IJCAI 2015

The Complexity of MAP Inference in Bayesian Networks Specified Through Logical Languages

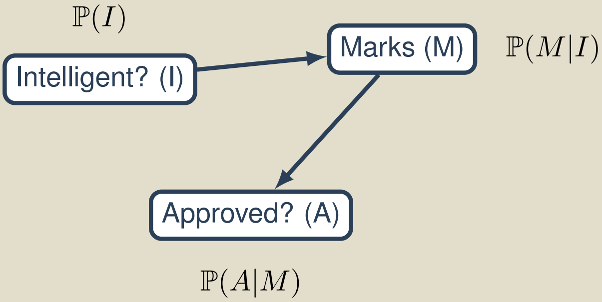
Denis D. Mauá — Universidade de São Paulo, Brazil

Cassio P. de Campos — Queen's University Belfast, UK

Fabio G. Cozman — Universidade de São Paulo, Brazil

Bayesian Network

- A DAG over a set of variables $X_1, \dots, X_n$
- A collection of local probability models $\mathbb{P}(X_i | \text{pa}(X_i))$
- Markov Condition: $\mathbb{P}(X_1, \dots, X_n) = \prod_i \mathbb{P}(X_i | \text{pa}(X_i))$



MAP Inference Problem

Given:

- Bayesian network $(G, \{\mathbb{P}(X_i|\text{pa}(X_i))\}_i)$
- Evidence $\mathbf{e} = \{E_1 = e_1, \ldots, E = e_m\}$
- MAP variables $\mathbf{M} \subseteq \{X_1, \ldots, X_n\} \setminus \{E_1, \ldots, E_m\}$

Compute

$$\max_{\mathbf{m}} \mathbb{P}(\mathbf{M} = \mathbf{m}, \mathbf{e}) = \max_{\mathbf{m}} \sum_{\mathbf{h}} \mathbb{P}(\mathbf{M} = \mathbf{m}, \mathbf{H} = \mathbf{h}, \mathbf{e})$$

Variants:

- DMAP: Decide if $\max_{\mathbf{m}} \mathbb{P}(\mathbf{M} = \mathbf{m}, \mathbf{e}) > k$ for given rational k
- SMAP: Select $\hat{\mathbf{m}}$ s.t. $\mathbb{P}(\mathbf{M} = \hat{\mathbf{m}}, \mathbf{e}) = \max_{\mathbf{m}} \mathbb{P}(\mathbf{M} = \mathbf{m}, \mathbf{e})$

MPE Inference Problem

Given:

- Bayesian network $(G, \{\mathbb{P}(X_i | \text{pa}(X_i))\}_i)$
- Evidence $\mathbf{e} = \{E_1 = e_1, \ldots, E = e_m\}$
- MAP variables $\mathbf{M} = \{X_1, \ldots, X_n\} \setminus \{E_1, \ldots, E_m\}$

Compute

$$\max_{\mathbf{m}} \mathbb{P}(\mathbf{M} = \mathbf{m}, \mathbf{e})$$

Variants:

- DMPE: Decide if $\max_{\mathbf{m}} \mathbb{P}(\mathbf{M} = \mathbf{m}, \mathbf{e}) > k$ for given rational k
- SMPE: Select $\hat{\mathbf{m}}$ s.t. $\mathbb{P}(\mathbf{M} = \hat{\mathbf{m}}, \mathbf{e}) = \max_{\mathbf{m}} \mathbb{P}(\mathbf{M} = \mathbf{m}, \mathbf{e})$

Complexity of Inference

Upper Bound:

Marginal and MPE inference can be performed in worst-case polynomial-time in networks of bounded treewidth

Chandrasekaran et al. 2008:

Provided that NP $\not\subseteq$ P/poly and the grid-minor hypothesis holds, there is **no** graphical property that if constrained makes (marginal) inference polynomial in high-treewidth networks

Kwisthout et al. 2010; Kwisthout 2014:

Unless the satisfiability problem admits a subexponential-time solution, there is no algorithm that performs (MAP or marginal) inference in worst-case subexponential time in the treewidth

Local Probability Models

Extensive Specification

Local models are given as tables of rational numbers

Intelligent?	Marks	$\mathbb{P}(M \mid I)$
yes	A	0.4
yes	B	0.5
$\vdots$	$\vdots$	$\vdots$
yes	D	0.1
no	A	0.1
no	B	0.2
$\vdots$	$\vdots$	$\vdots$
no	D	0.2

Local Structure

Structure that cannot be read off from the graph:

- Context-specific independence: e.g.,

$$\mathbb{P}(Y|X, Z = z_0) = \mathbb{P}(Y|Z = z_1)$$

$$\text{and} \qquad \mathbb{P}(Y|X, Z = z_1) \neq \mathbb{P}(Y|Z = z_1)\,.$$

- Determinism:

$$\mathbb{P}(Y|Z) = \begin{cases} 1, & \text{if } Y = f(Z), \\ 0, & \text{if } Y \neq f(Z). \end{cases}$$

- Noisy-or networks (e.g. QMR-DT)

Local Structure

Beyond The Treewidth Barrier

> *" It has long been believed (...) that exploiting the local structure of a Bayesian network can speed up inference to the point of beating the treewidth barrier. (...) [However,] we still do not have strong theoretical results that characterize the classes of networks and the savings that one may expect from exploiting their local structure."*
>
> *– A. Darwiche, 2010*

Local Structure

Can constraining the expressivity of the local probability models allow for tractable inference?

This Work

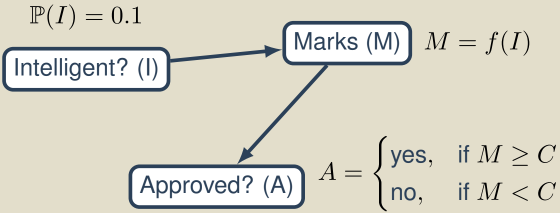
Complexity analysis of DMAP and DMPE in high-treewidth networks parameterized by the expresssivity of local probability models

Functional Bayesian Networks

Functional Bayesian Networks [Pearl 2000, Poole 2008]

Local probability models are

- arbitrary for root nodes (i.e. $\mathbb{P}(X) = \alpha$)
- deterministic for internal nodes (i.e. $X = f(\text{pa}(X))$)



Every Bayesian network can be converted into an equivalent functional Bayesian network (by adding new variables)

Results

There are tractable models of high treewidth...

E.g.: DMPE is in P when variables are Boolean, functions are logical conjunctions (AND) and evidence is positive (i.e. $E_i = \text{true}$)

...but they must be relatively simple

- DMPE is NP-complete when variables are Boolean and functions are logical conjunctions (evidence can be positive or negative)
- DMPE is NP-complete when variables are Boolean, functions are disjunctions (OR) and evidence is positive
- DMAP is NP^{PP}-complete when variables are Boolean, functions are disjunctions and evidence is positive
- DMAP is NP^{PP}-complete when variables are Boolean and functions are conjunctions (evidence is arbitrary)

Conclusion

- Continuation of previous work on complexity of marginal inference [Cozman and Mauá 2014]
- Some results showing tractable and intractable cases when parameters are "tied" (i.e., relational models)
- Meet me at poster session (poster #26)