

3.2 Decoder and Symbol Error Probability

3.19. Knowing the characteristics of the channel and its input, on the receiver side, we can use this information to build a “good” receiver.

We now consider a part of the receiver called the **(channel) decoder**. Its job is to guess the value of the channel input¹¹ X from the value of the received channel output Y . We denote this guessed value by \hat{X} .

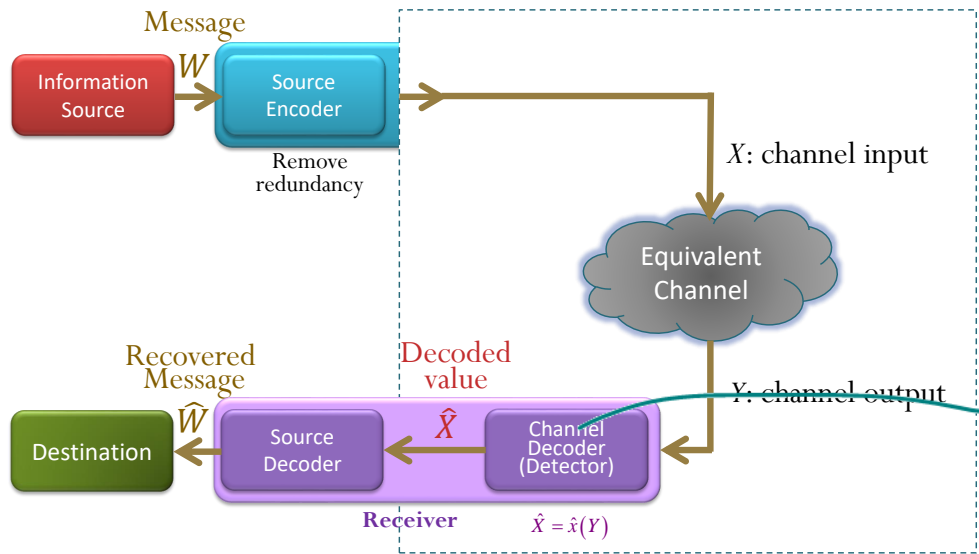


Figure 11: Adding a (channel) decoder to improve the performance of the system in Figure 4 considered in Section 3.1.

3.20. A “good” decoder is the one that (often) guesses correctly. So, our goal here is to

- maximize the probability of correct guessing
- minimize the probability of wrong guessing

Quantitatively, to measure the performance of a decoder, we define a quantity called the (symbol) error probability.

¹¹To simplify the analysis, we still haven’t considered the channel encoder. (It may be there but is included in the equivalent channel or it may not be in the system at all.)

Definition 3.21. The **(symbol) error probability**, denoted by $P(\mathcal{E})$, can be calculated from

$$P(\mathcal{E}) = P \left[\hat{X} \neq X \right].$$

- \mathcal{E} is the (decoding) error event.
- A dual (complementary) quantity is the probability of correct decoding:

$$P(\mathcal{C}) = P \left[\hat{X} = X \right] = 1 - P(\mathcal{E}).$$

- For channel with binary input, the error probability is the same as bit error rate (BER).

3.22. A “reasonable” decoder should make a guess based on all the information it has obtained. Here, the only information it can observe is the value of Y . Therefore, a decoder is a function of Y , say, $g(Y)$. Therefore, $\hat{X} = g(Y)$.

We use $\hat{x}(\cdot)$ to denote this function $g(\cdot)$; So, $\hat{X} = \hat{x}(Y)$. It is a deterministic function operating on the random channel output Y ; the randomness in the decoded value \hat{X} comes from the randomness in Y .

Definition 3.23. A “naive” decoder is a decoder that simply sets $\hat{X} = Y$. Equivalently,

$$\hat{x}_{\text{naive}}(y) = y.$$

- We may call this a “pass-through” decoder or a “do-nothing” decoder.
- It usually has poor performance. However, we shall see that, in some cases, this simple decoder can be optimal¹².

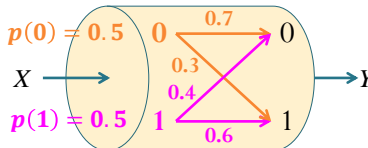
3.24. For general DMC, the error probability of the naive decoder is

$$\begin{aligned} P(\mathcal{E}) &= P \left[\hat{X} \neq X \right] = P \left[Y \neq X \right] = 1 - P \left[Y = X \right] \\ &= 1 - \sum_x P \left[Y = x, X = x \right] = 1 - \sum_x P \left[Y = x | X = x \right] P \left[X = x \right] \\ &= 1 - \sum_x Q(x|x) p(x) \end{aligned}$$

¹²Nothing has better error probability in that particular situation.

Example 3.25. Consider the BAC channel and input probabilities specified in Example 3.16. Find $P(\mathcal{E})$ when $\hat{X} = Y$.

(a) Method 1:



naive decoder

$$P(\mathcal{E}) = P[\hat{X} \neq X] = P[Y \neq X] = 0.15 + 0.20 = 0.35$$

$$Q = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \xrightarrow{\begin{matrix} \times 0.5 \\ \times 0.5 \end{matrix}} \begin{bmatrix} 0.35 & 0.15 \\ 0.20 & 0.30 \end{bmatrix}$$

$$P(\mathcal{C}) = P[\hat{X} = X] = P[Y = X] = 0.35 + 0.30 = 0.65$$

$$P(\mathcal{E}) = 1 - P(\mathcal{C}) = 1 - 0.65 = 0.35$$

(b) Method 2: To calculate the error probability directly, we write

$$\begin{aligned} P(\mathcal{E}) &= P[\hat{X} \neq X] = P[Y \neq X] \\ &= P[Y = 0, X = 1] + P[Y = 1, X = 0] \\ &= P[Y = 0 | X = 1] P[X = 1] + P[Y = 1 | X = 0] P[X = 0] \\ &= 0.4 \times \frac{1}{2} + 0.3 \times \frac{1}{2} = 0.35. \end{aligned}$$

(c) Method 3: Using the derived formula from 3.24,

$$\begin{aligned} P(\mathcal{E}) &= 1 - (Q(0|0)p(0) + Q(1|1)p(1)) \\ &= 1 - \left(0.7 \times \frac{1}{2} + 0.6 \times \frac{1}{2}\right) = 0.35. \end{aligned}$$

Example 3.26. Continue from Examples 3.12 and 3.18. Find the error probability $P(\mathcal{E})$ when a naive decoder is used with a DMC channel in which $\mathcal{X} = \{0, 1\}$, $\mathcal{Y} = \{1, 2, 3\}$, $Q = \begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.4 & 0.3 \end{bmatrix}$ and $\underline{p} = [0.2, 0.8]$.

$$Q = \begin{matrix} & \begin{matrix} x \backslash y & 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 0 \\ 1 \end{matrix} & \begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.4 & 0.3 \end{bmatrix} \end{matrix} \xrightarrow{\begin{matrix} \times 0.2 \\ \times 0.8 \end{matrix}} \begin{matrix} & \begin{matrix} 1 & 2 & 3 & y \backslash x \end{matrix} \\ \begin{matrix} 0 \\ 1 \end{matrix} & \begin{bmatrix} 0.10 & 0.04 & 0.06 \\ 0.24 & 0.32 & 0.24 \end{bmatrix} \end{matrix} = P$$

$$P(\mathcal{E}) = P[\hat{X} \neq X] = P[Y \neq X] = \sum (\quad) = \dots$$

$$P(\mathcal{C}) = P[\hat{X} = X] = P[Y = X] = 0.24$$

$$P(\mathcal{E}) = 1 - \quad = 0.76$$

Example 3.27. DIY Decoder: Consider a different decoder specified in the **decoding table** below. Find the error probability $P(\mathcal{E})$ when such decoder is used in Example 3.26.

y	$\hat{x}(y)$
1	0
2	1
3	0

$$\begin{aligned}
 P(\mathcal{C}) &= P[\hat{X} = X] \\
 &= 0.10 \\
 &\quad + 0.32 \\
 &\quad + 0.06 \\
 &= 0.48
 \end{aligned}$$

$$\begin{matrix}
 \hat{x} & 0 & 1 & 0 \\
 \downarrow & & & \\
 P = & \begin{matrix} 1 \\ 0 \\ 0 \end{matrix} \begin{bmatrix} 0.10 & 0.04 & 0.06 \\ 0.24 & 0.32 & 0.24 \end{bmatrix}
 \end{matrix}$$

$$P(\mathcal{E}) = 1 - P(\mathcal{C}) = 0.52$$

Example 3.28. Repeat Example 3.27 but use the following decoder

y	$\hat{x}(y)$
1	1
2	1
3	0

$$\begin{matrix}
 \hat{x} & 1 & 1 & 0 \\
 \downarrow & & & \\
 P = & \begin{matrix} 0 \\ 1 \\ 1 \end{matrix} \begin{bmatrix} 0.10 & 0.04 & 0.06 \\ 0.24 & 0.32 & 0.24 \end{bmatrix}
 \end{matrix}$$

$$P(\mathcal{C}) = 0.24 + 0.32 + 0.06 = 0.62$$

$$P(\mathcal{E}) = 1 - P(\mathcal{C}) = 1 - 0.62 = 0.38$$

Observation: For each column of the \mathbf{P} matrix, we circle the probability corresponding to the row of x that has the same value as $\hat{x}(y)$.

3.29. A recipe for finding $P(\mathcal{E})$ of any (DIY) decoder:

- Find the \mathbf{P} matrix by scaling each row of the \mathbf{Q} matrix by its corresponding $p(x)$.
- Write $\hat{x}(y)$ values on top of the y values for the \mathbf{P} matrix.
- For column y in the \mathbf{P} matrix, circle the element whose corresponding x value is the same as $\hat{x}(y)$.
- $P(\mathcal{C})$ = the sum of the circled probabilities. $P(\mathcal{E}) = 1 - P(\mathcal{C})$.