Chapter 2

Data Representation

OBJECTIVES

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

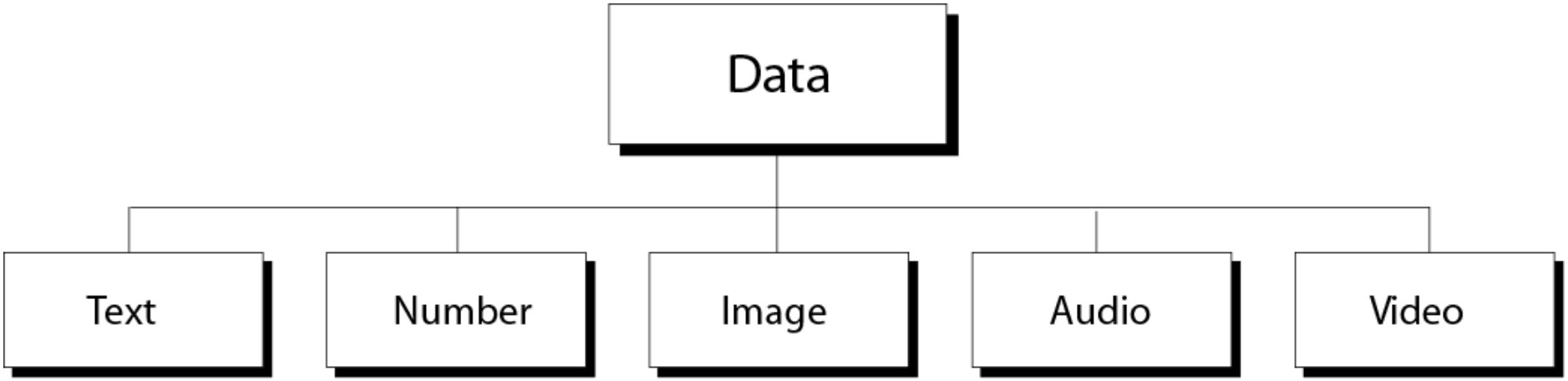
- Define data types.
- Visualize how data are stored inside a computer.
- Understand the differences between text, numbers, images, video, and audio.
- Work with hexadecimal and octal notations.

2.1

DATA TYPES

Figure 2-1

Different 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

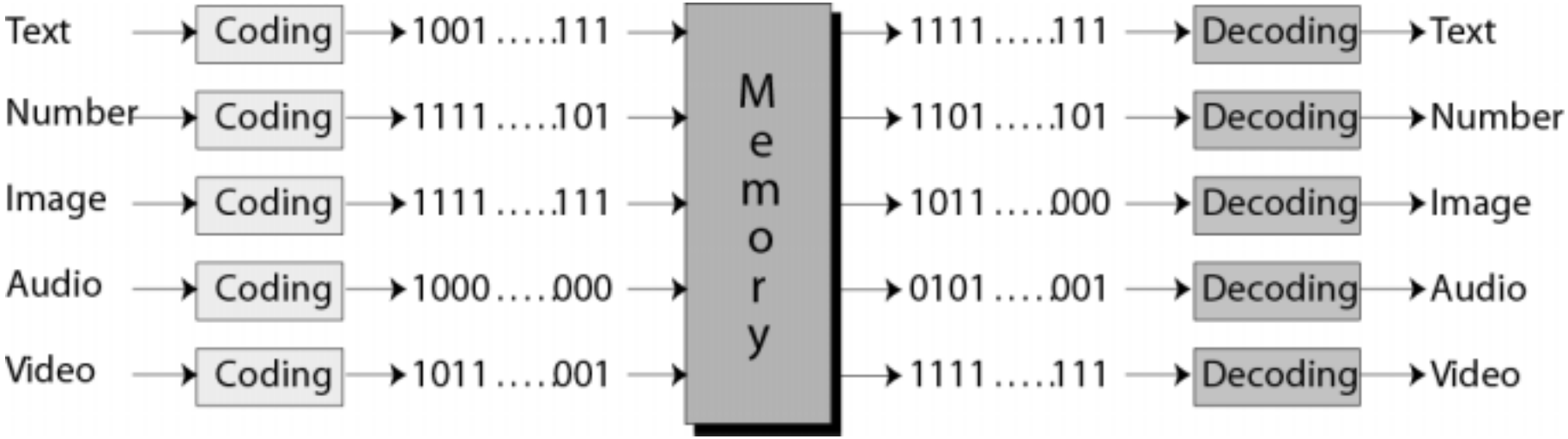
Figure 2-2

Uniform Data Representation inside a Computer: Bit pattern

bit (binary digit):
A smallest unit of data. It can be either 0 or 1.
bit pattern: a string of bits
Computer memory doesn’t know the type of data.
The data type is interpreted by programs or I/O devices.

Figure 2-3

Examples of bit patterns



A bit pattern of length 8 is called a byte.

1024 bytes=> 1 KB.

1024 KB=> 1 MB

1024 MB=> 1GB

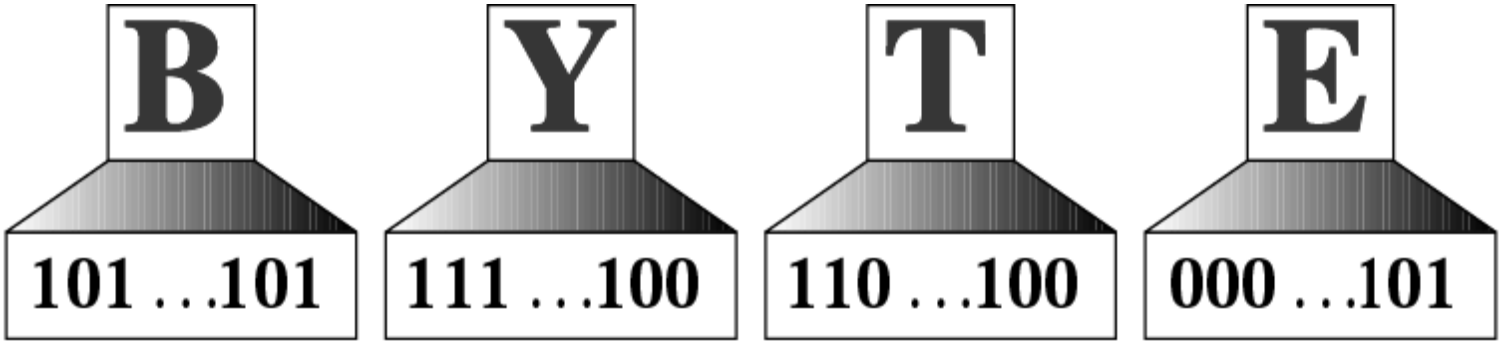
1024 GB=> 1TB

2.3

REPRESENTING DATA

Figure 2-4

Representing symbols using bit patterns



26 capital letters, 26 lower case letters, 10 digits, punctuations, etc.

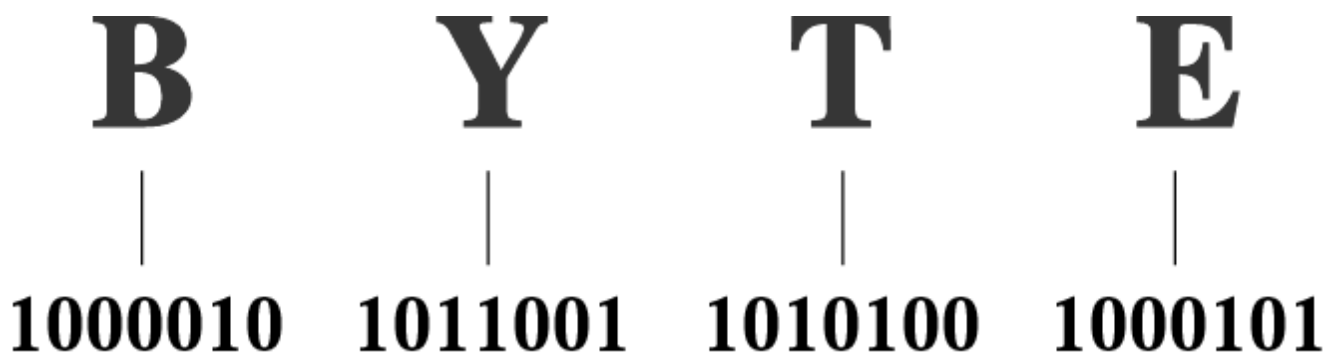
How long should a bit pattern is?

Table 2.1 Number of symbols and bit pattern length

Number of Symbols	***Bit Pattern Length***
	1
	2
	3
	4
	…
	7
	8
	…
	16

Figure 2-5

Representation of the word “BYTE” in ASCII code



ASCII: American Standard Code for Information Interchange
Use 7 bits for each symbol. (128 symbols)
See appendix A of the textbook for the ASCII table.


Ascii Table - ASCII character codes and html, octal, hex and decimal charts - Microsoft Internet Explorer
http://www.asciitable.com/
ASCII stands for American Standard Code for Information Interchange. Computers can only understand numbers, so an ASCII code is the numerical representation of a character such as 'a' or '@' or an action of some sort. ASCII was developed a long time ago and now the non-printing characters are rarely used for their original purpose. Below is the ASCII character table and this includes descriptions of the first 32 non-printing characters. ASCII was actually designed for use with teletypes and so the descriptions are somewhat obscure. If someone says they want your CV however in ASCII format, all this means is they want 'plain' text with no formatting such as tabs, bold or underscoring - the raw format that any computer can understand. This is usually so they can easily import the file into their own applications without issues. Notepad.exe creates ASCII text, or in MS Word you can save a file as 'text only'
Dec Hx Oct Html Chr
0 0 000 &#0; (null)
1 1 001 SOH (start of heading)
2 2 002 STX (start of text)
3 3 003 ETX (end of text)
4 4 004 EOT (end of transmission)
5 5 005 ENQ (enquiry)
6 6 006 ACK (acknowledge)
7 7 007 BEL (bell)
8 8 010 BS (backspace)
9 9 011 TAB (horizontal tab)
10 A 012 LF (NL line feed, new line)
11 B 013 VT (vertical tab)
12 C 014 FF (NP form feed, new page)
13 D 015 CR (carriage return)
14 E 016 SO (shift out)
15 F 017 SI (shift in)
16 10 020 DLE (data link escape)
17 11 021 DC1 (device control 1)
18 12 022 DC2 (device control 2)
19 13 023 DC3 (device control 3)
20 14 024 DC4 (device control 4)
21 15 025 NAK (negative acknowledge)
22 16 026 SYN (synchronous idle)
23 17 027 ETB (end of trans. block)
24 18 030 CAN (cancel)
25 19 031 EM (end of medium)
26 1A 032 SUB (substitute)
27 1B 033 ESC (escape)
28 1C 034 FS (file separator)
29 1D 035 GS (group separator)
30 1E 036 RS (record separator)
31 1F 037 US (unit separator)
Dec Hx Oct Html Chr
32 20 040   Space
33 21 041 ! !
34 22 042 " "
35 23 043 # #
36 24 044 $ $
37 25 045 % %
38 26 046 & &
39 27 047 ' '
40 28 050 ((
41 29 051))
42 2A 052 * *
43 2B 053 + +
44 2C 054 , ,
45 2D 055 - -
46 2E 056 . .
47 2F 057 / /
48 30 060 0 0
49 31 061 1 1
50 32 062 2 2
51 33 063 3 3
52 34 064 4 4
53 35 065 5 5
54 36 066 6 6
55 37 067 7 7
56 38 070 8 8
57 39 071 9 9
58 3A 072 : :
59 3B 073 ; ;
60 3C 074 < <
61 3D 075 = =
62 3E 076 > >
63 3F 077 ? ?
Dec Hx Oct Html Chr
64 40 100 @ @
65 41 101 A A
66 42 102 B B
67 43 103 C C
68 44 104 D D
69 45 105 E E
70 46 106 F F
71 47 107 G G
72 48 110 H H
73 49 111 I I
74 4A 112 J J
75 4B 113 K K
76 4C 114 L L
77 4D 115 M M
78 4E 116 N N
79 4F 117 O O
80 50 120 P P
81 51 121 Q Q
82 52 122 R R
83 53 123 S S
84 54 124 T T
85 55 125 U U
86 56 126 V V
87 57 127 W W
88 58 130 X X
89 59 131 Y Y
90 5A 132 Z Z
91 5B 133 [[
92 5C 134 \ \
93 5D 135]]
94 5E 136 ^ ^
95 5F 137 _ _
Dec Hx Oct Html Chr
96 60 140 ` `
97 61 141 a a
98 62 142 b b
99 63 143 c c
100 64 144 d d
101 65 145 e e
102 66 146 f f
103 67 147 g g
104 68 150 h h
105 69 151 i i
106 6A 152 j j
107 6B 153 k k
108 6C 154 l l
109 6D 155 m m
110 6E 156 n n
111 6F 157 o o
112 70 160 p p
113 71 161 q q
114 72 162 r r
115 73 163 s s
116 74 164 t t
117 75 165 u u
118 76 166 v v
119 77 167 w w
120 78 170 x x
121 79 171 y y
122 7A 172 z z
123 7B 173 { {
124 7C 174 | |
125 7D 175 } }
126 7E 176 ~ ~
127 7F 177 &#127; DEL
Source: www.asciitable.com
FREE orders over $25
Click here
amazon.com
Privacy Information
Sponsored Links
Debt Consolidation Loans Book Cheap Flights Purchase Discount Airfare Cheap Phone Cards
shop now AND SAVE!
Toys
amazon.com.
Click Here
Netgear MR814 802.11b Wi...
Netgear
Now $54.99!
Lexar Media 256 MB JumpD...
Lexar Media
Now $69.99!
Viking 256 MB CompactFla...
Viking Components
Now $69.99!
AsciiTable.com
不明的區域 (混合)
開始
Ascii Table - ASCII ...
index.doc - Microso...
Microsoft PowerPoi...
下午 02:44

In Class Exercise

- Find the ASCII representation for your English name.
- A (65) I (73) D (68) A (65)
- (1000001 1001001 1000100 1000001)

Features of ASCII

- 7-bit pattern from 0000000 to 1111111
- 0000000 represents null. (lack of character)
- 1111111 represents delete.
- The first 32 (from 0 to 31) are control characters.
- Difference between capital and lower case letter:
  - The sixth bit from the right
  - A:1000001
  - a: 1100001

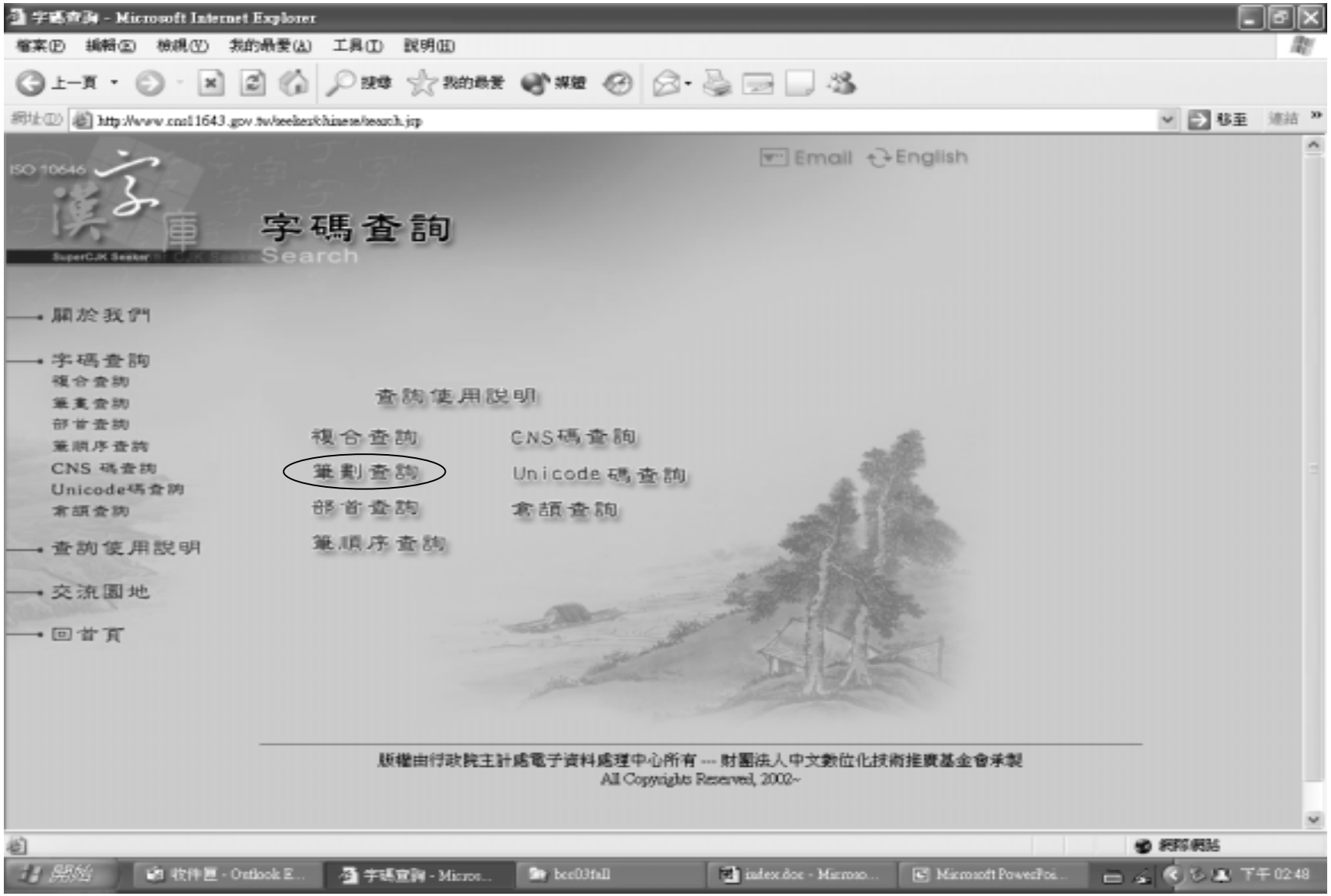
Other Representations.

- Extended ASCII
  - The leftmost bit of ASCII is 0.
  - Extended ASCII use 8 bits. (256 symbols.)
- EBCDIC (Extended binary coded decimal interchange code)
  - Developed by IBM
  - 8 bits to represents 256 symbols.

Other Representations.

- Unicode
  - Represent symbols other than English.
  - Use 16 bits (65536 symbols)
- ISO (International Organization of Standardization)
  - 32 bits
  - 4294967296 symbols.

http://www.cns11643.gov.tw/seeker/chinese/search.jsp



After Class Exercise

- 使用上述投影片的網址找中文名字的 unicode

Figure 2-6

Image representation methods

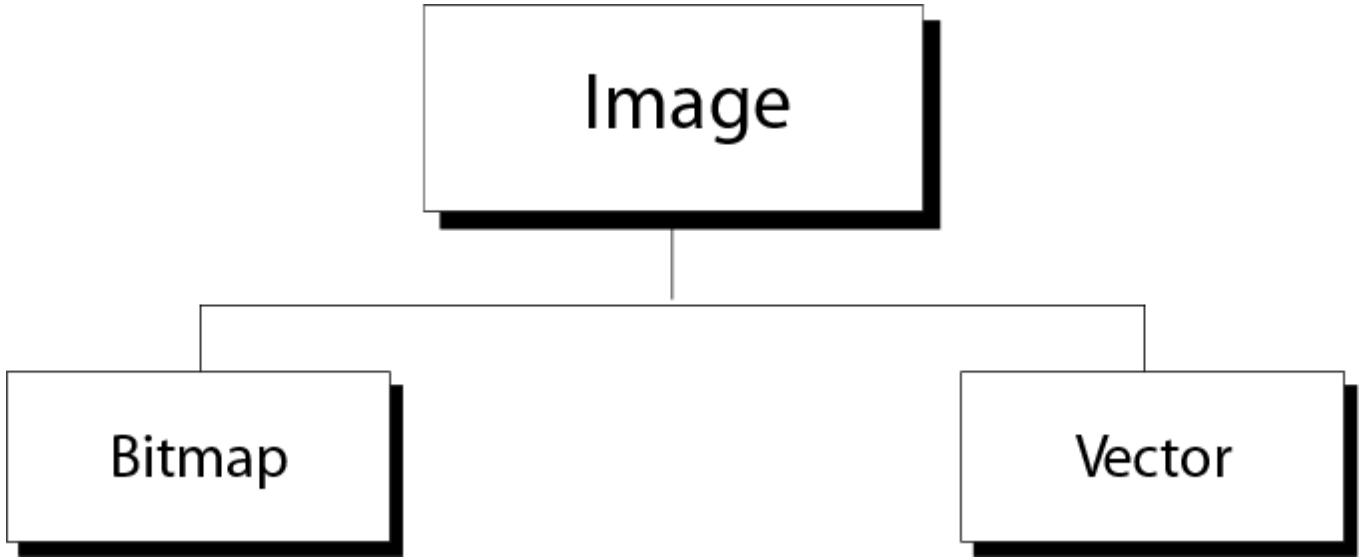
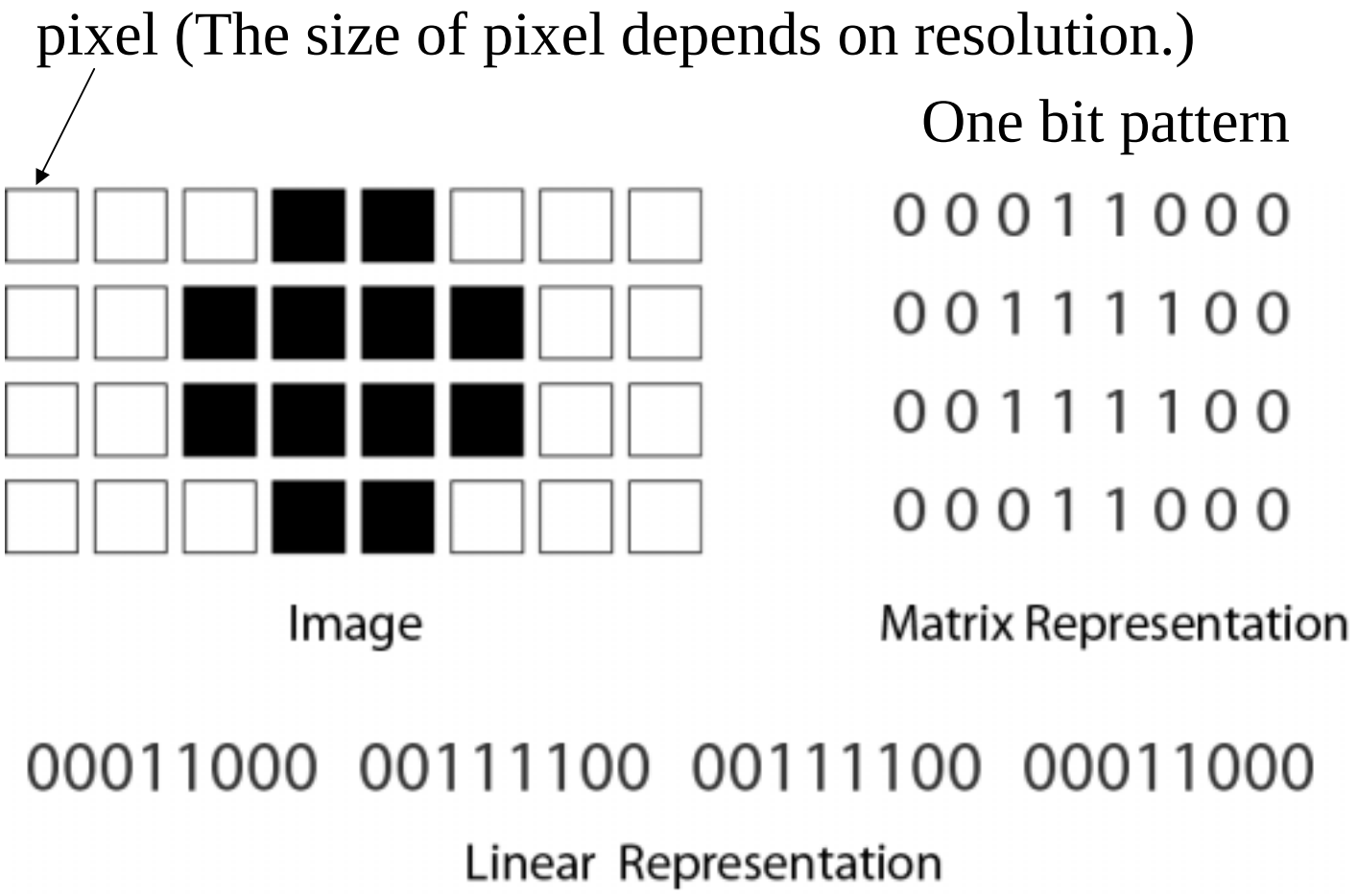


Figure 2-7

Bitmap graphic method of a black-and-white image



Representation of Images

- The baby's picture with smaller pixels - more detail.

- The baby's picture with 4 levels of gray.

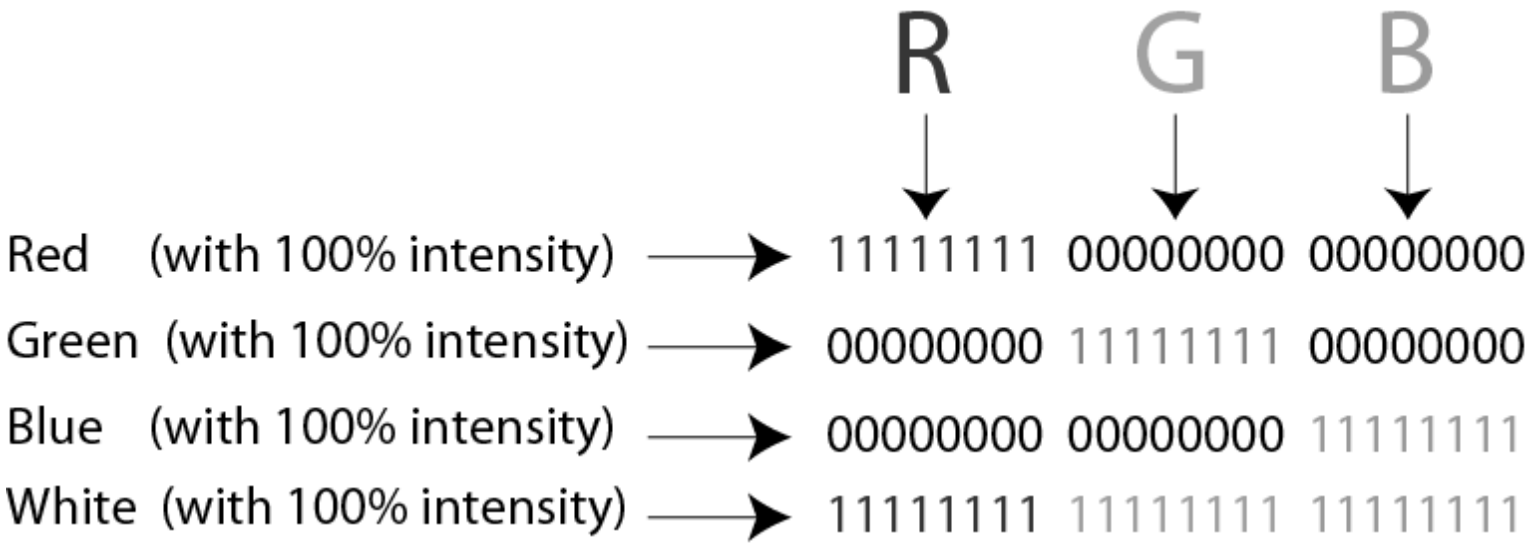
Representation of Images

- Photographic quality images have a **gray-scale.**
  - Several shades between black and white are used.
  - 4 level gray-scale means 4 shades are used.
    - Each pixel needs 2 bits:
    - 00 - represents white
    - 01 - represents light gray
    - 10 - represents dark gray
    - 11 - represents black
  - 256 level gray scale means
    - 8 bits per pixel are needed for 256 shades of gray

256 levels of gray

Figure 2-8

Representation of color pixels

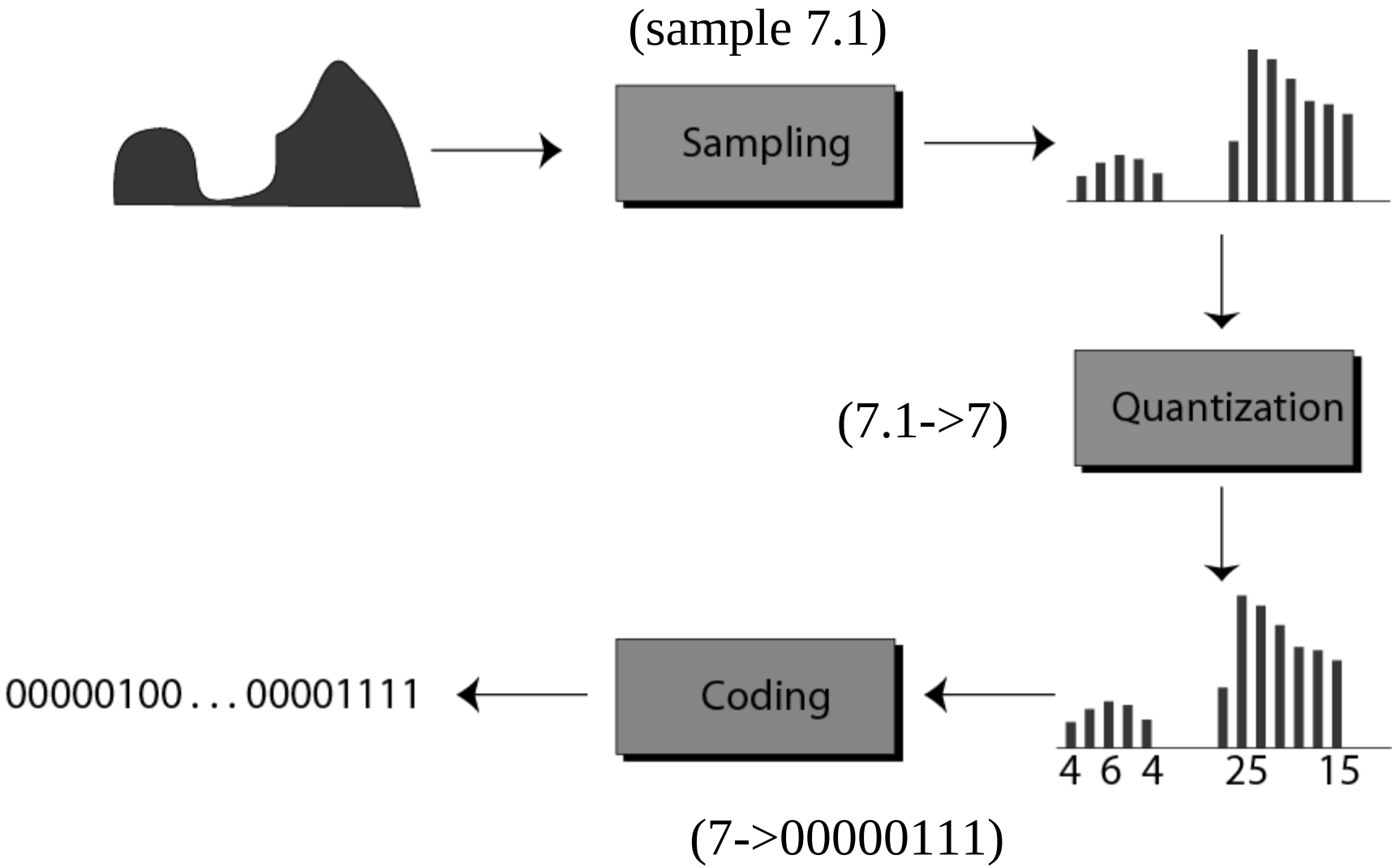


Vector Graphic

- Bit graphic causes problems when rescaling the image.
- Vector graphic decompose a graph into curves and lines. Each of the curves are represented by a formula.
- When rescaling, the computer reevaluating the formulas of the grpah.

Figure 2-9

Audio representation



Video

- Video is a representation of images in time.
- A movie is a series of images shown one after another.
- Each image is changed into a set of bit patterns and stored.
  - like MPEG file.

2.4

HEXADECIMAL NOTATION

Note:

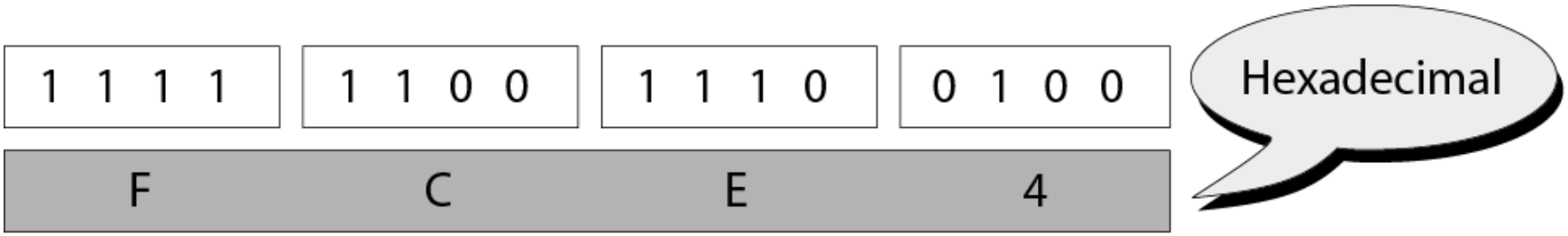
A 4-bit pattern can be represented by a hexadecimal digit, and vice versa.

Table 2.2 Hexadecimal digits

Bit Pattern	***Hex Digit***	*Bit Pattern*	***Hex Digit***
	0		8
	1		9
	2		A
	3		B
	4		C
	5		D
	6		E
	7		F

Figure 2-10

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 xCE2.

↑ Used to show this number is hexadecimal.

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 x0E2.

Example 3

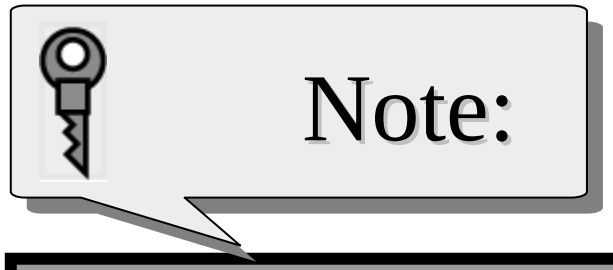
What is the bit pattern for x24C?

Solution

Write each hexadecimal digit as its equivalent bit pattern to get 001001001100.

2.5

OCTAL NOTATION



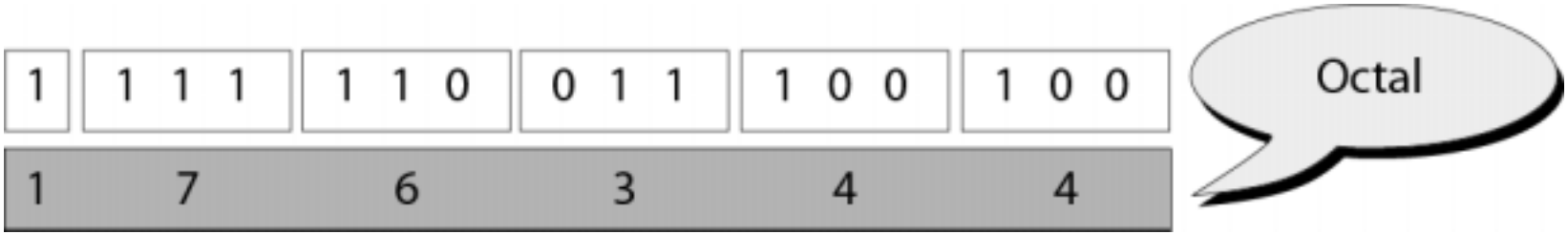
A 3-bit pattern can be represented by an octal digit, and vice versa.

Table 2.3 Octal digits

Bit Pattern	*Oct Digit*	*Bit Pattern*	*Oct Digit*
	0		4
	1		5
	2		6
	3		7

Figure 2-11

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.