MODULE C DATA REPRESENTATION · C8

Text and character sets: ASCII and Unicode: mark scheme

20 marks · Cambridge 0478 syllabus 1.2.1

How to use. One mark per creditworthy point up to the maximum shown. **/** separates alternative words; **//** separates alternative answers. Words in (brackets) are not needed for the mark. Every code, pattern and size was checked by program (Python). Codes are given, so none need remembering. A correct final answer with no working earns full marks, including where working is asked for. Never credit a reversed bit pattern.

QUESTION	ANSWER	MARKS	GUIDANCE
1(a) 0478 1.2.1	Computers can only process / store binary (0s and 1s) // computers are made of switches / logic gates that are on or off Each character is given a unique binary code // a character set is used to convert each character to binary	**2**	Any 2 from the points listed, max 2.
1(b) 0478 1.2.1	All the characters (and symbols) a computer can represent, each with its own unique (binary) code	**1**	Do not accept "the alphabet" alone.
2(a) 0478 1.2.1	68	**1**	Working: A 65, B 66, C 67, D 68.
2(b) 0478 1.2.1	01000011	**1**	Working: C = 67 = 64 + 2 + 1. Must be 8 bits, as the question asks. Do not accept 11000010 (reversed).
2(c) 0478 1.2.1	D: 01000100 A: 01000001 B: 01000010	**3**	Working: D = 68 = 64 + 4; A = 65 = 64 + 1; B = 66 = 64 + 2. 1 mark per correct row. Each must be 8 bits in the correct order.
3(a) 0478 1.2.1	01100100 = 100 = d 01101001 = 105 = i 01100111 = 103 = g	**3**	Word: **dig**. Working: 64 + 32 + 4 = 100; 64 + 32 + 8 + 1 = 105; 64 + 32 + 4 + 2 + 1 = 103; counting on from a = 97. 1 mark per correct character in the correct position. Upper case (DIG): max 2.
4(a) 0478 1.2.1	Unicode can represent more characters // a greater range of characters than ASCII Unicode includes other languages / alphabets // emojis; ASCII is mainly English characters Unicode uses more bits per character // ASCII uses 7 or 8 bits, Unicode up to 32	**2**	Any 2 from the points listed, max 2. Each difference must be clear about which is which. Do not accept "Unicode uses fewer bits".

QUESTION	ANSWER	MARKS	GUIDANCE
4(b) 0478 1.2.1	ASCII has a limited number of characters // 256 (or 128) characters ASCII cannot represent the characters of many (non-English) languages ASCII has no codes for emojis Unicode uses more bits per character … so it has a code for many more characters // languages and emojis So all users' messages / emojis display correctly	**3**	Any 3 from the points listed, max 3. Not linked to the languages or emojis in the scenario: max 2.
5(a) 0478 1.2.1	40 × 8 = 320 bits 320 ÷ 8 = 40 bytes	**2**	Correct answer 40 (bytes) with no working: 2. Accept 40 × 1 byte = 40.
5(b) 0478 1.2.1	160 (bytes)	**1**	Working: 40 × 32 = 1280 bits; 1280 ÷ 8 = 160 bytes // 40 × 4 bytes = 160.
5(c) 0478 1.2.1	The file is larger // four times larger	**1**	Do not accept "more characters" alone.

Total: 20 marks