CS 537
Section 7

Project 4

Michael Swift

The scenario:

- A museum was robbed of $300 million in art
- Last week, a hard disk belonging to the suspect was found
  - The suspect deleted files but did not wipe the disk
- You are part of the forensics team attempting to reconstruct the disk's contents. You have been given a region of the disk to reconstruct. So far other members of your team have determined that the file system was on a little- Endian machine running some form of UNIX with an inode structure

Known Inode Structure

```
#define N_DBLOCKS 10
#define N_IBLOCKS 4
  struct inode {
      int unknown; /* Unknown field */
      int protect; /* protection field */
      int nlink; /* Number of links to this file */
      int size; /* Number of bytes in file */
      int uid; /* Owner's user ID */
      int gid; /* Owner's group ID */
      int ctime; /* Time field */
      int mtime; /* Time field */
      int atime; /* Time field */
      int dblocks[N_DBLOCKS]; /* Pointers to data blocks */
      int iblocks[N_IBLOCKS]; /* Pointers to indirect blocks */
      int i2block; /* Pointer to doubly indirect block */
      int i3block; /* Pointer to triply indirect block */
  };
```

- Blocks are 1024 bytes
- The owner's UID and GID appear to be 18390 and 9921 respectively

Your task:

- You are given a 10 MB region of the disk and an unknown offset, containg inodes and da
- You need to:
  - Reconstruct any files in your assigned disk region
  - Produce a list of data blocks not in use by the above files
  - Identify the perpetrator and why you suspect him/her
  - Information about the files you found and how you found them

So, how do you do this?

- What hints do you have?

Inode structure

- You know that if you find the numbers matching the UID and GID, you have a good chance of finding an inode:

```
while (!feof(file)) {
    fread(&inode,sizeof(inode),1,file);
    If ((inode->uid == 18390) && (inode->gid == 9921))
      process_inode(inode);
    }
}
```


What other hints do you have?

- You don't know what chunk of the file you have, so you cannot directly use block numbers.
  - How can you figure out the offset of these blocks?
  - Can you identify the indirect block that corresponds to a file?
- You don't have the file names and extensions.
  - How can you figure out what types of files they are?

Magic Numbers

- Many files have "magic numbers" at the beginning that identify their type
  - See http://www.garykessler.net/library/file_sigs.html
- Example:

  $> hexdump < comic.jpg

  0000000 **ffd8 ffe0 0010 4a46 4946 0001** 0200 0064

  …

  $> hexdump < photo.jpg

  000000 **ffd8 ffe0 0010 4a46 4946 0001** 0201 0048

Logistics

- How can you break this into pieces?