

New Complexity Results for MAP in Bayesian Networks

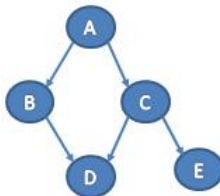
Cassio P. de Campos

Dalle Molle Institute for Artificial Intelligence
Switzerland



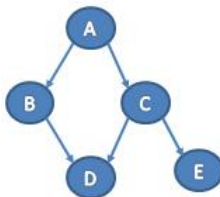
IJCAI, 2011

Bayesian nets



- ▶ Directed acyclic graph (DAG) with nodes associated to (categorical) random variables;
- ▶ Collection of conditional probabilities $p(X_i|\Pi_i)$ where Π_i denotes the parents of X_i in the graph (Π_i may be empty);
- ▶ Every variable is conditionally independent of its non-descendants given its parents (Markov condition).

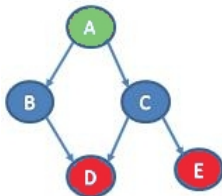
Bayesian nets



- In other words, it is a compact way based on (in)dependence relations to represent a joint probability distribution.

$$p(X_1, \dots, X_n) = \prod_i p(X_i | \Pi_i)$$

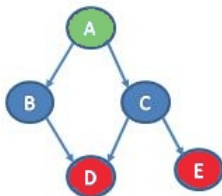
Belief Updating



- BU: given a set of queried variables X and their states x , evidence variables E and their states e , compute $p(X = x | E = e)$.

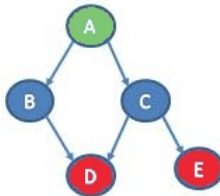
$$p(a|d, e) = \frac{p(a, d, e)}{p(d, e)} = \frac{\sum_{b,c} p(a, b, c, d, e)}{\sum_{a,b,c} p(a, b, c, d, e)}$$

Belief Updating - Decision version



- ▶ In terms of complexity, we can restrict ourselves to the computation of $p(X = x, E = e)$.
- ▶ D-BU: given a rational α , a set of queried variables X and their states x , evidence variables E and their states e , decide whether $p(x, e) > \alpha$.

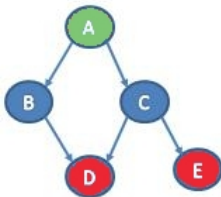
(Partial) Maximum a Posterior (MAP)



- ▶ MAP: given a set of queried variables X , evidence variables E and their states e , compute $\operatorname{argmax}_x p(x|e) = \operatorname{argmax}_x p(x, e)$.

$$\operatorname{argmax}_a p(a, d, e) = \operatorname{argmax}_a \sum_{b,c} p(a, b, c, d, e).$$

(Partial) Maximum a Posterior (MAP) - Decision version



- D-MAP: given a rational α , a set of queried variables X , evidence variables E and their states e , decide whether $\max_x p(x, e) > \alpha$.

Restricting Treewidth and Maximum cardinality

- ▶ MAP- z - w and D-MAP- z - w : same problems as before, but with two restrictions: z is a bound on the cardinality of any variable in the network, and w is a bound on the treewidth of the network. (The same definition can be used for the BU problem, which becomes BU- z - w and D-BU- z - w .)
- ▶ In order to express no bound, we use the symbol ∞ . E.g. D-MAP- ∞ - ∞ is the problem as defined earlier, and D-MAP- ∞ - w has a bound for treewidth, but not for cardinality.

Previous results

Complexity of this problems had been studied before, including the case of bounded treewidth.

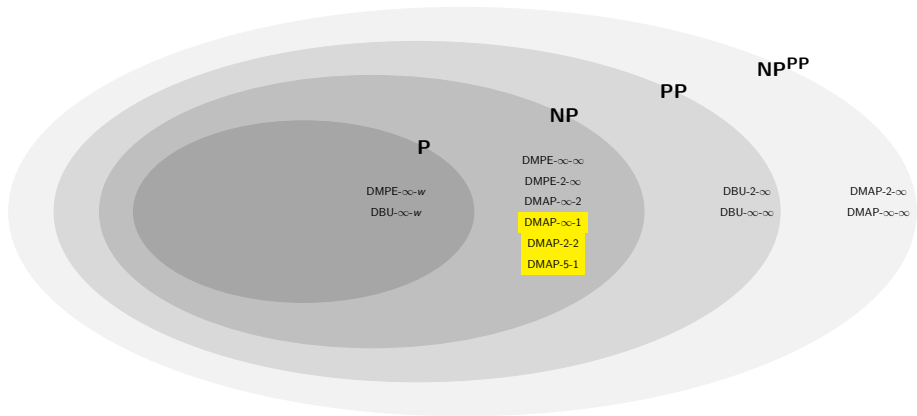
- ▶ D-BU- ∞ - ∞ is PP-complete, while D-BU- ∞ - w is in P. In fact, limiting the cardinality does not help: D-BU-2- ∞ is still PP-complete [Littman et al. 2001]. The functional versions are similar and discussed in [Roth 1996].
- ▶ D-MAP- ∞ - ∞ is NP^{PP}-complete, while D-MAP- ∞ - w is NP-complete [Park & Darwiche 2004].
- ▶ MAP- ∞ - w is also shown not to be in Poly-APX [Park & Darwiche 2004]. (Unless P=NP) It is shown that there is no polynomial time approximation that can achieve a 2^{b^ε} -factor approximation, for $0 < \varepsilon < 1$, b is the length of the input.

... but cardinality has been neglected so far

This paper presents new results for MAP that take cardinality into consideration.

- ▶ D-MAP-2-2 remains NP-complete (trick reduction from PARTITION).
 - ▶ This includes binary polytrees.
- ▶ D-MAP- ∞ -1 remains NP-complete (reduction from MAX-2-SAT using a naive-like structure) and D-MAP-5-1 is NP-complete too (reduction from PARTITION using an HMM-like structure).
 - ▶ This includes even simple trees.
- ▶ It is NP-hard to approximate MAP- ∞ -1 to any factor 2^{b^ϵ} (the construction comes from the naive-like structure, and uses similar arguments as in [Park & Darwiche 2004]).

Decision problems

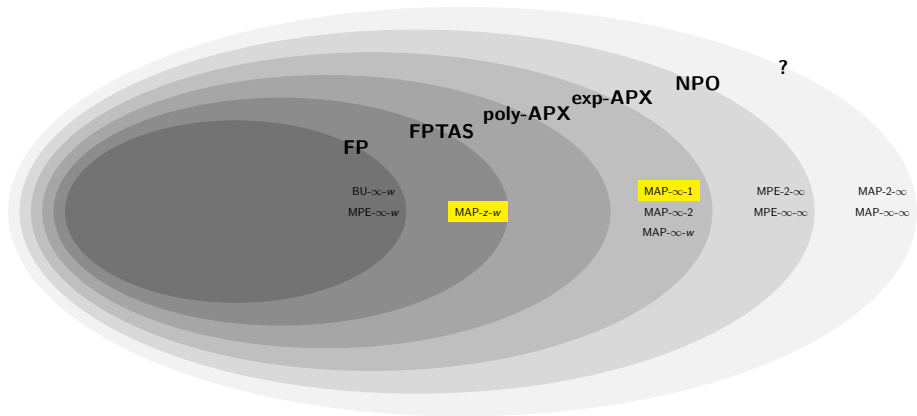


... and there is (some) hope ...

MAP- $z-w$ is hard, but has a FPTAS!

- ▶ We develop a Fully Polynomial-time Approximation Scheme for MAP when both treewidth and cardinality are bounded.
 - ▶ The idea is to compute all possible candidates and propagate them as in a BU inference, but keeping the number of candidates bounded by a polynomial in the length of the input (following ideas from [Papadimitriou and Yannakakis, 2000]).
- ▶ Previous inapproximability results are not contradicted: they had used variables with high cardinality.

Functional problems



Conclusions

This paper targets on understanding better the computational complexity of MAP.

- ▶ The problem is shown to remain hard in *binary polytrees* and *trees with bounded cardinality*.
- ▶ The problem is shown to be not approximable even in *trees* (without cardinality restrictions).
- ▶ An FPTAS is devised when both treewidth and cardinality are bounded.

Thanks

Thank you for your attention. Further questions: *cassiopc@acm.org*

Work partially supported by

- ▶ Project *Computational Life Sciences - Ticino in Rete*, Switzerland.
- ▶ Grant from the Swiss NSF n. 200020_134759/1.