18.413: Error-Correcting Codes Lab	February 5, 2004

Lecture 2

Lecturer: Daniel A. Spielman

2.1 Channels

A channel is given by a set of input symbols $a_1, \ldots, a_m$, a set of output symbols $b_1, \ldots, b_n$, and a set of transition probabilities $p_{i,j}$, where

$$p_{i,j} = \Pr\left[b_j \text{ is received} \middle| a_i \text{ is sent}\right].$$

Given a general channel, we now examine how to determine the probability that a particular input was transmitted given that a particular output was received. That is, we examine how to interpret the output of the channel. Our derivation follows from many applications of the law of conditional probability. We let x be the random variable corresponding to the input to the channel, and y be the output. We assume that x is chosen uniformly from $a_1, \ldots, a_m$.

$$
\begin{aligned}
\Pr\left[x = a_i \middle| y = b_j\right] &= \frac{\Pr\left[x = a_i \text{ and } y = b_j\right]}{\Pr\left[y = b_j\right]} \\
&= \frac{\Pr\left[y = b_j \middle| x = a_i\right] \Pr\left[x = a_i\right]}{\Pr\left[y = b_j\right]} \\
&= \frac{\Pr\left[y = b_j \middle| x = a_i\right] \Pr\left[x = a_i\right]}{\sum_k \Pr\left[y = b_j \text{ and } x = a_k\right]} \\
&= \frac{\Pr\left[y = b_j \middle| x = a_i\right] \Pr\left[x = a_i\right]}{\sum_k \Pr\left[y = b_j \middle| x = a_k\right] \Pr\left[x = a_k\right]} \\
&= \frac{\Pr\left[y = b_j \middle| x = a_i\right] (1/m)}{\sum_k \Pr\left[y = b_j \middle| x = a_k\right] (1/m)} \\
&= \frac{\Pr\left[y = b_j \middle| x = a_i\right]}{\sum_k \Pr\left[y = b_j \middle| x = a_k\right]} \\
&= \frac{p_{j,i}}{\sum_k p_{j,k}}.
\end{aligned}
$$

2.2 Capacity and Mutual Information

The *capacity* of a channel provides a sharp threshold: If one communicates over a channel using any code of rate greater than the capacity, then the probabiliy that one will have a communication error tends to 1. On the other hand, there exist codes of every rate less than capacity that drive

the probability of communication error to zero. It is easy to compute the capacity of symmetric channels.

Definition 2.2.1. *A channel is symmetric if*

- *For all i_1 and i_2, the vectors $(p_{i_1,1}, \ldots, p_{i_1,n})$ and $(p_{i_2,1}, \ldots, p_{i_2,n})$ are permutations of each other, and*

- *For all j_1 and j_2, the vectors $(p_{1,j_1}, \ldots, p_{m,j_1})$ and $(p_{1,i_2}, \ldots, p_{m,j_2})$ are permutations of each other.*

Most of the channels we consider will be symmetric.

Let x be the random variable uniformly chosen from $a_1, \ldots, a_m$ and let y be the random variable giving the output of the channel on input x. Then, the capacity of the channel is the mutual information of x and y, written $I(x; y)$ and defined by

$$I(x; y) \stackrel{\text{def}}{=} \sum_{i,j} \Pr\left[x = a_i \text{ and } y = b_j\right] \log_2 \left(\frac{\Pr\left[x = a_i \text{ and } y = b_j\right]}{\Pr\left[x = a_i\right] \Pr\left[y = b_j\right]} \right)$$

If a channel has a simple description, the one can compute $I(x; y)$ directly. Otherwise, one can estimate $I(x; y)$ by experiment if one can compute the quantity $\Pr\left[y = b_j | x = a_i\right] / \Pr\left[y = b_j\right]$: repeatedly choose x at random, generate y, compute

$$i(x; y) \stackrel{\text{def}}{=} \log_2 \left(\frac{\Pr\left[x = a_i \text{ and } y = b_j\right]}{\Pr\left[x = a_i\right] \Pr\left[y = b_j\right]} \right),$$

and take the average of all the $i(x; y)$ values obtained.

2.3 What I should have said

We will extend the definition of a symmetric channel to the following:

Definition 2.3.1. *A channel is symmetric if its output symbols can be partitioned into sets such that for each set S in the partition,*

- *For all i_1 and i_2, the vectors $(p_{i_1,j})_{j \in S}$ and $(p_{i_2,j})_{j \in S}$ are permutations of each other, and*

- *For all $j_1 \in S$ and $j_2 \in S$, the vectors $(p_{1,j_1}, \ldots, p_{m,j_1})$ and $(p_{1,i_2}, \ldots, p_{m,j_2})$ are permutations of each other.*

x2

In particular, this defition includes the "nice" channels that I defined in class, which satisfy:

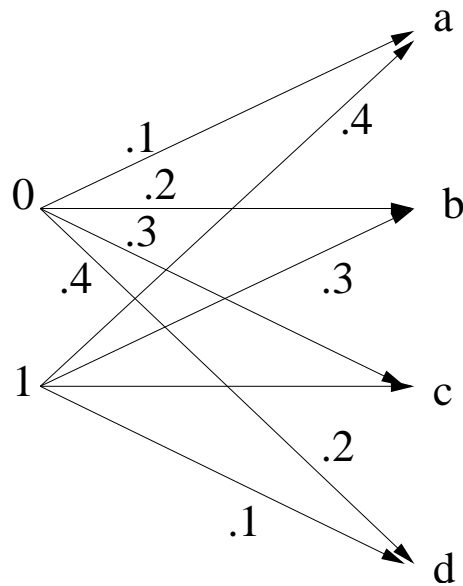- The channel has two input symbols, a_1 and $a_2 = 1$, and

Figure 2.1: A "nice" channel.

- The output symbols come in pairs, b_j and $b_{j'}$ such that $p_{1,j} = p(2, j')$ and $p_{2,j} = p(1, j')$.

For example, the following channel is symmetric:

The channels we will encounter in Small Project 1 are all nice.

Note that for any output symbol b_j of a nice channel, it is easy to compute $\Pr[y = b_j]$:

$$\Pr[y = b_j] = \Pr[x = 0 \text{ and } y = b_j] + \Pr[x = 1 \text{ and } y = b_j]$$
$$= \Pr[y = b_j | x = 0] \Pr[x = 0] + \Pr[y = b_j | x = 1] \Pr[x = 1].$$