ECE 771
Lecture 8 – More on Channel Capacity

Objective: To prove the converse to the channel coding theorem and a few other topics.

1 The converse to the coding theorem

We will begin with a special case: a zero-error channel, and show that this requires that the rate be less than the capacity. A zero error channel is one in which the sequence Y^n determines the input symbol W without any error, so $H(W|Y^n) = 0$. We can obtain a bound by assuming that W is uniformly distributed over the 2^{nR} input symbols, so $H(W) = nR$. Now

$$\begin{aligned} nR = H(W) &= H(W|Y^n) + I(W;Y^n) \\ &= I(W;Y^n) \\ &\leq I(X^n;Y^n) \qquad \text{data processing inequality} \\ &\leq \sum_{i=1}^{n} I(X_i;Y_i) \qquad \text{to be proved} \\ &\leq nC \qquad \text{definition of information channel capacity} \end{aligned}$$

from which we conclude that for a zero-error channel, $R \leq C$.

Recall that Fano's inequality related the probability of error to the entropy. Using the notation of the current context, it can be written as

$$H(X^n|Y^n) \leq 1 + P_e^{(n)} nR$$

We also observe (and could prove) that

$$I(X^n;Y^n) \leq nC$$

(Using the channel n times, the capacity per transmission is not increased.)

We now have the tools necessary to prove the converse to the coding theorem: any sequence of $(2^{nR}, n)$ codes with $\lambda^{(n)} \to 0$ must have $R \leq C$.

Proof Observe that $\lambda^{(n)} \to 0$ implies $P_e^{(n)} \to 0$. Let the symbols be drawn uniformly. Then

$$\begin{aligned} nR = H(W) &= H(W|Y^n) + I(W;Y^n) \\ &\leq H(W|Y^n) + I(X^n(W);Y^n) \qquad \text{data processing inequality} \\ &\leq 1 + P_e^{(n)} nR + I(X^n(W);Y^n) \qquad \text{Fano} \\ &\leq 1 + P_e^{(n)} nR + nC \qquad \text{the last observation} \end{aligned}$$

Rewriting,

$$P_e^{(n)} \geq 1 - \frac{C}{R} - \frac{1}{nR}.$$

If $R > C$, then for n sufficiently large, $P_e^{(n)}$ is bounded away from 0. $\square$

Hamming codes?

Feedback channels: the same capacity!

2 Joint source/channel coding theorem

We have seen that for a source with entropy $H(X)$, the data rate cannot be less than the entropy ($R > H$. We have also seen that we can transmit reliably at rates less than capacity ($R < C$). How do these two major theorems tie together?

That is, is it better to remove the redundancy (source coding), then put some back in (channel coding)? Or is there some kind of joint coding method that would work better?

The joint source/channel coding theorem says (in essence) that provided that a source has entropy $H < C$, then there is a code with $P_e^{(n)} \to 0$, and that (conversely) if $P_e^{(n)} \to 0$ then $H < C$. Note that the theorem is *asymptotic*. The proof of the theorem relies on AEP: we code only the typical sequences, and don't worry about the rest (forward). For the converse, we use (again) Fano.

Note that theorem is asymptotic: in practice, we have to deal with codes of finite length and take extra precautions.