Tutorial 6

Bitwise operators, binary files, and hex editing

Bitwise operators and masks in C

- 6 bit manipulation operators
- only work on integrals e.g. int or char
- & binary AND 101 & 110 = 100
- | binary OR 101 | 110 = 111
- ^ binary XOR 101 ^ 110 = 011
- ~ unary one's complement (NOT) ~101 = 010 (swap bits)
- << binary left shift 101 << 2 = 1010
- >> binary right shift 101 >> 2 = 1
- beware signed numbers have a sign bit (usually in position of **most significant bit**)

Bitwise operators and masks in C

- usually i have to write a binary example down to double-check (as in previous slide)

- octal or hexadecimal can also be used in C

    - octal prefix is 0   so $0177_8 = \mathbf{1}*64 + \mathbf{7}*8 + \mathbf{7} = 127_{10}$

        - 1 octal digit <-> 3 binary digits

    - hex prefix is 0x   so $0xFF_{16} = 15*16 +15*1 = 255_{10}$

        - 2 hex digits = 8 binary digits = 1 byte

- *some* compiler extensions allow binary with 0b prefix

Bitfield Masks

- Common use of bitwise operators: **bitfield** masks
- bitfields are a data structure
    - as an integral type - char for 8 bits, int for 32 etc
    - decide what you want each bit to mean as if it were a boolean **flag**
    - uses less data and only 1 variable for many flags

Using masks

```
#define SAMBA_MODE ( 1 << 0 )
#define DISCO_MODE (1 << 1 )
#define SHUFFLE_MODE ( 1 << 2 )
#define TOP_SECRET_MODE ( 1 << 3 )

void jukebox( unsigned char flags );

int main() {

    jukebox( SAMBA_MODE | SHUFFLE_MODE );
…
```

Usually enumerated types are better

```
typedef enum Genre {
  GENRE_POP = 0,
  GENRE_CLASSIC_HITS,
  GENRE_FUNK,
  GENRE_MAX
} Genre;

Genre songs_in_each_genre[GENRE_MAX];

void play_genre( Genre selection );

play_genre( GENRE_POP );
```

Hex is useful

- colours in HTML are in hex e.g. **FFFFFF**
    - 2 chars for **red**, 2 for **green**, 2 for **blue**
    - 255 vs. FF as plain-text chars saves 1 byte
- hex editing for inspecting binary files
    - install '**hexedit**' or a hex editor of some sort
- binary format *may* be smaller than ASCII
    - e.g. 4-byte binary float vs. text 10000024.0000023
    - harder for users to fiddle with (for better or worse)
    - hacking programs or patching screw-ups (ex. Wing Commander)
    - embed an image into a program

Typical Binary File

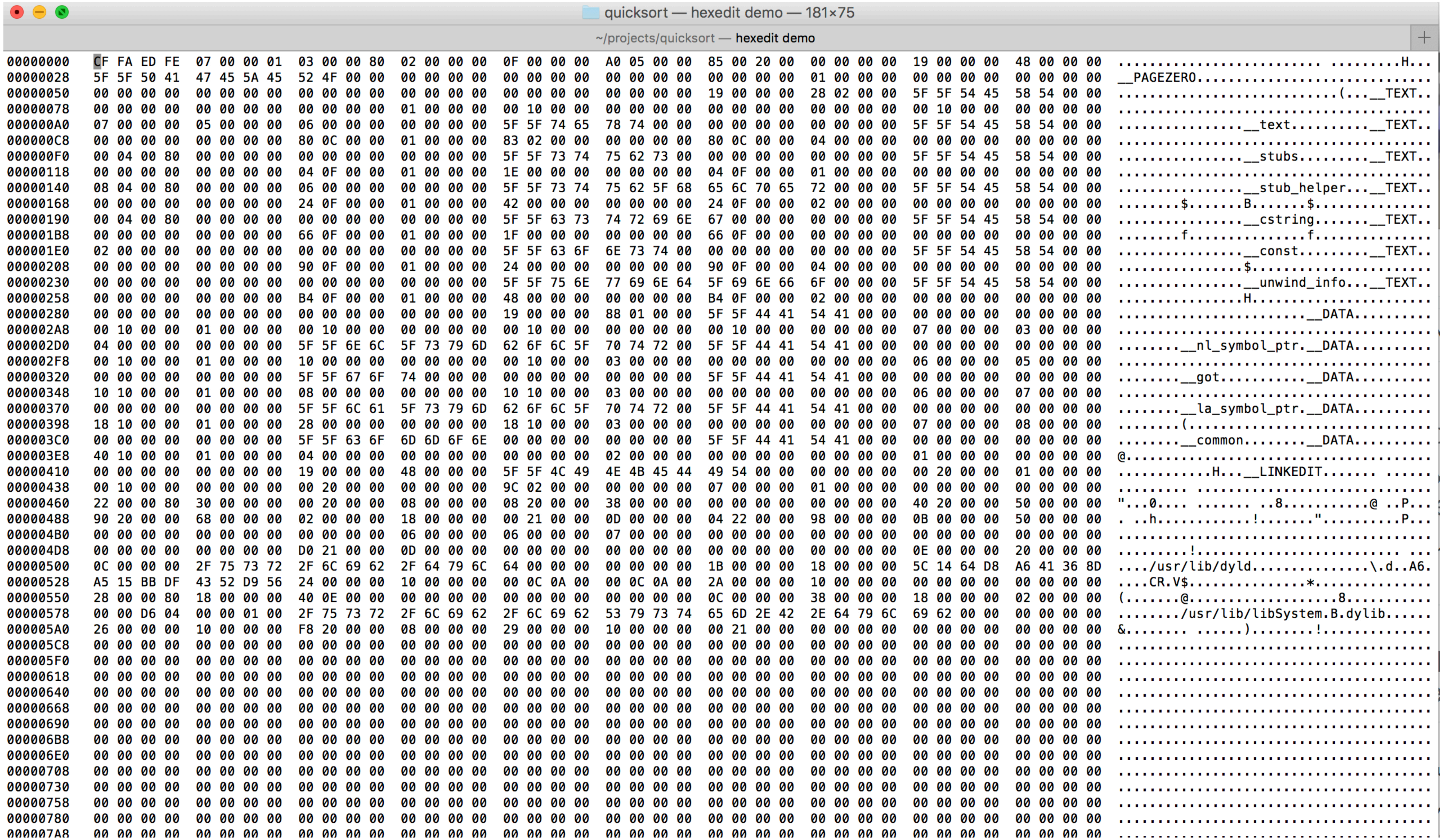
- Know your file format - specify this somewhere so you can read too
  - any header? e.g. format type or version number
    - ~some sort of char code so that it can show as plain text
  - **number of items** in next section e.g. integer with value **2**
  - **size** of data to follow e.g. **200** bytes
  - **200** bytes of **data**
  - **size** of next data e.g. **204** bytes
  - **204** bytes of **data**

Let's Write a Binary File, Hexedit, then read it

- `FILE* file_ptr = fopen("myfile.bin", "wb");`
- wb - write binary, rb - read binary
- `fwrite()` and `fread()` any memory or variable
- unfortunately - not reliable for read/write whole struct
- read and write assume same **endianness**
  - safe to assume **little-endian** bit order on modern machines
  - network protocols often use **big-endian**

BIG ENDIAN - The way people always broke their eggs in the Lilliput land

LITTLE ENDIAN - The way the king then ordered the people to break their eggs


quicksort — hexedit demo — 181×75
~/projects/quicksort — hexedit demo
+

```
00000000  CF FA ED FE  07 00 00 01  03 00 00 80  02 00 00 00  0F 00 00 00  A0 05 00 00  85 00 20 00  00 00 00 00  19 00 00 00  48 00 00 00  .......................... .........H...
00000028  5F 5F 50 41  47 45 5A 45  52 4F 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  01 00 00 00  00 00 00 00  00 00 00 00  __PAGEZERO..............................
00000050  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  19 00 00 00  28 02 00 00  5F 5F 54 45  58 54 00 00  ............................(...__TEXT..
00000078  00 00 00 00  00 00 00 00  00 00 00 00  01 00 00 00  00 10 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 10 00 00  00 00 00 00  ........................................
000000A0  07 00 00 00  05 00 00 00  06 00 00 00  00 00 00 00  5F 5F 74 65  78 74 00 00  00 00 00 00  00 00 00 00  5F 5F 54 45  58 54 00 00  ................__text..........__TEXT..
000000C8  00 00 00 00  00 00 00 00  80 0C 00 00  01 00 00 00  83 02 00 00  00 00 00 00  80 0C 00 00  04 00 00 00  00 00 00 00  00 00 00 00  ........................................
000000F0  00 04 00 80  00 00 00 00  00 00 00 00  00 00 00 00  5F 5F 73 74  75 62 73 00  00 00 00 00  00 00 00 00  5F 5F 54 45  58 54 00 00  ................__stubs.........__TEXT..
00000118  00 00 00 00  00 00 00 00  04 0F 00 00  01 00 00 00  1E 00 00 00  00 00 00 00  04 0F 00 00  01 00 00 00  00 00 00 00  00 00 00 00  ........................................
00000140  08 04 00 80  00 00 00 00  06 00 00 00  00 00 00 00  5F 5F 73 74  75 62 5F 68  65 6C 70 65  72 00 00 00  5F 5F 54 45  58 54 00 00  ................__stub_helper...__TEXT..
00000168  00 00 00 00  00 00 00 00  24 0F 00 00  01 00 00 00  42 00 00 00  00 00 00 00  24 0F 00 00  02 00 00 00  00 00 00 00  00 00 00 00  ........$.......B.......$...............
00000190  00 04 00 80  00 00 00 00  00 00 00 00  00 00 00 00  5F 5F 63 73  74 72 69 6E  67 00 00 00  00 00 00 00  5F 5F 54 45  58 54 00 00  ................__cstring.......__TEXT..
000001B8  00 00 00 00  00 00 00 00  66 0F 00 00  01 00 00 00  1F 00 00 00  00 00 00 00  66 0F 00 00  00 00 00 00  00 00 00 00  00 00 00 00  ........f...............f...............
000001E0  02 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  5F 5F 63 6F  6E 73 74 00  00 00 00 00  00 00 00 00  5F 5F 54 45  58 54 00 00  ................__const.........__TEXT..
00000208  00 00 00 00  00 00 00 00  90 0F 00 00  01 00 00 00  24 00 00 00  00 00 00 00  90 0F 00 00  04 00 00 00  00 00 00 00  00 00 00 00  ................$.......................
00000230  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  5F 5F 75 6E  77 69 6E 64  5F 69 6E 66  6F 00 00 00  5F 5F 54 45  58 54 00 00  ................__unwind_info...__TEXT..
00000258  00 00 00 00  00 00 00 00  B4 0F 00 00  01 00 00 00  48 00 00 00  00 00 00 00  B4 0F 00 00  02 00 00 00  00 00 00 00  00 00 00 00  ................H.......................
00000280  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  19 00 00 00  88 01 00 00  5F 5F 44 41  54 41 00 00  00 00 00 00  00 00 00 00  ........................__DATA..........
000002A8  00 10 00 00  01 00 00 00  00 10 00 00  00 00 00 00  00 10 00 00  00 00 00 00  00 10 00 00  00 00 00 00  07 00 00 00  03 00 00 00  ........................................
000002D0  04 00 00 00  00 00 00 00  5F 5F 6E 6C  5F 73 79 6D  62 6F 6C 5F  70 74 72 00  5F 5F 44 41  54 41 00 00  00 00 00 00  00 00 00 00  ........__nl_symbol_ptr.__DATA..........
000002F8  00 10 00 00  01 00 00 00  10 00 00 00  00 00 00 00  00 10 00 00  03 00 00 00  00 00 00 00  00 00 00 00  06 00 00 00  05 00 00 00  ........................................
00000320  00 00 00 00  00 00 00 00  5F 5F 67 6F  74 00 00 00  00 00 00 00  00 00 00 00  5F 5F 44 41  54 41 00 00  00 00 00 00  00 00 00 00  ........__got...........__DATA..........
00000348  10 10 00 00  01 00 00 00  08 00 00 00  00 00 00 00  10 10 00 00  03 00 00 00  00 00 00 00  00 00 00 00  06 00 00 00  07 00 00 00  ........................................
00000370  00 00 00 00  00 00 00 00  5F 5F 6C 61  5F 73 79 6D  62 6F 6C 5F  70 74 72 00  5F 5F 44 41  54 41 00 00  00 00 00 00  00 00 00 00  ........__la_symbol_ptr.__DATA..........
00000398  18 10 00 00  01 00 00 00  28 00 00 00  00 00 00 00  18 10 00 00  03 00 00 00  00 00 00 00  00 00 00 00  07 00 00 00  08 00 00 00  ........(...............................
000003C0  00 00 00 00  00 00 00 00  5F 5F 63 6F  6D 6D 6F 6E  00 00 00 00  00 00 00 00  5F 5F 44 41  54 41 00 00  00 00 00 00  00 00 00 00  ........__common........__DATA..........
000003E8  40 10 00 00  01 00 00 00  04 00 00 00  00 00 00 00  00 00 00 00  02 00 00 00  00 00 00 00  00 00 00 00  01 00 00 00  00 00 00 00  @.......................................
00000410  00 00 00 00  00 00 00 00  19 00 00 00  48 00 00 00  5F 5F 4C 49  4E 4B 45 44  49 54 00 00  00 00 00 00  00 20 00 00  01 00 00 00  ............H...__LINKEDIT....... ......
00000438  00 10 00 00  00 00 00 00  00 20 00 00  00 00 00 00  9C 02 00 00  00 00 00 00  07 00 00 00  01 00 00 00  00 00 00 00  00 00 00 00  ......... ..............................
00000460  22 00 00 80  30 00 00 00  00 20 00 00  08 00 00 00  08 20 00 00  38 00 00 00  00 00 00 00  00 00 00 00  40 20 00 00  50 00 00 00  "...0.... ....... ..8...........@ ..P...
00000488  90 20 00 00  68 00 00 00  02 00 00 00  18 00 00 00  00 21 00 00  0D 00 00 00  04 22 00 00  98 00 00 00  0B 00 00 00  50 00 00 00  . ..h............!......."..........P...
000004B0  00 00 00 00  00 00 00 00  00 00 00 00  06 00 00 00  06 00 00 00  07 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  ........................................
000004D8  00 00 00 00  00 00 00 00  D0 21 00 00  0D 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  0E 00 00 00  20 00 00 00  .........!.......................... ...
00000500  0C 00 00 00  2F 75 73 72  2F 6C 69 62  2F 64 79 6C  64 00 00 00  00 00 00 00  1B 00 00 00  18 00 00 00  5C 14 64 D8  A6 41 36 8D  ..../usr/lib/dyld...............\.d..A6.
00000528  A5 15 BB DF  43 52 D9 56  24 00 00 00  10 00 00 00  00 0C 0A 00  00 0C 0A 00  2A 00 00 00  10 00 00 00  00 00 00 00  00 00 00 00  ....CR.V$...................*...........
00000550  28 00 00 80  18 00 00 00  40 0E 00 00  00 00 00 00  00 00 00 00  00 00 00 00  0C 00 00 00  38 00 00 00  18 00 00 00  02 00 00 00  (.......@...................8...........
00000578  00 00 D6 04  00 00 01 00  2F 75 73 72  2F 6C 69 62  2F 6C 69 62  53 79 73 74  65 6D 2E 42  2E 64 79 6C  69 62 00 00  00 00 00 00  ......../usr/lib/libSystem.B.dylib......
000005A0  26 00 00 00  10 00 00 00  F8 20 00 00  08 00 00 00  29 00 00 00  10 00 00 00  00 21 00 00  00 00 00 00  00 00 00 00  00 00 00 00  &........ ......)........!..............
000005C8  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  ........................................
000005F0  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  ........................................
00000618  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  ........................................
00000640  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  ........................................
00000668  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  ........................................
00000690  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  ........................................
000006B8  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  ........................................
000006E0  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  ........................................
00000708  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  ........................................
00000730  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  ........................................
00000758  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  ........................................
00000780  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  ........................................
000007A8  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  00 00 00 00  ........................................
```

byte number
(in hex)

actual bytes
(in hex)

bytes as ASCII

Side Thoughts

- Binary files somewhat obscure your data
  - **Q.** How could you protect against hex-edit?
- **Q.** How could you tell if a user has edited the data?
  - e.g. detect cheating in game by map edit