

1 Pre-Check

This section is designed as a conceptual check for you to determine if you conceptually understand and have any misconceptions about this topic. Please answer true/false to the following questions, and if false, correct the statement to make it true:

- [1.1] Depending on the context, the same sequence of bits may represent different things.

true

- [1.2] It is possible to get an overflow error in Two's Complement when adding numbers

of opposite signs. *Overflow errors only occur when the correct result when the addition falls outside the range of $[-2^{n-1}, 2^{n-1}-1]$. Adding numbers of opposite signs will not result in numbers outside of this range.*

- [1.3] If you interpret a N bit Two's complement number as an unsigned number, negative numbers would be smaller than positive numbers.

false. In two's complement, the MSB is always 1 for a negative number. This means Every negative number in Two's complement, when converted to unsigned, will be larger than the positive number.

- [1.4] If you interpret an N bit Bias notation number as an unsigned number (assume there are negative numbers for the given bias), negative numbers would be smaller

than positive numbers. *In bias notation, we add a bias to the unsigned interpretation to create the value. Regardless of where we shift the range of representable values, the negative numbers, when converted to unsigned, will always stay smaller than the positive numbers. This is unlike Two's Complement.*

- [1.5] We can represent fractions and decimals in our given number representation formats (unsigned, biased, and Two's Complement).

false. Our current representation formats has a major limitation; we can only represent and do with integers. To successfully represent fractional values as well as numbers with extremely high magnitude beyond our current boundaries, we need another representation format.

2 Unsigned Integers

- [2.1] If we have an n -digit unsigned numeral $d_{n-1}d_{n-2}\dots d_0$ in radix (or base) r , then the value of that numeral is $\sum_{i=0}^{n-1} r^i d_i$, which is just fancy notation to say that instead of a 10's or 100's place we have an r 's or r^2 's place. For the three radices binary, decimal, and hex, we just let r be 2, 10, and 16, respectively.

Let's try this by hand.

- (a) Convert the following numbers from their initial radix into the other two common radices:

1. 0b10010011 147 0x93
2. 63 0b00111111 0x3F
3. 0b00100100 36 0x24
4. 0 0b00000000 0x0
5. 39 0b00100111 0x27

$$\begin{array}{lll}
 \textcircled{6.} & 437 & 0b1101010 \\
 & 7. & 0x1B5 \\
 & 0x0123 & 29
 \end{array}$$

(b) Convert the following numbers from hex to binary:

$$\begin{array}{ll}
 1. 0xD3AD & 0b110100111010101 \\
 2. 0xB33F & 0b1011001100111111 \\
 3. 0x7EC4 & 0b011111011000100
 \end{array}$$

2.2 Our preferred tool for writing large numbers is the IEC prefixing system, which is similar to scientific notation but with powers of 2 rather than 10:

$$\begin{array}{llll}
 \text{Ki (Kibi)} = 2^{10} & \text{Gi (Gibi)} = 2^{30} & \text{Pi (Pebi)} = 2^{50} & \text{Zi (Zebi)} = 2^{70} \\
 \text{Mi (Mebi)} = 2^{20} & \text{Ti (Tebi)} = 2^{40} & \text{Ei (Exbi)} = 2^{60} & \text{Yi (Yobi)} = 2^{80}
 \end{array}$$

For example, we would write 2^{81} as $2 * 2^{80} = 2 \text{ Yi}$.

(a) Write the following numbers using IEC prefixes:

$$\begin{array}{llll}
 \bullet 2^{16} & 64 \text{ Ki} & \bullet 2^{27} & 128 \text{ Mi} \\
 \bullet 2^{34} & 16 \text{ Gi} & \bullet 2^{61} & 2 \text{ Ei} \\
 & & \bullet 2^{43} & 8 \text{ Ti} \\
 & & \bullet 2^{47} & 128 \text{ Ti} \\
 & & \bullet 2^{36} & 64 \text{ Gi} \\
 & & \bullet 2^{59} & 512 \text{ Pi}
 \end{array}$$

(b) Write the following numbers as powers of 2:

$$\begin{array}{llll}
 \bullet 2 \text{ Ki} & 2^{11} & \bullet 512 \text{ Ki} & 2^{19} \\
 \bullet 256 \text{ Pi} & 2^{58} & \bullet 64 \text{ Gi} & 2^{36} \\
 & & & \bullet 16 \text{ Mi} & 2^{24} \\
 & & & \bullet 128 \text{ Ei} & 2^{67}
 \end{array}$$

3 Signed Integers

3.1 Unsigned binary numbers work for natural numbers, but many calculations use negative numbers as well. To deal with this, a number of different schemes have been used to represent signed numbers. Here are two common schemes:

Two's Complement:

- We can write the value of an n -digit two's complement number as $\sum_{i=0}^{n-2} 2^i d_i - 2^{n-1} d_{n-1}$.
- Negative numbers will have a 1 as their most significant bit (MSB). Plugging in $d_{n-1} = 1$ to the formula above gets us $\sum_{i=0}^{n-2} 2^i d_i - 2^{n-1}$.
- Meanwhile, positive numbers will have a 0 as their MSB. Plugging in $d_{n-1} = 0$ gets us $\sum_{i=0}^{n-2} 2^i d_i$, which is very similar to unsigned numbers.
- To negate a two's complement number: flip all the bits and add 1.
- Addition is exactly the same as with an unsigned number.
- Only one 0, and it's located at 0b0.

Biased Representation:

- The number line is shifted so that the smallest number we want to be representable would be 0b0...0.
- To find out what the represented number is, read the representation as if it was an unsigned number, then add the bias.
- We can shift to any arbitrary bias we want to suit our needs. To represent (nearly) as much negative numbers as positive, a commonly-used bias for N bits is $-(2^{N-1} - 1)$.

For questions (a) through (c), assume an 8-bit integer and answer each one for the case of an unsigned number, biased number with a bias of -127, and two's complement number. Indicate if it cannot be answered with a specific representation.

- (a) What is the largest integer? What is the result of adding one to that number?

1. Unsigned?	255	0
2. Biased?	128	-127
3. Two's Complement?	127	-128

- (b) How would you represent the numbers 0, 1, and -1?

1. Unsigned?	00000000	00000001	can't be represented
2. Biased?	01111111	10000000	01111110
3. Two's Complement?	00000000	10000001	11111111

- (c) How would you represent 17 and -17?

1. Unsigned?	00010001	can't be represented
2. Biased?	10010000	01101110
3. Two's Complement?	00010001	11101111

- [3.2] Prove that the two's complement inversion trick is valid (i.e. that x and $\bar{x} + 1$ sum to 0).

Note that for any x we have $x + \bar{x} = 0b1\dots1$. Adding $0b1$ to $0b1\dots1$ will cause the value to overflow meaning that $0b1\dots1 + 0b1 = 0b0 = 0$. Therefore, $x + \bar{x} + 1 = 0$

A straightforward hand calculation shows that $0b1\dots1 + 0b1 = 0$

- 3.3 We now have three major radices (or bases) that allow us to represent numbers using a finite amount of symbols: binary, decimal, hexadecimal. Why do we use each of these radices, and why are each of them preferred over other bases in a given context?

Decimal is the preferred radix for human hand calculations, likely related to the fact that humans have 10 fingers.

Binary numerals are particularly useful for computers. Binary signals are less likely to be garbled than higher radix signals, as there is more "distance" (voltage or current) between valid signals (HIGH and LOW). Additionally, binary signals are quite convenient to design circuits, as we'll see later in the course.

Hexadecimal numbers are a convenient shorthand for displaying binary numbers, owing to the fact that one hex digit corresponds exactly to four binary digits.

4 Arithmetic and Counting

- 4.1 Addition and subtraction of binary/hex numbers can be done in a similar fashion as with decimal digits by working right to left and carrying over extra digits to the next place. However, sometimes this may result in an overflow if the number of bits can no longer represent the true sum. Overflow occurs if and only if two numbers with the same sign are added and the result has the opposite sign.

- (a) Compute the decimal result of the following arithmetic expressions involving

6-bit Two's Complement numbers as they would be calculated on a computer.

Do any of these result in an overflow? Are all these operations possible?

$$1. 0b011001 - 0b000111 = 0b010010 = 18$$

$$2. 0b100011 + 0b111010 = 0b011101 = 29 \text{ overflow}$$

$$3. 0k3B + 0x06 = 0b111011 + 0b000110 = 0b000001 = 1$$

$$4. 0xFF - 0xAA \text{ need 8 bits}$$

$$5. 0b000100 - 0b001000 = 0b001100 + 0b111000 = 0b111100 = -4$$

- (b) What is the least number of bits needed to represent the following ranges using any number representation scheme?

$$1. 0 \text{ to } 256$$

$$2. 7 \text{ to } 56$$

$$3. 64 \text{ to } 127 \text{ and } -64 \text{ to } -127$$

$$4. \text{Address every byte of a 12 TiB chunk of memory}$$

$$5b-7+1=64 \quad 2^6=64$$

$$2 \times (127-64+1) = 128 \quad 2^7=128$$

$$44$$

- (c) How many distinct numbers can the following schemes represent? How many distinct positive numbers?

$$1. 10\text{-bit unsigned}$$

$$1024$$

$$1023$$

$$2. 8\text{-bit Two's Complement}$$

$$256$$

$$127$$

$$3. 8\text{-bit One's Complement}$$

$$255$$

$$127$$

$$4. 6\text{-bit biased, with a bias of } -30$$

$$64$$

$$33$$

$$5. 10\text{-bit sign-magnitude}$$

$$1024$$

$$511$$

$$0b111111 - 0b011110 = 0b100001 \\ = 32 + 1 = 33$$

$$30 = 0b011110$$

only one 0.