# **Problem Set #3**

## **Question 1**

Consider an alphabet with five letters, {a,b,c,d,e}, and suppose we know the frequencies fa​=0.32, fb​=0.25, fc​=0.2, fd​=0.18, and fe​=0.05. What is the expected number of bits used by Huffman's coding scheme to encode a 1000-letter document?

* 2230
* 2400
* 3000
* 3450

## **Question 2**

Under a Huffman encoding of nnn symbols, how long (in terms of number of bits) can a codeword possibly be?

* ln⁡n\ln nlnn
* n
* n−1n-1n−1
* 2​n

## **Question 3**

Which of the following statements holds for Huffman's coding scheme?

* If the most frequent letter has frequency less than 0.33, then all letters will be coded with at least two bits.
* A letter with frequency at least 0.5 might get encoded with two or more bits.
* If a letter's frequency is at least 0.4, then the letter will certainly be coded with only one bit.
* If the most frequent letter has frequency less than 0.5, then all letters will be coded with more than one bit.

## **Question 4**

Which of the following is true for our dynamic programming algorithm for computing a maximum-weight independent set of a path graph? (Assume there are no ties.)

* The algorithm always selects the maximum-weight vertex.
* As long as the input graph has at least two vertices, the algorithm never selects the minimum-weight vertex.
* If a vertex is excluded from the optimal solution of a subproblem, then it is excluded from the optimal solutions of all bigger subproblems.
* If a vertex is excluded from the optimal solution of two consecutive subproblems, then it is excluded from the optimal solutions of all bigger subproblems.

## **Question 5**

Recall our dynamic programming algorithm for computing the maximum-weight independent set of a path graph. Consider the following proposed extension to more general graphs. Consider an undirected graph with positive vertex weights. For a vertex v, obtain the graph G′(v) by deleting v and its incident edges from G, and obtain the graph G′′(v) from G by deleting v, its neighbors, and all of the corresponding incident edges from G. Let OOPT(H) denote the value of a maximum-weight independent set of a graph H. Consider the formula OPT(G)=max⁡{OPT(G′(v)),wv+OPT(G′′(v))}, where v is an arbitrary vertex of G of weight wv​. Which of the following statements is true?

* The formula is correct in path graphs but is not always correct in trees.
* The formula is always correct in trees, and it leads to an efficient dynamic programming algorithm.
* The formula is always correct in trees, but does not lead to an efficient dynamic programming algorithm.
* The formula is always correct in general graphs, and it leads to an efficient dynamic programming algorithm.