BP

July 31, 2013

1 Model

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

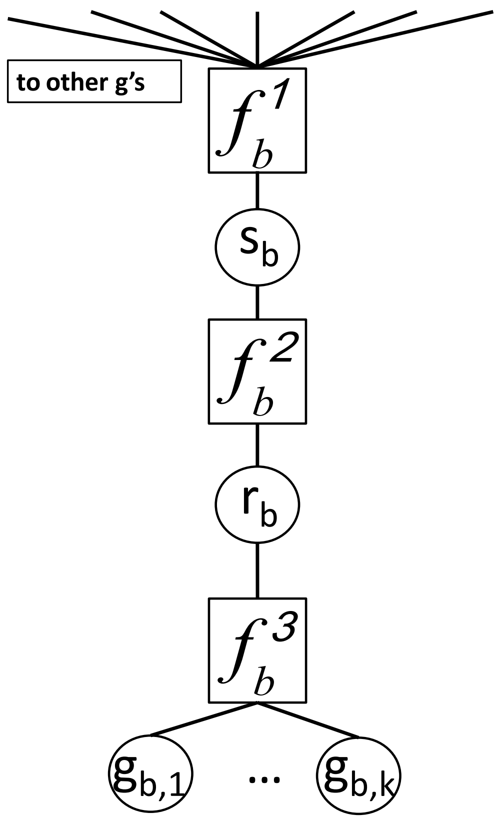

Figure 1: Factor graph

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 $s_b = 1$.

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

$$\begin{aligned}
\mu_{f_b^1 \to s_b}(s_b = 1) &= \sum_{\vec{g}} f_b^1(s_b = 1, \vec{g}) \prod_{g_{b',k}} \mu_{g_{b',k} \to f_b^1}(g_{b',k}) && (4)\\
&= \epsilon \times \prod_{g_{b',k}} \mu_{g_{b',k} \to f_b^1}(g_{b',k} = 0) && (5)\\
&\quad + \sum_{\vec{g}} \prod_{g_{b',k}} \mu_{g_{b',k} \to f_b^1}(g_{b',k}) - \prod_{g_{b',k}} \mu_{g_{b',k} \to f_b^1}(g_{b',k} = 0) \\
&= \epsilon \times \prod_{g_{b',k}} \mu_{g_{b',k} \to f_b^1}(g_{b',k} = 0) && (6)\\
&\quad + 1 - \prod_{g_{b',k}} \mu_{g_{b',k} \to f_b^1}(g_{b',k} = 0) \\
&= \prod_{g_{b',k}} \mu_{g_{b',k} \to f_b^1}(g_{b',k} = 0)(\epsilon - 1) + 1 && (7)\\
&= 1 - (1 - \epsilon) \prod_{g_{b',k}} \mu_{g_{b',k} \to f_b^1}(g_{b',k} = 0) && (8)\\
&= 1 - \mu_{f_b^1 \to s_b}(s_b = 0) && (9)
\end{aligned}$$

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

$$\mu_{s_b \to f_b^2}(s_b) = \mu_{f_b^1 \to s_b}(s_b) \tag{10}$$

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

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

$$
\begin{aligned}
\mu_{f_b^2 \to r_b}(r_b) &= \sum_{s_b} f_b^2(s_b, r_b)\mu_{s_b \to f_b^2}(s_b) && (11)\\
&= \sum_{s_b} P(r_b \mid s_b)\mu_{s_b \to f_b^2}(s_b) && (12)
\end{aligned}
$$

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

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

$$
\mu_{r_b \to f_b^3}(r_b) = \mu_{f_b^2 \to r_b}(r_b) \tag{13}
$$

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

$$
\begin{aligned}
\mu_{f_b^3 \to g_{b,k}}(g_{b,k}) &= \sum_{r_b} \sum_{k' \neq k} f_b^3(r_b, \vec{g}_b)\mu_{r_b \to f_b^3}(r_b) \prod_{k' \neq k} \mu_{g_{b,k} \to f_b^3}(g_{b,k'}) && (14)\\
&= \sum_{r_b} \mu_{r_b \to f_b^3}(r_b) \sum_{k' \neq k} P(\vec{g}_b \mid r_b) \prod_{k' \neq k} \mu_{g_{b,k} \to f_b^3}(g_{b,k'}) && (15)\\
&= \sum_{r_b} \mu_{r_b \to f_b^3}(r_b) \sum_{k' \neq k} P(\vec{g}_{b,k} \mid r_b) \prod_{k' \neq k} P(\vec{g}_{b,k'} \mid r_b)\mu_{g_{b,k} \to f_b^3}(g_{b,k'}) && (16)\\
&= \sum_{r_b} \mu_{r_b \to f_b^3}(r_b) P(\vec{g}_{b,k} \mid r_b) \prod_{k' \neq k} \sum_{g_{b,k'}} P(\vec{g}_{b,k'} \mid r_b)\mu_{g_{b,k} \to f_b^3}(g_{b,k'}) && (17)
\end{aligned}
$$

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

2.6 $\mu_{g_{b,k} \to f_b^1}$

$$
\mu_{g_{b,k} \to f_b^1}(g_{b,k}) = \mu_{f_b^3 \to g_{b,k}}(g_{b,k}) \prod_{b' \neq b} \mu_{f_{b'}^1 \to g_{b,k}}(g_{b,k}) \tag{18}
$$

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

$$
\mu_{f_b^1 \to g_{b',k}}(g_{b,k'}) = \sum_{s_b} \sum_{g \neq g_{b',k}} f_b^1(s_b, \mathbf{g})\mu_{s_b \to f_b^1}(s_b) \prod_{g \neq g_{b',k}} \mu_{g \to f_b^1}(g). \tag{19}
$$

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

$$
\begin{aligned}
\mu_{f_b^1 \to g_{b',k}}(g_{b',k} = 0) &= \mu_{s_b \to f_b^1}(s_b = 0) \times (1-\epsilon) \times \prod_{g \neq g_{b',k}} \mu_{g \to f_b^1}(g = 0) \qquad (20)\\
&+ \mu_{s_b \to f_b^1}(s_b = 1) \times \epsilon \times \prod_{g \neq g_{b',k}} \mu_{g \to f_b^1}(g = 0)\\
&+ \mu_{s_b \to f_b^1}(s_b = 1)\Big(\prod_{g \neq g_{b',k}} \sum_g \mu_{g \to f_b^1}(g) - \prod_{g \neq g_{b',k}} \mu_{g \to f_b^1}(g = 0)\Big)\\
&= \mu_{s_b \to f_b^1}(s_b = 0) \times (1-\epsilon) \times \prod_{g \neq g_{b',k}} \mu_{g \to f_b^1}(g = 0) \qquad (21)\\
&+ \epsilon \times \mu_{s_b \to f_b^1}(s_b = 1) \prod_{g \neq g_{b',k}} \mu_{g \to f_b^1}(g = 0)\\
&+ \mu_{s_b \to f_b^1}(s_b = 1)\Big(1 - \prod_{g \neq g_{b',k}} \mu_{g \to f_b^1}(g = 0)\Big)\\
&= \mu_{s_b \to f_b^1}(s_b = 0) \times (1-\epsilon) \times \prod_{g \neq g_{b',k}} \mu_{g \to f_b^1}(g = 0)\\
&+ \mu_{s_b \to f_b^1}(s_b = 1)\Big(\prod_{g \neq g_{b',k}} \mu_{g \to f_b^1}(g = 0)(\epsilon - 1) + 1\Big) \qquad (22)\\
&= \mu_{s_b \to f_b^1}(s_b = 0) \times (1-\epsilon) \times \frac{\prod_g \mu_{g \to f_b^1}(g = 0)}{\mu^0_{g_{b',k} \to f_b^1}(g = 0)} \qquad (23)\\
&+ \mu_{s_b \to f_b^1}(s_b = 1)\Big(\frac{\prod_g \mu_{g \to f_b^1}(g = 0)}{\mu^0_{g_{b',k} \to f_b^1}(g = 0)}(\epsilon - 1) + 1\Big)\\
&= \mu_{s_b \to f_b^1}(s_b = 0) \times \frac{\mu_{f_b^1 \to s_b}(s_b = 0)}{\mu^0_{g_{b',k} \to f_b^1}(g = 0)} \qquad (24)\\
&+ \mu_{s_b \to f_b^1}(s_b = 1) \times \Big(1 - \frac{\mu_{f_b^1 \to s_b}(s_b = 0)}{\mu^0_{g_{b',k} \to f_b^1}(g = 0)}\Big)
\end{aligned}
$$

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 $\mu_{s_b \to f_b^1}(s_b = 0)$.

$$
\begin{aligned}
\mu_{f_b^1 \to g_{b',k}}(g_{b,k'} = 1) &= \mu_{s_b \to f_b^1}(s_b = 1) \prod_{g \neq g_{b',k}} \sum_g \mu_{g \to f_b^1}(g) \qquad (25)\\
&= \mu_{s_b \to f_b^1}(s_b = 1) \qquad (26)
\end{aligned}
$$

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

$$\mu_{g_{b,k}\to f_b^3}(g_{b,k}) = \prod_{b'\in V(b)} \mu_{f_{b'}^3\to g_{b,k}}(g_{b,k}) \tag{27}$$

wher $V(b)$ is the after set of brick b. Note that $\mu_{f_{b'}^3\to g_{b,k}}(g_{b,k})$ needs to be "expanded" to the proper domain of $g_{b,k}$

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

$$\begin{aligned}
\mu_{f_b^3\to r_b}(r_b) &= \sum_{\vec{g}} f_b^3(\vec{g}, r_b) \prod_k \mu_{g_{b,k}\to f_b^3}(g_{b,k}) && (28)\\
&= \sum_{\vec{g}} \prod_k P(g_{b,k} \mid r_b) \prod_k \mu_{g_{b,k}\to f_b^3}(g_{b,k}) && (29)\\
&= \prod_k \sum_{g_{b,k}} P(g_{b,k} \mid r_b) \mu_{g_{b,k}\to f_b^3}(g_{b,k}) && (30)
\end{aligned}$$

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

$$\mu_{r_b\to f_b^2}(r_b) = \mu_{f_b^3\to r_b}(r_b) \tag{31}$$

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

$$\begin{aligned}
\mu_{f_b^2\to s_b}(s_b) &= \sum_{r_b} f_b^2(r_b, s_b)\mu_{r_b\to f_b^2}(r_b) && (32)\\
&= \sum_{r_b} P(r_b \mid s_b)\mu_{r_b\to f_b^2}(r_b) && (33)
\end{aligned}$$

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

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

$$\mu_{s_b\to f_b^(} s_b) = \mu_{f_b^2\to s_b}(s_b) \tag{34}$$