2.4 Markov Processes and Sources with Memory.

2.4.1 Markov Processes

— Let $A = \{a\}$ be the alphabet of a discrete source having $|A|$ symbols.

— Assume that the source emits a time sequence of symbols $(s_0, s_1, \ldots, s_t, \ldots)$

Def: j^{th}-Order Markov Process

A is called a j^{th}-order Markov Process if the conditional probability $P(s_t | s_{t-1}, s_{t-2}, \ldots, s_0)$ depends only on j previous symbols.

$$\Rightarrow \Pr(s_t | s_{t-1}, \ldots, s_0) = P(s_t | s_{t-1}, s_{t-2}, \ldots, s_{t-j})$$

Def: State of the Markov process at time t is the string

$$S_t = (s_{t-1}, \ldots, s_{t-j})$$

* A j^{th} order Markov Process has $N = |A|^j$ possible states

• Let $\pi_n(t)$ represent the probability of being in state n at time t.

$$\Rightarrow \Pi_t = \begin{bmatrix} \pi_0(t) \\ \pi_1(t) \\ \vdots \\ \pi_{N-1}(t) \end{bmatrix} \equiv \text{Probability Dist. of the system at time } t.$$

• Transition Probability Matrix

— At each state at time t, there are $|A|$ possible next states at time $t+1$.

— Let $P_{i|k}$ be the conditional prob. of going to state i from state k

$$k \xrightarrow{P_{i|k}} i$$

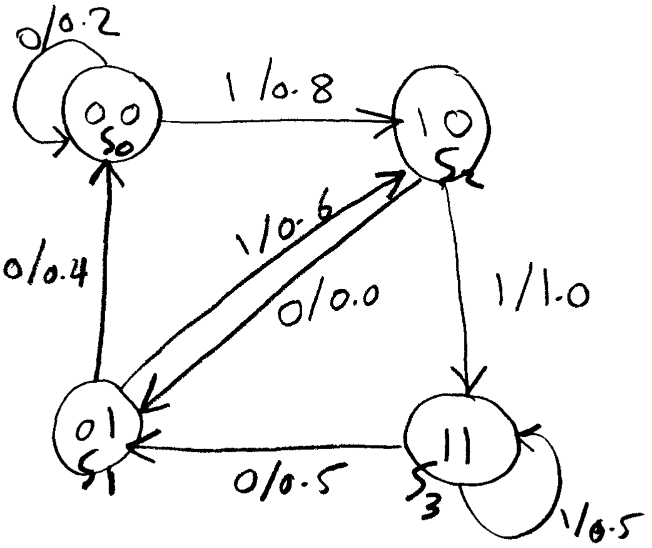
$\Rightarrow$ Transition Prob Matrix

$$P_{A|\Pi} = \begin{bmatrix} P_{0|0} & P_{0|1} & \cdots & P_{0|N-1} \\ P_{1|0} & P_{1|1} & \cdots & P_{1|N-1} \\ \vdots & & & \\ P_{N-1|0} & \cdots & & P_{N-1|N-1} \end{bmatrix}$$

• State Prob. Distribution

$$\Pi_{t+1} = P_{A|\Pi} \Pi_t$$

Example



The transition probability Matrix

$$P_{A|\Pi} = \begin{bmatrix} 0.2 & 0.4 & 0 & 0 \\ 0 & 0 & 0 & 0.5 \\ 0.8 & 0.6 & 0 & 0 \\ 0 & 0 & 1.0 & 0.5 \end{bmatrix}$$

— Assume that all states are equally Prob. at time zero. Find the state Prob. at time $t=1$?

$$\Pi_{t+1} = \begin{bmatrix} 0.2 & 0.4 & 0 & 0 \\ 0 & 0 & 0 & 0.5 \\ 0.8 & 0.6 & 0 & 0 \\ 0 & 0 & 1 & 0.5 \end{bmatrix} \begin{bmatrix} 0.25 \\ 0.25 \\ 0.25 \\ 0.25 \end{bmatrix}$$

$$\Pi_1^T = [0.15 \;\; 0.125 \;\; 0.35 \;\; 0.375]$$

2.4.2 Steady-State Probability and the Entropy Rate.

By Induction,

$$\Pi_t = (P_{A|\Pi})^t \, \Pi_0$$

Def: Ergodic Markov Process

For long t, Π_t approaches a steady-state

$$\Rightarrow \Pi_{t+1} = \Pi_t$$

See Example 2.4.3

Def: Entropy Rate of an ergodic Markov Process.

$$R = \lim_{t \to \infty} \frac{1}{t} H(A_0, A_1, \ldots, A_{t-1})$$

— At steady-state, the state probabilities will be Π_n and the entropy rate will be a function of the symbol Probabilities.

— Suppose we are in state S_n at time t,

$\Rightarrow$ The conditional entropy of the next symbol a is given by

$$H(A|S_n) = \sum_{a \in A} Pr(a/S_n) \log_2 \left(\frac{1}{Pr(a|S_n)} \right)$$

- Since each possible symbol a leads to a unique next state,

$$\Rightarrow H(A|S_n) = \sum_{i=0}^{N-1} P_{i|n} \log_2\left(\frac{1}{P_{i|n}}\right)$$

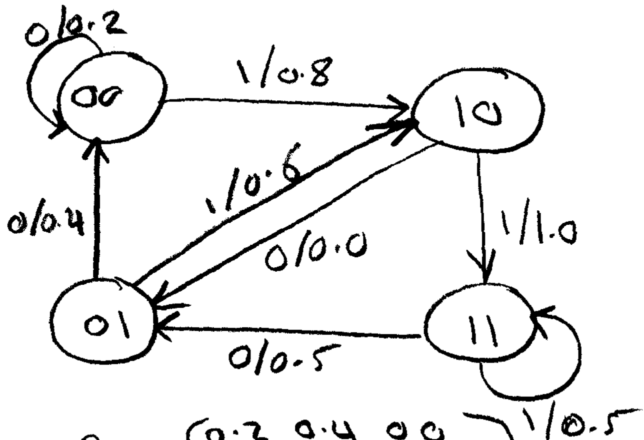
- For large t, the probability of being in state S_n is given by its steady-state prob. π_n

Average Rate $\Rightarrow$ $$R = \sum_{n=0}^{N-1} \pi_n H(A|S_n)$$

$$= \sum_{n=0}^{N-1} \pi_n \sum_{i=0}^{N-1} P_{i|n} \log_2\left(\frac{1}{P_{i|n}}\right)$$

- Therefore, the ~~steady~~ Entropy Rate of an ergodic Markov Process is a function only of its steady state prob. dist. and transition prob.

Example 2.4.4



$$P_{A|\pi} = \begin{bmatrix} 0.2 & 0.4 & 0 & 0 \\ 0 & 0 & 0 & 0.5 \\ 0.8 & 0.6 & 0 & 0 \\ 0 & 0 & 1 & 0.5 \end{bmatrix}$$

we need to find the steady-state prob. for which $\pi_{t+1} = \pi_t$

$$\Rightarrow \pi_{t+1} = P_{A|\pi}\, \pi_t$$

we get the following set of equations

$$\pi_0 = 0.2\pi_0 + 0.4\pi_1$$
$$\pi_1 = 0.5\pi_3$$
$$\pi_2 = 0.8\pi_0 + 0.6\pi_1$$
$$\pi_3 = \pi_2 + 0.5\pi_3$$

Also $\pi_0 + \pi_1 + \pi_2 + \pi_3 = 1$

Solving the above equations (see Example 2.4.3) we get

$$\pi_0 = \frac{1}{9}$$
$$\pi_1 = \pi_2 = \frac{2}{9}$$
$$\pi_3 = \frac{4}{9}$$

Thus, the Entropy Rate is

$$R = \sum_{n=0}^{3} \pi_n \sum_{i=0}^{3} P_{i|n} \log_2\left(\frac{1}{P_{i|n}}\right)$$

$$= \frac{1}{9}\left(0.2 \log_2 \frac{1}{0.2} + 0.8 \log_2 \frac{1}{0.8}\right)$$

$$+ \frac{2}{9}\left(0.4 \log_2 \frac{1}{0.4} + 0.6 \log_2 \frac{1}{0.6}\right)$$

$$+ \frac{2}{9}\log_2(1) + \frac{4}{9}\left[2(0.5)\log_2 2\right]$$

$$\boxed{R = 0.740}$$

The steady state symbol Prob. are:

$$Pr(0) = \sum_{n=0}^{3} \pi_n Pr(0|S_n) = \frac{0.2}{9} + \frac{0.4(2)}{9} + \frac{0.5(4)}{9}$$

$$= \frac{1}{3}$$

$$Pr(1) = 1 - Pr(0) = \frac{2}{3}$$

And

$$H(A) = \sum_{a=0}^{1} Pr(a) \log_2\left(\frac{1}{Pr(a)}\right) = 0.9183$$

Notice that your Entropy Rate is

$$R < H(A).$$

Thus, we manage to Reduce the Entropy of the source by using Memory and markov chains.

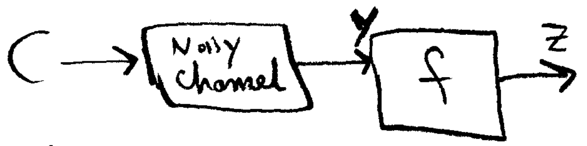
2.5 Markov Chains and Data Processing

- If the entropy rate of C exceeds the channel capacity $\Rightarrow$ information loss.

Q) Is there any kind of processing that can be performed on Y might Recover the lost information?

- The answer is NO.

<u>Why?</u> <u>Proof:</u>



Let $Z = f(Y)$ and the joint prob. be

$$P_{CYZ} = P_{CY} P_{Z|CY}$$

Since Z is a function of $Y \Rightarrow P_{Z|CY} = P_{Z|Y}$

$$\therefore P_{CYZ} = P_{CY} P_{Z|Y} \Rightarrow \text{Markov Chain}$$

- Define the conditional mutual information to be:

$$I(C;Y/Z) \equiv H(C|Z) - H(C|Y,Z)$$

"The reduction in our uncertainty of C due to our knowledge of Y when Z is given to us."

- Recall that $I(C;Y,Z) = H(C) - H(C|Y,Z)$

$$\Rightarrow I(C;Y,Z) = \underbrace{H(C) - H(C/Z)}_{I(C;Z)} + I(C;Y/Z)$$

we get,

$$(1) \quad I(C;Y,Z) = I(C;Z) + I(C;Y/Z)$$

since $I(C;Y,Z) = I(C;Z,Y)$

$$(2) \quad \Rightarrow I(C;Y,Z) = I(C;Y) + I(C;Z|Y)$$

and since

$$I(C;Z|Y) = H(C/Y) - H(C|Y,Z)$$

since $Z = f(Y)$, if we are given $Y \Rightarrow Z$ is completely determined $\Rightarrow H(C|Y,Z) = H(C|Y)$

$$\Rightarrow I(C;Z|Y) = 0$$

Combining (1) and (2)

$$\Rightarrow I(C;Z) + I(C;Y/Z) = I(C;Y)$$

Since $I(C;Y/Z) \geq 0$

with equality if $f(Y)$ is one-to-one and onto.

$$\Rightarrow I(C;Z) \leq I(C;Y)$$

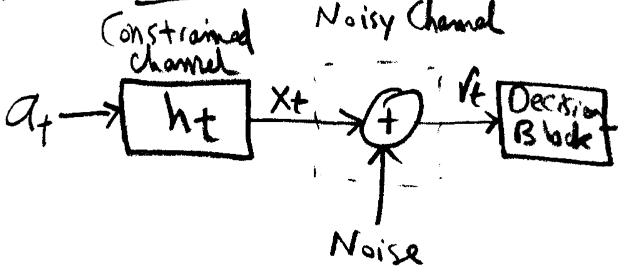
"Data Processing inequality"

<u>Theorem:</u> Additional Processing of the channel output Y can at ~~least~~ best result in no further loss of information and may even result in additional loss.

<u>Examples</u>: - Loss due to quantization errors
- Loss due to round off and truncation errors
- Many-to-One Mapping from Y to Z

$$Z = Y^2 \text{ or } Z = |Y|$$

2.6 Constrained Channels



2.6.2 Linear-Time-Invariant (LTI) channels

h_t is called the channel Impulse Responce.

The channel output is

$$x_t = \sum_{k=-\infty}^{\infty} h_k a_{t-k}$$

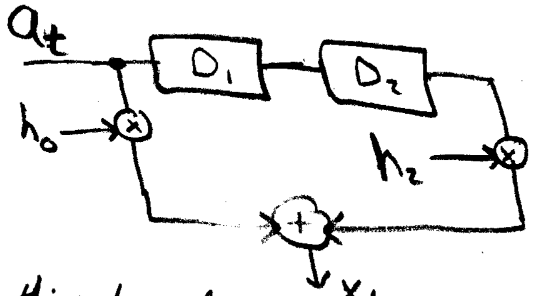
For a finite Impulse Responce {FIR} the channel can be modeled using a Markov Process.

The recieved noisy signal is $r_t = x_t + \eta_t$

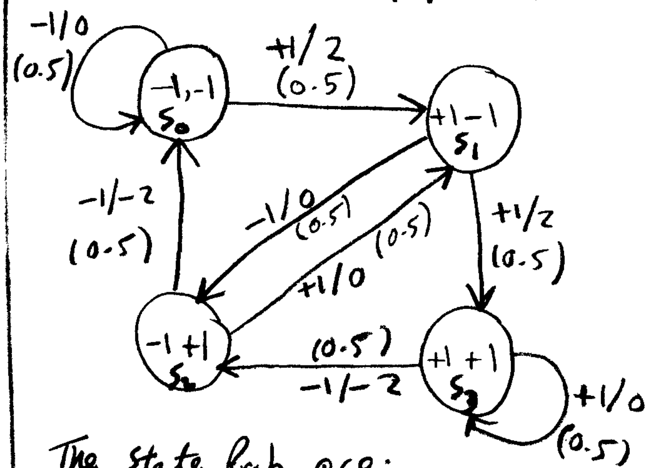
Example 2.6.1

- Let A be a memoryless binary source with equiprobable symbols $A = \{-1, +1\}$
- Let the band-limited channel have impulse response $\{h_0 = 1, h_1 = 0, h_2 = -1\}$
- Find the steady-state Entropy of the channel's output and the entropy rate of the sequence x_t.

$$x_t = h_0 a_t + h_2 a_{t-2}$$



this channel can be modeled by a 2nd-order Markov Process.



The state Prob. are:

$$\Pi_{t+1} = \begin{bmatrix} 0.5 & 0 & 0.5 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0 & 0.5 & 0 & 0.5 \end{bmatrix} \Pi_t$$

Solve the following equations to find the steady state prob.

$$\pi_0 = 0.5\,\pi_0 + 0.5\,\pi_2$$
$$\pi_1 = 0.5\,\pi_0 + 0.5\,\pi_2$$
$$\pi_2 = 0.5\,\pi_1 + 0.5\,\pi_3$$
$$\pi_3 = 0.5\,\pi_1 + 0.5\,\pi_3$$

Also, $\pi_0 + \pi_1 + \pi_2 + \pi_3 = 0$

$\xrightarrow{\text{Solution}}$ $\pi_i = 0.25$ for each state

The output symbol Prob at steady state are:

$$Pr(x_t) = \sum_{i=0}^{3} Pr(x_t|S_i)\,\pi_i$$

$\Rightarrow$ $Pr(-2) = 0.25 \times 0.5 + 0.25 \times 0.5 = 0.25$

$Pr(0) = 4 \times 0.25 \times 0.5 = 0.5$

$Pr(+2) = 2 \times 0.25 \times 0.5 = 0.25$

$\Rightarrow$ $H(x) = \sum_{x \in X} Pr(x) \log_2\left(\frac{1}{Pr(x)}\right)$

Entropy at steady state $= 1.5$

Also, Entropy Rate

$$R = \sum_{i=0}^{3} \pi_i \sum_{x \in X} Pr(x|S_i) \log_2\left(\frac{1}{Pr(x|S_i)}\right)$$

$= 1$