

# Information Theory — Tutorial 5

Iain Murray

October 30, 2010

*One of the questions from the class last week was on how to go about solving problems. Remember that looking at a hint or answer for a problem early will mean you tend to learn less. Use the hints and solutions wisely...*

1. **Mutual Information:** Prove that the mutual information  $I(X; Y) = H(X) - H(X|Y)$  is symmetric,  $I(X; Y) = I(Y; X)$ , (hence the term *mutual*) and that  $I(X; Y) \geq 0$ . This is MacKay Ex. 8.4, p140, which includes a big hint.
2. **Communication channels:** Find the mutual information for the binary symmetric (BSC), binary erasure (BEC) and Z channels as a function of the input distribution. Maximize the mutual information to find the capacity of each the channel. This question can be seen as a combination of several exercises in MacKay chapter 9. If this question is not straightforward, work through §9.4–5, pp. 148–151.
3. **Data processing inequality:** MacKay Ex. 8.9, p141. MacKay's answer, which is the standard approach also found in Cover and Thomas, uses technology we haven't covered in the lectures. It is possible to prove the inequality, albeit less neatly, using only methods we have covered. One strategy is given at the bottom of the page.

*Hint for final question: one relatively brute force way of showing the data processing inequality is to average  $D_{\text{KL}}[P(w|d) || P(w|r)]$  under  $P(d,r)$ .*