6.441 Transmission of Information	April 4, 2005
Lecture 13	
Lecturer: Madhu Sudan	*Scribe: Shaili Jain*

The lecture intro has been taken from Chung Chan's notes.

1 Break Up of Things We've Seen

We briefly summarize what we have seen so far in this course.

1.1 Phase I: The Tools

- Entropy
- Mutual Information
- AEP

1.2 Phase II: Exercise in Compression

- Source Coding and AEP
- Kraft's Inequality
- Shannon, Huffman, Lempel-Ziv Coding

1.3 Phase III

- Channel Coding
- Channel Capacity
  - Conditional probability distribution between X and Y models the channel
  - We used the ideas of random codign and maximum likelihood decoding
  - Joint AEP lets us say that we can get arbitrarily close to the channel capacity
  - Coding Theorem: For any discrete memoryless channel with capacity C, $\forall R < C$, there exists an encoding from $\{0,1\}^{Rn} \to X^n$, such that $P_{err} = Pr_{m \in \{0,1\}^{Rn}, Channel}[D(Channel(E(m)) \neq m] \to 0$
  - Converse: If $R > C$, then $P_{err} \to 1$

2 Error Exponent

In this lecture we show how to compute the error exponent over a binary symmetric channel using a random code ensemble and using maximum likelihood decoding (the same as minimum distance decoding in this case).

The general idea behind this lecture is based on the following facts: The probability of error of a discrete memoryless channel decays exponentially with n for a fixed rate below capacity. As n becomes large, the error exponent is representative of the quality of the code. Computing the error exponent turns out to be easier and more insightful than computing the probability of error exactly.

For the binary symmetric channel that has capacity, $C = 1 - H(p)$, we can write the error probability as follows: $Pr[Y^n = w] \geq p^n = 2^{-n \log \frac{1}{p}}$

Our goal is to come up with a general expression, where we conclude for any transmission $P_{err} > 2^{-nE_c(R)}$, where $E_c(R)$ is the error exponent. We are interested in how the error exponent $E_c(R)$ behaves. If $E_c(R) = 0$, then we know that the channel is virtually useless. On the other hand, if $E_c(R) = \infty$, then the channel is perfect. We know that $E_c(R) \leq \log_c \frac{1}{p}$, $\forall R > 0$, $\forall$ encoding, $\forall$ decoding (for the binary symmetric channel).

Today we study lower bounds for the error exponent for a random code ensemble.

Random Code For every $m \in \{0,1\}^{Rn}$ pick $E(m)$ uniformly and independently at random from $\{0,1\}^n$.

Decoding Scheme We use Maximum Likelihood Decoding (MLD). Given y, output m that maximizes $Pr[y \text{ received}|E(m) \text{ is transmitted}]$.

The channel capacity for a binary symmetric channel is $C = 1 - H(p)$.
If $R < C$, then we know that $R < 1 - H(p)$
Hence $H(p) < 1 - R$ and $p < H^{-1}(1 - R)$.
Let $P_R = H^{-1}(1 - R)$.
Clearly, $P_R > p$, so we can choose a τ such that $p < \tau < P_R$.

Type I error: $\triangle(E(m), y) \geq \tau n$ (This is the number of errors in transmission).

Pr[Type I error] $\to 0$

Type II error: $\exists\ m' \neq m$ such that $\triangle(E(m'), y) < \tau \cdot n$, where $\triangle$ denotes distance. (This is the case where y is to "close" to a codeword $E(m')$).

Pr[Type II error] $\to 0$ provided that $H(\tau) + R < 1$

We want a formula of the form:

$$Pr[\text{Error with MLD of Random Code}] \leq 2^{-(\ldots)\cdot n}$$

2.1 Maximum Likelihood Decoding

$$Pr[y \text{ received}|E(m) \text{ transmitted}] = p^d(1-p)^{n-d}$$

Let $\triangle(y, x) =$ the number of coordinates where x and y differ. Let $d = \triangle(y, E(m))$. The maximum likelihood decoding method chooses the message m such that $\triangle(y, E(m))$ is minimized.

Clearly, the MLD algorithm is unsuccessful if $\exists m' \neq m$ such that $\triangle(y, E(m')) < \triangle(y, E(m))$

Now consider the following scenario: $\exists \tau, m'$ such that $\triangle(y, E(m')) \leq \tau n \leq \triangle(y, E(m))$. Notice that the first inequality denotes a Type II error and the second inequality denotes a Type I error.

$$\begin{aligned} max_\tau[Pr[\triangle(y, E(m')) \leq \tau n \leq \triangle(y, E(m)))]] &\leq Pr[\text{Decoding Error}] \\ &\leq \Sigma_{\tau n=0}^{n}[\triangle(y, E(m')) \leq \tau n \leq \triangle(y, E(m))] \\ &\leq (n+1) max_\tau\{Pr[\triangle(y, E(m')) \leq \tau n \leq \triangle(y, E(m))]\} \end{aligned}$$

Hence we analyze the expression: $Pr[\triangle(y, E(m')) \leq \tau n \leq \triangle(y, E(m)))]$.

$$Pr[\text{Type I error}] = Pr[\triangle(y, E(m)) \geq \tau n] = \Sigma_{i=\tau n}^{n} \binom{n}{i} p^i (1-p)^{n-i}$$

$$\approx \binom{n}{\tau n} p^{\tau n}(1-p)^{n(1-\tau} \text{ assuming that } \tau > p$$

$$= 2^{-D(p||\tau)n}$$

$$Pr[\text{Type II error}] = 1 - (1 - \frac{2^{H(\tau)n}}{2^n})^{2^{Rn}-1}$$

$$\text{Fix } m', Pr[\triangle(E(m'), y) \leq \tau n] = \frac{2^{H(\tau)\cdot n}}{2^n}$$

$$Pr[\text{Type II error}] \approx 1 - 1 + 2^{Rn} \cdot \frac{2^{H(\tau)n}}{2^n}$$

since $R + H(\tau) < 1$. Thus we can write the probability of a Type II error as:

$$Pr[\text{Type II error } \approx 2^{-(1-H(\tau)-R)\cdot n}$$

$$Pr[\text{Type I and Type II error} = Pr[\text{Type I error}] \cdot Pr[\text{Type II error}]$$

$$\approx 2^{-[D(\tau||p+1-H(\tau)-R]\cdot n}$$

Notice that $1 - H(\tau) = D(\tau||\frac{1}{2})$, so we can write

$$Pr[\text{Type I and Type II error}] \approx 2^{-[D(\tau||p)+D(\tau||\frac{1}{2})-R]\cdot n}$$

Conclusion For a random code and maximum likelihood decoding, $P_{err} = 2^{-E_{RCE}(R)\cdot n}$, where E_{RCE} was derived to be:

$$E_{RCE}(R) = min_{p\leq\tau\leq P_R}\{D(\tau||p) + D(\tau||\frac{1}{2}) - R\}$$

Which choice of $\tau \in (p, P_R)$ minimizes $D(\tau||p) + D(\tau||\frac{1}{2})$?

We can find this by solving the following equation: $D'(\tau||p) = -D'(\tau||\frac{1}{2})$. We find that in this case, $\frac{\tau}{1-\tau} = \frac{\sqrt{p}}{\sqrt{1-p}}$. We find that when $\frac{\tau}{1-\tau} = \frac{\sqrt{p}}{\sqrt{1-p}}$, $D(\tau||p) + D(\tau||\frac{1}{2}) = 1 - \log(1 + 2\sqrt{p(1-p)}$.

A good reference for today's lecture is a paper by Barg and Forney entitled "Random codes: Minimum distances and error exponents" in IEEE Transactions on Information Theory in Sept 2002.