

EECE320 Data Lab: Manipulating Bits

1 Introduction

The purpose of this assignment is to become more familiar with bit-level representations of integers and floating point numbers. You'll do this by solving a series of programming "puzzles." Many of these puzzles are quite artificial, but you'll find yourself thinking much more about bits in working your way through them.

2 Logistics

This is an individual project. Clarifications and corrections will be posted on the course Piazza page. Grading will be done through in class questionnaire & electronic submission.

3 Handout Instructions

Start by copying `datalab1.tar` or `datalab2.tar` to a (protected) directory on a Linux machine in which you plan to do your work. Then give the command to extract a tar file, shown below is the example for `datalab1.tar`.

```
unix> tar xvf datalab1.tar.
```

This will cause a number of files to be unpacked in the directory. The only file you will be modifying and turning in is `bits.c`.

The `bits.c` file contains a skeleton for each of the 15 programming puzzles. Your assignment is to complete each function skeleton using only *straightline* code for the integer puzzles (i.e., no loops or conditionals) and a limited number of C arithmetic and logical operators. Specifically, you are *only* allowed to use the following eight operators:

```
! ~ & ^ | + << >>
```

A few of the functions further restrict this list. Also, you are not allowed to use any constants longer than 8 bits. See the comments in `bits.c` for detailed rules and a discussion of the desired coding style.

4 The Puzzles

This section describes the puzzles that you will be solving in `bits.c`.

Some of the puzzle are *bit manipulations* (treating the data as a vector of bits), some are focused on *two's complement arithmetic*, teaching bits as numbers. There are some extra credit problems that treat the bits as floating point numbers.

Each problem has a *Difficulty Rating* and *maximum number of operations*. The "Rating" field gives the difficulty rating (the number of points) for the puzzle, and the "Max ops" field gives the maximum number of operators you are allowed to use to implement each function. See the comments in `bits.c` for more details on the desired behavior of the functions - each one describes the intended semantics, the rating and the allowed number of operations. You may also refer to the test functions in `tests.c`. These are used as reference functions to express the correct behavior of your functions, although they don't satisfy the coding rules for your functions.

4.1 Floating-Point Operations

For Extra Credit, you can implement some common single-precision floating-point operations. In this section, you are allowed to use standard control structures (conditionals, loops), and you may use both int and unsigned data types, including arbitrary unsigned and integer constants. You may not use any unions, structs, or arrays. Most significantly, you may not use any floating point data types, operations, or constants. Instead, any floating-point operand will be passed to the function as having type *unsigned*, and any returned floating-point value will be of type *unsigned*. Your code should perform the bit manipulations that implement the specified floating point operations.

Table 1 describes a set of functions that operate on the bit-level representations of floating-point numbers. Refer to the comments in `bits.c` and the reference versions in `tests.c` for more information.

Name	Description	Rating	Max Ops
<code>float_neg(uf)</code>	Compute $-f$	2	10
<code>float_i2f(x)</code>	Compute (float) x	4	30
<code>float_twice(uf)</code>	Computer $2*f$	4	30

Table 1: Floating-Point Functions. Value f is the floating-point number having the same bit representation as the unsigned integer uf .

Functions `float_neg` and `float_twice` must handle the full range of possible argument values, including not-a-number (NaN) and infinity. The IEEE standard does not specify precisely how to handle NaN's, and the IA32 behavior is a bit obscure. We will follow a convention that for any function returning a NaN value, it will return the one with bit representation `0x7FC00000`.

The included program `fshow` helps you understand the structure of floating point numbers. To compile `fshow`, switch to the handout directory and type:

```
unix> make
```

You can use `fshow` to see what an arbitrary pattern represents as a floating-point number:

```
unix> ./fshow 2080374784

Floating point value 2.658455992e+36
Bit Representation 0x7c000000, sign = 0, exponent = f8, fraction = 000000
Normalized. 1.0000000000 X 2^(121)
```

You can also give `fshow` hexadecimal and floating point values, and it will decipher their bit structure.

5 Evaluation

The assignment is scored in three parts - you will receive 50 percent of the overall score for having correctly functioning functions handed in on time for both parts 1 and 2 of the datalab and 50 percent for being able to correctly answer questions about how your functions work in the in class questions, that will happen after part 2 is due.

Your correctness score will be computed out of a maximum of 62 points based on the following distribution:

34 of 62 points Correctness points (correct output) 28 of 62 points Performance points (not exceeding the maximum operations)

Correctness points. The 14 puzzles you must solve have been given a difficulty rating between 1 and 4, such that their weighted sum totals to 34. We will evaluate your functions using the `btest` program, which is described in the next section. You will get full credit for a puzzle if it passes all of the tests performed by `btest`, and no credit otherwise.

Performance points. Our main concern at this point in the course is that you can get the right answer. However, we want to instill in you a sense of keeping things as short and simple as you can. Furthermore, some of the puzzles can be solved by brute force, but we want you to be more clever. Thus, for each function we've established a maximum number of operators that you are allowed to use for each function. This limit is very generous and is designed only to catch egregiously inefficient solutions. You will receive two points for each correct function that satisfies the operator limit.

Extra Credit

You can receive up to 10% extra credit for this assignment by completing all of the floating point problems. You will likely have to read ahead to learn about floating point if you don't already know it, because we're going to cover it later; however, you're going to have to learn it someday...

Autograding your work

We have included some autograding tools in the handout directory — `btest`, `dlc`, and `driver.pl` — to help you check the correctness of your work.

- **btest**: This program checks the functional correctness of the functions in `bits.c`. To build and use it, type the following two commands:

```
unix> make
unix> ./btest
```

Notice that you must rebuild `btest` each time you modify your `bits.c` file.

You'll find it helpful to work through the functions one at a time, testing each one as you go. You can use the `-f` flag to instruct `btest` to test only a single function:

```
unix> ./btest -f bitAnd
```

You can feed it specific function arguments using the option flags `-1`, `-2`, and `-3`:

```
unix> ./btest -f bitAnd -1 7 -2 0xf
```

Check the file `README` for documentation on running the `btest` program.

- **dlc**: This is a modified version of an ANSI C compiler from the MIT CILK group that you can use to check for compliance with the coding rules for each puzzle. The typical usage is:

```
unix> ./dlc bits.c
```

The program runs silently unless it detects a problem, such as an illegal operator, too many operators, or non-straightline code in the integer puzzles. Running with the `-e` switch:

```
unix> ./dlc -e bits.c
```

causes `dlc` to print counts of the number of operators used by each function. Type `./dlc -help` for a list of command line options.

- **driver.pl**: This is a driver program that uses `btest` and `dlc` to compute the correctness and performance points for your solution. It takes no arguments:

```
unix> ./driver.pl
```

Your instructors will use `driver.pl` to evaluate your solution.

6 Handin Instructions

Upload your *bits.c* file to Turnin by the due date. You can upload your file multiple times, and we encourage you upload often so you don't run into problems at the last minute or lose your work. Make sure you upload the correct *bits.c* file for each part, since this assignment is being done in two parts.

It is also recommended you commit all your assignments to the provided individual course git repo, so if turnin has issues, I can grade your submission there.

7 Advice

- Don't include the `<stdio.h>` header file in your `bits.c` file, as it confuses `dlc` and results in some non-intuitive error messages. You will still be able to use `printf` in your `bits.c` file for debugging without including the `<stdio.h>` header, although `gcc` will print a warning that you can ignore.
- The `dlc` program enforces a stricter form of C declarations than is the case for C++ or that is enforced by `gcc`. In particular, any declaration must appear in a block (what you enclose in curly braces) before any statement that is not a declaration. For example, it will complain about the following code:

```
int foo(int x)
{
    int a = x;
    a *= 3;      /* Statement that is not a declaration */
    int b = a;   /* ERROR: Declaration not allowed here */
}
```

- You will need 32 bit libraries to compile this assignment on your machines, for Debian linux systems this can be installed via the `gcc-multilib` library.

8 The “Beat the Prof” Contest

For fun, we’re offering an optional “Beat the Prof” contest for each part that allows you to compete with other students and the instructor to develop the most efficient puzzles. The goal is to solve each Data Lab puzzle using the fewest number of operators. Students who match or beat the instructor’s operator count for each puzzle are winners!

To submit your entry to the contest, type:

```
unix> ./driver.pl -u ``Your Nickname``
```

Nicknames are limited to 35 characters and can contain alphanumerics, apostrophes, commas, periods, dashes, underscores, and ampersands. You can submit as often as you like. Your most recent submission will appear on a real-time scoreboard, identified only by your nickname. You can view the scoreboard by pointing your browser at

```
https://datalab1.bryancdixon.com  
https://datalab2.bryancdixon.com
```