Homework 5

1. Represent the following numbers as single-precision and double-precision IEEE floating-point numbers.

   a. 10.
   b. 10.5.
   c. 0.1.

2. Use the floating-point addition algorithm to compute $1.010_2 - .1001_2$ (in 4-bit precision).

3. This problem presents a way to reduce the number of addition steps in floating-point addition from three to two using only a single adder.

   a. Let A and B be integers of opposite signs, with a and b be their magnitudes. Show that the following rules for manipulating the unsigned numbers a and b gives $A + B$.

      (1) Complement one of the operands.
      (2) Using end around carry to add the complemented operand and the other (uncomplemented) one.
      (3) If there was a carry-out, the sign of the result is the sign associated with the uncomplemented operand.
      (4) Otherwise, if there was no carry-out, complement the result, and give it the sign of the complemented operand.

   b. Use the above to show how the steps in the floating-point addition algorithm can be performed using only a single addition.