Lecture 13, Friday w7, 2014-10-31

Things we proved:

- $I(X;Y) \geq 0$

- Entropy is concave in vector of probabilities

- $I(X;Y)$ is concave in input distribution $\mathbf{p}_x$

We noted without proof that $I(X;Y)$ is convex in channel parameters Q (non-examinable proof, Cover & Thomas §2.7). Be able to show with an example that averaging Q for different channels can easily decrease the mutual information.

Non-confusable inputs: For the noisy typewriter model, we proved that using the keys uniformly at random is an optimal input distribution. We then showed that identifying a set of *non-confusable* inputs can also achieve totally reliable communication at the capacity of the channel.

Extended channels: There are no non-confusable inputs for most of the model channels we've considered (BSC, BEC, Z). We'll look at the *Nth extension* of these channels, using the channel N times as one use of a channel with input in $\mathcal{A_X}^N$. If we only look at small extensions, some of the inputs may be less confusable, but it doesn't look like the noisy typewriter. We'll need to look at *large* blocks of symbols. Then we'll identify inputs $\mathbf{x}$ that are *nearly* non-confusable, the sets of typical outputs for the limited set of inputs we choose to use don't overlap.

Checksums: ISBN example. Checksums provide a quick check of data integrity. A single checksum doesn't allow *correction* of errors. Ideas we'll consider later: 1) Checksum a packet of data and throw it away if the check fails. Looks a bit like an erasure channel. 2) Use multiple different checks, and find the decoding that makes them all happy.

Check your progress

We're up to ISBNs in the 'week 7' slides. Mark anything that's unclear or that needs expanding on NB.

Is the optimal input distribution uniform for other channels we've considered, the BSC, BEC and Z channels? For those that it is, can you explain the proof?

Recommended reading

I haven't covered it all yet, but reading Chapter 9 of MacKay would help.