

# COM6115 L4

Number of participants: 125



## 1. Character encoding - Revision Choose all the correct answers:

20 correct answers  
out of 88 respondents



In UTF-32, characters are encoded in 4 bytes



60 votes



In UTF-16, characters are encoded in 16 bits.



44 votes

In UTF-8, characters are encoded in 1-4 bits.



45 votes

In UTF-16, texts with Latin-only characters are 4 times larger than UTF-8.



14 votes



In UTF-16, texts with Latin-only characters are 2 times larger than ASCII.



51 votes

UTF-32 contains surrogate pairs.



15 votes



## 2. Text compression - Revision Entropy sets ... for compression.

51 correct answers  
out of 78 respondents

an upper bound



27 votes



a lower bound



51 votes



### Text Compression - Revision Which of the following require a code tree for decompression? (more than one may be correct)

27 correct answers  
out of 83 respondents



Huffman Coding



70 votes

Canonical Huffman Codes



49 votes

Arithmetic Coding

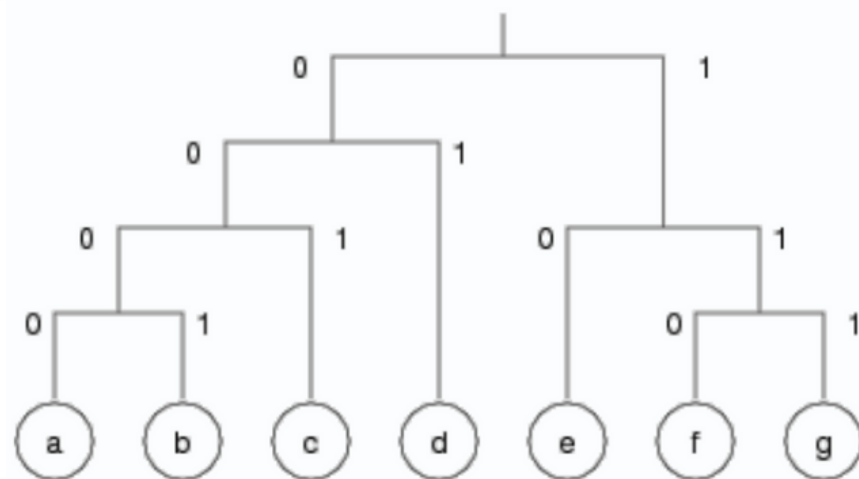


12 votes



### Text Compression - Revision Using Huffman 4. encoding and the given code tree, choose the correct answers:

31 correct answers  
out of 72 respondents

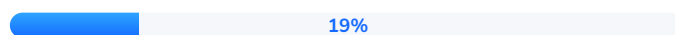


'bag' is coded as  
00010000111



52 votes

'bag' is coded as  
10000000111



14 votes



'faced' is coded as  
11000000011001



38 votes

'faced' is coded as  
01100001000110



19 votes