Iowa State University Lecturer: Aditya Ramamoorthy

EE 520: Topics in Communications: Multiuser Information Theory

Sep. 30, 2007 Scribe: Shizheng Li

Lecture 11

1 Source coding with side information

Consider the problem where two sources X and Y are encoded separately but only X is to be recovered without distortion. The model is given in the figure. The rate of X and Y are R_1 and R_2 respectively. X is encoded into $W_1 \in \{1, \ldots, 2^{nR_1}\}$ and Y is encoded into $W_2 \in \{1, \ldots, 2^{nR_2}\}$. Some simple cases are: 1) If $R_2 \geq H(Y)$, from Slepian-Wolf theorem, $R_1 \geq H(X|Y)$. 2) If $R_2 = 0$, from source coding theorem, $R_1 \geq H(X)$. In general, the rate region is given by

$$R_1 \geq H(X|U) \tag{1}$$

$$R_2 \geq I(Y;U) \tag{2}$$

where U is an auxiliary random variable with alphabet $\mathcal{U}$ such that $X \to Y \to U$ form a Markov chain and $|\mathcal{U}| \leq |\mathcal{Y}| + 2$.

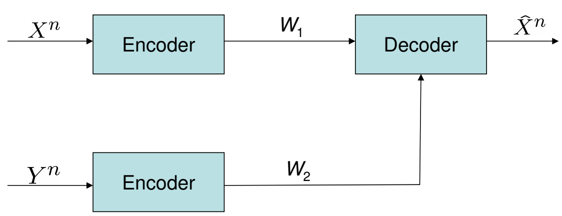

Figure 1: Source coding with side information problem model

2 Converse

In converse part, we shall show that for any encoding scheme that has a small probability of error, the rates should satisfy the equations above. We start from R_2.

$$\begin{aligned}
nR_2 &\geq H(W_2) \\
&= H(W_2) - H(W_2|Y^n) \text{ } W_2 \text{ is a function of } Y^n\text{, so } H(W_2|Y^n) = 0. \\
&= I(W_2;Y^n) \\
&= H(Y^n) - H(Y^n|W_2) \\
&= \sum_{i=1}^{n} H(Y_i) - H(Y_i|Y^{i-1}, W_2) \text{ By independency of } Y^i \text{ and chain rule} \\
&= \sum_{i=1}^{n} I(Y_i;U_i) \qquad (3)
\end{aligned}$$

(3) follows from the fact that we define $U_i = (Y^{i-1}, W_2)$. $X_i \to Y_i \to U_i$ form a Markov chain because the dependency between X_i and W_2 is only through Y_i (detailed proof can be done on a dependency graph).

Thus, $R_2 \geq \frac{1}{n}\sum_{i=1}^{n} I(Y_i; U_i)$.

Next, we prove for R_1.

$$
\begin{aligned}
nR_1 &\geq H(W_1) \\
&\geq H(W_1|W_2) \\
&= H(W_1|W_2) - H(W_1|W_2, X^n) \text{ } W_1 \text{ is a function of } X_n \\
&= I(W_1; X^n|W_2) \\
&= H(X^n|W_2) - H(X^n|W_1, W_2) \\
&\geq H(X^n|W_2) - n\varepsilon_n \qquad (4) \\
&= \sum_{i=1}^{n} H(X_i|X^{i-1}, W_2) - n\varepsilon_n \text{ Chain rule} \\
&\geq \sum_{i=1}^{n} H(X_i|X^{i-1}, Y^{i-1}, W_2) - n\varepsilon_n \\
&= \sum_{i=1}^{n} H(X_i|X^{i-1}, U_i) - n\varepsilon_n \qquad (5) \\
&= \sum_{i=1}^{n} H(X_i|U_i) - n\varepsilon_n \text{ because } X^{i-1} \to U_i \to X_i \text{ form a Markov chain.}
\end{aligned}
$$

(4) follows from Fano's inequality: $H(X^n|W_1, W_2) \leq n\varepsilon_n$.

(5) follows from the fact that we define $U_i = (Y^{i-1}, W_2)$. $X^{i-1} \to U_i \to X_i$ form a Markov chain because X^{i-1} and X_i are independent and conditioning on (Y^{i-1}, W_2) does not introduce dependency. This can be proved on a dependency graph in Fig.2. Note that if we remove $Y_1, Y_2, \ldots, Y_{i-1}$ and W_2, $X_1, X_2, \ldots, X_{i-1}$ and X_i are not connected. Thus, $R_1 \geq \frac{1}{n}\sum_{i=1}^{n} I(Y_i; U_i)$. We shall

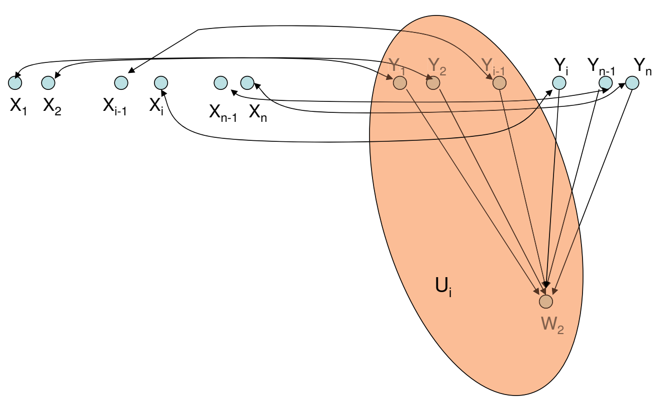

Figure 2: Dependency graph for showing $X^{i-1} \to U_i \to X_i$.

show that from $R_2 \geq \frac{1}{n}\sum_{i=1}^{n} I(Y_i;U_i)$ and $R_1 \geq \frac{1}{n}\sum_{i=1}^{n} I(Y_i;U_i)$, we can get $R_2 \geq I(Y;U), R_1 \geq H(X|U)$ by a time-sharing argument in the next lecture.