Information and coding theory

May 2017

Second part. Length : 2h

No GSM. No computer. No tablet. No smart devices.
If you use more than one sheet, number your pages.
If you do not answer a question, put the number of the questions and leave a blank space.
Put your name on every page.

Do not do unnecessary calculations. Be precise, complete and clear.

Q1 Entropy and mutual information measures

Let $\mathcal{X}$ and $\mathcal{Y}$ be two discrete random variables defined on a probability space $(\Omega, \mathcal{E}, P)$.

1. State the mathematical definitions of $H(\mathcal{X}), H(\mathcal{Y}), H(\mathcal{X}, \mathcal{Y}), H(\mathcal{Y}|\mathcal{X}), H(\mathcal{X}|\mathcal{Y}), I(\mathcal{X}; \mathcal{Y})$ and justify why each one of these quantities is non-negative.
2. State the numerical relations (inequalities and decompositions) between all these quantities and draw the generic Venn-diagram sumarizing them.
3. Consider the particular case where $\mathcal{X} \perp \mathcal{Y}$: justify which inequalities become equalities and show the corresponding Venn-diagram highlighting the situation.
4. Consider the particular case where $\mathcal{Y} = f(\mathcal{X})$: justify which inequalities become equalities and show the corresponding Venn-diagram highlighting the situation.

Q2 AEP Theorem

1. Define the notion of convergence in probability, and then state and prove the AEP theorem for a stationary and memoryless source (show why and how the law of large numbers can be applied).
2. Define the notion of typical set of sequences $A_\epsilon^{(n)}$, and state its 4 fundamental properties.
3. Discuss why this theorem (and its extensions) are fundamental tools in information theory, both in the context of source coding and in the context of channel coding.

Q3 Information capacity of a discrete channel

1. Define the notions of discrete channel, of discrete causal channel, of discrete causal and memoryless channel, of discrete causal, memoryless and stationary channel.
2. State the mathematical definition of the information capacity per channel usage of the discrete causal, memoryless and stationary channel; explain why this quantity only depends on the channel properties.
3. Present the binary symmetric channel example, and derive its information capacity as a function of the error probability p. Discuss in this perspective the achievable (error free) communication rates for $p = 0$, $p = 1$, $p = 0.5$.
4. **BONUS**: We consider the communication through a cascade of two binary symmetric channels, respectively of error probability p and q.

   What is the highest achievable (error free) communication rate in the two following settings:

   (a) the decoder has only access to the output of the second channel

   (b) the decoder has access to the output of both channels

   Justify your reasoning (pose $\mathcal{X}$ as the input of the first channel, $\mathcal{Y}$ as its output and also the input to the second channel, and $\mathcal{Z}$ as the output of the second channel). Discuss your findings.