INFORMATION AND CODING THEORY

Exercise sheet n° 2:

2-1. The random variables X and Y have the joint probability distribution $p(x,y)$:

(a) Calculate $H(X)$ and $H(Y)$.

(b) Calculate $H(X|Y)$ and $H(Y|X)$.

(c) Calculate $H(X,Y)$.

(d) Calculate $H(X{:}Y)$.

		$Y =$	
$P(X,Y)$	0	1	2
0	1/9	1/9	1/9
$X = $ 1	1/9	0	2/9
2	1/9	2/9	0

(e) Draw a Venn diagram for the quantities obtained in (a), (b), (c) and (d).

2-2. An urn contains red, white and black balls with probabilities p_R, p_W and p_B. We consider as a random variable X the color of a ball that is randomly drawn from the urn.

(a) Show that the entropy of drawing k balls with replacement is given by $H(X_1, X_2, ..., X_k) = kH(X)$ where $H(X)$ is the entropy of a single draw.

(b) If the balls are not replaced, and if one does not look at the color of the balls that were taken out, give an intuitive argument why $H(X_i) = H(X)$ where $H(X)$ is the entropy of an arbitrary draw. Without intuition, it might be useful to start solving (c) to (f).

(c) The balls are not replaced. Compare the probability to obtain colors c_1 at the first draw and c_2 at the second draw with the probability to obtain colors c_2 at the first draw and c_1 at the second draw.

(d) The balls are not replaced. Calculate the probability to obtain a red ball at the second draw.

(e) The balls are not replaced. What can one say about the marginal probabilities?

(f) The balls are not replaced. Compare the entropy of the second draw with the entropy of the first draw.

(g) Show that the entropy of drawing k balls without replacements is less than or equal to the entropy of drawing k balls with replacement.

2-3. The random variables X, Y and Z generate a Markov chain $X \to Y \to Z$ if the conditional probability of Z depends only on Y. Thus, $p(x,y,z) = p(x)p(y|x)p(z|y)$.

(a) Show that X and Z are conditionally independent given Y, $p(x,z|y) = p(x|y)p(z|y)$.

(b) Show that $H(X{:}Y) \geq H(X{:}Z)$ by using the result of (a) and the chain rule for the mutual entropy $H(X{:}Y,Z)$.

(c) The random variables X, Y and Z have probability distributions on sets $\mathcal{X}$ of cardinality n, $\mathcal{Y}$ of cardinality k and $\mathcal{Z}$ of cardinality m. Show that if $n > k$ and $m > k$ then $H(X{:}Z) \leq \log_2 k$.

(d) Explain what happens if $k = 1$.

2-4. We consider binary sequences of length n emitted by a source of independent and identically distributed random variables. Every bit is emitted indepedently of the previous bit

according to the binomial distribution where the probability to emit a "1" is given by p. We call *typical sequence* all sequences that contain k ones and $(n-k)$ zeros, with $k \simeq np$ (we round the mean number of ones to the closest whole number.)

(a) What is the probability that the source emits a particular typical sequence? Give an approximation to that probability as a function of the entropy of the source.

(b) What is the number of typical sequences, expressed as a function of the entropy of the source? Compare that number with the absolute number of sequences that can be emitted by the source. What is the approximate probability to emit a sequence that is typical? (Use the Stirling approximation $n! \approx n^n$)

(c) What is the most probable sequence that can be emitted by the source? Is it typical?

2-5. *Definition*: $\rho(x,y)$ is a *distance* if it satisfies the following properties $\forall x,y,z$:

   1. $\rho(x,y) \geq 0$

   2. $\rho(x,y) = \rho(y,x)$

   3. $\rho(x,y) = 0$ if and only if $x=y$

   4. $\rho(x,y) + \rho(y,z) \geq \rho(x,z)$

(a) Let $\rho(X,Y) = H(X|Y)+H(Y|X)$. Verify that $\rho(X,Y) = H(X,Y)-H(X{:}Y) = 2H(X,Y)-H(X)-H(Y)$, and interpret this relation with the help of a Venn diagram.

(b) Prove that $\rho(X,Y)$ satisfies all properties of a distance if the notation $X=Y$ means that there exists a bijection between X and Y.
Tip: To show the last property use the strong subadditivity.

2-6. Assume three binary random variables X, Y and Z. Suppose that these variables are uniformly distributed, that is $H(X) = H(Y) = H(Z) = 1$ bit. If one ignores Z, one can define the mutual information of X and Y, that is $H(X{:}Y)$. Equally, one can define the mutual information of X and Y, *conditional* on Z, i.e. $H(X{:}Y|Z)$. Finally, one defines $H(X{:}Y{:}Z) = H(X{:}Y) - H(X{:}Y|Z)$, that is the mutual information of three random variables. This entropy can be positive, negative or zero.

(a) Give an example for a distribution of X, Y and Z such that $H(X{:}Y{:}Z) > 0$.

(b) Give an exmaple for a distribution of X, Y and Z such that $H(X{:}Y{:}Z) < 0$.

Web site:

http://quic.ulb.ac.be/teaching/