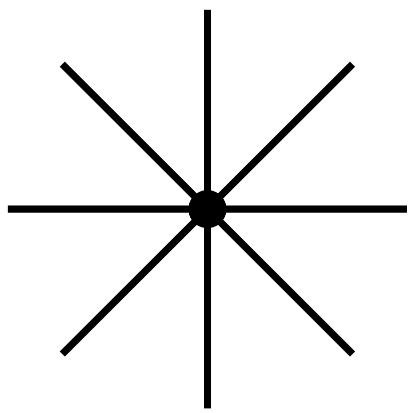
Solutions to Practice Final Examination

(Note: When the solutions refer to particular homework problems, it is speaking with respect to homework problems that were assigned in the year in which this exam was given – not the ones assigned during the Winter 2008-2009 instantiation of EE376A.)

1. **Graph.**
   What is the entropy rate of a random walk on the star graph with a central hub node and n edges:



Solution: Graph.
Note that the stationary distribution is $(\underbrace{1/2n, \ldots, 1/2n}_{n}, n/2n)$ and only the central hub node has non-zero conditional entropy. Hence, the entropy rate is $\frac{1}{2}\log n$.

2. **Optimal code when entropy is infinite.**
   Let X be integer valued with $H(X) = -\sum_{i=1}^{\infty} p_i \log p_i = \infty$. Thus the expected (binary)

description length $L = \sum p_i l_i$ is infinite, even for the Shannon ideal codeword length $l_i^* = \log \frac{1}{p_i}$.

Show nonetheless that $\{l_i^*\}$ is better than $\{l_i\}$ for any other instantaneous code in the sense that $\sum_{i=1}^{\infty} p_i(l_i - l_i^*) \geq 0$, for all $\{l_i\}$ satisfying the Kraft inequality.

Solution: Optimal code when entropy is infinite.

$$\begin{aligned}\sum_{i=1}^{\infty} p_i(l_i - l_i^*) &= \sum p_i \log \frac{p_i}{2^{-l_i}} \\ &= \sum p_i \log \frac{p_i}{r_i} - \log(\sum 2^{-l_i}) \\ &\geq 0,\end{aligned}$$

where $r_i = \frac{2^{-l_i}}{\sum 2^{-l_i}}$ and the last inequality follows from the nonnegativity of $D(p||r)$ and Kraft inequality, $\sum 2^{-l_i} \leq 1$.

3. **Huffman code.**

   (a) Find the binary Huffman code for $p = \{\frac{7}{20}, \frac{4}{20}, \frac{4}{20}, \frac{3}{20}, \frac{2}{20}\}$.

   (b) *Guess* the optimal (minimal expected description length) binary code for the integer valued random variable X, where $\Pr\{X = i\} = p^i q, \quad i = 0, 1, \ldots,$ and $p = .6$.

Solution: Huffman code.

(a) The Huffman tree for this distribution is

Codeword					
00	7	7	8	12	20
10	4	5	7	8	
11	4	4	5		
010	3	4			
011	2				

(b) Since $p^{i-1}q \geq p^i q \geq \sum_{k=i+1}^{\infty} p^k q = p^{i+1}$ for all i, the Huffman tree is

Codeword									
0	q	q	$\ldots$	q	q	$\ldots$	q	q	1
10	pq	pq	$\ldots$	pq	pq	$\ldots$	pq	p	
110	p^2q	p^2q	$\ldots$	p^2q	p^2q	$\ldots$	p^2		
$\vdots$									
$11\cdots10$	p^iq	p^iq	$\ldots$	p^iq	p^i				
$11\cdots110$	$p^{i+1}q$	$p^{i+1}q$	$\ldots$	p^{i+1}					
$\vdots$									

4. **Random walk.**
Let X_i be i.i.d. $\sim X$, where

$$X = \begin{cases} 1 & , \quad p \\ -1 & , \quad 1-p \end{cases}$$

Let $S_n = \sum_{i=1}^{n} X_i$ be the position of the random walk at time n

(a) Is $\{S_n\}_{n=1}^{\infty}$ stationary? Yes or no.

(b) Does $\{S_n\}$ have an entropy rate? If so, what is it?

Solution: Random walk.

(a) No. Whatever the initial distribution is, the distribution at the next step is different from it.

(b) Since X^n and S^n are in 1-1 correspondence, $H(X^n) = H(S^n)$ and

$$\begin{aligned} H(\mathcal{S}) &= \lim_n \frac{1}{n} H(S_1, \ldots, S_n) \\ &= \lim_n \frac{1}{n} H(X_1, \ldots, X_n) \\ &= \lim_n \frac{nH(p)}{n} \\ &= H(p). \end{aligned}$$

5. **Mutual information.**
Consider all probability mass functions on (U, X, Y) of the form $p(u, x, y) = p(u, x)p(y|x)$.

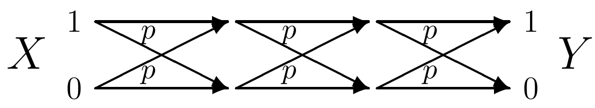
(a) Is $I(U;Y) \geq, =,$ or $\leq I(X;Y)$?

(b) Prove it.

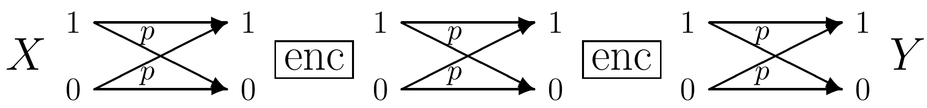
Solution: Mutual information.
Since $U \to X \to Y$, by the data processing inequality (Theorem 2.8.1 of the text), $I(U;Y) \le I(X;Y)$.

6. **Capacity.**
Find the capacities of the following channels.

(a) $Y = X \oplus Z_1 \oplus Z_2 \oplus Z_3$, where $X \in \{0, 1\}$ and Z_1, Z_2, Z_3 are independent Bernoulli(p).

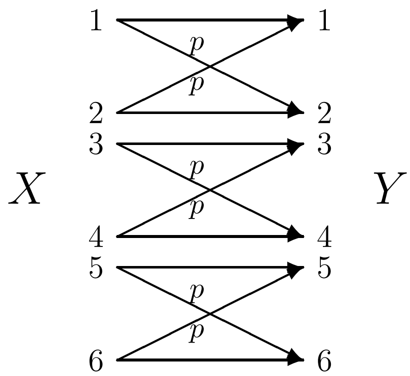


(b) Cascade of 3 BSC(p)'s with encoding and decoding between each.

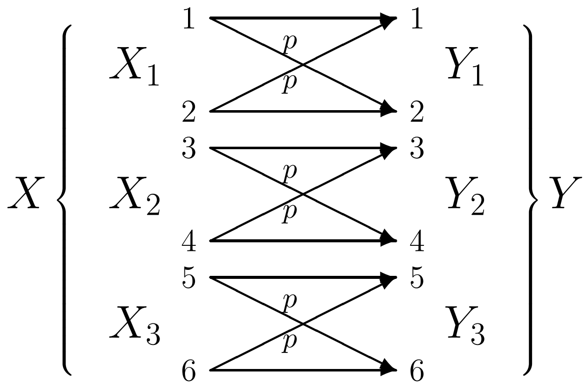


(c) $Y_i = (X_i \oplus Z_{1,i}, X_i \oplus Z_{2,i}, X_i \oplus Z_{3,i}), \qquad i = 1, 2, \ldots,$ where $Z_{k,i} \sim$ Bern(p).

(d) Parallel BSC's, send a symbol on only one.



(e) Parallel BSC's, send a symbol on each.



Solution: Capacity.

(a) As in Question 3(c) of Homework Set #8, this channel is equivalent to a single BSC with the transition probability $3p(1-p)^2+p^3 = 3p^2(1-p)+(1-p)^3$. Hence, the capacity is $1 - H(3p(1-p)^2 + p^3)$ with $p^*(x) = (1/2, 1/2)$.

(b) By the same argument as in Question 3(d) of Homework Set #8, the capacity is the minimum of the capacities of three channels, which is $\min(C_1, C_2, C_3) = 1 - H(p)$ with $p^*(x) = (1/2, 1/2)$.

(c) By symmetry, $p^*(x) = (1/2, 1/2)$. With this input distribution,

$$\begin{aligned}
I|_{p^*(x)} &= I(X;Y_1) + I(X;Y_2|Y_1) + I(X;Y_3|Y_1,Y_2) \\
&= H(Y_1) - H(Y_1|X) + H(Y_2|Y_1) - H(Y_2|X,Y_1) \\
&\quad +H(Y_3|Y_1,Y_2) - H(Y_3|X,Y_1,Y_2) \\
&= H(Y_1) - H(Y_1|X) + H(Y_2|Y_1) - H(Y_2|X) \\
&\quad +H(Y_3|Y_1,Y_2) - H(Y_3|X) \\
&= 1 - 3H(p) + \underbrace{H(Y_2|Y_1)}_{=H(2pq)} + \underbrace{H(Y_3|Y_1,Y_2)}_{=2pq+(p^2+q^2)H(\frac{pq}{p^2+q^2})} \\
&= 1 + 3p\log p + 3q\log q \\
&\quad -2pq\log 2pq - (p^2+q^2)\log(p^2+q^2) \\
&\quad +2pq + (p^2+q^2)\log(p^2+q^2) - pq\log pq - (p^3+q^3)\log(p^3+q^3) \\
&= 1 + 3p^2\log p + 3q^2\log q - (p^3+q^3)\log(p^3+q^3),
\end{aligned}$$

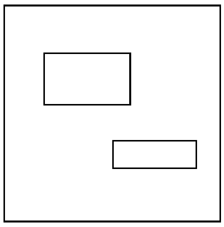
where $q = 1-p$.

(d) By the same argument as in Question 5 of Homework Set #8, the capacity is $\log_2(2^{C_1} + 2^{C_2} + 2^{C_3}) = \log 3 + 1 - H(p)$.

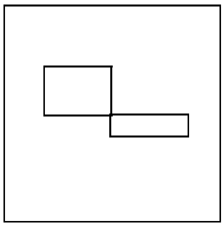
(e) By the same argument as in Quetsion 5 of Homework Set #7, the capacity is $C_1 + C_2 + C_3 = 3(1 - H(p))$.

7. **Kolmogorov complexity.**
Images are displayed on an $n \times n$ screen. What is the Kolmogorov complexity $K(x|n)$ of the following?

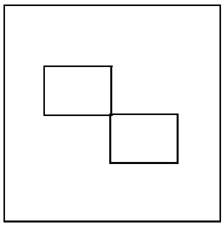
(a)

$x =$ 

(b)

$x =$ 

(c)

$x =$ 

Solution: Kolmogorov complexity.
Note that $K(\text{a single point on the screen}|n) \approx 2\log n + c$.

(a) With the upper-left corner and the lower-right corner, we can describe a rectangle. Hence, for two rectangles, $K(x|n) \approx K(4 \text{ points}|n) \approx 8\log n + c$.

(b) Assuming two rectangles meet at the center of the screen, we need to describe the upper-left corner of the rectangle on the left and the lower-right corner of the other rectangle. Hence, $K(x|n) \approx K(2 \text{ points}|n) \approx 4\log n + c$.

(c) Assuming two rectangles of the same shape meet at the center of the screen, we need to describe the upper-left corner of the rectangle on the left. Hence, $K(x|n) \approx K(1 \text{ point}|n) \approx 2\log n + c$.

8. **Random "20" questions.**
Let X be uniformly distributed over $\{1, 2, \ldots, m\}$. Assume $m = 2^n$. We ask random questions: Is $X \in S_1$? Is $X \in S_2$?...until only one integer remains. All 2^m subsets of $\{1, 2, \ldots, m\}$ are equally likely.

(a) How many deterministic questions are needed to determine X?

(b) Without loss of generality, suppose that $X = 1$ is the random object. What is the probability that object 2 yields the same answers for k questions as object 1?

(c) What is the expected number of objects in $\{2, 3, \ldots, m\}$ that have the same answers to the questions as does the correct object 1?

(d) Suppose we ask $n + \sqrt{n}$ *random* questions. What is the expected number of wrong objects agreeing with the answers?

(e) Use Markov's inequality $\Pr\{X \geq t\mu\} \leq \frac{1}{t}$, to show that the probability of error (one or more wrong objects remaining) goes to zero as $n \longrightarrow \infty$.

Solution: Random "20" questions.

(a) Obviously, Huffman codewords for X are all of length n. Hence, with n deterministic questions, we can identify an object out of 2^n candidates.

(b) Observe that the total number of subsets which include both object 1 and object 2 or neither of them is 2^{m-1}. Hence, the probability that object 2 yields the same answers for k questions as object 1 is $(2^{m-1}/2^m)^k = 2^{-k}$.

More information theoretically, we can view this problem as a channel coding problem through a noiseless channel. Since all subsets are equally likely, the probability the object 1 is in a specific random subset is $1/2$. Hence, the question whether object 1 belongs to the kth subset or not corresponds to the kth bit of the random codeword for object 1, where codewords X^k are Bern(1/2) random k-sequences.

Object	Codeword
1	$0110\ldots1$
2	$0010\ldots0$
$\vdots$	

Now we observe a noiseless output Y^k of X^k and figure out which object was sent. From the same line of reasoning as in the achievability proof of the channel coding theorem, i.e. joint typicality, it is obvious the probability that object 2 has the same codeword as object 1 is 2^{-k}.

(c) Let

$$1_j = \begin{cases} 1, & \text{object } j \text{ yields the same answers for } k \text{ questions as object 1} \\ 0, & \text{otherwise} \end{cases},$$

for $j = 2, \ldots, m$.

Then,

$$\begin{aligned} E(\#\text{ of objects in } \{2,3,\ldots,m\} \text{ with the same answers}) &= E(\sum_{j=2}^{m} 1_j) \\ &= \sum_{j=2}^{m} E(1_j) \\ &= \sum_{j=2}^{m} 2^{-k} \\ &= (m-1)2^{-k} \\ &= (2^n-1)2^{-k}. \end{aligned}$$

(d) Plugging $k = n + \sqrt{n}$ into (c) we have the expected number of $(2^n - 1)2^{-n-\sqrt{n}}$.

(e) Let N by the number of wrong objects remaining. Then, by Markov's inequality

$$\begin{aligned} P(N \geq 1) &\leq EN \\ &= (2^n-1)2^{-n-\sqrt{n}} \\ &\leq 2^{-\sqrt{n}} \\ &\to 0, \end{aligned}$$

where the first equality follows from part (d).