G.C.E. O/L Grade 10 ICT · Lesson 3

3.5 Coding Systems – BCD, EBCDIC, ASCII and Unicode

When you type the letter **A**, the computer does not store a picture of 'A' – it stores a **code number** in binary. A **coding system** is an agreed table that links characters to binary codes. In this note you will learn **BCD, EBCDIC, ASCII and Unicode**, where each is used, and their limitations.

1. Why do we need coding systems?

Computers understand only 0 and 1. To store letters, digits and symbols, each character is given a unique number. If all computers use the **same code**, a message typed on one device can be read correctly on another.

2. BCD – Binary Coded Decimal

BCD codes **each decimal digit separately in 4 bits**. It is used in digital clocks, calculators and electronic displays where digits are shown one by one.

Decimal	BCD	Pure binary
7	0111	111
59	0101 1001	111011
2026	0010 0000 0010 0110	11111101010

💡 **Note:** BCD of 59 (0101 1001) is NOT the same as binary of 59 (111011). BCD uses only codes 0000–1001; 1010–1111 are not used, so it wastes space and **cannot code letters**.

3. EBCDIC

EBCDIC (Extended Binary Coded Decimal Interchange Code) is an **8-bit** code made by **IBM** for its mainframe computers. It can code 256 characters. Example: 'A' = $C1_{16}$ = 11000001_2.

- Limitation: used mainly on IBM mainframes only – not compatible with ASCII.
- Letters are not in one continuous block, so sorting is harder.

4. ASCII

ASCII (American Standard Code for Information Interchange) is the most common code for English text. **Standard ASCII uses 7 bits → 2^7 = 128 characters** (codes 0–127). **Extended ASCII uses 8 bits → 256 characters**, adding symbols like é, £ and line-drawing characters.

© 2026 schoolict.net (Ruwan Mangala Suraweera) · All rights reserved

Character	Decimal	Binary (7-bit)	Hex
A	65	1000001	41
B	66	1000010	42
Z	90	1011010	5A
a	97	1100001	61
0	48	0110000	30
9	57	0111001	39
space	32	0100000	20

💡 **Exam shortcuts:** 'A' = 65, 'a' = 97 (lower case = upper case + 32), '0' = 48. So 'C' = 67, 'c' = 99 and '5' = 53.

Worked example: the word **CAT** in ASCII = 67 65 84 = 1000011 1000001 1010100.

- Limitation: only 128 (or 256) characters – no space for Sinhala, Tamil, Chinese or emoji.

5. Unicode

Unicode gives a unique code (code point) to **every character in every language**, plus symbols and emoji – over 150,000 characters. It is written as U+ followed by hex. The first 128 Unicode codes are the same as ASCII. It is stored using **UTF-8** (1–4 bytes per character) or UTF-16.

Character	Code point	UTF-8 bytes
A	U+0041	1
අ	U+0D85	3
ඛ	U+0D9A	3
😀	U+1F600	4

NEW 2026 **Sinhala in Unicode:** Sinhala letters live in the block **U+0D80 – U+0DFF** (e.g. අ = U+0D85). Thanks to Unicode you can type Sinhala on Facebook, WhatsApp and this website, and emoji like 😀 (U+1F600) work on every phone.

- Limitation: needs more storage than ASCII (2–4 bytes for many characters).
- Old software and devices may not show all characters (boxes □ appear).

6. Comparison

Code	Bits	Characters	Main use	Limitation
BCD	4 per digit	10 digits	Clocks, calculators	Digits only

© 2026 schoolict.net (Ruwan Mangala Suraweera) · All rights reserved

EBCDIC	8	256	IBM mainframes	Not compatible with ASCII
ASCII	7 (8 extended)	128 (256)	English text, keyboards	No other scripts
Unicode	8–32 (UTF-8: 1–4 bytes)	150,000+	All languages, emoji, web	Uses more space

Key Words

Coding system – An agreed table that gives each character a binary code.

Character – A letter, digit, symbol or space.

BCD – Binary Coded Decimal – each decimal digit coded in 4 bits.

EBCDIC – IBM's 8-bit character code for mainframes.

ASCII – 7-bit code for 128 characters; 'A' = 65.

Extended ASCII – 8-bit version of ASCII with 256 characters.

Unicode – Universal code covering all languages and emoji.

Code point – The number Unicode gives a character, e.g. U+0D85.

UTF-8 – A way of storing Unicode using 1–4 bytes per character.

Sinhala Unicode block – The range U+0D80–U+0DFF used for Sinhala letters.

Emoji – Picture characters such as 😀, coded in Unicode.

Compatibility – Ability of different systems to read the same data correctly.

Sample Q&A

1. What is a coding system?

> **Answer:** An agreed set of binary codes for characters, so all computers store and read text the same way.

2. Write 59 in BCD.

> **Answer:** 5 → 0101, 9 → 1001 → 0101 1001.

3. How many characters can 7-bit ASCII represent? Why?

> **Answer:** 128, because $2^7 = 128$.

4. The ASCII code of 'A' is 65. What is the code of 'E' and of 'e'?

> **Answer:** 'E' = 69; 'e' = 101 (add 32).

© 2026 schoolict.net (Ruwan Mangala Suraweera) · All rights reserved

5. Why was Unicode introduced?

> **Answer:** ASCII could not represent languages such as Sinhala, Tamil or Chinese. Unicode gives a code to every character in every language.

6. Which Unicode range is used for Sinhala?

> **Answer:** U+0D80 to U+0DFF.

MCQ

1. How many characters can standard 7-bit ASCII represent?

A. 65,536
B. 256
C. 7
D. 128

2. The ASCII code of 'A' is 65. What is the code of 'D'?

A. 100
B. 67
C. 68
D. 69

3. The ASCII code of lower-case 'a' is:

A. 97
B. 1
C. 32
D. 65

4. Which coding system was developed by IBM for mainframes?

A. Unicode
B. EBCDIC
C. ASCII
D. BCD

5. What is 47 in BCD?

A. 0100 0111
B. 0100 0110
C. 0111 0100
D. 101111

© 2026 schoolict.net (Ruwan Mangala Suraweera) · All rights reserved

6. Which coding system can represent Sinhala letters?

A. BCD
B. EBCDIC
C. Unicode
D. 7-bit ASCII

7. How many characters can extended ASCII represent?

A. 512
B. 128
C. 256
D. 1024

8. BCD is mainly used in:

A. Web page Sinhala text
B. Emoji on phones
C. Image compression
D. Digital clocks and calculators

9. A main limitation of ASCII is that it:

A. Uses 32 bits per character
B. Cannot represent the letter A
C. Cannot represent most non-English scripts
D. Works only on IBM mainframes

10. The Unicode code point of the Sinhala letter අ is:

A. U+0D85
B. U+1F600
C. U+0D00
D. U+0041

11. The ASCII code of the digit '0' is:

A. 65
B. 48
C. 0
D. 30

12. The word 'HI' is stored in ASCII as:

A. 71 72
B. 72 73
C. 104 105
D. 8 9

© 2026 schoolict.net (Ruwan Mangala Suraweera) · All rights reserved

13. Emoji such as 😀 can be stored because of:

A. Unicode
B. EBCDIC
C. BCD
D. 7-bit ASCII

14. In UTF-8, how many bytes does the Sinhala letter ඇ use?

A. 1
B. 4
C. 2
D. 3

15. The 7-bit ASCII code 1000010 represents:

A. 2
B. B
C. b
D. A

Structured Questions

1. (i) What does ASCII stand for? (1 mark)
(ii) How many bits does standard ASCII use and how many characters can it code? (2 marks)
(iii) State one limitation of ASCII. (1 mark) *(4 marks)*

> (i) American Standard Code for Information Interchange (ii) 7 bits; $2^7 = 128$ characters (iii) Cannot code Sinhala/Tamil/Chinese or emoji.

2. The ASCII code of 'A' is 65. (i) Write the ASCII code of 'G' in decimal. (1 mark) (ii) Convert it to 7-bit binary. (2 marks) (iii) What is the code of 'g'? (1 mark) *(4 marks)*

> (i) 71 (ii) 1000111_2 (iii) 103

3. (i) Write 2026 in BCD. (2 marks) (ii) Write 2026 in pure binary. (2 marks) (iii) Which uses more bits? (1 mark) *(5 marks)*

> (i) 0010 0000 0010 0110 (ii) 11111101010_2 (iii) BCD uses 16 bits; binary uses 11 bits, so BCD uses more.

4. Compare ASCII and Unicode under: (i) number of characters (ii) languages supported (iii) storage needed. (3 marks) *(3 marks)*

> (i) ASCII 128/256; Unicode 150,000+. (ii) ASCII – English only; Unicode – all languages incl. Sinhala, Tamil, plus emoji. (iii) ASCII 1 byte; Unicode (UTF-8) 1–4 bytes.

© 2026 schoolict.net (Ruwan Mangala Suraweera) · All rights reserved

5. A message is stored in 7-bit ASCII as: 1000011 1000001 1010100. (i) Convert each code to decimal. (2 marks) (ii) Decode the word. (1 mark) *(3 marks)*

> (i) 67, 65, 84 (ii) CAT

6. A student types a Sinhala message on WhatsApp with a 😀 emoji. (i) Which coding system makes this possible? (1 mark) (ii) Give the Unicode range for Sinhala. (1 mark) (iii) Give one reason EBCDIC is not used for this. (1 mark) *(3 marks)*

> (i) Unicode (UTF-8) (ii) U+0D80 – U+0DFF (iii) EBCDIC has only 256 codes and is used mainly on IBM mainframes – no Sinhala or emoji.

Summary

- A coding system gives every character a binary code.
- BCD codes each decimal digit in 4 bits (clocks, calculators).
- EBCDIC: 8-bit IBM mainframe code, 256 characters.
- ASCII: 7 bits = 128 characters; extended ASCII 8 bits = 256.
- ASCII: 'A' = 65, 'a' = 97, '0' = 48.
- Unicode covers all languages and emoji; written as U+hex.
- Sinhala Unicode block: U+0D80 – U+0DFF; stored with UTF-8.
- Limits: BCD digits only, EBCDIC IBM only, ASCII English only, Unicode uses more space.

MCQ answer key – 3.5

1. **D** – $2^7 = 128$.
2. **C** – A65 B66 C67 D68.
3. **A** – 'a' = 'A' + 32 = 97.
4. **B** – EBCDIC = IBM, 8-bit.
5. **A** – 4 → 0100, 7 → 0111.
6. **C** – Sinhala block U+0D80–U+0DFF.
7. **C** – 8 bits → $2^8 = 256$.
8. **D** – BCD shows decimal digits one by one.
9. **C** – Only 128/256 characters.
10. **A** – Sinhala block starts at U+0D80; අ = U+0D85.
11. **B** – '0' = 48 (30 in hex).
12. **B** – H = 72, I = 73.
13. **A** – 😀 = U+1F600.
14. **D** – Characters from U+0800 to U+FFFF use 3 bytes.
15. **B** – $1000010_2 = 66$ = 'B'.

© 2026 schoolict.net (Ruwan Mangala Suraweera). All rights reserved.