Họ và tên: Nguyễn Tiến Đạt - Vũ Nguyễn Duy Linh - Đặng Quang Khánh Linh -

Nguyễn Đỗ Đức Minh

Mã số sinh viên: 22520226 - 22520780 - 22520756 - 22520972

Lóp: IT007.014.1

HỆ ĐIỀU HÀNH BÁO CÁO LAB 6

CHECKLIST

6.4. BÀI TẬP THỰC HÀNH

	Câu 1	Câu 2	Câu 3	Câu 4	Câu 5
Trình bày giải thuật		\boxtimes	\boxtimes	\boxtimes	\boxtimes
Chụp hình minh chứng (chạy ít nhất 3 lệnh)		\boxtimes			
Giải thích code, kết quả					

Tự chấm điểm: 9.5

*Lưu ý: Xuất báo cáo theo định dạng PDF, đặt tên theo cú pháp:

<Tên nhóm>_LAB6.pdf

Link code: https://github.com/duylinh1510/LAB6-OS/blob/main/vinashell.c

6.4. BÀI TẬP THỰC HÀNH

1. Câu 1: Thực thi command trong tiến trình con

Trình bày giải thuật & giải thích code:

(Trong code đã ghi chú thích ở từng dòng, nên ở các câu hỏi sẽ chỉ ra vị trí và nêu cách thực hiện ý tưởng của code)

- Để hiện thị dấu nhắc ví dụ 'IT007> ', ta tạo 1 vòng lặp với điều kiện lặp là biến should_run (biến cờ hiệu, vòng lặp này sẽ luôn lặp cho tới khi biến cờ hiệu này mang giá trị bằng 0):

```
int should_run = 1; // Biến cờ để xác định chương trình có tiếp tục chạy hay không
while (should_run)
{
    printf("it007sh> "); // Dấu nhắc
    fflush(stdout); // Xóa bộ đệm
```

- Để người dùng nhập lệnh và thực thi lệnh đã nhập, ta lấy dữ liệu input từ người dùng. Dùng hàm strcspn tìm vị trí đầu tiên của '\n' trong chuỗi và thay bằng '\0' để kết thúc chuỗi tại đó nhằm loại bỏ ký tự xuống dòng:

```
char command[MAX_LINE];
fgets(command, MAX_LINE, stdin); // Nhập câu lệnh
command[strcspn(command, "\n")] = '\0'; // strcspn tìm vị trí đầu tiên có '\n'
```

- Bỏ qua câu lệnh rỗng nếu giá trị trả về của strlen(command) bằng 0. Kiểm tra câu lệnh nhập vào có phải là 'exit' hay không, nếu phải thì thực hiện tắt shell bằng cách gán giá trị của biến **should_run** = 0 → thoát khỏi vòng lặp while(should_run):

```
130
              // Bỏ qua nếu câu lệnh rỗng
              if (strlen(command) == 0)
131
                   continue;
132
133
               // Kiểm tra lệnh đặc biệt của người dùng nhập vào như "exit"
134
               if (strcmp(command, "exit") == 0)
135
136
137
                   should_run = 0;
                   continue;
138
139
```

- Tiếp đến là kiểm tra xem lệnh người dùng nhập vào có phải là lệnh 'HF' hay không và tiến hành xử lý nếu phải (sẽ trình bày chi tiết ở câu 2):

```
// Kiếm tra lệnh đặc biệt HF

char *hfPtr = strstr(command, "HF"); // Trả về con trỏ đến vị trí đầu tiên

int hfUsed = hfPtr == NULL ? 0 : 1; // Biến cờ để xác định người dùng đã nh

if (hfUsed)

{...
}

// Lưu lịch sử lại câu lệnh vừa nhập (nếu không phải lệnh HF)

if (!hfUsed)

{...
}

// Lưu lịch sử lại câu lệnh vừa nhập (nếu không phải lệnh HF)

if (!hfUsed)

{...
}
```

- Bước tiếp theo và cũng là ý chính của yêu cầu 1, chương trình gọi hàm fork() để tạo ra thêm 1 tiến trình con để thực thi các câu lệnh:

```
// Để ra tiến trình con để giao nhiệm vụ thực thi lệnh cho nó
pid = fork();
if (pid == -1)
{
    perror("fork() failed. Try again.");
    continue;
}
```

- * **Chú thích:** biến **pid** trong ảnh trên là một biến toàn cục được định nghĩa ở đầu file .C. Mục đích sử dụng của nó là phân biệt giữa các tiến trình với nhau (sẽ trình bày chi tiết về biến **pid** ở **câu 5**). Nếu pid = -1 thì hàm fork() đã gặp phải lỗi là ta gọi **continue** để vòng lặp while(should run) chạy lại từ đầu.
- Sau khi fork, nếu là tiến trình con, ta thực hiện một loạt các thao tác sau:

+ Từ câu lệnh người dùng nhập vào, tự động phát hiện nếu người dùng có ý đồ muốn sử dụng cơ chế pipeline nối tiếp nhiều câu lệnh với nhau (sẽ giải thích ở **câu 4**) và nếu có thì tách câu lệnh đó ra thành nhiều câu lệnh nhỏ phân tách bằng ký tự '|' rồi thêm chúng vào <u>một danh sách lệnh</u>. Nếu không thì vẫn thêm câu lệnh người dùng nhập vào <u>danh sách lệnh</u> trên:

```
// Khúc này là để tách chuỗi command thành nhiều sub-command bằng dấu '|'

char *token = strtok(command, "|");

while (token != NULL && cmd_count < MAX_COMMAND)

{...

206
}
```

+ Tiếp đó, thực thi từng câu lệnh trong danh sách lệnh đã thu thập được:

```
// Khúc này là để chạy từng sub-command. Nếu có nhiều hơn 1 sub-command thì cơ chế pipe sẽ có hiệu lực int prevfd = STDIN_FILENO; for (int i = 0; i < cmd_count; i++)

{...
243
244
244
245
}

// Khúc này là để chạy từng sub-command. Nếu có nhiều hơn 1 sub-command thì cơ chế pipe sẽ có hiệu lực
int prevfd = STDIN_FILENO; for (int i = 0; i < cmd_count; i++)

{...
}
exit(EXIT_SUCCESS);
```

* Chú thích: Vì vòng lặp for() ở ảnh trên liên quan đến cơ chế pipeline của câu 4 nên ở đây chỉ giải thích sơ lược như sau: điểm mấu chốt để có thể thực thi các câu lệnh trong danh sách lệnh là gọi hàm exec_cmd() được định nghĩa ở bên ngoài hàm main():

```
void exec_cmd(char *command)
   argsCount = 0;
   // Xử lý chuỗi nhập vào
   char *token = strtok(command, " "); // Tách chuỗi ký tự đầu tiên ra khi có dấu cách và lưu và
   while (token != NULL)
        args[argsCount++] = token; // args[i] bằng một chuỗi ký tự ngăn bởi dấu cách hay nói cách
       token = strtok(NULL, " "); // Tiếp tục tách chuỗi đến khi token là ký tự rỗng khi bị cắt
   args[argsCount] = NULL; // Từ cuối cùng sẽ là kết thúc chuỗi từ
    // Nếu lệnh là chuyển hướng đầu vào ra
    int input_fd, output_fd;
    int redirect_input = 0, redirect_output = 0;
    for (int j = 0; j < argsCount; j++)</pre>
        // Dấu hiệu chuyển hướng đầu ra của một tệp
        if (strcmp(args[j], ">") == 0)
            args[j] = NULL;
                               // Đánh dấu vị trí chuyển hướng là NULL
            redirect_output = 1; // Cờ chuyển hướng đầu ra
            if (args[j + 1] == NULL)
            { // Nơi đầu ra được chuyển hướng vào là NULL nghĩa là không có file output.
               printf("Missing output file\n");
                exit(EXIT_FAILURE); // Báo lỗi xong thoát
```

```
void exec_cmd(char *command)

argsCount = 0;

// Xử lý chuỗi nhập vào
char *token = strtok(command, " "); // Tách chuỗi ký tự đầu tiên ra khi có dấu cácl
while (token != NULL)

args[argsCount++] = token; // args[i] bằng một chuỗi ký tự ngăn bởi dấu cách ha
token = strtok(NULL, " "); // Tiếp tục tách chuỗi đến khi token là ký tự rỗng l
args[argsCount] = NULL; // Từ cuối cùng sẽ là kết thúc chuỗi từ
```

Tham số của hàm exec_cmd() sẽ là câu lệnh chúng ta nhập vào shell (nguyên vẹn) hoặc câu lệnh nhỏ được tách ra bởi ký tự '|'. Biến argsCount chính là để đếm số từ có trong câu lệnh đó, các từ ngăn nhau bởi dấu cách. Biến toàn cục args[] đã được tạo bên trên với mục đích là mỗi một từ trong câu lệnh sẽ được lưu vào với dạng một phần tử của của mảng này.

```
17 char *args[MAX_LINE / 2 + 1]; // Mảng để lưu trữ lệnh và các tham số
```

Trong quá trình tách các từ bằng dấu cách, nếu phát hiện ký tự chuyển hướng '<', '>', hàm exec_cmd sẽ có cơ chế xử lý riêng (giải thích ở **câu 3**).

```
for (int j = 0; j < argsCount; j++)
{
    // Dấu hiệu chuyển hướng đầu ra của một tệp
    if (strcmp(args[j], ">") == 0)
{
```

Sau khi đã tách câu lệnh thành các từ và lưu từng từ vào chuỗi args[], thì lúc này phần tử args[0] chính là câu lệnh ta đã nhập với **tham số** là **args[1]** đến **args[n]** \rightarrow Ta dùng hàm **execvp(args[0], args)** để thực hiện lệnh **args[0]** với mảng đối số cho lệnh là **args**. Ta kiểm tra giá trị trả về của execvp, nếu là -1 thì lệnh không tồn tại trong hệ thống \rightarrow tiến trình con được tạo ở hàm main in ra "Command does not..." \rightarrow tiến trình con kết thúc với mã EXIT_FAILURE (mã báo hiệu có lỗi xảy ra)

```
// Chay và kiểm tra lỗi
if (execvp(args[0], args) == -1)
99 {
    printf("Command does not exist.\n");
    exit(EXIT_FAILURE);
102 }
103 }
```

**

- Nếu là tiến trình cha, dùng lệnh waitpid(-1, NULL, 0) để đợi cho tiến trình con thực thi xong:

*Chú thích: waitpid(-1, NULL, 0): Theo quy định của hàm waitpid(), tham số '-1' là đơi tất cả tiến trình khác hoàn thành.

**

Khi con hoàn thành, cha cũng đã ở cuối vòng lặp while(should_run) nên tiếp tục chạy lại vòng lặp từ đầu, và quy trình đợi người dùng nhập lệnh, xử lý lệnh đặc biệt, xử lý pipeline, xử lý chuyển hướng input/output tiếp diễn cho đến khi người dùng nhập lệnh 'exit'.

Hình minh chứng & giải thích kết quả:

```
nguyentiendat-22520226@Nobara_PC:~$ ./LAB6
it007sh> ls
LAB6 LAB6.c cbonsai
it007sh> touch TEST
it007sh> ls
LAB6 LAB6.c TEST cbonsai
it007sh> rm -r TEST
it007sh> ls
LAB6 LAB6.c cbonsai
it007sh> ls
LAB6 LAB6.c cbonsai
it007sh> ls
CAB6 CAB6.c cbonsai
it007sh> exit
Goodbye!nguyentiendat-22520226@Nobara_PC:~$
```

GIẢI THÍCH:

- Câu lệnh đầu tiên nhập vào là lệnh **ls**, tự động phát hiện nếu người dùng có ý đồ muốn sử dụng cơ chế pipeline nối tiếp nhiều câu lệnh với nhau và nếu có thì tách câu lệnh đó ra thành nhiều câu lệnh nhỏ phân tách bằng ký tự '|' rồi thêm chúng vào một danh sách lệnh. Nếu không thì vẫn thêm câu lệnh người dùng nhập vào danh sách lệnh, tiếp đó ta thực thi các câu lệnh trong danh sách lệnh. Ở đây lệnh **ls** sẽ được thêm vào danh sách lệnh. Vì không có các câu lệnh con khác nào ngoài lệnh **ls** thì hê điều hành sẽ thực thi câu lênh **ls** (tức là liệt kê các file ở thự mục hiện hành).
- Tương tự như trên với lệnh touch TEST thì tách câu lệnh đó ra thành nhiều câu lệnh nhỏ rồi thêm chúng vào danh sách lệnh. Ta thực thi các câu lệnh trong danh sách lệnh bằng cách gọi hàm exec_cmd trong vòng lặp với hàm exec_cmd đã được định nghĩa, tham số của hàm exec_cmd() sẽ là câu lệnh chúng ta nhập vào shell (nguyên vẹn) hoặc câu lệnh nhỏ được tách ra bởi ký tự '|'. Biến argsCount chính là để đếm số từ có trong câu lệnh đó, các từ ngăn nhau bởi dấu cách. Biến toàn cục args[] đã được tạo bên trên với mục đích là mỗi một từ trong câu lệnh sẽ được lưu vào với dạng một phần tử của của mảng này. Sau khi đã tách câu lệnh thành các từ và lưu từng từ vào chuỗi args[], thì lúc này phần tử args[0] chính là câu lệnh touch ta đã nhập với tham số là args[1] đến args[n] → Ta dùng hàm execvp(args[0], args) để thực hiện lệnh touch với mảng đối số cho lệnh là TEST. Thì kết quả trả về cho ta là lệnh touch TEST giống như shell nguyên bản.
- rm -r TEST, tương tự như trên thì args[0] là lệnh rm với các đối số truyền vào là r và TEST. Thì kết quả trả về cho ta là lệnh rm -r TEST giống như shell nguyên bản.

2. Câu 2: Tạo tính năng sử dụng lại câu lệnh gần đây nhất

Trình bày giải thuật & giải thích code:

Ở đầu chương trình, ta tạo mảng **HF[HISTORY_SIZE**] để lưu lịch sử các câu lệnh và biến historyCount để đếm các câu lệnh nếu chưa đầy lịch sử (lịch sử chứa tối đa 5 câu lệnh gần nhất):

```
#define HISTORY_SIZE 5

12 #define MAX_LINE 80 // Số ký tự tối đa trên 1 dòng

13 #define MAX_COMMAND 4 // Số command tối đa được sử dụng trong pipe, ví dụ: 'ls | grep a | wc -l' là 3 command

14 #define DEFAULT_PID 1 // id mặc định để gán cho process chính khi hàm fork() chưa được gọi

15

16 char *HF[HISTORY_SIZE]; // Mảng để lưu lịch sử các câu lệnh

17 char *args[MAX_LINE / 2 + 1]; // Mảng để lưu trữ lệnh và các tham số của lệnh

18 int historyCount = 0;

19 int argsCount;
```

Ở trong hàm main, cứ mỗi khi nhập xong 1 câu lệnh

```
char command[MAX_LINE];
fgets(command, MAX_LINE, stdin);  // Nhập câu lệnh
command[strcspn(command, "\n")] = '\0'; // strcspn tìm vị trí đầu tiên có '\n' trong chuỗi;
```

Nếu câu lệnh không phải là 'exit', ta sẽ tiến hành kiểm tra xem người dùng có nhập lệnh 'HF' hay không, cách kiểm tra như sau:

```
// Kiểm tra lệnh đặc biệt HF

char *hfPtr = strstr(command, "HF"); // Trả về con trỏ đến vị trí đầu tiên của chuỗi "HF" trong chuỗi lớn "command"

int hfUsed = hfPtr == NULL ? 0 : 1; // Biến cờ để xác định người dùng đã nhập lệnh HF

if (hfUsed)
{
```

- Dùng hàm **strstr(command, "HF")**: hàm sẽ dò tìm trong command (câu lệnh người dùng nhập) trả về con trỏ tới vị trí đầu tiên của chuỗi "HF" trong câu lệnh đó. Nếu không tìm thấy, hàm trả về NULL. Giá trị trả về của strstr() được lưu vào biến con trỏ **char* hfPtr**.
- Nếu tìm thấy chuỗi "HF" trong command:
 - + Tạo biến **hfCount** với giá trị khởi tạo là 1 (nếu hfCount là 1 thì thực hiện lại câu lệnh gần nhất, nếu là 2 thì thực hiện câu lệnh gần thứ 2, tương tự với 3, 4, 5):

```
144
145 \rightarrow 146 int hfCount = 1;
```

+ Tiếp tục dùng lệnh strstr() để dò tìm từ "HF" trong các ký tự còn lại của câu lệnh (các ký tự chưa dò trước đó vì hàm strstr() dừng lại khi nó tìm được bất kỳ từ "HF" nào trong chuỗi) và cứ lặp lại cho đến khi không còn từ "HF" nào hoặc hfCount đã bằng với historyCount (số phần tử hiện đang có trong mảng lưu lịch sử HF):

```
while ((hfPtr = strstr(hfPtr + strlen("HF"), "HF")) != NULL && hfCount < historyCount) // Tiếp tự hfCount++;
```

*Chú thích: 'hfPtr + strlen("HF")' có ý nghĩa là ta lấy con trỏ đang trỏ tới từ "HF" được tìm thấy trước đó và dịch chuyển nó tới 2 đơn vị(độ dài của từ "HF"). Như vậy, hàm strstr() sẽ bắt đầu tìm kiếm từ "HF" mới ở phân đoạn của command mà chưa được tìm kiếm trước đó.

**

+ Bước cuối cùng là thay đổi giá trị của **command** hiện tại bằng câu lệnh gần nhất thứ n trong lịch sử (n = hfCount), sau đó in ra dòng chữ "it007sh>

command" để báo hiệu cho người dùng dễ nắm bắt được lệnh nào trong lịch sử sắp sửa được thực thi lại

```
149 strcpy(command, HF[historyCount - hfCount]); // Lấy câu lệnh từ
150 printf("it007sh> %s\n", command);
151 }
```

- **Nếu không tìm thấy chuỗi "HF" trong command**, thì **command** mà người dùng nhập vào sẽ được lưu vào lịch sử:

```
// Lưu lịch sử lại câu lệnh vừa nhập (nếu không phải lệnh HF)
else
{
    if (historyCount < HISTORY_SIZE) // Nếu lịch sử còn chỗ trống
        strcpy(HF[historyCount], command);
        historyCount++;
    }
else // Nếu lịch sử đã đẩy
{
        // Xóa lệnh đầu tiên và dịch chuyển các lệnh cũ lên trên
        for (int i = 0; i < HISTORY_SIZE - 1; i++)
        {
              strcpy(HF[i], HF[i + 1]);
        }
        strcpy(HF[HISTORY_SIZE - 1], command); // Lưu lệnh mới nhất vào cuối
        }
}</pre>
```

*Chú thích: Nếu lịch sử đã đầy, ta tiến hành xóa câu lệnh cũ nhất trong lịch sử. Vì mảng HF được tổ chức theo kiểu $FIFO \rightarrow HF[0]$ là lệnh cũ nhất \rightarrow Thực hiện phép dời mảng HF[i] = HF[i+1] để ghi đè lên $HF[0] \rightarrow$ Thêm command mới vào $HF[HISTORY_SIZE-1]$.

- ⇒ Đó là toàn bô quá trình xử lý History Feature của shell. **Tóm tắt** lai các bước xử lý:
 - Trong **command** của người dùng nhập, tìm và đếm số lần chữ "HF" xuất hiện
 - Nếu command xuất hiện n lần (n > 0, n < 5): Thay thế command bằng câu lệnh gần thứ n trong lịch sử. Nếu n = 0, giữ nguyên **command** và lưu vào lịch sử (nếu lịch sử đã đầy thì xoá câu lệnh cũ nhất rồi lưu).
 - Đưa **command** cho các phần tiếp theo của chương trình xử lý.

Hình minh chứng:

```
./vinashell
it007sh> echo hello1
hello1
it007sh> echo hello2
hello2
it007sh> echo hello3
hello3
it007sh> echo hello4
hello4
it007sh> echo hello5
hello5
it007sh> HF
it007sh> echo hello5
hello5
it007sh> HF HF
it007sh> echo hello4
hello4
it007sh> HF HF HF
it007sh> echo hello3
hello3
it007sh> HF HF HF HF
it007sh> echo hello2
hello2
it007sh> HF HF HF HF
it007sh> echo hello1
hello1
it007sh>
```

Nhập n lần HF để quay lại câu lệnh gần thứ n

Giải thích kết quả:

Ở ví dụ trong ảnh trên, ta thực thi một loạt các câu lệnh "echo hello1", "echo hello2", ..., "echo hello5". Sau khi thực thi xong, lúc này trong mảng HF đang chứa 5 câu lệnh theo thứ tư từ cũ nhất đến mới nhất:

HF = ["echo hello1", "echo hello2", "echo hello3", "echo hello4", "echo hello5"]

Có thể thấy khi ta gỗ lệnh "HF", chương trình duyệt lệnh mà ta vừa gỗ và tìm được 1 từ "HF" → thực hiện lại lệnh gần nhất trong lịch sử:

```
it007sh> HF
it007sh> echo hello5
hello5
```

. . .

Tương tự, khi ta gõ lệnh "HF HF HF HF", chương trình tìm thấy 5 từ "HF" và in ra câu lệnh gần thứ 5 trong lịch sử, cũng chính là câu lệnh cũ nhất còn lưu lại vì kích thước tối đa của lich sử là 5:

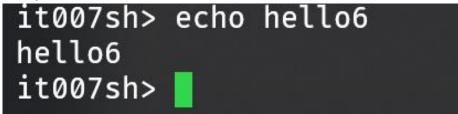
```
it007sh> HF HF HF HF HF it007sh> echo hello1
```

* Bonus:

Ngoài ra, vì chương trình đã được cài đặt để đếm số chữ "HF" không vượt quá kích thước tối đa của mảng HF[]. Nên dù ta có cố tình nhập 6 chữ "HF", chương trình cũng chỉ dừng lai ở chữ "HF" thứ 5:

```
it007sh> HF HF HF HF HF HF it007sh> echo hello1 hello1 it007sh>
```

Cuối cùng, để kiểm tra tính năng xóa lệnh cũ nhất khỏi lịch sử khi mảng HF[] đã đầy, biết rằng lúc này mảng HF[] đang có 5 phần tử. Nếu ta nhập câu lệnh "echo hello6" vào thì chắc chắn câu lệnh cũ nhất ("echo hello1") sẽ bị xóa ra khỏi mảng HF[] để chừa chỗ cho lênh "echo hello6":



Lúc này, mảng HF[] trở thành:

HF = ["echo hello2", "echo hello3", "echo hello4", "echo hello5", "echo hello6"]

Để kiểm chứng xem mảng HF[] có giống như vậy hay không, ta sẽ chạy một loạt các lệnh HF nối chuỗi nhau như ở phần trước để in ra các câu lệnh hiện có trong lịch sử theo thứ tự từ mới nhất đến cũ nhất:

```
it007sh> HF
it007sh> echo hello6
hello6
it007sh> HF HF
it007sh> echo hello5
hello5
it007sh> HF HF HF
it007sh> echo hello4
hello4
it007sh> HF HF HF
it007sh> echo hello3
hello3
it007sh> HF HF HF HF
it007sh> echo hello2
hello2
it007sh>
```

**

3. Câu 3: Chuyển hướng vào ra

Trình bày giải thuật, giải thích code:

Sau khi đã cắt câu lệnh thành các từ như ở câu 1 (được thực hiện ở hàm exec_cmd).

```
void exec_cmd(char *command)

argsCount = 0;

// Xử lý chuỗi nhập vào
char *token = strtok(command, " "); // Tách chuỗi ký tự đầu tiên ra
while (token != NULL)

args[argsCount++] = token; // args[i] bằng một chuỗi ký tự ngăn
token = strtok(NULL, " "); // Tiếp tục tách chuỗi đến khi token

args[argsCount] = NULL; // Từ cuối cùng sẽ là kết thúc chuỗi từ
```

Trước khi thực hiện execvp, ta kiểm tra trong chuỗi args[] có từ nào bằng với toán tử chuyển hướng hay không ('>' và '<'). Nếu có args[j] nào đó là '>' (tức là chuyển hướng đầu ra), ta sẽ đánh dấu vị trí chuyển hướng là NULL. Lúc này cờ chuyển hướng đầu ra sẽ là 1. Nếu args[j+1] là NULL thì câu lệnh này không có file đầu ra vậy nên ta sẽ báo lỗi và thoát. Nếu có tên file đầu ra, chương trình sẽ mở file đó và biến nó thành đầu ra chuẩn mới thay cho đầu ra chuẩn mặc định bằng lệnh dup2.

```
69 dup2(output_fd, STDOUT_FILENO); // Overwrite file STDOUT
```

Tương tự với args[j] là '<' (tức là chuyển hướng đầu vào).

```
Dấu hiệu chuyển hướng đầu ra của một tệp
                 (strcmp(args[j], ">") == 0)
                 args[j] = NULL; // Đánh dấu vị trí chuyển hướng là NULL
55
                 redirect_output = 1; // Cờ chuyển hướng đầu ra
                 if (args[j + 1] == NULL)
                 { // Nơi đầu ra được chuyển hướng vào là NULL nghĩa là không có file output.
                     printf("Missing output file\n");
                     exit(EXIT_FAILURE); // Báo lỗi xong thoát
                 // Tạo file args[i] nếu chưa có; file này sẽ thế chỗ đầu ra cho luồng chuẩn STDOUT
                 output_fd = open(args[j + 1], O_WRONLY | O_CREAT | O_TRUNC, 0644); // output_fd luu
                 if (output_fd < 0)</pre>
                     perror("System fail to execute.");
                     exit(EXIT_FAILURE); // Không mở được file nên báo lỗi rồi thoát
                 dup2(output_fd, STDOUT_FILENO); // Overwrite file STDOUT hiên tai bằng file output_f
                 close(output_fd);
```

Xử lý chuyển hướng đầu ra của một tệp

```
// Dấu hiệu chuyển hướng đầu vào của một tệp
else if (strcmp(args[j], "<") == 0)
   args[j] = NULL;
                        // Đánh dấu vị trí chuyển hướng là NULL
   redirect_input = 1; // Cờ chuyển hướng đầu vào
    if (args[j + 1] == NULL)
    { // Nơi đầu vào được chuyển hướng vào là NULL nghĩa là không có file input.
       printf("Missing input file.\n");
       exit(EXIT_FAILURE); // Báo lỗi xong thoát
    // Tạo file args[i] nếu chưa có; file này sẽ thế chỗ đầu vào cho file STDIN của terminal
   input_fd = open(args[j + 1], O_RDONLY); // input_fd luu tru giá tri file descriptor nên phía trên
   if (input_fd < 0)</pre>
       perror("System fail to execuse.");
        exit(EXIT_FAILURE); // Không mở được file nên báo lỗi rồi thoát
   dup2(input_fd, STDIN_FILENO); // Overwrite file STDIN hiện tại bằng file input_fd vừa tạo, luồng đ
   close(input_fd);
```

Xử lý chuyển hướng đầu vào của một tệp

Chup hình minh chứng:

```
ອ № main *1 ?9
                                                                                                       10:56:38 0
     /inashell
it007sh> cat test.txt
banana
orange
kiwi
hedieuhanh
apple
merry christmas
it007sh> sort < test.txt
apple
banana
hedieuhanh
kiwi
merry christmas
orange
it007sh> wc -l < test.txt
it007sh> echo "Hello world IT007, merry christmas & happy new year" > hello.txt
it007sh> cat hello.txt
"Hello world IT007, merry christmas & happy new year"
it007sh>
```

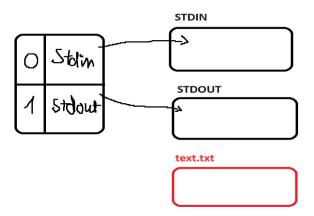
GIẢI THÍCH: (chỉ minh họa câu lệnh đầu, các câu lệnh sau chạy tương tự giải thuật đã trình bày trên)

B1: cắt lệnh thành các từ, ta có args[0] = "sort", args[1] = "<", args[2] = "test.txt"

B2: kiểm tra thấy args[1] là toán tử chuyển hướng đầu vào "<"

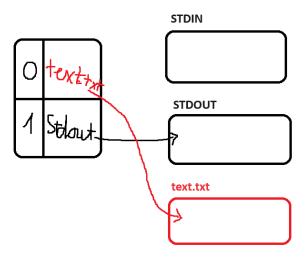
B3: kiểm tra thấy args[1+1] = args[2] = "text.txt" khác rỗng nên có file đầu vào là text.txt

B4: kiểm tra trong hệ thống đã tồn tại file tên args[2] = "text.txt" hay chưa, nếu chưa sẽ tạo file đó và mở file. Do mở file thành công nên không báo lỗi.



B5: tạo biến để lưu file descriptor của file vừa tạo là text.txt

B6: đổi file descriptor mặc định ở đầu đọc dữ liệu thành descriptor của file text.txt



B7: đóng file

Vậy là sau khi thực hiện lệnh, đầu vào lúc này của chương trình sẽ là file text.txt. Do vậy khi thực hiện lệnh sort, sort sẽ nhận đầu vào là dữ liệu sau:

```
it007sh> cat text.txt
banana
orange
kiwi
hedieuhanh
apple
merry christmas
```

Do đó khi sort xong, ta thấy các dữ liệu liệu của file text.txt đã được sắp xếp.

```
it007sh> sort < text.txt
apple
banana
hedieuhanh
kiwi
merry christmas
orange
```

4. Câu 4: Giao tiếp sử dụng cơ chế đường ống

Trình bày giải thuật:

Ở yêu cầu số 1, ta đã giải thích sơ lược về việc tiến trình con ở hàm main, dùng tiến trình con này để thực hiện **command** người dùng nhập; tiến trình cha chờ tiến trình con hoàn thành xong, sẽ thực hiện vòng lặp để xuất hiện dấu nháy. Cứ mỗi lần nhập một câu lệnh là ta sẽ đều tạo tiến trình con, đều xử lý pipeline câu lệnh đó dù nó có sử dụng cơ chế pipeline hay không. Với mỗi **sub-command** trong pipeline như vậy, chúng đều được quyền chuyển hướng đầu vào/ra theo ý muốn. Vì thế, ta phải đưa các từng **sub-command** này vào hàm **exec_cmd**() để xử lý chuyển hướng vào/ra riêng biệt cũng như thực thi chúng.

Bắt đầu giải thích từ hàm main, ta tạo một tiến trình con để xử lý lệnh nhập vào. Tiến trình này sẽ tách lệnh nhập từ input thành các đoạn lệnh nhỏ hơn với '|' là dấu ngăn cách và lưu vào chuỗi cmd_list. Nếu câu lệnh không có dấu '|' thì chuỗi cmd_list sẽ chỉ có 1 phần tử.

Ví dụ: 'ls | wc -l' khi được nhập vào sẽ có cmd_list[0] là 'ls' và cmd_list[1] là 'wc -l'. Nếu câu lệnh nhập vào chỉ có 'ls' thì cmd_list chỉ có 1 phần tử là cmd_list[0] = 'ls'.

```
// Tiến trình con
if (pid == 0)
   char cmd_list[MAX_COMMAND][MAX_LINE]; // Mảng để chứa các câu lệnh phân tách bằng
   int cmd_count = 0;
   // Khúc này là để tách chuỗi command thành nhiều sub-command bằng dấu '|'
   char *token = strtok(command, "|");
   while (token != NULL && cmd count < MAX COMMAND)
        // Remove leading and trailing spaces from the token
       while (*token == ' ' || *token == '\t')
           token++;
       int length = strlen(token);
       while (length > 0 && (token[length - 1] == ' ' || token[length - 1] == '\t'))
           token[length - 1] = '\0';
           length--;
        strcpy(cmd_list[cmd_count++], token); // luu subcommand vào câu lệnh cmd_list
        token = strtok(NULL, "|");
                                            // tìm subcommand tiếp theo
```

Khi tách các câu lệnh ngăn bởi dấu '|' (sub-command) ta cũng phải bỏ đi khoảng trắng ở phía trước các sub-command đó.

Sau khi đã tách thành các sub-command, ta gán cho biến prevfd bằng với đầu vào mặc định. Điều này là để sub-command đầu tiên sẽ có đầu vào mặc định. Khi thực hiện xong tiến trình, các prevfd trừ sub-command đầu tiên đều sẽ bị đổi đầu STDIN mặc định thành đầu pipefd[0]. Các sub-command trừ sub-command cuối cùng sẽ bị đổi đầu STDOUT mặc định thành đầu pipefd[1].

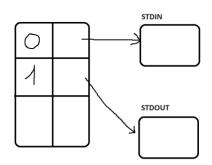
```
int prevfd = STDIN_FILENO;
```

Ta tạo vòng lặp với mỗi sub-command có được.

Trong mỗi sub-command ta sẽ phải chỉnh 2 thành phần đó là stdin và stdout của sub-command đó. Tạo một ống pipefd[2], đầu đọc là pipefd[0], đầu viết là pipefd[1]. Báo lỗi nếu không tạo được ống.

Ta tạo một tiến trình con để thực hiện điều chỉnh pipefd[0] và pipefd[1], tiến trình cha sẽ dùng để đổi giá trị prefd, cho sub-command sau dùng. Như vậy sub-command đầu tiên sẽ có STDIN mặc định, sub-command cuối cùng sẽ có STDOUT mặc định.

Nếu chỉ có một sub-command, lệnh dup2(prefd, STDIN_FILENO) sẽ không làm thay đổi đầu vào của sub-command này; do prefd đầu tiên đặt ngoài vòng lặp có giá trị là STDIN_FILENO. Lệnh dup2(pipefd[1], STDOUT_FILENO) cũng sẽ không được thực hiện. điều này khiến cho sub-command duy nhất được nhập này có đầu vào và đầu ra như hình bên dưới.



Với đầu vào và ra như trên, đây trở thành một câu lệnh bình thường mà không có ống dẫn, được thực hiện hàm exec_cmd bình thường (trong exec_cmd có hỗ trợ chuyển hướng đọc ở câu 2).

Khi có nhiều hơn một sub-command, điều này cũng có nghĩa trong câu lệnh nhập vào có ống dẫn.

Lúc này sub-command đầu tiên vẫn sẽ có đầu vào mặc định. Tuy nhiên đầu ra sẽ được chuyển vào ống viết pipefd[1]. Tiến trình cha của sub-command đầu tiên sẽ đổi prevfd thành đầu ống đọc pipefd[0].

```
229
230
231
231
232
232
233
234
235
else
{
    close(pipefd[1]);
    close(prevfd);
    prevfd = pipefd[0];
    waitpid(-1, NULL, 0); // Chờ tiến trình con hi sinh
}
```

Vòng lặp for làm cho tiến trình của sub-command thứ 2 được chạy.

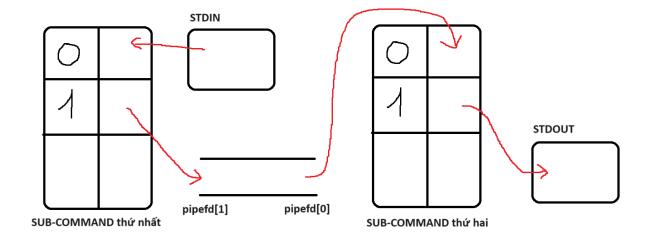
Lúc này, dup2(prevfd, STDIN_FILENO) sẽ làm cho sub-command thứ hai đổi file đầu vào mặc định thành ống pipefd[0]; do prevfd đã được đổi nhờ tiến trình cha của sub-command thứ nhất. Còn câu lệnh dup2(pipefd[1], STDOUT_FILENO) sẽ không được thực hiện do đã là sub-command cuối cùng.

```
else if (cpid == 0)

close(pipefd[0]);
dup2(prevfd, STDIN_FILENO);
if (i < cmd_count - 1)
dup2(pipefd[1], STDOUT_FILENO); // Chuyển hướng đầu ra của tất cả cmd trừ cmd cuối

exec_cmd(cmd_list[i]);
close(pipefd[1]);
exit(EXIT_SUCCESS); // Tiến trình con hi sinh sau khi chạy xong cmd
```

Kết quả sẽ như hình minh họa bên dưới:



Chup hình minh chứng & giải thích kết quả:

```
> ./vinashell
it007sh> sudo apt install cowsay
[sudo] password for ducmint:
Reading package lists ... Done
Building dependency tree ... Done
Reading state information ... Done
cowsay is already the newest version (3.03+dfsg2-8).
The following packages were automatically installed and are no longer required:
   python-pkg-resources python-setuptools
Use 'sudo apt autoremove' to remove them.
0 upgraded, 0 newly installed, 0 to remove and 195 not upgraded.
it007sh>
```

Tải chương trình 'cowsay'

```
11:10:09 0
  ./vinashell
it007sh> ls -l
total 40
-rw-r-r- 1 ducmint ducmint
                                  54 Dec 20 10:58 hello.txt
-rw-r-r- 1 ducmint ducmint 52 Dec 20 10:53 test.txt
-rwxr-xr-x 1 ducmint ducmint 17256 Dec 20 11:05 vinashell
 -rw-r-r- 1 ducmint ducmint 9918 Dec 20 11:04 vinashell.c
it007sh> ls -l | grep a
 -rwxr-xr-x 1 ducmint ducmint 17256 Dec 20 11:05 vinashell
 -rw-r—r— 1 ducmint ducmint 9918 Dec 20 11:04 vinashell.c
it007sh> ls -l | grep a
 -rwxr-xr-x 1 ducmint ducmint 17256 Dec 20 11:05 vinashell
 -rw-r-r- 1 ducmint ducmint 9918 Dec 20 11:04 vinashell.c
it007sh> cat hello.txt
"Hello world IT007. merry christmas & happy new year"
it007sh> cat hello.txt | cowsay
  "Hello world IT007, merry christmas &
  happy new year"
             (00)
it007sh> pwd | ls | grep h
hello.txt
vinashell
vinashell.c
```

Chạy các câu lệnh

*Bonus:

Chạy các câu lệnh

GIẢI THÍCH:

- Lấy ví dụ với câu lệnh **cat hello.txt** | **cow say**, theo như giải thích thuật toán ở phía trên, ta có:
 - 1. Chương trình sẽ tách input và chia tách dấu "|" thành các command và tham số:
 - Command 1: "cat"
 - Tham số của command 1: "hello.txt"
 - Command 2: "cowsay"
 - 2. Chương trình sẽ tạo một tiến trình con để thực hiện command "cat hello.txt" và chuyển output của command này thành input cho command "cowsay".
 - 3. Tiến trình con sẽ thực hiện command "cat hello.txt" để đọc nội dung của file "hello.txt".
 - 4. Output của command "cat hello.txt" sẽ được chuyển vào command "cowsay".

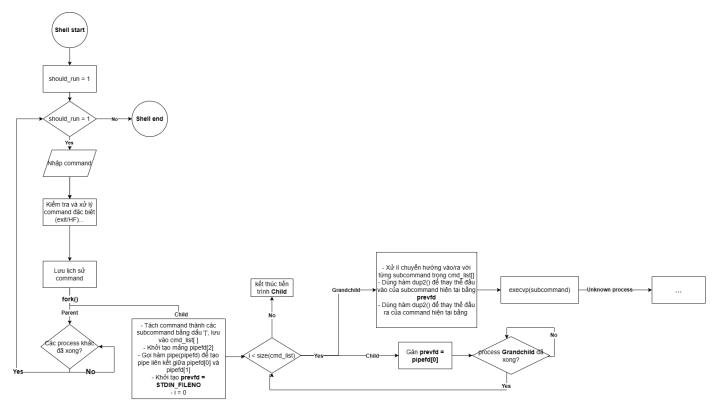
- 5. Command "**cowsay**" sẽ nhận input từ output của command trước đó và thực hiện công việc của nó. Output của ta sẽ là kết quả của lệnh **cowsay** với input là output của **cat hello.txt**.
- Lấy ví dụ với câu lệnh **pwd** | **ls** | **grep h**, theo như giải thích thuật toán ở phía trên, ta có:
 - 1. Chương trình sẽ tách input và chia tách dấu "|" thành các command và tham số:
 - Command 1: "pwd"
 - Command 2: "Is"
 - Command 3: "grep"
 - Tham số của command 3: "h"
 - 2. Chương trình sẽ tạo một tiến trình con để thực hiện command "**pwd**" và chuyển output của command này thành input cho command "**ls**".
 - 3. Tiến trình con sẽ thực hiện command "**pwd**" để xem thư mục hiện hành.
 - 4. Output của command "pwd" sẽ được chuyển vào command "ls".
 - 5. Command "**ls**" sẽ nhận input từ output của command trước đó và thực hiện công việc của nó. Nó sẽ thực hiện câu lệnh "**ls**" với input truyền vào là output của câu lệnh "**pwd**". Output lúc này sẽ được chuyển tới câu lệnh **grep.**
 - 6. Tiếp tới câu lệnh "grep" với input là **h** và output của bước 5, ta sẽ có kết quả của toàn bộ câu lệnh là "xuất ra các dòng/các file có ký tự "h" của thư mục hiện hành"

CÁC CÂU LỆNH CÒN LẠI CÁCH THỰC CHẠY TƯƠNG TỰ

5. Câu 5: Kết thúc lệnh đang thực thi

Trình bày giải thuật & giải thích code:

Để khái quát lại cách chạy của chương trình shell từ câu $1 \rightarrow 4$, ta có lưu đồ sau:



Lưu đồ giải thuật tổng quan câu 1 - 4

Nhìn vào lưu đồ trên, ta có thể nhận xét được rằng, khi người dùng gửi tín hiệu SIGINT bằng cách nhấn Ctrl + C, cơ chế xử lý mặc định của hệ điều hành sẽ kết thúc tất cả các tiến trình **Parent, Child, Grandchild/Unknown process**. Trong khi đó, tiến trình mà ta mong muốn kết thúc chỉ có **Unknown process**.

⇒ Ý tưởng giải thuật cho yêu cầu 5:

May mắn thay, ta có thể thay đổi cơ chế xử lý tín hiệu SIGINT của từng process bằng cách gọi hàm **signal(sig, func)** với **func** là những thao tác mà ta muốn tiến trình đó thực hiện khi nhận được **sig**. Lệnh signal() này sẽ được đặt ở đầu hàm main(), chính vì vậy nó xuất hiện trước toàn bộ các lệnh fork(), nên tất cả các tiến trình mới được sinh ra đều đã gọi hàm signal():

Hàm signal với tham số truyền vào là SIGINT (nhận tín hiệu Ctrl + C) và hàm sigint_handler

```
// Hàm để kill tiến trình con
void sigint_handler()

{
    if (pid == 0)
    {
        // Nếu là tiến trình con, in ra dòng chữ này
        printf("Ctrl + C pressed, quitting program...\n");
}
```

Hàm sigint_handler

Chi tiết hơn, ở giai đoạn <u>tiền thực thi</u> các sub-command (lúc người dùng chưa input command), chương trình chỉ có 1 tiến trình hiện hữu là Parent. Khi Parent nhận nhận được SIGINT, nó không phản hồi.

Sau đó, ở giai đoạn <u>thực thi</u> các sub-command (sau khi người dùng đã input command), các tiến trình hiện hữu là Parent, Child, GrandChild. Khi nhận được SIGINT:

- Tiến trình Parent: không phản hồi
- Tiến trình Child: in ra dòng chữ "Ctrl + C pressed..." (với ý nghĩa thông báo cho người dùng).
- Tiến trình Grandchild: Grandchild nhanh chóng gọi execvp(). Lúc này, code của nó sẽ bị thay thế thành code của sub-command và nó trở thành Unknown process. Dựa vào cơ chế mặc định của hệ điều hành thì chương trình Unknown process này bi kết thúc.

Bên cạnh đó, để hàm **sigint_handler** đã giới thiệu phía trên phân biệt giữa các tiến trình với nhau, ta dùng một biến toàn cục **pid_t pid** như sau:

```
#define DEFAULT_PID 1 // id mặc định để gán cho process chính khi hàm fork() chưa được gọi

pid_t pid = DEFAULT_PID;

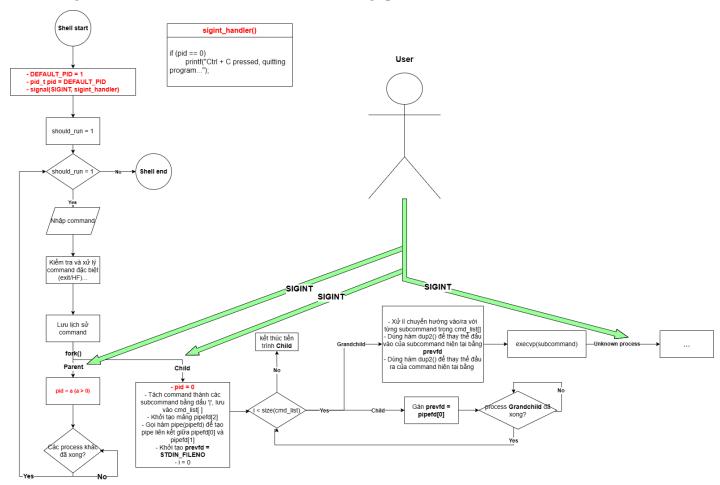
pid = fork();
```

Như vậy, **pid** của mỗi tiến trình trong quá trình chạy sẽ là:

- Ở giai đoạn tiền thực thi các sub-command:

- + Tiến trình Parent: (lý do gán với DEFAULT_PID là vì **pid** có giá trị mặc định là 0, điều này sẽ làm cho Parent in ra 'Ctrl + C pressed...' không mong muốn).
- Ở giai đoạn thực thi các sub-command:
 - + Tiến trình Parent: một số lớn hơn 0
 - + Tiến trình Child: 0
 - + Tiến trình Grandchild: 0 (vì Grandchild sinh ra từ Child nên copy lại biến **pid** của Child) nhưng nhanh chóng biến thành Unknown Process.

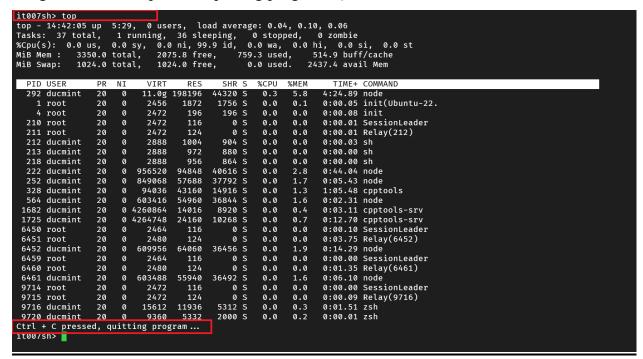
Tóm lại, giải thuật câu 5 có thể được biểu diễn thông qua lưu đồ sau:



Lưu đồ giải thuật của câu 5

Chụp hình minh chứng & giải thích kết quả:

Chạy lệnh 'top' để hiện thị một giao diện cho biết các tiến trình đang chạy trong hệ thống. Để thoát khỏi giao diện top và trở về shell, ta nhấn Ctrl + C (có thể thấy dòng chữ 'Ctrl + C pressed, quitting program..').



Nhấn Ctrl + C khi chương trình 'top' đang chạy

- Tiếp đến ta thử với lệnh 'cbonsai –live'. 'cbonsai' là chương trình tự động vẽ cây bonsai bằng các ký tự ngay trong terminal:



chương trình chonsai

```
it007sh> sudo apt install cbonsai
[sudo] password for ducmint:
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
cbonsai is already the newest version (1.2.1-1).
The following packages were automatically installed and are no longer required:
   python-pkg-resources python-setuptools
Use 'sudo apt autoremove' to remove them.
0 upgraded, 0 newly installed, 0 to remove and 195 not upgraded.
it007sh>
```

Tải chương trình chonsai

```
it007sh> cbonsai —live

Ctrl + C pressed, quitting program...
it007sh>
```

Nhấn Ctrl + C để thoát khỏi chonsai và trở lại shell

- Tương tự, thử với lệnh sort. Khi sort một file rất lớn, chương trình sort sẽ tốn nhiều thời gian và người dùng không thể thao tác với shell khi sort đang chạy. Ta có thể bấm tổ hợp phím Ctrl + C để ngắt ngang sort. Cùng lúc đó, tiến trình Parent sẽ không phản hồi, còn tiến trình Child sẽ in ra dòng chữ "Ctrl + C pressed, …".

```
it007sh> sort < very_large_file.txt
"CCtrl + C pressed, quitting program ...
it007sh>
```

Ngắt chương trình sort bằng tổ hợp Ctrl + C

Ở giai đoạn tiền thực thi (khi chưa input command), chương trình chỉ có 1 tiến trình duy nhất là tiến trình Parent mà tiến trình này được cài đặt để không có bất kỳ phản hồi nào với tín hiệu SIGINT nên nếu ta bấm Ctrl + C thì sẽ không có chuyện gì xảy ra:

Nhấn Ctrl + C khi chưa có lệnh nào được thực thi

Xin cảm ơn thầy đã đọc $\mathring{N}(@ - @) / H \acute{E} T$