Appendix – Some Mathematical Background

In this Appendix, all mathematical terminology used in the main text is defined, and some more background is given to the mathematical formalisms. The areas which are included are probability theory, information theory, algorithmic complexity, and network theory. All terminology which is defined is typeset in bold.

A Probability Theory

An **alphabet** $\mathscr{X}$ is a set of symbols, numeric or symbolic, continuous or discrete, finite or infinite. Symbolic alphabets are discrete; continuous alphabets are numeric and infinite. $|\mathscr{X}|$ denotes the size of set $\mathscr{X}$. An example of a numeric finite, discrete alphabet is the binary set $\{0,1\}$; an example of a symbolic, finite alphabet is a set of Roman letters $\{a,b,c,\ldots,z\}$. An example of a numeric infinite, discrete alphabet is that of the natural numbers $\mathbb{N}$, an example of a continuous alphabet is the set of real numbers $\mathbb{R}$. This book discusses only discrete alphabets.

A **discrete random variable** X is a discrete alphabet $\mathscr{X}$ equipped with a probability distribution $P(X) \equiv \{\Pr(X=x),\ x \in \mathscr{X}\}$. We denote the probabilities $\Pr(X=x)$ by $P(x)$ or sometimes, to avoid confusion, by $P_X(x)$. The **uniform distribution** of a set $\mathscr{X}$ is the distribution $P(x) = 1/|\mathscr{X}|$ for all $x \in \mathscr{X}$. For two discrete random variables, X and Y, the joint probabilities $\Pr(X=x, Y=y)$ on alphabet $\mathscr{X} \times \mathscr{Y}$ are denoted by $P(xy)$ or sometimes, to avoid confusion, by $P_{XY}(xy)$. The joint probability distribution induces a conditional probability distribution $P(x|y) \equiv \Pr(X=x|Y=y)$, which is a probability distribution on $\mathscr{X}$ conditioned on Y taking particular value $Y=y$. Any joint probability $P(xy)$ can be written as

$$P(xy) = P(x|y)P(y)\ . \tag{1}$$