Characters and ASCII

2.2.1 · Data · revision notes

How computers store text by giving every character a unique binary code, using 7-bit ASCII. Expect questions on how many characters 7 bits can represent, using the order of the codes to work out other codes, the number of bits needed for some text, and why ASCII is limited.

Key facts

- Character set: the characters a computer can represent, each with a unique binary code
- Codes are consecutive: A, B, C… and 0, 1, 2… follow each other
- Lower-case code = upper-case code + 32
- Spaces, punctuation and control characters are characters too
- 7-bit ASCII: 2^7 = 128 characters (codes 0 to 127)
- 'A' = 65, 'a' = 97, '0' = 48 (useful for checking)
- Size of text in bits = number of characters × bits per character

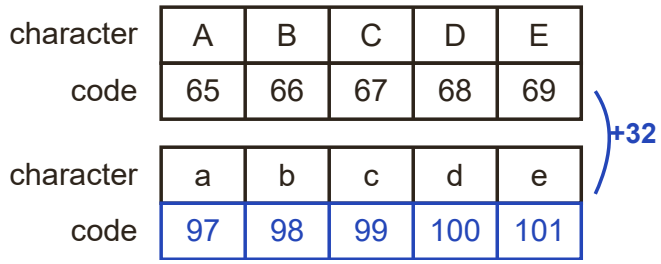
Notes

Character sets and ASCII

- Text is stored as a sequence of characters, and each character is given a unique binary code.
- A **character set** is the list of characters a computer can represent, with the code for each one. Computers must use the same character set for text to be shown correctly.
- **ASCII** uses **7 bits** for each character, so it has 2^7 = **128** codes (0 to 127).
- ASCII includes upper-case letters, lower-case letters, the digits 0 to 9, punctuation, symbols, the space, and control characters (such as new line) that are not printed.

How the codes are ordered

- The codes are in order, so you can count on from a code you are given: if 'A' is 65 (100 0001), then 'E' is 69 (100 0101).
- Upper-case letters come before lower-case letters: 'A' is 65 and 'a' is 97. Each lower-case letter's code is 32 more than its upper-case letter's code.

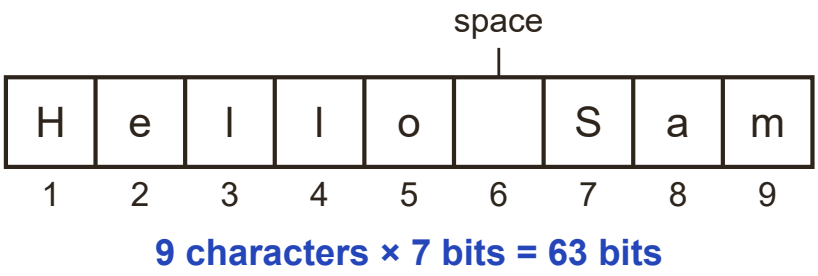


Count on from a code you know. Each lower-case letter is 32 more than its capital

- Digit characters have codes too: '0' is 48, up to '9', which is 57.
- The character '5' (code 53) is not the same as the number 5. A number typed in as text must be converted (e.g. with int() in Python) before you can calculate with it.
- You don't need to memorise the whole table: questions normally give you a code or a table to work from.

Storing text

- Bits for some text = number of characters × 7. Spaces and punctuation marks are characters too.
- Example: 'Hello Sam' has 9 characters (including the space), so it needs 9 × 7 = 63 bits.



Every character counts, including spaces and punctuation

- Limitation: 128 codes are not enough for accented letters (such as é), other alphabets (such as Greek, Arabic or Chinese) or emoji. Character sets that use more bits per character can represent many more characters, but every character then takes up more space. Grade 7+ only

How to answer each type of question

Work out a code or decode a message

1 to 2 marks | Grade 6

1. Find the code you are given as a starting point (e.g. 'K' = 75).
2. Count on or back through the alphabet or the digits.
3. Convert between denary and 7-bit binary if the question uses binary codes.

Example. The ASCII code for 'K' is 75.
(a) State the denary ASCII code for 'N'.
(b) Give the 7-bit binary ASCII code for 'H'.

Model answer.

(a) 78 (1)

(b) 'H' is 3 letters before 'K', so its code is 72 = 100 1000 (1)

Calculate the number of bits needed to store text

1 to 2 marks | Grade 5

1. Count every character, including spaces and punctuation.
2. Multiply by the number of bits per character (7 for 7-bit ASCII).
3. Give the unit (bits).

Example. A text message says `See you at 6!`
Calculate the number of bits needed to store the message using 7-bit ASCII.

Model answer.

13 characters, including the 3 spaces and the ! (1)

13 × 7 = 91 bits (1)

Explain a limitation of 7-bit ASCII

2 marks | Grade 7

1. Give the number of characters ASCII can represent (128).
2. Name what is missing in the context (accented letters, another alphabet, emoji).
3. If there is room, say that more bits per character are needed.

Example. A travel company wants its booking app to display place names in Greek and in Japanese. Explain why 7-bit ASCII is not suitable.

Model answer. 7-bit ASCII can only represent 128 different characters (1). It has no codes for Greek or Japanese characters, and Japanese alone needs thousands of characters, far more than 128 (1).

Identify the number of characters (multiple choice) 1 mark **Grade 4**

1. Work out 2^n for the number of bits given.
2. Watch out for the option that is one less (the largest code, not the number of codes).

Example. Identify the number of different characters that can be represented using 7-bit ASCII.
A 7
B 64
C 127
D 128

Model answer. D, 128 (2^7) (1)

Don't lose marks

- Forgetting to count spaces and punctuation when working out the size of some text.
- Saying 7-bit ASCII has 127 characters. There are 128 codes (0 to 127).
- Mixing up upper and lower case: 'h' does not have the same code as 'H'.
- Treating the digit character '9' as if it were the number 9 in a calculation.
- Using 8 bits per character when the question says 7-bit ASCII (or 7 bits when it says each character is stored in one byte).

More tips

Memory tricks

- 'A' is 65 and 'a' is 97: capitals first, then lower case, 32 apart.
- Upper-case letter code = 64 + its position in the alphabet (A = 1 gives 65; Z = 26 gives 90). Lower case is 96 + its position.
- Digit character code = 48 + the digit, so '7' is 55.
- When counting characters, point at each one in turn, including every space and punctuation mark.

Exam technique

- Use the codes the question gives you and show your counting, e.g. 'K = 75, so N = 78'.
- In 'explain' questions about ASCII's limits, give the number (128) and a specific example of a missing character.
- Check how many bits per character the question uses: 7 for 7-bit ASCII, 8 if it says one byte per character.

What each grade needs

What you need to be able to do, from the first marks up to the top grade.

Grade 3 **Define a character set**
A list of the characters a computer can represent, each with its own unique binary code.

Grade 4 **State how many characters 7-bit ASCII represents**
7 bits give 2^7 = 128 different codes, from 0 to 127.

Grade 5

Work out codes from a given code

Letters and digits have consecutive codes, so if 'A' is 65 then 'D' is 68.

Grade 5

Calculate the bits needed to store text

Multiply the number of characters (including spaces) by 7 for 7-bit ASCII.

Grade 6

Convert between binary codes and characters

Convert the 7-bit code to denary, then count on or back from a code you know.

Grade 7

Explain the limitations of 7-bit ASCII

128 codes cannot include accented letters, other alphabets or most symbols and emoji.

Grade 8

Explain why '7' and 7 are stored differently

The character '7' is stored as its ASCII code, 55 (011 0111), not as the binary number 7 (000 0111).

Free notes, Full packs, practice papers and mock exams at chhetriacademy.com. Written for this site; not taken from past papers. Stuck? 1-to-1 tutoring with a free 30-minute trial lesson.