

Information Theory

Manan Shah

August 22, 2016

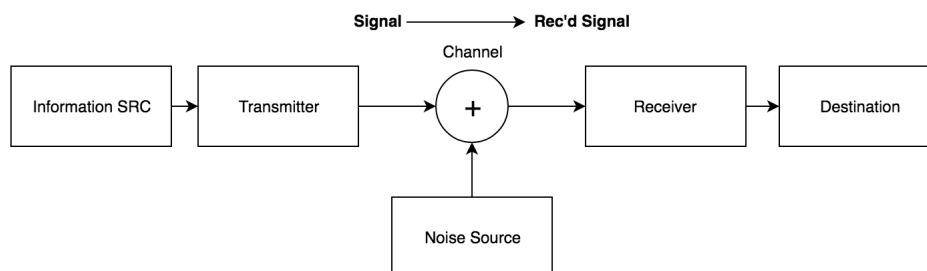
This document contains lecture notes from Harker’s Advanced Topics in Mathematics class in Information Theory, Parts I and II. These notes were taken using TeXShop and $\text{\LaTeX}2\epsilon$ and will be updated for each class.

1 Overview

Information may be conveyed in multiple ways; for example, by a person talking, a video, or pictures. Our goal is to model this kind of information as a system with one output and no input. The information is therefore a box that emits output probabilistically, and we will therefore be spending 1-2 weeks on probability, specifically regarding distribution functions and expectation. This is illustrated in the “information source” box—the job of the information source is to **provide the signal**.

We will next define the compression problem; having obtained numbers as a stream, the problem is to remove the redundancy in a stream. In particular, we will discuss both lossless and lossy compression. One of the most critical ideas in this course is that of entropy, and we will define the concept (in general, a notion of how much information is originating from the source). We will also discuss the sensibility of such a definition and potentially expand upon it. An overarching concept here will be our definition of the performance of a compression algorithm via the source compression theorem and the boundedness of compression algorithms. This is illustrated in the “transmitter” box—the job of the transmitter is to **remove all redundancy**.

We will subsequently discuss the channel circle, which we will model with conditional probability distribution functions (PDFs). The output of this box, the receiver signal, is input to the receiver box, which has the job of **reconstruction of the signal and the noise**; this is known as the coding problem. Much like the compression problem of the source, the channel has the capacity problem (how much information can be transmitted?). And similar to the compression theorem,



we will discuss the channel coding theorem, which will help us determine which parts of a code will require more redundancy for transmission.

We will finally discuss Shannon's theorem, in which he states that a communication with no loss and arbitrary cost (due to a constraint on the amount of bandwidth, power, etc.) is possible if the entropy of the source H is less than the capacity of the channel C . In our course, we will first discuss the pipeline for lossless compression and then move to lossy compression.