

Unix 实验报告

实验: 实验2目录树的遍历

专业: 计算机科学与技术

班级: 1 班

姓名: 姚怀聿

学号: 22920202204632

2022年10月21日

目 录

一 、	实验内容描述	2
_,	设计、实验构思	3
	构思:	3
	功能 1 的构思	3
	功能 2 的构思	4
	功能 3 的构思	5
	实现细节:	6
	功能1实现细节	8
	功能 2 实现细节	8
	功能3实现细节	10
三、	实验结果	11
	测试功能 1	12
	测试功能 2	12
	测试功能 3	13
四、	体会和建议	13
五、	完成人姓名及完成时间	14

一、 实验内容描述

本实验的目的是掌握与文件和目录树有关的系统调用和库函数。。

实验要求编写程序 myfind;

命令语法为:

- ./myfind <pathname> [-comp <filename> | -name <str>...]
 具体要实现的功能如下:
- 1. myfind <pathname> 的功能:

除了具有如程序 4-22 运行后所打印出的内容以外,还要输出在〈pathname〉目录子树下,文件长度不大于 4096 字节的常规文件在所有允许访问的普通文件中所占的百分比。程序不允许打印出任何路径名。

2. myfind <pathname> -comp <filename>的功能:

命令仅仅输出在〈pathname〉目录子树之下,所有与〈filename〉文件内容一致的文件的绝对路径名。不允许输出任何其它的路径名,包括不可访问的路径名。其中,〈filename〉是常规文件的路径名(非目录名,但是其路径可以包含目录)。

3. myfind <pathname> -name <str>····的功能:

命令输出在〈pathname〉目录子树之下,所有与〈str〉···序列中文件名相同的文件的绝对路径名。不允许输出不可访问的或无关的路径名。其中,〈str〉···是一个以空格分隔的文件名序列(不带路径)。

程序在实现时的要求如下:

- 1. 在以上三个功能实现时,〈pathname〉和〈filename〉均既可以是绝对路径名,也可以是相对路径名。〈pathname〉既可以是目录,也可以是文件,此时,目录为当前工作目录。
- 2. 尽可能地提高程序的效率
- 3. 避免因打开太多文件而产生的错误
- 4. 遍历目录树时,访问结点(目录项)的具体操作应当由遍 历函数 dopath 携带的函数指针参数决定

二、设计、实验构思

构思:

功能1的构思

功能一是本次实验三个功能中实现起来最简单的一个功能,只需要在程序 4-22 的基础上增加一个统计〈pathname〉目录子树下,文件长度不大于 4096 字节的常规文件的个数占所有允许访问的普通文件的百分比即可,具体实现时,我在全局增加了一个 int 类型的变量 nlit,用于记录遍历到的所有文件长度不大于 4096 字节的常规文件的个数。然后在 main 函数中用该个数除以所有允许访问的常规文件的总数量,得到功能要求的百分比,最后在 main 函数中输出。程序具体实现如下:

```
if(argc == 2) {
    ret = myftnd(argv[1], myfunc1, TYPE_P);
    ntot = nreg + ndtr + nblk + nchr + nftfo + nslink + nsock;
    if(ntot == 0) ntot = 1; // 遊免除数为0 所有文件所占的百分比部为0
    fprintf(stdout, "regular files = %5ld %9.2f %%\n", ndtr, ndtr * 100.0 / ntot);
    fprintf(stdout, "directories = x$ld %9.2f %%\n", ndir, ndir * 100.0 / ntot);
    fprintf(stdout, "block special = %5ld %9.2f %%\n", nblk, nblk * 100.0 / ntot);
    fprintf(stdout, "files = x$ld %9.2f %%\n", nfirfo, nfirfo * 100.0 / ntot);
    fprintf(stdout, "FIPOs = x$ld %9.2f %%\n", nfirfo, nfirfo * 100.0 / ntot);
    fprintf(stdout, "symbolic links = x$ld %9.2f %%\n", nslnk, nslink * 100.0 / ntot);
    fprintf(stdout, "sockets = x$ld %9.2f %%\n", nsock, nsock * 100.0 / ntot);
    fprintf(stdout, "scakets = x$ld %9.2f %%\n", nslink * 100.0 / ntot);
    fprintf(stdout, "scakets = x$ld %9.2f %%\n", nslink * 100.0 / ntot);
    fprintf(stdout, "scakets = x$ld %9.2f %%\n", nslink * 100.0 / ntot);
    fprintf(stdout, "scakets = x$ld %9.2f %%\n", nslink * 100.0 / ntot);
    fprintf(stdout, "scakets = x$ld %9.2f %%\n", nslink * 100.0 / ntot);
    fprintf(stdout, "scakets = x$ld %9.2f %%\n", nslink * 100.0 / ntot);
    fprintf(stdout, "scakets = x$ld %9.2f %%\n", nslink * 100.0 / ntot);
    fprintf(stdout, "scakets = x$ld %9.2f %%\n", nslink * 100.0 / ntot);
    fprintf(stdout, "scakets = x$ld %9.2f %%\n", nslink * 100.0 / ntot);
    fprintf(stdout, "scakets = x$ld %9.2f %%\n", nslink * 100.0 / ntot);
    fprintf(stdout, "scakets = x$ld %9.2f %%\n", nslink * 100.0 / ntot);
    fprintf(stdout, "scakets = x$ld %9.2f %%\n", nslink * 100.0 / ntot);
    fprintf(stdout, "scakets = x$ld %9.2f %%\n", nslink * 100.0 / ntot);
    fprintf(stdout, "scakets = x$ld %9.2f %%\n", nslink * 100.0 / ntot);
    fprintf(stdout, "scakets = x$ld %9.2f %%\n", nslink * 100.0 / ntot);
    fprintf(stdout, "scakets = x$ld %9.2f %\n", nslink * 100.0 / ntot);
    fprintf(stdout, "scakets = x$ld %9.2f %\n", nslink * 100.0 / ntot);
    fprintf(stdout, "scakets = x$ld % ystakets = x$ld % ysta
```

功能 2 的构思

功能2要求输出所有在〈pathname〉目录子树下,与〈filename〉 文件中的内容相同的所有文件的绝对路径。本功能我认为是 本次实验中,三个功能中最难以实现的一个功能,对于该功 能的实现,综合了上一次实验中的读、写文件。具体思路如 下:

新写一个 myfunc2 函数,在该函数中处理。函数 myfunc2 的函数原型和整体架构与 myfunc1 相同,其不同之处在于对文件和目录的处理上的不同。进入 myfunc2 函数,首先判断传入的是什么类型,这里只会有 3 个类型,分别是 FIND_F、FIND_DNR、FIND_NS,分别代表"文件"、"目录不能打开"、"无法 stat",这里不会有目录类型,如果是目录类型,则会在 dopath 函数中递归处理。

这其中最主要的是处理 FIND_F 类型,一旦发现该类型,说明递归到达叶子结点,我们需要对这个文件进行读写操作,然

后与之前在 main 函数中读入的〈filename〉中的内容做比较,如果相同,就输出该文件的绝对路径。

这其中,最麻烦的就是输出该文件的绝对路径这一功能;要得到该文件的绝对路径,在递归到该叶子结点之前,我们需要先记录下当前递归处理的文件所在的目录名称。每当递归到叶子结点的时候,就调用 chdir 函数将工作目录更改为 `fullpath`除去最后的文件名的路径;如果叶子结点的文件内容与〈filename〉文件内容相同,就调用 getcwd 函数得到当前工作目录的绝对路径,在回溯的时候重新将工作路径改变回之前的。实现的时候还要注意处理字符串,将当前文件的名称附加在绝对路径的最后输出。程序具体实现如下:

功能3的构思

功能 3 的实现相比功能 2 要更加简单一些,只需要在遍历到叶子结点文件的时候判断这个文件名与一开始输入的 str 序列中有没有名称相同的即可,一旦找到一个名称相同的文件,就直接输出这个文件的绝对路径(得到绝对路径方法与实现功能 2 大致相同),并直接退出循环。myfunc3 具体程序如下:

实现细节:

整体上实现时,适当定义全局变量,由于要实现三个功能,故定义了三个全局 int 变量,分别表示一个功能,基于程序 4-22 修改了 dopath 的函数原型,需要向 dopath 函数中传入两个参数,一个是使用的函数,另一个是当前的类型。程序截图如下:

```
17 #define TYPE_P 1
18 #define TYPE_C 2
19 #define TYPE_N 3 分别表示功能1、2、3
```

在 main 函数中判断应该调用哪种函数,程序截图如下:

```
if(argc == 2) {
    ret = myfind(argv[1], myfunc1, TYPE_P);
    ret = myfind(argv[1], myfunc1, TYPE_P;
    ret = myfind(argv[1], myfunc1, myfunc1, mock, nock, noc
```

其中,功能2的情况下, argc 应该等于4, 否则报错。

实现三个功能时都需要判断 pathname 是否是一个路径名,如果不是,则报错。

功能1实现细节

功能 1 的实现未涉及到太多细节,需要注意的是 nlit 不需要加到总数上。

功能2实现细节

功能2实现细节较多,分几点阐述:

(1) 判断〈filename〉是否是一个文件名,如果不是,则报错。

```
strcpy(c_filename, argv[3]);
//printf("2036 c_filename: %s\n", c_filename);
int fd = open(argv[3], O_RDONLY);
if(lstat(c_filename, &c_filestat) < 0) perror("stat()"), exit(1);
if(S_ISDIR(c_filestat.st_mode)) {
    fprintf(stderr, "%s is a directory, boolst should be a filename\n", c_filename);
    exit(1);
}
if((c_filelen = read(fd, c_filebuf, BUFSIZE)) < 0) {
    perror("read() );
    exit(1);
```

(2) dopath 函数:

如果是 FIND_D、FIND_NS、FIND_DNR 表示到达叶子结点;剩下一种则是 FIND_D,表示是还是一个目录,递归处理。

递归处理时注意不要处理`../`和`./`文件夹,这会产生回路,使用一个函数 path_noloop 判断是否会产生回路。

(3) path noloop 函数实现细节:

(4) 如果是 FIND_DNR 或者 FIND_NS

直接报错即可

(5) 如果是 FIND_F

表示是文件, 需要进一步处理。

在main函数中先将<filename>文件中的内容读入c_filebuf中

在 myfunc2 函数中将遍历到的文件中的内容读入 fullbuf 中

```
lf(type == FIND_F) {
   int fd;
   char* fullbuf = (char*)malloc(stzeof(char) * c_ftlelen); // 开辟一块和c_ftlelen一样大小的空间 用于存放从文件读入的数据
   char* tuppath = (char*)malloc(stzeof(char) * PATHSIZE);
   char* curpath = (char*)malloc(stzeof(char) * PATHSIZE);
   memset(fullbuf, 0, stzeof(fullbuf));
   if((fd = open(pathname, 0_RDONLY)) < 0) { // 将文件打开 并将文件描述符放在fd中
        fprintf(stderr, "%s: ", pathname);
        perror("");
   }
   if(read(fd, fullbuf, c_ftlelen) < 0) {
        fprintf(stderr, "%s: ", pathname);
        perror("");
   }
}</pre>
```

判断 c_filebuf 和 fullbuf 中的内容是否相同,若相同,则 更改当前的工作目录为 fullpath 中除去最后一个文件名的 目录,然后取得当前工作目录的绝对路径,最后改变回原样。

功能3实现细节

功能 3 在处理绝对路径时的细节与功能 2 实现时大致相同,这里不再赘述。

功能 3 输入进来的 str 是一个序列,里面存放的是一系列的 文件名,在 main 函数中处理该序列,将它们放入全局的一个 字符串数组中去。

全局字符串数组:

32 static char name_str[STRNUM][NAMESIZE];

对于递归遍历到的每一个文件名称,将它放入 namebuf 中去

```
if(type == FINO_F) { // 是文件
  int fd, i;
  char* namebuf = (char*)malloc(stzeof(char) * NAMESIZE); // 开辟一块和NAMESIZE一样大小的空间 用于存放文件名称
  char* tmppath = (char*)malloc(stzeof(char) * PATHSIZE);
  char* curpath = (char*)malloc(stzeof(char) * PATHSIZE);
  const_char* obs;
  pos = strrchr(pathname, '/');
  strcpy(namebuf, pos + 1);
```

namebuf 与全局定义的 name_str 字符串数组中的每一项比较,一旦找到相同的,就将该文件的绝对路径输出,然后跳

出循环、继续递归到下一个结点。

```
for(i = 0; i < str_num; i++) {
    if(strcmp(namebuf, name_str[i]) == 0) {
        char* pos2;
        strcpy(curpath, fullpath);
    pos2 = strrchr(curpath, '/');
    *pos2 = 0;
    if((tmppath = getcwd(NULL, 0)) == NULL) perror("getcwd(");
    if(chdir(curpath) < 0) perror("chdir()");
    if(cdir(tmppath) < 0) perror("chdir()");
    if(chdir(tmppath) < 0) perror("chdir()");
    fprintf(stdout, "%s%s\n", curpath, pos);
    break;
}

free(namebuf);
free(tmppath);
free(curpath);</pre>
```

三、 实验结果

源程序名	可执行程序名
myfind.c	myfind

编译生成可执行文件:

使用如下指令:

`make myfind`, 即可得到可执行文件`myfind`

```
rongrong@rongrong-virtual-machine:~/Desktop/huaiyuyao/unix/homework/homework_2$ ls
myfind.c myfind_glob.c
rongrong@rongrong-virtual-machine:~/Desktop/huaiyuyao/unix/homework/homework_2$ make myfind
cc myfind.c -o myfind
rongrong@rongrong-virtual-machine:~/Desktop/huaiyuyao/unix/homework/homework_2$ ls
myfind myfind.c myfind_glob.c
rongrong@rongrong-virtual-machine:~/Desktop/huaiyuyao/unix/homework/homework_2$
```

运行程序:

测试功能1

```
esktop/huaiyuyao/unix/homework/homework_2$ ./myfind ..
regular files
                            80.00 %
20.00 %
directories
                             0.00 %
0.00 %
block special
                      0
char special
                             0.00
                             0.00 %
FIF0s
                      0
symbolic links
                             0.00
sockets
                      0
                             0.00 %
                            33.33 %
<=4096Bytes
rongrong@rongrong-virtual-machine:~/Desktop/huaiyuyao/unix/homework/homework_2$ ./myfind ../..
regular files = 12
                            75.00 %
                            25.00 %
directories
                   4
0
0
0
block special
                             0.00 %
char special
                             0.00 %
FIF0s
                             0.00 %
symbolic links =
                             0.00 %
sockets
                             0.00 %
<=4096Bytes
                            31.25 %
rongrong@rongrong-virtual-machine:~/Desktop/huaiyuyao/unix/homework/homework_2$ ./myfind .
regular files
                             75.00 %
directories
                             25.00 %
block special
                             0.00 %
char special
                             0.00 %
FIF0s
                             0.00 %
symbolic links
                             0.00
sockets
                             0.00
```

经检验,该功能无误。

测试功能 2

首先使用 tree 命令看一下当前目录的结构:

其中 backup. c 和 myfind_glob. c 两文件中的内容相同, backup. c 是由 myfind_glob. c 文件拷贝到上一级目录中去的。



使用`./myfind .. -comp ../backup.c`

结果无误。

使用`./myfind .. -comp myfind_glob.c`

```
rongrong@rongrong-virtual-machine:~/Desktop/hualyuyao/unix/homework/homework_2$ ./myfind .. -comp myfind_glob.c
/home/rongrong/Desktop/hualyuyao/unix/homework/backup.c
/home/rongrong/Desktop/hualyuyao/unix/homework/homework_2/myfind_glob.c
```

结果无误。

```
rongrong@rongrong-vtrtual-machine:-/Desktop/huaiyuyao/unix/homework/homework_2$ ./myfind ../. .comp myfind_glob.c
/home/rongrong/Desktop/huaiyuyao/unix/homework/backup.c
/home/rongrong/Desktop/huaiyuyao/unix/homework/bomework_2/myfind_glob.c
rongrong@rong-virtual-machine:-/Desktop/huaiyuyao/unix/homework/homework_2$ ./myfind ../. .comp ../homework_2/myfind_glob.c
/home/rongrong/Desktop/huaiyuyao/unix/homework/backup.c
/home/rongrong/Desktop/huaiyuyao/unix/homework/backup.c
```

多次测试,结果均无误。

测试功能3

```
rongrong@rongrong-virtual-machine:~/Desktop/huaiyuyao/unix/homework$ ./homework_2/myfind ../.. -name myfind_glob.c
/home/rongrong/Desktop/huaiyuyao/unix/homework/homework_2/myfind_glob.c
rongrong@rongrong-virtual-machine:~/Desktop/huaiyuyao/unix/homework$ ./homework_2/myfind ../.. -name backup.c
/home/rongrong/Desktop/huaiyuyao/unix/homework/backup.c
```

多次测试,结果均无误。

四、 体会和建议

体会:本次实验大量运用了C语言中字符串处理的知识,让我感受到打好C语言的基础是多么的重要!以及这次的实验使用到了一个新的函数 getcwd,这个函数一开始用的时候不知道该怎么传参,得到什么内容,我使用 man 手册去查看该函数的使用方法,才知道这个函数的正确使用姿势。本次实验基本上只更改了 myfunc,而对其他的函数没有做太大的修改,这让我意识到程序的可拓展性也是很重要的。以及动态

分配内存要记得及时释放,否则极易造成内存泄露,这都是以前知道的知识,但是在这次实验中才真正感受到这些知识在实践中的重要性。

建议:希望以后上机课之前,老师能提前说明实验课需要做什么,比如要用到哪些知识点,以及书上哪些程序对本次实验是比较有帮助的,我们可以通过看书本上的例程去更好的实现实验要求。以及可能要用到的不熟悉的函数能为我们做详细的解释。

五、 完成人姓名及完成时间

完成人姓名	完成时间
姚怀聿	2022年10月24日