College of Computer & Information Science
Northeastern University
CS G399: Algorithmic Power Tools I

Fall 2007
Handout 2 November 2007
Scribe: Emrah Bayraktaroglu

Lecture Outline:

- Introduction to Information Theory (Discrete Memoryless Channels)
- Zero-Error Capacity of a Communication Channel
- Pentagon

In this lecture, we will give an introduction to Shannon's mathematical model of information and capacity of a channel building on that. We will also present Pentagon, a well-known case that some upcoming discussions on capacity will be based on.

1 Introduction to Information Theory

Let $\overline{x}$ be a vector of symbols that represents the input and $\overline{y}$ be the output vector we get when we send $\overline{x}$ over the channel C. It is highly possible that the channel would alter the input and cause $\overline{y}$ to be different that $\overline{x}$. That would lead to cases such as:

YOU WILL NOT ATTACK $\rightarrow$ YOU WILL NOW ATTACK

At this point, Shannon comes up with DMCs (Discrete Memoryless Channels) to model the channels and information sent over them.

1.1 Discrete Memoryless Channel

DMCs are discrete and memoryless, because we are working on the discrete domain of symbols or messages on a one at a time basis which makes it memoryless. The information theory we will discuss here is based on DMCs. Here we model information probabilistically, *i.e.* the amount of information according to this model depends on the possibility of sending out different messages/symbols. Let $\overline{x}$ be a vector and P_x be the possibility of sending out the symbol x, where $x \in \overline{x}$. Here we have *the surprisal* of symbol x as;

$$\log\left(\frac{1}{P_x}\right)$$

. The *entropy*, or *uncertainty* of $\overline{x}$ is modeled as;

$$H(\overline{x}) = \sum_x P_x \log\left(\frac{1}{P_x}\right)$$

As our main intuition is defining the amount of information as the possibility of sending different messages, *entropy* of a symbol space gives us the average amount of information we can get from a message/symbol.

Now it's time to define the information over the channel and capacity of a channel. We define it as the average amount of information we get about the input by observing the output. Let X be a random variable representing the symbol space for input and Y be the output random variable. The mutual information is;

$$I(X, Y) = H(X) - H(X|Y)$$

.

And the capacity of a channel is the maximum information we can send over the channel. Capacity C is given as;

$$C = max_x I(X, Y) \tag{1}$$

2 Zero-Error Capacity

Theorem 1. ***(Shannon's Noisy Channel Theorem)***

Every DMC's capacity C given by (1) is attainable asymptotically error-free.

Let's examine this theorem in two cases;

Binary Case: This is the case where we have only two possible inputs. The capacity in this case is:

$$\frac{C}{1 - H(P_b)}$$

where $H(P_b)$ is the *binary entropy function.*

General Case (Zero-Error Capacity): In the general case, we will incorporate graphs to define zero-error capacity.

Let $G = (E, V)$ be a graph, V set of vertices, which represent symbols. Given $a, b \in V$, $(a, b) \in E$ if the symbols a and b cannot be confused, i.e. transmitted at the same time. The clique $\omega(G)$ corresponds to the set of symbols that can be used simultaneously with zero-error. This gives us the capacity;

$$\log_2 \omega(G)$$

where we only send strings of length 1. As we want to move on to a general case where we use n instead of 1, we will have to use G^n instead of G, which is $G \times G \times \ldots \times G$ (n times). Let's recall the definition of cross-product:

$G' = G \times H = (E', V')$

$V' = \{(g_1 h_1) | g_1 \in V_g, h_1 \in V_h\}$

$E' = \{< (g_1 h_1), (g_2 h_2) > | < g_1, g_2 >\in E_g \vee < h_1, h_2 >\in E_h\}$

Zero-error capacity, also known as Shannon capacity is given as;

$$SC(G^n) = \lim_{n \to \infty} \frac{1}{n} \log_2 \omega(G^n)$$

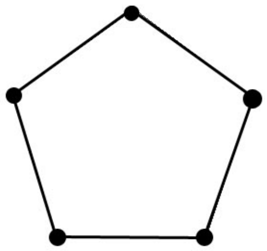

Figure 1: Representation of the pentagon (C_5)

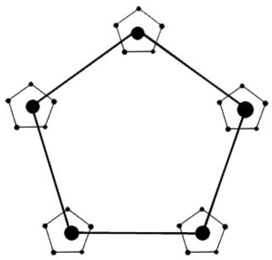

Figure 2: Representation for C_5^2

$$\approx \lim_{n \to \infty} \sqrt[n]{\omega(G^n)}$$

2.1 The Pentagon

This section will introduce the pentagon which will be used for capacity upper bound discussions in the next lecture. Pentagon is a graph C_5 that has five vertices each of which represent a symbol in the symbol domain connected in the shape of a pentagon. The figures 1 and 2 are representations for C_5 and (c_5^2), which yields (C_5) $\omega(C_5) = 2$ and $\omega(C_5^2) = 5$.