Examiners' commentary 2017–2018

CO3325 Data compression – Zone B

General remarks

Data Compression is a technical subject and the examination questions are often precise and specific. Each question may consist of a number of parts and each part may contain a number of sub-questions explicitly or implicitly. Questions are often labelled as "Question 1, Question 2", and so on. Parts are often labelled as "(a), (b)", and so on. Explicit sub-questions are often labelled as "i., ii." And so on. To achieve a good grade, you must understand the examination questions. For example, you would need to know at least how many sub-questions are required, and whether there are any further implicit tasks involved in each part or sub-question.

The examination can be viewed as a one-off written communication between you and the examiners. As such communication gives no second chance for clarification and is highly constrained by time, you would need to demonstrate your knowledge of the subject in the most efficient manner in the examination. For example, you may need to promptly decide on site if it would be better for you to draw a diagram with a short notation as your answer, or you would rather write a long paragraph of descriptions.

As Data compression is a level 6 course, you would also need to demonstrate your competence in problem solving. This includes sometimes sharing your interpretation of given specifics in questions. For example, you may leave a note if necessary to the examiners on your script to clarify certain issues to avoid any potential confusion, or to add assumptions to simplify your solutions.

The questions in the examination do not necessarily have an equal level of complexity. You would want to potentially secure as many marks as possible and as early as possible. For example, you may like to attempt "easy" questions first, or to give only itemised answers first and to add details later if time permits.

If you follow these guidelines as well as those in the subject guide and textbooks, your examination should be an enjoyable experience as it actually is a good opportunity for you to check the level of your knowledge and to celebrate your academic achievements.

The performance of the candidates this year was diverse. The examination grades were distributed from the high first class to the low failure.

Comments on specific questions

Question 1

a. There are three implied sub-questions in this part of the question. An easy way to answer the question is to divide your answers into three parts:
    i. Yes or No.
    ii. Justifiable reasons for your answer.
    iii. A supportive example.

b. This is a bookwork question consisting of two sub-questions and one

requirement for the second sub-question. An easy way to answer this part of the question is to first give an answer to each sub-question and then give an example for the second sub-question.

c. A good answer to this part of the question would include two sections to answer the implicit two sub-questions. The first section of your answer would include the explanation of the reasons and the second section would describe the approach of the Canonical and Minimum-variance Huffman coding.

d. An easy way to answer this part of the question is to present your answers in two sections. The first section would discuss the causes of the inefficiency with the calculation of the gap between the entropy and the average length of the code. The second section would demonstrate the differences by computing and analysing the compression efficiency of two cases, one is for the extended alphabet and the other for the original alphabet.

Question 2

a. This part of the question is straight forward. Your answer only needs to include a brief explanation and supporting example.

b. The answer to this part of the question should include three sections. The first section should include the run-length encoding. The second section should explain the meaning of each of the control characters. The third section should compute the entropy based on the frequency counts.

c. There are two sub-questions in this part of the question and you should answer them individually. The simplest way to answer the first sub- question would be to present either the encoding algorithm directly or in a flowchart. It would be fine if a program in a specific programming language is presented but this is not necessary. The second part of your answer should be a trace of the states of the required variables of the Adaptive Huffman encoding.

Question 3

a. This is a bookwork question. A good solution would consist of two parts, one is the explanation and the other is the illustration in a diagram.

b. A good answer to this part of the question would include two sections. The first section would include a brief description of the approach of the predictive encoding. The second section would include an example showing how the predictive encoding works. A better answer would include a compression technique as well as the predictive computation.

c. An easy way to answer this part of the question is to simply trace the value changes of the main variables and the dictionary for each iteration. A better answer would also highlight the final decoded message.