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

Denis D. Mauá
Cassio P. de Campos
Fabio G. Cozman

Universidade de São Paulo, Brazil
Queen’s University Belfast, UK
Universidade de São Paulo, Brazil
Bayesian Network

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

Intelligent? (I)  
Marks (M)  
Approved? (A)  

$\mathbb{P}(I)$  
$\mathbb{P}(M | I)$  
$\mathbb{P}(A | M)$
MAP Inference Problem

Given:

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

Compute

\[
\max_m \mathbb{P}(M = m, e) = \max_m \sum_h \mathbb{P}(M = m, H = h, e)
\]

Variants:

- **DMAP:** Decide if \(\max_m \mathbb{P}(M = m, e) > k\) for given rational \(k\)
- **SMAP:** Select \(\hat{m}\) s.t. \(\mathbb{P}(M = \hat{m}, e) = \max_m \mathbb{P}(M = m, e)\)
MPE Inference Problem

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

Compute

\[
\max_m \mathbb{P}(M = m, e)
\]

Variants:
- **DMPE**: Decide if \(\max_m \mathbb{P}(M = m, e) > k\) for given rational \(k\)
- **SMPE**: Select \(\hat{m}\) s.t. \(\mathbb{P}(M = \hat{m}, e) = \max_m \mathbb{P}(M = m, e)\)
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 | $P(M|I)$ |
|--------------|-------|----------|
| yes          | A     | 0.4      |
| yes          | B     | 0.5      |
| no           | A     | 0.1      |
| no           | B     | 0.2      |
| no           | D     | 0.2      |
Local Structure

Structure that cannot be read off from the graph:

- **Context-specific independence:** e.g.,

\[
P(Y|X, Z = z_0) = P(Y|Z = z_1)
\]

and

\[
P(Y|X, Z = z_1) \neq P(Y|Z = z_1).
\]

- **Determinism:**

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

- **Noisy-or networks** (e.g. QMR-DT)
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
Can constraining the expressivity of the local probability models allow for tractable inference?
Complexity analysis of DMAP and DMPE in high-treewidth networks parameterized by the expresssivity of local probability models
Functional Bayesian Networks

Local probability models are

- arbitrary for root nodes (i.e. \( P(X) = \alpha \))
- deterministic for internal nodes (i.e. \( X = f(pa(X)) \))

\[
P(I) = 0.1
\]

\[
\begin{align*}
\text{Intelligent? (I)} & \quad \text{Marks (M)} \\
\text{Approved? (A)} & \quad M = f(I)
\end{align*}
\]

\[
A = \begin{cases} 
\text{yes,} & \text{if } M \geq C \\
\text{no,} & \text{if } M < C
\end{cases}
\]

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 $\text{NP}^{\text{PP}}$-complete when variables are Boolean, functions are disjunctions and evidence is positive
- DMAP is $\text{NP}^{\text{PP}}$-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)