Poset Belief Propagation - Experimental Results

Jonathan Harel, Robert J. McEliece, and Ravi Palanki[1]
California Institute of Technology
Department of Electrical Engineering
Pasadena, CA 91125, USA
e-mail: {harel,rjm,ravi}@systems.caltech.edu

Abstract — **Poset belief propagation, or PBP, is a flexible generalization of ordinary belief propagation which can be used to (approximately) solve many probabilistic inference problems. In this paper, we summarize some experimental results comparing the performance of PBP to conventional BP techniques.**

I. Introduction

In [2], McEliece and Yildirim introduced a class of algorithms called *belief propagation on partially ordered sets*, or PBP. This general class includes as special cases, for example, ordinary belief propagation [3], the sum-product algorithm [4], generalized belief propagation [1], (all of these with and without loops), as well as many other instances whose effectiveness has not yet been investigated in detail. We summarize PBP and report the results of some experiments we have performed.

II. The MDP Problem

Technically, PBP is an algorithm for solving any *marginalized product density*, or MDP, problem:

Let $\{x_1, x_2, ..., x_n\}$ be a set of n variables taking values in the finite set $A = \{0, 1, ..., q \;\; 1\}$, and let $\mathcal{R} = \{R_1, R_2, ..., R_M\}$ be a collection of M sparse subsets of $[n] = \{1, 2, ..., n\}$. Now suppose we are given a set of "local potentials" $\{\alpha_R(\mathbf{x}_R) : R \in \mathcal{R}\}$. These kernels define a probability density function:

$$B(\mathbf{x}) = \frac{1}{Z} \prod_{R \in \mathcal{R}} \alpha_R(\mathbf{x}_R),$$

where Z is chosen so that $B(\mathbf{x})$ be normal. The problem is to compute, exactly or approximately, the *local marginal densities*, $\{B_R(\mathbf{x}_R)\}_{R \in \mathcal{R}}$ of the product density, where

$$B_R(\mathbf{x}_R) = \sum_{\mathbf{x} \setminus \mathbf{x}_R} B(\mathbf{x}).$$

III. Posets and PBP

Let P be a finite poset and let $H = H(P)$ be the Hasse diagram for P. Assume P is a junction poset for $\mathcal{R}$ (see [2] for details). Each vertex $\rho \in H(P)$ has associated with it a "belief table" $b_\rho(\mathbf{x}_\rho)$ initialized to $\prod_{R \in \mathcal{R}} \alpha_R(\mathbf{x}_R)$. An edge $e = (\rho, \sigma) \in H(P)$ is *inconsistent* if $b_\rho(\mathbf{x}_\rho)$ does not marginalize down to exactly $b_\sigma(\mathbf{x}_\sigma)$.

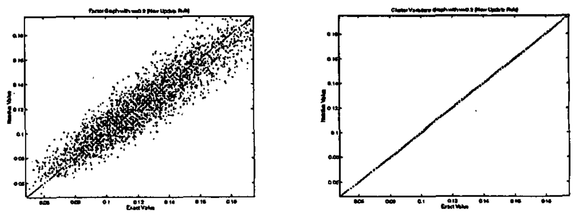
PBP proceeds as follows: for each edge $e = (\rho, \sigma)$ that is inconsistent, we define a correction table $\Delta_e(\mathbf{x}_\sigma)$ which when multiplied by the belief at σ will yield the belief at ρ. We then update the beliefs at all τ s.t. $\tau \geq \sigma$, but not $\tau \geq \rho$:

$$b_\tau(\mathbf{x}_\tau) \leftarrow b_\tau(\mathbf{x}_\tau) \cdot \Delta_e(\mathbf{x}_\sigma).$$

The hope is that when all the edges are consistent,

$$b_\rho(\mathbf{x}_\rho) \approx B_\rho(\mathbf{x}_\rho) = \sum_{\mathbf{x} \setminus \mathbf{x}_\rho} B(\mathbf{x}).$$

IV. Results



(a) Factor Graph (b) Cluster Variations

We found that by using different posets for the same given sets $\mathcal{R}$ and $\{\alpha_R\}$, we could vastly improve performance on small examples. It was discovered that the algorithm will often evolve too quickly, especially when $H(P)$ is deep, and so a damping coefficient w was employed in the update rule. In the above figures, we plot $\{(B_\rho(\mathbf{x}_\rho), b_\rho(\mathbf{x}_\rho))\}$ with $w = .2$ for randomly chosen kernels on posets of 5 variables. We notice in (a), the "Factor Graph", which carries out ordinary BP, the performance is significantly worse than in (b), in which the inference was carried out on a "Cluster Variations Graph", equivalent to GBP [1]. So far, we have not been able to reproduce this kind of improvement on much larger examples of inference problems. However, a tradeoff was observed in which we could gain some performance by sacrificing complexity. This exciting result raises some interesting possibilities concerning the class of PBP algorithms in between our endpoints. We remain confident that PBP may prove to be effective when ordinary BP is not.

References

[1] J. S. Yedidia, W. T. Freeman, and Y. Weiss, "Generalized belief propagation," pp. 689-695 in *Advances in Neural Information Processing Systems 13*, (2000) eds. Todd K. Leen, Thomas G. Dietterich, and Volker Tresp.

[2] R. J. McEliece and M. Yildirim, "Belief Propagation on Partially Ordered Sets," to appear in the IMA Volume "Mathematical Systems Theory in Biology, Communication, Computation, and Finance," D. Gilliam and J. Rosenthal, eds.

[3] J. Pearl, *Probabilistic Reasoning in Intelligent Systems.* San Francisco: Morgan-Kaufmann, 1988.

[4] F. R. Kschichang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inform. Theory*, vol. 47, no.2 (Feb. 2001), pp. 498-519.

[1] This research was supported by NSF grant no. CCR-0118670, and grants from Sony, Qualcomm, and Caltech's Lee Center for Advanced Networking.