Project C-Shell

Due: November 1, 2017 at 11:59pm

IMPORTANT: The TAs will start grading Shell 1 the day after it is due. Therefore, if you've handed in and might hand in again after the deadline, you need to run `cs0330_grade_me_late shell_1` *by the project deadline* to tell us not to grade your handin yet. *This is not a handin script.*

1 Introduction

Every year on October 31st, Monsters Incorporated hosts the Annual Halloween Human Harvest (AHHH): the largest scream harvesting event of the year and the one day monsters are allowed to venture outside the bedroom to scare on the streets of the human world. Last year, we generated enough screams to power the city of Monstropolis for a week (go us!), but our goal this year is to harvest one whole month of power. Your boss, Henry J. Waternoose III, has suggested transitioning from manual harvesting to a more automated process. You are a lowly monster programmer looking to make a name for yourself, and it is your job to write a computer shell to facilitate more efficient scream harvesting.

In this 2-part assignment, you will be writing your own “C” shell. A shell is typically used to allow users to run other programs in a friendly environment, often offering features such as command history and job control. Shells are also interpreters, running programs written using the shell’s language (shell scripts).

This assignment serves as an exercise in C string parsing and an introduction to system calls.

2 Assignment

Your task is as follows: your shell must display a prompt and wait until the user types in a line of input. It must then do some text parsing on the input and take the appropriate action. For example, some input is passed on to built-in shell commands, such as **`cd`**, while other inputs specify external programs to be executed by your shell.

Additionally, the command line may contain some special characters which will correspond to input/output file redirection. The shell must set up the appropriate files to deal with this.

As you know, users are far from perfect; a large part of this assignment will be supporting good error-checking while running your shell.

Install the project stencil by running

```
cs0330_install shell_1
```

2.1 Makefile

For this assignment, we have given you an outline for a Makefile that you can use to compile your program. This outline consists of a list of flags that we will use to compile your shell when grading. You are responsible for handling any warnings that these flags produce. The Makefile

also includes the target names that we would like you to use, but it does not include rules for any targets, and running **`make`** with this stencil Makefile will not compile your shell.

It is up to you to create a working Makefile for this assignment with what you learned from this week's lab. **We will be grading the Makefile you write on this assignment**, mainly for functionality and conciseness. Use variables wherever necessary to reduce repetition, and specify dependencies correctly so make knows to recompile executables if they've changed since the last make. Refer to the Makefile lab handout if you need help.

2.2 Files, File Descriptors, Terminal I/O

You have previously read from and written to files using the **`FILE`** struct and functions such as **`fopen()`** and **`fclose()`**. This struct and these functions provide a high-level abstraction for how file input and output actually works, obscuring lower-level notions such as file descriptors and system calls. In this assignment, you will be performing input using file descriptors and system calls instead of the high-level abstraction of **`fopen()`** and **`fclose()`** - though we are allowing you to use **`printf()`** and **`fprintf()`** as you normally would for output.

2.2.1 File Descriptors

At a lower level, file input and output is performed using *file descriptors*. A file descriptor is simply an integer which the operating system maps to a file location. The kernel maintains a list of file descriptors and their file mappings for each process. Consequently, processes do not directly access files using **`FILE`** structs but rather through the kernel by using file descriptors and low-level system calls.

Subprocesses inherit open files and their corresponding file descriptors from their parent process. As a result, processes started from within a normal UNIX shell inherit three open files: **`stdin`**, **`stdout`**, and **`stderr`**, which are assigned file descriptors **`0`**, **`1`**, and **`2`**[1] respectively. Since your shell will be run from within the system's built-in shell, it inherits these file descriptors; processes executed within your shell will then also inherit them. As a result, whenever your shell or any process executed within it writes characters to file descriptor **`1`** (the descriptor corresponding to **`stdout`**), those characters will appear in the terminal window.

2.2.2 open()

```
int open(const char *pathname, int flags, mode_t mode)
```

[1] The header file `unistd.h` defines macros `STDIN_FILENO`, `STDOUT_FILENO`, and `STDERR_FILENO` which correspond to those file descriptors. This is useful for making code more readable.

The **`open()`** system call opens a file for reading or writing, located at the relative (starting from the process working directory) or absolute (starting from the root directory, /) pathname, and returns a new file descriptor which maps to that file.

The other arguments for this system call are bit vectors which indicate how the file should be opened. In particular, **`flags`** indicates both status flags and access modes, allowing the user to determine the behavior of the new file descriptor. **`mode`** is used to determine the default permissions of the file if it must be created.

We recommend looking at the man pages (**`man 2 open`**) for more information.

File descriptors are opened lowest-first; that is, **`open()`** returns the lowest-numbered file descriptor available (i.e. currently not open) for the calling process. On an error, **`open()`** returns −1.

2.2.3 `close()`

```
int close(int fd)
```

`close()` closes an open file descriptor, which allows it to be reopened and reused later in the life of the calling process. If no other file descriptors of the calling process map to the same file, any system resources associated with that file are freed. **`close()`** returns 0 on success and −1 on error.

2.2.4 `read()`

```
ssize_t read(int fd, void *buf, size_t count)
```

`read()` reads up to `count` bytes from the given file descriptor (**`fd`**) into the buffer pointed to by **`buf`**. It returns the number of characters read and advances the file position by that many bytes, or returns −1 if an error occurred. **Check and use this return value**. It is otherwise impossible to safely use the buffer contents.

`read()` waits for input: it does not return until there is data available for reading. When reading from standard input, **`read()`** returns when the user types **`enter`** or **`CTRL-D`**. These situations can be distinguished by examining the contents of the buffer: typing **`enter`** causes a new-line character (**`\n`**) to be written at the end of the line, whereas typing **`CTRL-D`** does not cause any special character to appear in the buffer. You are allowed to assume that an input command ends with **`\n`**, as the demo does.

If a user types **`CTRL-D`** on a line by itself, **`read`** will return 0, indicating that no more data is available to be read—a condition called *end of file*. In this case, **your shell should exit**.

2.2.6 write()

```
ssize_t write(int fd, const void *buf, size_t count)
```

write() writes up to **count** bytes from the buffer pointed to by **buf** to the given file descriptor (**fd**). It returns the number of bytes successfully written, or -1 on an error.

While this is the lowest level and safest system call we can use to write to STDOUT and STDERR, you are not required to use **write()** for output on this assignment but may instead use **printf()** and **fprintf()** to write to STDOUT and STDERR. You will, however, have to use **read()** to read input.

2.2.7 printf()

```
int printf(const char *format, ...)
```

You're already familiar with **printf(),** which is similar to **write()**, writing formatted output to STDOUT. The primary difference is that a file descriptor doesn't needed to be specified when using **printf()**; its default file descriptor is STDOUT.

Note: if you're using **printf()** to write a string that doesn't end in a newline (hint: your prompt), you must use **fflush(stdout)** after **printf()** to actually write your output to the terminal.

2.3 Executing a Program

When a UNIX process executes another program, the process replaces itself with the new program entirely. As a result, in order to continue running, your shell must defer the task of executing another program to another process. Below is a list of system calls, functions, and shell commands useful to this project, related to executing a program:

2.3.1 fork()

```
pid_t fork(void)
```

First, you'll need to create a new process. This must be done using the system call **fork()**, which creates a new "child" process which is an exact replica of the "parent" (the process which executed the system call). This child process begins execution at the point where the call to **fork()** returns. **fork()** returns 0 to the child process, and the child's process ID (abbreviated pid) to the parent.

2.3.2 execv()

```
int execv(const char *filename, char *const argv[])
```

To actually execute a program, use the library function **`execv()`**. Because **`execv()`** replaces the entire process image with that of the new program, this function never returns if it is successful. Its arguments include **`filename`**, the full path to the program to be executed, and **`argv`**, a null-terminated[2] argument vector. Note that **`argv[0]`** MUST be the binary name (the final path component of **`filepath`**), NOT the full path to the program (which means you will have to do some processing in constructing **`argv[0]`** from **`filename`**).

As an example, the shell command **`/bin/echo Hello world!`** would have an argv that looks like this:

```
char *argv[4];
argv[0] = "echo";
argv[1] = "Hello";
argv[2] = "world!";
argv[3] = NULL;
```

See the which section to figure out how to get the full path to the program.

Here is an example of forking and executing **`/bin/ls`**, with error checking:

```
if (!fork()) {
    /* now in child process */
    char *argv[] = {"ls", NULL};
    execv("/bin/ls", argv);

    /* we won't get here unless execv failed */
    perror("execv");
    /* hint: man perror */

    exit(1);
}
/* parent process continues to run code out here */
```

2.3.3 `wait()`

```
pid_t wait(int *status)
```

Your shell should wait for the executed command to finish before displaying a new prompt and reading further input. To do this, you can use the **`wait()`** system call, which suspends execution of the calling process until a child process changes state (such as by terminating). If the status argument to wait is non-zero, details about that change of state will be stored in the memory

[2] An array for which argv[argc] is NULL, if argc is the number of entries in argv.

location addressed by status. You don't need that information in this assignment - if you pass **`wait`** the null pointer (0) then it will not store any data. Type **`man 2 wait`** into a terminal for further information.

2.3.4 `which`

```
which <program name>
```

In order to execute a program in your shell, you will need that program's full pathname. You will not be able to use only a shortcut, as you would in a bash terminal for programs such as **`ls, cat, xpdf, gedit`**, etc. To execute these programs from your shell, you must enter **`/bin/cat, /usr/bin/xpdf, /usr/bin/gedit`**, and so on. To find the full pathname for any arbitrary program, use **`which`**.

Example usage:

```
$ which cat
/bin/cat
$ which gedit
/usr/bin/gedit
```

For more information, see the **`man`** page for **`which`**. You can even use **`which`** in your shell, once you have determined its full path (type **`which which`** in a system terminal to find its full path)!

NOTE: You do not need to implement **`which`** for this assignment! It is described here as a resource to find full pathnames of programs that can be executed in your shell.

2.4 Built-In Shell Commands

In addition to supporting the spawning of external programs, your shell will support a few built-in commands. When a user enters a built-in command into your shell, your shell should make the necessary system calls to handle the request and return control back to the user. The following is a list of the built-in commands your shell should provide.

- `cd <dir>` : changes the current working directory.
- `ln <src> <dest>` : makes a hard link to a file.
- `rm <file>` : removes something from a directory.
- `exit` : quits the shell.

Note that we are only looking for the basic behavior of these commands. You do not need to implement flags to these commands such as **`rm -r`** or **`ln -s`**. You also do not need to support multiple arguments to **`rm`**, multiple commands on a single line, or shortcut arguments such as **`rm *`** or **`cd ~`**. Your shell should print out a descriptive error message if the user enters a malformed command.

2.4.1 UNIX System Calls for Built-In Functions

To implement the built-in commands, you will need to understand the functionality of several UNIX system calls. You can read the manual for these commands by running the shell command "**`man 2 <syscall>`**". It is highly recommended that you read all the man pages for these syscalls before starting to implement built-in commands.

```
int chdir(const char *path);
int link(const char *existing, const char *new);
int unlink(const char *path);
```

2.5 Prompt Format

While the contents of your shell's prompt are up to you, you must implement a particular feature in order to make your shell easier to grade. Specifically, you should surround the statement that prints your prompt with the C preprocessor directives **`#ifdef PROMPT`** and **`#endif`**, which will cause the compiler to include anything in between the two directives only when the **`PROMPT`** macro is defined.

For example, if you print your prompt with the statement `printf("33sh> ");` , you would replace it with the following:

```
#ifdef PROMPT
if (printf("33sh> ") < 0) {
    /* handle a write error */
}
#endif
```

Note: If you choose to use **`printf()`** to write your prompts, and not **`write()`**, there is an additional step you will have to take to get the prompt to show up in the terminal, because the prompt does not end in a newline. See the **`printf()`** section for more details.

Your Makefile should compile two different versions of your shell program: *`33sh`*, which compiles with `PROMPT` defined, and *`33noprompt`*, which compiles without `PROMPT` defined. If you do not remember how to compile your program with a macro defined, refer back to the maze solver Makefile or the Makefiles lab.

Any other user-defined writes to standard output (i.e. debugging printouts) from your shell should also be enclosed with the **`#ifdef PROMPT`** and **`#endif`** directives. Otherwise, the testing suite will not run correctly with your shell.

2.6 Input and Output Redirection

Most shells allow the user to redirect the input and output of a program, either into a file or through a *pipe* (a form of interprocess communication). For example, bash terminals allow you to send a program input from a file using <, send output from a program to a file using > or >>, and chain the output of a program to the input of another using |. Your shell will be responsible for redirecting the input and output of a program but not for chaining multiple programs together.

2.6.1 File Redirection

File redirection allows your shell to feed input to a user program from a file and direct its output into another file. You do not need to support redirection for built-in commands.

The redirection symbols ('<', '>', and '>>') can appear anywhere within a command in any order. For instance, the command **`echo hello > output.txt`** will write the results of **`echo hello`** to a new text file **`output.txt`**. You can visit the Linux's Redirection Definition page for specific examples and additional details of redirection.

- `< [path]` - Use file `[path]` as standard input (file descriptor 0).
- `> [path]` - Use file `[path]` as standard output (file descriptor 1). If the file does not exist, it is created; otherwise, it is truncated to zero length. (See the description of the `O_CREAT` and `O_TRUNC` flags in the `open(2)` man page.)
- `>> [path]` - Use file `[path]` as standard output. If the file does not exist, it is created; otherwise, output is appended to the end of it. (See the description of the `O_APPEND` flag in the `open(2)` man page.)

Your shell should also support error checking for input and output redirection. For example, if the shell fails to create the file to which output should be redirected, the shell must report this error and abort execution of the specified program. Additionally, it is illegal to redirect input or output twice (although it is perfectly legal to redirect input and redirect output). You can experiment with I/O redirection in the demo, which should serve as a model for the expected functionality of your shell.

2.6.2 Redirecting a File Descriptor

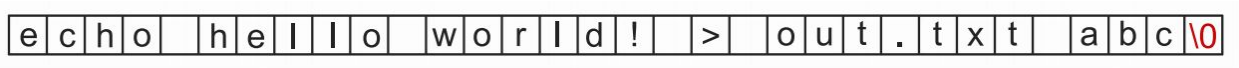
To make a program executed by your child process read input from or write output to a specific file, rather than use the default **`stdin`** and **`stdout`**, we have to redirect the **`stdin`** and **`stdout`** file descriptors to point to the specified input and output files. Luckily, the kernel's default behavior provides an elegant solution to this problem: when a file is opened, the kernel returns the smallest file descriptor available, regardless of its traditional association. Thus, if we close file descriptor 1 (**`stdout`**) and then open a file on disk, that file will be assigned file descriptor 1.

Then, when our program writes to file descriptor 1, it will be writing to the file we've opened rather than **`stdout`** (which traditionally corresponds to file descriptor 1 by default).

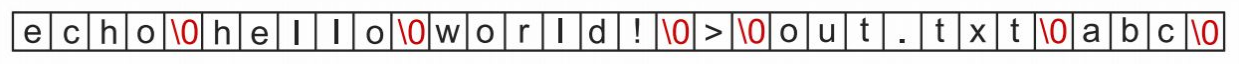
For the purposes of this project, we won't be concerned with restoring the original file descriptors for **`stdout`** and **`stdin`** in the child process as it won't affect your shell. If you're interested in the technically safer (but more complex) way to redirect files, check out the **`dup()`** and **`dup2()`** **`man`** pages.

3 Parsing the Command Line

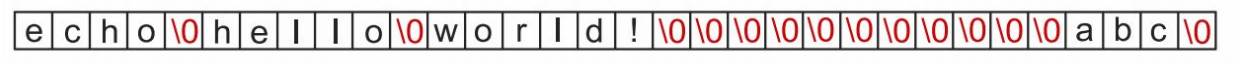
A significant part of your implementation will most likely be the command line parsing. Redirection symbols may appear anywhere on the command line, and the file name appears as the next word after the redirection symbol. One algorithm for parsing the command line is as follows:



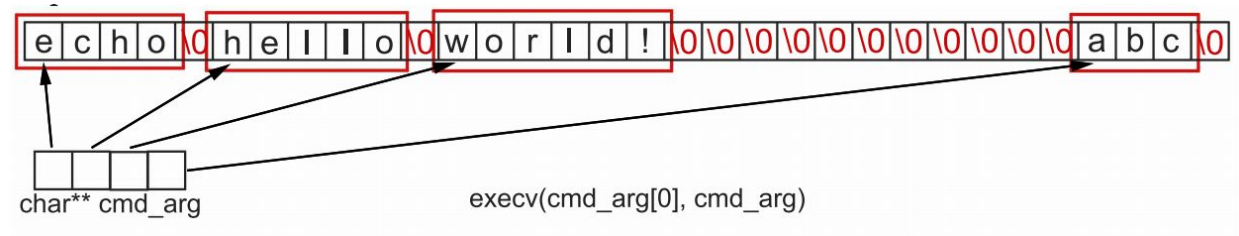
1. Split the line into words. The first word will be the command, and each subsequent word will be an argument to the command. Be sure to include the null characters so that execv can be given an array of `char *` and be still be able to find where each token ends.



2. Scan through the line for redirection symbols, keeping track of the input and output file names if they exist. Remove all traces of redirection from the command line (i.e. replace the relevant characters with the null character or spaces). Check for errors such as multiple redirection or missing filenames (i.e. a redirection token that is not followed by a filename) at this point.



3. Construct the `char**` array so that each index points to a character array from the original input buffer.



Symbols and words are separated by one or more spaces or tabs. Your shell must allow any number and combination of spaces and tabs when whitespace is required. Your shell must also support any number of spaces or tabs before the first token of the line, and after the last token.

Redirection characters will always be separated from arguments by spaces or tabs; they will not be immediately adjacent to the next or previous word. Special characters such as control characters should be treated just like alphanumeric characters and should not crash your shell. You do not need to special case quotes (in most shells quotes would group several words into a single argument that contains white space). Note that you are allowed to use many built-in C functions related to string parsing to help you do this; see section 4 for a list of allowed functions.

3.1 Invalid Command-Line Input

Be very careful to check for error conditions at all stages of command line parsing. When grading, we will run additional tests to check your code specifically for error handling on bad user input and will be deducting points if error messages aren't printed. Refer to our demo if you are unsure what is valid vs. invalid input.

Since the shell is controlled by a user, it is possible to receive bizarre input. For example, your shell should be able to handle all these cases (as well as many others):

```
33sh> /bin/cat < foo < gub
ERROR - Can't have two input redirects on one line.
33sh> /bin/cat <
ERROR - No redirection file specified.
33sh> > gub
ERROR - No command.
33sh> < bar /bin/cat
OK - Redirection can appear anywhere in the input.
33sh> [TAB]/bin/ls <[TAB] foo
OK - Any amount of whitespace is acceptable.
33sh> /bin/bug -p1 -p2 foobar
OK - Make sure parameters are parsed correctly.
```

You will not be held responsible if your input buffer is not big enough to handle user input. Use a large buffer length (e.g. 1024 bytes) and assume that the user will not enter more than that many characters.

You may assume that redirection characters are surrounded by whitespace.

4 Use of Library Functions

You should use the **`read()`** system call to read from file descriptors **`STDIN_FILENO`** (a macro defined as 0), **`STDOUT_FILENO`** (1), and **`STDERR_FILENO`** (2), which correspond to the file streams for standard input, standard output, and standard error respectively. You should use

the **write()** system call to write to **STDOUT_FILENO** or **STDERR_FILENO** OR the higher level non-system calls **printf()** (which doesn't require a specified file descriptor) and **fprintf()**.

You may use any syscall. Specifically, a system call is any function that can be accessed by using the shell command **man 2 <function>**. Do not use floating point numbers. If you have any questions about functions that you are able to use, please post your question on Piazza.

In order to avoid confusion, here is a list of allowed non-syscall functions. While use of these functions would be helpful in many implementations, it is by no means required.

opendir()	**str(n)cat()**	**tolower()**	**strerror()**
assert()	**exit()**	**atoi()**	**strtol()**
isalnum()	**isalpha()**	**iscntrl()**	**isdigit()**
islower()	**isprint()**	**ispunct()**	**isspace()**
isxdigit()	**malloc()**	**free()**	**realloc()**
memcmp()	**memcpy()**	**memmove()**	**memset()**
readdir()	**closedir()**	**perror()**	**(v)s(n)printf()**
str(n)cat()	**str(n)cmp()**	**str(n)cpy()**	**printf()**
strtoll()	**isgraph()**	**isupper()**	**fprintf()**
strlen()	**strpbrk()**	**strstr()**	**strtok()**
str(r)chr()	**str(c)spn()**	**toupper()**	**memchr()**
fflush()	**execv()**		

4.1 Error Handling

You are responsible for dealing with errors whenever you use the allowed system calls or **printf/fprintf()**. As this could get repetitive, you may want to make helper functions that will handle error-checking for you for frequently-used functions such as **printf()**. As with Maze, you are not required to error-check **fprintf()** when it is used to print an error message.

5 Support

We are providing you with a demo shell program and an automated testing program to use as you work on this project.

5.1 Demo

We have provided a demo implementation of Shell 1 to show you the expected behavior of the program. It is located in **`/course/cs0330/bin/`** with the name **`cs0330_shell_1_demo`**. There is also a no prompt version of the demo which is in **`/course/cs0330/bin/`** with the name **`cs0330_noprompt_shell_1_demo`**. You do not need to give it any arguments to run. Make sure you create an implementation both with and without a prompt, as previously described.

You should use the demo implementation as a reference for how edge cases you think of should be handled. The demo shell differs in some respects from the **`bash`** you know and love. Where they differ, emulate the demo rather than **`bash`** or another shell. For example, the **`cd`** command, when run without arguments in **`bash`**, changes directories to the user's home directory. Since you will have no way of knowing the user's home directory location, your **`cd`** implementation should emulate the demo's behavior for this case.

5.2 Tester

We have provided a test suite and testing program to test your shell. There are about 40 tests in **`/course/cs0330/pub/shell_1`**. The tester program will run some input through your shell, and then compare the output of your shell to the expected output. Each of the tests that this script will run represents input that C Shell should handle, either printing out an appropriate error or the output of a command, depending on the test. To use this script, run

```
cs0330_shell_1_test -s 33noprompt -u /course/cs0330/pub/shell_1
```

You must run the tester with the "no prompt" version of your shell - the extra characters printed by the prompt version will cause the test suite to fail. Please also note that if your **`33noprompt`** executable is not in your current directory, you will need to provide the fully-qualified path to the executable.

You can also run a single test by providing **`-t /course/cs0330/pub/shell_1/<testname>`** instead of the **`-u`** option.

Each test is a directory containing an input file, and an output and error file corresponding to the expected output of **`stdout`** and **`stderr`** respectively. Note that while most tests have their output hardcoded in their **`output`** file, some have this file generated at run time by a script called **`setup`**, also in the same folder. This shouldn't matter to you while working on this project, except if you are debugging an individual test failure where it would be useful to examine the expected outputs of the test. In these cases, make sure to look at **`setup`** so that you can see exactly how the output is constructed, if it differs from the hardcoded **`output`** file. When you run a test case, the **`setup`** is run first, and then commands in **`input`** are piped into your shell (the commands will

run in your shell), and then the tester checks if the output from your shell matches the content of the **`output`**.

The tester has some other options as well — run **`cs0330_shell_1_test -h`** to view.

If every test seems to be failing, your shell is likely printing extra information to **`stdout`** and/or **`stderr`**. Use the **`-v`** (verbose) option to check which. Also, each test is run with a 3-second timeout. If that seems to be happening for all of your tests, then your shell may not exit when it reaches **`EOF`** (when **`read()`** returns 0). Please make sure that this happens, since otherwise no test will pass.

5.3 Valgrind and Memory Safety

When grading, we will run your shell using a tool called **`valgrind`** to check it for memory safety errors. These errors include attempts to access uninitialized values and memory leaks, which occur when your program allocates memory, for example through **`malloc(),`** and then does not release those resources. We do not expect a baseline shell to use any functions that would cause potential for a memory leak, but in any case it is a good idea to get used to using **`valgrind`** to find memory safety problems. You will learn more about memory safety soon.

You can run your shell in valgrind using: **`valgrind --leak-check=full --track-origins=yes ./33sh`** . This will start your shell program. You should then run a series of commands as if you are testing your shell. When your program exits, **`valgrind`** will print a summary of any problems it encountered and where they are coming from.

5.4 Clean Up

It is important that you run **`cs0330_cleanup_shell`** (located in **`/course/cs0330/bin`**) every so often when working on your shell project. This kills any zombie processes, which if left running will eat up the computer's resources. This will be even more important when working on Shell 2.

6 GDB Tips for C Shell

As always, we recommend using GDB to help debug your project. Check out the GDB cheatsheet on the home page for more info!

6.1 Following Child Processes

When debugging your code to execute programs in C Shell it may be helpful to use GDB to verify that the programs are starting correctly. It's important to note that by default, GDB won't follow child processes started with **`fork()`**. This means that if you set a breakpoint on a line that

executes in the forked process (i.e. to make sure the arguments to **`execv()`** are formatted correctly), GDB won't break on that line.

However, you can change this behavior by running the GDB command **`set follow-fork-mode child`**. This tells GDB to debug the child process after a fork, and leave the parent process to run unimpeded. After setting this, GDB will break on breakpoints you set within a forked process. For more information, run **`help set follow-fork-mode`** within GDB.

6.2 Examining Memory in GDB

As you work on your command parsing logic, it may be helpful to use GDB to peek at an area of memory in your program, for instance the input buffer as you work with it to parse out the necessary tokens.

The simplest way of determining the value of a variable or expression in GDB is **`[p]rint <expression>`**, but sometimes you will want more control over how memory is examined. In these situations, the **`x`** command may be helpful.

The **`x`** command (short for "examine") is used to view a portion of memory starting at a specified address using the syntax **`x/(number)(format) <address>`**. For example, if you want to examine the first 20 characters after a given address, use **`x/20c <address>`** . If instead you want to examine the first 3 strings after a given address (remember that a string continues until the null character is encountered), use **`x/3s <address>`**.

Other useful format codes include **`d`** for an integer displayed in signed decimal, **`x`** for an integer displayed in hexadecimal, and **`t`** for an integer displayed in binary.

Note that the amount of memory displayed varies depending on the size of the specified format. **`x/4c <address>`** will print the first 4 characters after the given address, examining exactly 4 bytes of memory. **`x/4d <address>`** will print the first 4 signed integers after the given address, however this will examine exactly 16 bytes of memory (assuming the machine uses 4 byte integers).

7 Project Tips

We ~highly~ recommend **using gdb** (check out the GDB tips section above!) and going through the following checklist in order to get the most out of your C Shell experience:

- ❏ Read & parse input as recommended in Parsing the Command Line
- ❏ Handle child processes (get fork & execv working)
- ❏ Implement built-in commands
    - ❏ `cd <dir>` : changes the current working directory.

    - ❏ `ln <src> <dest>` : makes a hard link to a file.
    - ❏ `rm <file>` : remove something from a directory.
    - ❏ `exit` : quit the shell.
- ❏ Handle input/output redirection
- ❏ Implement error handling for syscalls and bad user input
- ❏ Comment, clean up, and think of useful abstractions for your code!

8 Minimum Requirements for Shell 2

Shell 1 is a project unto itself, but you will use most of your code for it again in next week's project, Shell 2.

For Shell 2, it is imperative that your code for Shell 1 can successfully **`fork`** and **`execv`** new child processes based on command-line input and the built-in **`cd`** command is functioning. Shell 2 does *not* require correct implementations of the built-in shell commands **`ln`**, **`rm`**, or **`exit`** or input/output redirection. In other words, you must **complete at least the first 2 tasks on the checklist (and also the `cd` built-in)** in the previous section before proceeding to Shell 2.

Here are some things to be aware of:

- A baseline Shell 1 implementation will *not* be released for work on Shell 2. All of the code you write for these two assignments will be your own.
- Late days used on Shell 1 will *not* also apply to Shell 2.
- If your Shell 1 project contains functionality errors, you may receive up to 50% of the points lost if those mistakes are fixed by the time of your Shell 2 handin. To request these points back, please document bug fixes in your README when you turn in Shell 2.
- If you hand in Shell 1 on time, we will give you feedback before the Shell 2 deadline (likely the morning of). As these projects are time-consuming to grade, we cannot guarantee that you will get a grade back for Shell 1 before the Shell 2 deadline if you hand in late. You will still be eligible for points back, but you may have to rely on your own testing to find errors.

9 Grading

Your grade for the first part of the shell project will be determined by the following categories, in order of precedence:

- *Functionality*: your shell should produce correct output.
- *Code Correctness*: your Makefile should work and your code should compile without warnings. Your code should be free of memory leaks and system calls should be used correctly. You must abide by the restrictions on library functions imposed in section 4—you *will* be penalized for using disallowed functions.

- *Error checking*: your shell should perform error checking on its input and display appropriate, informative error messages when any error occurs. Error messages should be written to standard error rather than standard output. Make sure you check the return value of each system call you use and handle errors accordingly.
- *Style*: your code will be evaluated for its style.

10 Handing In

To summarize, here is a list of features that a fully functioning shell would support:

- Continuously reads input from the user until it receives **`EOF`** (`Ctrl-D`)
- Executes programs and passes the appropriate arguments to those programs
- Supports 4 built in commands (**`cd, rm, ln, exit`**)
- Supports 3 file redirection symbols (**`<, >, >>`**), including both input and output redirection in the same line.
- Extensive error checking

To hand in the first part of your shell, run

```
cs0330_handin shell_1
```

from your project working directory. You must at a minimum hand in all of your code, the Makefile used to compile your program, and a README documenting the structure of your program, any bugs you have in your code, any extra features you added, and how to compile it.

If you wish to change your handin, you can do so by re-running the handin script.

Important note: *If you have handed in but plan to hand in again after the deadline, in addition to running the regular handin script (***`cs0330_handin shell_1`***), you must run* **`cs0330_grade_me_late shell_1`** *to inform us not to start grading you yet. You must run the script by the shell 1 deadline (11/1 at 11:59pm), because we will start grading the day after the project is due*.

If you do not run this script, the TAs will proceed to grade whatever you have already handed in, and you will receive a grade report with the rest of the class that is based on the code you handed in before we started grading.

If something changes, you can run the script with the **`--undo`** flag (before the project deadline) to tell us to grade you on-time and with the **`--info`** flag to check if you're currently on the list for late grading.