Rounding Guarantees for Message-Passing MAP Inference with Logical Dependencies

Stephen H. Bach,[1] Bert Huang,[1,2] and Lise Getoor[3]

[1] University of Maryland, College Park, [2] Virginia Tech, [3] University of California, Santa Cruz

1. MRFs with Logical Dependencies

Consider MAP inference in a Markov random field (MRF)

$$\arg\max_{\boldsymbol{x}} \; \boldsymbol{w}^\top \boldsymbol{\phi}(\boldsymbol{x})$$

where each variable is Boolean, each parameter is non-negative, and each potential is defined by the truth value of a logical clause:

$$\phi_j(\boldsymbol{x}) \triangleq \left(\bigvee_{i \in I_j^+} x_i \right) \vee \left(\bigvee_{i \in I_j^-} \neg x_i \right)$$

We refer to such MRFs as ***logical MRFs***.

MAP Inference in logical MRFs is NP-hard. [Garey et al., 1976]

We provide rounding guarantees for message-passing approximate MAP inference for logical MRFs

Examples of Dependencies in Logical MRFs

1. ***Implications***

$$\phi_j(\boldsymbol{x}) \triangleq \left(\bigwedge_{i \in I_j^-} x_i \right) \implies \left(\bigvee_{i \in I_j^+} x_i \right)$$

2. ***Submodular functions***

$$\phi_a(\boldsymbol{x}) \triangleq \neg x_1 \vee x_2 \qquad \phi_b(\boldsymbol{x}) \triangleq x_1 \vee \neg x_2$$

3. ***Supermodular functions***

$$\phi_a(\boldsymbol{x}) \triangleq x_1 \vee x_2 \qquad \phi_b(\boldsymbol{x}) \triangleq \neg x_1 \vee \neg x_2$$

2. Approximate MAP Inference for Logical MRFs

We consider two main approaches to approximate MAP inference:

1. ***Local consistency relaxations***

Introduce marginal distributions over variable and potential states, then constrain them to only be locally consistent

$$\arg\max_{(\boldsymbol{\theta},\boldsymbol{\mu}) \in \mathbb{L}} \sum_{j=1}^{m} w_j \sum_{\boldsymbol{x}_j} \theta_j(\boldsymbol{x}_j) \; \phi_j(\boldsymbol{x}_j)$$

where

$$\mathbb{L} \triangleq \left\{ \boldsymbol{\theta}, \boldsymbol{\mu} \geq \mathbf{0} \;\middle|\; \begin{array}{lr} \sum_{\boldsymbol{x}_j | x_j(i)=k} \theta_j(\boldsymbol{x}_j) = \mu_i(k) & \forall i,j,k \\ \sum_{\boldsymbol{x}_j} \theta_j(\boldsymbol{x}_j) = 1 & \forall j \\ \sum_{k=0}^{K_i - 1} \mu_i(k) = 1 & \forall i \end{array} \right\}$$

Advantage: Admits highly scalable message-passing algorithms

2. ***MAX SAT relaxations***

View as instance of MAX SAT, and relax as an LP that bounds expected truth value [Goemans and Williamson, 1994]

$$\arg\max_{\boldsymbol{y} \in [0,1]^n} \sum_{j=1}^{m} w_j \min \left\{ \sum_{i \in I_j^+} y_i + \sum_{i \in I_j^-} (1 - y_i), 1 \right\}$$

Round each variable with probability $p_i = \frac{1}{2} y_i^\star + \frac{1}{4}$ using the method of conditional probabilities

Advantage: Gives discrete solutions of guaranteed 3/4 quality

3. Equivalence Analysis

Theorem: For any logical MRF, the first-order local consistency relaxation of MAP inference is equivalent to the MAX SAT relaxation of Goemans and Williamson [1994].

Proof Technique:

Analyze the local consistency relaxation as a hierarchical optimization:

$$\max_{\boldsymbol{\mu} \in [0,1]^i} \sum_{j=1}^{m} \hat{\phi}_j(\boldsymbol{\mu}) \qquad \text{where} \qquad \hat{\phi}_j(\boldsymbol{\mu}) = \max_{\boldsymbol{\theta}_j | (\boldsymbol{\theta},\boldsymbol{\mu}) \in \mathbb{L}} w_j \sum_{\boldsymbol{x}_j} \theta_j(\boldsymbol{x}_j) \; \phi_j(\boldsymbol{x}_j)$$

Use the Karush-Kuhn-Tucker conditions to find value of $\hat{\phi}_j(\boldsymbol{\mu})$ for any setting of $\boldsymbol{\mu}$:

$$\hat{\phi}_j(\boldsymbol{\mu}) = w_j \min \left\{ \sum_{i \in I_j^+} \mu_i + \sum_{i \in I_j^-} (1 - \mu_i), 1 \right\}$$

4. Practical Implications

The equivalence of the two relaxations means that the advantages of each can be combined into a single technique:

1. Solve the local consistency relaxation with any of a number of scalable message-passing algorithms

2. Find a discrete solution of 3/4 quality by applying the rounding procedure of Goemans and Williamson [1994] to the optimal pseudomarginals $\boldsymbol{\mu}^\star$.

Scalable message-passing algorithms for finding $\boldsymbol{\mu}^\star$ include subgradient dual decomposition, the alternating direction method of multipliers (ADMM), and proximal methods

Supported by NSF grant no. IIS1218488, and IARPA contract D12PC00337