

# 操作系统实验（二）

南京大学软件学院

2016.4

## 实验重点

本次作业重点：熟悉掌握 Fat12 文件系统，*gcc + nasm* 联合编译实践以及了解实模式与保护模式的基本内容。

## 1 实验内容

### 1.1 编程读取 FAT12 文件

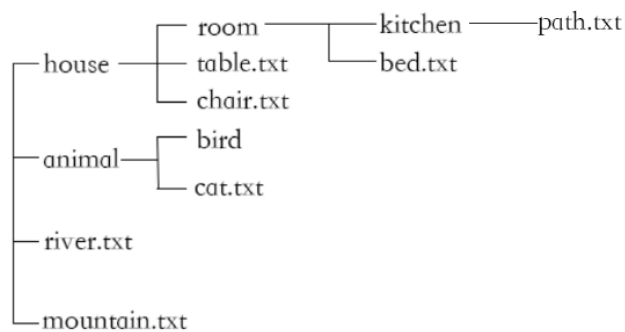
编写程序 `read_fat12`，读取 `a.img` 文件（该文件是 fat12 文件系统的软盘镜像）。

- 1 运行程序后，通过分析 fat12 文件系统，打印出所有文件。
- 2.1 打印完成后，要求能够获取用户输入文件路径（以回车结束），程序查询 Fat12 文件，分别对目录文件、普通文件、不存在的文件进行做相应的输出，具体规则见下文描述。
- 2.2 打印完成后，要求用户能够输入的指令“count 路径”，如果“路径”指向一个普通文件或不存在的文件，则输出提示。如果“路径”指向一个目录，则输出此目录和其下所有子目录所包含的文件和子目录数量，输出的顺序和缩进要能反应目录的父子关系。
- 2.1 和 2.2 是地位并列的两种功能，在程序运行和打印出所有文件以后，用户可以以任意顺序使用两种功能。

### 1.2 输入输出示例

比如对于如下目录结构的一个文件：

- 第 1 步的输出应该为：



```

house/room/kitchen/path.txt
house/room/bed.txt
house/table.txt
house/chair.txt
animal/bird
animal/cat.txt
river.txt
mountain.txt

```

- 第 2.1 步的输入输出与下例类似：
  - 输入 “house/room/”，显示如下内容：
 

```
house/room/kitchen/path.txt
house/room/bed.txt
```
  - 输入 “house/room/bed.txt”，而 bed.txt 里面的内容是 “Person”，显示如下内容：
 

```
Person
```
  - 输入 “house/room/beeeed.txt”，显示如下内容：
 

```
Unknown file
```
- 第 2.2 步的输入输出与下例类似：
  - 输入 “count house”，显示如下内容（file 和 directory 的单复数无需区分）：
 

```
house : 4 files, 2 directories
room: 2 files, 1 directory
kitchen: 1 file, 0 directory
```
  - 输入 “count house/room/bed.txt”，显示如下内容：
 

```
house/room/bed.txt is not a directory!
```

### 1.3 注意事项

- 程序运行平台：Linux，或 windows，或 mac，或 VMWare 等虚拟机中的 Linux。不是 Dosbox 或 Bochs！
- 本次实验语言：C + 汇编。
- 基本分 9 分，能正确打印超过 512B 的文件 10 分。
- 注意要求直接根据 fat12 文件系统格式直接读取 a.img 中的二进制内容，请不要先调用系统命令将 a.img 挂载再去调用系统命令遍历文件夹。
- 只需要支持纯英文字符即可，不用考虑中文字符。
- 程序应该对用户错误的输入做出恰当的提示，不能崩溃。
- 请将目录和普通文件用不同的颜色输出（不用按照示例输出中的颜色）。
- 程序只需要考虑接受一条输入，显示一条结果即可，不用考虑退出问题。
- 输入指令以回车符号结束，要求可以多次不断输入。
- 检查时会检查代码，可能的考察方式包括：
  - 要求进行 .img 文件的挂载，删除目录或文件，重新运行程序，即改变输入。
  - 要求对代码稍作修改后，比如对颜色等，重新 make 运行。
- 要求程序由两个源文件构成，**main.c**和 **my\_print.asm**，其中 main.c 是主程序，可以使用基本的 C 库。但是打印不能使用标准函数 printf，要求使用 my\_print.asm 中使用汇编编写的 my\_print 函数。
- 要求使用 makefile 编译链接两个文件，并作为作业的一部分提交。（如果是使用 mac 平台或者 windows 平台的同学请额外加 txt 文件说明下，其他同学默认使用 linux 平台）

### 1.4 补充说明

由于保护模式不容易理解，而且代码也比较死，所以这次代码作业没有去写保护模式的代码，而是探究了操作系统中的“文件系统”这一主题，正好 loader 部分也涉及到对 fat12 的理解。

同时 gcc 和 nasm 联合使用也是为之后的实验奠定基础。

## 1.5 加分项：链接实验

按照链接相关 PPT 中要求完成动态链接实验的同学可以获得加分。注意，必须通过实验手段验证每一步并进行解释，参照 PPT 中静态链接相关内容。

请提供你的实验过程截图，使用的所有源代码，将其组织成报告提交到 TSS。检查时请向助教主动演示，并回答助教的随机提问。

提示：可以使用 PPT 中的代码和 PPT 上提示的命令。

## 2 问题清单

在整个实验的过程中，无论是编程还是查资料，请各位同学注意思考以下问题，助教检查时会从中随机抽取数个题目进行提问，根据现场作答给出分数。请注意，我们鼓励自己思考和动手实验，如果能够提供自己的思考结果并辅助以相应的实验结果进行说明，在分数评定上会酌情考虑。

### 2.1 PPT 相关内容

1. 实模式下的寻址方式以及实模式的缺陷
2. 保护模式下的寻址过程：
  - 段寄存器中存储的是什么？GDT 是什么？LDT 是什么？如何区分 LDT 和 GDT？LDT 和 GDT 的区别是什么？如何定位到 Descriptor？Descriptor 的内容有哪些？
  - GDTR 中的内容是什么？LDTR 中存储的是什么？为什么 LDT 要放在 GDT 中？
3. 选择子的作用：
  - 选择子是什么？它的值存放在哪里？
  - 选择子里面的内容有哪些？
  - 为什么偏移地址大小是 13 位？
4. 描述符的作用：
5. GDTR/LDTR 的作用：
  - GDTR 的内容是什么？
  - LDTR 的内容是什么？
6. 根目录区大小一定么？扇区号是多少？为什么？
7. 数据区第一个簇号是多少？为什么？

8. FAT 表的作用？
9. 解释静态链接的过程。
10. 解释动态链接的过程。
11. 静态链接相关 PPT 中为什么使用 ld 链接而不是 gcc。
12. linux 下可执行文件的虚拟地址空间默认从哪里开始分配。

## 2.2 实验相关内容

1. BPB 指定字段的含义
2. 如何进入子目录并输出 (说明方法调用)
3. 如何获得指定文件的内容, 即如何获得数据区的内容 (比如使用指针等)
4. 如何进行 C 代码和汇编之间的参数传递和返回值传递
5. 汇编代码中对 I/O 的处理方式, 说明指定寄存器所存值的含义
6. 可以要求解释某些看不懂的代码 (我看不懂的话, 你得讲给我听)

## 3 参考资料

1. 《Orange'S: 一个操作系统的实现
2. [Introduction to NASM](#)
3. [An overview of FAT12](#)
4. [Dynamic Linking and Loading](#)