Problem Overview

The assignment is to create an encoding and decoding protocol to enable a student take a 20-question true/false test with a 10-bit message transmitted to him or her by a friend who is knowledgeable of the correct answers. The solution to each exam is a 20-bit binary. The encoder is aware of the actual solution, whereas the decoder is only able to get the 10-bit message and has to provide a complete set of answers. This is aimed at maximising the worst-case score, which is the minimum number of correct responses in all 20 possible exams.

High-Level Strategy

The test is divided into five blocks of four questions to be independent of each other. The blocks are coded individually with a small set of fixed representative patterns. Each block can only have four representative 4-bit patterns. These representatives are selected in a way that there is no possible block of four answers that is not different in at least one place by one of these representatives.

Two bits can be used to index each representative. A total of 5 blocks means that the length of the message is 5x 2=10 bits, which met the communication constraint. The complete decoding is associated with the choice of a representative in each block and its concatenation to create a 20-bit answer block.

Representative Patterns

The number of representatives of each block is four: [0,0,0,0], [0,0,0,1], [1,1,1,0], [1,1,1,1]

These patterns guarantee that the Hamming distance between any 4-bit block of true answers and any representative is at most 1.

Encoding Procedure

The encoder divides the 20-bit solution into five 4 bits blocks. It takes the true answers of each block and compares them with the four representatives and then picks the representative with the most matching ones (smallest Hamming distance). The chosen representative index (0-3) is then coded into two bits, and the message is attached to it. This needs to be repeated on all the five blocks to generate the final 10-bit message.

Decoding Procedure

The message received by the decoder is divided into five pairs of bits. All pairs are translated into indexes ranging between 0 and 3 that determine the representative of that block. The 4-bit representative is then placed in the correct entries of the 20-bit answer vector. It is a process that recreates the entire array of answers with the help of the received message only.

Worst-Case Performance Guarantee.

Every block is decoded by a representative which is not the actual block in a maximum of a single location. There are five blocks hence the number of incorrect answers is limited to five. Thus, the protocol ensures the following: 20-5=15 correct answers in the worst scenario.

Cheques and Counterchecks.

The implementation was tested on the given exhaustive tester that considered all 220 possible exams. The theoretical guarantee is always achieved as the tester records a worst-case score of 15.

This block-based system ensures a good balance between performance, simplicity and explainability. Although more sophisticated coding-theory methods can also achieve a reduced worst-case error, they usually require a more complicated encoding scheme. The solution offered has a high guaranteed score based on an effective and well-organised approach.

## References

1. Wikipedia, *Hamming Distance*.
   https://en.wikipedia.org/wiki/Hamming_distance
   (Used to define the measure of error between true and decoded answer vectors.)

2. Wikipedia, *Covering Code*.
   https://en.wikipedia.org/wiki/Covering_code
   (Provides background on covering codes and covering radius, which relate directly to the worst-case guarantee of the proposed scheme.)

3. R. Chapman, *Coding Theory: Definitions and Theorems*, University of Exeter Lecture Notes.
   (General background on binary codes and Hamming spaces.)