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 include an explanation:

1.1 Depending on the context, the same sets of bits may represent different things.

true, like in decimals vs. hex vs. octal, signed, unsigned

1.2 It is possible to get an overflow error in Two's Complement when adding numbers of opposite signs.

false. impossible $[-(2^{n-1}), 2^{n-1}-1]$

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. negative numbers are bigger because they start with 1.

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.

true. because in biased notation is like unsigned but 0 is in the middle.

2 Unsigned Integers

2.1 If we have an n-digit unsigned numeral $d_{n-1}d_{n-2}\ldots d_0$ 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. Recall that our preferred tool for writing large numbers is the IEC prefixing system:

Ki (Kibi) $= 2^{10}$ Gi (Gibi) $= 2^{30}$ Pi (Pebi) $= 2^{50}$ Zi (Zebi) $= 2^{70}$

Mi (Mebi) $= 2^{20}$ Ti (Tebi) $= 2^{40}$ Ei (Exbi) $= 2^{60}$ Yi (Yobi) $= 2^{80}$

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

$2^1 = 2$
$2^2 = 4$
$2^3 = 8$
$2^4 = 16$
$2^5 = 32$
$2^6 = 64$
$2^7 = 128$
$2^8 = 256$
$2^9 = 512$

1. 0b10010011   0x93 $1+2+2^4+2^7 = 147_{10}$
   (digit positions: 7 6 5 4 3 2 1 0, exponents 16, 128)

2. 63 $|||||_2$ 0x1F   *(3)*

3. 0b00100100 $4+2^5 = 36_{10}$ 0x24   *(positions 5, 2)*

4. 0   0x0 0_2

5. 39 100111_2 0x27

6. 437 $1 1 0 1 1 0 1 0 1_2$ 0x1B5   *(positions 8 7 6 5 4 3 2 1 0)*

7. 0x0123 $0001 0010 0011_2$ $1+2+2^5+2^8 = 35+256 = 291_{10}$ *(8·5)*

10 A 1010
11 B 1011
12 C 1100
13 D 1101
14 E 1110

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

1. 0xD3AD $110100111010 1101_2$

2. 0xB33F 1011001100111111_2

3. 0x7EC4 0111111011000100_2

(c) Write the following numbers using IEC prefixes:

- 2^{16} $2^6 \cdot 2^{10} = 64K$
- 2^{27} $2^7 \cdot 2^{20} = 128M$
- 2^{43} $2^3 \cdot 2^{40} = 8T$
- 2^{36} $2^6 \cdot 2^{30} = 64G$
- 2^{34} $2^4 \cdot 2^{30} = 16G$
- 2^{61} $2^1 \cdot 2^{60} = 2E$
- 2^{47} $2^7 \cdot 2^{40} = 128T$
- 2^{59} $2^9 \cdot 2^{50} = 512P$

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

- 2 Ki $2^1 \cdot 2^{10} = 2^{11}$
- 512 Ki $2^9 \cdot 2^{10} = 2^{19}$
- 16 Mi $2^4 \cdot 2^{20} = 2^{24}$
- 256 Pi $2^8 \cdot 2^{50} = 2^{58}$
- 64 Gi $2^6 \cdot 2^{30} = 2^{36}$
- 128 Ei $2^7 \cdot 2^{60} = 2^{67}$

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, but we will focus on two's complement, as it is the standard solution for representing signed integers.

- Most significant bit has a negative value, all others are positive. So the value of an n-digit two's complement number can be written as $\sum_{i=0}^{n-2} 2^i d_i - 2^{n-1} d_{n-1}$.

- Otherwise exactly the same as unsigned integers.

- A neat trick for flipping the sign of 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.

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? *511* *255, 0* *128*

   2. Biased? *511 − 127 = 384, −127*

   3. Two's Complement? *255 256/2−1 = 127, −128*

8 bit, highest ind is 7
overflow → go to smallest number

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

   1. Unsigned? *0:0   1:1   −1: N.A.*

   2. Biased? *0: 11111111   1: 10000000   −1: 11111110*

   3. Two's Complement? *0:0   1:01   −1: 11111111*
76543210   −128+127=−1

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

   1. Unsigned? *17: 10010   −17: NA*

   2. Biased? *17: 17 + 127 = 144 = 10010000   −17: −17+127 = 110 = 1101110*
5 4 3 2 1

   3. Two's Complement? *17: 010010   −17: 111111*
8 bits!
1 1 1 0 1 1 1 1
7 6 5 4 3 2 1 0

(d) What is the largest integer that can be represented by *any* encoding scheme that only uses 8 bits?

unsigned: 255

no such integer, we can choose any kind of mapping for a random scheme

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

17: 00010010
· 11101101
+
1 00000000

$X + \bar{X}$ is always all ones.
add 1 to if will overflow
→ 0

(f) Explain where each of the three radices shines and why it is preferred over other bases in a given context.

decimal: every day
hex: short
binary: machine understand

$$\underbrace{X + \underbrace{\bar{X} + 1}_{1}}_{} = 0$$

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

2. 0b100011 + 0b111010   *overflow*

3. 0x3B + 0x06

4. 0xFF − 0xAA
       6x6

Handwritten work:

1.
```
   011001
+  111001
 (1)010010
```

2.
```
  100011
+ 111010
(1)011101
```

3.
```
  00111011
+ 00000110
  01000001
```

4.
```
  11111111
+ 01010110
(1)01010101
```

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

1. 0 to 256   *9 8 bits unsigned. includes 256*   *Impossible! 8 bits*

2. -7 to 56   *64 nums → $2^6 = 64$ 6 bits*

3. 64 to 127 and -64 to -127   *$127 - 64 + 1 = 64$ $64 \cdot 2 = 128$ nums → 2^7 → 7 bits*

4. Address every byte of a 12 TiB chunk of memory

$12\,\text{TiB} = 12 \cdot 2^{40}$

$\approx 2^4 \cdot 2^{40} \approx 2^{44}$