Sequential Graph Matching with Sequential Monte Carlo

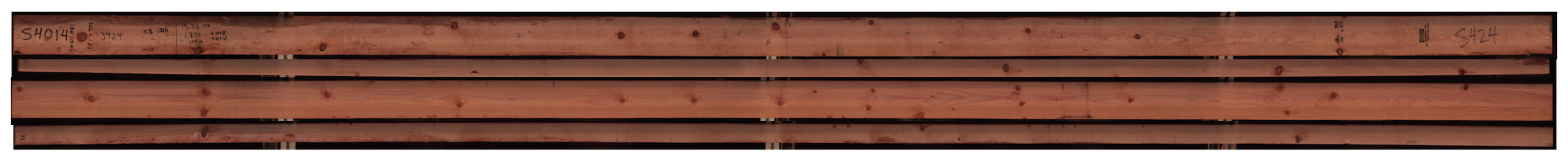
Seong-Hwan Jun[1] Samuel W.K. Wong[2] James V. Zidek[1] Alexandre Bouchard-Côté[1]

[1]Department of Statistics, University of British Columbia [2]Department of Statistics, University of Florida

Overview

- Plackett-Luce preference model applied to graph matching.
- Combinatorial sequential Monte Carlo for sampling graph matching.
- Supports both supervised and unsupervised learning via Monte Carlo Expectation Maximization.

Motivation: Knot Matching



Observation: the *knot faces* (vertices) on each of the four surfaces of a lumber.
Conditions:

1. Knot faces are remnants of a tree branch. Tree branches are represented by a hyperedge, which may contain $\{2, 3\}$ knot faces.
2. The knot faces from the same surface cannot be connected by a branch.

Objective: Represent a piece of lumber as a 4-partite hypergraph. Infer the latent branch structure via matching of the knot faces.

Sequential Graph Matching

Idea: Construct graph matching via sequence of decisions made by vertices amongst potential suitors (edges).

1. Select a vertex at random $\sigma : \{1, ..., n\} \to \{1, ..., n\}$.
2. Each vertex makes a decision amongst the suitors $\mathcal{D}(v_{\sigma(t)}|m_{t-1})$:
$$p(d_{v_{\sigma(t)}}|m_{t-1}, \sigma, \theta) = \frac{\exp\langle \phi(m_{t-1} + e_{d_{v_{\sigma(t)}}}), \theta\rangle}{\sum_{d' \in \mathcal{D}(v_{\sigma(t)}|m_{t-1})} \exp\langle \phi(m_{t-1} + e_{d'}), \theta\rangle},$$
3. Update matching state:
$$m_t = m_{t-1} + e_d = m_{t-1} \setminus \{d\} \bigcup \{d \bigcup \{v_{\sigma(t)}\}\}; \text{ for some } d \in \mathcal{D}(v_{\sigma(t)})$$

Inference

- Prior:
$$p(\theta) \propto \lambda \|\theta\|^2; p(\sigma) \propto 1$$
- Likelihood:
$$\ell(\theta) = \prod_{t=1}^{|V|} p(d_{v_{\sigma(t)}}|m_{t-1}, \sigma, \theta).$$
- Posterior:
$$p(\theta|d_\sigma, \sigma) \propto \ell(\theta)p(\sigma)p(\theta).$$
- **Unsupervised**: Jointly infer the latent matching and the parameters using Monte Carlo Expectation Maximization:
$$Q(\theta, \theta^t) = \sum_{d_\sigma, \sigma} p(d_\sigma, \sigma|\theta^t) \log p(\theta|d_\sigma, \sigma)$$
$$\approx \frac{1}{K}\sum_{k=1}^{K} \log p(\theta|d_{\sigma^k}, \sigma^k);\ (d_{\sigma^k}, \sigma^k) \sim p(d_\sigma|\sigma, \theta^t)p(\sigma)$$
$$\theta^{t+1} = \operatorname{argmax}_\theta Q(\theta, \theta^t)$$
- **Supervised**: Given labelled matchings, $\{m_i\}_{i=1}^I$, estimate the parameters. Sample $(\sigma^{i,k}, d_{\sigma^{i,k}}) \sim p(d_{\sigma^i}, \sigma^i|m_i, \theta^t)$.

Illustration: Pairwise decision model

Decision set formulation:

1. If $v_{\sigma(t)}$ is covered, $\mathcal{D}(v_{\sigma(t)}|m_{t-1}) = \{d\}$ such that $v_{\sigma(t)} \in d$, $d \in m_{t-1}$.
2. Else, $\mathcal{D}(v_{\sigma(t)}|m_{t-1}) =$
$$\{e \in m_{t-1} | e \cup \{v_{\sigma(t)}\} \text{ does not violate the 4-partite matching conditions}\} \bigcup$$
$$\{u : u \text{ is not from the same partition as } v_{\sigma(t)} \text{ and not convered}\}$$

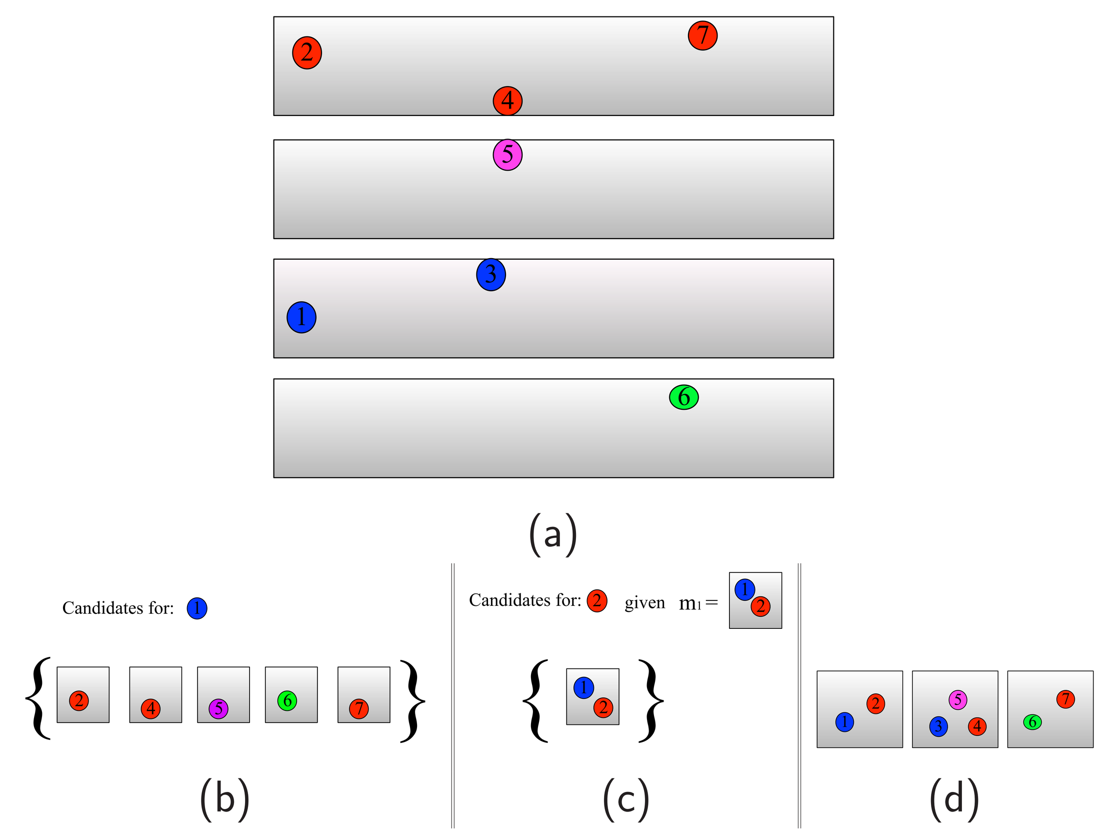


Figure 1: (a) 4-partite graph representing a piece of lumber. (b) The decision set for blue node #1. (c) The decision set for red node #2 given the decision made by blue node #1. (d) An example of a final matching.

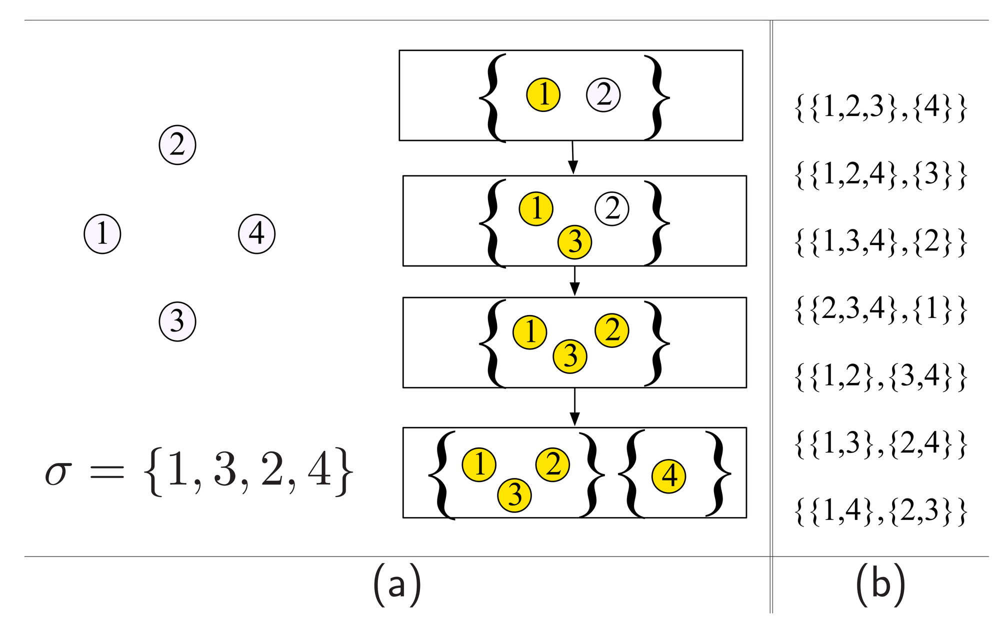


Figure 2: (a) A 4-partite graph with one vertex in each partition followed by an illustration of sequential matching process. Yellow vertices indicate the visited nodes. (b) All possible matching under the pairwise decision model.

Sampling Matching: Sequential Monte Carlo

Sequential Monte Carlo is useful for

- sampling of the sequence of decisions for both unsupervised and supervised learning.
- sampling from the predictive distribution.

For $r = 1, ..., |V|$:

1. (Resampling): $a_r^k \sim \text{Mult}(\bar{w}_{r-1,1}, \ldots, \bar{w}_{r-1,K})$
2. (Proposal): $s_{r,k} \sim \nu^+(\cdot|s_{r-1,a_r^k})$
3. (Weight computation): $w_{r,k} = \alpha(s_{r-1,a_r^k} \to s_{r,k})$
4. (Normalization): $\bar{w}_{r,k} = w_{r,k} / \sum_{k=1}^K w_{r,k}$

Combinatorial SMC for Sequential Graph Matching

- **Overcounting Problem**: Non-uniform number of paths for each state leads to biased sampling. See Figure 3 (a).
- **Backward kernel**: incorporating a simple backward kernel alleviates this problem [1]:
$$\nu^-(s|s') = 1[\nu^+(s'|s) > 0]|\mathcal{Q}(s')|^{-1}$$
$\mathcal{Q}(s)$ denotes the set of parent states for $s \in \mathcal{S}$.

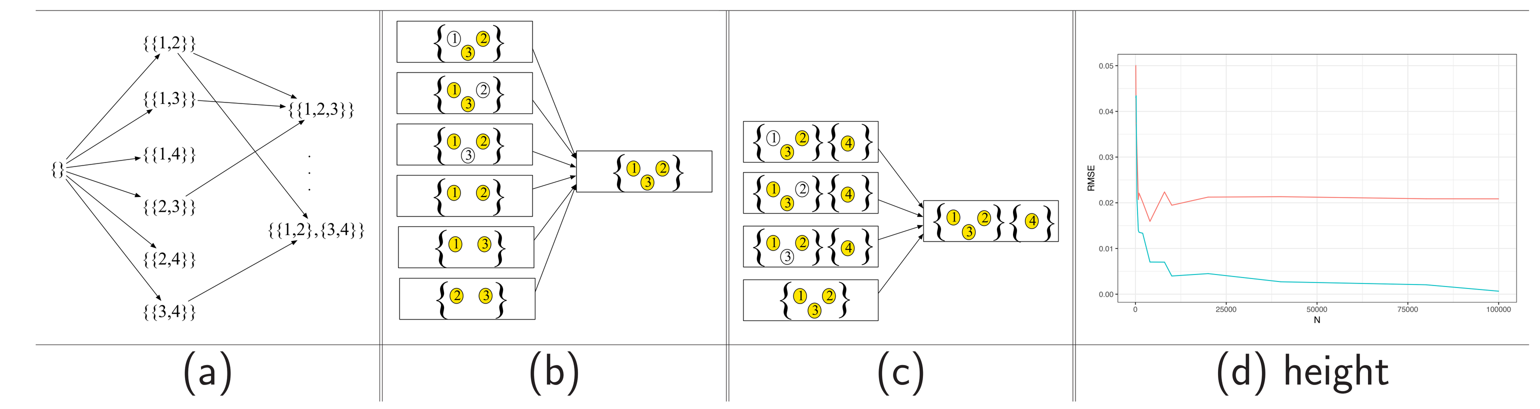


Figure 3: (a) 3 paths lead to $\{\{1, 2, 3\}\}$; 2 paths lead to $\{\{1, 2\}, \{3, 4\}\}$. (b) $\mathcal{Q}(s)$ for a state s containing a 3-matching where all three nodes have been visited. (c) $\mathcal{Q}(s)$ for a state containing a 3-matching and a singleton where all four nodes have been visited. (d) Sampling graph matching from uniform distribution with (blue) and without (red) overcounting correction.

Experimental Results

Knot matching experiments:

- Manually annotated 30 boards.
- Estimate the parameters and evaluated single sample prediction accuracy and quality of samples using Jaccard index via leave-one-out cross validation.

Image matching experiments: Single sample prediction accuracy evaluated on the CMU house dataset. Compared to the baseline of [2].

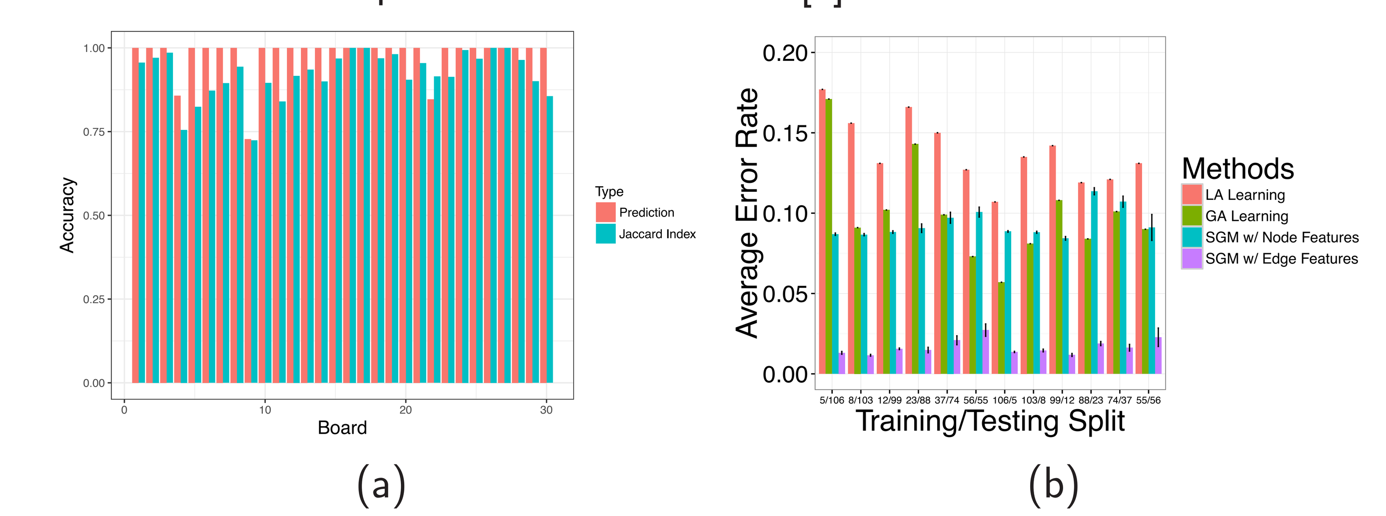


Figure 4: (a) Knot matching results. (b) Image matching results.

Conclusion

- Efficient parameter estimation for (hyper) graph matching.
- User configurable decision set to meet the problem constraints.
- Explored usability of sequential Monte Carlo for sampling graph matching.
- Our model admits full Bayesian analysis via particle MCMC framework.

References

[1] L. Wang, A. Bouchard-Côté, and A. Doucet. Bayesian phylogenetic inference using the combinatorial sequential Monte Carlo method. *Journal of the American Statistical Association*, 110:1362–1374, 2015.

[2] T. S. Caetano, J. J. McAuley, L. Cheng, Q. V. Le, and A. J. Smola. Learning graph matching. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(6):1048–1058, 2009.