

## Floating-Point Practice

Fill in the following table showing how the floating point numbers would be represented in binary, or the decimal value of the binary floating point numbers. Don't forget the exponent bias!

	Number	$s$	$e$	$f$
1	2			
2	1.5			
3	3			
4	-6			
5	-97.8125			
6	-1.333			
7	0.1			
8		1	0 1 1 0 0	1 0 0 0 0 0 0 0 0 0
9		0	1 0 0 1 0	0 1 0 0 0 0 0 0 0 0
A		1	1 0 0 1 0	0 1 1 0 0 0 0 0 0 0
B		0	0 1 1 1 1	0 0 0 0 0 0 0 0 0 0
C		1	1 0 1 1 0	1 1 1 1 1 1 1 0 0 0
D	-0			
E	NaN			
F	$\infty$			

## Normalized IEEE 754

$$(-1)^s \times 2^{(e - \text{Bias})} \times \left(1 + \frac{f}{2^m}\right)$$

$s$  0 if number is positive, 1 if number is negative

$e$  Between 0 and  $2^n - 1$  where  $n$  is the number of bits used to store the exponent

**Bias**  $2^{n-1} - 1$  where  $n$  is the number of bits used to store the exponent  
**Not Stored**

$f$  Integer numerator of the significand  
 $0 \leq f < 2^m$

$m$  Number of bits used to store  $f$   
**Not stored**

	$f$ Bits	$e$ Bits	Exponent Bias
<b>binary16</b>	10	5	$2^4 - 1 = 15$
<b>binary32</b>	23	8	$2^7 - 1 = 127$
<b>binary64</b>	52	11	$2^{10} - 1 = 1023$
<b>binary128</b>	112	15	$2^{14} - 1 = 16383$
<b>x87 80-bit</b>	64*	15	$2^{14} - 1 = 16383$