Assignment 2

8th March 2006

1. (2 points) Show that variational free energy

$$F(\theta) = \sum_h Q(h|\theta) \log \frac{Q(h|\theta)}{P(v,h)} \tag{1}$$

   can be expressed as

$$F(\theta) = -\log P(v) + KL(Q(h|\theta)||P(h|v)). \tag{2}$$

2. (a) (4 points) Consider a Boltzmann machine with the energy function given by $E(v,h) = -\sum_{i,j} W_{ij} v_i h_j - \sum_{j,k} U_{jk} h_j h_k$ Suppose we want to perform variational inference in this model. We choose to approximate the posterior distribution for any given data vector v by a distribution of the form $Q(h|\theta) = \prod_i \theta_i^{h_i}(1-\theta_i)^{1-h_i}$. That is we model the distribution over each hidden unit independently of all other hidden units. Write down the expression for the variational free energy using this $Q(h|\theta)$. Simplify the resulting expression as much as possible.

   (b) (3 points) Find the value of θ_i that minimizes $F(\theta)$ from 2(a), assuming that all θ_j for $i \neq j$ are fixed.

   (c) (1 point) Would you use variational inference in this model if $U_{jk} = 0$ for all j and k? Briefly justify your answer.