******The following questions are concerned with Chapter 1 of the notes - Basic Coding Theory.******

1 Let $C = \{00101, 11011, 10100, 10010\} \subseteq \{0,1\}^5$. Find $d(C)$. Give examples of words that do or do not have a unique nearest neighbour in C.

2 Consider the words GOTHS, HARES, HATES, MARES, MARKS, MATES, MATHS, MATEY, MITES, MOTHS, MYTHS, and RITES. Let these be the codewords of the (n, M_1, d_1) code $C_1 \subseteq \{A, B, \dots Z\}^5$. a) For each of the words PARKS, GOALS and DATES, find its nearest neighbour(s) in C_1.
b) Find n, M_1 and d_1. Now find a (n, M_2, d_2) code $C_2 \subseteq C_1$ such that $d(C_2) = 2$ and $|M_2| \geq 6$.
c) Find three codewords x, y, and z in C_1 such that $d(x, y) = d(x, z) + d(z, y)$.
d) Find three codewords x, y, and z in C_1 such that $d(x, y) < d(x, z) + d(z, y)$.

3 Let $C = \{01010, 10101, 11000, 11111\} \subseteq \{0,1\}^5$. Find $d(C)$.
How many errors can C detect? and how many can it correct?

4 Let $C = \{01234, 12340, 23401, 34012, 40123\} \subseteq \{0,1,2,3,4\}^5$. Find $d(C)$.
How many errors can C detect? and how many can it correct?

5 For fixed $n \geq 1$, how many binary $(n, 2, n)$ codes are there?

6 Let C be an (n, M, d) code with $n \geq d \geq 2$
a) Fix j with $1 \leq j \leq n$ and form C_1 by deleting the jth entry from each word in C.
Show that C_1 is a $(n-1, M, d)$ or $(n-1, M, d-1)$ code.
b) Form C_2 by deleting the last m entries of each word in C.
What can we say about the parameters of C_2 if $m < d$? How about if $m \geq d$?

7 A binary code with block length 4 is transmitted over a channel such that $P(1 \text{ received} \mid 0 \text{ sent}) = 0.1$ and $P(0 \text{ received} \mid 1 \text{ sent}) = 0.05$. Is this channel symmetric? If 0001 is sent what is the chance that 0110 is received?

8 Consider the code $C = \{c_1, c_2, c_3\} = \{000000, 110000, 111111\} \subseteq \{0,1\}^6$, and the words $w_1 = 010100$, $w_2 = 111100, w_3 = 110100, w_4 = 111110$.
a) Perform nearest neighbour decoding for each w_i. When is there no unique nearest neighbour?
b) C is sent over a binary symmetric channel with symbol-error probability p. For each $1 \leq i \leq 3$ and $1 \leq j \leq 4$, find $P(w_j \text{ received} \mid c_i \text{ sent})$.

9 For the binary code $C = \{0000, 1000, 1111\}$, the codeword 1111 is transmitted over a binary symmetric channel with symbol-error probability $p = 0.1$. We decode a received word to its unique nearest neighbour if it has one; otherwise we do not decode. What is the chance that the received word is decoded correctly? Incorrectly?

10 Consider the binary code $C = \{000, 111\}$. Suppose the codewords are transmitted over a binary symmetric channel with symbol-error probability p. Consider the following strategies:
(i) Complete decoding using nearest neighbour decoding.
(ii) Accepting a received word if it is in C but asking for retransmission otherwise.
For each strategy find the chance that, when we send 000, it is decoded correctly, perhaps after several transmissions. If $p = 0.1$, which method is more reliable? Should we therefore use this method?

11 The ternary code $C = \{01, 02, 20\}$ is transmitted over a ternary symmetric channel with error probability $p = 0.02$. We decode received words as the nearest neighbour if that is unique, and ask for retransmission otherwise.
a) If 02 is sent, what is the chance that it is received as a word in the code?
b) If 01 is sent, what is the chance that we ask for retransmission?
(Hint for part b): first find which received words do not have a unique nearest neighbour.)

12 Consider the codes $C_1 = \{0, 1, 2\}$ and $C_2 = \{000, 111, 222\} \subseteq \{0, 1, 2\}^3$, which are sent over a ternary symmetric channel with symbol-error probability p.
a) Find the minimum distances for C_1 and C_2. How many errors can C_1 and C_2 detect or correct?
b) For C_1 the codeword 0 is sent. What is the chance that the received word is decoded correctly under nearest neighbour decoding?
c) For C_2 the codeword 000 is sent. Determine the chance that the received word is decoded correctly if we do incomplete nearest neighbour decoding, where we only decode a received word x if $d(x, c) \leq 1$ for some $c \in C_2$ and do not do anything otherwise. What is the chance that we do not decode the received word at all?
d) Again, for C_2 the codeword 000 is sent, but now we accept only codewords as received words, and ask for retransmission otherwise. What is the chance that we receive a codeword the first time? What is the chance that we eventually decode the received word correctly, perhaps after several transmissions?
e) Now take $p = 0.1$ and compare the chance of failure for parts b), c) and d), where failure means that we decode either incorrectly or not at all, even after several transmissions.

13 Let the code $C = \{00000, 11111, 22222, 33333\} \subseteq \{0, 1, 2, 3\}^5$ be transmitted over a symmetric 4-ary channel with symbol-error probability $p = 0.1$. We assume that each codeword is equally likely to be sent.
a) Find the nearest neighbours of $w_0 = 00123$ and $w_1 = 00111$.
b) If $c_0 = 00000$ and $c_1 = 11111$, find $\mathbb{P}(w_j \text{ received} \mid c_i \text{ sent})$ for $0 \leq i, j \leq 1$.
c) Find $\mathbb{P}(w_j \text{ received})$ for $j = 0, 1$.
d) Find $\mathbb{P}(c_i \text{ sent} \mid w_j \text{ received})$ for $0 \leq i, j \leq 1$.
e) Comment on the following statement: "If 00000 is sent, we are as likely to receive 00111 as 00123. So if we decode 00123 to 00000, we should also decode 00111 to 00000."
f) Do $\mathbb{P}(00000 \text{ sent} \mid 00111 \text{ received})$ and $\mathbb{P}(11111 \text{ sent} \mid 00111 \text{ received})$ add up to 1? Should they?

14 Consider words of length 3 made using the alphabet $A = \{0, 1, \ldots, q-1\}$ where $q \geq 3$. Describe $S(000, r)$ for $r = 0, 1$ and 2. How many elements are there in each? Do those sets look like spheres if we identify the elements in A with $0, 1, \ldots, q-1$ in $\mathbb{R}$, and view all words in $\mathbb{R}^3$?

15 Let C be a ternary (4, 9, 3)-code. Show that C is perfect.

16 Let C be an $(n, M, 2t)$ code with $M > 1$. (In other words, $d(C)$ is even).
a) Given code words x and y such that $d(x, y) = 2t$, find a word z not in the code such that $d(x, z) = d(y, z) = t$.
b) Can z be in some $S(u, r)$ with u in the code and $r < t$?
c) Conclude that C cannot be a perfect code.