Lecture 01

Note Title 1/11/2012

Review of prob notation (assuming discrete random variables)

Capital X Y Z : random variable

small x y z : deterministic values

$P_X(x) = \text{Prob}(X = x)$ is a number / or a function w.r.t. x

$P_X(\cdot) : S_X \longmapsto [0,1]$ is a function

$P_{Y|X}(y|x) = \text{Prob}(Y = y \mid X = x)$ is a number

$P_{Y|X}(\cdot|\cdot) : S_Y \times S_X \longmapsto [0,1]$

$P_{Y|X}(y|\cdot) : S_X \longmapsto [0,1]$

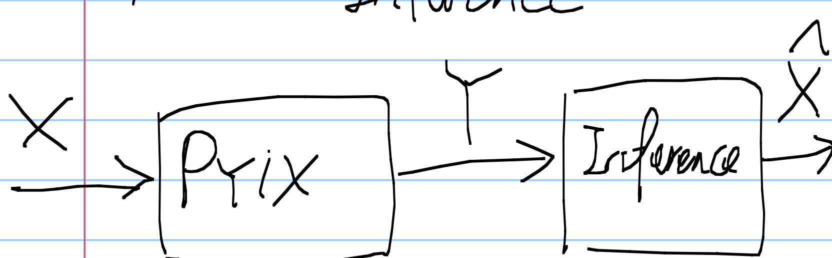
ECE 639 : ECC

ECE 642 : Info theory

ECE 645 : Estimation theory

Inference Method for codes on graphs

* Inference:



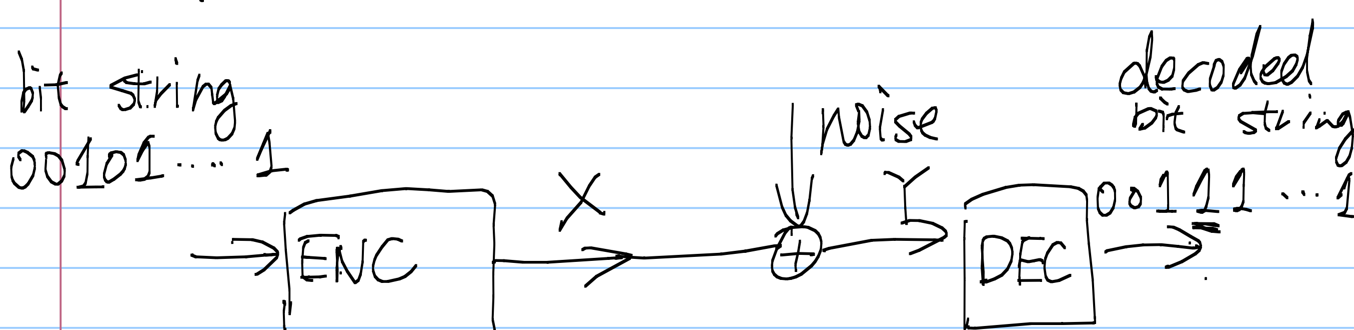
Deciding X by observing only Y.

① Also known as detection and estimation, hypothesis testing.

② Different structure of X decide the difficulty level of inference.

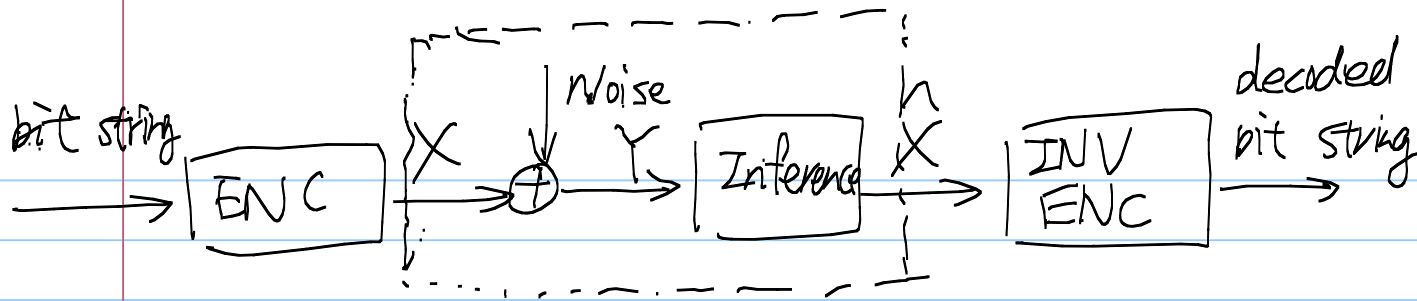
X: Transmitted Signal	X: The movement/position of an object	X: The network congestion level
Y: Received Signal	Y: The pixel image of a camera	Y: The reception status of a given session

* Error control codes:



① When there is no noise, DEC is simply the inverse of ENC. ✓

② When there is noise, the DEC can be decomposed as two parts



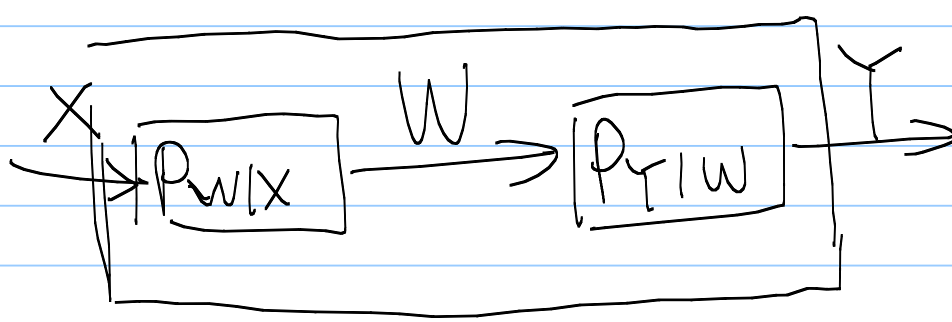
③ Studying new inference methods improves decoding

The difference between Error Control Coding & pure inference is that one can <u>design the structure of X</u> by designing $\boxed{\text{ENC}}$

* <u>Graphical models</u>

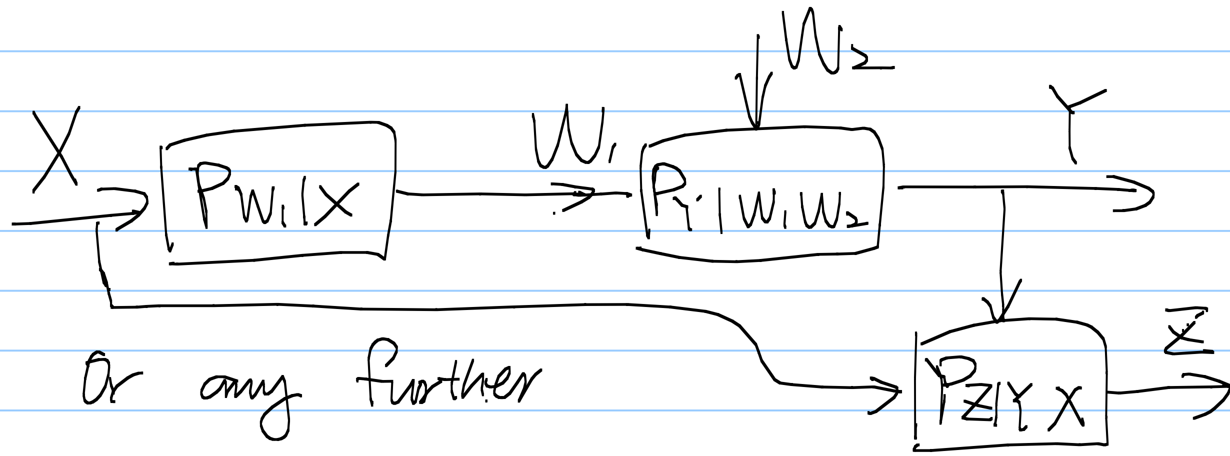
For many inference problems the relationship of X, Y follows some graphical structure.

<u>Exame</u>: The Markov chain



Knowing the underlying structure enables more efficient algorithms.

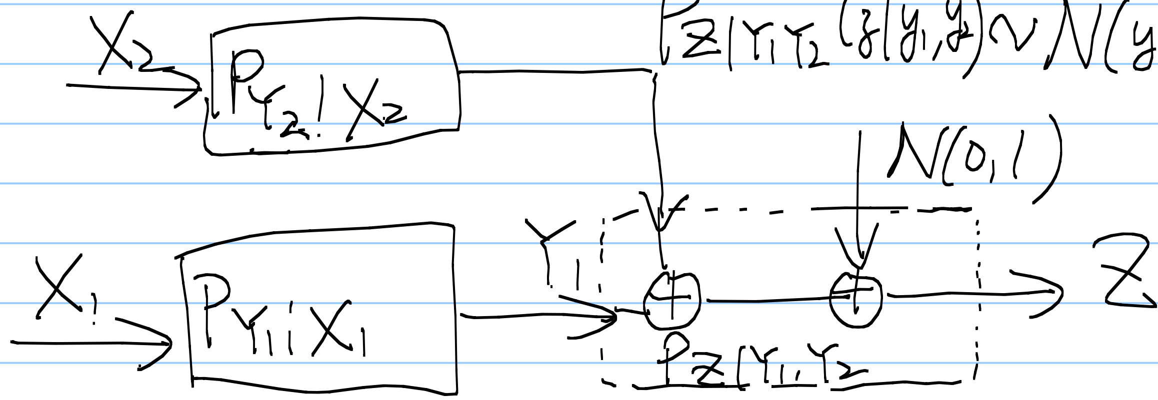
A popular model is the Markov Inference Networks:



Or any further concatenation.

For example we may have

$P_{Z|Y_1Y_2}(z|y_1,y_2) \sim N(y_1+y_2, 1)$

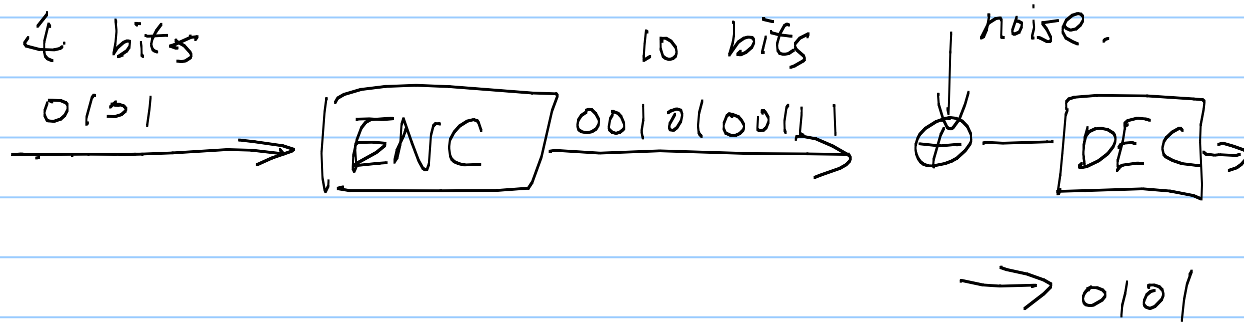


which is a multiple access channel.

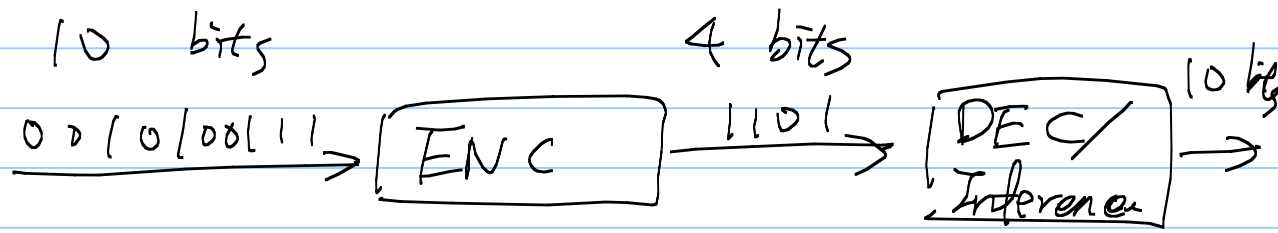
Computer Networks are another example, which is easily modelled as graphs.

* Applications ~~of~~ "coding"

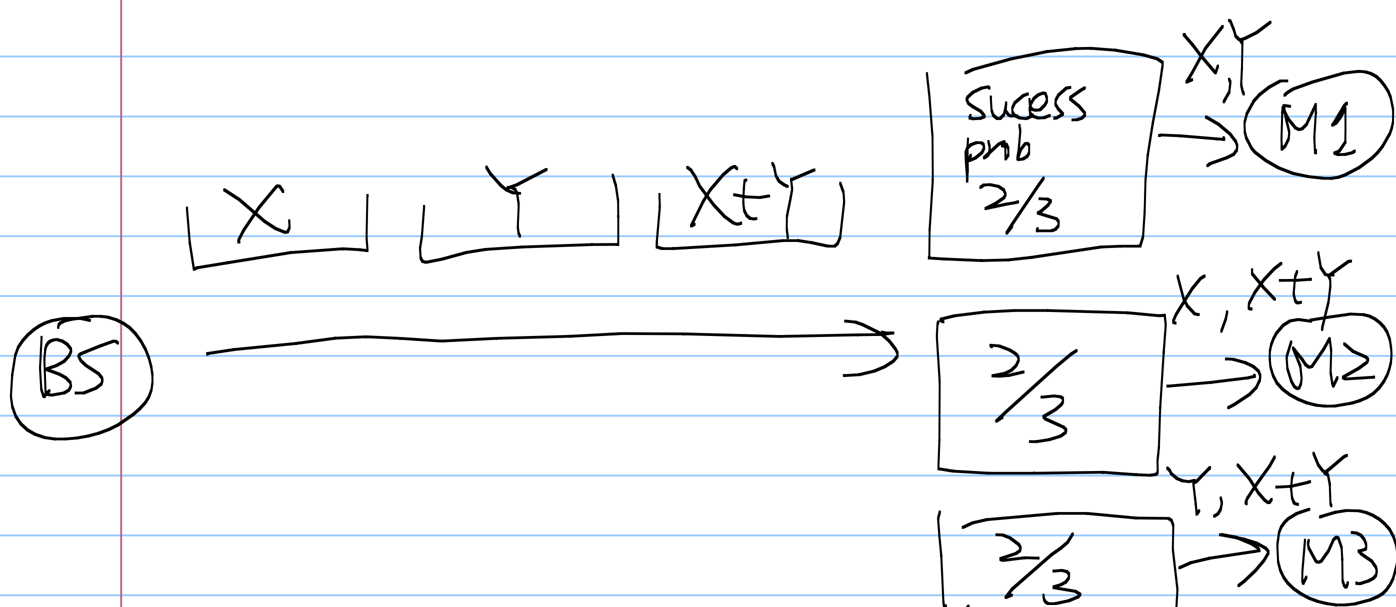
① Enhance error resilience



② Data compression



③ Enhance broadcast rates



any 2 received packets can recover both information bits X and Y

Topic 1: From inference to mutual information to channel capacity.

$\boxed{\text{An Inference Problem:}}$

Given $P_{Y|X}(\cdot|\cdot)$, $P_X(\cdot)$ and the observation $Y=y$

$$\xrightarrow{X} \boxed{P_{Y|X}(\cdot|\cdot)} \xrightarrow{Y=y}$$

Find $\hat{X}$ as a function of y.
Or find $P_{X|Y}(\cdot|y)$

Ex: X is the 4PAM signal.
Y is the output of additive white Gsn channel.

Then P_X is uniform on $\{-3,-1,1,3\}$

$$P_{Y|X} \sim N(x, \sigma^2)$$

A special case of the inference prob is the "hypothesis testing"

Two hypotheses.

H_0: $Y \sim P_Y$

H_1: $Y \sim Q_Y$

$P(H_0 \text{ is true}) = p_0$ $\quad P(H_1 \text{ is true}) = 1-p_0$

observe $Y=y$. Find $P(H_0 \text{ is true} \mid Y=y)$

It is indeed an inference problem

$X = 0$ or 1.

P_Y : $P_{Y|X}(\cdot \mid 0)$ $\qquad P(H_i \text{ is true})$

Q_Y : $P_{Y|X}(\cdot \mid 1)$ $\qquad = P(X=i)$

Solution of the HT:

The maximum a posteriori prob (MAP) detector:

$$\hat{X}_{MAP}(y) = \max_{x=0,1} P(X=x \mid Y=y)$$

By Bayes' Rule

$$P(X=0 \mid Y=y) = \frac{P(X=0, Y=y)}{P(Y=y)}$$

$$= \frac{P(Y=y \mid X=0)\, P(X=0)}{P(Y=y \mid X=0)\, P(X=0) + P(Y=y \mid X=1)\, P(X=1)}$$

By symmetry

$$\Rightarrow \text{If } \frac{P_{X|Y}(0|y)}{P_{X|Y}(1|y)} = \frac{P_{Y|X}(y|0)\cdot P_X(0)}{P_{Y|X}(y|1)\cdot P_Y(1)} \gtreqless 1$$

we choose $\hat{X}_{MAP}(y) = 0$ or 1

An equivalent form of the MAP detector is the likelihood ratio test.

$$\boxed{\text{Likelihood ratio} \quad L(y) = \frac{P_{Y|X}(y|0)}{P_{Y|X}(y|1)}}$$

$$\hat{X}_{MAP}(y) = \begin{cases} 0 & > \\ 0 \text{ or } 1 & \text{if } L(y) = \frac{P(X=1)}{P(X=0)} = \gamma \\ 1 & < \end{cases}$$

---

① The likelihood ratio test can be converted to the log likelihood ratio test by taking the logarithm

$$\log\left(\frac{P_{Y|X}(y|0)}{P_{Y|X}(y|1)}\right) \text{ vs. } \log\frac{P(X=1)}{P(X=0)}$$

② When we set $\gamma = 1$, & compare $L(y) \gtrless 1$, it becomes a maximum likelihood (ML) detector.

※ The MAP & ML detectors with 3 competing hypotheses

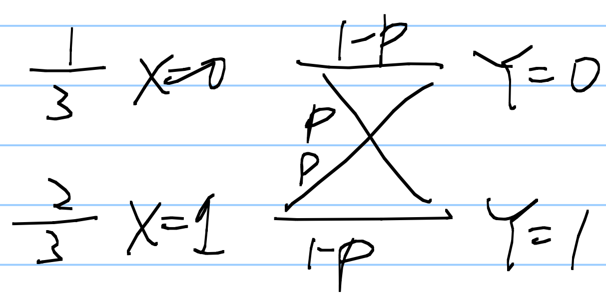
$x = 0, 1, 2$

$$\hat{X}_{MAP}(y) = \operatorname{argmax} P_{X|Y}(x|y)$$

$$\hat{X}_{ML}(y) = \operatorname{argmax} P_{Y|X}(y|x)$$

$\hat{X}_{MAP}(\cdot) = \hat{X}_{ML}(\cdot)$ when $P(X=0) = P(X=1) = P(X=2) = \frac{1}{3}$

Example: A binary symmetric channel with non-uniform prior distribution



Q: Find the $\hat{X}_{MAP}(y)$.

Ans: We need to determine $\hat{X}_{MAP}(y)$ as a function of y.

When $Y=0$

$$P_{X|Y}(x|0)=\begin{cases} \dfrac{\frac{1}{3}(1-p)}{\frac{2}{3}p+\frac{1}{3}(1-p)}=\dfrac{1-p}{1+p} & x=0 \\ \dfrac{2p}{1+p} & x=1 \end{cases}$$

When $Y=1$

$$P_{X|Y}(x|1)=\begin{cases} \dfrac{p}{2-p} & x=0 \\ \dfrac{2-2p}{2-p} & x=1 \end{cases}$$

The $\hat{X}_{MAP}(y)=\begin{cases} ? & \text{if } y=0 \\ ? & \text{if } y=1 \end{cases}$

We have three cases.

Case I

If $p<\frac{1}{3}$

$$\hat{X}_{MAP}(y)=\begin{cases} 0 & \text{if } y=0 \\ 1 & \text{if } y=1 \end{cases}$$

Case II

If $\frac{1}{3}<p<\frac{2}{3}$

$$\hat{X}_{MAP}(y)=1 \quad \text{for } y=0,1$$

Case III

If $\frac{2}{3}<p<1$

$$\hat{X}_{MAP}(y)=\begin{cases} 1 & \text{if } y=0 \\ 0 & \text{if } y=1 \end{cases}$$