# Belief Propagation and Free Energy

David Kappel

Advanced Signal Processing Seminar WS 2010

November 29, 2010

# Outline

Bayesian Networks, Markov Random Fields and Factor Graphs

Standard Belief Propagation

Free Energies

Summary

# Outline

Bayesian Networks, Markov Random Fields and Factor Graphs

Standard Belief Propagation

Free Energies

Summary

# Bayesian Networks

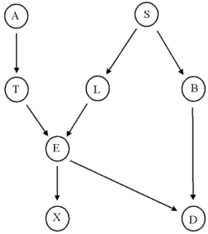
- Popular graphical model in AI literature.
- Directed graph structure.
- Applications: Medical diagnosis, language understanding, heuristic search ...

# "Asia" Medical Diagnosis

Medical system modelling statistical dependencies between symptoms, test results and diseases:

- A recent trip to Asia (A) increases the chances to tuberculosis (T).
- Smoking (S) is a risk factor for both lung cancer (L) and bronchitis (B).
- The presence of either (E) tuberculosis or lung cancer can be detected by an X-ray results (X), but the X-ray alone cannot distinguish between them.
- Dyspnoea (D) (shortness of breath) may be caused by bronchitis (B), or either (E) tuberculosis or lung cancer.

# "Asia" Medical Diagnosis




$$
\begin{aligned}
\mathrm{p}(\mathbf{x}) &= \mathrm{p}(x_A)\mathrm{p}(x_S)\mathrm{p}(x_T \mid x_A)\mathrm{p}(x_B \mid x_S) \times \\
&\quad \mathrm{p}(x_E \mid x_L, x_T)\mathrm{p}(x_D \mid x_B, x_E)\mathrm{p}(x_X \mid x_E)
\end{aligned} \tag{1}
$$

# Bayesian Networks

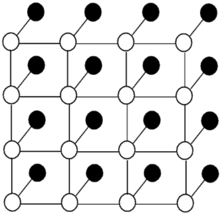
joint probability:

$$\mathrm{p}(x_1, x_2, \ldots, x_N) = \prod_{i=1}^{N} \mathrm{p}(x_i \mid Par\,(x_i)) \tag{2}$$

marginal probability:

$$\mathrm{p}(x_i) = \sum_{\{x\} \backslash x_i} \mathrm{p}(x_1, x_2, \ldots, x_N) \tag{3}$$

# Pairwise Markov Random Fields



- Undirected graph model.
- Low level computer vision, image $y_i$, scene $x_i$.

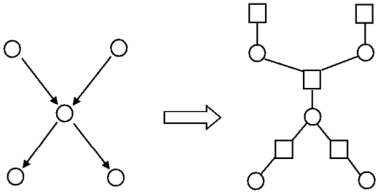
# Pairwise Markov Random Fields

- evidence function $\phi_i\left(x_i, y_i\right)$
- compatibility function $\psi_{ij}\left(x_i, x_j\right)$

joint probability:

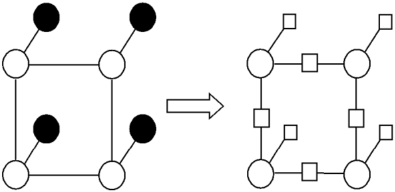
$$\mathrm{p}(\mathbf{x}, \mathbf{y}) = \frac{1}{Z} \prod_{(i,j)} \psi_{ij}\left(x_i, x_j\right) \prod_{i} \phi_i\left(x_i, y_i\right) \tag{4}$$

# Converting Bayesian Networks to Factor Graphs



Factor graph functions $\psi_a\left(\{x\}_a\right)$ directly correspond to Bayesian network probabilities $\mathrm{p}(x_i \mid Par\left(x_i\right))$.

# Converting Markov Random Fields to Factor Graphs



- Observed nodes replaced by single node functions $\phi_i\left(x_i, y_i\right)$
- Two node functions $\psi_{ij}\left(x_i, x_j\right)$ linking hidden nodes

# Outline

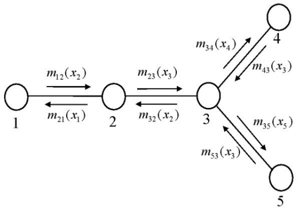
# Standard Belief Propagation

joint probability:

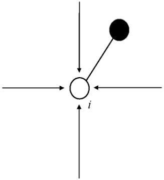
$$\mathrm{p}(\mathbf{x}) = \frac{1}{Z} \prod_{(i,j)} \psi_{ij}(x_i, x_j) \prod_i \phi_i(x_i) \tag{5}$$

marginal probability:

$$\mathrm{p}(x_i) = \sum_{\mathbf{x} \backslash x_i} \mathrm{p}(x_1, x_2, \ldots, x_N) \tag{6}$$



# Message Passing




$$b_i\left(x_i\right) = k\phi_i\left(x_i\right) \prod_{j \in N(x_i)} m_{ij}\left(x_i\right) \tag{7}$$

$$m_{ij}\left(x_i\right) \leftarrow \sum_{x_i} \phi_i\left(x_i\right) \psi_{ij}\left(x_i, x_j\right) \prod_{k \in N(x_i) \backslash j} m_{ki}\left(x_i\right) \tag{8}$$

# Outline

Bayesian Networks, Markov Random Fields and Factor Graphs

Standard Belief Propagation

Free Energies

Summary

# Free Energies

- A system of $N$ particles.
- Each which can in one of a discrete number of states.
- The state of the system: $\mathbf{x} = \{x_1, x_2, \ldots, x_N\}$
- Each state has a corresponding energy $E\left(\mathbf{x}\right)$

Boltzmann's law:

$$\mathrm{p}(\mathbf{x}) = \frac{1}{Z} e^{-E(\mathbf{x})/T} \tag{9}$$

# The Potts and Ising Models

- field $h_i(x_i, y_i) = \ln \phi_i(x_i, y_i)$
- interaction $J_{ij}(x_i, x_j) = \ln \psi_{ij}(x_i, x_j)$

Potts model energy:

$$E(\mathbf{x}) = -\sum_{(i,j)} J_{ij}(x_i, x_j) - \sum_i h_i(x_i, y_i) \tag{10}$$

Boltzmann law:

$$\mathrm{p}(\mathbf{x}) = \frac{1}{Z} e^{-E(\mathbf{x})/T}$$

$$\rightarrow \mathrm{p}(\mathbf{x}) = \frac{1}{Z} \prod_{(i,j)} \psi_{ij}(x_i, x_j) \prod_i \phi_i(x_i, y_i)$$

# The Helmholtz free energy

- Important quantity in statistical mechanics.
- Many techniques exist which give good approximations.

partition function (normalisation constant):

$$Z = \sum_{\mathbf{x} \in S} e^{-E(\mathbf{x})/T} \tag{11}$$

Helmholtz Free Energy:

$$F_H = -\ln Z \tag{12}$$

# Variational approach:

- True probability distribution $\mathrm{p}(\mathbf{x})$.
- Trail probability distribution $b(\mathbf{x})$.

Gibbs Free Energy:

$$F(b) = U(b) - H(b) \tag{13}$$

Average energy:

$$U(b) = \sum_{\mathbf{x}\in S} b(\mathbf{x})\, E(\mathbf{x}) \tag{14}$$

Variational entropy:

$$H(b) = -\sum_{\mathbf{x}\in S} b(\mathbf{x}) \ln b(\mathbf{x}) \tag{15}$$

# Gibbs Free Energy

$$
\begin{aligned}
F(b) &= U(b) - H(b) \\
&= \sum_{\mathbf{x}\in S} b(\mathbf{x})E(\mathbf{x}) + \sum_{\mathbf{x}\in S} b(\mathbf{x}) \ln b(\mathbf{x}) \\
&= \sum_{\mathbf{x}\in S} b(\mathbf{x}) \left( E(\mathbf{x}) + \ln b(\mathbf{x}) \right)
\end{aligned}
$$

using Boltzmann's law:

$$
\begin{aligned}
\mathrm{p}(\mathbf{x}) &= \frac{1}{Z} e^{-E(\mathbf{x})} \\
E(\mathbf{x}) &= -\ln Z - \ln \mathrm{p}(\mathbf{x}) \\
E(\mathbf{x}) &= F_H - \ln \mathrm{p}(\mathbf{x})
\end{aligned}
$$

# Gibbs Free Energy

$$
\begin{aligned}
F(b) &= \sum_{\mathbf{x} \in S} b(\mathbf{x}) \left( F_H - \ln \mathrm{p}(\mathbf{x}) + \ln b(\mathbf{x}) \right) \\
&= F_H + \sum_{\mathbf{x} \in S} b(\mathbf{x}) \ln \frac{b(\mathbf{x})}{\mathrm{p}(\mathbf{x})} \\
&= F_H + D_{KL}(b|p)
\end{aligned}
$$

With $D_{KL}(b|p)$, the Kullback-Leibler divergence.

# Gibbs Free Energy

- BP is equal to minimizing the variational free energy.
- If graph is loop-free the solution is exact.
- If graph has loops the results are equal to the Bethe approximation from statistical mechanics.

# Outline

# Summary

- Different disciplines have yielded different approaches and models.
- Yet, the methods are similar and can be converted into each other.
- Belief Propagation can be understood as a solution to an energy minimisation problem.

# Questions?

# References

- Belief propagation and its relation to free energy [1, 3].
- Further readings on graphical models [2].


Jonathan S. Yeddidia, William T. Freeman, and Yair Weiss.
Understanding belief propagation and its generalisation.
Technical report, Mitsubishi Electric Research Laboratories, 2002.

Christopher M. Bishop.
Pattern Recognition and Machine Learning.
Springer Science + Business Media, 2006.

Jonathan S. Yeddidia, William T. Freeman, and Yair Weiss.
Constructing free-energy approximations and generalized belief propagation algorithms.
IEEE Transactions on Information Theory, 51(7), 2005.