# PRCMP PL01
# Numeral systems and arithmetic

May 2023

## 1 Numeral systems

A *numeral system* is a mathematical notation to express *quantities* using *symbols* in a conventional manner. A numeral system must represent a set numbers that are meaningful to a specific purpose, providing that each number has a unique representation. Furthermore, it must reflect the arithmetic and algebraic structure of the numbers.

We were taught to represent quantities using a *positional system*, in which each symbol has its intrinsic value multiplied by the weight of its position in the number. For instance, we say that 123.5 means *one hundred and twenty three and a half* because we assume that the value of the digit 1 is 100 as it is located in the hundreds *order* (i.e. in the hundreds position). Our mental process is represented as:

$$123 = 100 + 20 + 3 + 0.5$$

Formally, we can write the number in a polynomial form, where each symbol is multiplied by the *base* of the numeral system – *i.e.* 10 – raised to the power of its *order*:

$$123_{10} = 1 \times 10^2 + 2 \times 10^1 + 3 \times 10^0 + 5 \times 10^{-1}$$

This rule applies to any base.

Things to remember:

- the *units* order (the position immediately to the left of the decimal point) is *zero*;
- the orders increase to the left (integer part);
- the orders decrease to the right (fractional part).

The *base* has a fundamental role in a numeral system, as it sets the weights of each order. The base indicates how many symbols (*i.e.* digits) the numeral system as available to represent a number. For instance, base-10 has 10 symbols – $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ – while base-2 has only two: $\{0, 1\}$. Every base uses the symbol 0 (zero) so, for any base $N$, the highest value of the available symbols is $(N - 1)$.

### Questions

1. Represent the following numbers in the polynomial form, according to their base, and determine their value in decimal:
   (a) $12345.091_{10}$
   (b) $640_8$
   (c) $1010_2$
   (d) $2102_3$
   (e) $241.3_5$

(f) $613.25_7$

(g) $C5F_{16}$

2. Write the sequence of numbers between...

   (a) ... $5_8$ and $11_8$, represented in base 8.

   (b) ... $0_2$ and $101_2$, represented in binary.

   (c) ... $9_{16}$ e $12_{16}$, represented in hexadecimal

3. Let us consider writing numbers in bases 2, 8, 10 and 16.

   (a) Write the greatest two-digit number for each of these bases.

   (b) Which number follows each of your previous answer?

   (c) Which number is the predecessor of 101 for each of these bases?

## 2 Decimal-to-binary conversions (and vice-versa)

Converting a number from any base to decimal is achieved by writing it in the polynomial form and evaluating the result. For instance, for the number $1011_2$ we determine the respective decimal representation in the following manner:

$$\begin{aligned}1011_2 &= 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 \\ &= 8 + 0 + 2 + 1 \\ &= 11_{10}\end{aligned}$$

Converting from decimal to another base requires a different process, that is divided into two phases:

- one phase for the integer part,
- one other phase for the fraccional part.

Converting the integer part is carried by successive integer divisions by the new base. The remainder of each division is the digit of the next higher order. For instance, converting $6_{10}$ into binary is done by taking the following steps:

**1st step** Order 0: $6/2 = 3$, remainder : **0**

**2nd step** Order 1: $3/2 = 1$, remainder : **1**

**3rd step** Order 2: $1/2 = 0$, remainder : **1**

**Result** $6_{10} = 110_2$

Converting the fractional part is carried by the new base's successive multiplications of the fraccional part. The integer part of each multiplication is the digit of the next lower order. For instance, converting $0.25_{10}$ into binary is done by taking the following steps:

**1st step** Order $-1$: $0.25 \times 2 = 0.50$, integer : **0**

**2nd step** Order $-2$: $0.50 \times 2 = 1.00$, integer : **1**

**3rd step** Order $-3$: $0.00 \times 2 = 0.00$, integer : **0**

**Result** $0.25_{10} = 0.010_2$

**Questions**

4. Convert the following decimal integers into binary.
   (a) 30
   (b) 49
   (c) 190
5. Convert the following integer binaries into decimal.
   (a) $1011_2$
   (b) $10\,1001_2$
   (c) $1000\,0111_2$
6. Compute a list of the powers of 2, from $2^0$ until $2^{10}$, both in decimal and in binary, and then answer the following questions without further calculations.
   (a) What is the decimal representation for binary $10\,0000_2$?
   (b) What is the decimal representation for binary $1\,1111_2$?
   (c) What is the value (in decimal) of the greatest binary number with 8 digits?
   (d) How many bits are required to represent the following decimal integers: 100, 200, 300, 500, 520?
7. Convert the following real numbers from binary into decimal.
   (a) $1101.01_2$
   (b) $11\,1001.0011_2$
   (c) $1010\,1101_2$
   (d) $10\,0110_2$
   (e) $0.1010_2$
   (f) $1010.1101_2$
8. Convert the following real numbers from decimal into binary (maximum resolution of 4 fractional bits).
   (a) $251.75_{10}$
   (b) $1020.64_{10}$
   (c) $0.942_{10}$
   (d) $7.654_{10}$

# 3 Binary/octal and binary/hexadecimal conversions

The binary base is the numeral system used by conventional computers because of many technological advantages. However, the human brain does not deal easily with the binary numeral system because binary numbers expand very quickly.

Engineers have become used to using octal and hexadecimal numeral systems as a more human-friendly way to work with binaries.

Base 8 and 16 have one interesting feature: both require an additional digit exactly on the same values that require an additional bit.

Base 8 is the third power of 2, so this pattern repeats for every group of 3 bits. This means that an octal symbol can be directly translated into a set of 3 bits, accordingly to

| | |
|---|---|
| $0_8$ — $000_2$ | $4_8$ — $100_2$ |
| $1_8$ — $001_2$ | $5_8$ — $101_2$ |
| $2_8$ — $010_2$ | $6_8$ — $110_2$ |
| $3_8$ — $011_2$ | $7_8$ — $111_2$ |

In the same way, base 16 is the fourth power of 2, so the pattern repeats for every group of 4 bits. This means that an hexadecimal symbol can be directly translated into a set of 4 bits, accordingly to

| | |
|---|---|
| $0_{16}$ — $0000_2$ | $8_{16}$ — $1000_2$ |
| $1_{16}$ — $0001_2$ | $9_{16}$ — $1001_2$ |
| $2_{16}$ — $0010_2$ | $A_{16}$ — $1010_2$ |
| $3_{16}$ — $0011_2$ | $B_{16}$ — $1011_2$ |
| $4_{16}$ — $0100_2$ | $C_{16}$ — $1100_2$ |
| $5_{16}$ — $0101_2$ | $D_{16}$ — $1101_2$ |
| $6_{16}$ — $0110_2$ | $E_{16}$ — $1110_2$ |
| $7_{16}$ — $0111_2$ | $F_{16}$ — $1111_2$ |

**Questions**

9. Convert the following binary numbers into octal and hexadecimal numbers.
   (a) $100\,1001_2$
   (b) $11\,1101.11_2$
   (c) $101\,1011.01_2$
10. Convert the following hexadecimal numbers into octal.
   (a) $3A_{16}$
   (b) $1CB_{16}$
   (c) $ABC.35_{16}$
11. Convert the following octal numbers into hexadecimal.
   (a) $37_8$
   (b) $4164_8$
   (c) $14.72_8$

# 4 Arithmetic operations

The algorithms for calculating arithmetic operations are universal, remaining the same in any numerical base.

**Questions**

12. Perform the following arithmetic operations, in binary.
   (a) $111\,0110_2 + 1111_2$
   (b) $1010_2 + 1\,1011_2$
   (c) $3322_8 + 332_8$
   (d) $11\,1010_2 - 1\,1001_2$
   (e) $3131_8 - 212_8$
   (f) $6584_{16} - CA_{16}$
   (g) $111\,0110_2 \times 1111_2$
   (h) $147_8 \times 31_8$
   (i) $F73_{16} \times D4_{16}$

(j) $1\,1100\,1010_2 \div 101_2$

(k) $231_8 \div 6_8$

(l) $158_{16} \div 10_{16}$

13. It is very easy to perform multiplication and division by powers of ten, in decimal. Check that the same rules apply with powers of two, in binary.

(a) $111\,0110_2 \times 10_2$

(b) $1010_2 \times 100_2$

(c) $111\,0110_2 \times 1000_2$

(d) $1\,1100\,1010_2 \div 10_2$

(e) $1\,0100\,1110_2 \div 100_2$

(f) $1\,1001\,1010_2 \div 1000_2$

# 5 Solutions

1. (a) $1 \times 10^4 + 2 \times 10^3 + 3 \times 10^2 + 4 \times 10^1 + 5 \times 10^0 + 0 \times 10^{-1} + 9 \times 10^{-2} + 1 \times 10^{-3}$
   $= 12345.091$

   (b) $6 \times 8^2 + 4 \times 8^1 + 0 \times 8^0 = 6 \times 64 + 4 \times 8 + 0$
   $= 416$

   (c) $1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 0 \times 2^0$
   $= 8 + 0 + 2 + 0$
   $= 10$

   (d) $2 \times 3^3 + 1 \times 3^2 + 0 \times 3^1 + 2 \times 3^0$
   $= 54 + 9 + 0 + 2$
   $= 65$

   (e) $2 \times 5^2 + 4 \times 5^1 + 1 \times 5^0 + 3 \times 5^{-1}$
   $= 50 + 20 + 1 + 0,6$
   $= 71,6$

   (f) $6 \times 7^2 + 1 \times 7^1 + 3 \times 7^0 + 2 \times 7^{-1} + 5 \times 7^{-2}$
   $= 294 + 7 + 3 + 0,285 + 0,102$
   $= 304,29$

   (g) $12 \times 16^2 + 5 \times 16^1 + 15 \times 16^0$
   $= 3072 + 80 + 15$
   $= 3167$

2. (a) $5_8, 6_8, 7_8, 10_8, 11_8$

   (b) $0_2, 1_2, 10_2, 11_2, 100_2, 101_2$

   (c) $9_{16}, A_{16}, B_{16}, C_{16}, D_{16}, E_{16}, F_{16}, 10_{16}, 11_{16}, 12_{16}$

3. (a) $11_2; 77_8; 99_{10}; FF_{16}$

   (b) $100_2; 100_8; 100_{10}; 100_{16}$

   (c) $100_2; 100_8; 100_{10}; 100_{16}$

4. (a) $1\,1110_2$

   (b) $11\,0001_2$

   (c) $1011\,1110_2$

5. (a) $11_{10}$

   (b) $41_{10}$

   (c) $135_{10}$

6. (a) $10\,0000_2 = 2^5 = 32$

   (b) $1\,1111 = 2^5 - 1 = 32 - 1 = 31$

   (c) $1111\,1111_2 = 2^8 - 1 = 255$

   (d) 100: 7 bits ($2^6 \leq 100 < 2^7$)
   200: 8 bits ($2^7 \leq 200 < 2^8$)
   300: 9 bits ($2^8 \leq 200 < 2^9$)
   500: 9 bits ($2^8 \leq 200 < 2^9$)
   1000: 10 bits ($2^9 \leq 200 < 2^10$)

7. (a) 13.25

   (b) 57.1875

   (c) 173

   (d) 38
   (e) 0.625
   (f) 10.8125

8. (a) $1111\,1011.11_2$
   (b) $11\,1111\,1100.1010_2$
   (c) $0.1111_2$
   (d) $111.1010_2$

9. (a) $111_8 = 49_{16}$
   (b) $75.6_8 = 3D.C_{16}$
   (c) $133.2_8 = 5B.4_{16}$

10. (a) $72_8$
   (b) $713_8$
   (c) $5274.152_8$

11. (a) $1F_{16}$
   (b) $874_{16}$
   (c) $C.E8_{16}$

12. (a) $1000\,0101_2$
   (b) $10\,0101_2$
   (c) $111\,1010\,1100_2$
   (d) $10\,0001_2$
   (e) $101\,1100\,1111_2$
   (f) $110\,0100\,1011\,1010_2$
   (g) $110\,1110\,1010_2$
   (h) $1010\,0000\,1111_2$
   (i) $1100\,1100\,1011\,0011\,1100_2$
   (j) $101\,1011_2$
   (k) $1\,1001_2$
   (l) $1\,0101_2$

13. (a) $1110\,1100_2$
   (b) $10\,1000_2$
   (c) $11\,1011\,0000_2$
   (d) $1110\,0101.0_2$
   (e) $101\,0011.10_2$
   (f) $11\,0011.010_2$