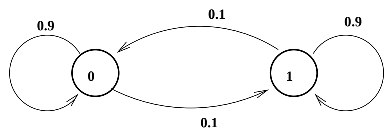
SySc/ECE Information Theory **10/26/04**
Preparation for Mid-Term **11/2/04**

1. Consider the Markov process described by the diagram:



   (a) Find the stationary probabilities (μ_0 and μ_1)

   (b) Calculate the entropy $H(X)$ of the stationary distribution.

   (c) Calculate the entropy *rate* $H(\mathcal{X})$. (You will find it easier to express entropies in terms of log(9) and log(10) (look at the worked example 1f).)

   (d) Calculate the probability distribution for the second extension, then calculate the entropy for that distribution.

   (e) Calculate the probability distribution and entropy for the third extension.

   (f) Calculate the probability distribution and entropy for the fourth extension.

s_i	p_i
0000	$5 \cdot 9^3 \cdot 10^{-4}$
0001	$5 \cdot 9^2 \cdot 10^{-4}$
0010	$5 \cdot 9^1 \cdot 10^{-4}$
0011	$5 \cdot 9^2 \cdot 10^{-4}$
0100	$5 \cdot 9^1 \cdot 10^{-4}$
0101	$5 \cdot 10^{-4}$
0110	$5 \cdot 9^1 \cdot 10^{-4}$
0111	$5 \cdot 9^2 \cdot 10^{-4}$
1000	$5 \cdot 9^2 \cdot 10^{-4}$
1001	$5 \cdot 9^1 \cdot 10^{-4}$
1010	$5 \cdot 10^{-4}$
1011	$5 \cdot 9^1 \cdot 10^{-4}$
1100	$5 \cdot 9^2 \cdot 10^{-4}$
1101	$5 \cdot 9^1 \cdot 10^{-4}$
1110	$5 \cdot 9^2 \cdot 10^{-4}$
1111	$5 \cdot 9^3 \cdot 10^{-4}$

$$
\begin{aligned}
H &= -2\sum_{k=0}^{3}\binom{3}{k}\cdot 5\cdot 9^k\cdot 10^{-4}\log\left(9^k\cdot 5\cdot 10^{-4}\right)\\
&= -\big(2\cdot 5\cdot 10^{-4}\log(5\cdot 10^{-4})+6\cdot 5\cdot 9^1\cdot 10^{-4}\log(5\cdot 9^1\cdot 10^{-4})\\
&\quad +6\cdot 5\cdot 9^2\cdot 10^{-4}\log(5\cdot 9^2\cdot 10^{-4})+2\cdot 5\cdot 9^3\cdot 10^{-4}\log(5\cdot 9^3\cdot 10^{-4})\big)\\
&= -\log(5\cdot 10^{-4})-5\cdot 10^{-4}\left(6\cdot 9+6\cdot 9^2\cdot 2+2\cdot 9^3\cdot 3\right)\log 9\\
&= 4\log 10-\log 5-2.7\log 9 = 1.668\text{nats} = 2.407\text{bits}
\end{aligned}
$$

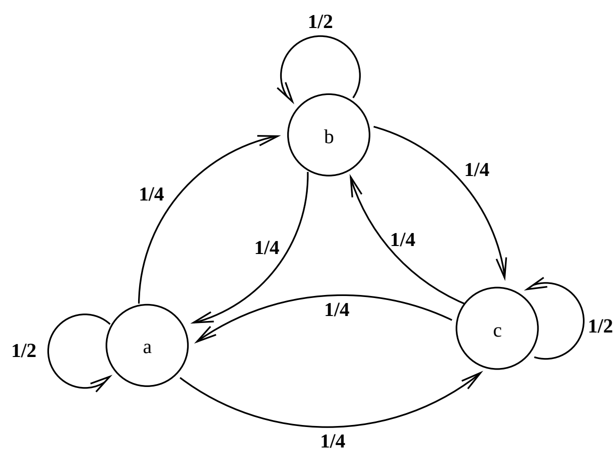
(g) Find radix 2 Huffman codes for single bits and for the extensions of length 2 and 3; calculate the average length of each Huffman code. Calculate the average number of coded output bits per input bit. Do the lengths of your Huffman codewords satisfy the Kraft inequality?

(h) Find radix 3 Huffman codes for the extensions of length 2 and 3; calculate the average length of each Huffman code. Also calculate the average number of coded output trits per input bit. Do the lengths of your Huffman codewords satisfy the Kraft inequality?

(i) In the limit of large n, what is the average length of the radix 2 Huffman code, for the n^{th} extension?

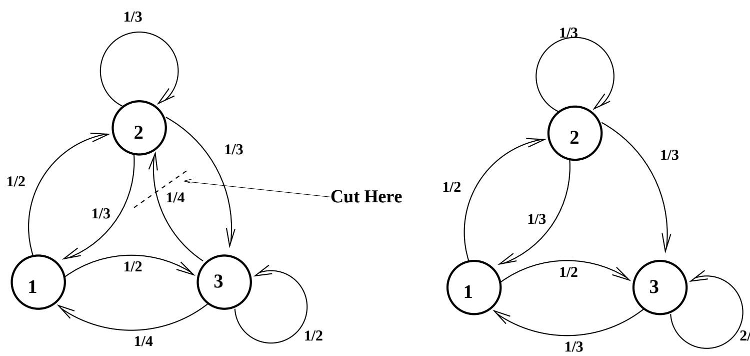
2. For the following Markov process

$$
\mathbf{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix}
$$



find the stationary distribution and the entropy rate $H(\mathcal{X})$.

3. Given a Markov process, one may *cut a branch* and redistribute the conditional probability proportionately among those remaining as in the diagrams below. Does such a process always decrease the entropy rate?



4. In problems 5 and 13 of chapter 4 the rate that the number of possible strings $N(t)$ grows was considered, i.e.,

$$h_T \equiv \lim_{t\to\infty} \frac{1}{t} \log N(t)$$

This quantity is sometimes called the *topological entropy.* Is it always possible to choose the conditional probabilities on the diagram of a Markov process so that the entropy rate is equal to topological entropy, i.e., $H(\mathcal{X}) = h_T$? Explain why or give a counter example.

5. Given five random variables with the following distributions

$$\begin{array}{rcl} P_{X_1} = \left(\frac{2}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right) & & H(X_1) \approx 1.922 \text{ bits} \\ P_{X_2} = \left(\frac{1}{5}, \frac{2}{5}, \frac{1}{5}, \frac{1}{5}\right) & & H(X_2) \approx 1.922 \text{ bits} \\ P_{X_3} = \left(\frac{1}{5}, \frac{1}{5}, \frac{2}{5}, \frac{1}{5}\right) & & H(X_3) \approx 1.922 \text{ bits} \\ P_{X_4} = \left(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{2}{5}\right) & & H(X_4) \approx 1.922 \text{ bits} \\ P_Y = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right) & & H(Y) = 2 \text{ bits,} \end{array}$$

a new variable is defined by $z = x_y$, in other words, the value of Y selects the X whose value is assigned to Z. Let me introduce the notation

$$\begin{aligned} E_Y H(P) &= E_Y H(X_Y) = \sum_{i=1}^{4} \Pr\{Y = i\} H(X_i) \\ H(E_Y P) &= H(Z) \\ &= -\sum_{j=1}^{4} \left(\sum_{i=1}^{4} \Pr\{Y = i\} \Pr\{X_i = j\} \right) \log \left(\sum_{i=1}^{4} \Pr\{Y = i\} \Pr\{X_i = j\} \right) \end{aligned}$$

(a) Calculate $H(Z)$ and compare to $E_Y H(X_Y)$.

(b) Consider the following relations:

$$\begin{aligned} H(Z) &\geq E_Y H(X_Y) \\ H(Z) &> E_Y H(X_Y) \\ H(Z) &\leq H(X_1) + H(X_2) + H(X_3) + H(X_4) \\ H(Z) &\leq H(X_1) + H(X_2) + H(X_3) + H(X_4) + H(Y) \end{aligned}$$

For each relation explain why it is true for all distributions of Y and X_i or find a counter example.

6. Explain the relationships amongst:
    - Kraft's inequality.
    - Instantaneous codes.
    - Uniquely decodable codes.
7. Discuss the strengths and weaknesses of Huffman coding with particular attention to effects of finite block size.