

1. This question is about the code and transmission channel described in the lecture notes on page 4 (the codewords are 000 and 111). Assuming that the codewords 000 and 111 are equally likely to be transmitted, prove that the probability we will decode correctly is indeed 97.2%.

Solution : We will decode correctly if and only if we have one of the following eight cases:

- 000 is sent and $y_1y_2y_3 \in \{000, 100, 010, 001\}$ is received.
- 111 is sent and $y_1y_2y_3 \in \{111, 110, 101, 011\}$ is received.

So we are asked to calculate probabilities of the form $P(A \cap B)$ where A stands for 000 (or 111) being sent and B stands for a certain $y_1y_2y_3$ being received.

Recall that

$$P(A \cap B) = P(A)P(B|A)$$

Since the codewords 000 and 111 are equally likely to be transmitted, we have that

$$P(A) = P(000 \text{ sent}) = P(111 \text{ sent}) = \frac{1}{2}$$

The probabilities $P(B|A)$ are listed in the table on page 4 in the lecture notes. We get

$$\begin{aligned} P(000 \text{ sent and } 000 \text{ received}) &= \frac{1}{2}(1-p)^3 \\ P(000 \text{ sent and } 100 \text{ received}) &= \frac{1}{2}p(1-p)^2 \\ P(000 \text{ sent and } 010 \text{ received}) &= \frac{1}{2}p(1-p)^2 \\ P(000 \text{ sent and } 001 \text{ received}) &= \frac{1}{2}p(1-p)^2 \\ P(111 \text{ sent and } 111 \text{ received}) &= \frac{1}{2}(1-p)^3 \\ P(111 \text{ sent and } 110 \text{ received}) &= \frac{1}{2}p(1-p)^2 \\ P(111 \text{ sent and } 101 \text{ received}) &= \frac{1}{2}p(1-p)^2 \\ P(111 \text{ sent and } 011 \text{ received}) &= \frac{1}{2}p(1-p)^2 \end{aligned}$$

Adding up all these probabilities, we get that the probability that we decode correctly is

$$(1-p)^3 + 3p(1-p)^2 = (1-p)^2(1+2p)$$

Substituting $p = 10\%$ into this expression, we find 97.2%. □

2. Let C be a binary code of length n . Suppose we have a symmetric channel with probability p of an error during transmission of a digit with $0 \leq p \leq 0.5$ (see page 3). Prove that Nearest Neighbor Decoding is indeed Maximum Likelihood Decoding.

Solution : Suppose we receive a word \mathbf{y} . In Nearest Neighbor Decoding, we look for the closest codeword \mathbf{c} to \mathbf{y} . The probability P that \mathbf{y} is received if \mathbf{c} is sent depends on $d_H(\mathbf{y}, \mathbf{c})$. If $k = d_H(\mathbf{y}, \mathbf{c})$ then

$$P = p^k(1-p)^{n-k}$$

as k digits are transmitted incorrectly and $n-k$ digits are transmitted correctly. If \mathbf{c}' is another codeword with $d_H(\mathbf{y}, \mathbf{c}') = l \geq k$ then the probability P' that \mathbf{y} is received if \mathbf{c}' is sent is

$$P' = p^l(1-p)^{n-l}$$

If we can show that $P \geq P'$ then it is more likely that \mathbf{c} was sent than \mathbf{c}' and we will have Maximum Likelihood Decoding.

If $p = 0$ then $P = P' = 0$. So we may assume that $0 < p < 0.5$. Dividing both sides by $p^k(1-p)^{n-l}$, we get

$$\begin{aligned} P \geq P' &\iff p^k(1-p)^{n-k} \geq p^l(1-p)^{n-l} \\ &\iff (1-p)^{n-k-(n-l)} \geq p^{l-k} \\ &\iff (1-p)^{l-k} \geq p^{l-k} \\ &\iff \left(\frac{1-p}{p}\right)^{l-k} \geq 1 \\ &\iff \frac{1-p}{p} \geq 1 && \text{since } l-k \geq 0 \\ &\iff 1-p \geq p \\ &\iff 1 \geq 2p \\ &\iff p \leq 0.5 \end{aligned}$$

3. Prove that there does not exist a binary one-error correcting code of length 6 with 9 codewords.

Proof : Suppose that C is a binary one-error correcting code of length 6 with 9 codewords. Then there are at least five codewords with the same digit in the first position. At least three of these five codewords, say $\mathbf{x}, \mathbf{y}, \mathbf{z}$, have the same digit in the second position.

Since C is one-error correcting, we know that $d(C) \geq 3$. So \mathbf{x} and \mathbf{y} (resp. \mathbf{x} and \mathbf{z} , resp. \mathbf{y} and \mathbf{z}) are different in at least three positions. Put

$$A = \{3 \leq i \leq 6 : x_i = y_i\}, B = \{3 \leq i \leq 6 : x_i = z_i\} \text{ and } C = \{3 \leq i \leq 6 : y_i = z_i\}$$

Then $A \cup B \cup C = \{3, 4, 5, 6\}$. Indeed, let $3 \leq i \leq 6$. If $x_i = y_i$ then $i \in A$. If $x_i \neq y_i$ then $\{x_i, y_i\} = \{0, 1\}$ since the code is binary and so $z_i \in \{x_i, y_i\}$ and thus $i \in B$ or $i \in C$. Hence

$$4 = |\{3, 4, 5, 6\}| = |A \cup B \cup C| \leq |A| + |B| + |C|$$

So $|A| \geq 2$ or $|B| \geq 2$ or $|C| \geq 2$. But then $d_H(\mathbf{x}, \mathbf{y}) \leq 2$ or $d_H(\mathbf{x}, \mathbf{z}) \leq 2$ or $d_H(\mathbf{y}, \mathbf{z}) \leq 2$, a contradiction.

Hence there does not exist a binary one-error correcting code of length 6 with 9 codewords. \square

4. Let n be an integer with $n \geq 2$. Let C be the binary code of length n consisting of all words containing an even number of ones. So if $n = 3$ then $C = \{000, 110, 101, 011\}$.

Find $|C|$ and $d(C)$.

Solution : Note that

$$\underbrace{000 \dots 000}_{n \text{ times}} \quad \text{and} \quad 11 \underbrace{00 \dots 00}_{n-2 \text{ times}}$$

are two distinct codewords. Since the distance between these two codewords is 2, we get that $d(C) \leq 2$.

Suppose that $d(C) = 1$. Then there exist distinct codewords $\mathbf{x}, \mathbf{y} \in C$ with $d_H(\mathbf{x}, \mathbf{y}) = 1$. So \mathbf{x} and \mathbf{y} are distinct in exactly one position, say the first position. Then $x_1 \neq y_1$ and $x_i = y_i$ for all $2 \leq i \leq n$. Since the code is binary, we have that either $(x_1, y_1) = (1, 0)$ or $(x_1, y_1) = (0, 1)$, say $(x_1, y_1) = (1, 0)$. Since \mathbf{x} is a codeword and $x_1 = 1$, we have that (x_2, \dots, x_n) contains an odd number of ones. Similarly, since \mathbf{y} is a codeword and $y_1 = 0$, we have that (y_2, \dots, y_n) contains an even number of ones. But that is a contradiction since $(x_2, \dots, x_n) = (y_2, \dots, y_n)$.

So $d(C) > 1$. Hence $d(C) = 2$.

Next, we count how many codewords there are. Consider the map

$$\theta : C \rightarrow \{0, 1\}^{n-1} : (x_1, x_2, \dots, x_{n-1}, x_n) \rightarrow (x_1, x_2, \dots, x_{n-1})$$

A codeword is of the form

$$x_1 x_2 \dots x_{n-1} x_n$$

where $x_i \in \{0, 1\}$ for all $1 \leq i \leq n$ and (x_1, \dots, x_n) contains an even number of ones. Given $x_1, \dots, x_{n-1} \in \{0, 1\}$ there exists a unique $x_n \in \{0, 1\}$ such that $x_1 \dots x_n \in C$: if (x_1, \dots, x_{n-1}) contains an odd number of ones then $x_n = 1$; if (x_1, \dots, x_{n-1}) contains an even number of ones then $x_n = 0$. This shows that θ is onto and one-to-one. Hence θ is a bijection. Thus

$$|C| = |\{0, 1\}^{n-1}| = 2^{n-1} \quad \square$$

5. Consider the ternary code $\{000, 111, 222\}$. Suppose we have a channel and a constant p such that for all distinct $a, b \in \{0, 1, 2\}$, we have that the probability $P(a|b)$ that the symbol a is received if the symbol b is transmitted, is equal to p (e.g., if $p = 5\%$ and we transmit the symbol 0, then there is a 5% chance we receive the symbol 1, a 5% chance we receive the symbol 2 and a 90% chance we receive the symbol 0).

We use the following decoding algorithm when receiving a word $y_1 y_2 y_3$:

- (a) If $y_1 y_2 y_3$ contains three different symbols, we declare a decoding error.
- (b) If $y_1 y_2 y_3$ contains at most two different symbols, we decode as aaa where a is the symbol that shows up at least twice in $y_1 y_2 y_3$.

So we declare a decoding error if we receive 120 but decode as 111 if we receive 121.

We transmit the codeword 000 and use the decoding algorithm to decode the received word.

- (a) Calculate the probability that we declare a decoding error. Evaluate this if $p = 5\%$.
- (b) Calculate the probability that we decode correctly as 000. Evaluate this if $p = 5\%$.

Solution : If $y_1y_2y_3$ is a ternary word then we denote by $P(y_1y_2y_3|000)$ the probability that we receive the word $y_1y_2y_3$ if we transmit the codeword 000. Note that the probability that we receive the symbol we transmit is $1 - 2p$.

(a) We will declare a decoding error if we receive any of the following words :

$$\{012, 021, 102, 201, 120, 210\}$$

If $y_1y_2y_3 \in \{012, 021, 102, 201, 120, 210\}$ then $P(y_1y_2y_3|000) = (1 - 2p)p^2$. So the probability that we declare a decoding error is

$$6(1 - 2p)p^2$$

For $p = 5\%$ this becomes

$$1.35\%$$

(b) We will decode correctly if we receive any of the following words :

$$\{000, 001, 002, 010, 020, 100, 200\}$$

We easily get that $P(000|000) = (1 - 2p)^3$ and $P(y_1y_2y_3|000) = (1 - 2p)^2p$ if $y_1y_2y_3 \in \{001, 002, 010, 020, 100, 200\}$. So the probability that we decode correctly is

$$6(1 - 2p)^2p + (1 - 2p)^3$$

For $p = 5\%$, this becomes

$$97.2\%$$