Channel Capacity Calculations

Karan Jayachandra

Personal Archive

October 28, 2022

1 Introduction

This report details the calculations of the channel capacity of a Binary Asymmetric Channel where both transition probabilities are independent of each other. A Binary Asymmetric Channel is a channel where a transmitted bit can turn into the other.

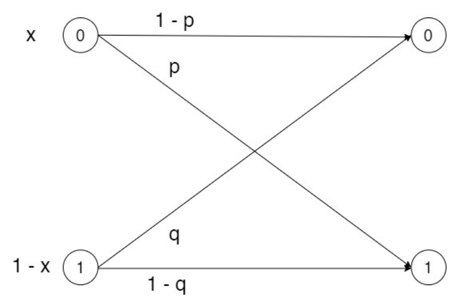

Figure 1: Channel Definition

In this report, it is assumed that the transition probability of a bit from 0 to 1 is p and from 1 to 0 is q. Also, the probability of the inputs bits themselves are x and $(1-x)$. We find the optimum input bit distribution for any channel characteristic i.e., p and q and thereby find the maximum allowable capacity for an arbitrarily slow error that asymptotically goes to 0.

In section 2, the capacity values are calculated theoretically. Section 3 shows the MATLAB implementation to verify the derivations and section 4 documents the results. Finally, section 5 looks the interpretation of the results and provides a subjective opinion on the performance of the channels.

From this point on, the Binary Asymmetric Channel where only one transition is possible is referred to as Z Channel and when two transitions are possible, the channel is referred to as Binary Asymmetric Channel or BSC. The Binary Symmetric Channel is also referred to as BAC.

2 Derivation

As we can see from figure 1, the transition probabilities and the input probabilities can definitively provide the distribution of the output symbols. Consider that the input symbol is X and the output symbol is Y. From simple probabilistic identities:

$$P(Y=0) = (1-p)x + (1-x)q \tag{1}$$

$$P(Y=1) = px + (1-x)(1-q) \tag{2}$$

From this, we can define the entropy in Y as follows:

$$
\begin{aligned}
H(Y) = -\left\{(1-p)x + (1-x)q\right\} \\
\log_2\left((1-p)x + (1-x)q\right) \\
-\left\{px + (1-x)(1-q)\right\} \\
\log_2\left(px + (1-x)(1-q)\right)
\end{aligned} \tag{3}
$$

The Mutual Information between Y and X define the information about Y that can be gleaned from X and vice versa. This is defined as:

$$I(X;Y) = H(Y) - H(Y \mid X) \tag{4}$$

To find the Conditional Entropy required to calculate the Mutual Information, we can sum over all the values that X and the Entropy of Y in each case.

$$H(Y \mid X) = \sum p(x)\ H(Y \mid X = x) \tag{5}$$

Substituting for the probabilities and the the Entropy of Y, we get:

$$H(Y \mid X) = xH(p) + (1-x)H(q) \tag{6}$$

Using equations 3 and 6 in 4, we find that the Mutual Information can be given by:

$$
\begin{aligned}
I(X;Y) = -\left\{(1-p)x + (1-x)q\right\} \\
\log_2\left((1-p)x + (1-x)q\right) \\
-\left\{px + (1-x)(1-q)\right\} \\
\log_2\left(px + (1-x)(1-q)\right) \\
- xH(p) - (1-x)H(q)
\end{aligned} \tag{7}
$$

The Channel Capacity is then described as the maximum of this over all distributions of X:

$$C = \max_x I(X;Y) \tag{8}$$

Setting the derivative to 0 of equation 8 provides the optimum value of the input distribution x as:

$$x_{opt} = \frac{1-(1+\alpha)q}{(1+\alpha)(1-p-q)} \tag{9}$$

where α is defined as:

$$\alpha = 2^{\frac{H(p)-H(q)}{1-p-q}} \tag{10}$$

This is the general form of the solution for any p and q. Substituting for x from equation 9 to 8. We can find the maximum capacity of the channel.

Based on the problem statement, $q = 0.26$. Substituting this in the equation, we find that the optimum value of x is:

$$x_{opt} = \frac{0.74 + 0.26\alpha}{(1+\alpha)(0.74 - p)} \tag{11}$$

where α is:

$$\alpha = \frac{H(p) - 0.25}{0.74 - p} \tag{12}$$

This gives the Capacity of the channel as:

$$\begin{aligned} C = -\left[\frac{1+0.52\alpha}{1+\alpha}\right]\log_2\left(\frac{1+0.52\alpha}{1+\alpha}\right) \\ -\left[\frac{1.48+\alpha}{1+\alpha}\right]\log_2\left(\frac{1.48+\alpha}{1+\alpha}\right) \\ +\frac{-0.12\alpha + 0.25p + 0.25p\alpha - 0.74H(p) + 0.26\alpha H(p)}{(1+\alpha)(0.74-p)} \end{aligned} \tag{13}$$

The above equation provides the Channel Capacity for any value of p.

3 Implementation

The above is implemented in MATLAB and the complete code is given in appendix **??**. The code follows the same equations as the given above. We look at three snippets which form the basis of this analysis. In the first case, we have the function that provides the optimum value of x as described in equation 9.

```
%% Finds the Optimal Binary Distribution
function X_opt = OptimalInput(p, q)
    H_p = (-p*log(p)-(1-p) ...
               * log(1-p))/log(2);
    H_q = (-q*log(q)-(1-q) ...
```

```
               * log(1-q))/log(2);
     alpha = 2^((H_p-H_q)/(1-p-q));
     X_opt = (1 - ((1+alpha)*q))...
               /((1+alpha)*(1-p-q));
end
```

As seen, the function takes the transition probabilities as it's input and provides the optimum input distribution to maximize the capacity of the channel. The second snippet, details the calculation of the Mutual Information provided the values of the probabilities of transition and the input distribution.

```
%% Finds the Mutual Information
function MI = MutualInformation(p, q, x)
     H_p = (-p*log(p)-(1-p)*log(1-p))/log(2);
     H_q = (-q*log(q)-(1-q)*log(1-q))/log(2);
     term_1 = -(((1-p)*x)+((1-x)*q))...
               *log(((1-p)*x)+((1-x)*q))/...
                                        log(2);
     term_2 = -((p*x)+((1-x)*(1-q)))...
               *log((p*x)+((1-x)*(1-q)))/...
                                        log(2);
     term_3 = - x * H_p;
     term_4 = - (1-x) * H_q;
     MI = term_1 + term_2 + term_3 + term_4;
end
```

The third shows the calculation of the capacity of known channel i.e., the Binary Symmetric Channel and the Z Channel.

```
%% Returns the Capacity of BSC
function Capacity = BSCCapacity(p)
     H_p = (-p*log(p)-(1-p)*log(1-p))...
                                        /log(2);
     Capacity = 1 - H_p;
end

%% Returns the Capacity of BAC
function Capacity = BACCapacity(p)
```

```
    Capacity = log(1+((1-p)*...
                      p^(p/(1-p))))/log(2);
end
```

The complete functionality of the analysis is described by the above three snippets.

4 Results

The results of sweeping over all values of p and q are given in the surface plot as shown in Figure 2. As expected, there are peaks when both the values are either 0 or 1. This is because of the the capacity is maximum where there is either no error or the channel acts as a logical NOT. This is also an ideal case, as a simple negation would provide no error. When the probabilities are 0.5, the channel behaves randomly and hence cannot be used as a good transmission medium. Hence there is a dip in the values along the diagonal. When the transition probabilities are 0 and 1 or vice-versa, all the bits received are the same. This also makes the channel impractical for transmission.

The sudden peaks seem to be due to MATLAB calculation practicalities. Figure 3 makes it easier to look at the capacity values over p and q and also highlights the specific value of q that is required for analysis.

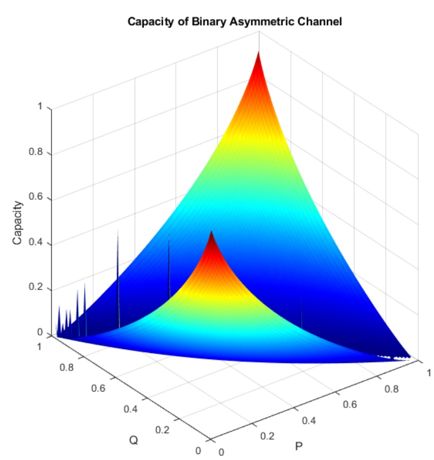

Figure 2: Sweeping Across P and Q

As seen, the channel capacity varies along p and has peaks at either transition probabilities as 0 and 1 which is the expected behavior as the transition probabilities are independent of

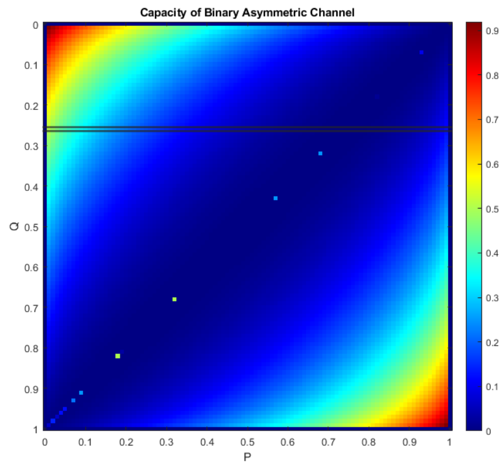

Figure 3: Sweeping Across P and Q

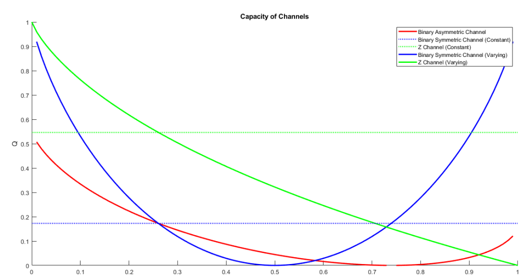

Figure 4: Comparing with Channels

each other. Once the transition probability of one of the input symbols is fixed. The capacity is maximized when the transition probability of the second symbol is either 0 or 1.

Figure 4 shows the comparison of the Capacity of a Binary Asymmetric Channel with independent transition probabilities with that of Binary Symmetric Channel and Z Channel. The dotted lines show the Capacity of the compared channels when the transition probabilities are fixed to 0.26 as required and the continuous lines show the Capacities when the probabilities are varying.

Consider the varying BSC, in general the Capacity of the channel is better compared to the channel under observation but when the transition probability is 0.5, the BSC has 0 Capacity while the BAC has a non-zero capacity. This is due to the skew in the probabilities of the the second transition which lets the user transmit with some level of confidence. The BSC with fixed probability performs well when compared to the BAC at higher transition probabilities because the value of q is very low, which makes the appearance of 1 at the output much more likely. This decreases the overall capacity.

The varying Z Channel performs well at lower levels as lower transition probabilities makes the channel closer to an ideal one. But as the transition probability gets close to 1, The BAC performs better and the Z Channel only transmits 0. The Z Channel with constant transition probability performs much better as 0.26 is comparatively a small transition probability and drops the capacity to around three-fourth the maximum value.

5 Discussion

The examination of the BAC shows that the channel describes the most general case of all other channels. When we look at the diagonal of figure 3 we observe the characteristics of the BSC. When we look at any of the side of the graph, we observe the Z Channel. All other channels are a subset of the Binary Asymmetric Channel. Due to it's asymmetry the BAC can be used to transmit albeit at lower levels of confidence at most values of p and q with the exception being the case when both of them have a probability of 0.5. This makes the channel completely random and the capacity drops to zero. In all other cases the BAC has some capacity and can transfer information at varying rates. The Channel Capacity is an accurate measure of the ability of a channel to relay information accurately. From section 2, the capacity has been derived for the most general of cases which can be modified and re-purposed for many other channels. The capacity of the BAC has two peaks when the transition probabilities are either 0 or 1 and it's capacity goes close to zero when either of them is around 0.5. A high entropy for the source provides a lot of information, but in the case of channels, the more deterministic channels (not when both cases provide the same output) have higher capacities. This is because they have higher reliability in transferring what was sent at the input and hence the output has almost full Mutual Information.