18.413: Error-Correcting Codes Lab February 10, 2004

Lecture 3

Lecturer: Daniel A. Spielman

3.1 Analysis of repetition code meta-channel

When we specialize our interpretation of the output of a channel to the meta channel formed by encoding using the repitition code and transmitting over another channel, we solve a fundamental problem of probability: *how to combine the results of independent experiments.*

That is, let w be a random varible taking values in $\{0,1\}$. Imagine encoding w using the repeat-2-times code to $(x_1, x_2) = (w, w)$, and transmitting x_1 and x_2 over a memoryless channel (so each transmission is independent). Equivalently, we could assume that x_1 is transmitted over one channel and x_2 is transmitted over another. Let y_1 and y_2 be the random variables corresponding to the outputs of the channel, let b_1 and b_2 be the values actually received, and let

$$
\begin{aligned}
p_1 &= \mathrm{P}\left[x_1 = 1 | y_1 = b_1\right], \text{ and} \\
p_2 &= \mathrm{P}\left[x_2 = 1 | y_2 = b_2\right].
\end{aligned}
$$

We would like to know the probability that $w = 1$ given both y_1 and y_2. As before, we will assume that w was uniformly distributed (half chance 0 and half chance 1). I think of each channel transmission as an experiment, and I now want to determine the probability that w was 1 given the results of both experiments.

By the theorem from last class, we have

$$
\mathrm{P}\left[w = 1 | y_1 = b_1 \text{ and } y_2 = b_2\right] = \frac{\mathrm{P}\left[y_1 = b_1 \text{ and } y_2 = b_2 | w = 1\right]}{\mathrm{P}\left[y_1 = b_1 \text{ and } y_2 = b_2 | w = 1\right] + \mathrm{P}\left[y_1 = b_1 \text{ and } y_2 = b_2 | w = 0\right]} \tag{3.1}
$$

To evaluate this probability, we first note that

$$
\mathrm{P}\left[y_1 = b_1 | w = 1\right] = \mathrm{P}\left[w = 1 | y_1 = b_1\right] \mathrm{P}\left[y_1 = b_1\right] / \mathrm{P}\left[w = 1\right] = p_1 \mathrm{P}\left[y_1 = b_1\right] / \mathrm{P}\left[w = 1\right].
$$

While we do not necessarily know $\mathrm{P}\left[y_1 = b_1\right]$, it will turn out not to matter.

Since the two channel outputs are independent given w, we have

$$
\begin{aligned}
\mathrm{P}\left[y_1 = b_1 \text{ and } y_2 = b_2 | w = 1\right] &= \mathrm{P}\left[y_1 = b_1 | w = 1\right] \mathrm{P}\left[y_2 = b_2 | w = 1\right] \\
&= \frac{p_1 \mathrm{P}\left[y_1 = b_1\right] p_2 \mathrm{P}\left[y_2 = b_2\right]}{\mathrm{P}\left[w = 1\right] \mathrm{P}\left[w = 1\right]}.
\end{aligned}
$$

Applying $\mathrm{P}[w = 0|y_1 = b_1] = 1 - \mathrm{P}[w = 1|y_1 = b_1]$, we can also comput

$$\mathrm{P}[y_1 = b_1 \text{ and } y_2 = b_2|w = 0] \, \frac{(1-p_1)\mathrm{P}[y_1 = b_1](1-p_2)\mathrm{P}[y_2 = b_2]}{\mathrm{P}[w = 0]^2}.$$

Combining these equations, and $\mathrm{P}[w = 0] = \mathrm{P}[w = 1] = 1/2$, we obtain

$$(3.1) = \frac{p_1 p_2}{p_1 p_2 + (1-p_1)(1-p_2)}.$$

In particular, the terms we don't know cancel!

3.2 Capacity of meta-channel

Consider the meta-channel obtained by encoding a bit w via the repeat-2-times code to obtain (x_1, x_2), and then passing these bits through the BSC_p to obtain (y_1, y_2). We will now compute the capacity of this meta-channel. We begin with the computation of the quantities that appear in the formula for $I(w; (y_1, y_2))$:

$$\begin{aligned}
\mathrm{P}[w = 1|(y_1, y_2) = (1,1)] &= \frac{(1-p)^2}{p^2 + (1-p)^2} \\
\mathrm{P}[w = 1|(y_1, y_2) = (1,0)] &= \frac{p(1-p)}{p(1-p) + (1-p)p} = 1/2 \\
\mathrm{P}[w = 1|(y_1, y_2) = (0,0)] &= \frac{p^2}{p^2 + (1-p)^2)} \\
\mathrm{P}[w = 0|(y_1, y_2) = (1,1)] &= \frac{p^2}{p^2 + (1-p)^2)} \\
\mathrm{P}[w = 0|(y_1, y_2) = (1,0)] &= \frac{p(1-p)}{p(1-p) + (1-p)p} = 1/2 \\
\mathrm{P}[w = 0|(y_1, y_2) = (0,0)] &= \frac{(1-p)^2}{p^2 + (1-p)^2}.
\end{aligned}$$

To compute the capacity, we must assume that $\mathrm{P}[w=1] = \mathrm{P}[w=0] = 1/2$, so we have

$$\begin{aligned}
i(w=1;(y_1,y_2)=(1,1)) &= \log_2\left(\frac{\mathrm{P}[w=1|(y_1,y_2)=(1,1)]}{\mathrm{P}[w=1]}\right), \\
&= \log_2\left(\frac{2(1-p)^2}{(1-p)^2+p^2}\right), \\
i(w=1;(y_1,y_2)=(1,0)) &= 0 \\
i(w=1;(y_1,y_2)=(0,0)) &= \log_2\left(\frac{2p^2}{(1-p)^2+p^2}\right), \\
i(w=0;(y_1,y_2)=(0,0)) &= \log_2\left(\frac{2(1-p)^2}{(1-p)^2+p^2}\right), \\
i(w=0;(y_1,y_2)=(1,0)) &= 0 \\
i(w=0;(y_1,y_2)=(1,1)) &= \log_2\left(\frac{2p^2}{(1-p)^2+p^2}\right).
\end{aligned}$$

We now compute $I(w;y_1,y_2)$ by summing over all events:

$$\begin{aligned}
I(w;y_1,y_2) &= \sum_{a,b_1,b_2} \mathrm{P}[w=a,y_1=b_1,y_2=b_2]\, i(w=a;y_1=b_1,y_2=b_2) \\
&= \left((1-p)^2+p^2\right)\left(1-H\left(\frac{p^2}{(1-p)^2+p^2}\right)\right).
\end{aligned}$$

3.3 Prior, Extrinsic, Posterior and Intrinsic Probabilities

It is unsatisfying to have to keep assuming that w is uniformly distributed just because we don't know how it is distributed. There is a way to avoid having to make this assuption. In the situation in which a variable w is chosen, and then experiments are performed that reveal information about w, such as passing w through a channel, we call the initial probability of $w=1$ the *prior* probability of $w=1$, usually written

$$\mathrm{P}^{prior}[w=1].$$

In general, when w can take one of many values $a_1,\ldots,a_m$, the prior distribution is the vector of prior probabilities

$$\left(\mathrm{P}^{prior}[w=a_1], \mathrm{P}^{prior}[w=a_2], \ldots, \mathrm{P}^{prior}[w=a_m]\right).$$

Our experiments reveal the *extrinsic* probability of $w=1$ given the outcome of the experiment. For example, if y is the output of a channel on input w, and b is the value received, then

$$\mathrm{P}^{ext}[w=1|y=b] \stackrel{\text{def}}{=} \frac{\mathrm{P}[y=b|w=1]}{\mathrm{P}[y=b|w=1]+\mathrm{P}[y=b|w=0]}$$

is the extrinsic probability of $w=1$ given the event $y=b$. Up to now, we have really been computing *extrinsic* probabilities. For example, when we derived the interpretation of the output of a channel, we really derived the extrinsic probability.

If you know the prior probability, then you can combine this knowledge with the extrinsic probability to achieve the *posterior probability*: then actual probability of $w = 1$ given the channel output. Treating the prior and extrinsic probabilities as independent observations, and applying the calculation of the previous section, we obtain

$$\mathrm{P}^{post}\left[w = 1 | y = b\right] = \frac{\mathrm{P}^{ext}\left[w = 1 | y = b\right] \mathrm{P}^{prior}\left[w = 1\right]}{\mathrm{P}^{ext}\left[w = 1 | y = b\right] \mathrm{P}^{prior}\left[w = 1\right] + \mathrm{P}^{ext}\left[w = 0 | y = b\right] \mathrm{P}^{prior}\left[w = 0\right]}.$$

A useful exercise would be to re-derive the probability that $w = 1$ given $y = b$ assuming that w is not uniformly distributed, and to observe that one obtains the above formula.

We will occasionally also see the term *intrinsic probability*. This will usually be treated in the same way as the prior, but will be distinguished from the prior in that it will often be determined from previous experiments.