

Operating Systems - 234123

Homework Exercise 1 - Wet

Teaching Assistant in charge:

Mohammad Agbarya

Assignment Subjects & Relevant Course material

Processes and IPC (inter-process communication)

Recitations 1-3 & Lectures 1-3, HW0

Introduction

A Unix shell is a command-line interpreter or shell that provides a command-line user interface for Unix-like operating systems [Wikipedia].

There are few well-known shells on Unix systems, one of the most used is Bash. Bash is a command processor that typically runs in a text window. Users can interact with the Bash shell by typing commands that cause some actions.

In this assignment, you are going to implement a “**smash**” (small shell) which will behave like a real Linux shell but it will support only a limited subset of Linux shell commands.

Please read the whole document before you start working.

Description

For simplicity, you can assume that only up to 100 processes can run simultaneously and the name of each process can contain up to 50 characters. In the smash.zip file attachment you will find a base code for handling commands (command.c), a shell code for smash.c, and a Makefile, you must complete the code found in command.c and add signal handling, the functions that handle and define the signal handling routine will be in the file siganls.c.

The program will work as follows:

- The program waits for commands that will be typed by the user and executes them (and back forth).
- The program can execute a small number of built-in commands, the commands will be listed below.
- When the program receives a command that is not one of the built-in commands (we call it an external command), it tries to run it like a normal shell. The way to run external commands will be described later.
- If one of the built-in commands was typed but with invalid parameters, the following error message should be printed:
smash error : > “command-line”
where command-line is the command line as typed by the user (note that the error message should be wrapped with double quotes)
you’ll see examples later on.
- Whenever an error occurs a proper error message should be printed and the program should get back to parse and execute next typed commands.
- While the program is waiting for the next command to be entered, it should print the following text:
smash>

Assumptions - you may assume the following:

- Each command appears on a separate line and cannot exceed 80 characters.
- You can assume that the number of arguments is up to 20.
- Any number of spaces can be used between words (in the same command line) and at the beginning of the line.
- Your smash should support up to 100 processes that can run simultaneously.

- The name of each process can contain up to 50 characters.

Detailed Description

As mentioned before, there are two kinds of commands to be executed by smash: **built-in commands** and **external commands**.

Built-in commands can be considered as features that shell propose for its users. Built-in commands usually run from the same code of the shell (same process) and they should not be forked and executed. Generally, built-in commands are simple and should not execute some external executable to achieve their goals (in most cases they will just run some query system call or maintain some internal data structures of the shell).

External commands usually require running an external executable. To that end, the shell runs external commands in a child process (i.e., shell will `fork` and `execv` to run these commands).

In shell terminology, a job is a process that the shell manages (processes forked by the shell process). Each job has its own sequential ID (which is assigned to the job by shell once it's inserted to the jobs list). since the job is also a separate process (forked by the shell) it also has a process ID (`PID`) in addition to the sequential ID. Please do not confuse between the sequential ID, which the shell assigns to the job and the process ID (`PID`) which is assigned to process by the kernel.

Shell job can be in one of three available states:

1) Foreground:

When you type a command in the shell terminal window, the command will start running and cause the shell to be blocked waiting for the command to finish (the shell waits for the [child] process to finish in this case). This is called a foreground job (which runs in the foreground and the shell waits for its completion to allow next commands to be typed)

2) Background:

When an ampersand symbol (&) is typed at the end of a command line, it means that the shell should run this command in the background. Running a job in the background means that it will not occupy the terminal window and the shell prompt will be displayed immediately to allow users to type new commands even though the background job is still running. Running a job in the background means that the shell will run it without waiting for its completion (without using a blocking wait system call). This is called a background job.

3) Stopped:

Users can stop running jobs in the foreground by pressing Ctrl+Z. Stopped jobs will remain stopped (their execution will be halted/stopped) until they get signaled with SIGOCONT then they will resume their execution.

Any job that either sent to the background (using the ampersand symbol (&)) or stopped (by pressing Ctrl+Z) will be added to the **jobs list** for allowing users to query and manage them later on by their associated sequential ID.

Please refer to bg, fg, and jobs commands for further details on how to manipulate and manage the jobs list.

Built-in commands:

Your smash should support and implement a limited number of shell commands (features). The commands you need to support will be listed below with a detailed description. Please note that these commands (the built-in commands) should be executed from the smash process (i.e., you should not fork the smash process to run them, but you will run them from smash code).

Here are the details of built-in commands that your smash should support them:

1) pwd

Description:

pwd command has no arguments.

pwd prints the full path of the current working directory. In the next command (cd command) will explain how to change the current working directory.

You may use [getcwd\(\)](#) function to retrieve current working directory.

Example:

```
smash> pwd
/home/magbarya/234123/homeworks/smash/build
smash>
```

Error handling:

If any number of arguments were provided with pwd then they will be ignored.

2) cd <path>

Description:

Change directory (cd) command receives a single argument <path> that describes the relative or full path to change current working directory to it. There is a special argument that cd can get which is "-". If "-" was specified as the only argument of cd command then it will change the current working directory to last working directory. That means, if the current working directory is X, then cd was executed to change the current working directory to Y, and then cd was called (again) with "-" then it should go back and set the current working directory to X.

You may use [chdir\(\)](#) system call to change the current working directory.

Example:

```
smash> cd -  
smash error: cd: OLDPWD not set  
smash> cd dir1 dir2  
smash error: cd: too many arguments  
smash> cd /home/magbarya  
smash> pwd  
/home/magbarya  
smash> cd ..  
smash> pwd  
/home  
smash> cd -  
smash> pwd  
/home/magbarya  
smash>
```

Error handling:

If more than one argument was provided, then cd command should print the following error message:

```
smash error: cd: too many arguments
```

If last working directory is empty and “cd -” was called (before calling cd with some path to change current working directory to it) then it should print the following error message:

```
smash error: cd: OLDPWD not set
```

If `chdir()` system call failed (e.g., `<path>` argument points to a non-existing path) then perror should be used to print proper error message (as described in Error Handling section).

3) history

Description:

Your smash should track the typed commands and save them in a history log.

history command should print the contents of the history log (which saves the executed commands in smash) formatted as will be explained.

The maximum capacity of the history log is 50 records maintained in a cyclic list. This means if the history log is full then it will go back to save new records from the beginning (in this case the first record will be the oldest one before overriding it with the new record).

Each record should contain a simple timestamp (a sequence number starting from 1) and the typed command line exactly as it was typed (with no processing).

If the same command was repeated consecutively (typed consecutively more than once) then it will be saved only once at history log but its sequence number will be incremented instead of inserting it more than once in the history log (see the example below).

Each log record should be printed as follows: first, the sequence number should be printed and padded with 5 white spaces, followed with double white spaces, and then

the command line as it was typed exactly with no processing. You can use the following code line to print a single log record:

```
cout << right << setw(5) << << [sequence-number] << " " << [command-line] << endl;
```

The history log should be printed starting from the oldest record to the newest one.

Example:

```
smash> history
  1 history
smash> cd ..
smash> cd -
smash> history
  1 history
  2 cd ..
  3 cd -
  4 history
smash> cd -
smash> cd -
smash> cd -
smash> history
  1 history
  2 cd ..
  3 cd -
  4 history
  7 cd -
  8 history
smash>
```

Error handling:

If any number of arguments were provided with history then they will be ignored.

4) jobs

Description:

jobs command prints the jobs list which contains:

1. unfinished jobs (which are running in the background). 2. stopped jobs (which were stopped by pressing Ctrl+Z while they are running).

the list should be printed in the following format:

if the job was stopped (see stopped jobs def. above)

```
[<job-id>] <command> : <process id> <seconds-elapsed> (stopped)
```

else

```
[<job-id>] <command> : <process id> <seconds elapsed>
```

where <seconds elapsed> is the seconds elapsed since the job was inserted to jobs list

(hint: think about using [time\(\)](#) system call and [difftime\(\)](#) library function) .

note: if the job was added again the timer should reset.

Example:

```
smash> sleep 100&
```

```
smash> sleep 200
^Zsmash: process 30902 was stopped
smash> jobs
[1] sleep 100& : 30901 18 secs
[2] sleep 200 : 30902 11 secs (stopped)
smash>
```

note: sleep is an external command, we'll see external commands in the next section.

Error handling:

If any number of arguments were provided with this command then they will be ignored.

5) kill -<signum> <job-id>

Description:

Kill command sends a signal which its number specified by <signum> to the job which its sequence ID in jobs list is <job-id> (same as job-id in jobs command). and prints a message reporting that the specified signal was sent to the specified job (see example below)

Example:

```
smash> kill -9 1
signal number 9 was sent to pid 30985
smash>
```

Error handling:

If job-id was specified with a job id which does not exist then the following error message should be reported:

```
smash error: kill: job-id <job-id> does not exist
```

If the syntax (number of arguments or the format of the arguments) is invalid then an error message should be printed as follows:

```
smash error: kill: invalid arguments
```

If the kill system call fails then report the failure using perror (as described in Error Handling section).

6) showpid

Description:

showpid command prints the smash pid.

Example:

```
smash> showpid
smash pid is 30903
smash>
```

Error handling:

If any number of arguments were provided with this command then they will be ignored.

7) fg <job-id>

Description:

fg command brings a stopped process or a process that runs in the background to the foreground.

fg command prints the command line of that job along with its pid (as can be seen in the example) and then sends to the required job SIGCONT signal and waits for it (hint: [waitpid](#)), which in effect will bring the requested process to run in the foreground.

The job-id argument is an optional argument. if it is specified then the specific job which its sequence number (as printed in jobs command) should be brought to the foreground. If job-id argument is not specified then the last added job to jobs list should be selected to be brought to the foreground.

side effects: After bringing the job to the foreground, it should be removed from the jobs list.

Example:

```
smash> sleep 100&
smash> sleep 200&
smash> sleep 300&
smash> sleep 400&
smash> sleep 500&
smash> jobs
[1] sleep 100& : 978 14 secs
[2] sleep 200& : 979 11 secs
[3] sleep 300& : 980 8 secs
[4] sleep 400& : 981 6 secs
[5] sleep 500& : 982 1 secs
smash> fg 5
sleep 500& : 982
```

Error handling:

If job-id was specified with a job id which does not exist then the following error message should be reported:

```
smash error: fg: job-id <job-id> does not exist
```

If fg was typed with no arguments (without job-id) but the jobs list is empty then the following error message should be reported:

```
smash error: fg: jobs list is empty
```

If the syntax (number of arguments or the format of the arguments) is invalid then an error message should be printed as follows:

```
smash error: fg: invalid arguments
```

8) bg <job-id>

Description:

bg command resumes one of the stopped processes in the background.

bg command should first print the command line of the job to be resumed and only then to resume it (by sending the SIGCONT signal) and continue running it in the background (i.e., there is no need to `wait` for it).

Job-id argument is an optional argument, if it is specified then the specific stopped job which its sequence number (as printed in jobs command) should be resumed in the background. If job-id argument is not specified then the last added stopped job [to jobs list] should be selected to continue running it in the background.

side effects: After resuming the stopped job in the background, the stopped mark should be removed from the jobs list.

Example:

```
smash> bg 4  
sleep 100 : 30986
```

Error handling:

If job-id was specified but does not exist in the jobs list then the following error message should be reported:

```
smash error: bg: job-id <job-id> does not exist
```

If job-id exists but it is for a job which is already running in the background (not stopped) then the following error message should be reported:

```
smash error: bg: job-id <job-id> is already running in the background
```

If bg was typed with no arguments (without job-id) but jobs list does not contain any stopped job to resume in the background then the following error message should be reported:

```
smash error: bg: there is no stopped jobs to resume
```

If the syntax (number of arguments or the format of the arguments) is invalid then an error message should be printed as follows:

```
smash error: bg: invalid arguments
```

9) **quit [kill]**

Description:

quit command exits the smash. If kill argument was specified (which is optional) then smash should kill (by sending SIGKILL signal) all of its unfinished and stopped jobs before exiting.

If kill option was specified then it should print (before exiting) the number of processes/jobs that were killed, their PIDs and command-lines (see the example for output formatting)

Example:

```
smash> quit kill  
smash: sending SIGKILL signal to 3 jobs:  
30959: sleep 100&  
30960: sleep 200  
30961: sleep 10&  
Linux-shell:
```

Error handling:

If any arguments were specified other than kill then they will be ignored.

10) cp <old-file-path> <new-file-path>

Description:

cp command copies a file from <old-file-path> which could be a relative or full path to an existing file to a new file specified by the <new-file-path> which could be a relative or full path.

Please note that this is a built-in command and not an external command, which means you must implement it inside smash code without using fork and execv (hint: cp command uses mainly three well-known system calls: open/read/write). Please note that you are required to check open/read/write system calls return code and act accordingly (i.e., your code should consider that write may not write all the data you asked to write and read may not read all the data you asked to read).

Implementing cp command as fork/execv (to run Linux cp utility) will not be accepted and doing that will cause **points reduction**, please avoid that.

Example:

```
smash> cp file1.txt file2.txt
smash: file1.txt was copied to file2.txt
smash>
```

Error handling:

If any of the system calls used to do the copy fails for any reason (including, file not found) should be reported using perror as described in the Error Handling section.

External commands:

Besides the built-in commands, the smash should support executing external commands that are not part of the built-in commands. External command is **any command** that is not a built-in command.

The smash should execute the external command and wait until the external command is executed.

Command line:

```
<command> [arguments]
```

Where:

command is the name of the external command/executable.

The arguments are optional, i.e., if they exist in the external command line then they should be passed as arguments of the external command.

You may assume that the number of arguments is up to 20.

How to run:

The external command can be provided as a simple command to run some executable with its arguments, for example:

```
a.out arg1 arg2
```

But, the external command could be provided in a complex way, i.e., with the use of some special characters, like "*" and "?". For example:

```
rm *.txt
```

TIP: In this case, you may prefer to run the external command using bash, which already has the mechanism to parse those complex commands and execute them, you can do that by calling bash with your received <command> [args] in the following form:

```
bash -c "<command> [args]"
```

For example:

```
bash -c "rm *.txt"
```

in this case smash will run bash (which its arguments in this case is ["-c", "rm *.txt"]) and bash will run the external command.

note¹: running bash means calling one of [exec](#) family syscalls on "/bin/bash" binary.

note²: no, you don't have to learn bash language for this assignment.

What is a complex command:

As we mentioned above, external commands should be executed using bash.

Any external command which includes one of the following signs should be considered as a complex command and to be executed using bash:

"*" or "?"

External commands in background:

External command can be executed in the background as in most Linux shells.

If a "&" sign is added to the end of a command line, then this command should be executed in the background.

For example:

```
ls -l &
```

When an external command ends with "&" (ignoring white spaces around it) it should be executed as explained above (in regular external commands section) with a minor change that smash should not wait for the command completion.

The command being executed in the background should be added to the **Jobs list** (as we saw sleep& example in jobs command).

Handling Ctrl+C and Ctrl+Z:

You smash should support catching Ctrl+C and Ctrl+Z and handle them as will be explained.

Ctrl+C causes the shell to send SIGINT to the process in foreground and Ctrl+Z causes the shell to send SIGTSTP to the process in foreground (note that from the point of view of Linux shell, the smash is the process who is running in foreground and it's not the process you are currently running from your smash, so the SIGINT is supposed to be sent to smash main process).

Your smash should route these two signals to the running process in the foreground. If there is no process running in the foreground then your smash should ignore them.

meaning:

Pressing Ctrl+C will send SIGINT to your smash and your smash should route it (send SIGINT signal) to the process which is running currently in the foreground, which will kill it.

Pressing Ctrl+Z will send SIGTSTP to your smash, and your smash should route it and stop the running process in the foreground (i.e., to send SIGSTOP signal to the process running in foreground).

The stopped process should be added to jobs list and marked as stopped as explained before (see jobs command).

When Ctrl+C is pressed, your smash should do the following:

- print the following information: **smash: got ctrl-C**
- send SIGKILL to the process in the foreground. If no process is running in the foreground then no signal will be sent.
- print the following information: **smash: process <foreground-PID> was killed**

When Ctrl+Z is pressed, your smash should do the following:

- print the following information: **smash: got ctrl-Z**
- add the foreground process to jobs list. If no process is running in the foreground then nothing will be added to jobs list.
- send SIGSTOP to the process in the foreground. If no process is running in the foreground then no signal will be sent.
- print the following information: **smash: process <foreground-PID> was stopped**

Example:

```
smash> sleep 1000
^Csmash: got ctrl-C
smash: process 31951 was killed
smash> sleep 1000
^Zsmash: got ctrl-Z
smash: process 31952 was stopped
smash> jobs
[1] sleep 1000 : 31952 2 secs (stopped)
smash>
```

important note:

Please note that you need to use [setpgrp](#) system call to change the group ID of all of your forked childs. This is necessary due to the fact that Linux shell may send SIGINT and SIGTSTP (in case Ctrl+C and Ctrl+Z was pressed) to your smash and **all** his children. This could happen because Linux shell sends the signals to all processes share the same group ID.

When you call fork() the created process (the child process) will have the same group ID as his parent unless you change it through `setpgrp()` system call. You need to call `setpgrp` right after `fork` in the child code, for example:

```
pid = fork();
if( pid == 0 ) {
```

```
// child process code goes here
setpgrp();
.....
} else ....
```

Pipes and I/O Redirection Support (20 bonus pts.)

Your smash code should support simple IO redirection and pipes features. You can assume, for simplicity, that each typed command could have up to one character of pipe or IO redirection. IO redirection characters that your smash should support: ">" and ">>". Pipe characters that your smash should support are: "|" and "|&".

How to use these pipes and redirection features:

- "command > output-file": using the ">" redirection character (as in this example) causes the command stdout file-descriptor to be redirected to output-file and writes any output produced by "command" to the given "output-file". If the output-file does not exist then it creates it and if it exists then it **overrides** its content.
- "command >> output-file": the same as ">" except that using ">>" will **append** command output to the given output-file if exists. If output-file does not exist then ">" and ">>" will produce the same result.
- "command1 | command2": using the pipe character "|" will produce a pipe, redirects command1 **stdout** to its write channel and command2 stdin to its read channel.
- "command1 |& command2": using the pipe character "|&" will produce a pipe, redirects command1 **stderr** to the pipe's write channel and command2 stdin to the pipe's read channel.

Your smash should have native support for pipes and IO redirection and they should be supported for built-in and external commands.

Example:

```
smash> showpid > pid.txt
smash> cat pid.txt
smash pid is 27882
smash> showpid >> pid.txt
smash> cat pid.txt
smash pid is 27882
smash pid is 27882
smash> ls -l > ls.txt
smash> cat ls.txt
total 140
-rwxrwxrwx 1 root root 6376 Nov 12 10:52 Makefile
-rwxrwxrwx 1 root root 146 Nov 12 14:11 ls.txt
-rwxrwxrwx 1 root root 38 Nov 12 14:11 pid.txt
-rwxrwxrwx 1 root root 82832 Nov 12 13:41 smash
smash> cat pid.txt | grep smash
smash pid is 27882
smash pid is 27882
smash> showpid | grep smash
smash pid is 27882
smash>
```

Error Handling:

- If a system call fails then your smash should use perror() function to report the failure.

The error message (to be provided to perror) should be as follows:

- **“smash error: <syscall name> failed”**

Where syscall-name is the name of the called method to execute the system call, for example, if a call to fork failed then should report about the failure this way:

```
perror(“smash error: fork failed”);
```

Important Notes and Tips

- The assignment will be graded by auto tests. Write your own tests and be sure that your system is working according to the above specification.
- Use the same setup you been required to use for HW0.
Please keep in mind, in case you are working with Azure, to stop your virtual machine (for not consuming all your allocated time on the cloud), also please remember that all virtual machines will shutdown automatically every day at 5:00 AM.
- First, try to understand exactly what your goal is.
- Figure out which data structures will serve you in the easiest and simplest way
- Write your own tests. We will check your assignment with our test program
- Start working on the assignment **as soon as possible**. The deadline is final, NO postponements will be given.

Questions & Answers

- The Q&A for the exercise will take place at a public forum Piazza **only**. Please **DO NOT** send questions to the private email addresses of the TAs.
- Critical updates about the HW will be published in **pinned** notes in the piazza forum. These notes are mandatory and it is your responsibility to be updated.

A number of guidelines to use the forum:

- Read previous Q&A carefully before asking the question; repeated questions will probably go without answers
- Be polite, remember that course staff does this as a service for the students
- You're not allowed to post any kind of solution and/or source code in the forum as a hint for other students; In case you feel that you have to discuss such a matter, please come to the reception hour
- When posting questions regarding **hw1**, put them in the **hw1** folder

Late Days

- Please **DO NOT** send postponement requests to the TA responsible for this assignment. Only the **TA in charge** can authorize postponements. In case you need a postponement, please fill out the attached form

<https://forms.gle/nqje6UoKXVbT5sW56>

Submission

You should create a zip file (use zip only, not gzip, tar, rar, 7z or anything else) containing the following files:

- 1) All source files you wrote **with no subdirectories of any kind**.
- 2) A file named **submitters.txt** which includes the ID, name and email of the participating students. The following format should be used:

Linus Torvalds linus@gmail.com 234567890 Ken Thompson ken@belllabs.com 345678901

Important Note: Make the outlined zip structure exactly. In particular, the zip should contain only the X files, without directories.

You can create the zip by running:

<code>zip XXX_YYY.zip <source_files> submitters.txt</code>
--

The zip should look as follows:

<pre>zipfile -+ Commands.h Commands.cpp signals.h signals.cpp smash.cpp Makefile submitters.txt +- other files</pre>
--

Important Note: when you submit, **retain your confirmation code and a copy of the file(s)**, in case of technical failure. Your confirmation code is **the only valid proof** that you submitted your assignment when you did.

Have a Successful Journey,

The course staff