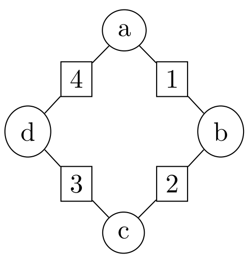
# Variational Distribution: Exam Example

*Amos Storkey*
`a.storkey@ed.ac.uk`

March 23, 2016

1. In this question we consider the use of a factor graph for inference Having absorbed the conditioned information, we are left with the following Factor Graph of binary $\{0, 1\}$ variables.




The graph has factor potentials $\phi_1(a, b, c, d) = \phi_1(a, b) = e^{ab}$, $\phi_2(a, b, c, d) = \phi_2(b, c) = e^{bc}$, $\phi_3(a, b, c, d) = \phi_3(c, d) = e^{cd}$ and $\phi_4(a, b, c, d) = \phi_4(d, a) = e^{-da}$, where $a, b, c$ take values 0 or 1 corresponding to each node.

(a) **Find an expression for $P(a, b)$ (up to a normalisation constant) using variable elimination. Show your working.**

**Answer:** To do this we need to eliminate nodes $c$ and $d$. Eliminating $c$ is easiest. There we have

$$P(a, b, d) = \sum_c P(a, b, c, d) = (1/Z)e^{ab}e^{-da} \sum_c e^{bc+cd} = e^{ab}e^{-da}[1 + e^{b+d}].$$

Now we eliminate $d$:

$$P(a, b) = \sum_d P(a, b, d) = (1/Z)e^{ab} \sum_d [e^{-da} + e^{b+d-da}] = (1/Z)e^{ab}[1 + e^b + e^{-a} + e^{b+1-a}]$$

So the distribution for $P(a, b)$ is

$$P(a, b) = (1/Z)e^{ab}[1 + e^b + e^{-a} + e^{b-a+1}]$$

(b) **Consider using a variational posterior distribution of the form $Q(a, b, c, d) = Q(a)Q(b)Q(c)Q(d)$. Let $q_a = Q(a = 1)$. Likewise $q_b = Q(b = 1)$, $q_c = Q(c = 1)$, $q_d + Q(d = 1)$. The Variational**

**Free Energy is written as**

$$F = -\sum_{a,b,c,d} Q(a,b,c,d)\left[\log\left(\prod_i \phi_i(a,b,c,d)\right) - \log Q(a,b,c,d)\right]$$

**Write the variational free energy for this system in terms of the parameters $q_a$, $q_b$ and $q_c$, $q_d$.**

**Answer:** Plugging in and using rules for logarithms of products, we have the free energy F is

$$\begin{aligned} F &= -q_a q_b - q_b q_c - q_c q_d + q_d q_a + q_a \log q_a + (1-q_a)\log(1-q_a) \\ &+ q_b \log q_b + (1-q_b)\log(1-q_b) + q_c \log q_c + (1-q_c)\log(1-q_c) + q_d \log q_d + (1-q_d)\log(1-q_d). \end{aligned}$$

Note we have used that $Q(a=0) = 1-q_a$, and $E_Q[ab] = q_a q_b$ etc.

(c) **Work out and write out one expression relating $q_a$, $q_b$ and $q_d$ that must hold at the minimum of the variational free energy.**

**Answer:** We need to relate $q_a$, to $q_b$ and $q_d$. Note from the graph that these variables are the market blanket of node $a$, so we can take derivative of $F$ w.r.t. $q_a$ to get an expression for these. At an optimum these derivatives must be zero, so we get (ignoring terms not containing $q_a$):

$$\frac{\partial F}{\partial q_a} = \frac{\partial}{\partial q_a}[-q_a q_b + q_d q_a + q_a \log q_a + (1-q_a)\log(1-q_a)] = \log q_a - \log(1-q_a) - q_b + q_d = 0$$

leading to $\log q_a - \log(1-q_a) = q_b - q_d$. This expression must hold at the optimum. Aside: In general we get 4 of these equations, and we solve this set of self consistent equations (usually iteratively) in order to find the optimum.

(d) **Using standard belief propagation, assume the initial clockwise message arriving at node 'a' is $(r, 1-r)$. What is the message arriving at node 'b' during clockwise belief propagation.**

**Answer:** This is a very simple case, where all nodes connect to only 2 nodes, and so no product is needed. The message is simply

$$\begin{pmatrix} e^{0\times 0} & e^{0\times 1} \\ e^{1\times 0} & e^{1\times 1} \end{pmatrix}\begin{pmatrix} r \\ (1-r) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & e \end{pmatrix}\begin{pmatrix} r \\ (1-r) \end{pmatrix} = \begin{pmatrix} 1 \\ r + (1-r)e \end{pmatrix}$$