Information Theory
First Midterm
October 9, 2018

1) State Kraft's theorem.

2) State the theorem called Chain rule.

3) What is the largest integer value of ℓ for which a prefix code of 8 codewords with respective lengths $1, 2, 3, 4, 5, 6, 7$, and ℓ over a binary alphabet does not exist? (Notice that we do not assume anything about the relation between the value of ℓ and the other given lengths.)

4) We roll a standard dice (i.e., one with $1, \ldots, 6$ dots on its 6 sides, respectively) until we get a result second time. That is we roll twice if the second roll results in the same number as the first one. We roll three times if the second roll gives a different result than the first one but the third roll has a result that is equal to the result of either the first or the second roll, etc. Let X denote the random variable that is the number of times we have to roll the dice according to the above rules. Give a binary prefix code encoding the outcome of X with optimal average length.

5) Let us have a fair and a possibly biased coin. When tossing the possibly biased coin we get a head with probability p and a tail with probability $1-p$, where $0 < p < 1$. For the fair coin head and tail are equally probable. We toss both coins independently of each other. Let X be the random variable that is equal to 1 if we get a head and to 0 if we get a tail when tossing the fair coin. Let Y be the identically defined random variable for the possibly biased coin and let Z be the modulo 2 sum of X and Y, that is $Z = X + Y \pmod 2$. Compare the values of the conditional entropies $H(Y|Z)$ and $H(Z|Y)$, that is decide which one is larger than the other. When will they be equal?

6) We roll two fair dice (like the one in problem 4 above) independently. Let Z denote the product of the two numbers rolled. For $i = 2, 3$, and 4, we denote by X_i the random variable which is 0 if Z is divisible by i and 1 otherwise. (For example, $X_2 = 0$ if Z is even and $X_2 = 1$ otherwise.) Calculate the entropy values $H(X_2), H(X_3), H(X_4)$ and the mutual information values $I(X_2, X_3)$ and $I(X_2, X_4)$.
(Formulas containing the binary entropy function of precisely given numbers like $h(\frac{9}{16})$ can be considered well determined and need not be calculated numerically. Nevertheless, obvious simplifications are considered important to make, for example, a 0 value should not be left there as a complicated sum.)

Sketches of solutions:

3) By the theorems of McMillan and Kraft such a code does not exist if and only if

$$\frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \frac{1}{2^5} + \frac{1}{2^6} + \frac{1}{2^7} + \frac{1}{2^\ell} > 1.$$

This is the case if and only if $\ell \leq 6$ (we would have equality for $\ell = 7$). So the requested largest number is $\ell = 6$.

4) We first have to find the probabilities with which x takes its possible values. By definition X cannot get a value smaller than 2 and it also cannot be larger than 7 (since we cannot roll seven different numbers with a dice).
$P(X = 2) = \frac{1}{6}$, as this is the probability that the second roll has the same result as the first one.

$P(X=3)=\frac{5}{6}\cdot\frac{1}{3}=\frac{5}{18}$, since the probability that the second roll is different from the first one is $\frac{5}{6}$ and the probability that the third roll is equal to one of the two numbers rolled at the first two rolls is $\frac{1}{3}$.
With similar logic, we obtain that
$P(X=4)=\frac{5}{6}\cdot\frac{2}{3}\cdot\frac{1}{2}=\frac{5}{18}=\frac{90}{324}$,
$P(X=5)=\frac{5}{6}\cdot\frac{2}{3}\cdot\frac{1}{2}\cdot\frac{2}{3}=\frac{5}{27}=\frac{60}{324}$,
$P(X=6)=\frac{5}{6}\cdot\frac{2}{3}\cdot\frac{1}{2}\cdot\frac{1}{3}\cdot\frac{5}{6}=\frac{25}{324}$,
$P(X=7)=\frac{5}{6}\cdot\frac{2}{3}\cdot\frac{1}{2}\cdot\frac{1}{3}\cdot\frac{1}{6}=\frac{5}{324}$.
Thus we need to find an optimal average length code, that is a Huffman code for the distribution $(\frac{1}{6},\frac{5}{18},\frac{5}{18},\frac{5}{27},\frac{25}{324},\frac{5}{324})=(\frac{54}{324},\frac{90}{324},\frac{90}{324},\frac{60}{324},\frac{25}{324},\frac{5}{324})$.
Applying the learnt algorithm to this distribution we obtain that the following code will be optimal:

$$X=2:110, X=3:00, X=4:01, X=5:10, X=6:1110, X=7:1111$$

5) We know that $H(Y|Z)+H(Z)=H(Y,Z)=H(Z|Y)+H(Y)$, so comparing $H(Y|Z)$ and $H(Z|Y)$ can be done via comparing $H(Y)$ and $H(Z)$. We know that $H(Y)=h(p)$. For Z we can calculate that $P(Z=0)=P(X=Y)=\frac{1}{2}\cdot p+\frac{1}{2}\cdot(1-p)=\frac{1}{2}$, therefore $H(Z)=h\left(\frac{1}{2}\right)=1\geq h(p)=H(Y)$. Thus we have $H(Y)\leq H(Z)$ and thus by $H(Y|Z)+H(Z)=H(Z|Y)+H(Y)$ we have $H(Y|Z)\leq H(Z|Y)$.

6) We have $X_2=1$ if both rolls are odd, and this event has probability $\frac{1}{4}$. Thus

$$H(X_2)=h\left(\frac{1}{4}\right).$$

Similarly, we have $X_3=1$ if neither rolls have a result divisible by 3, which has probability $\left(\frac{2}{3}\right)^2=\frac{4}{9}$, so

$$H(X_3)=h\left(\frac{4}{9}\right).$$

We have $X_4=1$ if either both rolls are odd, which is true for 9 of the possible 36 outcomes, or exactly one of them is even but not equal to 4, that can happen in $2\cdot3\cdot2=12$ ways, so altogether, the probability of $X_4=1$ is $\frac{9+12}{36}=\frac{7}{12}$. So

$$H(X_4)=h\left(\frac{7}{12}\right).$$

We observe, that $P(X_2=0|X_3=0)=P(X_2=0|X_3=1)=\frac{1}{2}=P(X_2=0)$, so knowing X_3 does not change the probabilities with which X_2 takes its values. This already shows
$H(X_2|X_3)=P(X_3=1)H(X_2|X_3=1)+P(X_3=0)H(X_2|X_3=0)=H(X_2)$, so

$$I(X_2,X_3)=H(X_2)-H(X_2|X_3)=0,$$

that is, X_2 and X_3 are independent.
If $X_4=0$ then we surely have $X_2=0$, too, since an integer divisible by 4 is also divisible by 2. Thus $P(X_2=0|X_4=0)=1$ and $P(X_2=1|X_4=0)=0$, implying $H(X_2|X_4=0)=0$. We also need the probabilities $P(X_2=0|X_4=1)$ and $P(X_2=1|X_4=1)$. We have already seen above that $X_4=1$ can happen in 21 ways out of the 36 possible outcomes of the two rolls. We have also seen that there are 9 of these 21 cases when neither number is even, that is $X_2=1$. Thus $P(X_2=1|X_4=1)=\frac{9}{21}=\frac{3}{7}$ and $P(X_2=0|X_4=1)=1-\frac{3}{7}=\frac{4}{7}$. So the requested mutual information is

$$I(X_2,X_4)=H(X_2)-H(X_2|X_4)=$$

$$h\left(\frac{1}{4}\right)-P(X_4=0)H(X_2|X_4=0)-P(X_4=1)H(X_2|X_4=1)=$$

$$h\left(\frac{1}{4}\right)-0-\frac{7}{12}h\left(\frac{3}{7}\right)=h\left(\frac{1}{4}\right)-\frac{7}{12}h\left(\frac{3}{7}\right).$$