Albert-Ludwigs-Universität Freiburg
Physikalisches Institut

Johannes Berg

Information Theory: From Statistical Physics to Quantitative Biology

2. exercise class – 12. November 2008

Applications of conditional and mutual entropies

1. Entropy of functions of a random variable

Consider a discrete random variable X and a function $g(X)$. Show that the information entropy of $g(X)$ is less than or equal to the information entropy of X. (20 Punkte)
Hint: Use the chain rule to express the joint entropy $H(X, g(X))$ once in terms of $H(g(X)|X)$, and once in terms of $H(X|g(X))$.

2. Entropy of sums

Let X and Y be random variables and consider their sum $Z = X+Y$. Show that $H(Z|X) = H(Y|X)$. Show that for independent X,Y $H(Y) \leq H(Z)$. (Addition of random variables adds uncertainty.) Find an example in which $H(Y) > H(Z)$. (This requires statistically dependent variables.) (20 Punkte)

3. The data processing inequality

a) We define the joint mutual information

$$I(X;Y,Z) = \sum_{x,y,z} p(x,y,z) \ln \frac{p(x,y,z)}{p(x)p(y,z)} \ . \tag{1}$$

Show that $I(X;Y,Z) = I(X;Y|Z) + I(X;Z)$ and similarly for conditioning with respect to Y. (30 Punkte)

b) The random variables X, Y, Z are said to form a *Markov chain* $X \to Y \to Z$ if

$$p(x,y,z) = p(z|y)p(y|x)p(x) \ . \tag{2}$$

Thus the conditional distribution of Z depends only on Y (and not on X). Markov chains play a key role in stochastic dynamics, where they are used to generate time series of causally related variables.

Use the results from a) to show that $I(X;Y) \geq I(X;Z)$. We will use this inequality in the context of gene expression networks modeled by Markov chains. (30 Punkte)