

Review Questions for Midterm One, Fall 2016

1. Convert decimal number 382.45 to octal number (base 8)
2. Considering numbers having a binary representation consisting of an infinite string of the form $0.yyyy\dots$, where y is a k -bit sequence. For example, the binary representation of $1/3$ is $0.01010101\dots$ ($y = 01$), while the representation of $1/5$ is $0.001100110011\dots$ ($y = 0011$).

A). Let Y be the decimal value of the infinite string $0.yyyyy\dots$. Give a formula in terms of y and k for the value Y represented by the infinite string.

Hint: Consider the effect of shifting the binary point k positions to the right.

B). What is the numeric value of the string for the following values of y ?

(a) 101

(b) 0110

(c) 010011

3. Around 250 B.C., the Greek mathematician Archimedes proved that $\frac{223}{71} < \pi < \frac{22}{7}$. Had he had access to a computer and the standard library `<math.h>`, he would have been able to determine that the single-precision floating point approximation of π has the hexadecimal representation `0x40490FDB`. Of course, all of these are just approximations, since π is not rational.

A. What is the fractional number denoted by this floating-point value?

B. What is the fractional binary representation of $\frac{22}{7}$?

C. At what bit position (relative to the binary point) do these two approximations to π diverge?

4. Determine the output of the following code segment (without running it):

Assume that a short is represented by 7 bits and an int is represented by 12 bits. What is the output generated by the following code segment:

```
int x = 63;
int y = -63;
short sx = (short) x;
short sy = (short) y;
```

`printf("%d , %d, %d, %d\n", x, y, (int)sx, (int)sy);` (10 points)

`printf("%x, %x, %x, %x\n", x, y, (int)sx, (int)sy);` (10 points)

`printf("%u, %u, %u, %u\n", x, y, (int)sx, (int)sy);` (10 points)

You may use a calculator to generate the values, but you must show you calculated them.

5. Consider the following two 7-bit floating point representations based on the IEEE floating point format. Neither has a sign bit - they can only represent non-negative numbers.

i). Format A.

- There are $k=3$ exponent bits. The exponent bias is 3.
- There are $n=4$ fraction bits.

ii). Format B.

- There are $k=4$ exponent bits. The exponent bias is 7.
- There are $n=3$ fraction bits.

Below, you are given some bit patterns in Format A, and your task is to convert them to the closest value in Format B. If necessary, you should apply the round-to-even rounding rule. In addition, give the values of numbers given by the Format A and Format B bit patterns. Give these as whole numbers (e.g., 17) or as fractions (e.g., $17/64$).

Format A		Format B	
Bits	Value	Bits	Value
011,0000			
101,1110			
010,1001			
110,1111			