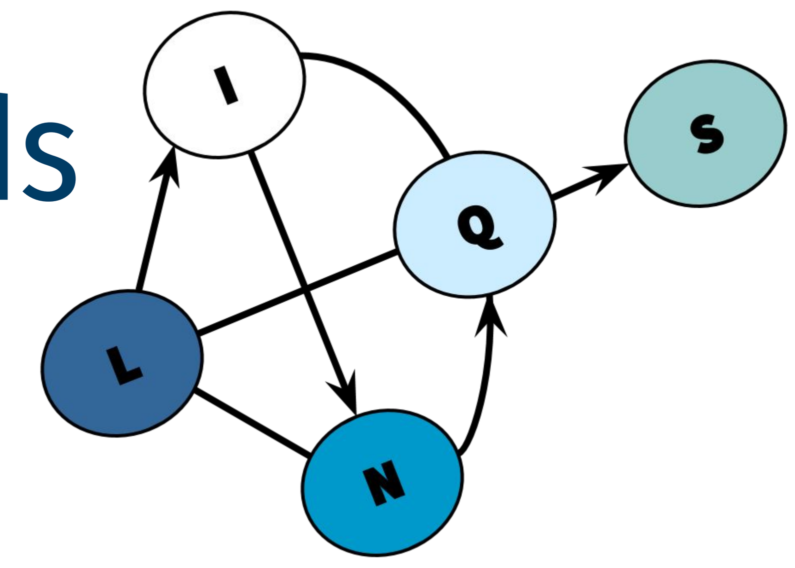
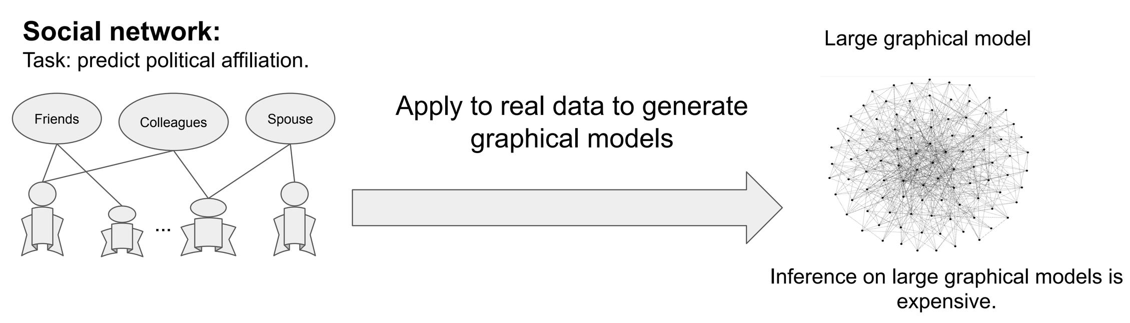
Lifted Hinge-Loss Markov Random Fields

Sriram Srinivasan+, Behrouz Babaki*, Golnoosh Farnadi*, & Lise Getoor+
+ University of California, Santa Cruz
* Polytechnique Montréal

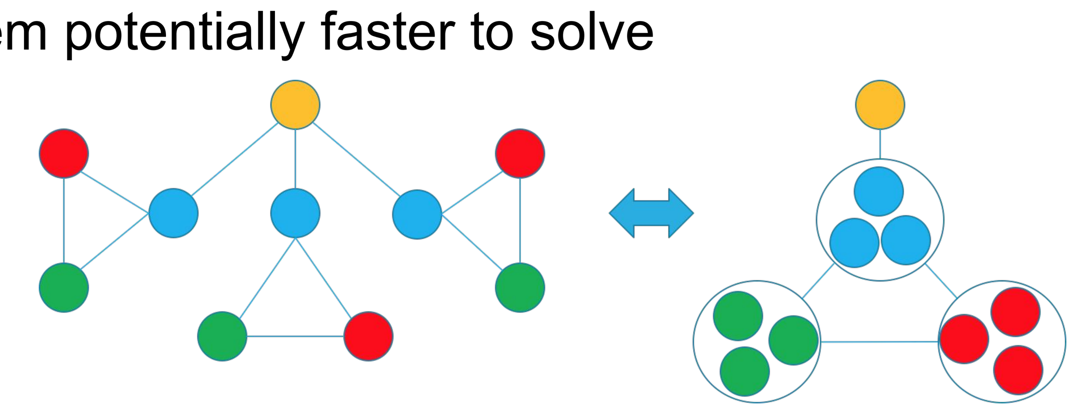


Lifted Inference

- Graphical models can be large and hard to perform inference



- Exploits symmetry to compress graphical mode
- Smaller problem potentially faster to solve



Contribution:

- Algorithm to combine power of lifting and convex objective of hinge-loss Markov random fields (HL-MRFs)
- Theoretical correctness of our approach
- Empirical analysis of impact of LHL-MRFs on different settings

Background

Hinge-loss Markov Random Field

- Probability distribution over continuous random variables.

$$P(\boldsymbol{y}|\boldsymbol{x}) = \frac{\exp\{-f_w(\boldsymbol{y},\boldsymbol{x})\}}{Z}$$

- Z is the normalization constant and

$$f_w(\boldsymbol{y},\boldsymbol{x}) = \sum_{i=1}^{m} w_i \phi_i(\boldsymbol{y},\boldsymbol{x})^{p_i}; where\ w_i \in R^+$$

$$\phi_i(y,x) = \max(l_i(\boldsymbol{y},\boldsymbol{x}),0)^{p_i}; where\ p_i \in \{1,2\}$$

- where ϕ is a hinge-loss potential. Inference problem can be written as:

$$argmax_{\boldsymbol{y}}\, P(\boldsymbol{y}|\boldsymbol{x}) = argmin_{\boldsymbol{y}}\, f_w(\boldsymbol{y},\boldsymbol{x})$$

- Optimization can be solved using alternating direction method of multipliers (ADMM).

Probabilistic Soft Logic

- Templating language for HL-MRFs
- Template rule grounded and converted to hinge-loss potentials.
- E.g.,

w: Friends(X,Y) ∧ Friends(Y, Z) ⟶ Friends(X, Z) ^p

Weight of rule

p = 1, linear loss
p = 2, quadratic loss

- Instantiate template rules with data: Bob, Dan, and Elsa

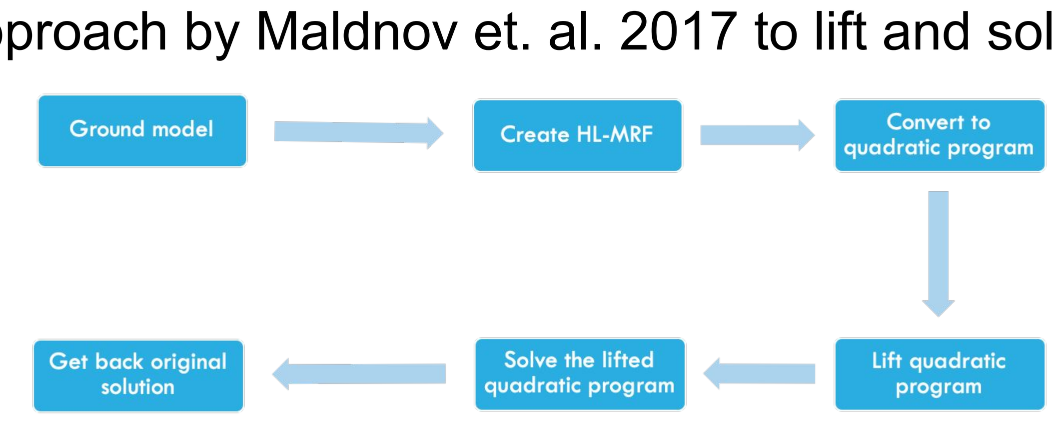
w: Friends(Bob,Dan) ∧ Friends(Dan, Elsa) ⟶ Friends(Bob, Elsa) ^2

- Converted to hinge-loss potential

$$\min((1-y_1)+(1-y_2)+y_3, 1) \quad y \in \{0,1\}$$

Solve via Lifted QP approach

- Applying approach by Maldnov et. al. 2017 to lift and solve HL-MRFs

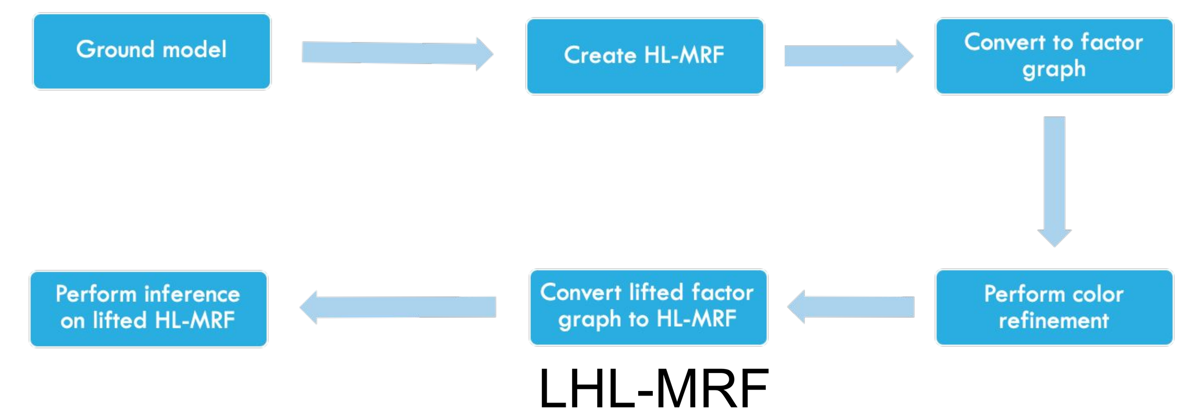


- Does not use the efficiencies of HL-MRFs

Lifted HL-MRFs

Approach:

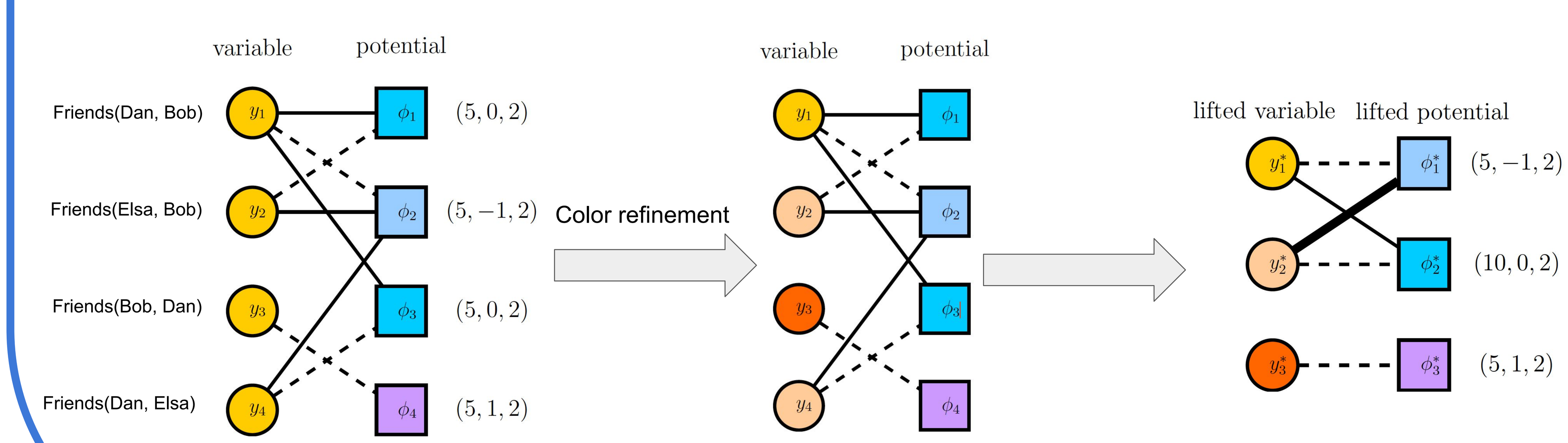
- Approach combines efficiencies from lifted inference and HL-MRFs



LHL-MRF illustration:

5: Friends(X,Y) ∧ Friends(Y, Z) ⟶ Friends(X, Z) ^2

Ground with Bob, Dan, and Elsa



Refer to paper for proof.

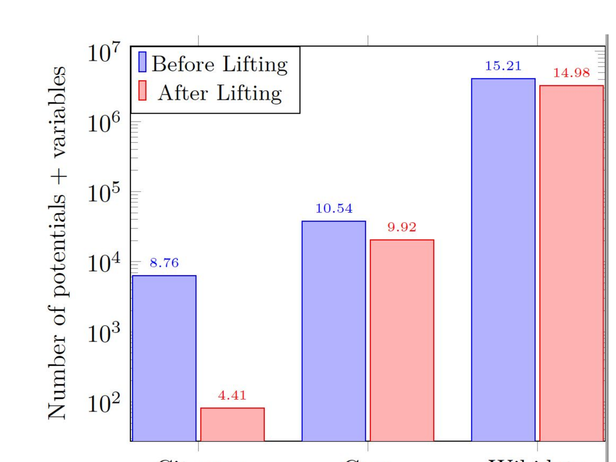
Results

Realworld datasets

- Citeseer: collective classification on citation dataset
  - 3312 papers and 4591 citations
- Cora: collective classification on citation dataset with
  - 2708 papers and 5429 citations
- Wikidata: entity-resolution on familial network
  - 418 families and 1844 family trees

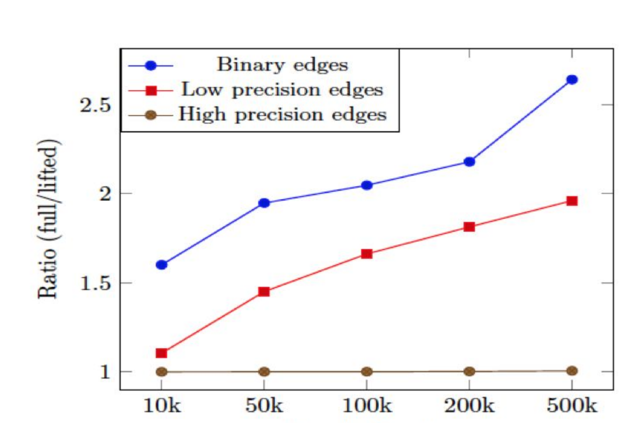
Datasets	HL-MRF (in sec)	LHL-MRF (solving) (in sec)	LHL-MRF (lifting) (in sec)	LHL-MRF (total) (in sec)
Wikidata	636.0	463.7	112.7	576.4
Cora	47.7	17.5	0.53	18.03
Citeseer	57.4	19.8	0.39	20.19

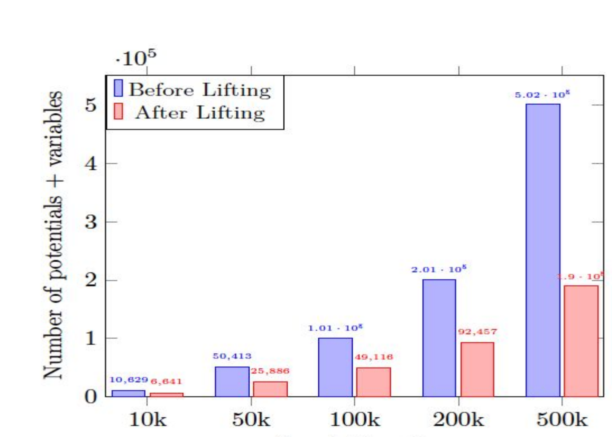
Table 1: Time taken to perform inference on different datasets.



Analysis On Synthetic data

- Link prediction dataset with 1000 nodes
- Compare effect of LHL-MRF on varied density and precision of observed links.





- Experiment to compare performance of LHL-MRF and LHL-MRF(Gurobi) for varied symmetry in the graph.