A Short Course in Information Theory — Outline

8 lectures by David MacKay (mackay@mrao.cam.ac.uk)

Mondays at 10.15, TCM seminar room, Mott Building, Cavendish Laboratory

Is it possible to communicate reliably from one point to another if we only have a noisy communication channel? How can the information content of a random variable be measured? This course will discuss the remarkable theorems of Claude Shannon, starting from the source coding theorem, which motivates entropy as the measure of information content, and culminating in the noisy channel coding theorem. Along the way we will study simple examples of codes for data compression and error correction.

This will be an informal course. All are welcome to attend. The level of presentation is intended to be appropriate for graduate students and final year undergraduates.

This outline is approximate and open to change. Feedback is welcome!

1 Perfect communication through a noisy channel

- The problem: reliable transmission of data from one point to another.
- The binary symmetric channel.
- Error correcting strategies: repetition; the (7,4) Hamming code.
- Shannon's Noisy Channel Coding Theorem: *Communication at finite rate can be achieved with arbitrarily small error probability.*
- Definition of Entropy, Conditional entropy, Mutual information.

2 Measuring the information content of a random variable

- Why Entropy?
- Shannon's source coding theorem: *N i.i.d. random variables each with entropy H can be compressed into more than NH bits with negligible loss of information, as $N \to \infty$; conversely if they are compressed into fewer than NH bits there is a dramatic fall-off of information.*

3 Source coding—data compression

- Block codes.
- Uniquely decodeable codes and the Kraft-MacMillan inequality.
- Lower bound on achieveable compression with a source code. Optimal code lengths.
- Huffman codes.

4 Arithmetic coding

- Arithmetic coding.

5 Noisy Channel Coding Theorem, part a

- Review of inference.
- Mutual information.
- The data processing theorem:
  data processing destroys information.
- Definition of channel capacity.

6 Noisy Channel Coding Theorem, part b

- Capacity of binary symmetric channel; of binary erasure channel; of Z channel.
- Random coding.

7 Possible topics for final lecture or two

- Neural networks as information storage devices.
- A variational free energy decoding algorithm.
- Theory of computation. Turing machines. Uncomputability and the halting problem.
- Complexity theory. NP-complete problems.
- Continuous channels and sources.

DJCM February 18, 1995.