BP

July 21, 2013

1 Model

The atomic unit of the factor graph is shown in Figure 1.

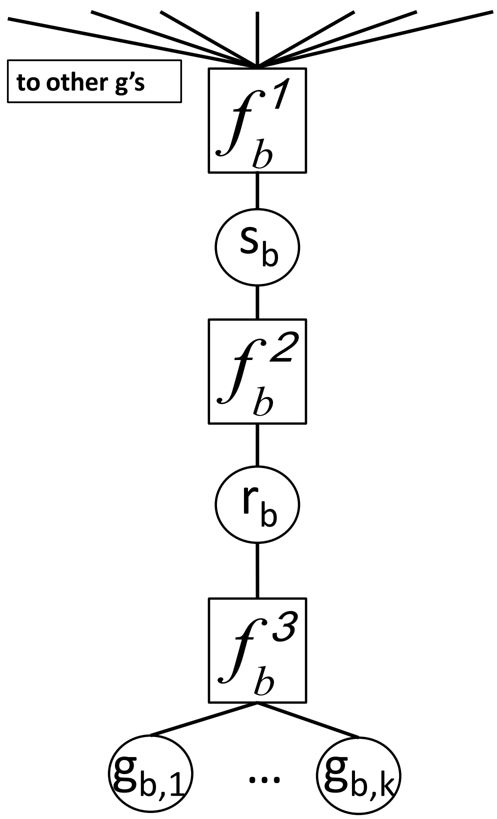

Figure 1: Factor graph

We require efficient computation of the messages.

2 Top-down messages

2.1 $\mu_{f_b^1 \to s_b}$

The general form of the message is given by

$$\mu_{f_b^1 \to s_b}(s_b) = \sum_{\vec{g}} f_b^1(s_b, \vec{g}) \prod_{g_{b',k}} \mu_{g_{b',k} \to f_b^1}(g_{b',k}) \tag{1}$$

We divide the cases into $s_b = 0$, and $s_b = 1$.

$$\begin{aligned} \mu_{f_b^1 \to s_b}(s_b = 0) &= \sum_{\vec{g}} f_b^1(s_b = 0, \vec{g}) \prod_{g_{b',k}} \mu_{g_{b',k} \to f_b^1}(g_{b',k}) &(2)\\ &= (1-\epsilon) \prod_{g_{b',k}} \mu_{g_{b',k} \to f_b^1}(g_{b',k} = 0) &(3) \end{aligned}$$

$$\begin{aligned} \mu_{f_b^1 \to s_b}(s_b = 1) &= \sum_{\vec{g}} f_b^1(s_b = 1, \vec{g}) \prod_{g_{b',k}} \mu_{g_{b',k} \to f_b^1}(g_{b',k}) &(4)\\ &= \epsilon \times \prod_{g_{b',k}} \mu_{g_{b',k} \to f_b^1}(g_{b',k} = 0) &(5)\\ &\quad + \sum_{\vec{g}} \prod_{g_{b',k}} \mu_{g_{b',k} \to f_b^1}(g_{b',k}) - \sum_{\vec{g}} \prod_{g_{b',k}} \mu_{g_{b',k} \to f_b^1}(g_{b',k} = 0) \\ &= \epsilon \times \prod_{g_{b',k}} \mu_{g_{b',k} \to f_b^1}(g_{b',k} = 0) &(6)\\ &\quad + 1 - \prod_{g_{b',k}} \sum_{g_{b',k}} \mu_{g_{b',k} \to f_b^1}(g_{b',k} = 0) \end{aligned}$$

where we have made use of the property that messages from a node/factor sum to 1

2.2 $\mu_{s_b \to f_b^2}$

if s_b has not been observed, this message is given by

$$\mu_{s_b \to f_b^2}(s_b) = \mu_{f_b^1 \to s_b}(s_b) \tag{7}$$

otherwise, it is a constant message that indicates its observed state.

2.3 $\mu_{f_b^2 \to r_b}$

$$\begin{aligned} \mu_{f_b^2 \to r_b}(r_b) &= \sum_{s_b} f_b^2(s_b, r_b) \mu_{s_b \to f_b^2}(s_b) &(8)\\ &= \sum_{s_b} P(r_b \mid s_b) \mu_{s_b \to f_b^2}(s_b) &(9) \end{aligned}$$

s_b only has two values, so this is easy to compute.

2.4 $\mu_{r_b \to f_b^3}$

$$\mu_{r_b \to f_b^3}(r_b) = \mu_{f_b^2 \to r_b}(r_b) \tag{10}$$

2.5 $\mu_{f_b^3 \to g_{b,k}}$

$$\begin{aligned}
\mu_{f_b^3 \to g_{b,k}}(g_{b,k}) &= \sum_{r_b} \sum_{k' \neq k} f_b^3(r_b, \vec{g}_b) \mu_{r_b \to f_b^2}(r_b) \prod_{k' \neq k} \mu_{g_{b,k} \to f_b^3}(g_{b,k'}) && (11)\\
&= \sum_{r_b} \mu_{r_b \to f_b^2}(r_b) \sum_{k' \neq k} P(\vec{g}_b \mid r_b) \prod_{k' \neq k} \mu_{g_{b,k} \to f_b^3}(g_{b,k'}) && (12)\\
&= \sum_{r_b} \mu_{r_b \to f_b^2}(r_b) \sum_{k' \neq k} P(\vec{g}_{b,k} \mid r_b) \prod_{k' \neq k} P(\vec{g}_{b,k'} \mid r_b) \mu_{g_{b,k} \to f_b^3}(g_{b,k'}) && (13)\\
&= \sum_{r_b} \mu_{r_b \to f_b^2}(r_b) P(\vec{g}_{b,k} \mid r_b) \prod_{k' \neq k} \sum_{g_{b,k'}} P(\vec{g}_{b,k'} \mid r_b) \mu_{g_{b,k} \to f_b^3}(g_{b,k'}) && (14)
\end{aligned}$$

The number of values r_b can take should be small (no more than, say 10? 20?), so this should be possible to compute.

3 Bottom-up messages

3.1 $\mu_{f_b^1 \to g_{b',k}}$

Recall that f_b^1 connects to other bricks. We denote one such brick by b'.

Because f_b^1 only cares about whether $g_{b',k}$ points to b, we reduce the domain of $g_{b',k}$ to $0, 1$ which means "doesn't point to b" and "points to b", respectively. The messages can be "expanded" by $g_{b',k}$ to its original domain when $g_{b',k}$ needs to send a message to f_b^3.

The general form of this message is

$$\mu_{f_b^1 \to g_{b',k}}(g_{b,k'}) = \sum_{s_b} \sum_{g \neq g_{b',k}} f_b^1(s_b, \mathbf{g}) \mu_{s_b \to f_b^1}(s_b) \prod_{g \neq g_{b',k}} \mu_{g \to f_b^1}(g). \tag{15}$$

We divide the cases into $g_{b',k} = 0$, and $g_{b',k} = 1$

$$
\begin{aligned}
\mu_{f_b^1 \to g_{b',k}}(g_{b',k}=0) &= \mu_{s_b \to f_b^1}(s_b=0) \times (1-\epsilon) \times \prod_{g \neq g_{b',k}} \mu_{g \to f_b^1}(g=0) && (16)\\
&\quad + \mu_{s_b \to f_b^1}(s_b=1) \times \epsilon \times \prod_{g \neq g_{b',k}} \mu_{g \to f_b^1}(g=0) \\
&\quad + \mu_{s_b \to f_b^1}(s_b=1)\Big(\prod_{g \neq g_{b',k}} \sum_{g} \mu_{g \to f_b^1}(g) - \prod_{g \neq g_{b',k}} \mu_{g \to f_b^1}(g=0)\Big) \\
&= \mu_{s_b \to f_b^1}(s_b=0) \times (1-\epsilon) \times \prod_{g \neq g_{b',k}} \mu_{g \to f_b^1}(g=0) && (17)\\
&\quad + \epsilon \times \mu_{s_b \to f_b^1}(s_b=1) \prod_{g \neq g_{b',k}} \mu_{g \to f_b^1}(g=0) \\
&\quad + \mu_{s_b \to f_b^1}(s_b=1)\Big(1 - \prod_{g \neq g_{b',k}} \mu_{g \to f_b^1}(g=0)\Big) \\
&= \mu_{s_b \to f_b^1}(s_b=0) \times (1-\epsilon) \times \prod_{g \neq g_{b',k}} \mu_{g \to f_b^1}(g=0) \\
&\quad + \mu_{s_b \to f_b^1}(s_b=1)\Big(\prod_{g \neq g_{b',k}} \mu_{g \to f_b^1}(g=0)(\epsilon-1)+1\Big) && (18)\\
&= \mu_{s_b \to f_b^1}(s_b=0) \times (1-\epsilon) \times \frac{\prod_g \mu_{g \to f_b^1}(g=0)}{\mu^0_{g_{b',k} \to f_b^1}(g=0)} && (19)\\
&\quad + \mu_{s_b \to f_b^1}(s_b=1)\Big(\frac{\prod_g \mu_{g \to f_b^1}(g=0)}{\mu^0_{g_{b',k} \to f_b^1}(g=0)}(\epsilon-1)+1\Big) \\
&= \mu_{s_b \to f_b^1}(s_b=0) \times \frac{\mu_{s_b \to f_b^1}(s_b=0)}{\mu^0_{g_{b',k} \to f_b^1}(g=0)} && (20)\\
&\quad + \mu_{s_b \to f_b^1}(s_b=1)\Big(\frac{\mu_{s_b \to f_b^1}(s_b=0)}{\mu^0_{g_{b',k} \to f_b^1}(g=0)}(\epsilon-1)+1\Big) \\
&= \frac{\mu_{s_b \to f_b^1}(s_b=0)^2}{\mu^0_{g_{b',k} \to f_b^1}(g=0)} && (21)\\
&\quad + \frac{\mu_{s_b \to f_b^1}(s_b=1)\mu_{s_b \to f_b^1}(s_b=0)(\epsilon-1) + \mu^0_{g_{b',k} \to f_b^1}(g=0)}{\mu^0_{g_{b',k} \to f_b^1}(g=0)} \\
&= \frac{\mu_{s_b \to f_b^1}(s_b=0)^2 + \mu_{s_b \to f_b^1}(s_b=1)\mu_{s_b \to f_b^1}(s_b=0)(\epsilon-1) + \mu^0_{g_{b',k} \to f_b^1}(g=0)}{\mu^0_{g_{b',k} \to f_b^1}(g=0)} && (22)
\end{aligned}
$$

where $\mu^0_{g_{b',k} \to f_b^1}$ means the old message sent by $g_{b',k}$, and we have made use of the property that messages from a node/factor sum to 1. Note that we can compute this message without having to explicitly examine all the other $g'_{b,k}s$; we re-used the computation from computing $\mu_{s_b \to f_b^1}(s_b = 0)$.

$$\mu_{f_b^1 \to g_{b',k}}(g_{b,k'} = 1) = \mu_{s_b \to f_b^1}(s_b = 1) \prod_{g \neq g_{b',k}} \sum_g \mu_{g \to f_b^1}(g) \tag{23}$$

$$= \mu_{s_b \to f_b^1}(s_b = 1) \tag{24}$$

3.2 $\mu_{g_{b,k} \to f_b^3}$

$$\mu_{g_{b,k} \to f_b^3}(g_{b,k}) = \prod_{b' \neq b} \mu_{f_{b'}^3 \to g_{b,k}}(g_{b,k}) \tag{25}$$

Note that $\mu_{f_{b'}^3 \to g_{b,k}}(g_{b,k})$ might need to be "re-expanded" to the proper domain of $g_{b,k}$.

3.3 $\mu_{f_b^3 \to r_b}$

$$\mu_{f_b^3 \to r_b}(r_b) = \sum_{\vec{g}} f_b^3(\vec{g}, r_b) \prod_k \mu_{g_{b,k} \to f_b^3}(g_{b,k}) \tag{26}$$

$$= \sum_{\vec{g}} \prod_k P(g_{b,k} \mid r_b) \prod_k \mu_{g_{b,k} \to f_b^3}(g_{b,k}) \tag{27}$$

$$= \prod_k \sum_{g_{b,k}} P(g_{b,k} \mid r_b) \mu_{g_{b,k} \to f_b^3}(g_{b,k}) \tag{28}$$

3.4 $\mu_{r_b \to f_b^2}$

$$\mu_{r_b \to f_b^2}(r_b) = \mu_{f_b^3 \to r_b}(r_b) \tag{29}$$

3.5 $\mu_{f_b^2 \to s_b}$

$$\mu_{f_b^2 \to s_b}(s_b) = \sum_{r_b} f_b^2(r_b, s_b) \mu_{r_b \to f_b^2}(r_b) \tag{30}$$

$$= \sum_{r_b} P(r_b \mid s_b) \mu_{r_b \to f_b^2}(r_b) \tag{31}$$

These messages should be trivial to compute, since the set of values r_b is small.

3.6 $\mu_{s_b \to f_b^1}$

$$\mu_{s_b \to f_b^(} s_b) = \mu_{f_b^2 \to s_b}(s_b) \tag{32}$$