Indian Institute of Technology Kharagpur
Course: MA41024/MA60020 Information and Coding Theory
Spring Semester 2020-21
Time : 45 minutes
Class Test - II

Declaration:

- **Each question carries 2 marks.**
- **NO query will be entertained during the examination.**
- **There may be multiple options correct for a problem. Full marks is given only when all the correct options are identified.**
- **Once a problem is passed, it will not appear in your screen again and hence if a problem appears in your screen then identify the correct option and then go for the next problem.**

---

1. Let X be a random variable corresponding to a memoryless information source with $H(X) = 2$ bits. Let C be a prefix code corresponding to the source letters of the channel whose information source is X having average codeword length 5 bits. Consider the random variable corresponding to the generation of a source sequences of length 10, denoted by $\overline{X}$. Then the entropy and average codeword length of $\overline{X}$ are —— and —— respectively.

   Ans. 20 and 50

   Watch Lecture 11, 1:30:54

   Tolerance: 0.001

2. Consider a source alphabet $S = \{a, b, c, d, e\}$ with $p(a) = 0.1$, $p(b) = 0.15$, $p(c) = 0.2$, $p(d) = 0.26$, $p(e) = 0.29$. Then the expected codeword-length of the Huffman code for these source letters is —— (up to 1 decimal digits)

   Ans. 2.4

   Codewords: $a = 1111$, $b = 1110$, $c = 110$, $d = 10$, $e = 0$ Tolerance: 0.1

3. Consider the following channel (in the picture) with input random variable X and output random variable Z for $\lambda_1 = 1/2$ and $\lambda_2 = 1/2$. Then the capacity of the channel is —- bits

   Ans. 0

   Tolerance: 0.001

   The capacity of the channel is $1 - H_2(1/2) = 0$ bits since the channel capacity is $1 - H_2(\lambda_1 * \lambda_2)$ where $\lambda_1 * \lambda_2 = \lambda_1(1 - \lambda_2) + (1 - \lambda_1)\lambda_2$

4. The shortest possible codeword length of a source letter for a uniquely decodable code is —— bits when the source letters are with probabilities $1/2, 1/4, 1/8, 1/16, 1/32, 1/32$.

   Ans. 1.94

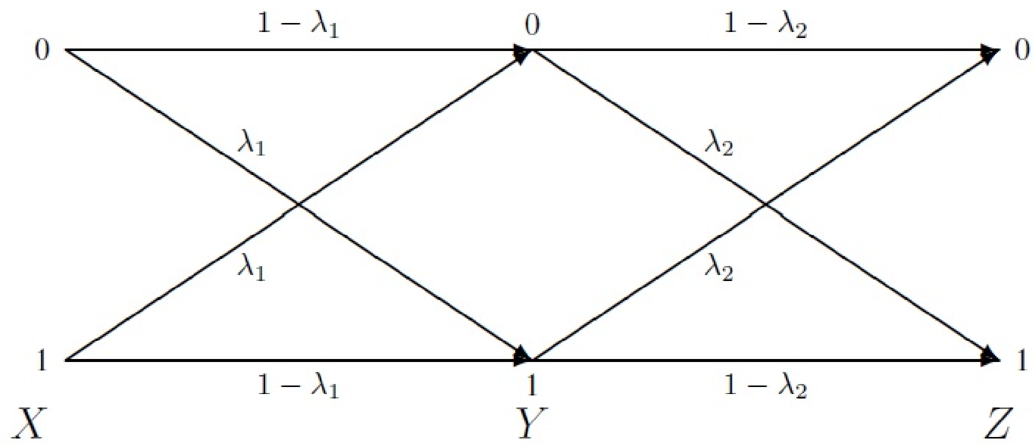

Figure 1: Channel

Tolerance: 0.1

Shannon's Source Coding Theorem tells us that the entropy of the distribution is the lower bound on average code length, in bits per source letter, and the entropy is $31/16 = 1.94$

5. Which one of the following codes is Huffman for some probability distribution of the corresponding source letters?

   (a) $\{00, 01, 10, 110\}$

   (b) $\{0, 10, 11\}$

   (c) $\{01, 10\}$

   Ans. (b)

   The code $\{00, 01, 10, 110\}$ can be shortened to $\{00, 01, 10, 11\}$ without losing its prefix property, and therefore is not optimal, so it cannot be a Huffman code. The code $\{01, 10\}$ can be shortened to $\{0, 1\}$ without losing its prefix property. The code in (b) is optimal corresponding to the probability distribution: $1/2, 1/4, 1/4$.

6. Consider a channel with input random variable X and output random variable Y. Suppose number of source letters is 3 and somehow it is known that $H(X|Y) = 2.1$ bits. Now consider guessing the value of the input from the knowledge of output. Then the the probability of making an error in the guess is at least —-

   Ans.

   From Fano's lemma we know that $P_e \log_2(M-1) + H_2(P_e) \geq H(X|Y)$, where P_e is the probability of error and K is the number of source letters Then $P_e \geq H(X|Y) - 1 = 1.1$

   Tolerance: 0.001

7. Consider a memoryless binary symmetric channel with crossover probability 1/2, and source letters a, b with $p(a) = 1/2$. Then the probability that (a, a) will not be received at the receiver's end after transmitting a sequence (a, a) through the channel is —-

   Ans. 0.75

   Tolerance: 0.001

8. In the process of construction of a prefix code from a given sequence $2, 2, 2, 3, 3$ of ternary codeword lengths, the number of unused terminal nodes in the full tree is —–

   Ans. 25

   Follows from the definition of the full tree

   Tolerance: 0.001

9. Consider a DMC with channel capacity 0.2 bits, and the source is a random variable X with three source letters a, b, c and $p(a) = 1/4, p(b) = 1/4$. Consider sending a source sequences of 5 letters through the the channel with 10 times uses of the channel. Then the of average error probability for a output source sequence $\overline{y}$ will be different from its corresponding source sequence $\overline{x}$ is at least —–

   Ans. 0.1

   Note that $\langle P_e \rangle \log(M-1) \geq H(X) - \frac{N}{L}C - H(\langle P_e \rangle) \geq H(X) - \frac{N}{L}C - 1$, where C is channel capacity, L is the length of the source sequence, N is the number of uses of the channel, $\langle P_e \rangle$ is the average error. Then the desired result follows from the fact that $H(X) = 1.5$ bits, $H(\langle P_e \rangle) \leq 1$, $N = 10$, $L = 5$, $C = 1/5$. (Lecture 15)

   Tolerance: 0.001

10. Let X be the random variable corresponding to the information source of a channel with source letters $a_1, a_2, \ldots, a_K$ with $p(a_1) = 1/8$. Let C be a prefix binary code for the source letters such that the average codeword length satisfies $\overline{n} \leq H(X) + 1$. Then the length of the codeword for a_1 can be at most —- bits.

    Ans. 3

    Watch Lecture 11, 1:15:25

    Tolerance: 0.001