

Project No. 1: A Simplified Parallel (multi-process) Make Program

PURPOSE

- Practicing UNIX system calls.
- Practicing multi-process programming.

DESCRIPTION

In this project, you will implement a parallel (multi-process) version of a simplified make program whose behavior should be (almost) the same as the UNIX make utility. Your program should be able to take advantage of the computing power in the current multi-core processors. The simplified make program organizes a project by performing operations based on the dependencies specified in a makefile. The makefile can have three types of components: macros, target rules, and inference rules. The format for a macro is

```
string1=string2
```

which defines *string1* to be *string2*. The reference to $\$string1$ or $\$(string1)$ will be replaced with *string2* during the execution of the make program. The format for a target rule is as follows.

```
Target [target] ... : [prerequisite] [prerequisite] ...
<tab>command1
<tab>command2
...
```

Multiple targets and prerequisites can be specified in the rules. Each of the targets and prerequisites is a string. The format of the commands will be described later. The format of an inference rule is as follows:

```
Target:
<tab>command1
<tab>command2
...
```

The target must be of the form $.s_1$ or $.s_1.s_2$. For a inference rule with the $.s_1$ target, the rule applies to make the target file X from the source file $X.s_1$. For a inference rule with the $.s_1.s_2$ target, the rule applies to make the target file $X.s_2$ from the source file $X.s_1$. Each of the commands associated with a rule can be anything that the shell can execute, including I/O redirection, pipe, and background execution, and the sequential execution of multiple commands.

The simplified make program should support the following features:

- The program should behave like *make*: checking all dependencies (and the timestamps of the related files) in the makefile and performing operations only when needed.
- The default makefiles are *mymake1*, *mymake2*, and *mymake3* in the current directory in order. The default can be overwritten with an “-f” flag. The command line can allow making multiple targets.
- The program should support all three components in the makefile for the UNIX make utility: macros, target rules, and inference rules.

- The makefile should allow comments: everything that follows character '#' is a comment.
- Commands associated with inference rules can use two special symbols: '\$@' for the target without suffix and '\$<' for the source.
- Commands associated with rules can use defined macros in the form of either '\$string' or '\$(string)', as well as the two special symbols (inference rules only).
- Each command line to be supported can be in one of the following forms:
 - a simple command with command line arguments.
 - multiple commands separated by ';'. These commands are to be executed sequentially in order.
 - the 'cd' command (making sense only in multiple commands).
 - multiple piped commands: *cmd1* | *cmd2* | *cmd3* | *cmd4*.
 - a command to be executed in background ('&').
 - a command with redirected I/O ('>' and '<').
- The program should stop executing when the execution of a command failed.
- The paths to search for the commands in a relative path are stored in an environment variable MYMAKEPATH, which has the same format as the *PATH* variable in *tcsh*.
- The program should allow circular dependencies in the makefile.

To take advantage of multi-core processors, the program must use multiple processes to make targets. In particular, if there are more than three targets to be made, your program must fork at least three processes to make the targets. Your program should also try to balance the load among the processes. The processes must be synchronized in some way to enforce the dependence.

DEADLINES AND MATERIALS TO BE HANDED IN

Due date: **Sept. 22** and **Oct. 4**. A sequential version of the code is due Sept. 22 (demo after class). The parallel version of the code is due Oct. 4 (demo after class). You should clean and tar your project directory and send it to me. In the directory, you should have all the needed files to create the executables of the program. On both *program* and *linprog*, a 'make' command should create the executable. You should also have a README file describing how to compile and run your program and the known bugs in your program. You should also have a makefile that will (1) automatically generate the executable by issuing a 'make' command under the directory, (2) compile the C files with -Wall -ansi -pedantic, (3) clean the directory by issuing 'make clean', and (4) recompile only the related files when a file is modified.

You need to hand in a hard-copy of your programs, makefiles and the README file in the due date when you demonstrate your project.

GRADING POLICY

A program with compiling errors will get 0 point. A program that cannot run any command in the makefile will get 0 point. A program with any 'system' routine in it will get 0 point (the 'system' routine is NOT allowed in the project). Any thing that only works on one platform will get 50% of the points.

1. proper README file (5)
2. proper makefile file(for compiling the simplified make utility) (5)
3. using the correct default makefile, the “-f” flag (5)
4. running one rule with one simple command (5)
5. allowing multiple commands in a line (;) (5)
6. allowing the 'cd' command (the changed directory is only effective in the execution of that line) (5)
7. running multiple piped commands (|) (5)
8. allowing background execution (5)
9. allowing I/O redirection in a command (> and <) (5)
10. Searching each relative path with paths in MYMAKEPATH (10)
11. supporting macros (5)
12. supporting inference rules, \$@ and \$ < symbols (5)
13. supporting comments (#) (5)
14. supporting simple dependencies with multiple rules (5)
15. supporting advanced dependencies with multiple rules (5)
16. parallel version with correct synchronization (5)
17. simple load balancing mechanism (5)
18. Overall similar behavior as the UNIX make (5)
19. proper self-guided demo (5)

MISCELLANEOUS

- The implementation of this program will likely be more than 1000 lines of dense code (500- for parsing and dependency checking and 500+ for executing commands in different forms). Please start the project as early as you can.
- It is normal for system calls to have unexpected behavior. You can either call it a bug or a feature. As a result, you might need to work your way around if that occurs to your program — this is part of the project.
- You can make any assumptions about the format (syntax) of the file as long as the format does not limit the commands' functionality and contradict our specification. Your program is considered correct as long as you can handle the correct inputs.
- There are mainly three modules in this program (1) the module to load the makefile into some internal data structures that can be easily manipulated, (2) the module to check the dependencies and act accordingly (this is likely a recursive routine), and (3) the execution module that runs one line of commands at a time. Initially, you should not pay too much attention to (1) and (2), rather, you need to work on (3) and create a make program that can handle makefiles that are similar to sample0.mk. The right order to implement the modules is (3) → (1) → (2) (if you don't know where to start).
- Your program should work with all examples in sample.tar, as well as all the makefile in examples/lect2/*. More makefiles will be created to test your submission (you want to create some for your own testing).
- There are many ways that the make utility can be parallelized. One simple method is to have a process to read in the rules and determine all targets that need to be made, as well as their dependencies. This can be represented by a DAG. After that, you can partition the DAG in some way, and have different processes to make the target in different part of the DAG. Synchronization among processes can be realized by pipes.
- All programs submitted will be checked by a software plagiarism detection tool.
- 10 points for the second unknown bug.