MODULE 2: Data Representation

Lecture 2.5
Character Codes

Prepared By:

- Scott F. Midkiff, PhD
- Luiz A. DaSilva, PhD
- Kendall E. Giles, PhD

Electrical and Computer Engineering

Virginia Tech

Lecture 2.5 Objectives

- Describe the purpose of binary coded decimal (BCD) representation and other decimal codes
- Convert decimal values to and from binary coded decimal
- Describe the purpose and basic features of ASCII, EBCDIC, and Unicode character codes
- Using a table, be able to convert to and from ASCII character codes

Decimal Codes

- Sometimes we want to operate directly on decimal values
  - Avoids conversions
  - Special-purpose systems
- Binary coded decimal (BCD) is the usual scheme
  - Four bits per decimal digit (0000=0 to 1001=9)
  - Six combinations not used (1010 to 1111)
- There are also other schemes

Example Decimal Codes (1)

	BCD	2-4-2-1	Excess-3	Gray
0	0000	0000	0011	0010
1	0001	0001	0100	0110
2	0010	0010	0101	0111
3	0011	0011	0110	0101
4	0100	0100	0111	0100
5	0101	1011	1000	1100
6	0110	1100	1001	1101
7	0111	1101	1010	1111
8	1000	1110	1011	1110
9	1001	1111	1100	1010

Example Decimal Codes (2)

- 2421 and Excess-3 are self-complementary codes
  - Complementing the bits gives the 9's complement
  - Excess-3 example: 6 = 1001, 3 = 0110
- BCD and 2421 are weighted codes
  - For 2421, each position is assigned a weight (2-4-2-1)
  - Standard BCD is an 8-4-2-1 code
- Gray code is an unweighted code with the special property that adjacent values differ in just one bit
  - This property makes it a cyclic code

BCD Examples (1)

- Represent $(4509)_{10}$ as a BCD value
  - 4 = 0100
  - 5 = 0101
  - 0 = 0000
  - 9 = 1001
- So, $(4509)_{10} = (0100\ 0101\ 0000\ 1001)_{BCD}$
- Note that this representation is less compact than radix-2
- Combinations 1010, 1011, 1100, 1101, 1110, 1111 are “wasted” in BCD

BCD Examples (2)

- What is $(0100\ 1001\ 1000\ 0100\ )_{BCD}$ in decimal?
  - $(0100)_2 = 4$
  - $(1001)_2 = 9$
  - $(1000)_2 = 8$
  - $(0100)_2 = 4$
- So, $(0100\ 1001\ 1000\ 0100\ )_{BCD} = (4984)_{10}$

CHECK POINT

As a checkpoint of your understanding, please pause the video and make sure you can do the following:

- Represent $(5490)_{10}$ as a BCD value

If you have any difficulties, please review the lecture video before continuing.

Non-Numeric Data

- Computers need to represent more than just numbers
  - Alphabetic characters and other symbols
  - Numerals (versus numbers)
  - Special control codes (e.g., backspace, carriage return)
- ASCII is the most widely used code for this purpose
- Unicode is growing in use
- EBCDIC used in IBM mainframe systems

ASCII (1)

- ASCII: American Standard Code for Information Interchange
- Widely used for exchange of data and internal representation
- Seven bits (sometimes 8 to extend the character set or for 1 error check bit)
- Used for
  - “Printing” characters (0...9a...zA...Z!@#$...)
  - “Control” characters (CR, BS, SP, ESC, ...)

ASCII (2)

- Example ASCII codes

A	100 0001	$(41)_{16}$
a	110 0001	$(61)_{16}$
DEL	111 1111	$(7F)_{16}$
ESC	001 1011	$(1B)_{16}$
0	011 0000	$(30)_{16}$
9	011 1001	$(39)_{16}$

EBCDIC

- EBCDIC: Extended Binary Coded Decimal Interchange Code
- Less widely used than ASCII; used mainly in IBM mainframes
- 8-bit code
- Has more control characters than ASCII

Unicode

- ASCII and EBCDIC support the Latin character set typically used in computers, but cannot represent many alphabets used in the world
- Unicode designed to address this problem and is now an industry standard
- UTF-16 is a 16-bit code that represents many more characters than ASCII or EBCDIC
- Can represent many special printable characters (such as à, ê, æ, ü, etc.)

CHECK POINT

As a checkpoint of your understanding, please pause the video and make sure you can do the following:

- What is the binary value and hexadecimal value for the ASCII character “9”?

If you have any difficulties, please review the lecture video before continuing.

Summary

- Computers and special-purpose digital systems sometimes represent decimal values in decimal form
    - Four bits used to represent a decimal digit (0, 1, …, 9)
    - Intel x86 math co-processor architecture supports 20-digit (80 bit) "Packed Decimal" format
- Other weighted and unweighted codes may also be used to represent decimal values
- Computers must represent non-numeric data including characters and special control codes
- Commonly used non-numeric codes: ASCII (7 bits), EBCDIC (8 bits), Unicode (16 bits)

MODULE 2: Data Representation

Lecture 2.5
Character Codes

Prepared By:

- Scott F. Midkiff, PhD
- Luiz A. DaSilva, PhD
- Kendall E. Giles, PhD

Electrical and Computer Engineering

Virginia Tech