

Midterm Examination

1. (20 points) **Huffman coding.**

(a) (10 points) Find the binary Huffman code for the probability mass function

$$\mathbf{p} = \left(\frac{9}{39}, \frac{8}{39}, \frac{7}{39}, \frac{5}{39}, \frac{4}{39}, \frac{3}{39}, \frac{2}{39}, \frac{1}{39} \right).$$

(b) (10 points) Now find the 3-ary ($D = 3$) Huffman code for \mathbf{p} .

2. (20 points) **Minimum expected description length code.**

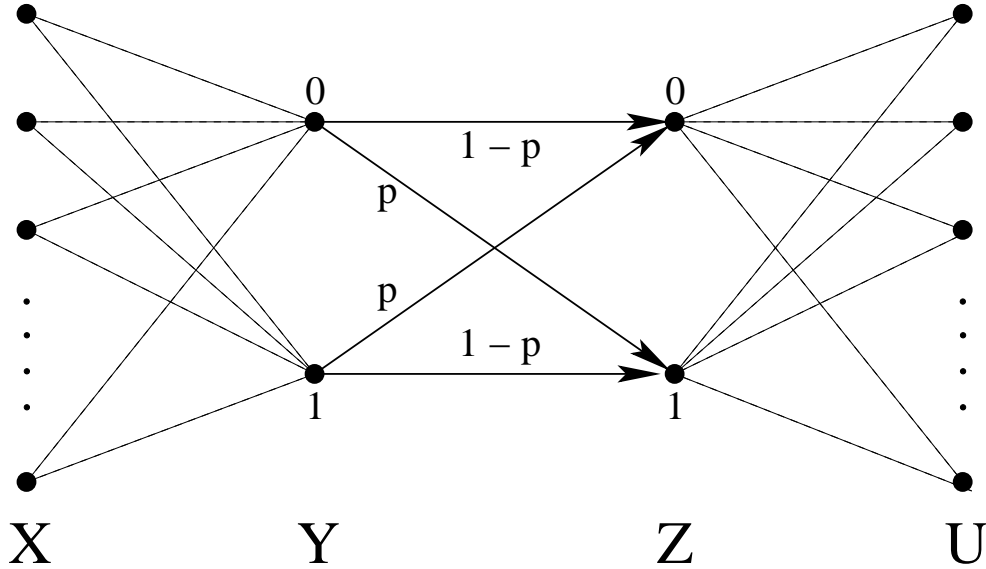
Which of these codes cannot be optimal instantaneous codes for any \mathbf{p} :

(a) (10 points) 00,10,11

(b) (10 points) 00,01,100,101,11

3. (30 points) **Cascade.**

Consider four random variables $X, Y, Z,$ and U . Let the alphabet sizes $|\mathcal{Y}| = |\mathcal{Z}| = 2$, and let the alphabet sizes $|\mathcal{X}|$ and $|\mathcal{U}|$ be arbitrary. Let $p(z|y)$ be as shown in the figure below. Let $p(y|x)$ and $p(u|z)$ be arbitrary.



Here $X \rightarrow Y \rightarrow Z \rightarrow U$ forms a Markov chain.

- (a) (20 points) Show $I(X; U) \leq 1 - H(p)$.
- (b) (10 points) What is a necessary condition on $p(y), y \in \{0, 1\}$, such that the bound $I(X; U) \leq 1 - H(p)$ can be achieved?

4. (30 points) **Repeated histories.**

Let X_1, X_2, \dots be i.i.d. $\sim p(x)$.

- (a) (20 points) Find the limiting behavior of

$$(p^r(X_1, X_2, \dots, X_n))^{\frac{1}{n}}$$

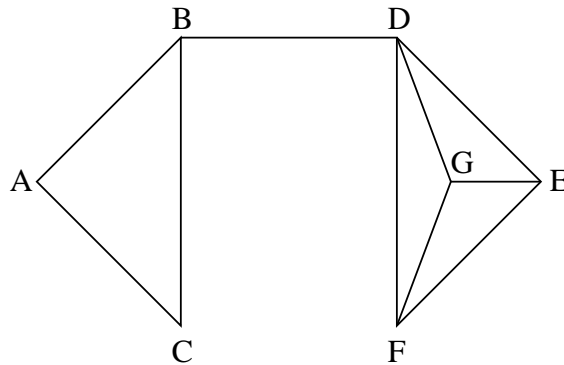
as $n \rightarrow \infty$.

- (b) (10 points) Find the expected value

$$E \frac{1}{p(X_1, X_2, \dots, X_n)}$$

of $\frac{1}{p(X^n)}$.

5. (70 points) **Random walks.**



Assume the stationary distribution on a random walk on this graph.

- (a) (10 points) Find $\Pr(\{X_1 = B\})$.
- (b) (10 points) Find $\Pr(\{X_1X_2X_3X_4X_5 = BABDE\})$.
- (c) (10 points) Find $\Pr(\{X_1X_2X_3X_4X_5 = EDBAB\})$.
- (d) (10 points) Is $\Pr(\{X^n = x_1x_2 \dots x_n\})$ always equal to $\Pr(\{X^n = x_nx_{n-1} \dots x_1\})$?
Yes or No

You are offered 7-for-1 fair odds on gambles on the state of the walk. Assume log-optimal (maximal growth rate) gambling.

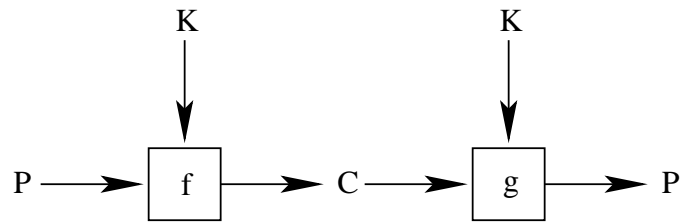
- (e) (10 points) What proportion of your wealth should you gamble on $X_1 = B$?
- (f) (10 points) Assuming a log-optimal betting scheme, in which each bet is based on the past states, and assuming $S_0 = 1$, how much wealth S_5 would you achieve on the sequence

$$X^5 = BABDE \quad ?$$

- (g) (10 points) Again assuming a log-optimal betting scheme, find $E \log S_n$.

6. (20 points) **Shannon cryptography.**

Consider the following relations between random variables P , K , and C :



- i. The plaintext P and the key K are independent of each other.
- ii. The ciphertext is a deterministic function of the plaintext and key: $C = f(P, K)$.
- iii. The plaintext is a deterministic function of the ciphertext and key: $P = g(C, K)$.

Show that the following relationship holds:

$$H(K|C) = H(K) + H(P) - H(C)$$

We conclude that if the ciphertext C gives no information about the key K , i.e., if $H(K|C) = H(K)$, then

$$H(C) = H(P).$$