Undefined symbols for architecture x86_64:
  "doubler(int)", referenced from:
    _main in doubler_example.o
ld: symbol(s) not found for architecture x86_64
clang: error: linker command failed with exit code 1 (use -v to see invocation)
C02LN00GFD58:330 hank$
Problem is fixed with C++...

```plaintext
C02LN00GFD58:330 hank$ nm doubler.o
000000000000000048 s EH_frame0
000000000000000000 T __Z7doublerf
000000000000000060 S __Z7doublerf.eh
C02LN00GFD58:330 hank$ nm doubler_example.o
000000000000000068 s EH_frame0
000000000000000032 s L_.str
U __Z7doubleri
000000000000000000 T _main
000000000000000080 S _main.eh
U _printf
C02LN00GFD58:330 hank$ nm doubler
C02LN00GFD58:330 hank$ nm doubler_example
doubler.c   doubler_example.c
doubler.o   doubler_example.o
c
C02LN00GFD58:330 hank$ nm doubler_example
000000000000000068 s EH_frame0
000000000000000032 s L_.str
U _doubler
000000000000000000 T _main
000000000000000080 S _main.eh
U _printf
```
Mangling

• Mangling refers to combing information about the return type and arguments and “mangling” it with function name.
  – Way of ensuring that you don’t mix up functions.
• Causes problems with compiler mismatches
  – C++ compilers haven’t standardized.
  – Can’t take library from icpc and combine it with g++. 
C++ will let you overload functions with different types

```c
float doubler(float f) { return 2*f; }
int doubler(int f) { return 2*f; }
```

```
C02LN00GFD58:330 hank$ cat t.c
float doubler(float f) { return 2*f; }
int doubler(int f) { return 2*f; }
C02LN00GFD58:330 hank$ gcc -c t.c
```

```
t.c:2:5: error: conflicting types for 'doubler'
int doubler(int f) { return 2*f; }
^
```

```
t.c:1:7: note: previous definition is here
float doubler(float f) { return 2*f; }
^
```

1 error generated.
```
C02LN00GFD58:330 hank$ g++ -c t.C
C02LN00GFD58:330 hank$
```
C++ also gives you access to mangling via “namespaces”

```c
#include <stdio.h>

namespace CIS330
{
    int GetNumberOfStudents(void) { return 56; };
}

namespace CIS610
{
    int GetNumberOfStudents(void) { return 9; };
}

int main()
{
    printf("Number of students in 330 is %d, but in 610 was %d\n",
            CIS330::GetNumberOfStudents(),
            CIS610::GetNumberOfStudents());
}
```

Functions or variables within a namespace are accessed with “::”
C++ also gives you access to mangling via “namespaces”

The “using” keyword makes all functions and variables from a namespace available without needing “::”. And you can still access other namespaces.

```cpp
namespace CIS610
{
    int GetNumberOfStudents(void) { return 9; }
}

using namespace CIS330;

int main()
{
    printf("Number of students in 330 is %d, but in 610 was %d\n",
           CIS610::GetNumberOfStudents(),
           CIS610::GetNumberOfStudents());
}
```

```bash
C02LN00GFD58:330 hank$ g++ cis330.C
C02LN00GFD58:330 hank$ ./a.out
Number of students in 330 is 56, but in 610 was 9
```
Outline

• Announcements/Review
• Project 1C
• File I/O, Redirection, Pipes
• Project 2A
• Symbols and mangling (if time)
• Killing and suspending jobs (if time)
• Web pages (if time)
• If you want to kill a program that your shell is currently running, then it is easy:
  – press Control-C

• What happens?
  – OS will send a signal to the program.
  – Programs can catch the signal and ignore, but most programs just die.
Killing programs, part 2

• What if a program is running, but it is not running as part of your shell? How to kill it then?

  – “ps”: process status

```
  hank$ ps
    PID  TTY     TIME CMD
  3518 ttys000  0:00.18 -bash
  25240 ttys002  0:00.53 -bash
  5708 ttys003  0:00.42 -bash
  29442 ttys003  0:00.02 vi qual         
  6080 ttys004  0:00.08 -bash
  30107 ttys004  0:00.01 vi /Users/hank/UO_SVN/CV/hrc_cv_faculty.bib
  9028 ttys005  0:00.04 -bash
  25964 ttys006  0:00.08 -bash
  27120 ttys010  0:00.09 -bash
  27329 ttys010  0:00.01 vi temp         
  27217 ttys012  0:00.01 -bash
```

PID = Process ID
Every program has a Process ID

Once you know the PID, you can use the Unix command “kill”: 
kill 29442
Killing programs, part 3

- “kill -9 <PID>” → the “I really mean it” signal

The following are some of the more common signals you might encounter and want to use in your programs:

<table>
<thead>
<tr>
<th>Signal Name</th>
<th>Signal Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIGHUP</td>
<td>1</td>
<td>Hang up detected on controlling terminal or death of controlling process</td>
</tr>
<tr>
<td>SIGINT</td>
<td>2</td>
<td>Called if the process receives an interrupt signal (Ctrl-C).</td>
</tr>
<tr>
<td>SIGQUIT</td>
<td>3</td>
<td>Terminate signal: terminates the given process</td>
</tr>
<tr>
<td>SIGFPE</td>
<td>8</td>
<td>Floating-point exception detected and will not perform</td>
</tr>
<tr>
<td>SIGKILL</td>
<td>9</td>
<td>Kill signal: forcefully terminates the given process and sends SIGKILL</td>
</tr>
<tr>
<td>SIGALRM</td>
<td>14</td>
<td>Alarm Clock signal (used for timers)</td>
</tr>
<tr>
<td>SIGTERM</td>
<td>15</td>
<td>Software termination signal (sent by kill by default).</td>
</tr>
</tbody>
</table>

nohup: “no hangup”  
nohup program1  
(Program 1 will run even when you log out)

Killing programs, part 4

List of Signals:

There is an easy way to list down all the signals supported by your system. Just issue `kill -l` command and it would display all the supported signals:

```
$ kill -l
1) SIGHUP  2) SIGINT  3) SIGQUIT  4) SIGILL
5) SIGTRAP  6) SIGABRT  7) SIGBUS  8) SIGFPE
9) SIGKILL 10) SIGUSR1 11) SIGSEGV 12) SIGUSR2
13) SIGPIPE 14) SIGALRM 15) SIGTERM 16) SIGSTKFLT
17) SIGCHLD 18) SIGCONT 19) SIGSTOP 20) SIGTSTP
21) SIGTTIN 22) SIGTTOU 23) SIGURG 24) SIGXCPU
25) SIGXFSZ 26) SIGVTALRM 27) SIGPROF 28) SIGWINCH
29) SIGIO 30) SIGPWR 31) SIGSYS 34) SIGRTMIN
35) SIGRTMIN+1 36) SIGRTMIN+2 37) SIGRTMIN+3 38) SIGRTMIN+4
39) SIGRTMIN+5 40) SIGRTMIN+6 41) SIGRTMIN+7 42) SIGRTMIN+8
43) SIGRTMIN+9 44) SIGRTMIN+10 45) SIGRTMIN+11 46) SIGRTMIN+12
47) SIGRTMIN+13 48) SIGRTMIN+14 49) SIGRTMIN+15 50) SIGRTMAX-14
51) SIGRTMAX-13 52) SIGRTMAX-12 53) SIGRTMAX-11 54) SIGRTMAX-10
55) SIGRTMAX-9 56) SIGRTMAX-8 57) SIGRTMAX-7 58) SIGRTMAX-6
59) SIGRTMAX-5 60) SIGRTMAX-4 61) SIGRTMAX-3 62) SIGRTMAX-2
63) SIGRTMAX-1 64) SIGRTMAX
```

The actual list of signals varies between Solaris, HP-UX, and Linux.

Backgrounding

• “&”: tell shell to run a job in the background
  – Background means that the shell acts as normal, but the command you invoke is running at the same time.

• “sleep 60” vs “sleep 60 &”

When would backgrounding be useful?
Suspending Jobs

• You can suspend a job that is running
  
  Press “Ctrl-Z”

• The OS will then stop job from running and not schedule it to run.

• You can then:
  
  – make the job run in the background.
    • Type “bg”
  
  – make the job run in the foreground.
    • Type “fg”

  – like you never suspended it at all!!
Outline

• Announcements/Review
• Project 1C
• File I/O, Redirection, Pipes
• Project 2A
• Symbols and mangling (if time)
• Killing and suspending jobs (if time)
• Web pages (if time)
Web pages

• ssh –l <user name> ix.cs.uoregon.edu
• cd public_html
• put something in index.html
• it will show up as
  http://ix.cs.uoregon.edu/~<username>
Web pages

• You can also exchange files this way
  – scp file.pdf <username>@ix.cs.uoregon.edu:~/public_html
  – point people to http://ix.cs.uoregon.edu/~<username>/file.pdf

Note that ~/public_html/dir1 shows up as http://ix.cs.uoregon.edu/~<username>/dir1
(“~/dir1” is not accessible via web)