Data Representation

OBJECTIVES

After reading this chapter, the student should be able to:

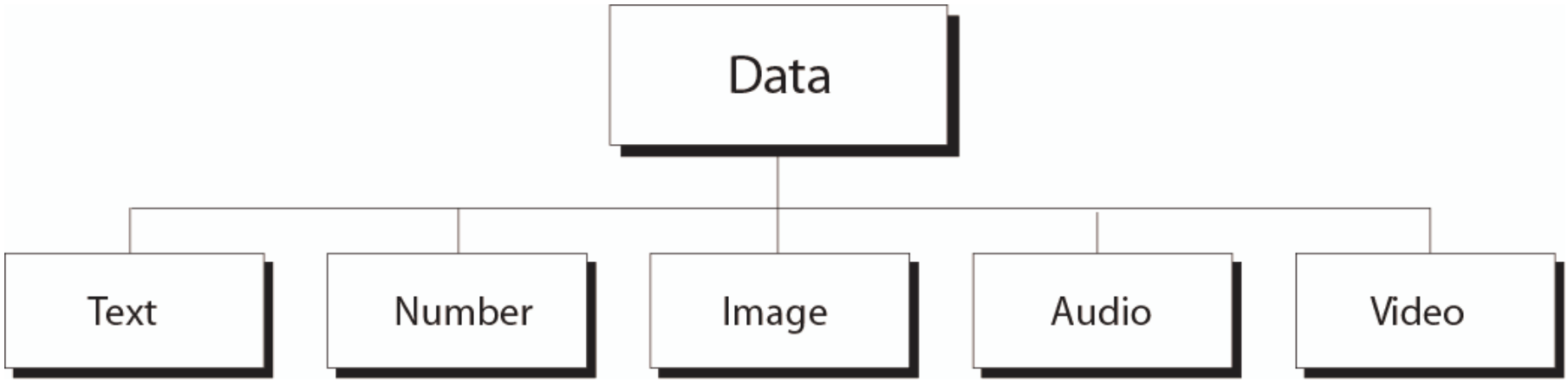
- Define data types.
- Visualize how data are stored inside a computer.
- Understand the differences between text, numbers, images, video, and audio.
- Work with hexadecimal and octal notations.

2.1

DATA TYPES

Different types of data

Numbers, text, images, audio, and video are all forms of data. Computers need to process all types of data.



Note:

The computer industry uses the term "multimedia" to define information that contains numbers, text, images, audio, and video.

2.2

DATA INSIDE THE COMPUTER

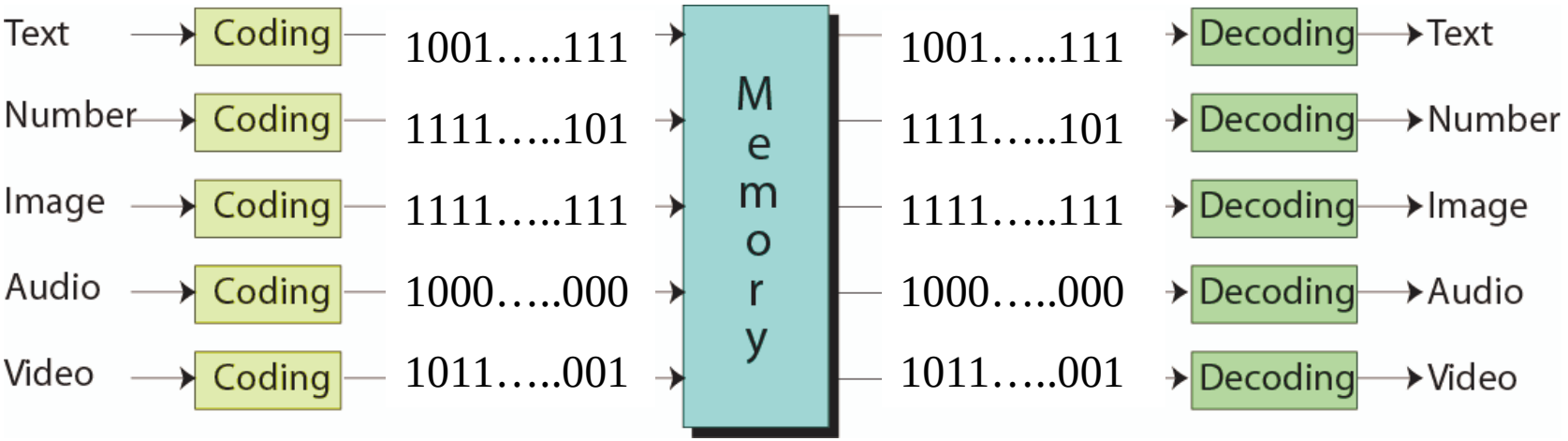
Bit pattern

- All data types from outside a computer are transformed into a uniform representation called a bit pattern for processing by computers.
- A bit is the smallest unit of data that can be stored in a computer
- A switch, with its two states of on and off, can represent a bit(0 or 1)
- A bit pattern is a sequence of bits that can represent a symbol

1 0 0 0 1 0 1 0 1 0 1 1 1 1 1

Examples of bit patterns

- Data are coded when they enter a computer and decoded when they are presented to the user

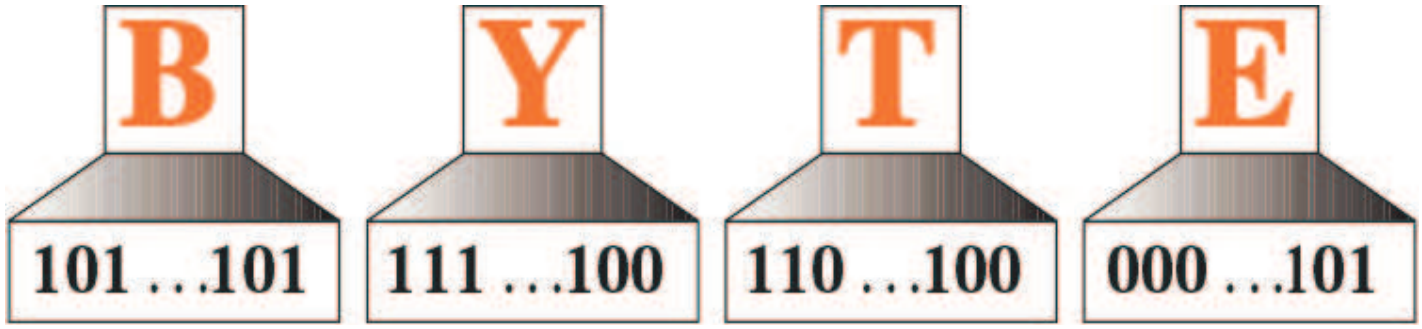


- A bit pattern of length 8 is called a byte
- Byte also been used to measure the size of memory or other storage devices

2.3

REPRESENTING DATA

Representing symbols using bit patterns



How many bits are needed in a bit pattern to represent a symbol in a language?

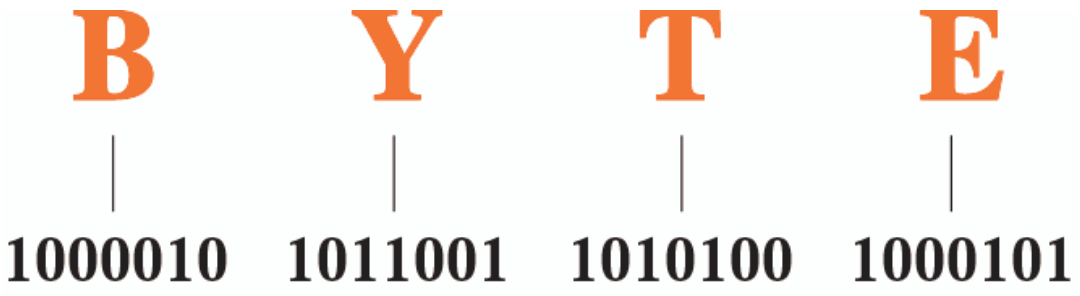
⇒Depend on how many symbols are in the set

Number of symbols and bit pattern length

Number of Symbols	Bit Pattern Length
2	1
4	2
8	3
16	4
…	…
128	7
256	8
…	…
65,536	16

Codes

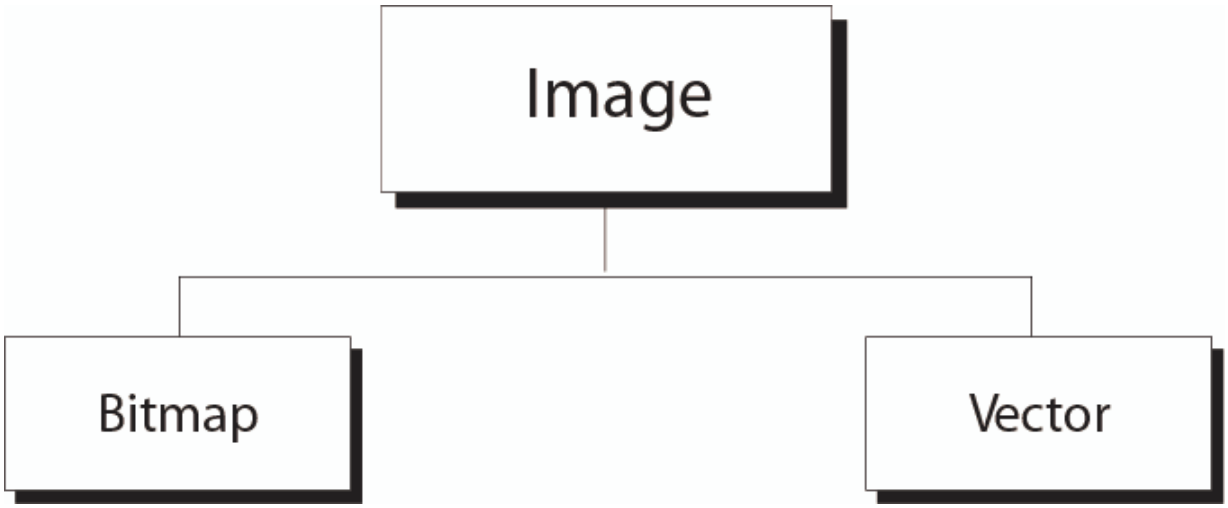
- Code: different sets of bit patterns have been designed to represent text symbols
  - ASCII: popular code for symbols
    - American Standard Code for Information Interchange
  - EBCDIC: used in IBM mainframes
    - Extended Binary Coded Decimal Interchange Code
  - Unicode: 16-bit code, allow a greater number of symbols
  - ISO: 32-bit code, allow a greater number of symbols
- Coding is the process of transforming data into a bit pattern



Some features of ASCII

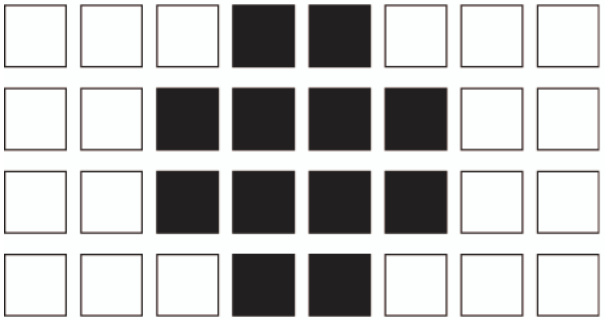
- ASCII uses a 7-bit pattern: 0000000 ~ 1111111
- 0000000 → null character; 1111111 → delete character
- There are 31 control (nonprintable) characters
- The numeric characters (0~9) are coded before letters
- There are several special printable characters
- The uppercase letters (A~Z) come before the lowercase letters (a~z)
- The upper and lowercase characters are distinguished by only 1 bit. (A →1000001; a →1100001)
- There are six special characters between the upper and lowercase letters

Image representation methods



Bitmap graphic method

- Image is divided into a matrix of pixels (picture elements), each pixel is assigned a bit pattern
- The size of the pixel depends on what is called the resolution



Image

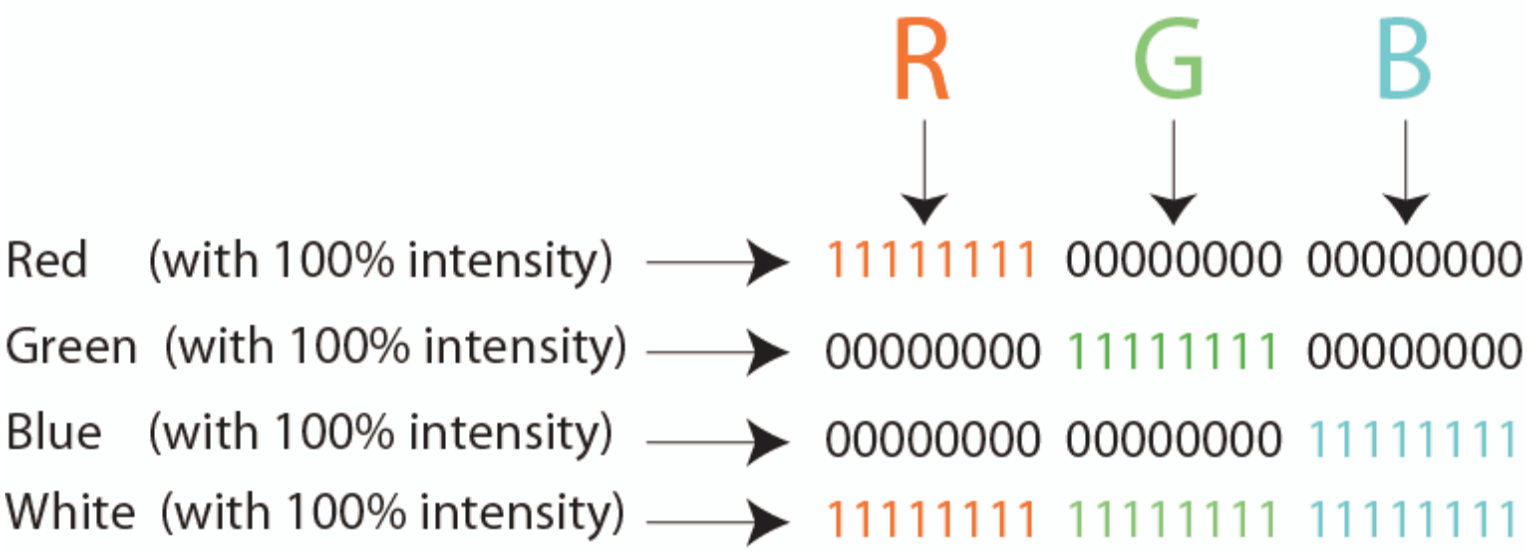
0 0 0 1 1 0 0 0
0 0 1 1 1 1 0 0
0 0 1 1 1 1 0 0
0 0 0 1 1 0 0 0

Matrix Representation

00011000 00111100 00111100 00011000

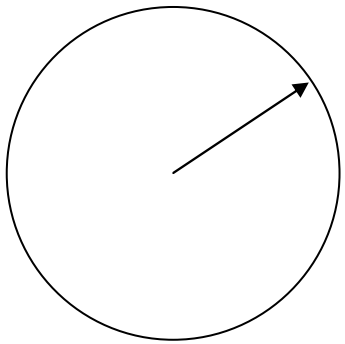
Linear Representation

Representation of color pixels



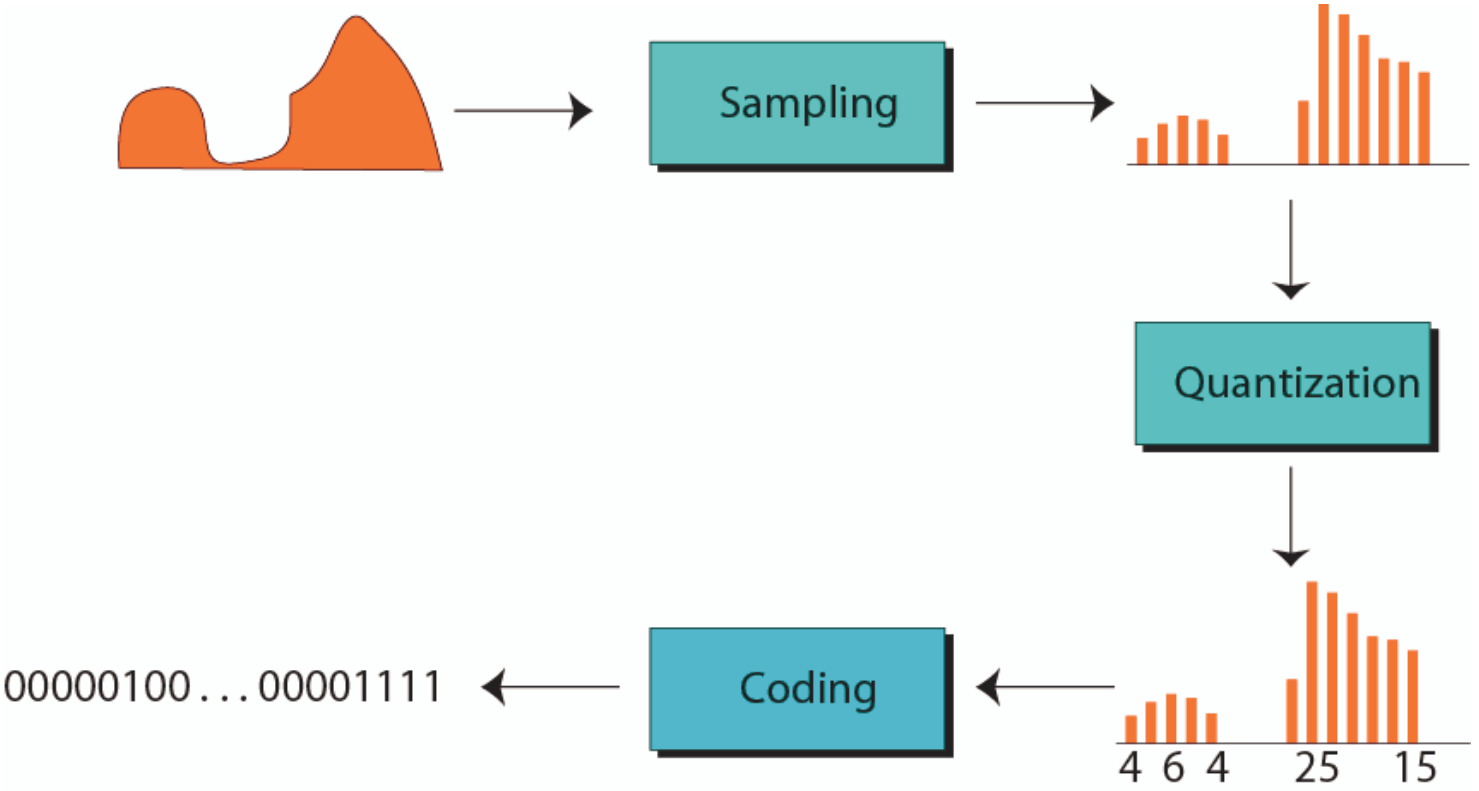
Vector graphic

- If you want to rescale the bitmap graphic image, which creates a ragged or grainy look
- An image is decomposed into a combination of curves and lines
- Each curve or line is represented by a mathematical formula
- Each time an image is drawn, the formula is reevaluated



Audio representation

Audio data are transformed to bit patterns through sampling, quantization, and coding



Video

- Video is a representation of images (called frames) in time, the combination of the images represents the video
- A movie is a series of frames shown one after another to create the illusion of motion
- Today video is normally compressed: MPEG

2.4

HEXADECIMAL NOTATION

Note:

A 4-bit pattern can be represented by a hexadecimal digit, and vice versa.

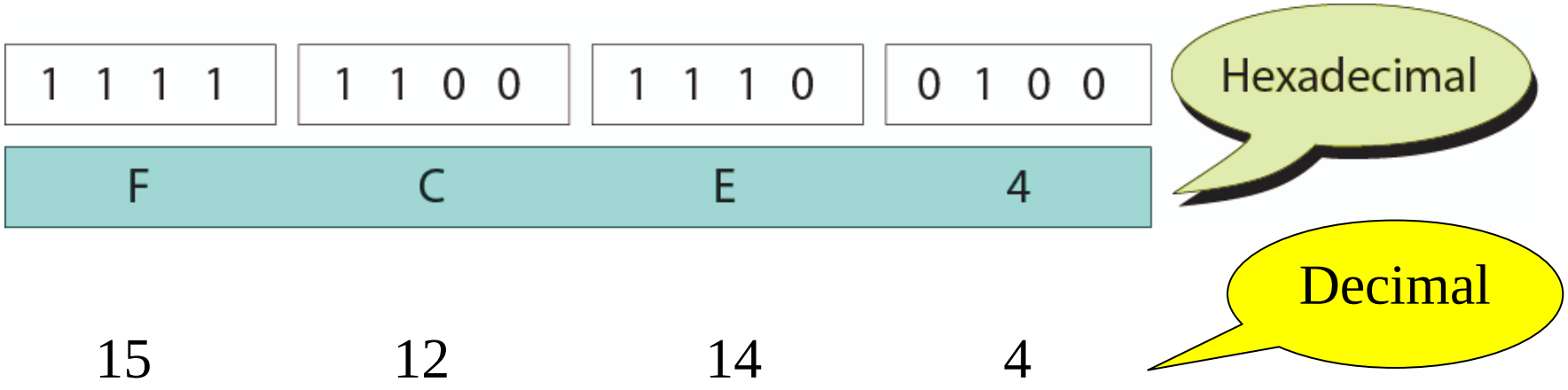
Hexadecimal digits

Bit Pattern	Hex Digit	Bit Pattern	Hex Digit
0000	0	1000	8
0001	1	1001	9
0010	2	1010	A
0011	3	1011	B
0100	4	1100	C
0101	5	1101	D
0110	6	1110	E
0111	7	1111	F

People find it difficult to manipulate bit patterns. Writing a long stream of 0s and 1s is tedious and prone to error

Binary to hexadecimal and

Hexadecimal to binary transformation



Example 1

Show the hexadecimal equivalent of the bit pattern 1100 1110 0010.

Solution

Each group of 4 bits is translated to one hexadecimal digit. The equivalent is 0xCE2.

Example 2

Show the hexadecimal equivalent of the bit pattern 0011100010.

Solution

Divide the bit pattern into 4-bit groups (from the right). In this case, add two extra 0s at the left to make the number of bits divisible by 4. So you have 000011100010, which is translated to 0x0E2.

Example 3

What is the bit pattern for 0x24C?

Solution

Write each hexadecimal digit as its equivalent bit pattern to get 0010 0100 1100.

2.5

OCTAL NOTATION

Note:

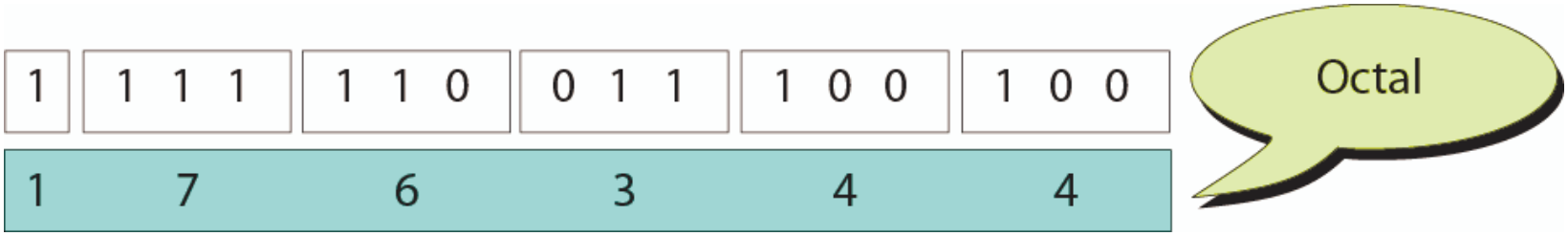
A 3-bit pattern can be represented by an octal digit, and vice versa.

Octal digits

Bit Pattern	Oct Digit	Bit Pattern	Oct Digit
000	0	100	4
001	1	101	5
010	2	110	6
011	3	111	7

Binary to octal and

Octal to binary transformation



Example 4

Show the octal equivalent of the bit pattern 101110010.

Solution

Each group of 3 bits is translated to one octal digit. The equivalent is 0562, o562, or 562_8.

Example 5

Show the octal equivalent of the bit pattern 1100010.

Solution

Divide the bit pattern into 3-bit groups (from the right). In this case, add two extra 0s at the left to make the number of bits divisible by 3. So you have 001100010, which is translated to 142_8.

Example 6

What is the bit pattern for 24_8?

Solution

Write each octal digit as its equivalent bit pattern to get 010100.

Example 7

What is the decimal equivalent for 24_8?

Solution

- $2 * 8 + 4 = 20.$
- 010100_2
  $= 0 * 2^5 + 1 * 2^4 + 0 * 2^3 + 1 * 2^2 +$
  $0 * 2^1 + 0 * 2^0$
  $= 16 + 4 = 20.$

HW

- Find out the EBCDIC code of ‘C’, ‘S’, ‘I’, ‘E’, and compare them with their ASCII codes.
- Find out the EBCDIC code of ‘0’, ‘5’, ‘9’, and compare them with their ASCII codes.
- Describe the difference which you observed between the design of EBCDIC and ASCII tables.

如夢令

昨夜雨疏風驟，濃睡不消殘酒。

試問捲簾人，卻道海棠依舊。

知否？知否？

應是綠肥紅瘦。