Belief Propagation

Srikumar Ramalingam

School of Computing

University of Utah

Reference

- Most or all slides are adapted from the following paper:

Jonathan S. Yedidia, Message-passing Algorithms for Inference and Optimization: “Belief Propagation” and “Divide and Concur”

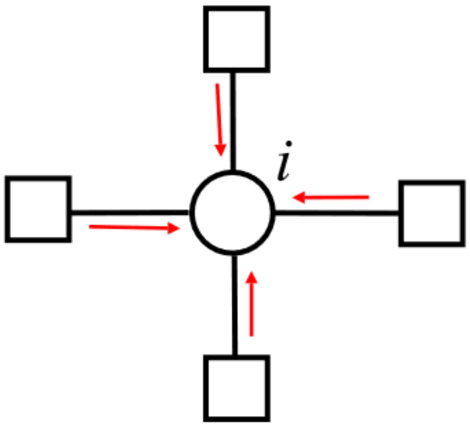
http://people.csail.mit.edu/andyd/CIOG_papers/yedidia_jsp_preprint_princeton.pdf

Please read the paper till Section 7.

“Messages” in BP or message passing algorithms.

- “States” or “Labels” refer to the different discrete values of the variables.
- A message is what a variable tells its neighbors the cost for it to be in different states. Note that a message is not “states” of the variables.
- The size/dimension of a message is same as the number of states the associated variable can take.

Belief update rule

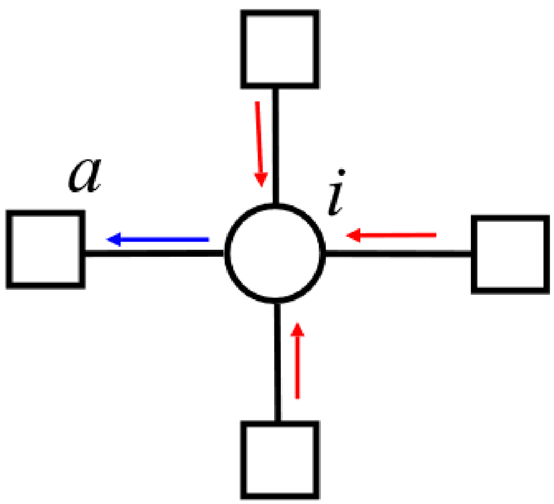


$$b_i(x_i) = \sum_{a \in N(i)} m_{a \to i}(x_i)$$

"belief" "messages"

The belief update rule for the min-sum BP algorithm says that the belief at a variable node is simply the sum of incoming messages from neighboring factor nodes.

Variable-to-factor message update rule

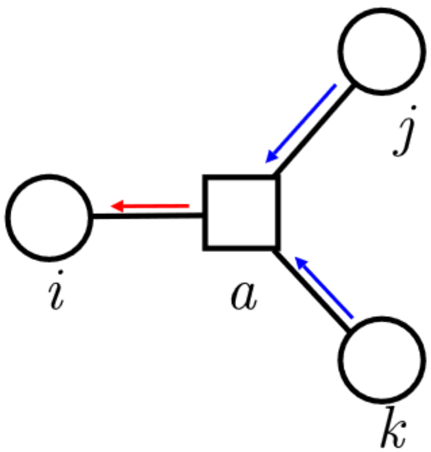


$$m_{i \to a}(x_i) = \sum_{b \in N(i) \setminus a} m_{b \to i}(x_i)$$

$$m_{i \to a}(x_i) = b_i(x_i) - m_{a \to i}(x_i)$$

The variable-to-factor message update rule in min-sum BP says that the outgoing (blue) message is the sum of all the incoming (red) messages on edges other than the edge of the outgoing message.

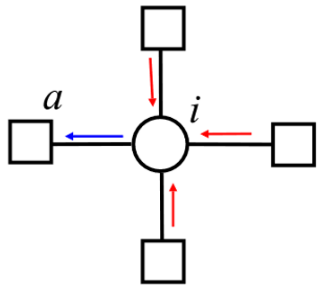
Factor to variable message updating rule



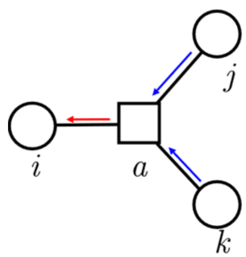
$$m_{a \to i}(x_i) = \min_{x_j, x_k} \left[C_a(x_i, x_j, x_k) + m_{j \to a}(x_j) + m_{k \to a}(x_k) \right]$$

The update rule for a message from factor to a variable depends on the local cost function, and the incoming variable-to-factor messages on other edges.

2 Message updating rules



$$m_{i\to a}(x_i) = \sum_{b\in N(i)\backslash a} m_{b\to i}(x_i)$$

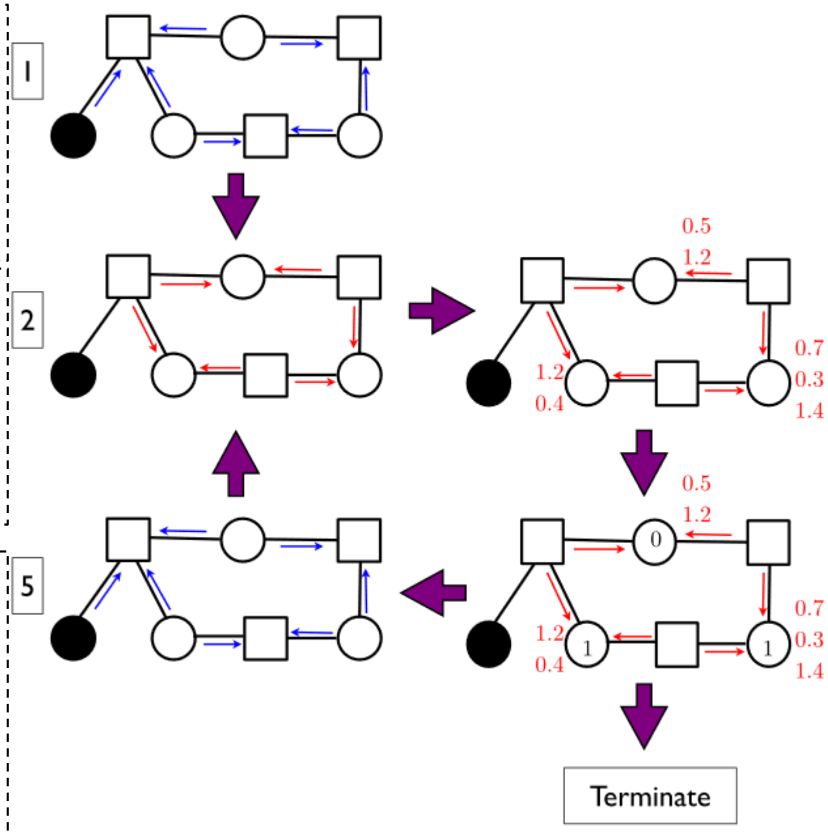


$$m_{a\to i}(x_i) = \min_{X_a\backslash x_i}\left[C_a(X_a) + \sum_{j\in N(a)\backslash i} m_{j\to a}(x_j)\right]$$

Outline of Message Passing Algorithms

1. Messages from variable nodes to factor nodes (in blue) are initialized to random or non-informative values.

2. The factor nodes compute from the incoming messages new outgoing messages (in red).



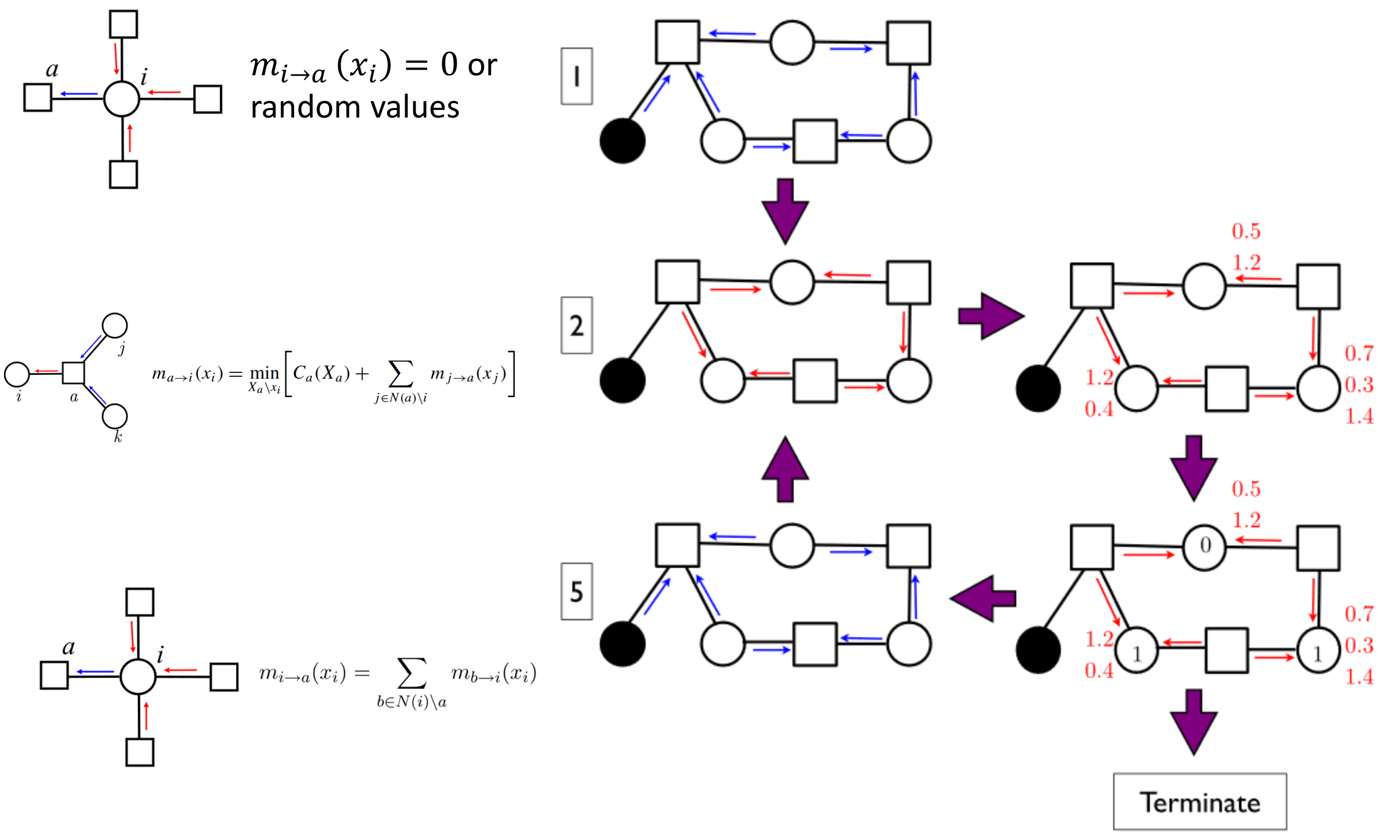
3. The messages are converted into beliefs, which in BP are generally represented as a cost for each possible state (the red numbers)

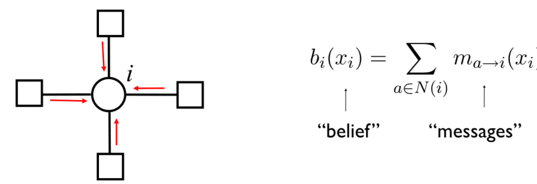
4. The beliefs are thresholded to their lowest cost (represented by the number inside the variable node), and a termination is checked.

5. The beliefs and incoming messages are used to compute new outgoing messages from the variable nodes, and then one returns to step 2, and the cycle continues.

Source: Yedidia

Outline of Message Passing Algorithms





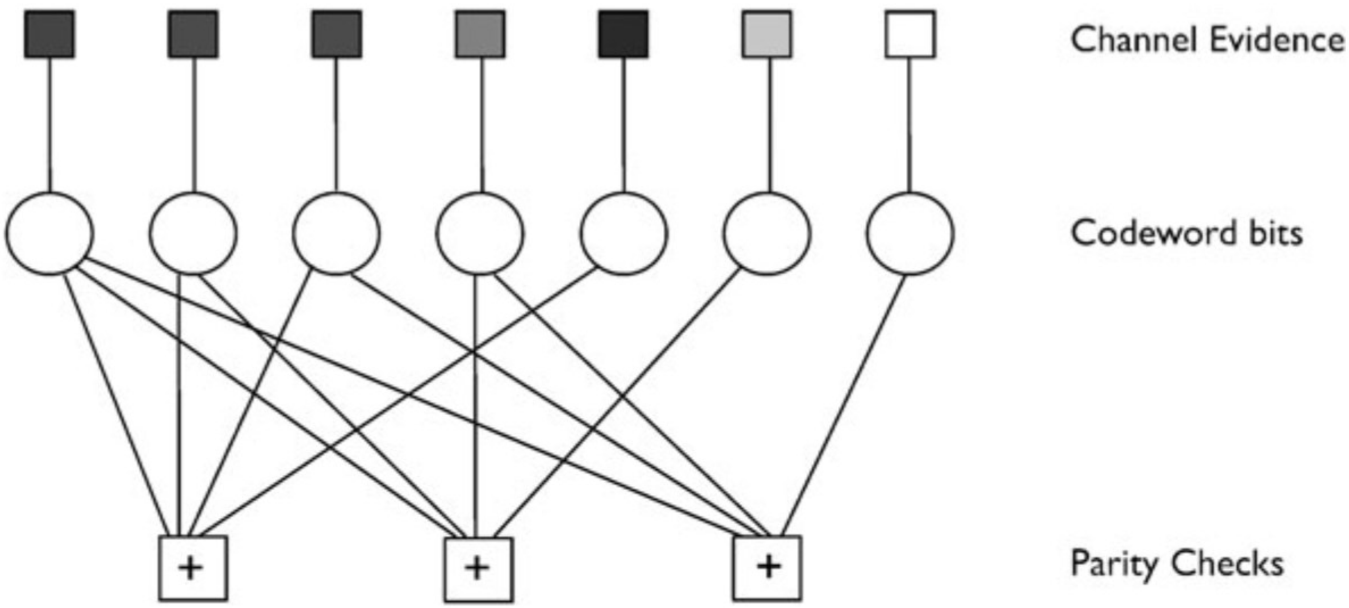
4. The beliefs are thresholded to their lowest cost (represented by the number inside the variable node), and a termination is checked.

Source: Yedidia

What is the termination condition?

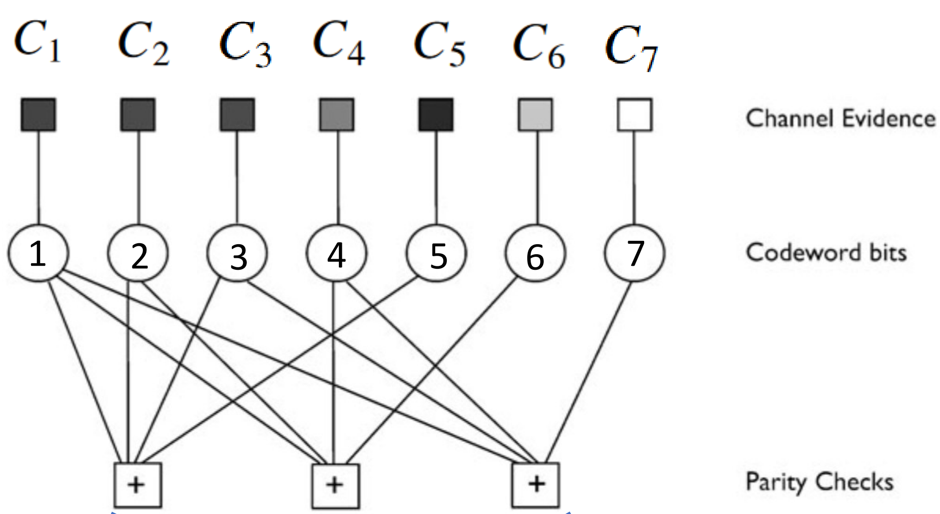
- Check whether the guess or the beliefs have changed from previous iterations
- Check whether maximum number of iterations has been reached.

Hamming Code



A factor graph for the (N = 7 , k = 4) Hamming code, which has seven codeword bits, of which the left-most four are information bits, and the last three are parity bits.

Hamming Code



$C_B(x_1, x_2, x_4, x_6)$

$C_A(x_1, x_2, x_3, x_5)$ $C_C(x_1, x_3, x_4, x_7)$

- 0 or infinity for parity check costs

$C_1(x_1 = 0) = 0.0;$ $C_1(x_1 = 1) = 3.0$

$C_2(x_2 = 0) = 0.0;$ $C_2(x_2 = 1) = 2.0$

$C_3(x_3 = 0) = 0.0;$ $C_3(x_2 = 1) = 2.5$

$C_4(x_4 = 0) = 0.0;$ $C_4(x_2 = 1) = 5.4$

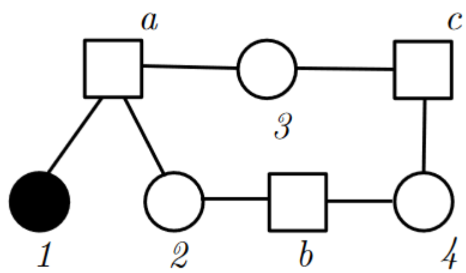
$C_5(x_5 = 0) = 0.0;$ $C_5(x_2 = 1) = 4.0$

$C_6(x_6 = 0) = 0.2;$ $C_6(x_2 = 1) = 0.0$

$C_7(x_7 = 0) = 0.7;$ $C_7(x_2 = 1) = 0.0$

Factor Graphs (Using Energy or Cost functions)

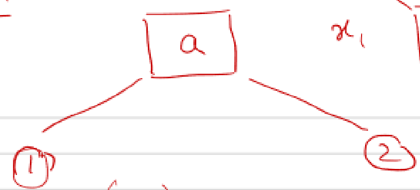
x_1	x_2	x_3	C_a
0	0	0	∞
0	0	1	0
0	1	0	0
0	1	1	∞
1	0	0	0
1	0	1	∞
1	1	0	∞
1	1	1	0



x_3	x_4	C_c
0	0	0.4
0	1	1.9
0	2	0.2
1	0	4.9
1	1	0.3
1	2	2.4

x_2	x_4	C_b
0	0	1.2
0	1	1.7
0	2	3.2
1	0	1.9
1	1	0.6
1	2	1.4

Problem 1



C_a

$x_1 \backslash x_2$		
	0.3	0.2
	0.5	0.1

① $m_{1\to a} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ $\quad m_{2\to a} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

② $m_{a\to 1} = \begin{bmatrix} \min_{x_2}\left(C_a(0,x_2) + m_{2\to a}(x_2)\right) \\ \min_{x_2}\left[C_a(1,x_2) + m_{2\to a}(x_2)\right] \end{bmatrix} = \begin{pmatrix} 0.2 \\ 0.1 \end{pmatrix}$

$m_{a\to 2} = \begin{bmatrix} \min_{x_1}\left(C_a(x_1,0) + m_{1\to a}(x_1)\right) \\ \min_{x_1}\left(C_a(x_1,1) + m_{1\to a}(x_1)\right) \end{bmatrix} = \begin{pmatrix} 0.3 \\ 0.1 \end{pmatrix}$

③ $b_1 = \begin{pmatrix} 0.2 \\ 0.1 \end{pmatrix}$ $\quad b_2 = \begin{pmatrix} 0.3 \\ 0.1 \end{pmatrix}$

④ $x_1 \leftarrow 1$ $\quad x_2 \leftarrow 1$

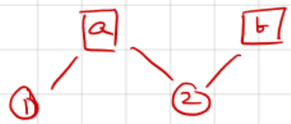
⑤ $m_{1\to a} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ $\quad m_{2\to a} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

⑥ Same as Step 2.

⑦ $b_1 = \begin{pmatrix} 0.2 \\ 0.1 \end{pmatrix}$ $\quad b_2 = \begin{pmatrix} 0.3 \\ 0.1 \end{pmatrix}$

⑧ $x_1 \leftarrow 1$, $\quad x_2 \leftarrow 1$

The state assignments have not changed from Step ④. So terminate.



C_a

	x_2	
x_1	0.3	0.2
	0.5	0.1

C_b

x_2	
	0.1
	0.6

Goal: State assignment

$$x_1, x_2 \in \{0, 1\}$$

$$x_1, x_2 = \underset{x_1, x_2}{\arg\min} \left[C_a(x_1, x_2) + C_b(x_2) \right]$$

BP:

① $m_{1\to a} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ $m_{2\to a} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ $m_{2\to b} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

② $m_{a\to 1} = \begin{bmatrix} \min\limits_{x_2} [C_a(0, x_2) + m_{2\to a}(x_2)] \\ \min\limits_{x_2} [C_a(1, x_2) + m_{2\to a}(x_2)] \end{bmatrix} = \begin{pmatrix} 0.2 \\ 0.1 \end{pmatrix}$

$m_{a\to 2} = \begin{bmatrix} \min\limits_{x_1} [C_a(x_1, 0) + m_{1\to a}(x_1)] \\ \min\limits_{x_1} [C_a(x_1, 1) + m_{1\to a}(x_1)] \end{bmatrix} = \begin{pmatrix} 0.3 \\ 0.1 \end{pmatrix}$

$m_{b\to 2} = \begin{bmatrix} \min C_b(0) \\ \min C_b(1) \end{bmatrix} = \begin{bmatrix} 0.1 \\ 0.6 \end{bmatrix}$

③ $b_1 = \begin{pmatrix} 0.2 \\ 0.1 \end{pmatrix}$ $b_2 = \begin{pmatrix} 0.3 \\ 0.1 \end{pmatrix} + \begin{pmatrix} 0.1 \\ 0.6 \end{pmatrix} = \begin{pmatrix} 0.4 \\ 0.7 \end{pmatrix}$

④ $x_1 \leftarrow 1$, $x_2 \leftarrow 0$

⑤ $m_{1\to a} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $m_{2\to a} = \begin{pmatrix} 0.1 \\ 0.6 \end{pmatrix}$, $m_{2\to b} = \begin{pmatrix} 0.3 \\ 0.1 \end{pmatrix}$

⑥

(6) $m_{a\to 1} = \begin{bmatrix} \min_{x_2} \left(C_a(0, x_2) + m_{2\to a}(x_2) \right) \\ \min_{x_2} \left(C_a(1, x_2) + m_{2\to a}(x_2) \right) \end{bmatrix} = $ ([redacted]) 0.4 0.6

$m_{a\to 2} = \begin{bmatrix} \min_{x_1} \left(C_a(x_1, 0) + m_{1\to a}(x_1) \right) \\ \min_{x_1} \left(C_a(x_1, 1) + m_{1\to a}(x_1) \right) \end{bmatrix} = \begin{pmatrix} 0.3 \\ 0.1 \end{pmatrix}$

$m_{b\to 2} = \begin{bmatrix} 0.1 \\ 0.6 \end{bmatrix}$

(7) $b_1 = $ ([redacted]) 0.4 0.6 $b_2 = \begin{pmatrix} 0.4 \\ 0.7 \end{pmatrix}$

(8) $x_1 \leftarrow 0, \quad x_2 \leftarrow 0$

(9) $m_{1\to a} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad m_{2\to a} = \begin{pmatrix} 0.1 \\ 0.6 \end{pmatrix} \quad m_{2\to b} = \begin{pmatrix} 0.3 \\ 0.1 \end{pmatrix}$

$\longrightarrow$ Same as (5)

(10) Same as step (6)

(11) Same as step (7)

(12) $x_1 \leftarrow 0, \quad x_2 \leftarrow 0$

Terminate